# COARSE-GRAINING RECURSIVE COMPUTATION: CONTINUATION-RELATIVE INTERFACES AND FAILURE MODES

**Anonymous authors**
Paper under double-blind review

## ABSTRACT

A learned substitute can replace a run of blocks or iterations in a deep or recursive network at lower cost. Replacement methods fit the substitute to reproduce the output of what it replaces and accept it on that fidelity. Fidelity, however, is measured at the replacement boundary, and usefulness is decided after it, by the *continuation*: the computation that runs on the substitute's output, and whether it still solves the task at lower total cost. We give an evaluation contract that judges a replacement by its continuation. The contract separates boundary-state agreement, path statistics, executed computation, and task loss, and states what agreement at each level certifies about the next; it decomposes the continuation gap exactly into a fixed-action term and a reallocation term; and it requires the control that replacement methods omit, the reduced network trained from scratch on the task. Under the contract, fidelity is neither necessary nor sufficient, and three settings show both directions. On GPT-2, a substitute trained through the continuation with no fidelity target beats a published block-replacement method's fidelity optimum at all five boundaries tested in distribution; under distribution shift the fidelity optimum keeps the lower loss. On a pretrained recursive model, the continuation-trained substitute is $0.36$–$0.37$ nats below the un-adapted model in distribution at $24\%$ fewer recursion MACs, with up to thirty times the fidelity optimum's boundary error. On an eight-expert prototype over six synthetic data families, a from-scratch control beats every arm that inherits the trained weights on all 60 family, model, and split combinations, and anchoring expert occupancy to the original path cuts the distance from $0.78$ to $0.26$ without consistently lowering error. A replacement claim thereby becomes a question its continuation can answer: what it still achieves, at what cost, against which control.

## 1 INTRODUCTION

Repeated neural computation evolves a representation through recursive updates, solver iterations, or denoising steps; in a sparse-expert model a router also selects which modules act on the current state. A learned *macro* can replace several such updates at lower cost, but a cheaper step is useful only if the computation that follows still performs the task. We therefore ask a single question: *when can a segment of recursive computation be replaced by a cheaper macro while preserving task performance under a specified continuation?*

Fidelity is measured at the replacement boundary, and usefulness is decided after it. Two paths can agree at a boundary state yet route differently afterwards; equal occupancies (per-expert activation mass over an interval) can hide different temporal orders; a predicted path statistic is not the computation actually executed; and speed and quality must be assessed jointly. We therefore judge a replacement *continuation-relatively*: by what the declared continuation still achieves with the macro's output, at what total cost, against controls trained on the same deployed graph. Short of exact agreement on every input, which would trivially preserve the fine path's continuation error, measured fidelity (boundary-state or occupancy agreement) is neither necessary nor sufficient, and we show both directions experimentally.

Contributions:



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.

- **An evaluation contract** that fixes the replacement boundary, the information a replacement may retain, the declared continuation, and the legal deployment information, with four checkable pass rules (section 4).
- **Conditional analysis** for state closure, path composition, sparse support, and continuation loss, including an exact decomposition of the continuation gap into a fixed-action term and a reallocation term; the mathematics is elementary and its value lies in delimiting what each diagnostic can certify (section 5).
- **Controlled evidence** (section 6). On a learned eight-expert prototype over six synthetic data families, the DIRECT controls train the deployed graph directly and have no trained fine path to match. Matched to the development-selected arm's total training MACs and updates, DIRECT-COMPUTE beats that arm's own fine path on 58 of 60 parent–split pairs, and DIRECT-DATA beats every structured arm on all 60 (section J). An occupancy anchor to the fine path cuts the macro's occupancy distance to it by a factor of $2.3$–$3.0$ without consistently lowering its continuation error. The anchor and a composition penalty together raise error at every weight tested; added alone, the composition penalty raises error ($+13.7\%$, 5/5 parents) and the anchor does not ($-0.6\%$, 2/5; section L). On GPT-2, a substitute trained with no fidelity target beats the least-squares fidelity optimum of a published block-replacement method's substitute class on in-distribution loss at all five boundaries tested (section N); on a pretrained recursive model it is below the un-adapted model and the model's own coarse step in distribution (section O). Neither the isolated $2.22\times$ speedup nor an input-only selector achieves the declared quality-matched savings.

## 2 Related Work and Novelty Boundary

**Compressing state evolution.** Distillation, consistency, shortcut, and mean-flow methods reduce repeated evaluations of learned dynamics (Lipman et al., 2022; Liu et al., 2022; Salimans & Ho, 2022; Song et al., 2023; Frans et al., 2024; Geng et al., 2025); recursive and looped flow maps, Universal Transformers, DEQs, LoopFormer, and Latent Flow Transformers reuse or compress computation across recurrence or depth (Huang et al., 2026b; Lee et al., 2026; Yoo et al., 2026; Suleymanzade et al., 2026; Dehghani et al., 2018; Bai et al., 2019; Lin et al., 2026; Jeddi et al., 2026; Wu et al., 2025). These works compress state evolution; we ask what a replacement must preserve for a declared continuation.

**Replacing, pruning, and skipping computation.** The closest empirical setting is depth pruning with replacement: LLM-Streamline, ReplaceMe, and Ghosted Layers remove a block of layers and fit a lightweight, linear, or activation-aligned substitute to the removed block's input–output map (Chen et al., 2025; Shopkhoev et al., 2025; Yun et al., 2026); expert pruning keeps the experts that best reconstruct a layer's output (Lu et al., 2024), and BERT-of-Theseus trains smaller successor modules inside the frozen original network with the task loss alone (Xu et al., 2020). Mixture-of-Depths learns per-token block skipping (Raposo et al., 2024), and CALM calibrates hidden-state early exits to agree with the full model or to bound task risk relative to it (Schuster et al., 2022); our prototype selectors read only the raw input, and the recursive-model gate reads only the boundary state reached before the decision. Our contract asks what a substitute must preserve and adds the control the replacement methods omit, training the same reduced graph from scratch. That control has precedent: pruned architectures retrained from scratch match or beat fine-tuned pruned weights (Liu et al., 2019), in LLMs only for coarse structured pruning given the full pipeline's tokens (Xu et al., 2026), the budget DIRECT-DATA receives here, and distillation from a misspecified teacher separates mimicry from task performance (Hara & Hino, 2026), the separation our occupancy anchor exhibits inside a recursive expert graph.

**Recurrence, coadaptation, and budgeted selection.** Sparse MoEs route inputs to subsets of experts (Shazeer et al., 2017; Fedus et al., 2021); Mixture-of-Recursions, LoopMoE, and The Thinking Pixel combine recurrence with conditional computation (Bae et al., 2025; Chen et al., 2026b; Sun et al., 2026). SFTN and SAR adapt the teacher or its router to the student (Park et al., 2021; Kim et al., 2025); System-1.5 and recursive shortcut reasoning learn shortcuts with two-stage self-distillation or local repair (Wang et al., 2025; Shankar, 2026). Routing signatures from normalized activation counts (Avinash, 2026) and same-layer expert specialization (Li et al., 2025) do not capture directed temporal order; DYNAMOD jointly fits a cheap predictor and gate under a

prediction-cost budget (Nan & Saligrama, 2017). Value Equivalence preserves specified decision queries and CQM preserves action comparisons under a fixed continuation (Grimm et al., 2020; Chen et al., 2026a); neither covers arbitrary future menus or budgets. That fidelity to a replaced computation is neither necessary nor sufficient for downstream usefulness has precedents: value equivalence (Grimm et al., 2020) and value-aware model learning (Farahmand et al., 2017) for dynamics models, objective mismatch (Lambert et al., 2020) empirically, Stanton et al. (2021) for distillation, and model stitching (Csiszárik et al., 2021) and end-to-end sparse dictionaries (Braun et al., 2024) for modules spliced into frozen networks; we instantiate that separation in two pretrained models and in a sparse-expert prototype with an executable route, an input-only selector, and same-graph DIRECT controls (mechanism-by-mechanism overlaps: table 28, section T).

## 3 Continuation-Relative Interfaces

Four objects organize the interface: *boundary state*, *retained path statistic*, *actual execution*, and *continuation usefulness*; their separation is scope, not novelty, since a recorded path statistic is not automatically a sufficient state, an executed route, or quality evidence.

**Fine recursion and boundary state.** For positive integers $E, K$, a fine program with $E$ experts takes $K$ steps of width $h = 1/K$:

$$z_{i+1} = F_i(z_i, c, b_i), \qquad \pi_i = R_i(z_i, c, b_i) \in \Delta^{E-1}, \quad 0 \le i < K. \tag{1}$$

Here $z_i$ is visible state, $c$ fixed context, $b_i$ closure history. The interval map $\Psi_{s,\Delta}$ composes them for $s = jh$, $\Delta = nh$, $j \ge 0, n \ge 1$, $j + n \le K$. If $z$ is sufficient, a macro predicts $(z_{s+\Delta}, S_I)$ from $(z_s, c, s, \Delta)$: $S_I$ is a path statistic of the interval. Otherwise it must also accept and propagate sufficient $b$, giving product-state dynamics on $(z, b)$, not merely an extra output head (equations (15) and (16)).

**Occupancy and composition.** For an interval $I = [s, s+\Delta]$ covering $n$ equally spaced steps with per-step routes $p_i$ (soft: $p_i = \pi_i$; hard: one-hot), $A_I = \sum_{i=0}^{n-1} p_i$ and $Q_I = A_I/n$. Lowercase $q, p$ denote generic simplex occupancies. Soft responsibilities $p_i = \pi_i$ are probability-valued; activations with unrenormalized capacity drops are sub-probability: $\mathbf{1}^\top Q_I = 1 - D_I/n$, $D_I = \sum_i (1 - \mathbf{1}^\top p_i)$. Step-count mass $A_I$ corresponds to duration mass $hA_I = \Delta Q_I$. With context suppressed, $\Delta = \Delta_1 + \Delta_2$, and $z_1 = \Psi_{s,\Delta_1}(z)$, positive admissible intervals obey

$$\Psi_{s,\Delta}(z) = \Psi_{s+\Delta_1,\Delta_2}(z_1), \tag{2}$$

$$\Delta Q^\star_{s,\Delta}(z) = \Delta_1 Q^\star_{s,\Delta_1}(z) + \Delta_2 Q^\star_{s+\Delta_1,\Delta_2}(z_1). \tag{3}$$

Here $Q^\star_{s,\Delta}(z)$ is the teacher's (the fine program's; the macro is the student) normalized interval occupancy, and the second interval uses the *reached successor*, not the original input; with memory these laws use the augmented state.

Occupancy $S_I^{(0)} = (n, A_I)$ records amounts, not order. The optional ordered statistic $S_I^{(1)} = (n, A_I, B_I, \ell_I, r_I, D_I)$, with $B_I = \sum_{i=0}^{n-2} p_i p_{i+1}^\top$ and $(\ell_I, r_I) = (p_0, p_{n-1})$, composes as $B_{IJ} = B_I + B_J + r_I \ell_J^\top$ for adjacent intervals (equation (13)); the boundary term cannot be recovered by averaging transition frequencies, and the prototype does *not* train $S^{(1)}$.

**Execution and continuation (figure 1).** The occupancy head predicts the path statistic of the skipped fine interval; a separate route $\rho_\theta$ determines which experts run (the prototype executes one continuation expert). Fix $(x, y)$ and frozen parameters: from a shared encoding, fine prefix $\mathcal{F}_\vartheta$ and macro $\mathcal{C}_\varphi$ reach the same replacement boundary. For a common finite action menu $\mathcal{A}$, let $o_j(a)$ denote the shared-readout output after prefix $j \in \{F, M\}$ and action $a$; each router chooses $a_j$ from its own reached state, not from $y$ or the menu outputs. Corresponding actions require shared semantics or fixed calibration alignment, never per-example rematching (section D).

**An illustrative construction.** A hand-built discrete construction makes the distinction exact (section G): for $k \in \{2, 3, 4\}$ blocks whose continuation must answer $A_i \oplus B_i$ or $B_i$ by phase, exhaustive enumeration shows that the one-bit summary $A_i \oplus B_i$ reaches zero risk at cost $2.70/3.25/3.80$, whereas the equal-width summary $A_i$ keeps risk $0.25$ at those costs; it is no evidence about learned macros.

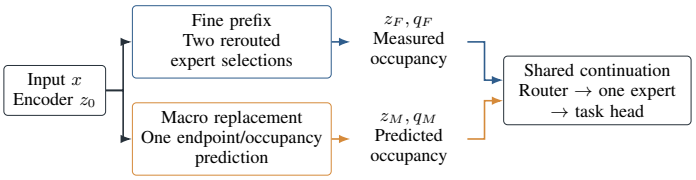



Figure 1: Continuation-relative replacement. A fine recursive prefix and its macro reach the same replacement boundary before a shared continuation and readout; evaluated quantities include endpoint/occupancy discrepancies, per-action responses, task error, and execution cost.

## 4 The Eight-Expert Prototype and Its Controls

**Eight-expert prototype.** The prototype is the executable, sparse version of the interface; table 1 lists its training arms. An encoder maps $x \in \mathbb{R}^4$ to $z_0 \in \mathbb{R}^{16}$; the fine prefix makes two internally rerouted hard selections from eight shared residual experts (hidden width 48), producing $z_F$ and occupancy $q_F = \frac{1}{2}(\mathrm{e}_{a_0} + \mathrm{e}_{a_1})$; a width-24 macro with softmax occupancy head predicts $(z_M, q_M) = C(z_0, 0, 2)$. Phase and span are step counts $j, n$ with $K = 2$, halved to $s = j/2$, $\Delta = n/2$ like the prefix phase $t/2$. Both branches feed state, occupancy, and normalized phase 1 to the same continuation router; the chosen expert and the two-output head read the state only. The macro never executes its skipped prefix; $q_M$ is anchored to that interval's occupancy only in arms with the occupancy anchor and is otherwise a learned router input.

**Training is not deployment.** Argmax deployment executes only the selected expert at each routing location. Training evaluates the finite continuation menu with probability-weighted individual task losses and a hard-forward straight-through prefix surrogate; neither is an exact gradient of hard dispatch. The base objective combines fine/macro task losses, endpoint MSE, and load balancing; the full objective adds occupancy anchoring (to the stop-gradient current fine occupancy), state/occupancy composition, and menu-response matching with uniform or detached task-relevance weights and no direct gradient through continuation-router logits. Only span 2 (the fine steps the deployed macro replaces) is fine-anchored, unit spans are factorization auxiliaries, and the macro accepts no incoming routing history (section P).

**The outer selector; budgets and fairness.** The bank and both complete paths are frozen before an input-only outer selector is fitted to predict fine/macro losses from raw $x$; development data fix thresholds before test scoring, and online selection is blind to targets and executes only the chosen path.

Both DIRECT controls train the identical fast graph from raw initialization; forward-affine multiply-accumulate (MAC) matching excludes backward, nonlinear, optimizer, memory, and dispatch work. Both DIRECT controls use uninterrupted Adam while the structured arms reset it, and fine pretraining leaves the macro untrained, so a structured arm's macro gets 1,200 updates to DIRECT's 2,400 or 7,910; each standalone method is charged pretraining, with the shared fit counted once.

**Parents, splits, and the candidate.** A *parent* is one independently initialized and pretrained fine model from which every arm starts, sharing its minibatch stream, with disjoint training, development, in-distribution (IID), and range-shift (enlarged input range) test draws; the *original family* is a two-branch piecewise target on $[-1, 1]^4$ (section P). Two development parents select relevance weighting, making RELEVANT the *candidate*, before five fresh confirmation parents; all 55 fits (eleven arms, five parents) were frozen before test access. The seven arms that inherit the parent's weights are the *structured* (inherited) arms; the two DIRECT rows are controls.

**Pass rules.** A replacement claim passes the contract only if all four rules hold, each decidable from the reported run. (a) *Shared interface*: passes iff macro and fine path reach the same replacement boundary and feed one declared continuation with one action menu. (b) *Same-graph controls*: passes iff controls trained from scratch on the same deployed graph are reported at named budgets (here the candidate's updates and MACs). (c) *Quality tolerance*: passes iff continuation error on every test

split is within a declared tolerance of the macro's own fine path and of continued fine-only training (here $2\%$, the $2\%$ *rule*). (d) *Legal-input cost*: passes iff end-to-end cost, measured with a selector that reads only legal deployment information, is below the fine path's.

Table 1: Arms: trainable parts, objective beyond the base task, and role. All arms share data and the deployed evaluation graph.

| Arm | Trainable parts | Objective beyond base | Role |
|---|---|---|---|
| JOINT-END | all | none (task + endpoint + load balance) | joint baseline |
| SIG | all | + occupancy anchor + composition | structure only |
| RESP-ONLY | all | + menu-response matching | continuation matching only |
| SHORTCUT-CONT | all | + state composition + own-selected output MSE | repair-style shortcut control |
| RELEVANT | all | + occupancy + composition + relevance-weighted response | **candidate** (chosen on development parents) |
| ROUTER-RELEVANT | macro + routers | full RELEVANT objective | macro-and-router adaptation |
| FROZEN-RELEVANT | macro only | full RELEVANT objective | frozen-bank adaptation |
| DIRECT-DATA / - COMPUTE | fast graph from scratch | task + load balance (no endpoint) | unrestricted same-graph control (matched updates / matched MACs) |
| Continued fine-only; pre-trained fine | fine path | task + load balance | references |

## 5 Conditions and What They Predict

Each elementary condition below fixes what one diagnostic can and cannot certify, with its empirical status in the prototype. Norms are fixed and bounds finite; full statements and proofs are in section E.

**Proposition 1** (Six conditions and their empirical status)**.** *(i)* ***Closure with omitted memory.*** *At fixed visible $(z, c, s, \Delta)$, if two admissible hidden boundary states $b, b'$ of an augmented sufficient state $(z, b)$ produce endpoints separated by $d_{\mathrm{sep}} > 0$, every deterministic state-only predictor has worst-case error at least $d_{\mathrm{sep}}/2$ on this pair (proposition 2).* Status: *vacuous here, since the prefix router reads only $(z_t, t)$ and $d_{\mathrm{sep}} = 0$.*

*(ii)* ***State-composition bound.*** *On positive valid intervals obeying the teacher process law, suppose the student composition defect is at most $\delta_z$, subinterval errors are uniformly at most $\varepsilon_1, \varepsilon_2$ on relevant states,* including the predicted first successor*, and the second teacher map is $L_2$-Lipschitz there. Full-interval error is then at most $\delta_z + \varepsilon_2 + L_2\varepsilon_1$ (proposition 3). Sequential error can amplify when the relevant Lipschitz constant exceeds $1$ (proposition 4).* Measured: *two composed unit-span macros give $6.25\times$ (IID) and $4.30\times$ (range shift) the span-2 macro's continuation MSE, a downstream diagnostic, not the bound's state norm; the unit-span maps are not fine-anchored, so $\varepsilon_1, \varepsilon_2$ are uncontrolled, and the trained composition penalties are mean squared errors, not these pointwise defects.*

*(iii)* ***Endpoint-to-route non-identifiability.*** *At a fixed input, let expert outputs be columns of $V \in \mathbb{R}^{d\times E}$ and $q$ be an interior simplex point. If $\{a : Va = 0, \mathbf{1}^\top a = 0\}$ is nontrivial, as guaranteed for $E > d + 1$, infinitely many nearby probability routes give the same $Vq$ (lemma 1).* Status: *not applicable; the prefix selections are hard, so $q_F$ is not interior.*

*(iv)* ***Occupancy does not determine order.*** *For noncommuting maps $A(z) = 2z$ and $B(z) = z+1$, both orders have occupancy $(1/2, 1/2)$ but endpoints $2z + 1$ and $2z + 2$ (proposition 6). The ordered $S^{(1)}$ distinguishes this pair.* Status: *untested.*

*(v)* ***Sparse-support floor.*** *For integers $E, N \geq 1$, $1 \leq k \leq E$, $N$ top-$k$ macrosteps activate at most $m = \min(E, Nk)$ experts. With aligned identities and probability occupancies, the minimum $\mathrm{TV}(p, q) = \frac{1}{2}\|p - q\|_1$ over all $p$ supported on at most $m$ experts is $1 - \sum_{e=1}^{m} q_{(e)}$, with $q_{(e)}$ sorted decreasingly (proposition 8). Pre-capacity top-$k$ sets coincide if the teacher logit margin is $\gamma > 0$ and $\|\widehat{g} - g\|_\infty < \gamma/2$; $k = E$ needs no margin (proposition 7).* Status: *not applicable; the floor bounds executed occupancy, and the macro executes no expert in the replaced interval ($q_M$ is a predicted head), so it cannot explain the candidate's $0.31$ distance $\mathrm{TV}(q_M, q_F)$.*

*(vi)* ***Fixed-action term, reallocation term, and regret.*** *For the frozen finite-menu interface, with finite losses, define*

$$\ell_j(a) = \tfrac{1}{2}\left\|o_j(a) - y\right\|_2^2, \qquad m_j = \min_{a\in\mathcal{A}} \ell_j(a), \qquad \mathfrak{r}_j = \ell_j(a_j) - m_j \geq 0. \tag{4}$$

*(With two-dimensional outputs this equals the per-coordinate MSE of the tables; other dimensions differ by a fixed factor.) The oracle minimum $m_j$ is diagnostic only. The exact split is*

$$\begin{aligned}\ell_M(a_M) - \ell_F(a_F) &= \underbrace{\ell_M(a_F) - \ell_F(a_F)}_{\textit{fixed-action term}} + \underbrace{\ell_M(a_M) - \ell_M(a_F)}_{\textit{reallocation term}} \\ &= m_M - m_F + \mathfrak{r}_M - \mathfrak{r}_F. \end{aligned} \tag{5}$$

*If $\max_{a\in\mathcal{A}} |\ell_M(a) - \ell_F(a)| \leq \varepsilon_\ell$, then $|m_M - m_F| \leq \varepsilon_\ell$. Sufficient conditions for* every *action are $\|o_M(a) - o_F(a)\|_2 \leq \varepsilon_o$ and $\|o_F(a) - y\|_2 \leq B_{\text{res}}$, giving $\varepsilon_\ell = B_{\text{res}}\varepsilon_o + \varepsilon_o^2/2$ (proposition 10). The first line takes $a_F$ as reference; the mirror split is equally exact, so term comparisons in table 2 refer to $a_F$. A minibatch mean response loss supplies no uniform coverage or regret control, and the reallocation term can offset the fixed-action term.* Measured: *table 2.*

That measured fidelity is neither necessary nor sufficient is the empirical finding of section 6, consistent with these conditions rather than a consequence of them: exact agreement of $(z, q)$ at the replacement boundary gives zero gap, but short of it (vi) leaves the reallocation term uncontrolled, since a small state change can switch the chosen action, and the measured statistics never certify exact agreement.

## 6 Measured Fidelity Is Neither Necessary Nor Sufficient

**In parent means, response matching goes with a smaller fixed-action term, not a smaller reallocation term.** The decomposition is a frozen-model diagnostic (equation (5)), not evidence of a training-causal property of the router. RESP-ONLY attains the smallest positive fixed-action term ($+0.36$), consistent with its objective, yet its reallocation term is $+0.48$ and its total gap $+0.83$: average response matching supplies neither uniform coverage nor a regret guarantee. The candidate's reallocation term is about half its fixed-action term and positive, and its macro-side regret is $5.21$ (table 2); the negative mean reallocation terms of SHORTCUT-CONT and JOINT-END rest on one parent (seed 3413: $-13.95$ and $-5.21$ against fixed-action terms of $+16.05$ and $+5.76$); without it both means are positive. The oracle-menu floor $3.95$ versus realized $9.16$ bounds what any action choice on the same states could achieve; ROUTER-RELEVANT's negative gap reflects a fine path degraded by router-only adaptation ($14.43 \to 28.97$).

**Factor A raises error at every weight tested; added alone, the composition penalty raises it and the anchor does not.** JOINT-END, SIG, RESP-ONLY, and RELEVANT (prespecified arms; grouping post hoc) form a paired $2 \times 2$ design on the objective (factor A: occupancy anchor plus state/occupancy composition; factor R: menu-response matching, relevance-weighted in RELEVANT), sharing initialization, pretraining, and minibatch stream within each parent; table 3 reports log-MSE main effects. On the original family, adding A raises continuation error on every parent in both splits ($+20.5\%$ IID), adding R lowers it on five of five IID parents ($-3.0\%$), and the interaction is below $2\%$. One cell is not a pure contrast (RELEVANT weights its response term by relevance), so the pure simple effect is JOINT-END$\to$SIG: worse on $5/5$ parents in both splits. A declared post-hoc weight sweep (section I) finds no helpful regime: SIG is worse than JOINT-END on $5/5$ IID parents at every weight, the penalty rising from $+11.2\%$ to $+22.5\%$. A declared post-hoc decomposition adds each factor-A term alone to JOINT-END at two budgets: the composition penalty raises IID error alone ($+13.7\%$, $5/5$ parents) and the anchor does not ($-0.6\%$, $2/5$) while cutting occupancy distance $0.78 \to 0.26$; together they interact ($+21.2\%$; section L). Nor is fidelity necessary: the DIRECT controls have no trained fine path, yet DIRECT-COMPUTE, matched to the candidate's total training MACs, has lower continuation error than the candidate's own fine path at 2,840 instead of 5,176 forward MACs on 58 of 60 parent–split pairs, and DIRECT-DATA, matched to its total updates, has lower error than every structured arm on all 60 (section J). These margins depend on DIRECT's budget, and the opposite orderings of the five trainable arms by occupancy

Table 2: Continuation-gap decomposition, original family, IID, mean over five parents ($\times 10^{-3}$; positive Gap means the macro is worse than its fine path). Gap = fixed-action + reallocation (exact; rounded); *Regret* is the macro side's excess loss over its own oracle menu; *Id. agr.* is next-expert agreement with the fine path; bold cells are discussed in the text. †DIRECT controls have no trained fine path. Range-shift split: section H.

| Arm | Fine | Macro | Gap | Fixed-act. | Realloc. | Regret | Id. agr. |
|---|---|---|---|---|---|---|---|
| JOINT-END | 6.91 | 7.92 | $+1.01$ | $+1.79$ | $-0.78$ | 4.64 | 0.948 |
| SIG | 6.19 | 9.52 | $+3.32$ | $+3.08$ | $+0.25$ | 5.53 | 0.944 |
| RESP-ONLY | 6.89 | 7.72 | $+0.83$ | $\mathbf{+0.36}$ | $+0.48$ | 4.49 | 0.960 |
| SHORTCUT-CONT | 6.22 | 8.94 | $+2.73$ | $+5.07$ | $-2.34$ | 5.26 | 0.918 |
| RELEVANT (candidate) | 6.44 | 9.16 | $+2.72$ | $+1.83$ | $\mathbf{+0.89}$ | 5.21 | 0.948 |
| ROUTER-RELEVANT | 28.97 | 20.20 | $-8.77$ | $-3.97$ | $-4.81$ | 8.12 | 0.807 |
| FROZEN-RELEVANT | 14.43 | 24.48 | $+10.04$ | $+9.41$ | $+0.64$ | 10.95 | 0.822 |
| DIRECT-DATA† | n/a | 5.54 | n/a | n/a | n/a | 3.53 | n/a |
| DIRECT-COMPUTE† | n/a | 3.87 | n/a | n/a | n/a | 2.69 | n/a |

fidelity and by continuation error (Spearman $-1.00$; table 13) hold only at the confirmation budget (section K). Factor A leaves endpoint error nearly unchanged (JOINT-END $0.032$, SIG $0.031$); the candidate has the lowest endpoint error of the seven structured arms ($0.024$) yet higher continuation error than three of them.

**Five further data families replicate the DIRECT losses; the factor effects replicate only in part.** Five further families (*quadrant*, *iterated*, *deep6*, *deep8*, *chaotic*; section J) were run under protocols committed before their runs, with every hyperparameter, arm, budget, and the candidate inherited. The candidate loses to both DIRECT controls on every parent of every family and split (rule (b)'s controls), and the macro/own-fine ratio stays far outside the $2\%$ rule ($1.12$–$1.95$); factor A is positive on $5/5$ and $4/5$ parents of the first two families ($+7$–$10\%$), on $5/5$ of the deeper maps ($+8.5$–$19.4\%$), and on $3/5$ of chaotic parents ($+2.7$–$2.8\%$); the weight sweep was run on the original family only. The response-matching benefit is weaker than on the original family ($-2.0\%$ to $+0.4\%$, lowering error on at most $4/5$ parents in any family and split).

**The separation appears in a pretrained language model.** A prospectively declared instantiation on GPT-2 small (section N) removes blocks $[l, l{+}k)$ at five boundaries and fits, within ReplaceMe's residual linear substitute class and with its least-squares objective as one arm (not a reproduction of that pipeline), six objectives with the continuation frozen. The closed-form fidelity optimum has lower boundary MSE than every continuation-trained substitute at every boundary (its Adam-fitted twin FID-SGD is marginally lower at three of five). Yet a substitute trained through the continuation with no fidelity target (the DIRECT analogue restricted to the substitute's parameters; no from-scratch control of the full reduced graph was run on the language models) has lower in-distribution loss at 5 of 5 boundaries (two seeds at four, one at $(4, 4)$), and is below the unmodified model at the three two-block boundaries. From the fidelity optimum, continuation training raises boundary MSE and lowers loss at every checkpoint of every boundary. Under shift (LAMBADA) the fidelity optimum keeps the lower loss at 5 of 5; refitting on a broad web corpus cuts that gap from $0.52$ to $0.11$ nats and reverses it against the arm that matches the unmodified model's output distribution (END-KL), at the cost of the in-distribution gain: the reversal is a property of the fitting distribution, not of the objective.

**The separation appears on a pretrained recursive model, with the largest margin.** A second declared instantiation (section O) uses LoopFormer (Jeddi et al., 2026), three shared blocks applied eight times, and replaces iterations 2–3 by the same residual linear map with iterations 4–7 frozen; skipping the two iterations costs $+0.075$ nats, and the model's own coarse step (one block call with $dt{=}2/8$) costs $+0.009$. The fidelity optimum (FID-LS) again has the lowest boundary MSE ($0.53$) and a higher loss than the unmodified model ($3.897$ versus $3.857$); CONT-ZERO has thirty times the boundary MSE ($16.1$) and a loss of $3.502$ on both seeds, $0.36$ nats below the un-adapted model and below the model's own coarse step (NATIVE-$dt$, $3.866$), at $24\%$ fewer recursion MACs. From the fidelity optimum, CONT-LS raises boundary MSE and lowers loss at all six checkpoints on both seeds (seeds vary minibatch order only), ending at $3.487$. The unmodified model and NATIVE-$dt$ are not fitted on WikiText-2 and no same-data adapter on the unmodified graph was run; the margin is absent when the substitute is fitted on a broad corpus (section O). Under shift the order reverses as

Table 3: Original and first two replication families, identical frozen protocol (five parents each; *original; the other two were declared and run afterwards without re-choosing the candidate; three further families: table 12). Columns 3–4 count parents on which the candidate's error exceeds DIRECT-DATA's and JOINT-END's (DIRECT-COMPUTE's: $5/5$ everywhere); factor effects are geometric-mean changes in macro MSE (positive = higher error; parentheses: parents on which A raises and R lowers error); the last column is the macro/own-fine ratio of parent-mean MSE ($\leq 1.02$ passes the $2\%$ rule).

| Family | Split | >DIRECT-D | >JOINT | A effect | R effect | Macro/fine |
|---|---|---|---|---|---|---|
| original* | IID | 5/5 | 5/5 | $+20.5\%$ (5/5) | $-3.0\%$ (5/5) | 1.42 |
| original* | Shift | 5/5 | 5/5 | $+17.3\%$ (5/5) | $-3.5\%$ (4/5) | 1.29 |
| quadrant | IID | 5/5 | 5/5 | $+9.6\%$ (5/5) | $+0.4\%$ (2/5) | 1.47 |
| quadrant | Shift | 5/5 | 5/5 | $+7.4\%$ (5/5) | $+0.0\%$ (3/5) | 1.43 |
| iterated | IID | 5/5 | 3/5 | $+7.2\%$ (4/5) | $-1.4\%$ (2/5) | 1.51 |
| iterated | Shift | 5/5 | 4/5 | $+9.7\%$ (4/5) | $+0.2\%$ (1/5) | 1.95 |

before (LAMBADA: NATIVE-$dt$ 3.229, unmodified 3.225, CONT-ZERO 3.551, CONT-LS 3.502); a broad fitting corpus halves that gap at the cost of the in-distribution gain.

**Fixed replacement fails the utility screen: rules (c) and (d).** The candidate (per-arm MSE: table 26) beats FROZEN-RELEVANT and ROUTER-RELEVANT on every parent of both splits but loses to JOINT-END and both DIRECT controls on every parent of the original family (IID MSE $15.76\%$, $65.52\%$, and $137.06\%$ higher), failing the prespecified screen ($5\%$ lower MSE than JOINT-END and both DIRECT controls); SHORTCUT-CONT and RESP-ONLY also have lower mean MSE. Macro IID MSE is $42.19\%$ above its own coadapted fine path, so the $2\%$ rule fails; the batch-one latency ratio is $0.451$ ($2.22\times$), a speed–quality trade-off, not quality-matched acceleration.

**A label-informed oracle could save $43\%$ of work; no input-only selector does on the prototype.** An oracle that sees the labels and picks the macro on a fixed number of examples, charged the selector's 96 MACs, chooses it on 10,180/10,240 development examples at MSE $0.007167$ (declared limit $0.007174$, $1.02\times$ the fine path's development MSE), saving $43.01\%$ affine work: *development oracle headroom*, not achieved selector performance. The input-only selectors are a $4 \to 16 \to 2$ tanh MLP (96 MACs) and a linear $4 \to 2$ control (8 MACs), 50 fits in total (five arm banks, five parents, two kinds) with development-frozen thresholds (section P).

The legal primary selector (table 4) increases own-fine error by $8.70\%$ and latency by $23.78\%$ (rule (d) fails); its $9.71\%$ affine saving misses the declared $10\%$ target, and the linear control trades $10.32\%$ lower latency for $15.11\%$ higher error. Both fail range-shift tolerance, and every primary selector policy fails pooled quality on both splits. On the recursive model a linear gate on the boundary state (a hidden state, not the raw input), fitted on validation, meets rule (d)'s cost criterion by MAC count in distribution ($63\%$ usage: WikiText-2 test 3.561 against 3.857, $15.5\%$ of recursion MACs saved), but the ungated substitute meets it with lower loss and a larger saving, and the declared selector hypothesis (no worse than the unmodified model on both held-out sets) was not met (section O).

## 7 Discussion and Limitations

**What the failures distinguish.** Fidelity, continuation usefulness, and deployment cost are separate screens: wins against restricted adaptation and losses against the same-graph DIRECT controls answer different questions (section 6). DIRECT shares the deployed graph, so any advantage concerns learning structure, supervision, or optimization, not a larger function class. In this prototype the candidate's fine path beats DIRECT-COMPUTE on at most one of five parents on every family, including three declared to search for a task where depth matters (section J), and at every fine-pretraining budget up to $16\times$ (section K); it beats only a from-scratch graph charged the replacement's 1,200 updates alone (DIRECT-ADAPT, $26\%$ of that phase's MACs). The prototype therefore tests whether the contract detects a replacement that should not be made, not whether a worthwhile one can be learned. A weak fine path suggests a rival reading of factor A, that anchoring to a poor teacher must hurt (Hara & Hino, 2026); but the anchor alone raises error on only two of five parents (section L), and factor A raises error on 13 of the 17 parent–split pairs where the candidate's

Table 4: Input-only selectors on the prototype, IID test, mean over five parents (ratios to the candidate's own fine path; latency and MACs include selector, partition, and dispatch). Quality requires MSE ratio $\leq 1.02$ against both fine references; neither row passes.

| Selector | MSE | MSE/fine | MAC saving | Latency/fine |
|---|---|---|---|---|
| MLP (primary) | 0.007006 | 1.0870 | 9.71% | 1.2378 |
| Linear (control) | 0.007419 | 1.1511 | 19.73% | 0.8968 |

own fine path beats DIRECT-DATA (section J) and at every pretraining budget (section K). Both directions thus hold on the prototype: the arms with no trained fine path (DIRECT) are the most useful, and higher occupancy fidelity does not consistently bring lower continuation error (table 13). GPT-2 shows both directions in a trained model, where in-distribution gain and transfer trade off through the fitting corpus (section N); the recursive model shows them at one boundary with a threefold larger in-distribution margin over an un-adapted model, and a boundary-state gate there adds nothing to the ungated substitute (section O).

**Limits and unexecuted extensions.** The $43.01\%$ oracle saving on the prototype is headroom that no input-only selector there recovers; the failed two-unit-span composition leaves multiscale replacement open. The synthetic study has six data families with a common input distribution, five parents each, one replacement boundary, and batch-one CPU timing; the language-model studies have two models, one substitute class, five and one boundaries, and three and two fitting corpora (sections N and O); no fitted substitute yet matches the unmodified model on every held-out set simultaneously, and the recursive model shares one block with no experts. Together they establish a failure mode and a set of controls, not their prevalence in trained sparse-expert models: no trained mixture-of-experts model was tested, and the only routed setting is the prototype (a protocol for that replication is declared but not executed, section M); CPU timings do not predict GPU throughput. Scale-up, larger expert banks, and horizon-wide controllers remain unexecuted (section S); compression may expose routes to hardware observation (Ding et al., 2025) (section R).

## 8 Conclusion

Whether a macro can replace part of a recursive computation is a question about the computation that follows; short of exact agreement, how closely the macro reproduces what it replaced does not decide it. Under the contract, measured fidelity is neither necessary nor sufficient. Not necessary: DIRECT-COMPUTE, with no trained fine path, beats the candidate's own fine path on 58 of 60 parent–split pairs; on GPT-2 a no-fidelity substitute beats the fidelity optimum in distribution at all five boundaries; on a pretrained recursive model it is below the un-adapted model in distribution. Not sufficient: an occupancy anchor cuts occupancy distance $2.3$–$3.0\times$ without lowering continuation error, and the most faithful arm is among the least accurate. On the prototype no input-only selector passes the quality and cost rules together, and on the recursive model a boundary-state gate adds nothing to the ungated substitute. Learning continuation-sufficient summaries in a real domain is open.

**AI use statement.** In this work, we used generative AI tools (chat assistants and coding agents) to develop the conceptual framework and refine hypotheses; formulate propositions and write and check proofs; design and critique experiments, controls, and synthetic data families; write the research code; run the experiments and compute the reported numbers; interpret results; and translate planning notes into English drafts. Qualitative analysis and data cleaning do not apply. We also used them for literature search, citation checks, figures, drafting, editing, formatting, and review. The authors confirmed the research question and set the goals of each revision. We reviewed the AI-assisted work at each revision, with AI assistance for the detailed checks: AI agents recomputed reported numbers with separate scripts and checked citations against their sources, and independent AI review passes checked the proofs. We take responsibility for the final content of this work, including text, claims, or artifacts produced with the aid of generative AI.

**Reproducibility statement.** Proofs and setup are in sections E, G, J and P; the declared GPT-2 protocol is in section N. The supplement holds the source and protocols, which retrain every run from its seed, and per-run JSON outputs, but not the per-example arrays; without retraining, only analyses of the JSON outputs and the discrete construction rerun. Retraining is deterministic on one platform but not bitwise portable across

platforms: a rerun of the original setting on another runtime reproduces the declared rule statistics, with per-value relative differences of median $7 \times 10^{-7}$ and maximum $13.4\%$ (section K).

## REFERENCES


Mynampati Sri Ranganadha Avinash. Task-conditioned routing signatures in sparse mixture-of-experts transformers, 2026. URL https://arxiv.org/abs/2603.11114v1. Preprint, version 1, 11 March 2026.

Sangmin Bae, Yujin Kim, Reza Bayat, Sungnyun Kim, Jiyoun Ha, Tal Schuster, Adam Fisch, Hrayr Harutyunyan, Ziwei Ji, Aaron Courville, and Se-Young Yun. Mixture-of-recursions: Learning dynamic recursive depths for adaptive token-level computation, 2025. URL https://arxiv.org/abs/2507.10524.

Shaojie Bai, J. Zico Kolter, and Vladlen Koltun. Deep equilibrium models, 2019. URL https://arxiv.org/abs/1909.01377.

Stella Biderman, Hailey Schoelkopf, Quentin Anthony, Herbie Bradley, Kyle O'Brien, Eric Hallahan, Mohammad Aflah Khan, Shivanshu Purohit, USVSN Sai Prashanth, Edward Raff, Aviya Skowron, Lintang Sutawika, and Oskar van der Wal. Pythia: A suite for analyzing large language models across training and scaling, 2023. URL https://arxiv.org/abs/2304.01373.

Kevin Black, Noah Brown, Danny Driess, Adnan Esmail, Michael Equi, Chelsea Finn, Niccolo Fusai, Lachy Groom, Karol Hausman, Brian Ichter, Szymon Jakubczak, Tim Jones, Liyiming Ke, Sergey Levine, Adrian Li-Bell, Mohith Mothukuri, Suraj Nair, Karl Pertsch, Lucy Xiaoyang Shi, James Tanner, Quan Vuong, Anna Walling, Haohuan Wang, and Ury Zhilinsky. $\pi_0$: A vision-language-action flow model for general robot control, 2024. URL https://arxiv.org/abs/2410.24164.

Dan Braun, Jordan Taylor, Nicholas Goldowsky-Dill, and Lee Sharkey. Identifying functionally important features with end-to-end sparse dictionary learning. In *Advances in Neural Information Processing Systems*, volume 37, 2024. URL https://arxiv.org/abs/2405.12241.

Junlin Chen, Ruijie Wang, and Jianxin Li. Counterfactual quotient models: Learning what actions change, not what the world does, 2026a. URL https://arxiv.org/abs/2608.22092v1. Preprint, version 1, 22 August 2026.

Wenkai Chen, Tianshu Li, Wenyong Huang, Yichun Yin, Lifeng Shang, and Chengwei Qin. Loop-MoE: Unifying iterative computation with mixture-of-experts for language modeling, 2026b. URL https://arxiv.org/abs/2606.04438v2. Version 2, 27 August 2026.

Xiaodong Chen, Yuxuan Hu, Jing Zhang, Yanling Wang, Cuiping Li, and Hong Chen. Streamlining redundant layers to compress large language models. In *International Conference on Learning Representations*, 2025. URL https://arxiv.org/abs/2403.19135.

Cheng Chi, Zhenjia Xu, Siyuan Feng, Eric Cousineau, Yilun Du, Benjamin Burchfiel, Russ Tedrake, and Shuran Song. Diffusion policy: Visuomotor policy learning via action diffusion, 2023. URL https://arxiv.org/abs/2303.04137.

Karl Cobbe, Vineet Kosaraju, Mohammad Bavarian, Mark Chen, Heewoo Jun, Lukasz Kaiser, Matthias Plappert, Jerry Tworek, Jacob Hilton, Reiichiro Nakano, Christopher Hesse, and John Schulman. Training verifiers to solve math word problems. *arXiv preprint arXiv:2110.14168*, 2021. URL https://arxiv.org/abs/2110.14168.

Adrián Csiszárik, Péter Kőrösi-Szabó, Ákos K. Matszangosz, Gergely Papp, and Dániel Varga. Similarity and matching of neural network representations. In *Advances in Neural Information Processing Systems*, volume 34, 2021. URL https://arxiv.org/abs/2110.14633.

Mostafa Dehghani, Stephan Gouws, Oriol Vinyals, Jakob Uszkoreit, and Łukasz Kaiser. Universal transformers, 2018. URL https://arxiv.org/abs/1807.03819.

Ruyi Ding, Tianhong Xu, Xinyi Shen, Aidong Adam Ding, and Yunsi Fei. MoEcho: Exploiting side-channel attacks to compromise user privacy in mixture-of-experts LLMs, 2025. URL https://arxiv.org/abs/2508.15036.

Amir-massoud Farahmand, André Barreto, and Daniel Nikovski. Value-aware loss function for model-based reinforcement learning. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, volume 54 of *Proceedings of Machine Learning Research*, pp. 1486–1494, 2017. URL https://proceedings.mlr.press/v54/farahmand17a.html.

William Fedus, Barret Zoph, and Noam Shazeer. Switch transformers: Scaling to trillion parameter models with simple and efficient sparsity, 2021. URL https://arxiv.org/abs/2101.03961.

Kevin Frans, Danijar Hafner, Sergey Levine, and Pieter Abbeel. One step diffusion via shortcut models, 2024. URL https://arxiv.org/abs/2410.12557.

Niklas Funk, Julen Urain, Joao Carvalho, Vignesh Prasad, Georgia Chalvatzaki, and Jan Peters. ActionFlow: Equivariant, accurate, and efficient policies with spatially symmetric flow matching, 2024. URL https://arxiv.org/abs/2409.04576.

Zhengyang Geng, Mingyang Deng, Xingjian Bai, J. Zico Kolter, and Kaiming He. Mean flows for one-step generative modeling, 2025. URL https://arxiv.org/abs/2505.13447.

Christopher Grimm, André Barreto, Satinder Singh, and David Silver. The value equivalence principle for model-based reinforcement learning. In *Advances in Neural Information Processing Systems*, volume 33, 2020. URL https://proceedings.neurips.cc/paper/2020/hash/3bb585ea00014b0e3ebe4c6dd165a358-Abstract.html.

Kazuyuki Hara and Hideitsu Hino. Knowledge distillation under teacher misspecification: An order-parameter analysis of the gap between teacher mimicry and task performance. *arXiv preprint arXiv:2608.29472*, 2026. URL https://arxiv.org/abs/2608.29472.

Haidong Huang, Xixin Zhao, Yaohua Zhou, Jiayu Song, Jiayi Zhang, Jun Ma, Haiyue Zhu, and Xiaocong Li. CoRDE: Concept-prior routed diffusion experts for structural generalization in robot manipulation, 2026a. URL https://arxiv.org/abs/2606.21935.

Jiahe Huang, Sihan Xu, Sharvaree Vadgama, and Rose Yu. Recursive flow matching, 2026b. URL https://arxiv.org/abs/2605.26535.

Ahmadreza Jeddi, Marco Ciccone, and Babak Taati. LoopFormer: Elastic-depth looped transformers for latent reasoning via shortcut modulation. In *International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=RzYXb5YWBs.

Gyeongman Kim, Gyouk Chu, and Eunho Yang. Every expert matters: Towards effective knowledge distillation for mixture-of-experts language models, 2025. URL https://arxiv.org/abs/2502.12947v1. Preprint, version 1, 18 February 2025.

Moo Jin Kim, Karl Pertsch, Siddharth Karamcheti, Ted Xiao, Ashwin Balakrishna, Suraj Nair, Rafael Rafailov, Ethan Foster, Grace Lam, Pannag Sanketi, Quan Vuong, Thomas Kollar, Benjamin Burchfiel, Russ Tedrake, Dorsa Sadigh, Sergey Levine, Percy Liang, and Chelsea Finn. OpenVLA: An open-source vision-language-action model, 2024. URL https://arxiv.org/abs/2406.09246.

Nathan Lambert, Brandon Amos, Omry Yadan, and Roberto Calandra. Objective mismatch in model-based reinforcement learning. In *Proceedings of the 2nd Conference on Learning for Dynamics and Control*, volume 120 of *Proceedings of Machine Learning Research*, pp. 761–770, 2020. URL https://proceedings.mlr.press/v120/lambert20a.html.

Chanhyuk Lee, Jaehoon Yoo, Manan Agarwal, Sheel Shah, Jerry Huang, Aditi Raghunathan, Seunghoon Hong, Nicholas M. Boffi, and Jinwoo Kim. Flow map language models: One-step language modeling via continuous denoising, 2026. URL https://arxiv.org/abs/2602.16813.

Pingzhi Li, Morris Yu-Chao Huang, Zhen Tan, Qingquan Song, Jie Peng, Kai Zou, Yu Cheng, Kaidi Xu, and Tianlong Chen. Leave it to the experts: Detecting knowledge distillation via MoE expert signatures, 2025. URL https://arxiv.org/abs/2510.16968v1. Preprint, version 1, 19 October 2025.

Junchao Lin, Zenan Ling, Jingwen Xu, and Robert C. Qiu. Consistency deep equilibrium models, 2026. URL https://arxiv.org/abs/2602.03024.

Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matt Le. Flow matching for generative modeling, 2022. URL https://arxiv.org/abs/2210.02747.

Bo Liu, Yifeng Zhu, Chongkai Gao, Yihao Feng, Qiang Liu, Yuke Zhu, and Peter Stone. LIBERO: Benchmarking knowledge transfer for lifelong robot learning, 2023. URL https://arxiv.org/abs/2306.03310.

Xingchao Liu, Chengyue Gong, and Qiang Liu. Flow straight and fast: Learning to generate and transfer data with rectified flow, 2022. URL https://arxiv.org/abs/2209.03003.

Zhuang Liu, Mingjie Sun, Tinghui Zhou, Gao Huang, and Trevor Darrell. Rethinking the value of network pruning. In *International Conference on Learning Representations*, 2019. URL https://arxiv.org/abs/1810.05270.

Xudong Lu, Qi Liu, Yuhui Xu, Aojun Zhou, Siyuan Huang, Bo Zhang, Junchi Yan, and Hongsheng Li. Not all experts are equal: Efficient expert pruning and skipping for mixture-of-experts large language models. *arXiv preprint arXiv:2402.14800*, 2024. URL https://arxiv.org/abs/2402.14800.

Stephen Merity, Caiming Xiong, James Bradbury, and Richard Socher. Pointer sentinel mixture models, 2016. URL https://arxiv.org/abs/1609.07843.

Niklas Muennighoff, Luca Soldaini, Dirk Groeneveld, Kyle Lo, Jacob Morrison, Sewon Min, Weijia Shi, Pete Walsh, Oyvind Tafjord, Nathan Lambert, Yuling Gu, Shane Arora, Akshita Bhagia, Dustin Schwenk, David Wadden, Alexander Wettig, Binyuan Hui, Tim Dettmers, Douwe Kiela, Ali Farhadi, Noah A. Smith, Pang Wei Koh, Amanpreet Singh, and Hannaneh Hajishirzi. OLMoE: Open mixture-of-experts language models, 2024. URL https://arxiv.org/abs/2409.02060.

Feng Nan and Venkatesh Saligrama. Dynamic model selection for prediction under a budget, 2017. URL https://arxiv.org/abs/1704.07505.

Octo Model Team, Dibya Ghosh, Homer Walke, Karl Pertsch, Kevin Black, Oier Mees, Sudeep Dasari, Joey Hejna, Tobias Kreiman, Charles Xu, Jianlan Luo, You Liang Tan, Lawrence Yunliang Chen, Pannag Sanketi, Quan Vuong, Ted Xiao, Dorsa Sadigh, Chelsea Finn, and Sergey Levine. Octo: An open-source generalist robot policy, 2024. URL https://arxiv.org/abs/2405.12213.

Denis Paperno, Germán Kruszewski, Angeliki Lazaridou, Quan Ngoc Pham, Raffaella Bernardi, Sandro Pezzelle, Marco Baroni, Gemma Boleda, and Raquel Fernández. The LAMBADA dataset: Word prediction requiring a broad discourse context. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, 2016. URL https://arxiv.org/abs/1606.06031.

Dae Young Park, Moon-Hyun Cha, Changwook Jeong, Daesin Kim, and Bohyung Han. Learning student-friendly teacher networks for knowledge distillation. In *Advances in Neural Information Processing Systems*, volume 34, 2021. URL https://proceedings.neurips.cc/paper/2021/hash/6e7d2da6d3953058db75714ac400b584-Abstract.html.

Guilherme Penedo, Hynek Kydlíček, Loubna Ben allal, Anton Lozhkov, Margaret Mitchell, Colin Raffel, Leandro Von Werra, and Thomas Wolf. The FineWeb datasets: Decanting the web for the finest text data at scale. *arXiv preprint arXiv:2406.17557*, 2024. URL https://arxiv.org/abs/2406.17557.

Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. Language models are unsupervised multitask learners. *OpenAI technical report*, 2019.

David Raposo, Sam Ritter, Blake Richards, Timothy Lillicrap, Peter Conway Humphreys, and Adam Santoro. Mixture-of-depths: Dynamically allocating compute in transformer-based language models. *arXiv preprint arXiv:2404.02258*, 2024. URL https://arxiv.org/abs/2404.02258.

Tim Salimans and Jonathan Ho. Progressive distillation for fast sampling of diffusion models, 2022. URL https://arxiv.org/abs/2202.00512.

Tal Schuster, Adam Fisch, Jai Gupta, Mostafa Dehghani, Dara Bahri, Vinh Q. Tran, Yi Tay, and Donald Metzler. Confident adaptive language modeling. In *Advances in Neural Information Processing Systems*, 2022. URL https://arxiv.org/abs/2207.07061.

Shiv Shankar. Learning shortcut models for efficient recursive reasoning. In *Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics (Volume 4: Student Research Workshop)*, pp. 1351–1360. Association for Computational Linguistics, 2026. URL https://aclanthology.org/2026.acl-srw.120.pdf.

Noam Shazeer, Azalia Mirhoseini, Krzysztof Maziarz, Andy Davis, Quoc Le, Geoffrey Hinton, and Jeff Dean. Outrageously large neural networks: The sparsely-gated mixture-of-experts layer, 2017. URL https://arxiv.org/abs/1701.06538.

Dmitriy Shopkhoev, Ammar Ali, Magauiya Zhussip, Valentin Malykh, Stamatios Lefkimmiatis, Nikos Komodakis, and Sergey Zagoruyko. ReplaceMe: Network simplification via depth pruning and transformer block linearization. *arXiv preprint arXiv:2505.02819*, 2025. URL https://arxiv.org/abs/2505.02819.

Yang Song, Prafulla Dhariwal, Mark Chen, and Ilya Sutskever. Consistency models, 2023. URL https://arxiv.org/abs/2303.01469.

Samuel Stanton, Pavel Izmailov, Polina Kirichenko, Alexander A. Alemi, and Andrew Gordon Wilson. Does knowledge distillation really work? In *Advances in Neural Information Processing Systems*, volume 34, 2021. URL https://proceedings.neurips.cc/paper/2021/hash/376c6b9ff3bedbbea56751a84fffc10c-Abstract.html.

Ayhan Suleymanzade, Chanhyuk Lee, Floor Eijkelboom, Nicholas M. Boffi, İsmail İlkan Ceylan, and Jinwoo Kim. Thinking with looped flows, September 2026. URL https://arxiv.org/abs/2609.11801v1.

Yuwei Sun, Yuxuan Yao, Hui Li, and Siyu Zhu. The thinking pixel: Recursive sparse reasoning in multimodal diffusion latents, 2026. URL https://arxiv.org/abs/2604.25299.

Xiaoqiang Wang, Suyuchen Wang, Yun Zhu, and Bang Liu. System-1.5 reasoning: Traversal in language and latent spaces with dynamic shortcuts. In *Advances in Neural Information Processing Systems*, volume 38, 2025. doi: 10.52202/085713-2191. URL https://papers.nips.cc/paper_files/paper/2025/hash/5e66b9f710a8e38351da690f3806cba9-Abstract-Conference.html.

Yen-Chen Wu, Feng-Ting Liao, Meng-Hsi Chen, Pei-Chen Ho, Farhang Nabiei, and Da shan Shiu. Latent flow transformer, 2025. URL https://arxiv.org/abs/2505.14513v1. Version 1, 20 May 2025.

Canwen Xu, Wangchunshu Zhou, Tao Ge, Furu Wei, and Ming Zhou. BERT-of-Theseus: Compressing BERT by progressive module replacing. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*, 2020. URL https://arxiv.org/abs/2002.02925.

Yufeng Xu, Taiming Lu, Kunjun Li, Jiachen Zhu, Mingjie Sun, and Zhuang Liu. Small LLMs: Pruning vs. training from scratch. *arXiv preprint arXiv:2606.14150*, 2026. URL https://arxiv.org/abs/2606.14150.

Jiayu Yang, Tao Yang, Xiang Chang, Fei Chao, Changjing Shang, and Qiang Shen. PAMAE: Phase-aware-MoE action experts towards reliable flow-matching vision-language-action policies, 2026. URL https://arxiv.org/abs/2606.27144.

Jaehoon Yoo, Wonjung Kim, Floor Eijkelboom, Chanhyuk Lee, Nicholas M. Boffi, Seunghoon Hong, and Jinwoo Kim. Self-conditioned flow map language models via fixed-point flows, 2026. URL https://arxiv.org/abs/2607.00714.

Vincent-Daniel Yun, Junhyuk Jo, Sai Praneeth Karimireddy, and Sunwoo Lee. Ghosted layers: Unconstrained activation alignment for recovering layer-pruned LLMs. *arXiv preprint arXiv:2605.15491*, 2026. URL https://arxiv.org/abs/2605.15491.

Qinglun Zhang, Zhen Liu, Haoqiang Fan, Guanghui Liu, Bing Zeng, and Shuaicheng Liu. FlowPolicy: Enabling fast and robust 3d flow-based policy via consistency flow matching for robot manipulation, 2024. URL https://arxiv.org/abs/2412.04987.

# APPENDIX

The appendix expands proofs, protocols, and supporting experiments; earlier exploratory studies whose results are not in the supplement are summarized in section Q and excluded from the evidence.

## A PROBLEM FORMULATION

### A.1 FINE RECURSIVE CONDITIONAL COMPUTATION

This section defines the signature-supervised formulation (a *signature* is the retained path statistic of the main text), separating its formal interface from the implementations tested later. The eight-expert prototype of the main text (section P) instantiates lowest-order occupancy and a shared continuation, not every extension below. Let $c \in \mathcal{C}$ denote external conditioning, $d, E, K$ be positive integers, and normalized compute time be $s \in [0, 1]$. There are $E$ experts, and the fine path (called the *teacher* in these formal sections) uses $K$ updates of width $h = 1/K$:

$$\begin{aligned} z_{i+1} &= F_i(z_i, c, b_i), \\ \pi_i &= R_i(z_i, c, b_i) \in \Delta^{E-1}, \qquad i = 0, \dots, K-1. \end{aligned} \tag{6}$$

Here $z_i \in \mathcal{Z} \subseteq \mathbb{R}^d$ is the visible hidden, action, or physical state; $\pi_i$ is a soft expert-responsibility vector; and $b_i$ denotes any additional route boundary or history variable needed to make the recursion closed. A memoryless router is the special case in which $b_i$ is absent. A hard top-$k$ implementation, with integer $1 \le k \le E$, additionally produces a normalized multi-hot activation vector $\alpha_i$. If capacity drops are recorded without renormalization, it is a sub-probability vector with dropped mass $d_i = 1 - \mathbf{1}^\top \alpha_i \ge 0$.

When $z$ is a sufficient state, the teacher transition over an interval is

$$\Psi_{s,\Delta}(z; c) = F_{j+m-1} \circ \cdots \circ F_j(z; c), \quad s = jh, \quad \Delta = mh. \tag{7}$$

Here $j \ge 0$ and $m \ge 1$ are integers with $j + m \le K$. Because $F_i$ may depend on $i$, $\Psi$ is a two-parameter process. It satisfies

$$\Psi_{s,\Delta_1+\Delta_2} = \Psi_{s+\Delta_1,\Delta_2} \circ \Psi_{s,\Delta_1}, \tag{8}$$

for positive, grid-aligned $\Delta_1, \Delta_2$ with $s + \Delta_1 + \Delta_2 \le 1$. A one-parameter semigroup is a special autonomous case. If $z$ is not sufficient, equations (7) and (8) apply instead to the augmented state $x = (z, b)$ and its closed updates, including the update of $b$; proposition 2 quantifies why this distinction matters. The continuous idealization below uses valid positive intervals, not an assertion that the discrete teacher is defined between grid points.

### A.2 ROUTE PATHS AND RETAINED SIGNATURES

For a nonempty discrete interval $I$ containing $n \ge 1$ route vectors $p_0, \dots, p_{n-1}$, where $p_i$ can be either $\pi_i$ or $\alpha_i$, define the lowest-order unnormalized mass and step-averaged occupancy

$$A_I = \sum_{i=0}^{n-1} p_i, \qquad Q_I = A_I/n. \tag{9}$$

For soft responsibilities, $Q_I \in \Delta^{E-1}$. With recorded capacity drops it is instead sub-probability, with total mass $1 - D_I/n$ for $D_I$ defined below. Here $A_I$ counts activation mass per step; for width $h$ and duration $\Delta = nh$, the corresponding duration mass is $hA_I = \Delta Q_I$. For a measurable continuous probability-valued responsibility path on $[s, s+\Delta]$, with $\Delta > 0$, the duration-normalized occupancy is

$$Q^\star_{s,\Delta}(z; c) = \frac{1}{\Delta} \int_s^{s+\Delta} \pi_r(z_r, c)\, dr \in \Delta^{E-1}. \tag{10}$$

Occupancy records how much each expert was used but not when. The unnormalized mass is additive; hence normalized occupancy obeys the exact weighted cocycle

$$\begin{aligned} &(\Delta_1 + \Delta_2) Q^\star_{s,\Delta_1+\Delta_2}(z; c) \\ &\quad = \Delta_1 Q^\star_{s,\Delta_1}(z; c) + \Delta_2 Q^\star_{s+\Delta_1,\Delta_2}\big(\Psi_{s,\Delta_1}(z; c); c\big). \end{aligned} \tag{11}$$

The successor state in the second term matters because routing is state dependent.

Occupancy is the minimal signature $S_I^{(0)} = (n, A_I)$. When order matters, we use the optional first-order signature

$$\begin{aligned} S_I^{(1)} &= (n, A_I, B_I, \ell_I, r_I, D_I), \\ B_I &= \sum_{i=0}^{n-2} p_i p_{i+1}^\top, \\ (\ell_I, r_I) &= (p_0, p_{n-1}), \end{aligned} \tag{12}$$

where $D_I = \sum_i (1 - \mathbf{1}^\top p_i)$ records dropped mass. The empty sum gives $B_I = 0$ for a singleton interval. If nonempty adjacent intervals $I$ and $J$ are concatenated,

$$\begin{aligned} n_{IJ} &= n_I + n_J, & A_{IJ} &= A_I + A_J, \\ D_{IJ} &= D_I + D_J, & (\ell_{IJ}, r_{IJ}) &= (\ell_I, r_J), \\ B_{IJ} &= B_I + B_J + r_I \ell_J^\top. \end{aligned} \tag{13}$$

The cross-boundary outer product is indispensable. Define $C_I = B_I/(n-1)$ for $n > 1$, and set $C_I = 0$ for $n = 1$ as a bookkeeping convention; a singleton has no observed adjacent pair. Retaining $n_I$ makes the normalized transition frequency compose as

$$C_{IJ} = \frac{(n_I - 1)C_I + (n_J - 1)C_J + r_I \ell_J^\top}{n_I + n_J - 1}. \tag{14}$$

The prototype and the earlier soft-mixture studies (section Q.1) use $S^{(0)}$ only; the earlier four-expert continuation studies (section Q.3) trained no path-signature head. The algebra of $S^{(1)}$ has executable invariant checks, but no reported implementation trains an ordered-transition signature head. Its learned empirical value remains part of the unexecuted program in section S.

### A.3 Observable model versus product-state closure

There are two mathematically distinct targets. If $z_s$ is sufficient, a coarse interval model produces an endpoint and a path observable,

$$(z_s, c, s, \Delta) \longmapsto \big(z_{s+\Delta}, S_{[s,s+\Delta]}\big). \tag{15}$$

The signature summarizes what happened inside the interval; it is not a second dynamic state. If boundary route state $b_s$ carries predictive information beyond $z_s$, closure requires

$$(z_s, b_s, c, s, \Delta) \longmapsto \big(z_{s+\Delta}, b_{s+\Delta}, S_{[s,s+\Delta]}\big). \tag{16}$$

Only equation (16) justifies the phrase *product-state dynamics*. This prevents a conceptual overclaim in which an interval average is treated as if it were automatically Markovian. Feeding a predicted occupancy to a continuation router, as the prototype does, defines a particular learned interface; it does not establish closure for arbitrary incoming routing history or arbitrary replacement schedules.

### A.4 Compression target

Given teacher states $\{z_i\}_{i=0}^K$ and per-update routes $\{\pi_i, \alpha_i\}_{i=0}^{K-1}$, learn one parameter-shared family that accepts admissible $(s, \Delta)$ and predicts

$$\begin{aligned} \widehat{z}_{s+\Delta} &= \Phi_\theta(z_s, c, s, \Delta), \\ \widehat{S}_{s,\Delta} &= \widehat{S}_\theta(z_s, c, s, \Delta). \end{aligned} \tag{17}$$

A distinct execution route $\rho_\theta$ chooses which coarse student experts are evaluated. In this formulation, a positive integer deployment budget $N$ specifies an externally chosen partition into $N$ positive intervals whose lengths sum to one. The discrete-teacher experiments require grid-aligned endpoints, hence $N \le K$. This is not a learned budget controller. The target is a well-behaved family across admissible budgets, not zero error at one schedule. When route memory is required, the model input and output must be augmented as in equation (16). Fewer model evaluations are an algorithmic quantity; wall-clock improvement requires end-to-end measured execution, including dispatch when sparse. The prototype's separately fitted selector chooses between two fixed complete paths at one boundary; it does not learn the arbitrary partition controller described by this broader goal.

### A.5 THREE TIMES AND THREE CAPACITIES

In robot-policy language, three clocks must not be conflated: physical control time, generative-flow time within one action prediction, and internal recursive depth. The formulation compresses whichever recursive coordinate is declared; it does not automatically accelerate the other two. Likewise, physical parameter capacity, active conditional capacity per update, and effective depth from weight reuse are distinct. Reusing a block does not create stored parameters, and unselected experts are not computationally free on every hardware stack.

Table 5: Core notation. Every symbol is defined at first substantive use; this table is a lookup aid.

| Symbol | Meaning |
|---|---|
| $c$; $s, \Delta$ | external context; compute start and interval length |
| $z_s, b_s$ | visible state; optional route/history state needed for closure |
| $E$; $\pi_s$; $\alpha_s$; $d_s$ | expert count; soft responsibilities; hard activations; dropped mass |
| $\Psi_{s,\Delta}$; $\Phi_\theta$ | fine teacher transition; learned coarse state map |
| $A_I, Q_I$ | unnormalized route mass; normalized occupancy |
| $B_I, C_I, \ell_I, r_I$ | adjacent-transition mass/frequency; first/last route |
| $\widehat{S}_\theta$; $\rho_\theta$ | predicted teacher-path signature; actual student execution route |
| $\delta_z, \delta_q$ | state-composition and occupancy-cocycle defects |
| $N, k$ | deployed model evaluations; active experts per coarse evaluation |

## B THE SIGNATURE-SUPERVISED FORMULATION

This section specifies the signature-supervised formulation; the eight-expert prototype (section P) implements its lowest-order occupancy together with a shared continuation. The earlier soft-mixture implementations and four-expert continuation studies are summarized, and excluded from the evidence, in sections Q.1 and Q.3. The prototype's selected-only timings say nothing about sparse execution of the soft-mixture models.

### B.1 ONE OPERATOR, TWO ROUTING ROLES

Let $u_\theta = \mathrm{Enc}_\theta(z, c, e_s(s), e_\Delta(\Delta))$ combine state, context, start-time encoding, and interval-size encoding. The model has a *signature head*, which predicts the teacher path statistic to be preserved, and an *execution router*, which determines the student computation:

$$\widehat{S}_\theta = G_\theta^{\mathrm{sig}}(u_\theta), \tag{18}$$

$$\rho_\theta = H_\theta(u_\theta, \widehat{S}_\theta), \qquad v_\theta = \sum_{e=1}^{E} \rho_{\theta,e} V_{\theta,e}(u_\theta), \tag{19}$$

$$\Phi_\theta(z, c, s, \Delta) = z + \Delta v_\theta. \tag{20}$$

$\widehat{S}_\theta$ predicts what the fine teacher did inside the skipped interval; $\rho_\theta$ specifies what the coarse student actually executes. They can share features without being equated. Sparse deployment would require $H_\theta$ to implement top-$k$ dispatch and capacity handling. The earlier soft-mixture studies (section Q.1) use a tied soft special case,

$$\widehat{S}_\theta = \widehat{Q}_\theta = \rho_\theta = \mathrm{softmax}(G_\theta(u_\theta)/\tau), \tag{21}$$

with temperature $\tau > 0$, and evaluate all experts. They therefore test identification and recursive-evaluation compression, not sparse-system acceleration. A direct endpoint, distributional head, or action chunk may replace the residual parameterization in equation (20); these alternatives do not describe additional trained models here.

### B.2 TEACHER ANCHORS AT MULTIPLE SCALES

For equal-dimensional vectors, write $\mathrm{mse}(v, w) = \|v - w\|_2^2 / \dim(v)$ for mean squared coordinate error. For a sampled teacher interval $I = (z_s, c, s, \Delta)$, the endpoint and lowest-order signature

losses are

$$\mathcal{L}_{\text{jump}} = \mathbb{E}_I[\text{mse}(\Phi_\theta(I), \Psi_{s,\Delta}(z_s; c))], \tag{22}$$

$$\mathcal{L}_{Q\text{-anchor}} = \mathbb{E}_I\left[D_A\left(\widehat{Q}_\theta(I), Q^\star_{s,\Delta}(z_s; c)\right)\right]. \tag{23}$$

$D_A$ is an explicitly selected anchor discrepancy with prediction first and target second. In the earlier soft-mixture studies it is the stabilized teacher-target cross-entropy $D_A(q,p) = -\sum_e p_e \log(q_e + 10^{-8})$; route composition below instead uses coordinate MSE. Other specified discrepancies are possible, but not interchangeable numerical implementations. If student experts inherit teacher identities, coordinatewise comparison supplies a declared convention. If identities are learned independently, one global matching must be fitted on disjoint calibration data; per-example rematching is invalid.

For a partition $\Delta = \Delta_1 + \Delta_2$, define

$$\widehat{z}_1 = \Phi_\theta(z, c, s, \Delta_1), \tag{24}$$

$$\widehat{z}^{\text{comp}}_{12} = \Phi_\theta(\widehat{z}_1, c, s + \Delta_1, \Delta_2), \tag{25}$$

$$\widehat{Q}^{\text{comp}}_{12} = \frac{\Delta_1}{\Delta}\widehat{Q}_\theta(z, c, s, \Delta_1) + \frac{\Delta_2}{\Delta}\widehat{Q}_\theta(\widehat{z}_1, c, s + \Delta_1, \Delta_2). \tag{26}$$

The cross-scale losses are

$$\mathcal{L}_{\text{state-comp}} = \mathbb{E}\,\text{mse}(\Phi_\theta(z, c, s, \Delta), \widehat{z}^{\text{comp}}_{12}), \tag{27}$$

$$\mathcal{L}_{Q\text{-cocycle}} = \mathbb{E}D_C\left(\widehat{Q}_\theta(z, c, s, \Delta), \widehat{Q}^{\text{comp}}_{12}\right), \tag{28}$$

where $D_C = \text{mse}$ in the earlier soft-mixture studies. A stop-gradient on the composed target changes the optimization path rather than the desired fixed-point identity; its stability benefit is not guaranteed. For a fixed teacher-fidelity target, cocycle consistency alone cannot select the intended occupancy, because it also admits a constant, uninformative occupancy. Teacher anchoring supplies that target in this construction. This is not a claim that teacher-path anchoring is necessary for every task-useful, jointly adapted macro system. The empirical coordinate-MSE averages in equations (27) and (28) are not the pointwise norm and total-variation defects assumed in propositions 3 and 5. Small sampled penalties do not establish uniform bounds on all relevant true and predicted states.

For the ordered signature in equation (12), $G^{\text{sig}}_\theta$ additionally predicts $(\widehat{B}, \widehat{\ell}, \widehat{r}, \widehat{D})$, and its anchor and composition losses use equation (13) exactly. In particular, the transition target must include $\widehat{r}_I\widehat{\ell}_J^\top$; a weighted average of $C_I$ and $C_J$ is not correct. Because this head has not yet been trained in the reported experiments, we give the equations and executable invariant tests but make no empirical claim for it.

### B.3 Total objective and model controls

The lowest-order signature objective is

$$\begin{aligned}\mathcal{L} = {} & \lambda_J\mathcal{L}_{\text{jump}} + \lambda_Q\mathcal{L}_{Q\text{-anchor}} \\ & + \lambda_S\mathcal{L}_{\text{state-comp}} + \lambda_C\mathcal{L}_{Q\text{-cocycle}} \\ & + \lambda_T\mathcal{L}_{\text{task}} + \lambda_B\mathcal{L}_{\text{balance}} + \lambda_X\mathcal{L}_{\text{execution}}.\end{aligned} \tag{29}$$

All weights $\lambda$ are nonnegative. $\mathcal{L}_{\text{task}}$ is optional downstream supervision; $\mathcal{L}_{\text{balance}}$ controls load or capacity; and $\mathcal{L}_{\text{execution}}$ aligns the actual coarse route with the chosen signature-dependent dispatch policy when they are not tied. Load balance, teacher fidelity, and sparse execution are different objectives and are reported separately. The earlier soft-mixture studies set these three optional weights to zero.

Comparisons must support the particular gain being claimed. We distinguish: (i) equal full-parameter state-only and joint models; (ii) an equal active-FLOP control; (iii) a dense capacity

control; and (iv) equal-data and equal-total-training-compute tracks. The earlier soft-mixture studies provided only (i) and a smaller dense reference; the other controls belong to the unexecuted extensions in section S. The prototype instead includes equal-data and affine-forward-MAC-budget DIRECT controls, not a fully matched backward/optimizer/training-cost track. Complete-system utility and the effect of a particular auxiliary feedback remain different estimands.

In the prototype, the signature prediction is not tied to the one continuation expert actually executed. Its occupancy anchor uses coordinate MSE against the current fine path with stop-gradient on that target; it is not the fixed-teacher cross-entropy anchor above. State/occupancy composition and response penalties differentiate both operands. These implementation changes are specified in section P, rather than silently treated as the same numerical objective as equation (29).

### B.4 Training curriculum

The earlier soft-mixture studies trained on the valid dyadic interval lengths $h, 2h, 4h, \ldots$ jointly from the start; they did not execute a progressive warm-start curriculum. A possible future curriculum is:

1. collect fine traces $(z_i, \pi_i, \alpha_i, d_i)$ and choose the smallest signature justified by the task;
2. sample intervals and compute exact endpoint and signature targets;
3. warm-start small intervals with task, jump, and signature anchors;
4. introduce larger intervals and the matching state/signature composition losses;
5. calibrate actual top-$k$ dispatch and profile fixed inference budgets on target hardware.

For such a future comparison, signature complexity and curriculum must be selected on development data and frozen before final evaluation. Adding higher-order statistics after inspecting test failures would invalidate the confirmatory comparison. The prototype uses a different schedule: fine pretraining followed by joint adaptation of span-2 and composed span-1 macros. Only span 2 is fine-anchored; the unit-span maps are factorization auxiliaries, not calibrated one-step teacher maps. Its risk selector is trained only after those blocks are frozen. Neither schedule implements generative flow matching.

### B.5 Why the formulation is domain-general

Only the compressed coordinate changes across domains; the interface of state, context, compute time, interval length, endpoint, retained path statistic, and execution route is shared. In language models, $z$ is a residual-stream state and $s$ is loop depth. In robot action generation, $z$ is a noisy action chunk or policy latent and $s$ is generative time. In scientific computing, $z$ is a discretized field and $s$ is solver time or resolution. In a recurrent world model, $z$ is a latent belief and $s$ is rollout depth. Universality of performance, architecture, or optimal signature is not claimed.

## C Analysis

We state finite, conditional claims about the signature-supervised formulation. The following elementary algebra, information limits and perturbation bounds delimit its scope; they are not claimed as individually new general theorems. Proofs appear in section E. State-space norms are fixed but arbitrary; state Lipschitz constants use that norm, whereas routing total variation explicitly uses $\ell_1$. Error bounds and Lipschitz constants are nonnegative and finite. Every compared teacher or learned map must be defined on the relevant true and predicted states. The defects $\delta_z$ and $\delta_q$ below are pointwise bounds in the stated state norm and routing total variation, respectively, not empirical mean squared coordinate errors. A finite-sample mean composition loss does not by itself supply these bounds or the uniform approximation assumptions.

### C.1 Computation-space dynamics and coarse-grained closure

Let $r$ denote a continuous computation coordinate. The dynamical-system interpretation does not require it to be physical time: the underlying discrete coordinate may index layer reuse, reasoning depth, denoising progress, solver resolution, or recurrent rollout. Suppose an absolutely continuous

idealization of the fine computation has velocity almost everywhere

$$\dot{z}_r = \sum_{e=1}^{E} \pi_e(r) f_e(z_r, c, r). \tag{30}$$

Assume $\pi(r)$ is measurable and probability-valued, and every weighted field $\pi_e(r)f_e(z_r, c, r)$ is integrable over a valid interval of positive length $\Delta$. Its exact interval increment can be written

$$\begin{aligned} z_{s+\Delta} - z_s &= \Delta \sum_{e=1}^{E} Q_e^\star \, \bar{f}_e, \\ \bar{f}_e &= \frac{\int_s^{s+\Delta} \pi_e(r) f_e(z_r, c, r)\, dr}{\int_s^{s+\Delta} \pi_e(r)\, dr}, \end{aligned} \tag{31}$$

with $\bar{f}_e = 0$ when its denominator is zero. The occupancy $Q_e^\star$ is therefore analogous to a phase fraction, while $\bar{f}_e$ is that phase's mean impulse. The analogy is useful but limited: occupancy alone does not determine the impulse, expert updates may not commute, and eliminating fine states can introduce memory. The learned map $\Phi_\theta(z, c, s, \Delta)$ is an effective coarse operator, a closure model for those unresolved fine-scale effects. Predicting $\widehat{S}_\theta$ exposes structured information about the eliminated path instead of forcing the endpoint to absorb everything. It becomes a dynamic variable only if propagated as part of a sufficient augmented state. In this bounded sense, the signature-supervised formulation performs learned multiscale coarse-graining of computational dynamics.

**Proposition 2** (Closure lower bound for omitted route memory)**.** *Fix visible input $(z, c, s, \Delta)$. Suppose two admissible hidden boundary states $b$ and $b'$ produce endpoints $y$ and $y'$ with $\|y - y'\| = d_{\mathrm{sep}} > 0$. Then every deterministic state-only coarse map $\chi(z, c, s, \Delta)$ has worst-case error at least $d_{\mathrm{sep}}/2$ over these two continuations.*

This is an information limit rather than an optimization failure. At the fixed visible input and interval, exact deterministic prediction requires the endpoint to be constant across admissible hidden boundary states. Propagating a sufficient $b$ in equation (16) is one remedy. A distributional predictor can represent unresolved uncertainty but does not identify the realized hidden continuation; predicting a non-propagated occupancy does not, by itself, remove this ambiguity.

**Proposition 3** (One partition: state error)**.** *Suppose the teacher obeys equation* (8) *on valid positive intervals. Let the student's state-composition defect at $(z, s, \Delta_1, \Delta_2)$ be at most $\delta_z$. Suppose its two subinterval approximation errors are at most $\varepsilon_1$ and $\varepsilon_2$ uniformly on the relevant states, including the predicted first successor, and the second teacher map is $L_2$-Lipschitz there. Then*

$$\begin{aligned} &\|\Phi_{\theta;s,\Delta_1+\Delta_2}(z) - \Psi_{s,\Delta_1+\Delta_2}(z)\| \\ &\quad \le \delta_z + \varepsilon_2 + L_2\varepsilon_1. \end{aligned} \tag{32}$$

Thus composition consistency connects errors across resolutions, but it does not select a designated teacher: $\delta_z = 0$ is compatible with a consistently wrong family of maps for that fidelity target.

**Proposition 4** (Sequential accumulation)**.** *For a positive integer $N$ and $N$ valid positive intervals, start teacher and student at the same state. If every learned local map has uniform teacher error at most $\varepsilon$ on the reached states and every relevant teacher map is $L$-Lipschitz there, the deployed endpoint error is at most*

$$\varepsilon \sum_{j=0}^{N-1} L^j = \begin{cases} \varepsilon N, & L = 1, \\ \varepsilon(L^N - 1)/(L - 1), & L \neq 1. \end{cases} \tag{33}$$

The bound explains why local accuracy may fail under long recursion when $L > 1$, and why direct coarse jumps can be attractive even if a local model is accurate.

**Proposition 5** (Routing occupancy error)**.** *For probability-valued occupancies, let $D(q, p) = \frac{1}{2}\|q - p\|_1$. Let $\Delta = \Delta_1 + \Delta_2$ for valid positive intervals and let the learned route cocycle defect be at most $\delta_q$. Suppose subinterval route errors are at most $\eta_1, \eta_2$ when evaluated at the true states, and the second-interval route predictor is $K_q$-Lipschitz between the true and predicted successor states under $D$. If the first predicted state has error at most $\varepsilon_1$, then*

$$D(Q_{\theta;s,\Delta}, Q^\star_{s,\Delta}) \le \delta_q + \frac{\Delta_1\eta_1 + \Delta_2(\eta_2 + K_q\varepsilon_1)}{\Delta_1 + \Delta_2}. \tag{34}$$

State and routing errors are therefore coupled even though their concatenation laws differ.

**Lemma 1** (Endpoint-only routing is not identifiable)**.** *At a fixed input, suppose the expert outputs form columns of $V \in \mathbb{R}^{d\times E}$ and let $q$ lie in the relative interior of the probability simplex, so every $q_e > 0$. If the affine nullspace $\{a : Va = 0, \mathbf{1}^\top a = 0\}$ is nontrivial (guaranteed when $E > d + 1$), then infinitely many distinct routing vectors produce the same mixture output $Vq$ in a neighborhood of $q$.*

This algebraic non-uniqueness explains why endpoint loss cannot, in general, recover designated teacher responsibilities. Route consistency alone likewise need not select the intended teacher route. These are teacher-identification statements, not requirements that a useful task solution have a unique expert decomposition.

**Proposition 6** (Occupancy cannot recover expert order)**.** *There exist two expert sequences with identical occupancy and different endpoints. In particular, for scalar maps $A(z) = 2z$ and $B(z) = z+1$, the paths $(A, B)$ and $(B, A)$ both have occupancy $(1/2, 1/2)$, but produce $B(A(z)) = 2z+1$ and $A(B(z)) = 2z + 2$.*

Thus $S^{(0)}$ is not a sufficient statistic for noncommuting computation. The boundary-corrected transition statistic in $S^{(1)}$ distinguishes these two paths, although even a first-order signature need not identify arbitrary long operator words.

**Proposition 7** (Top-$k$ stability)**.** *Let $E$ be a positive integer, let integer $1 \le k \le E$, and let $g, \widehat{g} \in \mathbb{R}^E$ be teacher and student routing logits. If $k < E$, suppose the teacher margin between its $k$th and $(k + 1)$st largest logits is $\gamma > 0$ and $\|\widehat{g} - g\|_\infty < \gamma/2$; then their pre-capacity top-$k$ expert sets coincide. For $k = E$, all experts are selected and no boundary margin is needed.*

Occupancy agreement alone does not imply this pointwise condition. A model may match interval fractions while permuting their temporal order. Equality at the perturbation threshold does not guarantee a strict ranking, and subsequent capacity drops require a separate dispatch analysis.

**Proposition 8** (Sparse-support fidelity floor)**.** *Let $E, N$ be positive integers and integer $1 \le k \le E$. Suppose a teacher occupancy $q \in \Delta^{E-1}$ has entries sorted as $q_{(1)} \ge \cdots \ge q_{(E)}$. A student using $N$ macrosteps with top-$k$ routing can activate at most $m = \min(E, Nk)$ distinct experts. For a common, aligned expert index set, among all probability vectors $p$ supported on at most $m$ experts,*

$$\min_{p\in\Delta^{E-1}:\,|\mathrm{supp}(p)|\le m} \mathrm{TV}(p, q) = 1 - \sum_{e=1}^{m} q_{(e)}. \tag{35}$$

This is a lower bound for actual execution occupancy, not for a signature head that may predict all $E$ coordinates while activating fewer experts. The optimization relaxes further constraints on occupancies attainable by particular routing weights or schedules; its minimizer need not be executable by every student. Exact execution-occupancy fidelity may be impossible at an aggressive $(N, k)$ budget even when the state endpoint remains accurate. Capacity drops further require comparing sub-probability activation mass and reporting overflow, rather than silently applying the probability-normalized formula to dropped mass.

**Proposition 9** (Effect of route error on a mixture update)**.** *At a common state, let every expert update satisfy $\|f_e\| \le M$. For responsibilities $p, q \in \Delta^{E-1}$,*

$$\|\textstyle\sum_e (q_e - p_e) f_e\| \le 2M\,\mathrm{TV}(q, p). \tag{36}$$

*Over an interval of positive length $\Delta$, with a common state/field evaluation path and uniform bound $M$, the norm of the integrated discrepancy contributed by routing alone is at most $2M\Delta$ times the mean TV error.*

This bound is the perturbative counterpart of equation (31). It does not include subsequent state divergence or altered expert functions. The mixture interpretation is not a claim that neural routing obeys a physical conservation law.

### C.2 Expert identity and permutation

Tied expert functions supply a shared coordinate interpretation; teacher initialization supplies a labeling convention but does not alone guarantee persistent functional identity after independent

training. An unanchored expert bank generally admits label permutations, without implying that permutation is its only possible non-identifiability. Let $\mathcal{P}_E$ be the set of permutation matrices and let $\mathcal{D}_{\mathrm{cal}}$ be a disjoint, nonempty calibration set. A consistent evaluation fixes

$$P^\star \in \arg\min_{P\in\mathcal{P}_E} \frac{1}{|\mathcal{D}_{\mathrm{cal}}|} \sum_{i\in\mathcal{D}_{\mathrm{cal}}} D(q_i, Pp_i),$$
$$D_{\mathrm{aligned}}(q,p) = D(q, P^\star p). \tag{37}$$

The calibration objective and deterministic tie rule must be recorded; a confusion-based assignment may use a different stated calibration objective. $P^\star$ remains fixed across test examples and schedules. Minimizing separately for each test pair would measure a different, more permissive object. In the prototype, the macro's occupancy head and the fine path index the same shared expert bank, so their coordinates coincide and no alignment is fitted. Functional equivalence under a continuation class, discussed next, is distinct from relabeling experts.

# D Continuation-Valid Replacement and Joint Design

The preceding path-signature analysis concerns what a coarse operator records about a fine program. A complementary question is what information its output must retain for *subsequent computation*. These questions need not have the same answer: a useful replacement need not reproduce a unique internal path, and an accurate path signature does not alone certify useful continuation. We formalize this distinction using a fixed-horizon interface. The eight-expert prototype (section P) combines this interface with predicted lowest-order occupancy; the integration does not include ordered-transition learning, arbitrary-horizon control, or all capabilities of the broader signature formulation.

## D.1 A common continuation interface

Let $x$ be an input, $y$ its evaluation target, $\vartheta$ the shared encoder, expert-bank, router, and readout parameters, and $\varphi$ the macro parameters. These parameters are distinct from the signature student's $\theta$ above. With $u_0 = \mathrm{Enc}_\vartheta(x)$, define

$$u^F = \mathcal{F}_\vartheta(u_0), \qquad u^M = \mathcal{C}_\varphi(u_0). \tag{38}$$

The fine program $\mathcal{F}_\vartheta$ reroutes between its internal steps; the macro $\mathcal{C}_\varphi$ replaces this prefix. Both end at the same declared logical stage. For a common finite, nonempty menu $\mathcal{A}$ of deployable continuation actions, let $o_j(a)$ be the output of applying action $a$ and the shared readout to $u^j$, for $j \in \{F, M\}$. The router selects $a_j$ using its own reached state, not $y$ or all candidate outputs. In the prototype, $\mathcal{A}$ contains eight physical experts, and the continuation router receives a hidden state plus interval occupancy. For that restricted interface, $u^j$ denotes the pair; this notation does not assert that occupancy is sufficient for arbitrary future computations.

For a fixed input–target pair and frozen parameter version, write

$$\ell_j(a) = \tfrac{1}{2}\,\|o_j(a) - y\|_2^2\,,$$
$$m_j = \min_{a\in\mathcal{A}} \ell_j(a), \qquad \mathfrak{r}_j = \ell_j(a_j) - m_j \geq 0. \tag{39}$$

Here $m_j$ is an offline oracle diagnostic, not a target-aware deployment policy; $\mathfrak{r}_j$ measures the actual router's excess loss over that menu. Comparing corresponding actions requires shared action semantics. Independently learned banks require a declared alignment or another explicitly defined comparison, not per-test-example rematching. More general equivalence notions likewise depend on a specified policy or query class, rather than arbitrary futures (Grimm et al., 2020; Chen et al., 2026a).

**Proposition 10** (Oracle-menu shift under uniform output fidelity)**.** *Under the finite-menu definitions above, with finite losses and finite nonnegative bounds $\varepsilon_\ell, \varepsilon_o, B_{\mathrm{res}}$,*

$$\ell_{\mathrm{M}}(a_{\mathrm{M}}) - \ell_{\mathrm{F}}(a_{\mathrm{F}}) = \big[\ell_{\mathrm{M}}(a_{\mathrm{F}}) - \ell_{\mathrm{F}}(a_{\mathrm{F}})\big] + \big[\ell_{\mathrm{M}}(a_{\mathrm{M}}) - \ell_{\mathrm{M}}(a_{\mathrm{F}})\big] \tag{40}$$
$$= m_{\mathrm{M}} - m_{\mathrm{F}} + \mathfrak{r}_{\mathrm{M}} - \mathfrak{r}_{\mathrm{F}}. \tag{41}$$

*If* $\max_{a\in\mathcal{A}} |\ell_{\mathrm{M}}(a) - \ell_{\mathrm{F}}(a)| \le \varepsilon_\ell$*, then*

$$|\ell_{\mathrm{M}}(a_{\mathrm{M}}) - \ell_{\mathrm{F}}(a_{\mathrm{F}}) - (\mathfrak{r}_{\mathrm{M}} - \mathfrak{r}_{\mathrm{F}})| \le \varepsilon_\ell. \tag{42}$$

*In particular, if for every* $a \in \mathcal{A}$*,* $\|o_{\mathrm{M}}(a) - o_{\mathrm{F}}(a)\|_2 \le \varepsilon_o$ *and* $\|o_{\mathrm{F}}(a) - y\|_2 \le B_{\mathrm{res}}$*, one may take* $\varepsilon_\ell = B_{\mathrm{res}}\varepsilon_o + \varepsilon_o^2/2$*. Moreover, the fixed-action term satisfies* $|\ell_{\mathrm{M}}(a_{\mathrm{F}}) - \ell_{\mathrm{F}}(a_{\mathrm{F}})| \le \varepsilon_\ell$*, and if in addition* $a_{\mathrm{M}} = a_{\mathrm{F}}$ *then* $|\ell_{\mathrm{M}}(a_{\mathrm{M}}) - \ell_{\mathrm{F}}(a_{\mathrm{F}})| \le \varepsilon_\ell$*; exact boundary agreement gives* $\varepsilon_o = 0$ *and* $a_{\mathrm{M}} = a_{\mathrm{F}}$*, hence zero gap. Consequently* $|\ell_{\mathrm{M}}(a_{\mathrm{M}}) - \ell_{\mathrm{M}}(a_{\mathrm{F}}) - (\mathfrak{r}_{\mathrm{M}} - \mathfrak{r}_{\mathrm{F}})| \le 2\varepsilon_\ell$*: uniform fidelity ties the reallocation term to the regret difference but does not control it.*

*Proof.* Adding and subtracting $\ell_{\mathrm{M}}(a_{\mathrm{F}})$ gives equation (40); substituting the two regret definitions gives equation (41). Testing a fine minimizer in the macro menu gives $m_{\mathrm{M}} \le m_{\mathrm{F}} + \varepsilon_\ell$, and exchanging the two menus gives the reverse inequality. Finally, with $e_a = o_{\mathrm{M}}(a) - o_{\mathrm{F}}(a)$,

$$\ell_{\mathrm{M}}(a) - \ell_{\mathrm{F}}(a) = \langle o_{\mathrm{F}}(a) - y, e_a\rangle + \tfrac{1}{2}\|e_a\|_2^2 .$$

Cauchy–Schwarz yields the sufficient bound. The fixed-action bound is the premise evaluated at $a_{\mathrm{F}}$; with $a_{\mathrm{M}} = a_{\mathrm{F}}$ the reallocation term vanishes. Since the reallocation term is the gap minus the fixed-action term, the triangle inequality with equation (42) and the fixed-action bound gives the $2\varepsilon_\ell$ bound. □

The split in equation (40) takes the fine action as reference; the mirror split $[\ell_{\mathrm{M}}(a_{\mathrm{M}}) - \ell_{\mathrm{F}}(a_{\mathrm{M}})] + [\ell_{\mathrm{F}}(a_{\mathrm{M}}) - \ell_{\mathrm{F}}(a_{\mathrm{F}})]$ is equally exact, so comparisons between the two terms (table 2) refer to the fine-action reference.

These are elementary conditional identities and a standard sufficient bound, not a claimed new theorem. They separate replacement error from decision quality: the second term in equation (40) may be negative, so changing the selected expert can compensate for replacement. A suitable router-logit margin can guarantee identity preservation, but equal identities are neither necessary nor, by themselves, sufficient for task-quality preservation. A minibatch mean response loss neither supplies the uniform premise nor controls unseen inputs, untested actions, or the learned router's regret. Repeated replacement requires additional coverage and error-propagation assumptions; no arbitrary-horizon or arbitrary-budget guarantee follows.

If cost is part of the decision, the same algebra applies to $\ell_j(a) + \lambda_c\kappa_j(a)$ for declared nonnegative cost $\kappa_j$ and conversion weight $\lambda_c$. A task-loss bound alone does not bound this augmented quantity. Operation counts and measured latency are distinct choices of cost, and training or profiling all candidates cannot be omitted when reporting their respective costs.

### D.2 Functional non-uniqueness is not feedback redundancy

Different states, expert orders, or selections can realize the same useful continuation. However, equality for a fixed continuation class need not survive an expanded action menu, altered budget, or changed task. Exact matching is also not a quality certificate: if every corresponding output equals an arbitrarily bad prediction, response discrepancy is zero while task loss remains arbitrarily large. Conversely, both routers can select a correct action while unused candidate outputs differ arbitrarily. Thus full candidate-response matching is neither necessary for preserving the realized task result nor sufficient for making that result useful.

The prototype's response term compares corresponding candidate outputs, uniformly or with task-relevance weights, without router-logit alignment. It permits changed fine paths but does not implement invariance to every functionally equivalent expert decomposition, and its retained endpoint, occupancy, and composition penalties do not implement arbitrary functional equivalence either. Nor does functional non-uniqueness show that different auxiliary gradient paths compensate for one another in training; a causal explanation of that kind must match an intervention that was performed.

### D.3 What joint design can, and cannot, establish

The proposed organizational object is a coupled fine program, executable macro, and shared continuation whose parameters can adapt together. Its potential value is that compressibility and useful expert organization can influence each other. Merely enlarging a fixed-teacher optimization feasible

set proves only that its ideal minimum cannot increase; this set-inclusion fact is not a new theorem or an optimization guarantee. Moreover, the DIRECT control has the same fast deployment computation graph. The candidate therefore has no larger deployed function family than an unrestricted DIRECT optimizer. Any advantage over that control must be established in finite-budget learning, data efficiency, robustness, or measured quality–cost tradeoffs, not asserted as an expressivity separation.

The prototype does not establish such an advantage. Its predeclared complete-system comparison (section P) integrates occupancy and continuation learning, then evaluates a separately trained frozen-bank selector, and fails the declared utility screens. This is evidence about the complete tested arms, not a requirement that one feedback edge always win. The selector is not jointly trained with the expert bank and macro, and no flow-matching objective is present. The prototype covers six synthetic families (section J) and one-step continuation; it establishes neither ordered-signature learning nor arbitrary-horizon, arbitrary-budget, or cross-domain validity.

# E PROOFS

For compactness, context $c$ and signature-model parameters $\theta$ are suppressed. These proofs concern the conditional statements in the signature analysis, not optimization success or new continuation-model guarantees. Intervals are nonempty and valid; uniform approximation assumptions cover the predicted states used below.

*Proof of proposition 2.* Let $\widehat{y} = \chi(z, c, s, \Delta)$. The triangle inequality gives

$$d_{\mathrm{sep}} = \|y - y'\| \le \|y - \widehat{y}\| + \|\widehat{y} - y'\|\,.$$

Consequently at least one of the two errors is at least $d_{\mathrm{sep}}/2$. □

*Proof of proposition 3.* Let $\widehat{z}_1 = \Phi_{\theta;s,\Delta_1}(z)$ and $z_1^\star = \Psi_{s,\Delta_1}(z)$. Add and subtract the learned two-step composition and use the teacher process law:

$$\begin{aligned}
&\|\Phi_{\theta;s,\Delta_1+\Delta_2}(z) - \Psi_{s,\Delta_1+\Delta_2}(z)\| \\
&\le \|\Phi_{\theta;s,\Delta_1+\Delta_2}(z) - \Phi_{\theta;s+\Delta_1,\Delta_2}(\widehat{z}_1)\| \\
&\quad + \|\Phi_{\theta;s+\Delta_1,\Delta_2}(\widehat{z}_1) - \Psi_{s+\Delta_1,\Delta_2}(\widehat{z}_1)\| \\
&\quad + \|\Psi_{s+\Delta_1,\Delta_2}(\widehat{z}_1) - \Psi_{s+\Delta_1,\Delta_2}(z_1^\star)\| \\
&\le \delta_z + \varepsilon_2 + L_2\varepsilon_1.
\end{aligned}$$

The first term is the student composition defect, the second uses the uniform second-interval approximation assumption, and the third uses Lipschitz continuity and the first-interval error bound. □

*Proof of proposition 4.* Let $e_j$ be the error after $j$ deployed intervals. One learned update contributes at most $\varepsilon$, while propagation through the next teacher update multiplies the previous error by at most $L$. Thus $e_{j+1} \le \varepsilon + Le_j$ and $e_0 = 0$. Unrolling gives $e_N \le \varepsilon \sum_{j=0}^{N-1} L^j$, with the stated closed forms. □

*Proof of proposition 5.* Let $q_{12}$ be the learned full occupancy and $\widetilde{q}_{12}$ the weighted combination of learned subinterval occupancies, evaluated at the learned successor state. The cocycle defect gives $D(q_{12}, \widetilde{q}_{12}) \le \delta_q$. By convexity of the $\ell_1$ norm and the exact teacher cocycle,

$$D(q_{12}, q_{12}^\star) \le \delta_q + \frac{\Delta_1}{\Delta} D(q_1, q_1^\star) + \frac{\Delta_2}{\Delta} D(q_2(\widehat{z}_1), q_2^\star(z_1^\star)).$$

The last term is at most $D(q_2(\widehat{z}_1), q_2(z_1^\star)) + D(q_2(z_1^\star), q_2^\star(z_1^\star)) \le K_q\varepsilon_1 + \eta_2$. Substitution yields equation (34). □

*Proof of lemma 1.* Consider the augmented matrix

$$A = \begin{bmatrix} V \\ \mathbf{1}^\top \end{bmatrix} \in \mathbb{R}^{(d+1)\times E}.$$

By assumption its nullspace contains a nonzero vector $a$ satisfying $Va = 0$ and $\mathbf{1}^\top a = 0$; the rank–nullity theorem guarantees this whenever $E > d + 1$. For an interior simplex point $q$, sufficiently small positive and negative $t$ preserve nonnegativity of $q+ta$, while $\mathbf{1}^\top(q+ta) = 1$ and $V(q+ta) = Vq$. Varying $t$ in a nonzero interval gives infinitely many distinct routes with the same endpoint update. □

*Proof of proposition 6.* Both length-two paths use each expert once, so equation (9) gives $(1/2, 1/2)$ for both. Direct composition gives $B(A(z)) = (2z) + 1$ and $A(B(z)) = 2(z + 1) = 2z + 2$, which differ by one for every $z$. □

*Proof of proposition 7.* For $1 \le k < E$, any teacher-selected index $i$ and unselected index $j$ satisfy $g_i - g_j \ge \gamma$. The perturbation obeys $\widehat{g}_i - \widehat{g}_j \ge g_i - g_j - 2\left\|\widehat{g} - g\right\|_\infty > 0$. Therefore every teacher-selected logit remains above every unselected logit, so the pre-capacity top-$k$ sets coincide. The case $k = E$ is immediate because the selected set is the whole bank. □

*Proof of proposition 8.* Since $1 \le m \le E$, every feasible probability vector has a nonempty support set $S$ with $|S| \le m$. Let $r = \sum_{e \notin S} q_e$. Because $p_e = 0$ outside $S$, the outside contribution to $\|p - q\|_1$ is $r$. Probability conservation implies $\sum_{e \in S}(p_e - q_e) = r$, so the inside absolute deviation is at least $r$. Hence $\mathrm{TV}(p, q) \ge r$. The smallest possible omitted mass is attained by retaining the $m$ largest entries, giving $r^\star = 1 - \sum_{e=1}^{m} q_{(e)}$. Equality in this relaxed probability-vector class is achieved by preserving those retained entries and adding $r^\star$ to any one retained coordinate. Therefore its minimum TV is $r^\star$; no attainability by a more constrained routing schedule is asserted. □

*Proof of proposition 9.* By the triangle inequality,

$$\left\| \sum_e (q_e - p_e) f_e \right\| \le \sum_e |q_e - p_e| \, \|f_e\| \le M \|q - p\|_1 = 2M \, \mathrm{TV}(q, p).$$

Applying the integral triangle inequality and then integrating the pointwise bound along the common state/field evaluation path gives the second statement. Divergent state trajectories would require additional terms. □

## F Additional Technical Remarks

**Normalized versus unnormalized routing measure.** For the measurable continuous probability-valued responsibility path in equation (10), the duration mass $A_{s,\Delta} = \Delta Q_{s,\Delta}$ is additive and $Q$ remains in the simplex. This duration mass differs from the discrete step-count mass $A_I = \sum_i p_i$: with step width $h$, $n$ steps, and $\Delta = nh$, one has $\Delta Q_I = hA_I$. When capacity drops are recorded without renormalization, $Q_I$ is sub-probability with total mass $1 - D_I/n$, not a simplex point; the corresponding missing duration mass is $hD_I$. Normalized occupancies obey the weighted identity in equation (11), rather than $Q_{12} = Q_1 + Q_2$. Probability-normalized TV claims must not be silently applied to dropped activation mass.

**Ordered-signature composition.** For nonempty adjacent route paths, every adjacent pair lies either wholly in the first interval, wholly in the second, or crosses their single boundary. These three disjoint cases yield $B_{IJ} = B_I + B_J + r_I \ell_J^\top$ in equation (13); dividing by the positive total number $n_I + n_J - 1$ of adjacent pairs yields equation (14). For a singleton, $B_I = 0$ and the bookkeeping choice $C_I = 0$ ensures $(n_I - 1)C_I = 0$, so the formula still applies. Concatenating three intervals gives the same two cross-boundary terms under either parenthesization, which verifies associativity. Empty intervals are outside this signature definition because their first/last routes are undefined.

**Non-autonomous versus autonomous maps.** If every fine update is generated by an autonomous flow, start time can be omitted and $\Psi_{\Delta_1+\Delta_2} = \Psi_{\Delta_2} \circ \Psi_{\Delta_1}$ is a semigroup. Looped networks with iteration embeddings, scheduled diffusion dynamics, and time-dependent control are generally non-autonomous, so the start index must remain explicit.

**Stochastic teachers.** For Markov stochastic recursion on a sufficient state, the process law holds for transition kernels under Chapman–Kolmogorov composition. Routing occupancy may be defined pathwise and then averaged, or directly as a conditional expected measure. With finite second moments, the population minimizer of unrestricted deterministic mean-square prediction is the conditional mean and can erase multimodality; a distributional predictor is needed if preserving that conditional distribution is the objective. This population characterization does not guarantee that a restricted trained student reaches the minimizer.

## G Illustrative Discrete Construction: Protocol and Results

This appendix is the verifiable companion to the illustrative discrete construction in the main text. Its scope is the discrete construction alone: it says nothing about the learned prototype, the same-graph DIRECT controls, or any general method.

**Files and digests.** The supplement holds the protocol (`SCALE_K_PROTOCOL.md`) and, under `scale-k/`, the runner, `RESULT.json`, the per-input trace, and the run log. SHA-256 digests of the result file and of the checksum list (each digest on its own line): `RESULT.json` `f81a82352446c619ef24154ef204f1841fdf081e1271bcf4267139482e8eefd7` `SHA256SUMS.txt` `abb286d9d5c33c34ea0a2135ee7a71466e5731dc487bb2409cbff57428c04dea`. The remaining files' digests verify against `SHA256SUMS.txt` in the same directory. Rerunning the runner reproduces the per-input trace byte-identically and every numerical field of `RESULT.json`.

**Determinism; no seeds.** The runner exhaustively enumerates the full population ($2^{3k}$ atoms: 64/512/4096) and *both* controller classes (all 81 restricted policies; the saturation class decomposes exactly by block because risk and fee are additive). There is no gradient fit, no sampling, no seed, no held-out split, and no model selection: the declared evaluation is the whole population, so "a badly fitted baseline" is not available as an explanation of any gap.

**Plant and tariff.** A discrete continuation-horizon family indexed by $k \in \{2, 3, 4\}$: $k$ blocks, each carrying bits $(A_i, B_i)$ and a phase bit $P_i$; the continuation answers one query per block, $Y_i = A_i \oplus B_i$ if $P_i = 0$ else $B_i$. The macro reads all blocks and writes a $k$-bit summary $Z$ (one bit per block, identical width for every arm); at query $i$ the continuation reads $Z_i$, $P_i$, and $i$ free, while any uncovered data bit costs one paid expert call that returns the correct $Y_i$ (no arm observes $Y$). Declared tariff: stop-path fee 1.60, per-expert call 1.10. Arms include summary C ($Z_i = A_i \oplus B_i$: free on $P_i = 0$ queries, one paid call on $P_i = 1$; $Z_i = B_i$ would tie it), equal-width generic G ($Z_i = A_i$, independent of $Y_i$ in both phases), always-expert AE, always-shallow AS, and richest legal controller S (saturation; may depend on block index). Restricted controllers map $(Z_i, P_i)$ to $\{\text{stop-emit-0}, \text{stop-emit-1}, \text{call}\}$ and are fully enumerated.

**Prespecified pass criteria (G1–G3) and outcomes.** The pass criteria were fixed in the protocol before the runner was written. G1 (quality, $\varepsilon$=0): C reaches risk 0 at every $k$. G2 (mechanism): at C's risk-0 fees 2.70/3.25/3.80 for $k = 2, 3, 4$, the matched-fee generic retains risk 0.25. G3 (frontier): C's risk-0 fee is strictly below the generic/always-expert risk-0 fee (3.80/4.90/6.00; gap 1.10/1.65/2.20, per-query gap constant $\approx 0.55$, the preregistered linear-in-$k$ shape). All three criteria hold at every $k$. Two prespecified degeneracies: always-expert coincides with the generic risk-0 fee ($G_{\text{risk0}} \equiv \text{AE}$), so G2 and G3 are not independent statements; and saturation S ties C on risk and fee (S ties C), where an S win would mean that S sees information C lacks. The two prespecified claims are (i) G's risk stays strictly positive at C's fee, and (ii) C's risk-0 fee is strictly below AE's. Pass criteria and degeneracy notes are those recorded in `RESULT.json`; no $\varepsilon$ was changed after the run.

**What this appendix does not claim.** No result about the learned prototype, no DIRECT comparison, no repair of the prototype's quality failures, and no statement about flow matching or mixture-of-experts systems in general.

Table 6: Illustrative discrete construction (result files in the supplement's `scale-k/` directory). Risk 0 fees for C; generic at those fees keeps risk 0.25; S ties C; $G_{\text{risk0}} \equiv \text{AE}$.

| $k$ (atoms) | C fee (risk 0) | G at C fee (risk) | S vs C |
|---|---|---|---|
| 2 (64) | 2.70 | 0.25 | tie |
| 3 (512) | 3.25 | 0.25 | tie |
| 4 (4096) | 3.80 | 0.25 | tie |

# H CONTINUATION-GAP DECOMPOSITION: SHIFT SPLIT AND PARENT-LEVEL SPREAD

Figure 2(A) plots the parent-level IID spread behind the main-text fixed-path table.

Same decomposition as table 2, on the distribution-shift split (mean over five parents, $\times 10^{-3}$). DIRECT rows report macro loss, regret, and agreement only, as in the main table. The machine-readable source, including the per-parent values and the hash-verified artifact list, is `analysis/continuation_decomposition_results.json` in the supplement; the tables in this section are reproduced from that JSON; re-running the script additionally requires the per-example arrays, which exceed the supplement size limit; their SHA-256 digests are included so a rerun can be checked against the reported values.

Table 7: Continuation-gap decomposition, shift split ($\times 10^{-3}$). *Sign* counts, of five parents, are those with positive reallocation; per-parent values for both splits are in the supplement JSON.

| Arm | Fine | Macro | Gap | Fixed-act. | Realloc. | Sign | Regret |
|---|---|---|---|---|---|---|---|
| JOINT-END | 10.99 | 12.17 | +1.18 | +2.37 | −1.19 | 2/5 | 4.98 |
| SIG | 11.36 | 14.03 | +2.68 | +3.72 | −1.04 | 4/5 | 6.06 |
| RESP-ONLY | 10.91 | 11.58 | +0.67 | +0.99 | −0.31 | 2/5 | 4.67 |
| SHORTCUT-CONT | 10.72 | 13.47 | +2.75 | +5.83 | −3.08 | 3/5 | 5.77 |
| RELEVANT (candidate) | 10.64 | 13.77 | +3.13 | +2.54 | +0.60 | 4/5 | 5.93 |
| ROUTER-RELEVANT | 35.26 | 26.70 | −8.57 | −4.12 | −4.45 | 0/5 | 10.43 |
| FROZEN-RELEVANT | 20.64 | 30.66 | +10.02 | +10.11 | −0.09 | 2/5 | 13.08 |
| DIRECT-DATA$^\dagger$ | n/a | 8.92 | n/a | n/a | n/a | n/a | 4.74 |
| DIRECT-COMPUTE$^\dagger$ | n/a | 6.08 | n/a | n/a | n/a | n/a | 3.59 |

**Per-parent factorial contrast.** Table 8 lists the four structured arms per parent behind table 3. Effects are on log macro MSE: $A = \frac{1}{2}[(\log \text{SIG} - \log \text{JOINT}) + (\log \text{RELEVANT} - \log \text{RESP})]$ and $R = \frac{1}{2}[(\log \text{RESP} - \log \text{JOINT}) + (\log \text{RELEVANT} - \log \text{SIG})]$, reported as relative percentages. RELEVANT differs from RESP-ONLY in factor A and in relevance weighting of the response term, so $A$ partly absorbs that weighting; the simple effect JOINT-END→SIG (neither cell has a response term) is the purer contrast and is worse on $5/5$ parents in both splits. Script: `analysis/factorial_contrast.py`.

Table 8: Per-parent macro continuation MSE ($\times 10^{-3}$) for the $2 \times 2$ structured design and the per-parent main effects.

| Split | Parent | JOINT-END | SIG | RESP-ONLY | RELEVANT | $A$ | $R$ |
|---|---|---|---|---|---|---|---|
| IID | 3407 | 5.67 | 8.57 | 5.71 | 8.11 | +46.4% | −2.4% |
| IID | 3413 | 7.04 | 7.87 | 6.90 | 7.45 | +9.8% | −3.6% |
| IID | 3433 | 7.58 | 9.15 | 7.56 | 8.88 | +19.1% | −1.6% |
| IID | 3449 | 9.63 | 11.09 | 9.06 | 10.74 | +16.8% | −4.5% |
| IID | 3461 | 9.66 | 10.92 | 9.36 | 10.64 | +13.3% | −2.8% |
| Shift | 3407 | 10.29 | 12.82 | 9.73 | 12.24 | +25.2% | −5.0% |
| Shift | 3413 | 13.36 | 15.26 | 13.55 | 15.23 | +13.3% | +0.6% |
| Shift | 3433 | 11.94 | 14.69 | 11.54 | 14.61 | +24.8% | −2.0% |
| Shift | 3449 | 12.73 | 14.15 | 11.73 | 13.75 | +14.2% | −5.4% |
| Shift | 3461 | 12.55 | 13.26 | 11.35 | 13.01 | +10.0% | −5.8% |

# I SENSITIVITY OF FACTOR A TO ITS LOSS WEIGHT (DECLARED POST HOC)

The paired $2 \times 2$ contrast (table 3) fixes the occupancy-anchor and composition weight at the pre-specified $0.1$. To test whether a smaller weight recovers the loss, the protocol in the supplement (`sensitivity_A_01/PROTOCOL.md`, hash recorded in the result file before execution) declared a sweep of $\lambda \in \{0.02, 0.05, 0.2\}$ (anchor and composition weights set equal) for the SIG arm, with the primary statistic "parents on which SIG$_\lambda$ is worse than JOINT-END" and the decision rule "no helpful regime if every $\lambda$ has $\geq 4/5$ IID parents worse". Everything else is identical to the confirmation runs: the same five fine-parent checkpoints, the same adaptation minibatch stream, the same test draws, and the same source files imported unchanged. The sweep ran under an older torch build (2.8.0; the confirmation runs used 2.14.0); a parity control refit SIG at $\lambda = 0.1$ on parent 3407 in that environment and reproduced the confirmation IID macro MSE exactly ($0.0085695$, absolute difference $0$). The parity check covers one parent, one arm, and the IID split; a same-protocol rerun on another runtime (section K) reproduces that fit to within $10^{-6}$ relative but moves parent 3461's SIG macro MSE by $4.2\%$ (IID) and $4.8\%$ (range shift), so per-parent margins of a few percent (e.g. $+2.6\%$ for parent 3413, IID, $\lambda = 0.02$) are within cross-runtime differences. Table 9 gives the outcome: the penalty is monotone in $\lambda$ and no weight tested is neutral. This is post-hoc and one synthetic family; it does not exclude a helpful regime below $0.02$ or under a different continuation. The sweep moves the anchor and composition weights together; at $\lambda = 0.1$ the rise is attributed to the term added alone (the composition penalty raises error alone, $+13.7\%$ on 5/5 parents, and the anchor alone does not, $-0.6\%$ on 2/5; together they interact, $+21.2\%$), and which term drives its growth with $\lambda$ is untested (section L).

Table 9: SIG$_\lambda$ versus JOINT-END, macro continuation MSE, five paired parents. Relative effect is the geometric mean of per-parent ratios.

| $\lambda$ | IID effect | IID parents worse | Shift effect | Shift parents worse | Source |
|---|---|---|---|---|---|
| 0.02 | $+11.2\%$ | $5/5$ | $+7.8\%$ | $5/5$ | rerun |
| 0.05 | $+16.0\%$ | $5/5$ | $+10.4\%$ | $4/5$ | rerun |
| 0.1 | $+21.6\%$ | $5/5$ | $+15.5\%$ | $5/5$ | confirmation arm |
| 0.2 | $+22.5\%$ | $5/5$ | $+18.7\%$ | $5/5$ | rerun |

# J MULTI-FAMILY REPLICATION (DECLARED POST HOC)

The confirmation study uses one synthetic input family. To test whether its conclusions are a property of that function, two structurally different synthetic families were declared and run after the original under the frozen runner (`train.py` byte-identical, `core.py` changed only in the data generator; `multifamily_01/PROTOCOL.md` in the supplement; the protocol, both family definitions, and the replication rule were committed before any run). *quadrant* is a four-region piecewise map with saturating nonlinearities and a different branch geometry; *iterated* is a two-step iterated map $z_{k+1} = \tanh(A_b z_k + c_b)$ with branch-dependent $A_b$, so the target is itself the result of a short recursive computation. Everything else is inherited unchanged: architecture, all objective weights, data sizes, step counts, learning rate, all eleven arms, the DIRECT budget rule, the timing protocol, the freeze order, and RELEVANT as the primary without re-selection. The declared replication rule (candidate worse than DIRECT-DATA on $\geq 4/5$ parents and factor A positive on $\geq 4/5$ parents, IID, on both families) is met. Per-family fixed-path means are in table 10; the IID continuation-gap decomposition is in table 11. Batch-one macro/fine latency ratio is $0.475$ (quadrant) and $0.475$ (iterated) with all timing parity checks passing.

Across all seven structured arms and all six families, DIRECT-DATA (the candidate's 2,400 updates) has the lower macro MSE on all 60 parent–split pairs and DIRECT-COMPUTE on 59; the exception is iterated, IID, parent 5461, where RESP-ONLY ($0.01469$) is below DIRECT-COMPUTE ($0.01489$). Given only the 1,200 adaptation updates, the from-scratch graph beats every structured arm on only $1/5$ IID original-family parents (section K). What replicates: the candidate loses to both DIRECT controls on every parent of every family and split; the two restricted-adaptation arms lose to the candidate on every parent; the isolated macro speedup persists ($\approx 2.1\times$) while the ratio of parent-mean macro MSE to parent-mean own-fine MSE is far outside the $2\%$ rule ($1.43$–$1.95$); factor A is positive on $5/5$ (quadrant) and $4/5$ (iterated) parents. What does not replicate: the small

benefit of response matching (factor R) seen on the original family is absent on both new families (R main effect $+0.4\%$ and $-1.4\%$ IID, reducing error on only $2/5$ IID parents in either family), and on *iterated* the candidate is worse than JOINT-END on only $3/5$ IID parents. The reallocation term is negative on the quadrant family for every structured arm except FROZEN-RELEVANT (the router partially compensates a large replacement error, $\mathfrak{r}$ near $40 \times 10^{-3}$), and positive for the candidate on iterated ($4/5$ parents). These are synthetic families with a common input distribution and output dimension; they are not a real domain.

Table 10: Fixed-path results on the two new families (MSE, mean over five parents; same rows as the main-text table).

| Arm (macro) or fine reference | quadrant IID | quadrant shift | iterated IID | iterated shift |
|---|---|---|---|---|
| Pretrained fine reference | 0.093523 | 0.101661 | 0.024047 | 0.037823 |
| Continued fine-only | 0.060141 | 0.067100 | 0.011832 | 0.016615 |
| Candidate's coadapted fine path | 0.066664 | 0.070645 | 0.012356 | 0.018408 |
| JOINT-END macro | 0.089037 | 0.093989 | 0.017629 | 0.032582 |
| SIG macro | 0.096632 | 0.099983 | 0.018670 | 0.035291 |
| RESP-ONLY macro | 0.088933 | 0.093510 | 0.017218 | 0.032206 |
| SHORTCUT-CONT macro | 0.094799 | 0.097910 | 0.018796 | 0.033667 |
| RELEVANT macro (candidate) | 0.097732 | 0.100755 | 0.018717 | 0.035837 |
| ROUTER-RELEVANT macro | 0.136945 | 0.146289 | 0.033644 | 0.058613 |
| FROZEN-RELEVANT macro | 0.141470 | 0.151572 | 0.042214 | 0.070139 |
| DIRECT-DATA macro | 0.044928 | 0.052502 | 0.011932 | 0.023200 |
| DIRECT-COMPUTE macro | 0.027293 | 0.035360 | 0.007781 | 0.013154 |

Table 11: IID continuation-gap decomposition on the two new families ($\times 10^{-3}$; columns as in the main-text decomposition table except Gap and identity agreement; Sign = parents with positive reallocation). †DIRECT: no trained fine path.

| Family | Arm | Fine | Macro | Fixed-act. | Realloc. | Sign | Regret |
|---|---|---|---|---|---|---|---|
| quadrant | JOINT-END | 65.26 | 89.04 | $+39.12$ | $-15.34$ | 0/5 | 41.55 |
| quadrant | SIG | 64.97 | 96.63 | $+46.41$ | $-14.74$ | 0/5 | 45.28 |
| quadrant | RESP-ONLY | 65.21 | 88.93 | $+31.89$ | $-8.17$ | 1/5 | 37.88 |
| quadrant | SHORTCUT-CONT | 65.29 | 94.80 | $+46.83$ | $-17.32$ | 0/5 | 43.92 |
| quadrant | RELEVANT | 66.66 | 97.73 | $+37.38$ | $-6.31$ | 1/5 | 42.35 |
| quadrant | ROUTER-RELEVANT | 103.89 | 136.94 | $+35.01$ | $-1.96$ | 2/5 | 56.93 |
| quadrant | FROZEN-RELEVANT | 93.52 | 141.47 | $+47.60$ | $+0.35$ | 3/5 | 57.71 |
| quadrant | DIRECT-DATA† | n/a | 44.93 | n/a | n/a | n/a | 24.59 |
| quadrant | DIRECT-COMPUTE† | n/a | 27.29 | n/a | n/a | n/a | 15.37 |
| iterated | JOINT-END | 12.58 | 17.63 | $+5.26$ | $-0.21$ | 2/5 | 12.94 |
| iterated | SIG | 13.16 | 18.67 | $+4.48$ | $+1.02$ | 2/5 | 13.51 |
| iterated | RESP-ONLY | 13.68 | 17.22 | $+3.02$ | $+0.52$ | 3/5 | 11.96 |
| iterated | SHORTCUT-CONT | 12.87 | 18.80 | $+6.75$ | $-0.82$ | 2/5 | 13.40 |
| iterated | RELEVANT | 12.36 | 18.72 | $+4.06$ | $+2.30$ | 4/5 | 13.15 |
| iterated | ROUTER-RELEVANT | 42.85 | 33.64 | $-3.26$ | $-5.95$ | 0/5 | 15.39 |
| iterated | FROZEN-RELEVANT | 24.05 | 42.21 | $+14.87$ | $+3.30$ | 5/5 | 18.41 |
| iterated | DIRECT-DATA† | n/a | 11.93 | n/a | n/a | n/a | 9.27 |
| iterated | DIRECT-COMPUTE† | n/a | 7.78 | n/a | n/a | n/a | 6.45 |

**Search for a family where the fine path is worth replacing (declared post-hoc).** A further objection to the above is that the candidate's fine path is itself dominated by DIRECT-COMPUTE, so nothing is lost by never executing the fine computation and the replacement premise is empty. A second protocol (`multifamily_02/PROTOCOL.md` in the supplement; the protocol and all three family definitions were committed before any run) therefore searched for a family where depth matters: *deep6* and *deep8* iterate the branch-dependent map $z \mapsto \tanh(A_b z + c_b)$ for $K = 6$ and $K = 8$ steps, and *chaotic* applies four steps of a branch- and input-dependent logistic map ($r \in [3.6, 3.9]$). The declared qualification criterion: a family is worth replacing on if, on IID, the candidate's own fine path has lower MSE than DIRECT-COMPUTE's macro on at least four of five parents. No family qualifies (table 12): the fine path beats DIRECT-COMPUTE on $0/5$ IID parents on all three families (at most $1/5$ on any family and split), the candidate again loses to both DIRECT controls on every parent, both DIRECT controls stay below every structured arm on all 30 of these parent–split pairs, and factor A is positive on $5/5$ deep6/deep8 parents but on only $3/5$ ($+2.7\%$ IID, $+2.8\%$ range shift) of chaotic parents.

Table 12: Search families of the second protocol (five parents each; declared before any run). Column 3: parents on which the candidate's own fine path undercuts DIRECT-COMPUTE (qualification requires $\geq 4/5$ on IID); columns 4–6: parents on which the candidate's macro error exceeds each control's; column 7: factor-A main effect (geometric mean; parents with a positive effect); last column: macro/own-fine ratio of parent means (2% rule: $\leq 1.02$).

| Family | Split | Fine<DIR-C | >DIR-D | >DIR-C | >JOINT | A effect | Macro/fine |
|---|---|---|---|---|---|---|---|
| deep6 | IID | 0/5 | 5/5 | 5/5 | 5/5 | $+19.4\%$ (5/5) | 1.60 |
| deep6 | Shift | 0/5 | 5/5 | 5/5 | 5/5 | $+14.8\%$ (5/5) | 1.55 |
| deep8 | IID | 0/5 | 5/5 | 5/5 | 5/5 | $+12.8\%$ (5/5) | 1.46 |
| deep8 | Shift | 1/5 | 5/5 | 5/5 | 5/5 | $+8.5\%$ (5/5) | 1.67 |
| chaotic | IID | 0/5 | 5/5 | 5/5 | 3/5 | $+2.7\%$ (3/5) | 1.15 |
| chaotic | Shift | 0/5 | 5/5 | 5/5 | 3/5 | $+2.8\%$ (3/5) | 1.12 |

**Factor A stratified by teacher strength (declared post-hoc).** Because the fine path each anchored arm anchors to (its own, under stop-gradient) is weaker than DIRECT-DATA on most parents, the factor-A penalty admits a rival reading: anchoring to a poor teacher must hurt. Section L tests the anchor directly: added alone, it raises error on only $2/5$ original-family parents, whereas the composition penalty raises it on $5/5$. As a descriptive check across families, the 60 parent–split pairs of the six families were split by whether the candidate's own fine path has lower continuation MSE than DIRECT-DATA's macro on that parent and split (`analysis/teacher_strength_stratified.py`, same log-MSE main-effect formula as table 3). On the 17 pairs where the fine path is the stronger model (two original, seven iterated, two deep6, two deep8, four chaotic; none on quadrant), factor A raises continuation error on 13 of 17 (geometric mean $+8.3\%$), and the simple contrast JOINT-END→SIG is worse on the same 13; on the 43 pairs where the fine path is weaker, on 41 of 43 ($+12.0\%$). The penalty is smaller in the first stratum but present on 13 of its 17 pairs, which come from 11 parents; the four exceptions are two parents (iterated 5449 and chaotic 8407) in both splits, and 11 of the 17 pairs are iterated or chaotic. Within families, the first stratum has the smaller effect in four of the five families that have both strata (not deep6), on one to five parents each. The stratifier is an outcome of an anchored arm, not a pre-treatment covariate, so the split is descriptive: the penalty is present in both strata, but the split does not separate fine-path strength from family. The fine path beats DIRECT-COMPUTE on only 2 of 60 pairs, so the same split against DIRECT-COMPUTE is uninformative; section K varies teacher strength directly through the fine pretraining budget.

**Five-arm fidelity ordering.** Within the five fully trainable arms of the original family at the confirmation budget, ordering by occupancy fidelity to the fine path is the inverse of ordering by continuation error on both splits (table 13): the Spearman correlation of the two orderings is $-1.00$ on arm means and averages $-0.88$ (IID) and $-0.86$ (range shift) over parents. The ordering largely restates the factor-A contrast: the two anchored arms (SIG and RELEVANT, occupancy TV $0.31$) form one nearly tied cluster and the three unanchored arms ($0.71$–$0.79$) another, so the correlation adds little beyond the paired $2 \times 2$ design. Both anchored arms also carry the composition penalty, which section L identifies as the source of the factor-A rise; ANCHOR-ONLY, more faithful than either (occupancy distance $0.264$ versus $0.309$ for SIG on the same runtime), has higher error than JOINT-END on only $2/5$ parents. With $4$–$16\times$ longer fine pretraining the full ordering does not recur (Spearman on arm means $-0.2$, $-0.6$, $+0.2$, IID), while the anchored pair keeps the higher mean error (section K).

Table 13: The five fully trainable arms of the original family, least-to-most faithful (occupancy total-variation distance to the fine path) and best-to-worst continuation MSE ($\times 10^{-3}$, parent means); the two orders coincide at this budget, not after longer fine pretraining (section K), and both anchored arms (REL, SIG) also carry the composition penalty (section L). Spearman: arm means / mean per-parent. Arm-mean values: `analysis/continuation_decomposition_results.json`; per-parent correlations: `analysis/fidelity_rank_five_arms.json`.

| Split | Occupancy TV (→ more faithful) | Continuation MSE (→ worse) | Spearman |
|---|---|---|---|
| IID | RESP .79, JE .78, SC .72, REL .31, SIG .31 | RESP 7.72, JE 7.92, SC 8.94, REL 9.16, SIG 9.52 | $-1.00$ / $-0.88$ |
| Shift | RESP .78, JE .78, SC .71, REL .31, SIG .31 | RESP 11.58, JE 12.17, SC 13.47, REL 13.77, SIG 14.03 | $-1.00$ / $-0.86$ |

# K LONGER FINE PRETRAINING (DECLARED POST HOC)

Pretraining the fine model up to $16\times$ longer lowers the mean error of the pretrained fine reference and of the candidate's own fine path, but at no budget does that fine path beat DIRECT-COMPUTE or DIRECT-DATA on four of five parents, and factor A raises continuation error on at least four of five parents at every budget. The study addresses an objection to table 12: the fine model receives 1,200 pretraining updates and every inherited arm then adapts for 1,200 updates after an Adam reset, while DIRECT-DATA trains for 2,400 uninterrupted updates, so the fine path might lose only because of its update budget, and the factor-A penalty might be specific to an under-trained fine model (section L separates its two terms). The study varies only the fine pretraining budget on the original family and asks whether the candidate's fine path then beats the DIRECT controls on at least four of five parents and, at any budget where it does, whether factor A still raises macro error.

**Declared protocol.** The protocol (`teacher_regime_01/PROTOCOL.md` in the supplement), the four condition files, the training and analysis scripts, and the interpretation rules were committed before any run on the confirmation parents; the study was declared and run after both multi-family studies (section J), and before the commit only a timing run on a reserved seed that enters no result was executed. The four conditions pretrain the fine model for 1,200, 4,800, 9,600, and 19,200 updates ($\times 1$, $\times 4$, $\times 8$, $\times 16$ the confirmation budget; $\times 16$ was admitted by a timing rule fixed before the timing run). Adaptation stays at 1,200 updates, and the architecture, objective weights, data sizes, batch size, learning rate, Adam reset, the eleven arms of the confirmation study, and the candidate are unchanged. The five confirmation parents are reused and paired by seed across conditions; their IID and range-shift draws were opened in the confirmation study at $\times 1$, so the study is a paired budget ladder on known parents, not a fresh confirmation. The two DIRECT rules of the main text scale with the budget: DIRECT-DATA receives the candidate's update count (2,400, 6,000, 10,800, 20,400) and DIRECT-COMPUTE the candidate's forward-affine MACs including fine pretraining (7,910, 17,722, 30,805, 56,970 updates; none truncated). Two given-teacher controls, which charge only the replacement budget as when a trained fine model already exists, are added: DIRECT-ADAPT receives the 1,200 adaptation updates (update-matched but given $26\%$ of the adaptation phase's MACs, so favourable to the candidate), and DIRECT-ADAPT-COMPUTE receives 4,640 updates, matching the adaptation phase's MACs. All four DIRECT arms start from the parent's raw initialization and read the leading rows of one minibatch stream, so they are checkpoints of one from-scratch trajectory, and both given-teacher controls are byte-identical across conditions. Each condition has 13 fits per parent (65 per condition). All conditions ran on one runtime (Linux x86-64, PyTorch 2.14.0 on CPU, one thread per process, deterministic algorithms) that differs from the confirmation study's, so the $\times 1$ condition is the same-runtime baseline.

**Declared rules.** The rules use IID; the range shift is reported next to every statement and cannot override it. A condition qualifies if the candidate's fine path has lower MSE than DIRECT-COMPUTE (primary; the criterion of table 12), DIRECT-DATA (secondary), or DIRECT-ADAPT (given-teacher) on at least $4/5$ parents; a given-teacher qualification is called MAC-matched only if it also holds against DIRECT-ADAPT-COMPUTE on at least $4/5$ parents. Factor A is the paired $2 \times 2$ main effect of table 3. The declared outcomes are (a) no condition qualifies; (b) some condition qualifies and in every qualifying condition factor A raises IID error on at least $4/5$ parents, stated as "the fidelity-inducing terms raise macro error also in a regime where the fine path is worth replacing" (the protocol's wording; factor A comprises the occupancy anchor and the composition penalty, section L) with each qualifying condition and each control beaten named; (c) some qualifying condition has factor A lowering IID error on at least $4/5$ parents, so the factor-A result is regime-dependent (takes precedence over (b)); (d) otherwise, inconclusive. A manipulation check requires the pretrained fine reference to have lower IID MSE than at $\times 1$ on at least $4/5$ parents. In a qualifying condition, a candidate macro that beats a DIRECT control on at least $4/5$ parents is reported as a positive result against that control.

**Outcome.** The declared rules give outcome (b), reached only through the given-teacher control DIRECT-ADAPT (table 14). All 260 fits completed on the first attempt with every test output finite. The manipulation check holds at $\times 4$, $\times 8$, and $\times 16$ ($5/5$ parents each): the pretrained fine reference's mean IID MSE falls from $14.43$ to $6.00$, $3.17$, and $3.54 \times 10^{-3}$, and the candidate's own fine path from $6.43$ to $3.77 \times 10^{-3}$. No condition qualifies under the primary or the secondary criterion: the

Table 14: Fine pretraining budget on the original family (five parents per condition; IID is the declared primary split). Means are MSE $\times 10^{-3}$: pretrained fine reference, candidate's own fine path, and DIRECT-COMPUTE macro. Columns 6–9: parents on which the candidate's own fine path has lower MSE than DIRECT-COMPUTE (C), DIRECT-DATA (D), DIRECT-ADAPT (A), and DIRECT-ADAPT-COMPUTE (AC); a condition qualifies with $\geq 4/5$ on IID against C, D, or A. Macro<A: parents on which the candidate's macro has lower MSE than DIRECT-ADAPT. Factor A: geometric-mean main effect (parents with a positive effect). Macro/fine: ratio of parent means (2% rule: $\leq 1.02$). DIRECT-ADAPT and DIRECT-ADAPT-COMPUTE means are the same in every condition (9.09 and 3.71 IID; 14.23 and 6.60 range shift).

| | | Mean MSE ($\times 10^{-3}$) | | | Own fine path below | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Pretraining | Split | Ref. fine | Own fine | DIR-C | C | D | A | AC | Macro<A | Factor A | Macro/fine |
| 1,200 ($\times$1) | IID | 14.43 | 6.43 | 3.87 | 0/5 | 1/5 | 4/5 | 0/5 | 2/5 | +20.2% (5/5) | 1.43 |
| 4,800 ($\times$4) | IID | 6.00 | 5.09 | 2.59 | 0/5 | 0/5 | 5/5 | 0/5 | 4/5 | +15.5% (5/5) | 1.48 |
| 9,600 ($\times$8) | IID | 3.17 | 4.33 | 2.34 | 0/5 | 0/5 | 5/5 | 1/5 | 5/5 | +13.1% (5/5) | 1.28 |
| 19,200 ($\times$16) | IID | 3.54 | 3.77 | 1.80 | 0/5 | 1/5 | 5/5 | 2/5 | 5/5 | +13.6% (4/5) | 1.35 |
| 1,200 ($\times$1) | Shift | 20.64 | 10.65 | 6.08 | 1/5 | 1/5 | 5/5 | 0/5 | 3/5 | +16.7% (5/5) | 1.30 |
| 4,800 ($\times$4) | Shift | 8.90 | 8.65 | 4.08 | 0/5 | 1/5 | 5/5 | 0/5 | 4/5 | +15.9% (5/5) | 1.29 |
| 9,600 ($\times$8) | Shift | 5.83 | 7.03 | 3.72 | 0/5 | 0/5 | 5/5 | 2/5 | 5/5 | +9.8% (5/5) | 1.32 |
| 19,200 ($\times$16) | Shift | 5.10 | 6.63 | 3.70 | 0/5 | 0/5 | 5/5 | 2/5 | 5/5 | +11.8% (4/5) | 1.31 |

candidate's fine path is below DIRECT-COMPUTE on $0/5$ IID parents at every budget and below DIRECT-DATA on at most $1/5$, while DIRECT-COMPUTE, whose MAC budget includes the fine pretraining, also improves (from 3.87 to $1.80 \times 10^{-3}$). Every condition qualifies under the given-teacher criterion (below DIRECT-ADAPT on $4/5$, $5/5$, $5/5$, and $5/5$ IID parents), none of them MAC-matched (below DIRECT-ADAPT-COMPUTE on $0/5$, $0/5$, $1/5$, and $2/5$). Factor A raises IID error on $5/5$, $5/5$, $5/5$, and $4/5$ parents ($+20.2\%$, $+15.5\%$, $+13.1\%$, $+13.6\%$) and lowers it on at most one parent in any condition, so outcome (c) does not apply. With the control named, the declared statement is: the factor-A terms (the occupancy anchor plus the composition penalty, which the protocol calls the fidelity-inducing terms; section L attributes most of the rise to the composition penalty at $\times 1$ and $\times 4$) raise macro error also where the fine path beats the update-matched from-scratch control DIRECT-ADAPT (1,200 updates, 26% of the adaptation phase's MACs), at $\times 1$, $\times 4$, $\times 8$, and $\times 16$. This qualification is not MAC-matched, and it already holds at $\times 1$, so longer pretraining does not create it; under the criterion of table 12 (DIRECT-COMPUTE) no condition is worth replacing. The range shift agrees: factor A raises error on $5/5$, $5/5$, $5/5$, and $4/5$ parents ($+9.8\%$ to $+16.7\%$), and the candidate's fine path is below DIRECT-ADAPT on $5/5$ parents in every condition.

**Candidate and secondary outcomes.** Among the DIRECT controls, the candidate's macro beats only DIRECT-ADAPT: it is below DIRECT-ADAPT on $4/5$, $5/5$, and $5/5$ IID parents at $\times 4$, $\times 8$, and $\times 16$ (a positive result for the recipe against DIRECT-ADAPT only; $2/5$ at $\times 1$), and it is worse than each of DIRECT-ADAPT-COMPUTE, DIRECT-DATA, and DIRECT-COMPUTE on all 40 parent–split pairs (four conditions, five parents, two splits). It is also worse than JOINT-END and RESP-ONLY on at least $4/5$ IID parents at every budget. The budget of the from-scratch control decides whether it beats every structured arm: at $\times 1$, DIRECT-ADAPT has lower macro MSE than all seven structured arms on $1/5$ IID and $0/5$ range-shift parents (none at longer pretraining), and DIRECT-ADAPT-COMPUTE on $5/5$ in both splits ($3/5$ to $5/5$ at longer pretraining). Its macro/own-fine ratio stays outside the 2% rule (1.28–1.48 IID). Three descriptive facts qualify the ladder: the pretrained fine reference and DIRECT-DATA have higher mean IID MSE at $\times 16$ than at $\times 8$ (3.54 versus 3.17 and 2.99 versus $2.24 \times 10^{-3}$); at $\times 8$ and $\times 16$ the candidate's coadapted fine path has a higher mean MSE than the pretrained fine reference it starts from (4.33 versus 3.17 and 3.77 versus $3.54 \times 10^{-3}$) and than continued fine-only training (3.04 and $2.73 \times 10^{-3}$), neither of which is below DIRECT-ADAPT-COMPUTE on more than $3/5$ IID parents at any budget (continued fine-only at $\times 16$: $5/5$ range-shift parents); and factor R lowers IID error on $4/5$, $4/5$, $4/5$, and $3/5$ parents ($-0.8\%$ to $-6.3\%$). Over the five fully trainable arms, the Spearman correlation between occupancy distance and macro MSE on arm means is $-1.0$ at $\times 1$ but $-0.2$, $-0.6$, and $+0.2$ at $\times 4$, $\times 8$, and $\times 16$ (IID; range shift $-1.0$, $-0.5$, $-0.3$, $-0.1$). What persists in every condition and split is the factor-A cluster contrast: the two anchored arms (SIG, RELEVANT; occupancy distance 0.31–0.40) have a higher mean macro MSE than the three unanchored arms (0.66–0.77); both anchored arms also carry the composition penalty (section L). The batch-one macro/fine latency ratio

is $0.35$–$0.45$ with every timing parity check passing; conditions ran two at a time on two CPUs and three of them also overlapped other jobs, which affects wall-clock measurements only, so the latency ratios (the $\times 16$ value in particular) are indicative.

**Reproduction and scope.** The $\times 1$ condition reproduces the confirmation study's rule statistics across runtimes: over 200 per-parent values (eleven arms, fine and macro paths, both splits), the median relative difference is $7 \times 10^{-7}$ and the maximum $13.4\%$ (the fine path of ROUTER-RELEVANT, one IID parent); 27 of the 140 values of the seven structured arms differ by more than $1\%$; every DIRECT value differs by less than $10^{-5}$ relative; the primary and secondary qualification counts and the factor-A parent count coincide, and the factor-A effect is $+20.2\%$ here versus $+20.5\%$ in the confirmation study. An independent reimplementation of the analysis from the per-parent outputs reproduces every count, effect, mean, and correlation above. Only the pretraining update count varies: constant-learning-rate Adam, the reset, the width, and the adaptation budget are fixed, so the study bears on the update budget up to $16\times$, not on learning-rate schedules, wider fine models, or longer adaptation.

## L Separating the Two Factor-A Terms (Declared Post Hoc)

On the original family at the prespecified weight, most of the factor-A penalty comes from the composition penalty: added alone to JOINT-END, the composition penalty raises IID continuation error on $5/5$ parents at both fine pretraining budgets tested ($+13.7\%$ and $+15.1\%$), whereas the occupancy anchor alone lowers the occupancy distance to the fine path on $5/5$ parents and raises continuation error on $2/5$ (geometric-mean change $-0.6\%$ and $-4.4\%$); with the composition penalty present, the anchor raises error further at 1,200 updates but not at 4,800. Factor A (table 3) adds two terms, each with weight $0.1$, to the JOINT-END objective. The occupancy anchor $\mathrm{mse}(q_M, q_F)$, with a stop-gradient on $q_F$, pulls the span-2 macro's predicted occupancy toward that of its own fine path; it is the fidelity term. The composition penalty $\mathrm{mse}(z_M, z_S) + \mathrm{mse}(q_M, q_S)$, where $(z_S, q_S)$ are the state and averaged occupancy after the two chained unit-span macro calls $C(z_0, 0, 1)$ and $C(\cdot, 1, 1)$, ties the span-2 macro to its own unit-span maps; it is self-consistency of the macro family, not fidelity to the fine path. The $2 \times 2$ design of section 6 varies the two terms together. This study asks which of them, added alone at its weight in SIG, raises continuation error (the MSE of the hard-deployed macro path).

**Declared protocol.** The protocol (`factorA_decomp_01/PROTOCOL.md` in the supplement), the two condition files, the training and analysis scripts, and the interpretation rules were committed before any run on the confirmation parents. The study was declared after the paper's original results, including the factor-A result that motivated it, and after the $\times 4$ condition of section K had finished (its $\times 1$ condition had not started); no result file of that study was read before the declaration. Two arms are added: ANCHOR-ONLY (JOINT-END plus the occupancy anchor) and COMP-ONLY (JOINT-END plus the composition penalty), each term at weight $0.1$ as in SIG, so JOINT-END, ANCHOR-ONLY, COMP-ONLY, and SIG form a $2 \times 2$ design over (anchor, composition) without response matching. The SIG objective equals JOINT-END plus the two single-term changes exactly, in value and in gradient (checked on a reserved seed before the declaration). Everything else is inherited from section K: the five confirmation parents; fine pretraining for 1,200 updates (the confirmation budget, $\times 1$; primary) or 4,800 updates ($\times 4$; robustness); for every arm, 1,200 adaptation updates after an Adam reset from the same fine parent on the same minibatch stream; the architecture, objective weights, and data sizes; and the runtime (Linux x86-64, PyTorch 2.14.0 on CPU, one thread per process, deterministic algorithms). RELEVANT is fitted because the training script requires the candidate arm; it enters no statistic or rule. REFERENCE, JOINT-END, SIG, and RELEVANT run unchanged code, so SIG versus JOINT-END repeats a known contrast: the study is a paired decomposition on known parents, not a fresh confirmation, and its new information is ANCHOR-ONLY and COMP-ONLY. Before the protocol commit, only code checks and a timing run on a reserved seed that enters no result were executed. The training script prints development-set (DEV) MSE values in its event log; during the timing run the DEV MSE values of REFERENCE, JOINT-END, and SIG on the reserved seed were seen, after the analysis script implementing every statistic and rule had been written, and no part of the design was changed afterwards (`PROTOCOL.md`, Section 10). During the confirmation runs, after the protocol commit, two

DEV MSE values of one confirmation parent (COMP-ONLY at 1,200 and SIG at 4,800 updates) were displayed; no statistic or rule of this study uses DEV values.

**Declared statistics and rules.** Per parent, the contrast of an arm with JOINT-END is the difference in log continuation MSE, reported as a geometric-mean change with the number of parents on which the arm has higher error. The $2 \times 2$ effects follow the convention of table 8: with $\ell_{ac}$ the log MSE of the arm with anchor $a$ and composition $c$ (JOINT-END $\ell_{00}$, ANCHOR-ONLY $\ell_{10}$, COMP-ONLY $\ell_{01}$, SIG $\ell_{11}$), the anchor main effect is $\frac{1}{2}[(\ell_{10} - \ell_{00}) + (\ell_{11} - \ell_{01})]$, the composition main effect $\frac{1}{2}[(\ell_{01} - \ell_{00}) + (\ell_{11} - \ell_{10})]$, and the interaction $\frac{1}{2}[(\ell_{11} - \ell_{01}) - (\ell_{10} - \ell_{00})]$. Three diagnostics are computed per test row against each arm's own fine path and averaged: occupancy distance $\mathrm{TV}(q_M, q_F)$, endpoint error $\overline{(z_M - z_F)^2}$ over rows and state coordinates, and composition error $\overline{(z_M - z_S)^2}$, reported with $\mathrm{TV}(q_M, q_S)$. The rules use IID; the range shift is evaluated with the same rules and reported next to every statement without overriding it. With $k_A$, $k_C$, and $k_S$ the IID parents on which ANCHOR-ONLY, COMP-ONLY, and SIG have higher continuation error than JOINT-END, the declared outcomes are D1 ($k_A \geq 4$, $k_C \leq 3$): "the occupancy anchor, the fidelity term, raises error"; D2 ($k_C \geq 4$, $k_A \leq 3$): "the composition penalty, not the fidelity term, raises error", in which case the paper states that the anchor alone does not raise error on at least $4/5$ parents and stops attributing the factor-A harm to fidelity; D3 (both $\geq 4$): "both terms raise error"; D4 (otherwise): neither term alone raises error consistently, "only jointly" if $k_S \geq 4$ and inconclusive otherwise. The 1,200-update condition decides the statement; if the 4,800-update condition gives the same outcome, the statement covers both budgets, and otherwise the dependence on the budget is stated. For the anchor alone, with $f$ the parents on which ANCHOR-ONLY has lower occupancy distance than JOINT-END and $j$ those with both lower distance and higher error, "more faithful and less accurate" holds if $j \geq 4$, does not hold if $j \leq 3$ and $f \geq 4$, and has its premise unmet if $f \leq 3$. The manipulation checks require ANCHOR-ONLY to have lower occupancy distance, and COMP-ONLY lower composition error, than JOINT-END on at least $4/5$ IID parents; a failed check would be attached to the statement without changing it.

Table 15: Each factor-A term added alone to JOINT-END: original family, 1,200 fine pretraining updates (the confirmation budget; primary condition), IID split, five paired parents. Macro MSE: continuation error of the hard-deployed macro path, parent mean ($\times 10^{-3}$). Change: geometric-mean change in macro MSE versus JOINT-END (parents with higher error). Occupancy distance $\mathrm{TV}(q_M, q_F)$, endpoint error $\overline{(z_M - z_F)^2}$ ($\times 10^{-3}$), and composition error $\overline{(z_M - z_S)^2}$: parent means (parents below JOINT-END). †Descriptive; enters no rule.

| Arm | Macro MSE | Change | Occupancy distance | Endpoint error | Composition error |
|---|---|---|---|---|---|
| JOINT-END | 7.88 | – | 0.780 | 31.95 | 0.731 |
| ANCHOR-ONLY | 7.82 | $-0.6\%$ (2/5) | 0.264 (5/5) | 31.63 (4/5) | 0.811 (0/5) |
| COMP-ONLY | 8.91 | $+13.7\%$ (5/5) | 0.725 (5/5) | 31.64 (3/5) | 0.0059 (5/5) |
| SIG | 9.43 | $+21.2\%$ (5/5) | 0.309 (5/5) | 31.27 (3/5) | 0.0067 (5/5) |
| RELEVANT† | 9.20 | – | 0.310 | 24.14 | 0.0074 |

Table 16: As table 15, range-shift split, 1,200 fine pretraining updates.

| Arm | Macro MSE | Change | Occupancy distance | Endpoint error | Composition error |
|---|---|---|---|---|---|
| JOINT-END | 12.20 | – | 0.775 | 36.21 | 0.767 |
| ANCHOR-ONLY | 12.27 | $+0.7\%$ (3/5) | 0.272 (5/5) | 35.96 (4/5) | 0.848 (0/5) |
| COMP-ONLY | 13.47 | $+10.6\%$ (4/5) | 0.711 (5/5) | 33.13 (4/5) | 0.0063 (5/5) |
| SIG | 13.91 | $+14.1\%$ (4/5) | 0.309 (5/5) | 32.89 (4/5) | 0.0071 (5/5) |
| RELEVANT† | 13.82 | – | 0.312 | 25.38 | 0.0078 |

**Outcome.** The declared rules give D2 in both conditions and on both splits (tables 15 to 18). All 50 fits (five arms, five parents, two conditions) completed on the first attempt with no non-finite value. Both manipulation checks hold on $5/5$ parents in both conditions: ANCHOR-ONLY lowers occupancy distance from $0.780$ to $0.264$ at 1,200 updates and from $0.719$ to $0.310$ at 4,800, and COMP-ONLY lowers composition error from $0.731$ to $0.0059$ and from $0.466$ to $0.0059$, with $\mathrm{TV}(q_M, q_S)$ also lower on $5/5$ parents. At 1,200 updates, IID, COMP-ONLY has higher continuation error than JOINT-END on $5/5$ parents ($+13.7\%$), SIG on $5/5$ ($+21.2\%$), and ANCHOR-ONLY on $2/5$ ($-0.6\%$); at 4,800 updates the counts are $5/5$ ($+15.1\%$), $4/5$ ($+13.0\%$), and $2/5$

Table 17: As table 15, IID split, 4,800 fine pretraining updates (robustness condition).

| Arm | Macro MSE | Change | Occupancy distance | Endpoint error | Composition error |
|---|---|---|---|---|---|
| JOINT-END | 6.83 | – | 0.719 | 35.85 | 0.466 |
| ANCHOR-ONLY | 6.53 | $-4.4\%$ (2/5) | 0.310 (5/5) | 35.86 (2/5) | 0.471 (1/5) |
| COMP-ONLY | 7.90 | $+15.1\%$ (5/5) | 0.713 (3/5) | 35.64 (3/5) | 0.0059 (5/5) |
| SIG | 7.80 | $+13.0\%$ (4/5) | 0.305 (5/5) | 36.72 (1/5) | 0.0068 (5/5) |
| RELEVANT$^\dagger$ | 7.55 | – | 0.325 | 26.07 | 0.0071 |

Table 18: As table 15, range-shift split, 4,800 fine pretraining updates (robustness condition).

| Arm | Macro MSE | Change | Occupancy distance | Endpoint error | Composition error |
|---|---|---|---|---|---|
| JOINT-END | 10.04 | – | 0.729 | 39.80 | 0.488 |
| ANCHOR-ONLY | 9.90 | $-1.1\%$ (2/5) | 0.314 (5/5) | 39.44 (3/5) | 0.488 (1/5) |
| COMP-ONLY | 11.64 | $+15.1\%$ (5/5) | 0.724 (4/5) | 37.93 (4/5) | 0.0062 (5/5) |
| SIG | 11.76 | $+16.5\%$ (5/5) | 0.304 (5/5) | 39.50 (1/5) | 0.0069 (5/5) |
| RELEVANT$^\dagger$ | 11.17 | – | 0.335 | 30.27 | 0.0072 |

($-4.4\%$). Post hoc paired $95\%$ $t$-intervals over the five parents (log MSE, back-transformed) are $[-11.7\%, +11.8\%]$ for ANCHOR-ONLY and $[+0.2\%, +29.0\%]$ for COMP-ONLY at 1,200 updates, and $[-13.5\%, +5.7\%]$ and $[+2.1\%, +29.7\%]$ at 4,800, so at 1,200 updates an anchor effect of about $\pm 12\%$ is not excluded. The range shift gives the same outcome: COMP-ONLY is higher on $4/5$ and $5/5$ parents, ANCHOR-ONLY on $3/5$ and $2/5$. In the $2 \times 2$ design the IID composition main effect is $+17.7\%$ ($5/5$ parents positive) at 1,200 and $+16.6\%$ ($5/5$) at 4,800 updates (range shift $+12.0\%$ and $+16.4\%$; $4/5$ and $5/5$), the anchor main effect is $+2.9\%$ ($3/5$) and $-3.1\%$ ($2/5$), and the interaction is $+3.6\%$ ($5/5$) and $+1.3\%$ ($1/5$); SIG's IID error exceeds the multiplicative combination of the two single-term effects by $7.2\%$ and $2.7\%$. Post hoc and descriptive: added to COMP-ONLY rather than to JOINT-END, the anchor raises IID error by $6.6\%$ ($4/5$ parents) at 1,200 updates and changes it by $-1.8\%$ ($1/5$) at 4,800 (range shift $+3.2\%$, $5/5$; $+1.2\%$, $3/5$); COMP-ONLY accounts for $67\%$ of SIG's IID log-MSE rise over JOINT-END at 1,200 updates ($76\%$ range shift) and for all of it at 4,800 ($92\%$ range shift). The declared statement, made for both pretraining budgets, is: the composition penalty, not the fidelity term, raises error, and the occupancy anchor alone does not raise error on at least $4/5$ parents.

**Fidelity and accuracy of each term.** For the anchor alone, $f = 5$ and $j = 2$ on IID in both conditions, so "more faithful and less accurate" does not hold: the anchor makes the macro more faithful in occupancy on $5/5$ parents without higher continuation error on at least $4/5$ of them (range shift: $j = 3$ and $j = 2$, with $f = 5$). ANCHOR-ONLY also has lower endpoint error than JOINT-END on $4/5$ IID parents at 1,200 and on $2/5$ at 4,800 updates. The composition penalty is not a fidelity term, yet at 1,200 updates COMP-ONLY also lowers occupancy distance slightly on $5/5$ parents ($0.725$ versus $0.780$) and has higher error on the same five (descriptive); at 4,800 updates it lowers occupancy distance on $3/5$. For SIG the same counts give $j = 5$ and $j = 4$. Occupancy fidelity to the fine path therefore remains insufficient in the sense that lowering occupancy distance by a factor of 2.3 to 3.0 does not lower continuation error on at least four of five parents (ANCHOR-ONLY is lower on $3/5$ in both conditions), but the single-term contrasts do not support the reading that making the macro more faithful raises its error: at the prespecified weight on the original family, at both budgets, most of the factor-A harm reported in table 3 and section K is attributed to the composition penalty, although with the composition penalty present the anchor raises error further at 1,200 updates (Outcome paragraph), and the RESP-ONLY→RELEVANT half of the factor-A main effect is not decomposed. ANCHOR-ONLY, the most faithful arm at 1,200 updates and nearly tied with SIG at 4,800 ($0.310$ versus $0.305$), has a lower parent-mean IID error than JOINT-END in both conditions ($7.82$ versus $7.88$ and $6.53$ versus $6.83 \times 10^{-3}$), so the fidelity–error ordering of table 13, whose two most faithful arms carry both terms, does not show that fidelity itself raises error. ANCHOR-ONLY's macro/own-fine ratio of parent-mean MSE is $1.15$ (IID) and $1.12$ (range shift) at 1,200 updates and $1.11$ and $1.17$ at 4,800, so this macro fails the $2\%$ rule on both splits at both budgets.

**Reproduction and scope.** The REFERENCE, JOINT-END, SIG, and RELEVANT checkpoints are bitwise identical, tensor by tensor and by file hash, to those of the $\times 1$ and $\times 4$ conditions of section K (40 of 40: four arms, five parents, two conditions), and an independent recomputation from the per-parent outputs reproduces every count, effect, and diagnostic above. On this runtime

the simple contrast JOINT-END→SIG at 1,200 updates is worse on $5/5$ IID and $4/5$ range-shift parents, versus $5/5$ on both splits in the confirmation study (section 6). The study covers the original family, the prespecified weight $0.1$, and two pretraining budgets. It does not separate the state and occupancy parts of the composition penalty, does not decompose factor A within the response arms (RELEVANT versus RESP-ONLY), and does not test which term drives the rise of the penalty with the joint weight in section I. The anchor main effect is not consistent in sign (positive on $3/5$ parents at 1,200 and $2/5$ at 4,800 updates), and the interaction, positive on $5/5$ parents at 1,200 updates, is positive on $1/5$ at 4,800.

## M Prospective Replication Protocol (Not Executed)

This protocol specifies how the controls of this paper would be repeated on a publicly trained Mixture-of-Experts language model (for example Switch-Base-8 or OLMoE). It has *not* been executed and makes no claim about the outcome; its purpose is to fix, before any result exists, how the contract would be tested on a model with learned routing; the two pretrained models of sections N and O have no expert router.

**Replacement boundary.** Choose a contiguous block of $k$ routed expert updates inside the deployed inference graph. The fine path is the original computation; the macro is trained to emit, from the boundary input, the state the continuation consumes. The declared continuation is everything downstream of the boundary: the router, the remaining experts, and the readout.

**Controls.** (i) *DIRECT*: train a model of the same deployed graph from scratch on the same data, at matched parameter updates (DIRECT-DATA) and at matched compute (DIRECT-COMPUTE). (ii) *Restricted adaptation*: adapt only the macro, or the macro and routers, holding the expert bank, encoder, and readout fixed. (iii) *Fidelity diagnostics*: endpoint state error, occupancy total variation, next-expert agreement, and all-candidate response MSE are logged per checkpoint, before any continuation evaluation.

**Quality and selector rules.** A replacement passes only if continuation quality stays within a predeclared tolerance of the fine path (e.g. $2\%$ relative perplexity or loss increase) on held-out text from a distribution distinct from training. Selectors must be legal: input-only at inference, fitted after the macro is frozen, and evaluated on the held-out distribution. Latency claims require end-to-end measured execution including dispatch, on the target hardware.

**Sequencing and freezing.** Five seeds; parents (in the language-model setting, five independent pretraining checkpoints or five distinct domains) are frozen before any macro training. The analysis script, seeds, tolerance, and arm list are committed before the first training run; all arms, including those that beat the candidate, are reported.

**Success criterion for a positive claim.** A macro-replacement method claim requires: continuation error within tolerance on every parent, at a measured end-to-end speedup, with the fidelity diagnostics reported whether or not they moved. Absent that, the correct conclusion is the one this paper reports for its prototype: a failure mode with result files.

## N The Contract on a Pretrained Language Model (Declared Before Running)

The synthetic prototype cannot say whether the separation between measured fidelity and continuation usefulness survives in a trained model. This appendix reports a small instantiation of the contract on GPT-2 small (12 blocks, width 768; Radford et al., 2019) with a *published* replacement objective. The protocol, arms, metrics, predictions, and falsification criterion were committed before any run (supplement `realmodel_01/PROTOCOL.md`); the learning rate was chosen once on the WikiText-2 validation split; the test splits were opened only by the frozen runner. Nothing in the synthetic evidence changed. Pass 2 (seed 1 and FID-SGD) ran on a second machine of identical type under a protocol amendment declared before the run (torch 2.14 there versus 2.8 on the

primary); its boundary MSE agrees with the primary machine to the fourth decimal (differences at most $3 \times 10^{-4}$), below every decisive margin (smallest $6 \times 10^{-3}$ NLL).

**Replacement boundary and substitute.** Blocks $[l, l{+}k)$ are removed and replaced by one position-wise residual linear map $h \mapsto h + Wh + b$ on the residual stream entering block $l$, the substitute class of ReplaceMe (Shopkhoev et al., 2025); blocks $l{+}k, \ldots, 11$, the final LayerNorm, and the LM head are the continuation and, like every other weight, are frozen. Five boundaries $(l, k) \in \{(6, 2), (4, 2), (8, 2), (4, 4), (6, 4)\}$ are the replication unit. Fitting data are WikiText-2 train (819,200 sampled token positions per SGD arm); IID evaluation is WikiText-2 test (32 non-overlapping 1024-token windows); range-shift evaluation is LAMBADA (Paperno et al., 2016) (1,000 passages: all-token NLL and teacher-forced last-word accuracy).

**Arms.** All arms share the deployed graph and the substitute class; only the objective differs. DROP: identity (no substitute). FID-LS: closed-form ridge least squares of $h_{l+k}$ on $h_l$, the exact fidelity optimum within the class and the ReplaceMe objective. FID-SGD: the same objective by Adam, matching the CONT arms' optimizer and data. CONT-ZERO: next-token NLL through the frozen continuation from zero initialization, with no fidelity target (the same-graph DIRECT control). CONT-LS: the same NLL objective started from the FID-LS solution, with held-out checkpoints at steps $0, 25, 50, 100, 200, 400$. END-KL: $\mathrm{KL}(p_{\mathrm{full}} \,\|\, p_{\mathrm{replaced}})$ at the output. SGD arms: Adam, 400 steps of $4 \times 512$ tokens, 20 warmup steps, learning rate $3 \times 10^{-4}$ (selected on validation by a 100-step CONT-ZERO run and reused unchanged for FID-SGD and END-KL), clip $1.0$; seeds set the minibatch stream only. FID-LS is fitted in closed form on 200,000 training tokens with ridge $10^{-3}$.

**Metrics.** Boundary fidelity is the per-coordinate MSE between substitute output and $h_{l+k}$ on WikiText-2 test; endpoint fidelity is the mean per-token $\mathrm{KL}(p_{\mathrm{full}} \| p_{\mathrm{replaced}})$; usefulness is held-out NLL (WikiText-2 test, LAMBADA) and LAMBADA last-word accuracy.

**Declared predictions.** P1 (not sufficient): on $\geq 4/5$ boundaries FID-LS has lower boundary MSE than every CONT-ZERO seed and higher NLL on both test sets. P2 (not necessary): on $\geq 4/5$ boundaries every CONT-ZERO seed has lower IID NLL than FID-LS with boundary MSE at least as large. P3 (trajectory): along CONT-LS, boundary MSE rises while IID NLL falls from step 0 to 400 on $\geq 4/5$. P4: pooling arms and checkpoints within a boundary, endpoint KL correlates positively with IID NLL on $\geq 4/5$, boundary MSE not positively on $\geq 3/5$. Falsification: FID-LS with IID NLL no higher than CONT-ZERO on $\geq 3$ boundaries.

**Results. P1 (fidelity not sufficient).** Final count **0/5** as pre-registered (both test sets): the IID clause holds on **5/5** boundaries (FID-LS is the most faithful by boundary MSE and still worse on WikiText-2 NLL than CONT-ZERO; at $(6, 2)$: MSE $1.52$ versus $3.01$, NLL $3.640$ versus $3.331$), while the shift clause holds on **0/5**: under LAMBADA the fidelity optimum keeps the lower NLL on every boundary, and the accuracy ordering reverses on only $2/5$. END-KL, the most faithful arm by output KL ($0.167$ at $(6, 2)$), is likewise not the most useful on IID ($3.576$).

**P2 (fidelity not necessary).** Holds at $(6, 2)$: CONT-ZERO, trained with no fidelity target through the same frozen continuation, reaches lower IID NLL ($3.331$) with larger boundary MSE ($3.01$) than FID-LS ($3.640$, $1.52$), and also beats the full model ($3.444$). The final count is **5 of 5** boundaries, both seeds where run.

**P3 (trajectory). 5/5** boundaries: along CONT-LS from step 0 to 400, boundary MSE rises while IID NLL falls at every checkpoint (at $(6, 2)$: $1.52 \to 2.55$ and $3.640 \to 3.325$; table 20); CONT-LS@400 ends at or below CONT-ZERO on every boundary, so the fidelity optimum serves as an *initialization* but not as a *criterion*. **P4 (pooled correlations).** P4a fails as declared, **0/5** with the sign reversed: pooled within a boundary, endpoint KL correlates *negatively* with IID NLL (Spearman $-0.19$ to $-0.88$; the arms closest at the output are the worst at the task), while P4b holds **5/5** (boundary MSE not positively correlated; $-0.25$ to $-0.59$).

**Falsification criterion.** Not triggered on any boundary (0/5): FID-LS IID NLL is higher than CONT-ZERO's everywhere; the criterion fires only if the reverse holds on $\geq 3$.

**The LAMBADA reversal.** Continuation fitting on 819,200 WikiText-2 token positions is domain adaptation toward the IID test distribution, so it need not transfer to LAMBADA; the fidelity-fitted arms degrade least under the shift (END-KL 3.869, FID-LS 3.971, versus 4.129). The reading consistent with the paper's thesis: usefulness is relative to the *declared* continuation and distribution, and neither direction of the fidelity–usefulness separation is distribution-free.

Table 19: GPT-2 small, blocks $[l, l+k)$ replaced by a residual linear map; held-out metrics, $(2)$ marks the mean of two seeds; other cells are single-seed. Boundary MSE is per coordinate at $h_{l+k}$; KL is per token; NLL in nats. Lower is better except accuracy. Full model: WikiText-2 test NLL 3.444, LAMBADA NLL 3.611. $(4, 4)$ CONT-ZERO: seed 1 only (its machine-1 seed-0 run was not executed; no prediction turns on it).

| Boundary / arm | MSE | KL | WT NLL | LAM NLL | LAM acc |
|---|---|---|---|---|---|
| (6,2) DROP | 2.77 | 0.418 | 3.879 | 3.974 | 0.234 |
| (6,2) FID-LS | 1.52 | 0.247 | 3.640 | 3.971 | 0.264 |
| (6,2) FID-SGD | 1.41 | 0.246 | 3.636 | 3.947 | 0.270 |
| (6,2) CONT-ZERO (2) | 3.00 | 0.553 | 3.332 | 4.137 | 0.282 |
| (6,2) CONT-LS@400 (2) | 2.55 | 0.531 | 3.326 | 4.202 | 0.281 |
| (6,2) END-KL (2) | 2.32 | 0.168 | 3.576 | 3.872 | 0.273 |
| (4,2) DROP | 1.24 | 0.538 | 4.015 | 3.929 | 0.186 |
| (4,2) FID-LS | 0.75 | 0.264 | 3.670 | 3.958 | 0.221 |
| (4,2) FID-SGD | 0.70 | 0.255 | 3.661 | 3.928 | 0.225 |
| (4,2) CONT-ZERO (2) | 1.56 | 0.549 | 3.322 | 4.054 | 0.233 |
| (4,2) CONT-LS@400 (2) | 1.35 | 0.550 | 3.322 | 4.071 | 0.240 |
| (4,2) END-KL (2) | 1.16 | 0.161 | 3.574 | 3.816 | 0.241 |
| (8,2) DROP | 7.25 | 0.403 | 3.914 | 3.951 | 0.271 |
| (8,2) FID-LS | 3.20 | 0.253 | 3.659 | 3.973 | 0.312 |
| (8,2) FID-SGD | 3.27 | 0.260 | 3.659 | 3.965 | 0.323 |
| (8,2) CONT-ZERO (2) | 6.81 | 0.589 | 3.354 | 4.281 | 0.303 |
| (8,2) CONT-LS@400 (2) | 5.09 | 0.540 | 3.341 | 4.410 | 0.303 |
| (8,2) END-KL (2) | 5.66 | 0.184 | 3.597 | 3.860 | 0.287 |
| (4,4) DROP | 5.13 | 1.383 | 4.910 | 4.538 | 0.070 |
| (4,4) FID-LS | 2.48 | 0.697 | 4.064 | 4.439 | 0.136 |
| (4,4) FID-SGD | 2.47 | 0.713 | 4.070 | 4.455 | 0.142 |
| (4,4) CONT-ZERO | 4.96 | 0.792 | 3.604 | 4.454 | 0.180 |
| (4,4) CONT-LS@400 (2) | 3.47 | 0.750 | 3.582 | 4.511 | 0.169 |
| (4,4) END-KL (2) | 4.54 | 0.442 | 3.822 | 4.197 | 0.175 |
| (6,4) DROP | 13.70 | 1.344 | 5.010 | 4.691 | 0.124 |
| (6,4) FID-LS | 5.97 | 0.695 | 4.075 | 4.551 | 0.197 |
| (6,4) FID-SGD | 6.14 | 0.725 | 4.086 | 4.524 | 0.242 |
| (6,4) CONT-ZERO (2) | 11.75 | 0.894 | 3.672 | 4.667 | 0.184 |
| (6,4) CONT-LS@400 (2) | 7.83 | 0.807 | 3.638 | 4.803 | 0.208 |
| (6,4) END-KL (2) | 11.05 | 0.490 | 3.887 | 4.249 | 0.180 |

Table 20: CONT-LS trajectory at $(6, 2)$, seed 0: the NLL objective started from the FID-LS solution (step 0); P3 holds: MSE rises while IID NLL falls at every checkpoint. "—": not completed at this writing.

| step | 0 | 25 | 50 | 100 | 200 | 400 |
|---|---|---|---|---|---|---|
| boundary MSE | 1.52 | 1.65 | 1.79 | 1.95 | 2.21 | 2.55 |
| WT test NLL | 3.640 | 3.463 | 3.415 | 3.371 | 3.349 | 3.325 |
| end KL | 0.247 | 0.354 | 0.422 | 0.432 | 0.502 | 0.523 |

**Fitting corpus (declared post hoc, after the five-boundary study).** The shift reversal above admits two readings: continuation training learns something that does not transfer, or it adapts the substitute to the one domain it was fitted on. A third declared protocol (supplement `vision_transfer_01/PROTOCOL.md`; hypotheses, predictions, and falsification committed before any run) separates them by refitting every arm at boundary $(6, 2)$, with the same runner, learning rate, steps, and 819,200 sampled positions, on two further corpora: *web*, the first 3,000 documents of FineWeb-Edu (Penedo et al., 2024); and *mix*, WikiText-2 train and *web* in equal token

shares. A third held-out set, the first 1,000 GSM8K test problems rendered as text (Cobbe et al., 2021), is a second shift that neither corpus resembles. Table 21 reports the result. On *web*, the no-fidelity substitute's gap to the unmodified model under shift falls from $0.52$ to $0.11$ nats and its loss under shift drops below the output-fidelity arm's ($3.72$ versus $3.79$; the order was reversed on WikiText-2), while its in-distribution advantage disappears ($3.72$ versus $3.44$); *mix* keeps the in-distribution advantage ($3.37$) and about two thirds of the shift improvement ($3.76$). Of the five declared predictions, one held (P3: on *web*, the no-fidelity substitute beats output-fidelity matching under shift) and four failed: P1 (on *web*, the no-fidelity substitute below the unmodified model in distribution: $3.72$ versus $3.44$), P2 (within $0.10$ nats under shift: missed by $0.005$–$0.012$, one CONT-LS seed met it), P4 (*mix* keeping both), and P5 (GSM8K within $0.15$ nats: gap $0.185$–$0.213$). The declared hypothesis that a single fitted substitute would be no worse than the unmodified model on both held-out sets was not met (closest: CONT-LS on *mix*, $3.35$ in distribution, $3.75$ under shift). The falsification criterion (worse than the untrained deletion under shift) was not triggered. What transfers is therefore not fidelity to the original computation as such, but whatever the continuation objective learns from a fitting distribution broad enough to cover the shift; the breadth of that distribution trades against in-distribution gain.

Table 21: Boundary $(6, 2)$ refitted on three corpora (held-out NLL, nats; two-seed means where two seeds were run; FID-LS and END-KL single seed). Full model: WikiText-2 test $3.444$, LAMBADA $3.611$, GSM8K $2.908$. WikiText-2 rows are refits with the same runner and seeds, which reproduce the five-boundary study's $(6, 2)$ values.

| Corpus | Arm | Boundary MSE | WikiText-2 test | LAMBADA | GSM8K |
|---|---|---|---|---|---|
| WikiText-2 | FID-LS | 1.52 | 3.640 | 3.971 | 3.371 |
| WikiText-2 | END-KL | 2.32 | 3.576 | 3.869 | 3.192 |
| WikiText-2 | CONT-ZERO | 3.00 | 3.332 | 4.137 | 3.609 |
| WikiText-2 | CONT-LS | 2.55 | 3.326 | 4.202 | 3.760 |
| web | FID-LS | 1.66 | 3.765 | 3.850 | 3.210 |
| web | END-KL | 2.46 | 3.676 | 3.785 | 3.134 |
| web | CONT-ZERO | 2.73 | 3.723 | 3.720 | 3.107 |
| web | CONT-LS | 2.13 | 3.722 | 3.711 | 3.099 |
| mix | FID-LS | 1.47 | 3.661 | 3.870 | 3.232 |
| mix | END-KL | 2.36 | 3.598 | 3.788 | 3.132 |
| mix | CONT-ZERO | 2.86 | 3.370 | 3.758 | 3.180 |
| mix | CONT-LS | 2.23 | 3.358 | 3.748 | 3.173 |

# O The Contract on a Pretrained Recursive Model (Declared Before Running)

Every real-model result above replaces two of twelve distinct blocks in GPT-2: segment replacement in a stack, not in a recursion. This appendix reports the same contract on a pretrained *recursive* language model, LoopFormer (Jeddi et al., 2026): three shared transformer blocks applied eight times, each pass conditioned on a time $t = i/8$ and a step size $dt$ (278M parameters; `armenjeddi/LoopFormer-3block-8iterations`). The recursion is load-bearing: WikiText-2 test NLL is $3.857$ with eight iterations and $3.932$ when iterations 3–4 are skipped (table 22, DROP). The model also contains a native coarse step: one call of the shared block with $dt{=}2/8$ in place of iterations 3–4 (NATIVE-$dt$) costs only $+0.009$ nats. The protocol, arms, hypotheses, and falsification were committed before any run (supplement `loopformer_01/PROTOCOL.md`); the learning rate ($3 \times 10^{-4}$) was chosen on WikiText-2 validation; the test splits were opened only by the frozen runner, which reproduces the model's own forward pass bit-exactly when the removed iterations are executed. Two machines ran the arms: seed 0 of every arm, NATIVE-$dt$, DROP, and CONT-LS seed 1 on the second machine; CONT-ZERO seed 1 on the first (section N provenance rule). The CONT-LS seed-1 file records a runner commit not archived in the repository; it was reproduced from the frozen runner at the declaring amendment and is marked as such in the supplement.

**Boundary, substitute, arms.** Boundary $(i, k) = (2, 2)$: iterations 2–3 of eight are removed; iterations 0–1 are the prefix, iterations 4–7 with the final norm and head are the frozen continuation.

The substitute is the residual linear map $x \mapsto x + Wx + b$ on the 2,048-dimensional residual stream (4.2M MACs per token against 151M per iteration; taking it saves $24\%$ of the recursion's MACs). Arms as in section N (400 steps of $4 \times 256$ tokens on WikiText-2 train, learning rate $3 \times 10^{-4}$ selected on validation by a 100-step CONT-ZERO run and reused for FID-SGD and END-KL; FID-LS in closed form on 100,000 tokens with ridge $10^{-2}$; seeds 0 and 1, varying minibatch order only, for CONT-ZERO and CONT-LS; END-KL and FID-SGD seed 0 only), plus NATIVE-$dt$, which needs no training. Of the three declared boundaries, $(2, 2)$ is complete; $(4, 2)$ has DROP, NATIVE-$dt$, FID-LS, and CONT-ZERO seed 0 (same ordering: CONT-ZERO 3.527 versus FID-LS 3.921 and the un-adapted model 3.857, boundary MSE 96.5 versus 2.8); $(2, 4)$ was not run before submission. The declared H4 (one replaced segment against uniform coarsening to four $dt$=1/4 steps) was not evaluated under the protocol.

**Results (table 22).** Both directions of the thesis hold at this boundary, with the largest in-distribution margin in the paper (two data-order seeds; no same-data adapter on the unmodified graph was run). *Not sufficient*: FID-LS has the lowest boundary MSE of every arm (0.53; NATIVE-$dt$ 0.56) and an output KL ten times smaller than the continuation arms', yet a higher NLL than every arm except FID-SGD and DROP. *Not necessary*: the two continuation-trained substitutes have boundary MSE 16.1 (CONT-ZERO) and 2.15 (CONT-LS), thirty and four times FID-LS's, and NLL 3.502 and 3.487 on both seeds (seeds agree to 0.003), below the unmodified model (3.857) by 0.37 nats and below the model's own native step (3.866). The learned substitute therefore beats the coarse step the model was trained with. Started from the fidelity optimum, continuation training raises boundary MSE and lowers NLL at every checkpoint on both seeds (table 23). Under shift the order reverses again: on LAMBADA, NATIVE-$dt$ (3.229) and FID-LS (3.301) are within 0.08 of the unmodified model (3.225) while CONT-LS is 0.28 worse.

**Transfer (declared amendment).** Refitting on FineWeb-Edu (first 10,000 documents, same steps and learning rate) gives CONT-LS LAMBADA 3.364 (gap to the unmodified model 0.14, from 0.28) but WikiText-2 test 3.972, above the unmodified model: on this model the broad corpus shrinks the shift gap by about half and removes the in-distribution gain, the same trade-off as on GPT-2 (table 21). One seed; the declared target (within 0.10 under shift with the in-distribution gain kept) was not met.

**Selector (declared amendment).** A gate reading only $x_2$ (the residual stream entering the boundary, a hidden state rather than the raw input), fitted on WikiText-2 validation per-token labels and with usage thresholds frozen there, routes each token to the substitute or to the two native iterations (table 24). On WikiText-2 test the gated model is below the un-adapted model at every usage tested (linear gate, $63\%$ usage: 3.561 versus 3.857, $15.5\%$ of the recursion's MACs saved by count; no latency was measured), so it meets rule (d)'s cost criterion in distribution; but the ungated substitute meets it with lower loss (3.487) and a larger saving ($24\%$), so every gated point is dominated in distribution, and the gate's only possible value is under shift, where it recovers part of the substitute's loss (mlp gate, $21\%$ usage: LAMBADA 3.292 against 3.520 ungated and 3.225 unmodified) without reaching the unmodified model. The declared selector hypothesis S1 (some usage at which the gated model is no worse than the unmodified model on both held-out sets) was therefore not met. The per-token oracle at $63\%$ usage is 3.307. Gate cost: 2,048 MACs (linear).

**What this establishes.** The contract's two directions and the trajectory result on a pretrained recursive model at one boundary, with the model's own coarse step as an additional reference that the learned substitute is below in distribution. It does not establish that the substitute beats an equally adapted model (no same-data adapter was run), a useful selector, sparse-expert routing (LoopFormer shares one block and has no experts), transfer under shift, or scale; one boundary complete of three declared, one model.

Table 22: LoopFormer, iterations 2–3 replaced; held-out metrics; two-seed means marked (2). DROP skips the iterations without advancing time; NATIVE-$dt$ is the model's own block applied once with $dt{=}2/8$. Boundary MSE per coordinate at $x_4$; KL per token; NLL in nats. Unmodified model: WikiText-2 test 3.857, LAMBADA 3.225.

| Arm | Boundary MSE | End KL | WikiText-2 test | LAMBADA | Acc. |
|---|---|---|---|---|---|
| DROP | 18.42 | 0.076 | 3.932 | 3.281 | 0.324 |
| NATIVE-$dt$ | 0.56 | 0.005 | 3.866 | 3.229 | 0.366 |
| FID-LS | 0.53 | 0.062 | 3.897 | 3.301 | 0.348 |
| FID-SGD | 7.24 | 0.332 | 4.172 | 3.530 | 0.250 |
| END-KL | 15.09 | 0.054 | 3.881 | 3.276 | 0.356 |
| CONT-ZERO (2) | 16.09 | 0.604 | 3.502 | 3.551 | 0.343 |
| CONT-LS (2) | 2.15 | 0.580 | 3.487 | 3.502 | 0.353 |

Table 23: CONT-LS trajectory on LoopFormer, boundary MSE / WikiText-2 test NLL at each checkpoint, both seeds. Step 0 is the fidelity optimum.

| Seed | 0 | 25 | 50 | 100 | 200 | 400 |
|---|---|---|---|---|---|---|
| 0 | 0.53/3.897 | 0.82/3.644 | 1.05/3.585 | 1.31/3.540 | 1.66/3.507 | 2.16/3.487 |
| 1 | 0.53/3.897 | 0.82/3.623 | 1.07/3.589 | 1.32/3.545 | 1.69/3.512 | 2.14/3.487 |

Table 24: Input-only gate on LoopFormer with the seed-0 CONT-LS substitute; thresholds frozen on validation for nominal usage $q$; realized usage and recursion-MAC saving on WikiText-2 test. Unmodified model: WikiText-2 test 3.857, LAMBADA 3.225; substitute alone: 3.487 / 3.520.

| Gate | $q$ | Usage | MAC saving | WikiText-2 test | LAMBADA |
|---|---|---|---|---|---|
| linear | 0.3 | 40% | 9.8% | 3.645 | 3.351 |
| linear | 0.5 | 50% | 12.4% | 3.604 | 3.380 |
| linear | 0.7 | 63% | 15.5% | 3.561 | 3.414 |
| $2048 \to 64 \to 1$ | 0.3 | 22% | 5.5% | 3.687 | 3.292 |
| $2048 \to 64 \to 1$ | 0.5 | 30% | 7.3% | 3.647 | 3.315 |
| $2048 \to 64 \to 1$ | 0.7 | 89% | 22.0% | 3.504 | 3.488 |

Table 25: LoopFormer, iterations 2–3, arms refitted on FineWeb-Edu (seed 0).

| Arm | Boundary MSE | End KL | WikiText-2 test | LAMBADA | Acc. |
|---|---|---|---|---|---|
| FID-LS | 0.65 | 0.084 | 3.949 | 3.290 | 0.360 |
| END-KL | 15.14 | 0.065 | 3.913 | 3.284 | 0.318 |
| CONT-ZERO | 15.57 | 0.147 | 3.998 | 3.407 | 0.322 |
| CONT-LS | 1.36 | 0.116 | 3.972 | 3.364 | 0.334 |

# P THE EIGHT-EXPERT PROTOTYPE: PROTOCOL, CONTROLS, AND SELECTOR

The prototype implements the lowest-order retained path statistic (occupancy) and the shared continuation interface in one trained model. It has eight experts, predicts an interval occupancy, and trains a span-conditioned macro with composition and continuation objectives. Its selector is fitted *after* these blocks are frozen, so the prototype integrates occupancy and continuation learning but does not jointly adapt the expert bank, macro, and selector.

## P.1 EXACT COMPUTATION AND LEARNING INTERFACE

In this section, $z$ denotes a width-16 hidden representation and $q \in \Delta^7$ an interval occupancy. For an input $x \in \mathbb{R}^4$, an encoder gives $z_0$. The fine prefix executes two shared, internally rerouted residual experts:

$$a_t = \arg\max_e \rho_e(z_t, t), \qquad z_{t+1} = E_{a_t}(z_t), \qquad q_F = \tfrac{1}{2}\big(\mathbf{e}_{a_0} + \mathbf{e}_{a_1}\big), \quad t \in \{0, 1\}, \tag{43}$$

where $\mathbf{e}_a$ is the one-hot vector for expert $a$ and $z_F = z_2$. The prefix router $\rho$ encodes phase $t/2$; macro phase and span inputs are likewise divided by two. Each expert has a width-48 hidden layer. The width-24 macro receives the starting hidden state, phase, and span, and predicts $(z_M, q_M) = C(z_0, 0, 2)$. Its occupancy head is normalized by softmax. Macro deployment does not execute the skipped prefix to obtain $q_F$. The predicted $q_M$ describes that skipped interval; it is not the occupancy of the single continuation expert actually executed by the fast path.

Both branches use the same subsequent router, expert bank, and two-output head:

$$p_b = \mathrm{softmax}\big(W[z_b, q_b, 1] + \beta\big), \quad a_b = \arg\max_e p_{b,e}, \quad o_e^b = H(E_e(z_b)), \qquad b \in \{F, M\}. \tag{44}$$

Here the final coordinate is the normalized common continuation phase. The common router has weights $W$ and bias $\beta$, and $H$ is the shared readout. Deployment evaluates only $o^b_{a_b}$. Training instead evaluates the finite candidate menu and minimizes $\sum_e p_{b,e}\|o_e^b - y\|_2^2/2$, averaged over examples. Prefix routing uses a hard-forward straight-through surrogate; the continuation uses this soft expected loss during training and hard argmax at deployment. Neither surrogate is an exact gradient of hard dispatch.

The joint base objective contains fine and macro task losses, $0.05$ times fine–macro endpoint MSE, and $0.01$ times load balancing. The integrated objective adds, each with coefficient $0.1$:

1. an occupancy anchor against stop-gradient current $q_F$;
2. state and occupancy composition MSE between $C(z_0, 0, 2)$ and two unit-span macros, with the second evaluated at the first macro's successor and phase 1, and the two occupancies averaged;
3. candidate-response discrepancy, either uniformly averaged over the eight experts or weighted by detached $w_e \propto \exp[-\|o_e^F - y\|_2^2/(2T)]$, with $T = 0.1$.

MSE averages over the coordinates of the respective object. Composition and response penalties differentiate both operands. The response penalty has no direct gradient through the continuation-router logits; that router learns through task and load terms. The fine occupancy target changes with the coadapted fine program. It is not an immutable teacher route or a conserved physical quantity.

Only the span-2 map is anchored to a fine endpoint and occupancy. The two span-1 maps are factorization auxiliaries, not individually calibrated fine one-step replacements. Moreover, $C(z, s, \ell)$ does not take an arbitrary incoming routing history. These facts rule out interpreting this prototype as closed product-state dynamics over arbitrary replacement schedules.

## P.2 PROSPECTIVE DESIGN AND COMPLETE CONTROLS

Two development parents, 3301 and 3307, use 4,096 training and 1,024 development examples, 600 fine-pretraining updates and 600 adaptation updates. Only the uniform and relevance-weighted full arms are eligible for selection. Their mean hard-macro development MSEs are $0.016450$ and $0.016220$, respectively, fixing the latter (`RELEVANT`) before confirmation; these values are recorded

in `protocol_confirmation.json`, whereas the development fits themselves are not in the supplement. The two development parents are not included in the final five-parent means.

Confirmation parents are 3407, 3413, 3433, 3449, and 3461. Each uses 8,192 training, 2,048 development, 4,096 independent identically distributed (IID) test, and 4,096 range-shift test examples. Inputs are drawn uniformly from $[-1, 1]^4$. With $b_*(x) = \sin(2x_1) + x_2/2$, the two-output target is $y = (\sin(2x_3) + x_1x_2,\ x_3 + x_4^2)$ if $b_*(x) \geq 0$ and $y = (\cos(2x_3) - x_1x_2,\ x_3 - x_4^2)$ otherwise; the range shift multiplies the first two input coordinates by $1.25$ before $y$ is computed, and deployment observes neither the branch nor the target. Fresh, disjoint random streams are used, but this is still one synthetic population family, not a new real domain. Fine pretraining and adaptation each take 1,200 Adam updates at learning rate $0.002$, batch size 128, with final checkpoints only. Parent initializations and minibatch streams are paired across applicable arms.

`JOINT-END` uses the joint base. `SIG` adds occupancy and composition; `RESP-ONLY` adds uniform response coupling without those signature terms. `SHORTCUT-CONT` adds state-only composition and agreement between each branch's own selected next output, allowing different expert identities. It is a local repair-aware shortcut adaptation, not a reproduction of a published language-model benchmark. `ROUTER-RELEVANT` and `FROZEN-RELEVANT` use the full RELEVANT objective but train macro plus routers, or macro alone, respectively; encoder, experts, and head stay fixed in both. `REFERENCE` is the pretrained fine model (the pretrained fine reference of the main text), and `ADAPTED-FINE` continues fine-only task learning from it (continued fine-only in the main text).

Both `DIRECT` arms start from the common raw initialization, never from a freely supplied pretrained fine model, and use exactly the candidate's fast deployment graph. Their summary coordinates may learn any task-useful code within the same simplex parameterization; they have no occupancy supervision. `DIRECT-DATA` receives 2,400 updates; `DIRECT-COMPUTE` receives 7,910, matching the candidate's standalone *forward-affine MAC* budget up to integer rounding (shortfall $0.01047\%$). A multiply–accumulate (MAC) is an analytical affine-operation count here, not full training FLOPs: backward passes, nonlinearities, optimizer work, memory, and dispatch are not included. DIRECT uses uninterrupted Adam; inherited arms reset Adam after pretraining. These differences are explicit, not exact total-training-cost equivalence. Each standalone method is charged its pretraining, while the shared parent fit is counted once in the total compute.

All 55 confirmation fits and development-dependent choices were frozen before either test distribution was opened. Five trained parents, not individual examples or timing repeats, are the uncertainty unit. Every parent result and descriptive paired interval is in the supplement (`confirmation_01/`, `core_analysis_01/RESULTS.json`).

### P.3 Fixed replacement: implementation succeeds, utility screen fails

Table 26: Fixed-path results for the eight-expert prototype on the original family: MSE averaged over five independently trained parents (lower is better). Fast paths cost 2,840 forward-affine MACs per example, fine paths 5,176. Every control is shown, including those that win.

| Arm (macro) or fine reference | IID MSE | Range-shift MSE |
|---|---|---|
| Pretrained fine reference | 0.014432 | 0.020641 |
| Continued fine-only | 0.009712 | 0.012471 |
| Candidate's coadapted fine path | 0.006445 | 0.010635 |
| JOINT-END macro | 0.007917 | 0.012174 |
| SIG macro | 0.009520 | 0.014035 |
| RESP-ONLY macro | 0.007721 | 0.011579 |
| SHORTCUT-CONT macro | 0.008943 | 0.013466 |
| **RELEVANT macro (candidate)** | **0.009164** | **0.013767** |
| ROUTER-RELEVANT macro | 0.020199 | 0.026695 |
| FROZEN-RELEVANT macro | 0.024476 | 0.030664 |
| DIRECT-DATA macro | 0.005537 | 0.008918 |
| DIRECT-COMPUTE macro | 0.003866 | 0.006075 |

Table 26 reports every fixed-path control.

The candidate beats both restricted-adaptation arms in all five parents on both distributions, but loses to JOINT-END and both DIRECT controls in every parent on both distributions (table 26 reports the

ratios). The prespecified whole-system screen, requiring at least $5\%$ improvement over each of six controls, therefore fails. SHORTCUT-CONT and RESP-ONLY also have lower mean MSE on both distributions; unlike the JOINT-END and DIRECT comparisons above, no every-parent claim is made for these two mean rankings. These outcomes do not require interpreting any single auxiliary edge as indispensable, nor isolate occupancy as the unique cause of the regression.

The candidate macro's IID error sits above its own fine path and below continued fine-only training (table 26); comparing only against the latter would hide compression loss in the coadapted system. The declared replacement criterion requires staying within $2\%$ of *both* references. The measured batch-one mean-of-parent-median latency ratio yields the fine/macro speed reported in section 6, but the quality criterion fails: a speed–quality trade-off, not quality-matched acceleration.

The two-unit-span composition has IID MSE $0.057287$ and range-shift MSE $0.059216$, versus $0.009164$ and $0.013767$ for the span-2 macro. An implemented composition penalty has not produced a task-valid multiscale family in this run. Uncalibrated unit spans are a design limitation; the observations alone do not establish a unique causal explanation.

### P.4 Dynamic headroom versus a deployable selector

Before test scoring, a separate subprotocol fits frozen-bank selectors. For development example $i$, let $L_F(i)$ and $L_M(i)$ be the two complete path losses and $d_i = L_M(i) - L_F(i)$. At exactly $k$ macro choices, a label-informed oracle has optimal empirical loss

$$\overline{L}_F + \frac{1}{N}\sum_{j=1}^{k} d_{(j)}, \qquad d_{(1)} \le \cdots \le d_{(N)}. \tag{45}$$

The exchange argument is immediate: exchanging a selected larger difference for an unselected smaller one cannot increase loss. This standard cardinality bound is an offline diagnostic, not a deployable controller or novel theorem. Equal-size development sets give equal parent weights; the constraint concerns global average cost, not per-input deadlines.

Including 96 selector MACs per input, the candidate's development oracle satisfies the prespecified headroom criterion (counts, MSE, and saving as in section 6), so all selector controls were run. It neither predicts attainable test performance nor licenses label access at deployment.

Each selector sees only raw $x$, predicts $(L_F, L_M)$, and selects macro when its predicted loss difference is below a fixed threshold. The primary multilayer perceptron (MLP) is $4 \to 16 \to 2$ with tanh and 96 MACs; a mandatory linear $4 \to 2$ control uses 8 MACs. Both train for 600 Adam updates on risk labels computed on the training split only, by executing both frozen paths offline. This label-generation work is charged. Candidate, JOINT-END, SHORTCUT-CONT, ROUTER-RELEVANT, and FROZEN-RELEVANT each receive both selector kinds for every parent: 50 fits in total. One pooled development threshold per arm bank and selector kind maximizes saving subject to the $2\%$ quality tolerance against both own fine and continued fine-only. Four additional development-quantile thresholds are frozen for descriptive operating curves, not test selection.

Table 4 reports both candidate selector rows.

The primary selector lowers error relative to always taking the macro but misses the declared saving target; the simpler linear selector trades more error for latency reduction (table 4 reports all quantities). Both also fail the range-shift quality tolerance; range-shift latency was not measured. All ten primary policies (five arm banks, two selector kinds) fail the pooled quality requirement on each test distribution. Fallback constant-fine policies retain their actual selector overhead and can fail quality relative to continued fine-only. The supplement (`dynamic_01/`, `dynamic_analysis_02/`) holds all control and operating-point results, not only the candidate rows.

**Scope.** The prototype study covers integration, fixed-path testing against strong controls, offline headroom analysis, and a legal selected-only selector; it does not establish a quality-matched whole-system advantage. Broader schedules, a real second domain, and generative flow matching are untested here, neither established nor ruled out. A useful next hypothesis must explain when correct replacement decisions are cheaply learnable and when composition remains task-valid. The oracle gap does not prove irreducible uncertainty or a unique causal defect.

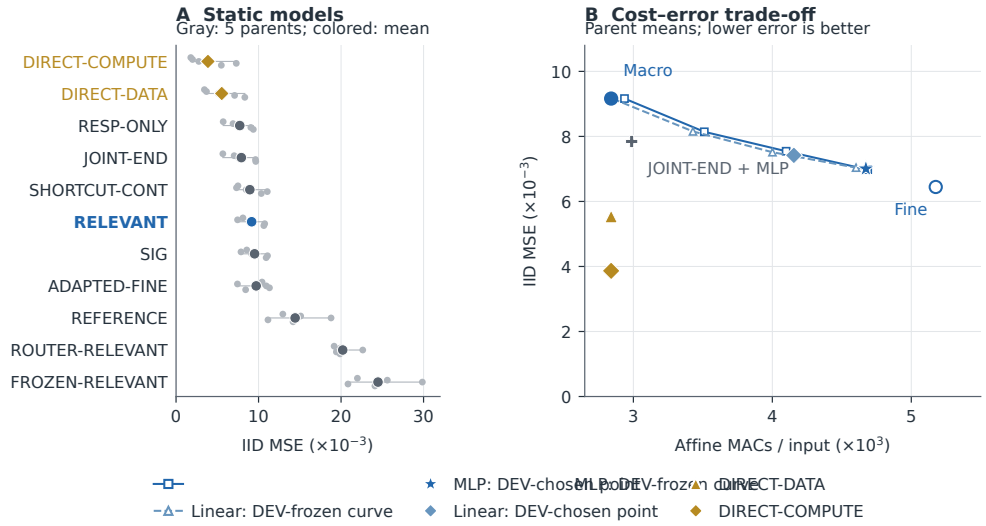



Figure 2: **Complete controls and the measured quality–work trade-off.** (A) Five parent-level IID MSEs and their mean for each fixed-path arm. (B) Candidate static paths, frozen-threshold selector operating points, and strong direct controls; means use the same five parents. Selector curve points use development-frozen thresholds, not test-retuned quantiles. Affine MACs exclude backward work and non-affine deployment overhead; these curves are not measured-latency frontiers. No label-informed test oracle or inferred population confidence band is shown. Label key: ADAPTED-FINE = continued fine-only; REFERENCE = pretrained fine reference; DEV = development split (DEV-chosen point: the threshold chosen on development data; DEV-frozen curve: thresholds fixed on development data).

**Verification and reproducibility.** After all fits were frozen, verification scripts recomputed 1,945,600 per-example MSEs of the fixed paths (`core_analysis_01/RESULTS.json`) and 409,600 per-example selector training labels (`dynamic_analysis_02/DYNAMIC_RESULTS.json`). Check counts describe software/numerical coverage, not independent scientific replications. The frozen protocols, per-parent summaries, thresholds, timing samples, verification scripts, and source hashes are in the supplementary archive; per-example prediction arrays and checkpoints exceed the size limit and are represented by their SHA-256 digests.

## Q Earlier Exploratory Studies (Excluded from the Evidence)

Two groups of earlier studies shaped the prototype's design, but no claim in this paper rests on them. Their code, configurations, and per-seed results come from an earlier code base that is not in the supplement, and several of their summary statistics cannot be regenerated from what was retained. This section therefore states what each group tested and what it taught, without quantitative results. Table 27 lists every construction in this appendix with its evidence status.

### Q.1 Soft-mixture signature studies

The first studies trained the signature-supervised formulation (section B) on two small analytic systems: a two-dimensional state driven by a context-conditioned mixture of four vector fields over eight integrator steps, and an eight-instruction recursive program over four contractive nonlinear operators whose instruction sequence is given as context. Each compared a local one-step expert model, a smaller dense shortcut, a state-only multi-width expert model, and the signature-supervised model with endpoint, occupancy-anchor, and composition losses, over three seeds. All expert models used the tied soft special case (equation (21)) and evaluated every expert, so these studies concern

Table 27: Constructions in this appendix and their evidence status. Only rows marked as evidence support claims in the paper; their numbers are regenerable from the supplement.

| Construction | What is learned and executed | Status |
|---|---|---|
| Soft-mixture signature studies (sections Q.1 and Q.2) | Multi-width endpoint and occupancy predictor; tied soft mixtures that evaluate every expert | Excluded; results not in the supplement |
| Four-expert continuation studies (section Q.3) | Rerouted two-step prefix, macro, one selected continuation expert; selected-only deployment | Excluded; results not in the supplement |
| Eight-expert prototype (sections H to J and P) | Endpoint and occupancy macro, shared continuation, input-only selector fitted after freezing | Evidence |
| Discrete construction (section G) | Exhaustively enumerated summaries and controllers | Evidence |
| Ordered signature; broader scheduling (section S) | Boundary-corrected transition statistic; horizon-wide control | Algebra and invariant checks only; never trained |

occupancy identification, not sparse execution or speed. They taught three things that the main text builds on: endpoint-only training leaves routing underdetermined (lemma 1), whereas an occupancy anchor recovers much more of the fine path's allocation; the anchor's effect on endpoint quality depends on the system; and a small occupancy-cocycle defect can come from a nearly constant router, so composition consistency has to be read together with fidelity to the fine path.

### Q.2 Cross-domain signature study

A follow-up applied the same formulation to an analytic pick–transport–place simulation (not a robot), to a small weight-shared recursive Transformer with four experts trained on WikiText-2 (Merity et al., 2016), and to an evaluation-schedule probe of a public looped Transformer without an expert router (Jeddi et al., 2026). It asked whether joint state–route training improves the quality–compute frontier, improves fidelity to the fine path's occupancy, transfers to unseen interval partitions, and helps more as depth and expert specialization grow. Occupancy fidelity improved in both local systems; frontier and held-out-partition outcomes were mixed and depended on the metric; and no benefit of depth, or of depth combined with specialization, appeared. Random streams were not paired across variants, the complete ablation used a single seed, and the pooling behind several summaries was not recorded, so the study is descriptive and excluded from the evidence. It motivated keeping route fidelity, task quality, and cost as separate measurements.

### Q.3 Four-expert continuation studies

Two studies tested whether a fine program, its executable macro, and the immediate continuation can adapt together. They used the prototype's population (section P) with a four-expert bank, a rerouted two-step fine prefix, a residual macro, and one selected continuation expert; deployment executed only selected experts. The first study (three parents) compared frozen and jointly adapted objectives that match the endpoint, the candidate losses (VALUE), or the candidate outputs and centered router scores (RESPONSE), together with shuffled-correspondence, detached-gradient, extra-trained-fine, dense, and same-graph DIRECT controls. The second (five fresh parents) crossed the auxiliary objective (VALUE or RESPONSE) with whether its gradient reached the shared parameters or only the macro, while base gradients trained the shared system in every arm. In both studies the directly trained fast graph had lower mean macro error than the jointly adapted VALUE and RESPONSE arms, and the RESPONSE macro, although faster than its own fine path, stayed outside the $2\%$ quality tolerance of that path on every parent. In the first study, fine and macro chose the same continuation expert on most inputs without preserving quality; in the second, the gradient-recipient intervention showed no stable RESPONSE-specific benefit. These outcomes motivated the prototype's same-graph DIRECT controls, reported on every parent, and its split of the continuation gap at a fixed continuation choice (proposition 10). Their protocols and result files are not in the supplement, so neither study is part of the evidence.

## R Supplemental Interface and Comparison Caveats

**Comparisons within the prototype.** The earlier soft-mixture and four-expert studies are excluded from the evidence (section Q). Within the prototype, applicable arms share each parent's initialization and minibatch stream, but equal numerical weights on different discrepancies do not equalize their strength or backward computation. Useful next-expert allocation relative to a development-selected fixed expert, and a useful prefix relative to deleting it, would not establish an optimal expert count or specialization superior to direct learning. Hard-dispatch deployment is trained with richer all-candidate information and surrogate routing gradients (section P).

**Scope of retained information.** Teacher identity can have operational meaning, but preserving its routes can also preserve its faults; functional replacement need not recover identical expert identities. The ordered transition statistic distinguishes some orders, not arbitrary operator words. The support floor concerns occupancy in the same expert coordinates, not task-quality impossibility (section E). The prototype does not accept arbitrary incoming routing history. Hidden-history dependence would require a sufficient boundary variable or distributional model before claiming closed product-state dynamics. The oracle-menu bound in proposition 10 assumes bounded per-action mismatch and bounds the gap only up to the regret difference; empirical mean response loss does not establish its premise.

**Limits of analogies and extensions.** Neural coarse-graining asserts neither thermodynamic conservation nor a universal variational principle. A sequential agent interface would have to account for asynchronous tools, permissions, side effects, and variable durations; these change the mathematical object and are not validated by the one-step construction. Shared recursion increases effective depth, not stored parameter capacity or guaranteed large-model knowledge. Physical control time, generative-flow time, and recursive depth remain distinct scopes (section A.5), and fewer evaluations imply a latency improvement only after end-to-end measured execution, including dispatch when sparse.

**Application motivation, not reproduced systems.** Diffusion Policy, ActionFlow, and FlowPolicy motivate iterative action generation (Chi et al., 2023; Funk et al., 2024; Zhang et al., 2024); Octo, OpenVLA, and $\pi_0$ are generalist policy precedents (Octo Model Team et al., 2024; Kim et al., 2024; Black et al., 2024). PAMAE and CoRDE use structured expert routing in policy generation (Yang et al., 2026; Huang et al., 2026a). These are not head-to-head reproductions or evidence of robot deployment here. OLMoE and Pythia are sparse/dense language-model resources (Muennighoff et al., 2024; Biderman et al., 2023), not matched systems against which this prototype establishes superiority.

**A separate gating diagnostic.** An exploratory diagnostic on a separate continuous task compared a cost-aware gated model ($F_\lambda$) with DIRECT controls under one feature-cost functional applied identically to every arm. $F_\lambda$ was ahead of a DIRECT control with its gates forced open but not consistently ahead of a gated DIRECT control given the same sparsity prior, so the diagnostic does not separate the structured objective from that prior. Its result files are not in the supplement, and it supports no claim in this paper.

## S Proposed Extensions (Not Executed)

None of the extensions below has been run, and none is evaluated in this paper. They state what a broader claim than the prototype's would require, in addition to repeating the prototype's controls on a trained sparse-expert language model (section M). Every choice below would be fixed on development data before any confirmatory evaluation.

**Language-model mechanism study.** A looped sparse Transformer (untied prefix and suffix blocks around shared recursive sparse blocks repeated eight times, with eight experts and top-2 routing) would compare the signature-supervised objective with untied and recursive dense models, the uncompressed sparse teacher, early exit, endpoint-only and state-only-composition training, route-anchor-only training, and dense controls matched separately by active FLOPs and by stored parameters, at one, two, four, and eight evaluations.

**Retained path statistic.** Development data would select the smallest adequate statistic: occupancy; occupancy with first and last route; boundary-corrected transitions (equation (13)); or a learned path summary.

**Causal value of routing.** The primary intervention would permute actual top-$k$ assignments while preserving marginal expert counts, top-$k$ cardinality, and capacity pressure, holding checkpoint, input, starting state, and expert bank fixed and letting the hidden trajectory change.

**Second real domain and systems.** A policy-learning benchmark such as LIBERO (Liu et al., 2023) would supply a real second domain, and a sparse-execution claim would require end-to-end latency percentiles, throughput, memory, and dispatch overhead measured on a named target accelerator.

**Training cost.** An equal-total-training-compute track would count fine-path tracing, direct and composed calls, backward passes, and tuning. A finite break-even deployment count

$$N_{\mathrm{break}} = \frac{C_{\mathrm{train,full}} - C_{\mathrm{train,base}}}{c_{\mathrm{infer,base}} - c_{\mathrm{infer,full}}} \tag{46}$$

exists only when both the additional training cost and the quality-matched inference saving are positive.

# T SCOPE OF CLAIMS

Table 28: Closest mechanism overlaps (none reproduced head-to-head); claim scope in section T.

| Prior work | Existing mechanism | Present focus |
|---|---|---|
| LoopMoE | Shared iterative experts; conditional routing | Internal replacement and subsequent computation |
| Latent Flow Transformer | Transport replacing a layer block | Path statistics and the continuation menu |
| SFTN / SAR | Student-friendly teacher / student-aware router | Rerouted prefix and macro share later experts |
| LLM-Streamline / ReplaceMe / Ghosted Layers | Removed block replaced by a learned or closed-form map fit to boundary hidden states | Routed continuation after replacement; same-graph from-scratch (DIRECT) control |

Table 29: Analytical claims and their support. Established statements refer only to the stated assumptions. Prototype results are in table 30.

| Claim | Evidence and status | Boundary |
|---|---|---|
| State/occupancy/ordered-signature composition | Exact identities and executable checks | Signature sufficiency and learned benefit do not follow. |
| Omitted-memory and sparse-support limitations | Elementary counterexamples and conditional bounds | Not universal task-quality impossibility results. |
| Continuation loss decomposition and sufficient bound | Section D; inline exact proof | Standard scaffold, fixed menu and assumptions; no learned generalization certificate. |
| Earlier soft-mixture and four-expert studies | Qualitative summary only (section Q); results not in the supplement | Excluded from the evidence; no claim rests on them. |
| Bounded contribution and prior-work boundary | Continuation-relative interface, conditional analysis, and controlled comparisons; table 28 locates mechanism overlap | The fine/macro/continuation assembly and four-object/three-clock organization are not claimed novel; no priority certificate. |
| General empirical superiority | No matched large-model, real-robot, or new real second-domain continuation study | Untested; no state-of-the-art or universal-acceleration claim is made. |

**Scope.** These tables support the main text's bounded contribution: a continuation-relative evaluation contract, conditional analysis, and controlled evidence. The prior-work boundary is in table 28. The work does not establish broad empirical superiority.

Table 30: Eight-expert prototype, original family (section P): five training parents and a pre-specified range shift. Values are in the supplement (`core_analysis_01/RESULTS.json`, `dynamic_analysis_02/DYNAMIC_RESULTS.json`).

| Claim | Evidence and status | Boundary |
|---|---|---|
| Integrated replacement interface executes | Independently checked sparse occupancy/macro/continuation implementation; 55 core and 50 frozen-bank selector fits | Fixed eight-expert, top-1 menu; not joint controller–bank coadaptation. |
| Broader adaptation can help | Candidate macro beats both restricted-adaptation controls in all five parents on both splits | Ordinary joint training and both DIRECT controls beat it in every parent; no unique causal attribution. |
| Quality-matched fixed compression | IID macro MSE $0.00916408$ versus own fine $0.00644494$; $0/5$ parent quality passes on each split | Measured fixed-macro batch-one CPU speedup versus own fine; the joint quality criterion against both fine references fails. |
| Whole-method superiority | IID candidate/control MSE ratios are $1.158$ (joint), $1.655$ (DIRECT-DATA), and $2.371$ (DIRECT-COMPUTE) | The six-control superiority screen fails; forward-affine budget matching is not total training-FLOP matching. |
| Finite-menu selection headroom | Development-label-informed oracle reaches the declared quality target with $43.01\%$ affine-MAC saving | Optimistic offline bound, not a deployable or held-out achievement. |
| Cheap learned selection succeeds | IID MLP: $8.70\%$ more error than own fine, $9.71\%$ proxy saving, $23.78\%$ slower; linear: $15.11\%$ more error, $19.73\%$ proxy saving, $10.32\%$ faster | Both fail the quality target; timings are IID-only ratios of mean parent p50s. |
| Useful multiscale deployment | IID two-span-1 MSE $0.0572872$ versus span-2 MSE $0.00916408$ | Negative; individual span-1 maps were not fine-anchored. |
| Universal or flow-based advantage | No flow-matching, real second-domain, or broader-scheduling experiment was run | Untested; neither established nor ruled out. |