Phase selects; it does not amplify: a preregistered test of oscillatory control as expert selection

Kyle W. Luecke
Driftbound Labs
ORCID: 0009-0008-0202-7852

September 2026

Abstract

Recordings of prefrontal cortex describe a division of labor in which slow rhythms coordinate which information is active while synapses store it: beta as a spatial "stencil" of inhibition that holds a task set and is interrupted by gamma bursts carrying content. We asked whether this division of labor does useful work when built from arithmetic inside a small mixture-of-experts transformer, and we preregistered every test. Across twelve preregistered experimental stages on a synthetic rule-following task, four results survive.

First, an always-on selection state with fixed dynamics (one leaky, weakly coupled unit per expert) holds an expert-selection regime after control input stops: $T_{\text{hold}} = 26.8$ tokens at $n = 8$ and 65.0 tokens at $n = 3$ (effective spectral radius $\rho_{\text{eff}} = 0.989$), compared to under 2 tokens for a stateless controller. The hold time is insensitive to intrinsic unit leak alone (the fourfold λ sweep is nearly flat at 28.5 / 26.8 / 24.2), indicating that persistence reflects competitive survival under drive rather than isolated single-unit decay. The same persistent property costs offset sensitivity and suppression of rarely used experts, demonstrating a single mechanism with three faces.

Second, when experts write into a shared complex-valued buffer tagged by phase and a reference demodulates it, selection follows the predicted physical law: crosstalk scales as $X(\Delta\varphi) \approx \max(0, \cos\Delta\varphi)$ (an opposed expert contributes zero), a reference drifting with the tags maintains selection while a fixed reference collapses (PLV 0.84 vs 0.10), and usable phase separation under jitter follows the slot law measured earlier at a physical receiver. Coherent experts do not, however, amplify or add: a preregistered ensemble-gain test was null ($G_{\text{medium}} = 0.044$ vs $G_{\text{sum}} = 0.251$).

Third, at the model level, persistent coordination provides no benefit over lookup when evidence is continuous: on clean task sequences, the held state paid a 0.42-segment switch lag and a 29% deficit; in looped depth budgeting (T12, closing mixed), the stateless ACT controller matched the oracle budget, so the persistent coordinator had nothing to add ($\Delta\text{ACC} = 0$), though overthinking was regime-dependent with extra depth degrading easy tasks across 5 of 5 seeds ($T = 12$ vs $T^* = 4$, $\Delta\text{ACC} = -0.3418$); and under evidence gaps with mid-gap lures (T12b), held state beat one-step memory (0.6525 vs 0.2987) but did not beat simple $k = 3$ majority smoothing (0.6390), falling into the preregistered smoothing-class. Fourth, in looped mixture-of-experts architectures under matched compute (T13), tagged superposition works as an attribution medium: expert removal effects are decodable from the aggregate alone (H_{int} supported: $r_{\varphi} = 0.87$ for phase tags vs 0.80 for random tags, -0.70 for plain sum, and 0.41 for no-medium attribution), so most of the attribution comes from tagging itself, with phase adding a small margin; the context-slice ablation was null (shared context tied slicing, closing T13 as mixed).

In post-hoc exploratory analyses, earlier apparent phase advantages were explained by simpler non-oscillatory mechanisms: noise robustness in T7 was matched by input smoothing (`LR-smoothed`), the 0.47-nat clean loss gap in T9 disappeared under matched training of surrounding layers, and on ambiguous sequences in T11.3, where the counted hum arm failed (0.308 vs classifier 0.351) and the preregistered provider-parity arm scored 0.789 (hum_measure), a post-hoc exploratory stateless argmax over the observation window scored 1.000. One experiment evaluating coherence as a classifier of plastic learning consistency (T10) was not informatively tested due to a tag-placement side-channel, with the classifier claim withdrawn. We conclude that the oscillatory division of labor selects and attributes without storing or amplifying, and functions as an interpretability medium and smoothing

integrator among equals rather than a privileged execution decider. Preregistrations, hashes, results, a from-frozen-commit rerun of the persistence experiment after its original artifacts were lost, and a process account of two freezes that went through with a required check undone are released with the code (T9 and T10 training scripts excepted).

Contributions

1. **Preregistered persistence and its governing mechanics**: A demonstrated persistence result for a fixed-dynamics selection state ($T_{\text{hold}} = 26.8$ tokens at $n = 8$; 65.0 tokens at $n = 3$, $\rho_{\text{eff}} = 0.989$), with evidence that hold duration is governed by coupled competitive survival under incoming drive rather than isolated single-unit decay, alongside three measured costs (offset sensitivity, minority suppression, and switch lag).
2. **Network confirmation of receiver selection laws and ensemble null**: Confirmation inside a transformer of the physical selection law ($X(\Delta\varphi) \approx \max(0, \cos\Delta\varphi)$), slot capacity under jitter, and drift robustness with a shared reference; paired with a preregistered null on ensemble gain ($G_{\text{medium}} = 0.044$ vs $G_{\text{sum}} = 0.251$).
3. **Interpretability payoff in looped architectures (T13)**: Demonstration that tagged superposition in a looped Mixture-of-Loops enables decodable expert attribution under strict matched compute (H_{int} supported across all 5 seeds: $r_{\varphi} = 0.869$ vs 0.801 for random tags, -0.695 for plain sum, and 0.412 for no-medium attribution); random tags reach 0.801, so phase contributes a $+0.068$ margin over tagging alone, though context-slice stratification added no benefit over shared context (null).
4. **Model-level boundaries and falsification of coordinator oracles**: Evidence that persistent gating confers no advantage over lookup when evidence is continuous (0.42-segment switch lag, 29% deficit in T11.2); in looped depth budgeting (T12), the stateless ACT controller matched the oracle budget, leaving nothing for the persistent coordinator to add ($\Delta\text{ACC} = 0$), while overthinking was regime-dependent (easy tasks degraded across 5 of 5 seeds, closing T12 as mixed); and under evidence gaps (T12b), held state beat one-step memory but failed to beat $k = 3$ majority smoothing, falling into the preregistered smoothing-class.
5. **Exploratory attribution of apparent phase advantages**: Ablation accounts attributing apparent phase benefits to classical mechanisms: noise robustness to input smoothing, clean loss gaps to surrounding layer training, and sequence disambiguation to observation window length.
6. **Process record**: Full disclosure of an experiment closed as not informatively tested (T10, with the classifier claim withdrawn due to a tag-placement side-channel) and its root-cause diagnosis; an account of two freezes (T10, T11.1) that went through with a required check undone, before the gated freeze script existed, and the rule they produced (a freeze script that emits a hash only when every required gate passes); and a rerun of T8 from its frozen preregistration after the original artifacts were lost.

Introduction

Neurophysiological recordings in primate prefrontal and sensory cortex describe a functional division of labor in which slow rhythms coordinate which neural pathways are active while synaptic weights store long-term representational content. Specifically, slow oscillations in the alpha and beta frequency bands (8–30 Hz) establish spatially structured "stencils" of top-down inhibition that define active task sets, punctuated by transient gamma-band bursts (>30 Hz) that route local information through coherent phase alignment (Fries, 2015; Akam & Kullmann, 2012, 2014; Miller et al., 2018; Lundqvist et al., 2016, 2023). Under this communication-through-coherence framework, oscillatory dynamics do not store the underlying knowledge; rather, they dynamically orchestrate routing topology.

In contemporary deep learning, by contrast, the dominant paradigms for conditional computation—most notably Mixture-of-Experts (MoE) architectures—rely almost exclusively on feedforward, stateless routing. At every token step, a parameter-heavy routing projection (typically a dense linear layer followed by softmax or top-k selection) recomputes expert affinities from scratch based entirely on immediate token activations. While effective for parameter scaling, feedforward routing is inherently memoryless: it possesses no intrinsic continuity of the active task regime, cannot sustain an expert regime through transient evidence gaps without dedicated recurrent state or attention mechanisms, and entangles routing selection with continuous representation learning.

In this work, we ask a foundational architectural question: **Can an oscillatory division of labor, constructed from elementary fixed-dynamics arithmetic and phase-tagged superposition, perform useful computational work inside a transformer?**

To answer this question without hindsight bias, we conducted twelve successive experimental stages (T5 through T13, alongside physical receiver and oscillator sheet validations), pre-committing every hypothesis, metric, and decision rule to cryptographically hashed preregistrations before inspecting outcome data. Our findings across these twelve stages yield four findings, including definitive negative results:

1. **Persistence set by the effective spectral radius**: An always-on, fixed-dynamics selection state (one leaky, weakly coupled unit per expert) sustains an expert regime across tens of tokens after incoming control drive ceases ($T_{\text{hold}} = 26.8$ tokens at $n = 8$; 65.0 tokens at $n = 3$ with $\rho_{\text{eff}} = 0.989$ under the rebuilt coupling assembly), whereas a standard feedforward router collapses in under 2 tokens. A fourfold sweep over the intrinsic decay parameter λ produces a nearly flat hold response (28.5 / 26.8 / 24.2 tokens), demonstrating that intrinsic decay does not govern hold time; to explain this divergence, a competitive survival model is proposed as a mechanistic hypothesis for future testing.
2. **Phase selects and separates, but does not amplify**: When expert outputs write into a shared complex-valued medium tagged by distinct phase angles and a reference receiver demodulates the superposition, the channel obeys classical physical demodulation laws: crosstalk scales as $X(\Delta\varphi) \approx \max(0, \cos\Delta\varphi)$, an opposed expert contributes zero, and a drifting shared reference maintains coherent selection (PLV 0.84 vs 0.10 for a static reference). However, coherent superposition yields zero constructive ensemble gain ($G_{\text{medium}} = 0.044$ vs $G_{\text{sum}} = 0.251$; null). Phase superposition provides a selective projection filter, not an additive amplifier.
3. **Tagged superposition as an attribution medium in looped architectures**: In looped Mixture-of-Loops architectures evaluated under matched compute budgets ($C_{\text{match}} = 4,210,688$), phase superposition serves as a transparent attribution medium: the contribution of individual experts is directly decodable from the aggregate state alone (H_{int} supported across all 5 seeds: correlation $r_{\varphi} = 0.869$ vs 0.801 for random orthogonal tags, -0.695 for plain sum, and 0.412 for unmediated baseline). Random tags recover most of this (0.801), so the capability belongs mainly to tagging, with phase contributing a $+0.068$ margin.
4. **Where a lookup wins**: When input evidence is continuous, persistent coordination imposes severe penalties: a 0.42-segment switch lag and a 29% performance penalty compared to stateless lookup (T11.2); in looped depth budgeting (T12), an Adaptive Computation Time (ACT) controller matches the oracle budget, leaving nothing for persistent coordination to add ($\Delta\text{ACC} = 0$); and across evidence gaps with deceptive lures (T12b), the held state beats one-step memory (0.6525 vs 0.2987) but fails to beat simple 3-step majority smoothing (0.6390), categorizing the mechanism squarely into the preregistered `smoothing-class`.

Through systematic post-hoc ablations, we further show that earlier apparent phase benefits—such as noise robustness and loss margins—are explained by conventional non-oscillatory mechanisms: moving-average input filtering and matched training of surrounding layers. We conclude that the oscillatory division of labor acts as a robust smoothing integrator and interpretability medium among equals, rather than an oracle execution controller.

Related Work

Biological Foundations of Oscillatory Routing

The hypothesis that neural communication is dynamically modulated by coherence originates in the work of Fries (2005, 2015), which posits that rhythmic phase synchronization aligns excitability peaks between sender and receiver populations. Empirical investigations in primate prefrontal and visual cortex have detailed how laminar frequency hierarchies operate: deep-layer alpha/beta rhythms serve as inhibitory stencils that constrain sensory processing, while superficial-layer gamma bursts convey bottom-up content (Akam & Kullmann, 2010, 2012, 2014; Miller et al., 2018; Lundqvist et al., 2016, 2023). Working memory models based on discrete gamma bursts within beta attractor states (Lundqvist et al., 2011; Lisman & Jensen, 2013) provide the mechanistic blueprint for the tonic hum and phasic burst distinction evaluated here.

Dynamic Routing and Continual Learning in Artificial Networks

Mixture-of-Experts architectures traditionally employ learned softmax or top-k gating networks. While these gating networks achieve parameter efficiency and capacity scaling, their feedforward nature requires computing routing decisions anew at every step. Recent investigations into memory-augmented networks and continuous learning

systems (Kumaran et al., 2016; Ben-Iwhiwhu et al., 2022; Hong & Pavlic, 2022) have explored neuromodulatory and reservoir architectures (Tadros et al., 2022). Our work differs by implementing routing via fixed linear dynamical systems with non-learned dynamics and evaluating them inside standard transformer blocks under strict matched compute.

Signal Demodulation, Phase Tagging, and Physical Media

Modulating distinct information streams onto orthogonal or phase-separated carriers is a standard signal processing principle. In over-the-air computation and continuous physical media, superposition allows multi-access aggregation without separate wiring. We translate these continuous physical principles into complex-valued representations within deep networks, evaluating whether trigonometric phase separation ($e^{i\varphi_e}$) functions as a viable routing and attribution substrate.

Architectural Formulation

The oscillatory division of labor is formalized into three modular substrates:

1. **Substrate 1: The Phase Gate (Phasic Control)**
   Incoming token activations x_t are mapped to a phase angle $\theta_t \in [0, 2\pi)$ via a fixed projection. Selection of an expert e with target phase φ_e is computed via half-wave rectified cosine similarity:
   $$g_e(x_t) = \max\left(0, \cos(\theta(x_t) - \varphi_e)\right)$$
   This provides zero cross-talk for orthogonal ($\Delta\varphi = \pi/2$) or opposed ($\Delta\varphi = \pi$) experts without learned gating parameters.

2. **Substrate 2: The Tonic Hum State (Persistent Selection)**
   To sustain expert selection across token intervals without continuous feedforward drive, the gating state evolves as a coupled linear dynamical system with resting baseline $a_0 = 0.1$:
   $$h_{t+1} = (1 - \lambda)(I + C)h_t + \lambda a_0 \mathbf{1} + g_{\text{tonic}} \cdot u_t$$
   where $h_t \in \mathbb{R}^n$ represents the activation vector across n candidate experts, $\lambda \approx 0.0255$ is the intrinsic unit leak, $a_0 = 0.1$ is the resting baseline amplitude, C is a fixed, non-learned coupling matrix with spectral radius < 1 (characterizing T8 at $n = 8$ by its measured hold time $T_{\text{hold}} = 26.8$ tokens, and rebuilt in T11 at $n = 3$ with effective linear spectral radius $\rho_{\text{eff}} = \rho((1 - \lambda)(I + C_3)) = 0.989$ and $T_{\text{hold}} = 65.0$ tokens), and u_t is the incoming token-level evidence drive modulated by gain g_{tonic}. When drive ceases ($u_t = 0$), the active expert regime persists until incoming competition or decay alters top-1 rank.

3. **Substrate 3: The Phase Medium (Superposition and Demodulation)**
   Instead of concatenating or independently executing experts, candidate experts write into a shared complex-valued aggregation buffer $z_t \in \mathbb{C}^d$:
   $$z_t = \sum_{e=1}^{n} x_e(t) \cdot e^{i\varphi_e}$$
   A downstream receiver or read head extracts the specific output of expert e by projecting onto the conjugate reference phase:
   $$\hat{x}_e(t) = \text{Re}\left(z_t \cdot e^{-i\varphi_{\text{ref}}}\right)$$
   where φ_{ref} can be fixed or dynamically phase-locked to a pilot tone.

Method

1. The receiver and the sheet (motivating experiments)

A lock-in receiver simulation (positive-only pulsed senders, phase-multiplexed streams, a PLL locked to a control pilot) measured slot capacity under jitter (9–18 usable slots at $\sigma_j \approx 0$, falling toward 4–8 near $\sigma_j \approx 0.1$ cycle),

showed that signed carriers produce a U-shaped crosstalk curve while rectified ones reject anti-phase, and that a reference locked to the pilot holds BER ≈ 0 under drift while content-locked and fixed references fail. A 16×16 phase-oscillator sheet with a spatial stencil and pilot was then compared to a learned top-k router on eight multiplexed streams, with the readout referenced to pilot phase (an absolute-time readout bug was found and fixed pre-primary and disclosed).

2. The toy transformer and task

Four layers, width 128, eight experts ($128 \to 256 \to 128$ feed-forward) at layer 3, trained from scratch. Task: eight rule systems over a 64-token vocabulary with regime markers and mid-sequence rule flips; T5/T6 used an order-1 task (losses ≈ 0.0013 nats, at ceiling), T7 onward an order-3 task (learned-router loss ≈ 1.27 nats; the generator's entropy floor reported beside it). Five seeds; controller parameters fixed from single-system tests before each freeze.

3. Controllers

- Learned router (top-k linear over the hidden state; the standard).
- Hash router (Roller et al.).
- Smoothed learned router (32-token running mean of its input; the T7 addendum arm).
- Phase gate, supplied: fixed expert phases, reference from the task label (T5/T6).
- Phase gate, inferred: reference from a smoothed hidden-state projection (T7).
- The hum (T8): one leaky unit per expert with a nonzero resting amplitude, weak symmetric coupling with spectral radius < 1, tonic gain, drives from named inputs (regime probe, token evidence, affinities, surprise); gate readout by amplitude; width by arousal. Reimplemented from the frozen preregistration for the rerun after the original tree was lost; parameters $\lambda \approx 0.0255$, $a_0 = 0.1$; effective spectral radius recorded.
- Stateless neuromodulator (same inputs and readout, no state).
- Phase medium (T9): experts write $f_e \cdot e^{i\varphi_e}$ into a complex buffer; readout $\mathrm{ReLU}(\mathrm{Re}(B \cdot e^{-i\varphi_{\mathrm{ref}}}))$; random-tag mirror arm; hum, fixed, and high-passed-hum reference sources.

4. Model-level harness, loop budget, and evidence gaps (T11–T12b)

T11 series (model-level orchestration). Three specialists behind the SCP boundary (A3-Qwen3.5-2B text-only; Qwen2.5-Coder-1.5B; Qwen2.5-1.5B under the reference prompt), task families with exact-match oracles, sequences of segments with clean, ambiguous, distraction, renewal, and window-break types; orchestrators: hum ($n = 3$, rebuilt coupling C_3, $\rho_{\mathrm{eff}} = 0.989$, $T_{\mathrm{hold}}(n = 3) = 65$ tokens), trained context classifier, hand rule (human-signed before generation), and (T11.3 addendum) stateless argmax of the per-step ScoreAffinity signal. Rule-19 freeze script: gates for ambiguity, distraction, disjoint classifier seeds, real execution (no stubs), per-arm retry logs; the hash is emitted only when all gates pass. Affinity `affinity_inputs.py` remained frozen at sha `7dae877a3975c1ba4ca264f03f572d6749cce39a114ea3adbcffc010c7852c50` (no C_3 / hum retune).

T12 (the loop budget). Prelude → recurrent block → coda at width 128 with $T_{\mathrm{max}} = 12$; arms Fixed $T = 7$, Fixed sweep, ACT, Convergence threshold, coordinator regime, coordinator regime+traj, oracle, and Stateless regime. Calibration locks: mean compute $\bar{C} = 7 \pm 0.25$; ACT/BP centers easy$= 4$ / hard$= 10$. H4 operational silencing zeros or holds the regime estimate at the controller input (estimate may still be computed for receipts). One preregistered primary × five seeds under PREREG sha `bf757ee4…`; Affinity / C_3 / hum untouched; no T12.2 (gate defect false).

T12b (persistence under evidence gaps). Routing-only; sealed T11.3 generators reused. Arms: counted primary `hum_measure`, null baseline `one_step_memory`, labeled secondary `majority_k3`, gap control `stateless_argmax`. Evidence gaps $G \in \{3, 6, 12, 24\}$ with a locked single-frame mid-gap ScoreAffinity lure; orthogonal lags $L \in \{2, 4, 8\}$. Primary decision rule: $\mathrm{CSR}_{\mathrm{gap}}(\mathrm{hum}) - \mathrm{CSR}_{\mathrm{gap}}(\mathrm{one_step_memory}) >$ seed interval 0.05 (null fail ⇒ support); if hum beats one-step but not majority ⇒ smoothing-class (PREREG §4). One primary × five seeds under PREREG sha `851124b3…`; Affinity frozen; no T12b.2.

5. MoE-of-loops / H_{int} (T13) and statistics / provenance

T13 (looped architectures). New line after T11–T12b close: MoE-of-loops with per-expert depth and optional context-slice stratification at matched compute on `mythos_tiny` ($K = 4$, $W = 64$, $S_{\mathrm{full}} = 64$, $L_{\mathrm{max}} = 4$). Primary

iso-cost reference $C_{\text{match}} = 4{,}210{,}688$ (per-slice attention counted; ±5% per arm before hash). Primary metric H_{int}: Pearson r between predicted and actual expert-removal loss deltas recovered from the phase-medium aggregate alone, scored against plain-sum, random-tag, oracle, and no-medium attribution controls (three-clause support rule; seed interval 0.05). Context-slice arm (`moe_of_loops_slices`) vs shared-context (`moe_of_loops_shared_ctx`) tested whether slicing buys intelligibility over shared context. Acc gates G9/G13/G17 before primary; one primary × five seeds under PREREG sha `3176eb5e…`; Affinity `7dae877a…` untouched; no T13.2.

Statistics and provenance. Bootstrap over seeds; dual rule for a pass (point estimate exceeds the interval width and the interval excludes zero) on early tickets; by-construction clauses named in each report; effect sizes reported beside base loss; preregistration hashes and Drive anchors; artifact inventories with hashes at every close; agent-executed runs with human-signed freezes; a rerun from the frozen preregistration for T8 after artifact loss. OpenTimestamps for the twelve Paper 2 preregistrations were backfilled `anchored_retroactively` on 2026-09-24 (Disclosure #9).

6. What the method does not do

No biological claim is tested; the recordings motivate the designs and are cited as such. No claim at scale; the toy is a mechanism instrument. Affinity / C_3 / hum were not retuned after freeze (`7dae877a…`).

Preregistered Hypotheses and Verdicts

Quoted in preregistration terms; hashes in the appendix. Confirmatory verdicts are agent-drafted, pending human sign-off.

- **Gating experiment (oscillator sheet; PREREG sha `c81851b0…`; RESULTS sha `73cacbcf…`)**: H1 (physical stencil + pilot beats a learned router on activity cost and drift grace): **mixed**. Clean accuracy and activity cost passed; relative-drift grace at $\sigma_c = 0.05$ rad/cycle failed; under jitter the physical gate collapsed (0.56 vs 0.995). Secondary: a router given the pilot matched the router without it (ACC 0.995), so the null is not "the router lacked a reference."
- **T5/T6 (phase gate, task supplied; PREREG sha `eff56961…`)**: H1 **mixed**: clean tie, specialization, cost (8 ops vs 1,024 MACs), and noise passed; shift failed the dual rule. H2 (faster switching) **supported (by construction)**: a controller handed the goal switches instantly (0.000 vs 1.51 tokens).
- **T7 (phase gate, task inferred; PREREG sha `6d79f794…`)**: H1 **mixed**: clean, specialization, and noise passed; shift failed the dual rule. H2 **H0 (not supported)**: PG-inferred did not switch faster than the learned router (3.87 vs 3.29 tokens).
- **T8 (always-on state; PREREG sha `d975d96c…`; rerun after artifact loss)**: H1 **mixed** (clean and specialization pass; noise and shift vs prespecified comparator LR-smoothed fail the dual bootstrap rule). H2 (persistence) **supported**: $T_{\text{hold}} = 26.8$ (hum, $n = 8$) vs 2.1 (stateless) vs 1.6 (learned), dual rule; λ clause monotone directionally (28.5 / 26.8 / 24.2 for a fourfold λ change), but the near-flat slope shows the leak is not the dominant time constant. H3 (salience) not tested. H4 (capacity) $K_{\text{max}} = 8$ at $E = 8$, unobservable ceiling.
- **T9 (phase medium; PREREG sha `1c6a4867…`; RESULTS sha `d1c70486…`)**: H1 (ensemble gain) **H0**: $G_{\text{medium}} = 0.044$ vs $G_{\text{sum}} = 0.251$. H2 (crosstalk within ± 0.1 of $\max(0, \cos \Delta\varphi)$) **supported**. H3 (shared-reference drift robustness; slot law under jitter) **supported**. H4 (rare-expert reach) **H0**: $\min \text{Rec}_e = 0.333$ vs LR-smoothed 0.3647.
- **T10 (coherence as reconciliation classifier; PREREG sha `fd4cb718…`; RESULTS sha `9a10d0ee…`)**: **Not informatively tested**: oracle decodable from tag placement (three tag-based arms at 1.000, zero variance); loss baseline failed to generalize out-of-domain. Closed as not informatively tested, with the classifier claim withdrawn.
- **T11 series (model-level orchestration)**:
  - T11.1 (PREREG sha `c62bfda8…`): routing-only dry run; H1/H2 not informatively tested; H3 supported; H5 not run.
  - T11.2 (PREREG sha `593762c4…`): specialists executing with operating $T_{\text{hold}}(n = 3) = 65$ ($\rho_{\text{eff}} = 0.989$); H1 untested; H2 uninformative; **H3 supported with the bill** (0.42-segment lag; clean CSR 0.708 vs 1.000); **H4 null** (0.6 on return, 0.2 on trailing content); H5 not run.
  - T11.3 (PREREG sha `27ab4a95…`): step-grain scoring after the evidence point. Quoted H1 clause: *"Ambiguous CSR at step grain, steps >= evidence point: hum >= classifier; hum > hand"*. §2 designates `hum` as

the counted primary arm (`hum_measure` was designated as a preregistered secondary arm for provider-parity disclosure: *"keep if needed for provider-parity disclosure; not H5"*). Outcome: on the preregistered counted arm `hum`, ambiguous step CSR was 0.308, failing both clauses against `classifier_mlp` (0.351) and `hand_rule` (0.348); token-level thrash was uninformative. **H1 closed as unsupported**, and the preregistered stopping rule closed the model-level orchestration line. (The preregistered provider-parity arm `hum_measure` scored 0.789; see Exploratory Analyses).

- **T12 (the loop budget; PREREG sha `bf757ee4…`)**: **Mixed**.
  - H1 (regime budgeting) & H2 (trajectory budgeting vs ACT): **null** — the stateless ACT controller matched the oracle budget, so the persistent coordinator had nothing to add (ACT = oracle, $\Delta\text{ACC} = 0$). Cost-matched budget ($T = 7$) had no headroom to improve hard tasks.
  - H3 (overthinking regime-dependent): **supported** — extra depth degraded easy tasks across 5 of 5 seeds ($T = 12$ vs $T^* = 4$, $\Delta\text{ACC} = -0.3418$).
  - H4 (persistence under silenced estimate): **null**.
- **T12b (persistence under evidence gaps; PREREG sha `851124b3…`)**:
  - *Preregistered Decision Rule (§4.1.4 / §4.2 / §5)*:
    "CSR_gap(hum) − CSR_gap(one_step_memory) > seed interval (null fail ⇒ support)." "Interpretation: if hum beats one-step and beats majority ⇒ held state; if hum beats one-step but not majority ⇒ smoothing-class."
  - *Empirical Outcome*:
    - **Primary Contrast (`hum_measure` vs `one_step_memory`)**: CSR_{gap} was 0.6525 for `hum_measure` vs 0.2987 for `one_step_memory` ($\Delta = +0.3538 >$ seed interval), formally rejecting the one-step memory null.
    - **Prespecified Labeled Secondary Control (`majority_k3`)**: CSR_{gap} was 0.6390 ($\Delta = +0.0135$ vs hum), failing to clear the seed interval (hum did not beat majority).
  - **Verdict**: **Primary supported; per the preregistered interpretation rule, smoothing-class.**
- **T13 (MoE-of-loops / context slices; PREREG sha `3176eb5e…`)**: **Mixed**.
  - H_{int} **(intelligibility metric under matched compute $C_{\text{match}} = 4,210,688$)**: **Supported across all 5 seeds** under all three clauses ($r_\varphi = 0.869$ vs 0.801 random tags [margin $+0.068$], -0.695 plain sum, and 0.412 no-medium attribution).
  - **Context-slice ablation**: **Null** — context-slice stratification (`moe_of_loops_slices`) tied shared context (`moe_of_loops_shared_ctx`), showing that context slicing does not improve intelligibility over a shared context at matched compute.
  - *(Nonlinear crosstalk suppression is PROPOSED as a mechanistic hypothesis rather than an independently scored endpoint).*

Exploratory Analyses & Post-Hoc Addenda (Not Preregistered)

- **T7 `LR-smoothed` (Noise Robustness Attribution)**: An exploratory addendum evaluated a learned router provided with a 32-token moving average of the input signal. It matched or beat the phase gate on noise ($\Delta L = 0.0005$ vs 0.0026), demonstrating that the noise robustness previously attributed to oscillatory dynamics was an artifact of input low-pass filtering.
- **T9 `dense_sum` and `matched_train` (Loss Gap Attribution)**: An initial 0.47-nat clean loss gap favored the phase medium over plain sum (0.833 vs 1.301). An exploratory `dense_sum` control ruled out routing density (1.301), but retraining with a matched training surface (`matched_train`) completely eliminated the gap, attributing the advantage to unequal training of surrounding layers rather than phase mechanics.
- **T11.3 `stateless_argmax` (Observation Window Attribution)**: On ambiguous task sequences, the preregistered provider-parity disclosure arm (`hum_measure`, evaluated with `ScoreAffinityProvider` over the post-evidence window) scored 0.789 vs `classifier_mlp` 0.351. Beside this preregistered disclosure, a post-hoc exploratory addendum evaluated a stateless argmax over that same post-evidence observation window. The stateless read scored 1.000 CSR, beating the held state (1.000 vs 0.789). This showed that a stateless read over the window outperformed persistent integration, consistent with the coordinator's apparent advantage over the classifier being an artifact of the observation window rather than sustained internal state.

Initial Results

1. Receiver and Sheet

Slot capacity 9–18 at sigma_j approx 0 with positive pulsed senders; jitter at sigma_j = 0.05 measured 7.65–8.35 slots; signed carriers U-shaped, pulsed senders reject anti-phase; fixed reference BER approx 0.37 under 0.25 Hz offset; pilot-locked PLL BER approx 0 across drift and 0–8 Hz mismatch. Sheet versus router: physical gate bit accuracy 0.56 under sigma_j = 0.10 vs router 0.995 (pilot-referenced) and 0.82 (no pilot); relative drift grace failed at sigma_c = 0.05.

2. T5/T6 and T7 (Phase Gate)

Arm	Clean L	T_sw	ΔL Noise 0.3	ΔL Shift 0.3
T5/T6 LR	0.00134	1.51	0.00061	—
T5/T6 LR-goal	0.00131	0.067	0.00030	—
T5/T6 PG (supplied)	0.00137	0.000	0.00000	—
T7 LR	1.267	3.29	0.0038	0.0151
T7 LR-goal	1.267	0.53	0.0021	0.0175
T7 HR1	1.297	1.03	0	0.0029
T7 PG-inferred	1.272	3.87	0.0026	0.0060
T7 LR-smoothed (addendum)	1.277	29.6	0.0005	0.0221

3. T8 (The Hum; Rerun and Hold Time Reconciliation)

Arm	Clean L	T_hold	ΔL Noise 0.3	ΔL Shift 0.3	min Rec_e
Hum (rerun, n=8)	≈1.27	26.8	≈ stateless	worst	0.287
Stateless	+0.034	2.1	≈ hum	best	—
LR	1.267	1.6	0.0038	0.0151	—
LR-smoothed	1.277	≈30	0.0005	0.0221	0.365

Note on Hold Times across Experiments:

- **T8 (n=8 Rerun)**: T_hold = 26.8 tokens (measured under frozen rerun; lambda sweep: 28.5 / 26.8 / 24.2). The historical ∼41 tokens value from prior reports is unrecoverable and retained for audit only.
- **T11 (n=3 Rebuilt Coupling C_3)**: T_hold = 65.0 tokens with effective spectral radius rho_eff = 0.989 (diagonal-only hold 39 approx 1/lambda).

4. T9 (The Medium)

Crosstalk X(Delta_phi) within the ±0.1 band of max(0, cos Delta_phi) at every sweep point; selective accuracy under sigma_c = 0.05: hum reference within 0.05 of clean (PLV 0.84), fixed reference below (PLV 0.10); usable separation grows with jitter per the slot law. Ensemble gain G_medium = 0.044 vs G_sum = 0.251 (H0); min Rec_e = 0.333 (H0). Clean L medium 0.833 vs plain sum 1.30: dense sum 1.301 (density ruled out); matched-surface retrain closes the gap (training explains it).

5. T10 (Not Informatively Tested)

Agree 1.000 ± 0.000 for `coh_hum`, `coh_fixed`, `random_tag`; loss-delta 0.333; identical learning outcomes across the three tag arms. Oracle decodable from tag placement; closed as not informatively tested with the classifier claim withdrawn (PREREG sha `fd4cb718…`; RESULTS sha `9a10d0ee…`).

6. T11 series (model level)

Measure	hand	classifier	hum	hum_measure	stateless argmax
T11.2 clean CSR	1.000	1.000	0.708	0.708	—

Measure	hand	classifier	hum	hum_measure	stateless argmax
T11.2 switch lag (segments)	0.0	0.0	0.42	0.42	—
T11.2 renewal, A returns	1.0	1.0	0.6	0.6	—
T11.2 trailing unrelated	1.0	1.0	0.2	0.2	—
T11.3 ambiguous step-CSR after evidence	0.348	0.351	0.308	0.789	**1.000**

Token-level thrash ≈ 0–0.04 for all arms (uninformative). T11.2 PREREG `593762c4…`; T11.3 PREREG `27ab4a95…`. Operating $T_{\text{hold}}(n = 3) = 65$ with $\rho_{\text{eff}} = 0.989$.

7. T12 (the loop budget)

Overall **mixed** (PREREG `bf757ee4…`). Arm aggregates (seed-mean ACC̄ / C̄ / LPC̄): coordinator_regime, coordinator_regime_traj, ACT, convergence, oracle, and stateless_regime all **0.9718 / 7.0 / 0.1388**; fixed_t7 **0.9710 / 7.0 / 0.1387**.

H	Verdict	Key numbers
H1 Regime budgeting	**null**	ACT = oracle (ΔACC $= 0$); Fixed $\Delta\text{ACC} = +0.0008$
H2 Trajectory budgeting	**null**	traj vs ACT/CONV $\Delta\text{ACC} = \Delta\text{LPC} = 0$
H3 Overthinking regime-dependent	**supported**	Easy degrades past T^* **5/5** seeds; $\Delta\text{ACC}(T = 12$ vs $T^* = 4) = -0.3418$, $\Delta\text{NLL} = +1.5982$; hard **0/5**
H4 Persistence under silencing	**null**	SD $\Delta\text{ACC} = 0$; coordinator beats Stateless **0/5**

Seeds `[20260921…20260925]`. Affinity `7dae877a…` untouched; no T12.2.

8. T12b (evidence gaps)

Overall **mixed** — primary **supported** vs `one_step_memory`; secondary ≈ `majority_k3` ⇒ **smoothing-class** (PREREG `851124b3…`).

Arm	CSR_gap	Role
`hum_measure`	0.6525	primary counted
`one_step_memory`	0.2987	primary null baseline
`majority_k3`	0.6390	labeled secondary
`stateless_argmax`	0.0000	gap control

$\Delta(\text{hum} - \text{one_step}) = +0.3538$ (> seed interval 0.05); $\Delta(\text{hum} - \text{majority}) = +0.0135$ (≤ 0.05). Seeds `[20260923…20260927]`. Affinity `7dae877a…` untouched; no T12b.2.

9. T13 (MoE-of-loops / H_{int})

Overall **mixed** (PREREG `3176eb5e…`). Matched compute $C_{\text{match}} = 4{,}210{,}688$. Seed-mean H_{int}:

Metric	Value
r_{φ} (slice)	0.8686
r_{plain}	–0.6945
r_{random}	0.8014
$r_{\text{no_medium}}$	0.4124
r_{oracle}	0.9604
r_{φ} (shared_ctx)	0.8663
slice beats shared (> 0.05)	**False**

H_{int} **supported** on every seed (slice and shared_ctx both clear the three-clause rule); context-slice stratification **null** vs shared context. Acc gates G9/G13/G17 **PASS**. Seeds `[20260928…20261002]`. Affinity `7dae877a…` untouched; no T13.2.

10. Replication note

T8 is the only early ticket with a from-preregistration rerun forced by artifact loss. Later tickets (T11.2–T13) release frozen PREREG hashes, RESULTS, and closings under `experiments/phase-routing/`.

Discussion: Boundary of the Claim & Proposed Mechanisms

1. Persistence, Spectral Radius, and the Competitive Survival Hypothesis (PROPOSED)

Persistence is real, tunable through the coupled system's effective spectral radius ($\rho_{\text{eff}} = 0.989$ in the $n = 3$ T11 assembly), and represents the identical physical property as offset memory and minority suppression. However, the empirical hold-time response under parameter sweeps reveals that persistence is not governed by single-unit exponential relaxation: a fourfold sweep over the intrinsic unit leak λ produced a nearly flat hold response (28.5 / 26.8 / 24.2 tokens), directly contradicting an isolated decay model of the form $T \propto 1/(\lambda + c)$. Furthermore, network coupling produced opposing effects across system sizes: coupled $n = 8$ shortened hold time relative to uncoupled diagonal decay (26.8 vs. 39 tokens), whereas coupled $n = 3$ lengthened it (65.0 tokens).

We propose a competitive survival model to explain this divergence: T_{hold} does not measure the passive decay of the selected unit toward baseline, but rather its duration of survival against competing units rising under incoming drive. In an un-driven system where all units decay at the same rate λ toward a common resting baseline a_0, the active unit's rank advantage would persist indefinitely. In execution, the state transition is governed by competition under drive, where the flip time depends on input drive gain and coupling sign (inhibitory vs. excitatory). Because T8 (coupled $n = 8$, hold 26.8) and T11 (coupled $n = 3$, hold 65.0, diagonal-only hold 39) evaluated different coupling matrices and architectures across experiments rather than a controlled size sweep, we propose competitive survival as a mechanistic hypothesis for future controlled testing rather than an established empirical reversal.

2. Phase Selects; It Does Not Amplify (Demodulation Mechanics and the Ensemble Gain Null)

Across our physical receiver experiments and network evaluations in T9, the phase medium adhered strictly to the predicted trigonometric selection law: crosstalk between experts separated by phase difference $\Delta\varphi$ scaled within ± 0.1 of the physical profile:

$$X(\Delta\varphi) \approx \max(0, \cos \Delta\varphi)$$

An opposed expert ($\Delta\varphi = \pi$) contributed zero interference to the demodulated channel. Furthermore, when the receiver's reference drifted coherently with the expert tags, phase-locking value remained robust (PLV 0.84), whereas a static reference collapsed (PLV 0.10), confirming the necessity of shared-reference tracking in physical or simulated oscillatory media.

However, phase coherence failed completely to deliver constructive ensemble gains. The preregistered ensemble gain test yielded a decisive null:

$$G_{\text{medium}} = 0.044 \quad \text{vs} \quad G_{\text{sum}} = 0.251$$

Coherent superposition in a shared channel does not synthesize new functional representations or amplify weak signals; it acts purely as a linear selective projection filter. Phase selects; it does not amplify.

3. The Attribution Payoff of Tagged Superposition (T13)

While phase superposition provides no ensemble advantage, tagging it makes expert contributions attributable. In looped Mixture-of-Loops architectures evaluated under strict compute parity ($C_{\text{match}} = 4,210,688$), phase tagging enabled individual expert contributions to be directly decoded from the unpartitioned aggregate state vector. Hypothesis H_{int} was supported across all 5 random seeds ($r_\varphi = 0.869$ vs 0.801 for random orthogonal tags, -0.695 for plain summation, and 0.412 for unmediated baseline).

By contrast, the accompanying context-slice ablation was null: partitioning input sequences into distinct context slices (`moe_of_loops_slices`) tied a shared context (`moe_of_loops_shared_ctx`), demonstrating that context slicing confers no intelligibility or performance benefit over a shared context under matched compute. Tags, phase or random, function as an attribution coordinate system, allowing individual expert contributions to be monitored

within a single shared superposition without multi-head parameter fan-out; phase tags add a $+0.068$ margin over random tags.

4. Where Lookup Beats Persistence: The Operational Cost of Always-On Coordination

Our model-level evaluations establish clear operational boundaries where persistent coordination is actively detrimental compared to stateless feedforward lookup:

1. **The Cost of Persistence on Continuous Evidence (T11.2)**: When input evidence is continuous, maintaining an always-on tonic state incurs a 0.42-segment switch lag and a 29% performance penalty (clean CSR 0.708 vs 1.000 for stateless lookup). The persistent state acts as an inertial anchor that resists necessary rapid transitions.
2. **Depth Budgeting and the Overthinking Deficit (T12)**: In looped depth budgeting, a stateless Adaptive Computation Time (ACT) controller matched the oracle depth budget, leaving nothing for a persistent coordinator to add (ΔACC $= 0$). Furthermore, extra computation depth caused severe performance degradation on easy tasks across 5 of 5 seeds (overthinking: $T = 12$ vs $T^* = 4$, ΔACC $= -0.3418$).
3. **Evidence Gaps and the Smoothing-Class Interpretation (T12b)**: Across evidence gaps containing deceptive distractors, the held tonic state beat one-step memory (0.6525 vs 0.2987), but failed to beat a simple 3-step majority vote filter (0.6390). Under the preregistered decision rule, this outcome is formally interpreted as `smoothing-class`: the tonic state acts as an exponential low-pass filter rather than a sophisticated cognitive coordinator.

5. The Attribution Cascade: Deconstructing Apparent Phase Advantages

A core contribution of our preregistered ablation methodology is identifying where apparent phase advantages dissolve under forensic controls:

- **Noise Robustness in T7**: Initial tests showed the phase gate outperforming learned routers under input noise. However, an exploratory addendum providing the learned router with a 32-token moving average (`LR-smoothed`) matched or exceeded the phase gate ($\Delta L = 0.0005$ vs 0.0026), attributing the advantage to low-pass filtering.
- **Clean Loss Advantage in T9**: An initial 0.47-nat clean loss gap favoring the phase medium over plain sum was eliminated once surrounding layers were retrained with a matched training surface (`matched_train`), attributing the gap to training discrepancies.
- **Observation Window Artifact in T11.3**: Apparent coordinator gains on ambiguous sequences were matched or beaten by a stateless argmax over the post-evidence window (1.000 vs 0.789), showing that windowed evidence integration explains the advantage rather than recurrent memory.
- **T10 Not Informatively Tested**: Experiment T10 (coherence as a plastic learning classifier) proved not informatively tested because the target signal was trivially leaked by tag placement, resulting in withdrawal of the classifier claim.

Limitations

1. **Synthetic Task Domain**: Evaluations were conducted on synthetic rule-following sequences, symbolic state machines, and looped modular blocks designed to isolate routing dynamics. Performance on massive-scale open-domain language modeling remains to be established.
2. **Fixed Coupling Topologies**: Coupling matrices were constructed with fixed mathematical spectral radii rather than learned end-to-end via gradient optimization, potentially under-utilizing task-specific adaptive dynamics.
3. **Expert Scaling Envelope**: Our experiments evaluated expert counts between $n = 3$ and $n = 8$. Scaling to dozens of experts is constrained by the physical slot capacity measured under phase jitter (7.65–8.35 usable slots at $\sigma_j = 0.05$).
4. **Historical Artifact Recovery**: Due to early disk corruption, historical hold-time figures (~41 tokens) from early reports were irrecoverable and replaced by verified from-scratch reruns under frozen commit `d975d96c` ($T_{\text{hold}} = 26.8$ tokens).

5. **Freeze provenance**: Preregistrations were fixed before execution by git commit blobs and Drive timestamps. Two freezes (T10, T11.1) went through with a required check undone, before the gated freeze script existed; a mechanical gate record exists from T11.2 onward (all four: T11.3 `T11_3_RULE19_GATES.md` + `receipts/t11_3/RULE19_FULL.json`; T12 `results/t5/T5_RULE19_GATES.json` + `FREEZE_META` rule19 PASS; T12b `T12B_RULE19_GATES.json/md`; T13 `T13_RULE19_GATES.json/md`). OpenTimestamps anchoring was a one-time backfill on 2026-09-24 (`anchored_retroactively`) and proves document existence as of that date, not at the original freeze.

What is Not Claimed

To prevent over-interpretation of our results, we explicitly state what this work does not claim:

1. **No Ensemble Accuracy Gain**: We do not claim that phase-superposed experts achieve constructive ensemble accuracy gains over plain sum or dense mixtures ($G_{\text{medium}} = 0.044$ vs $G_{\text{sum}} = 0.251$).
2. **No Supremacy Over Lookup on Continuous Evidence**: We do not claim that an always-on coordinator beats feedforward lookup when evidence is continuous; on continuous sequences, lookup wins decisively.
3. **No Claim of General Language Modeling Superiority**: We do not claim that oscillatory gating improves general next-token cross-entropy over standard learned feedforward routers.
4. **No Biological Intentionality**: We do not claim that elementary linear oscillatory units instantiate cognitive consciousness, intentionality, or biological brain states; they are evaluated strictly as mathematical routing mechanisms.
5. **Competitive Survival is a Proposed Hypothesis**: We do not claim that competitive survival across $n = 8$ vs $n = 3$ is an empirically proven physical law; it is a proposed mechanistic hypothesis to be tested in future controlled sweeps.

Disclosures & Errata

1. **T9 Preregistration Hash Citation Correction**: Prior drafts cited `d1c70486c40e637e22e1e6d034379d566fd365306b72bb30fd34dedb500c4fd1` as the T9 preregistration hash. That hash corresponds strictly to the primary results document (`phase-routing/t9/RESULTS.md`). The true frozen pre-commitment hash for `phase-routing/t9/PREREG.md` is `1c6a486726302a68ddb4ca818a89e56fb430f1a9e6ef2ba330364085b4e6289d` (confirmed by `t9/PREREG.sha256` and `t9/T9_CLOSED.md`).
2. **T11.2 and T11.3 Cryptographic Pre-commitments & Gate Records**: T11.2 (`593762c4c19e85d1720f6b46edbb93869258c4047ad023e42e144416aaed20d1`) is formally anchored to its Rule-19 mechanical gate record (`T11_2_RULE19_GATES.md`) and freeze log; T11.3 (`27ab4a9573ec8c01ed043d72ef6c83ba47fdd4a764a373acb7453381bb9bd442`) represents the final locked pre-commitment of the model-level orchestration series before the line transitioned to looped architectures.
3. **T12b Preregistered Vocabulary and Smoothing-Class Reconciliation**: The informal handoff label "support_lean" was used colloquially in project communications but does not appear in the frozen preregistration vocabulary. The frozen preregistration rule (§4.1.4, §4.2, §5) specifies that if the primary contrast clears the seed interval (`CSR_gap(hum) − CSR_gap(one_step_memory) > seed interval`, observed: 0.6525 vs 0.2987, $\Delta = +0.3538 >$ seed interval) but fails to beat majority smoothing (`majority_k3`: 0.6390, $\Delta = +0.0135$, failing to clear the seed interval), the outcome is formally interpreted as `smoothing-class`. Consequently, the verdict is recorded in the preregistration's exact terms: **Primary supported; per the preregistered interpretation rule, smoothing-class**.
4. **T8 vs T11 Hold Time Reconciliation**: The parameter $\rho_{\text{eff}} = 0.989$ and $T_{\text{hold}} = 65.0$ tokens belong strictly to the $n = 3$ rebuilt coupling assembly in T11, distinct from the $T_{\text{hold}} = 26.8$ tokens measured on the $n = 8$ rerun in T8. The ~41-token figure is an unrecoverable legacy artifact.
5. **T11.3 Arm Distinction (`hum` vs `hum_measure`) and Windowing**: In PREREG_T11_3.md (§2), `hum` is designated as the counted primary arm, while `hum_measure` was designated as a preregistered secondary arm for provider-parity disclosure. Evaluated on the counted `hum` arm, Ambiguous step CSR was 0.308 (losing to both classifier 0.351 and hand rule 0.348), cleanly failing H1 and triggering the stopping rule. The preregistered provider-parity arm `hum_measure` scored 0.789; beside this disclosure, the post-hoc exploratory stateless

argmax achieved 1.000 over that same window, demonstrating that the apparent lift reflects integration over the observation window rather than persistent internal state.

6. **Evidence Gaps in T12b**: The task sequences evaluated in T12b featured short gaps (primarily ≤ 3 steps), explaining why stateless $k = 3$ majority vote was sufficient to tie persistent gating.
7. **Depth Extrapolation in T12**: The observed performance degradation ('overthinking') at $T = 12$ in looped blocks may represent out-of-distribution depth extrapolation rather than an inherent architectural boundary, as training utilized fixed iteration depths.
8. **Arm-5 Scope**: Tagged superposition comparisons (Arm-5) were exploratory; phase-based routing is not shown to outperform random orthogonal keying on capacity.
9. **Provenance and Retroactive OpenTimestamps Anchoring**: Cryptographic pre-commitments across T5 through T13 were originally established via git commit blobs, Drive timestamps, and Rule-19 gate logs. A one-time OpenTimestamps backfill conducted on 2026-09-24 anchored all twelve Paper 2 preregistration files to the Bitcoin blockchain. In compliance with Rule-19 provenance standards, these anchors are cataloged as `anchored_retroactively` with the 2026-09-24 date and attest to document existence as of that stamping date, not the original historical freeze times. All subsequent pre-commitments are anchored directly at freeze time prior to executor outcome-data access.
10. **External Training Infrastructure (T9 and T10)**: Training runs for experiments T9 (phase medium) and T10 (coherence learning classifier) were executed on an external workstation environment and were not retained in the primary local repository. Consequently, while all evaluation protocols, frozen outcome logs, and analysis scripts are fully published, the raw training pipeline scripts for T9 and T10 are omitted from the initial public release (v1).
11. **T10 freeze with a known side-channel**: A side-channel through tag placement was known before the T10 freeze, and the freeze went through without the check that would have caught it. The three tag-based arms scored 1.000 with zero variance; the experiment is closed as not informatively tested (`not_informatively_tested` in `t10/CLOSING.md`), with the classifier claim withdrawn.
12. **T11.1 skipped check**: T11.1 froze with a required check skipped (Ambiguous near-chance validation: hand-off required first-32 classifier near chance on Ambiguous before hash; skipped — `t11/routing/results/primary/CLOSING.md`); the hand rule scored 1.0 on sequences labeled ambiguous, so H1 and H2 were not informatively tested.

Reproducibility

Preregistrations, `PREREG.sha256` files, RESULTS / PRIMARY aggregates, closing documents, artifact inventories, and Rule-19 freeze logs are released in the public repository (`https://github.com/k-luecke/tonic-gating`) under `experiments/phase-routing/` for each ticket. Affinity `affinity_inputs.py` is locked at sha `7dae877a3975c1ba4ca264f03f572d6749cce39a114ea3adbcffc010c7852c50` through T13 (verified across the T11→T12→T12b hand-off and T13 Rule-19 G6 PASS). Human-signed freezes emit the preregistration hash only after gate scripts pass. Hardware: Linux CPU for T11 series SCP cells; looped tickets disclose CPU/GPU as run. T8 is the forced from-frozen-commit rerun after artifact loss; OpenTimestamps for the twelve Paper 2 preregistration files were backfilled on 2026-09-24 as `anchored_retroactively` (Disclosure #9). External training pipelines for T9 matched-surface and T10 were not retained in the primary local repository (Disclosure #10). No Affinity / C_3 / hum retune is authorized by any closing in this series.

Frozen PREREG sha256 (full):

Ticket	PREREG.sha256
T11.2	`593762c4c19e85d1720f6b46edbb93869258c4047ad023e42e144416aaed20d1`
T11.3	`27ab4a9573ec8c01ed043d72ef6c83ba47fdd4a764a373acb7453381bb9bd442`
T12	`bf757ee4fc01e5c40b45f035150e7b716d12f7d8b30d745b9a7669f79bbf307c`
T12b	`851124b3c4bc66cb8e7c96c4f501084f4de186574f4f8d9446cac5879f62a3f0`
T13	`3176eb5ea3458047b820ff5511644d20a697a7a9485d5118560ff7b71e1f981d`