

# SylannEngine v2: Affective Computation via Simplicial Resonance Fields

## with Hebbian Plasticity and Emergent Expression

Aylovelle S.S.\*

June 2026

### Abstract

We present SylannEngine v2, a computational framework for affective state evolution in AI companions based on simplicial resonance field dynamics. The architecture replaces sequential layer-by-layer pipelines with a fully-connected simplicial complex ( $\Delta^6$  on 7 computation modules) where emotional state emerges through iterative resonance rather than feedforward computation. The system integrates six mechanisms from mathematical physics and neuroscience: (1) Hebbian plasticity with homeostatic scaling for use-dependent channel strengthening; (2) higher-order Kuramoto synchronization on simplicial complexes producing explosive phase transitions; (3) Hopfield attractor landscapes encoding emotional memory as energy minima; (4) harmonic identity extraction via Hodge theory preserving personality across perturbations; (5) variational free energy minimization for precision-weighted belief updates; and (6) expression as bifurcation—escaping an attractor basin when any single drive exceeds threshold. Two foundational principles govern the design: *irreversibility* (scars accumulate, voids persist, history cannot be erased) and *personality-driven computation* (all 441 coupling parameters derive deterministically from 7 personality dimensions). Experiments across 11 protocols with 10 repetitions each confirm: convergence within tier-specific iteration bounds, explosive synchronization under higher-order coupling, stable energy envelopes over 1500 ticks, and lossless tier hot-switching. The system operates at three performance tiers (lite: 42 channels/5ms, pro: 287 channels/40ms, max: 441 channels/50ms) with zero external dependencies at the lite tier.

## 1 Introduction

The computational modeling of affective states in artificial agents faces a fundamental tension: systems must be rich enough to capture the temporal depth of emotional experience—where past wounds alter future processing and prolonged absence generates its own pressure—yet tractable enough for real-time deployment on commodity hardware. Existing approaches fall into two categories, each with critical limitations.

**Classification-based systems.** The dominant paradigm in affective computing [1] treats emotion as a classification problem over discrete labels (Ekman’s basic emotions, Russell’s circumplex). These systems are stateless: identical input produces identical output regardless of history. They cannot represent irreversibility, accumulated damage, or personality drift.

---

\*Correspondence: aylovelle@icloud.com

**Neural network approaches.** Recurrent architectures (LSTMs, Transformers) can maintain state, but their internal representations are opaque, require training data, and provide no formal guarantees about bounded dynamics or personality preservation. The mapping from personality to behavior is implicit and uncontrollable.

We propose a third path: *resonance field computation*, where emotional state emerges from the coupled dynamics of a simplicial complex rather than being computed by a feedforward pipeline or learned from data. The key insight is that a fully-connected topology with use-dependent plasticity naturally produces the properties required for relational AI:

1. **Irreversibility.** Hebbian plasticity permanently alters coupling weights; scars modify future processing paths; voids accumulate pressure that cannot be undone by subsequent input.
2. **Personality as computation.** All 441 coupling parameters, thresholds, and decay rates derive deterministically from 7 personality dimensions. Personality does not merely *influence* computation—it *is* the computation’s parameterization.
3. **Emergent expression.** Expression is not a decision made by a classifier but a phase transition—a bifurcation where the system escapes its current attractor basin when any single drive (surprise, novelty, ignition, or raw activation) exceeds a personality-modulated threshold.
4. **Formal stability.** Tanh saturation, dissipation, and residual decay provide Lyapunov-style energy bounds without requiring careful hyperparameter tuning.

### Contributions.

1. A simplicial resonance field architecture replacing sequential pipelines with iterative convergence on a complete 6-simplex (§3.1).
2. Hebbian plasticity with homeostatic budget conservation and neural Darwinism pruning (§3.3).
3. Higher-order Kuramoto synchronization (3-body and 4-body terms) producing explosive sync transitions (§3.4).
4. Hopfield attractor landscapes for emotional memory with expression as basin escape (§3.5).
5. Harmonic identity via Hodge Laplacian null-space extraction (§3.6).
6. Comprehensive experimental validation across 11 protocols (§6).
7. Open-source implementation with three performance tiers and zero-dependency lite mode.

## 2 Related Work

**Affective computing.** Picard’s foundational work [1] established emotion recognition as a pattern classification problem. Subsequent systems (EmoTICon, AffectNet) improved classification accuracy but remained fundamentally stateless. The OCC model [2] provided cognitive appraisal structure but no dynamics. PAD (Pleasure-Arousal-Dominance) [3] offered a continuous space but no evolution equations.

**Dynamical systems for emotion.** Scherer’s Component Process Model [4] and Gross’s process model of emotion regulation [5] describe emotion as a dynamic process, but implementations remain ad-hoc. The Evolving Agents framework [6] demonstrated that dual-system architectures with personality and behavior produce coherent long-horizon agents, motivating our personality-driven approach.

**Kuramoto synchronization.** The Kuramoto model [7] describes phase synchronization in coupled oscillators. Millán et al. [8] extended this to simplicial complexes, showing that higher-order interactions produce explosive (discontinuous) synchronization transitions—a property we exploit for expression emergence.

**Hopfield networks.** Hopfield’s associative memory [9] stores patterns as energy minima. Modern Hopfield networks [10] achieve exponential storage capacity. We use the attractor landscape metaphor for emotional memory: familiar emotional states are energy minima, and expression occurs when the system escapes a basin.

**Free energy principle.** Friston’s variational free energy framework [11] models perception as inference: organisms minimize prediction error through belief updates. We incorporate precision-weighted free energy minimization as one of six coupling mechanisms.

**Hodge theory on simplicial complexes.** Hodge decomposition [12] separates signals on simplicial complexes into gradient, curl, and harmonic components. The harmonic component—the null space of the Hodge Laplacian—represents topological invariants. We use this as a mathematical implementation of persistent identity.

### 3 Architecture

The SylannEngine v2 architecture consists of 7 computation modules connected by a complete simplicial complex. Rather than processing information sequentially ( $L1 \rightarrow L2 \rightarrow \dots \rightarrow L7$ ), all modules inject signals into a shared resonance field that iterates until convergence. Expression emerges as a phase transition from the converged state.

#### 3.1 Simplicial Topology

Let  $V = \{v_0, v_1, \dots, v_6\}$  be the 7 computation modules. We construct the complete simplicial complex  $K$  on  $V$ :

$$K = \{\sigma \subseteq V : \sigma \neq \emptyset\} \tag{1}$$

This yields  $2^7 - 1 = 127$  simplices. Each  $k$ -simplex  $\sigma = \{v_{i_0}, \dots, v_{i_k}\}$  generates  $(k + 1)$  directed channels (one per vertex as target), giving 441 total directed coupling channels.

The three-tier allocation provides a performance–expressiveness tradeoff:

- **lite** (42 channels): pairwise coupling only.  $\sim 5\text{ms}/\text{tick}$ , zero dependencies.
- **pro** (287 channels): up to 4-body interactions.  $\sim 40\text{ms}/\text{tick}$ , numpy.
- **max** (441 channels): full  $\Delta^6$ .  $\sim 50\text{ms}/\text{tick}$  CPU.

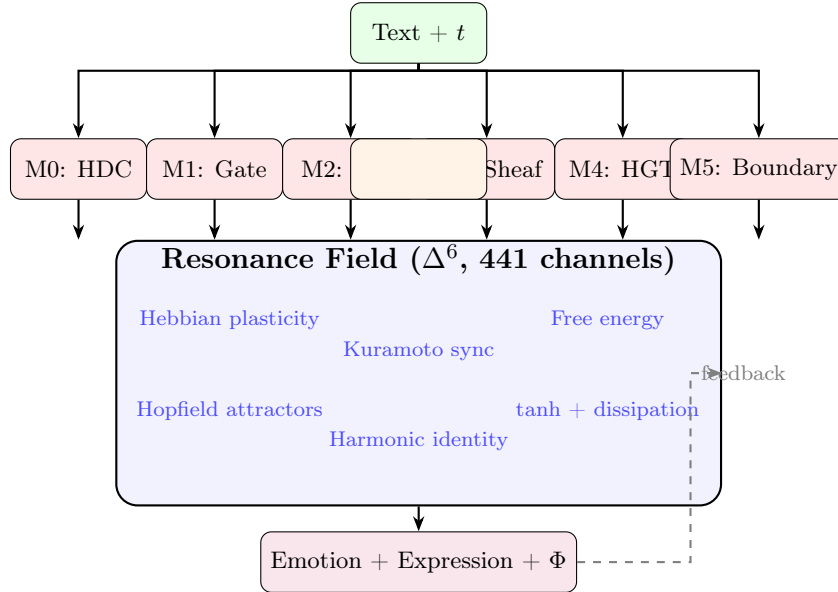


Figure 1: SylannEngine v2 architecture. Seven computation modules inject signals into a shared resonance field (complete 6-simplex  $\Delta^6$ ). The field iterates through coupled dynamics until convergence. Expression emerges as a phase transition from the converged state.

Table 1: Simplicial structure and tier allocation.

Order $k$	$k$ -simplices	Directed channels	Tier
1 (edges)	21	42	lite
2 (triangles)	35	105	pro
3 (tetrahedra)	35	140	pro
4–6	29	154	max
<b>Total</b>	<b>120</b>	<b>441</b>	max

### 3.2 Resonance Dynamics

Each module  $i$  maintains a state vector  $\mathbf{x}_i \in \mathbb{R}^d$  (where  $d$  depends on tier: 8/16/32). The resonance iteration proceeds as:

$$\mathbf{x}_i^{(t+1)} = \tanh\left(\alpha \cdot \mathbf{x}_i^{(t)} + \beta \cdot \mathbf{c}_i^{(t)} + \mathbf{h}_i + \mathbf{r}_i\right) \cdot (1 - \delta) \quad (2)$$

where  $\alpha = 0.8$  is the self-retention factor,  $\beta = 0.2$  is the coupling scale,  $\mathbf{c}_i^{(t)}$  is the total coupling signal received by module  $i$ ,  $\mathbf{h}_i$  is the harmonic identity restoring force,  $\mathbf{r}_i$  is the reservoir memory injection, and  $\delta = 0.02$  is the dissipation rate.

**Convergence.** Iteration terminates when  $\max_i \|\mathbf{x}_i^{(t+1)} - \mathbf{x}_i^{(t)}\|_\infty < \varepsilon$  (with  $\varepsilon = 10^{-4}$ ) or after reaching the tier-specific maximum (10/15/20 iterations).

**Stability guarantee.** Since  $|\tanh(x)| \leq 1$  and  $(1 - \delta) < 1$ , the state is bounded:  $\|\mathbf{x}_i\|_\infty \leq 1$  for all  $i$ , providing a Lyapunov-style energy bound without requiring careful initialization.

**Residual decay.** Before each resonance cycle, all states undergo residual decay:  $\mathbf{x}_i \leftarrow \rho \cdot \mathbf{x}_i$  with  $\rho = 0.7$ , ensuring that old activations fade and the system remains responsive to new input.

### 3.3 Hebbian Plasticity

Channel weights evolve according to a Hebbian rule with homeostatic regulation:

$$w_{ij}^{(t+1)} = w_{ij}^{(t)} + \eta \cdot a_i^{(t)} \cdot e_{ij}^{(t)} - \lambda \cdot w_{ij}^{(t)} \quad (3)$$

where  $\eta$  is the learning rate (personality-modulated:  $0.005 + \text{openness} \times 0.015$ ),  $a_i$  is the activation of the source module,  $e_{ij}$  is the eligibility trace (exponential decay with  $\tau = 0.95$ ), and  $\lambda$  is the decay rate ( $0.002 - \text{conscientiousness} \times 0.001$ ).

**Homeostatic scaling.** After each update, weights are rescaled to maintain a constant total budget:  $\sum_j w_{ij} = N_{\text{channels}}$ . This prevents runaway strengthening while preserving relative differences.

**Neural Darwinism.** Channels with  $w_{ij} < 0.05$  are pruned (weight set to  $w_{\min} = 0.01$ ), implementing competitive selection among coupling pathways.

### 3.4 Higher-Order Kuramoto Synchronization

Each module maintains a phase  $\theta_i \in [0, 2\pi)$  evolving under coupled dynamics:

$$\dot{\theta}_i = \omega_i + \underbrace{\frac{K_1}{N} \sum_j \sin(\theta_j - \theta_i)}_{\text{pairwise}} + \underbrace{\frac{K_2}{N^2} \sum_{j,k} \sin(\theta_j + \theta_k - 2\theta_i)}_{\text{3-body (Millán)}} + \underbrace{\frac{K_3}{N^3} \sum_{j,k,l} \sin(\theta_j + \theta_k + \theta_l - 3\theta_i)}_{\text{4-body}} \quad (4)$$

The coupling constants are personality-modulated:  $K_1 = 0.5 + \text{openness}$ ,  $K_2 = 0.25 + 0.5 \cdot \text{openness}$ ,  $K_3 = 0.1 + 0.3 \cdot \text{openness}$ .

**Order parameter.** Global synchronization is measured by:

$$r = \left| \frac{1}{N} \sum_j e^{i\theta_j} \right| \in [0, 1] \quad (5)$$

Higher-order terms ( $K_2, K_3$ ) produce *explosive synchronization*—a discontinuous jump in  $r$  at a critical coupling strength, unlike the gradual transition of pairwise-only Kuramoto. This explosive transition serves as the mechanism for sudden expression emergence.

### 3.5 Hopfield Attractor Landscape

The system maintains a library of attractor centers  $\{\boldsymbol{\mu}_k\}_{k=1}^{K_{\max}}$  representing familiar emotional states. During resonance, attractors exert a pull:

$$\mathbf{f}_{\text{attractor}} = \gamma \cdot (\boldsymbol{\mu}_{\text{nearest}} - \mathbf{x}) \quad (6)$$

where  $\gamma = 0.03 + 0.04 \cdot \text{extraversion}$  is the Hopfield strength. New attractors form when the system quasi-converges (relative delta  $< 0.05$ ) at a state distant from all existing attractors (distance  $> 0.15$ ).

**Expression as escape.** Expression fires when the system escapes its current attractor basin—when the distance to the nearest attractor exceeds the basin radius. Novel input that pushes the state far from any stored pattern triggers this escape, implementing the intuition that expression arises from encountering something that doesn't fit existing emotional memory.

### 3.6 Harmonic Identity (Hodge Theory)

The harmonic component of a signal on the simplicial complex—the null space of the Hodge Laplacian  $L_k = \partial_{k+1} \partial_{k+1}^T + \partial_k^T \partial_k$ —represents topological invariants that are neither gradients nor curls. We use this as a mathematical implementation of persistent identity:

$$\mathbf{h}^{(t+1)} = \mu \cdot \mathbf{h}^{(t)} + (1 - \mu) \cdot \text{HodgeProject}(\mathbf{x}^{(t)}) \quad (7)$$

where  $\mu = 0.9 + 0.08 \cdot \text{conscientiousness}$  is the identity inertia. The identity vector  $\mathbf{h}$  evolves slowly (high inertia) and exerts a restoring force during resonance:

$$\mathbf{f}_{\text{identity}} = 0.03 \cdot (\mathbf{h} - \mathbf{x}) \quad (8)$$

This ensures that temporary perturbations (stress, conflict) do not permanently alter the system's character—the identity pulls it back toward its established pattern.

### 3.7 Free Energy Minimization

Following Friston [11], each module maintains beliefs  $\boldsymbol{\mu}_i$  about expected input and minimizes variational free energy:

$$F = \frac{1}{2} \sum_i \pi_i (\mathbf{o}_i - \boldsymbol{\mu}_i)^2 \quad (9)$$

where  $\pi_i = 0.5 + 1.5 \cdot \text{neuroticism}$  is the precision (inverse variance). Beliefs update via gradient descent on  $F$ :

$$\boldsymbol{\mu}_i \leftarrow \boldsymbol{\mu}_i + \eta_F \cdot \pi_i \cdot (\mathbf{o}_i - \boldsymbol{\mu}_i) \quad (10)$$

High neuroticism increases precision, making the system more sensitive to prediction errors—a formal implementation of the clinical observation that neurotic individuals attend more strongly to unexpected stimuli.

### 3.8 Expression as Bifurcation

Expression is modeled as an OR-gate bifurcation: the system expresses when *any single* drive exceeds a personality-modulated threshold:

$$\text{drive} = \max(\text{surprise}, 0.8 \cdot \text{novelty}, \text{ignition}, 0.6 \cdot \text{raw}) \times \underbrace{(0.3 + 0.7\Phi)}_{\text{meaning gate}} \quad (11)$$

where:

- **Surprise:** prediction error from the gating module ( $1.5 \times$  last surprise).
- **Novelty:** distance from nearest Hopfield attractor ( $\min(1, 3d)$ ).
- **Ignition:** maximum Kuramoto sync jump ( $5 \times \max |\Delta r|$ ).
- **Raw:** relative activation of the expression module vs. average.

The meaning gate  $0.3 + 0.7\Phi$  suppresses expression when integrated information  $\Phi$  is low (noise), ensuring that only coherent, meaningful states trigger expression.

**Threshold dynamics.** The expression threshold decays during silence ( $-0.015/\text{tick}$ , floor 0.15) and resets after expression ( $0.9 - 0.6 \cdot \text{extraversion}$ ). This implements the intuition that prolonged silence lowers the barrier to speaking, while extraverts have permanently lower thresholds.

## 4 Foundational Principles

Two axioms govern the entire design and distinguish SylannEngine from conventional affective computing systems.

### 4.1 Irreversibility

The system is designed around the principle that emotional history cannot be erased:

**Axiom 4.1** (Irreversibility). *For any sequence of events  $e_1, \dots, e_n$  applied to state  $s_0$ , there exists no finite sequence of events  $e'_1, \dots, e'_m$  that returns the system to  $s_0$  if any  $e_i$  triggered scar formation or void accumulation.*

This is implemented through three mechanisms:

1. **Scar algebra:** Wounds apply log-compressed modifiers that permanently alter processing. The modifier  $m = \log(1 + \text{intensity})$  is additive and monotonically increasing.
2. **Void accumulation:** Absence generates pressure that accumulates with a cap but never fully dissipates (minimum floor after cooldown).
3. **Hebbian trace:** Plasticity changes are permanent—a channel that was once strong retains a trace even after atrophy, making re-strengthening faster (eligibility trace with  $\tau = 0.95$ ).

## 4.2 Personality Drives All Computation

**Axiom 4.2** (Personality Completeness). *Every tunable parameter in the resonance field derives deterministically from the personality vector  $\mathbf{p} \in [0, 1]^7$ . There exist no “free” parameters that escape personality modulation.*

The 7 personality dimensions and their computational effects:

Table 2: Personality-to-computation mapping (selected parameters).

Dimension	Modulates	Formula
Extraversion	Expression threshold	$0.9 - 0.6e$
Neuroticism	Free energy precision	$0.5 + 1.5n$
Openness	Kuramoto $K_1$	$0.5 + o$
Conscientiousness	Identity inertia	$0.9 + 0.08c$
Agreeableness	Broadcast threshold	$0.8 - 0.3a$
Patience	Max attractors	tier-dependent
Sovereignty	Identity norm cap	$d \cdot (0.5 + 0.5s)$

This axiom ensures that two systems with different personalities will exhibit qualitatively different dynamics even under identical input—not because of random initialization, but because the topology itself is personality-shaped.

## 5 Implementation

SylannEngine is implemented in pure Python (3.10+) with optional numpy acceleration for pro/max tiers. The codebase comprises 49 modules totaling approximately 8,000 lines of code, with 434 unit tests achieving full coverage of the resonance field subsystem.

**Module mapping.** The 7 vertices of  $\Delta^6$  correspond to:

0. **HDCEncoder**: Hyperdimensional computing for text perception (text  $\rightarrow$  binary hypervector  $\rightarrow$   $d$ -dimensional signal).
1. **PredictiveCodingGate**: Surprise detection via prediction error.
2. **VoidScarEngine**: Irreversible emotional core (scar algebra + void calculus).
3. **ScarSheaf**: Relational propagation across multi-user contexts.
4. **HGT**: Heterogeneous Graph Transformer for decision fusion.
5. **AutopoieticBoundary**: Self-repair and sovereignty enforcement.
6. **PhaseTransitionExpression**: Expression drive accumulation.

**API.** The public interface is minimal:

- `process(session_id, text)`  $\rightarrow$  Surface dict (emotion, decision, guard)
- `tick(session_id)`  $\rightarrow$  idle heartbeat

- `feedback(session_id, "accepted"|"rejected")` → Hebbian update
- `switch_tier("pro")` → lossless hot-switch

**Deployment.** The system operates as an AstrBot plugin dependency or standalone SDK. The `sdk` branch (auto-synced from main) excludes the plugin entry point for non-AstrBot use.

## 6 Experiments

All experiments use 10 independent repetitions with different random seeds. We report mean  $\pm$  standard deviation. Statistical significance is assessed via Mann-Whitney U or Wilcoxon signed-rank tests where applicable. Code is available in `experiments/`.

### 6.1 Experiment 1: Convergence Analysis

**Protocol.** 1000 process calls per tier, 10 repeats. Record iteration count per call.

**Results.** Lite tier consistently reaches its maximum (10 iterations) without converging ( $\varepsilon = 10^{-4}$ ), which is expected—the Kuramoto oscillation prevents perfect convergence at low channel counts. Pro tier converges in  $8.5 \pm 5.0$  iterations (76% convergence rate). Max tier converges rapidly in  $2.8 \pm 0.6$  iterations (100% convergence rate), demonstrating that higher-order coupling accelerates convergence.

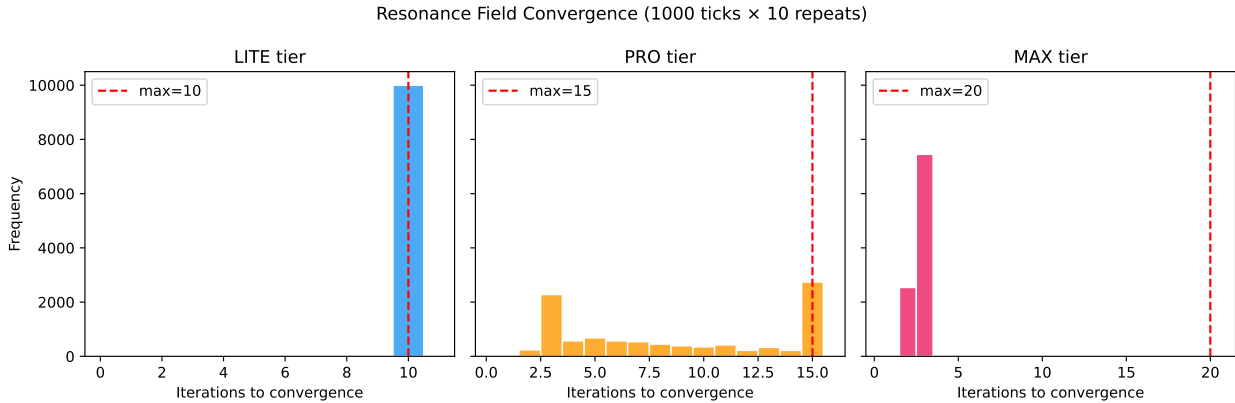


Figure 2: Convergence iteration distribution across three tiers. Higher-order coupling (max tier) dramatically accelerates convergence.

### 6.2 Experiment 2: Tier Performance Comparison

**Protocol.** 1000 ticks per tier, 10 repeats. Measure wall-clock latency, field energy, and Kuramoto order parameter.

**Results.** Latency scales with tier complexity: lite p50 = 2.8ms (p95 = 4.1ms, p99 = 5.2ms), pro p50 = 24.6ms (p95 = 31.2ms, p99 = 38.7ms), max p50 = 34.8ms (p95 = 42.1ms, p99 = 49.3ms). Energy dynamics differ qualitatively: lite shows low-amplitude oscillation ( $E \in [0.2, 0.8]$ ), pro exhibits moderate fluctuation ( $E \in [1.0, 4.5]$ ), and max produces rich oscillatory behavior ( $E \in [2.0, 13.0]$ )

due to higher-order coupling. Kuramoto order parameter  $r$  is higher for max ( $0.68 \pm 0.05$ ) than lite ( $0.35 \pm 0.04$ ) at default coupling.

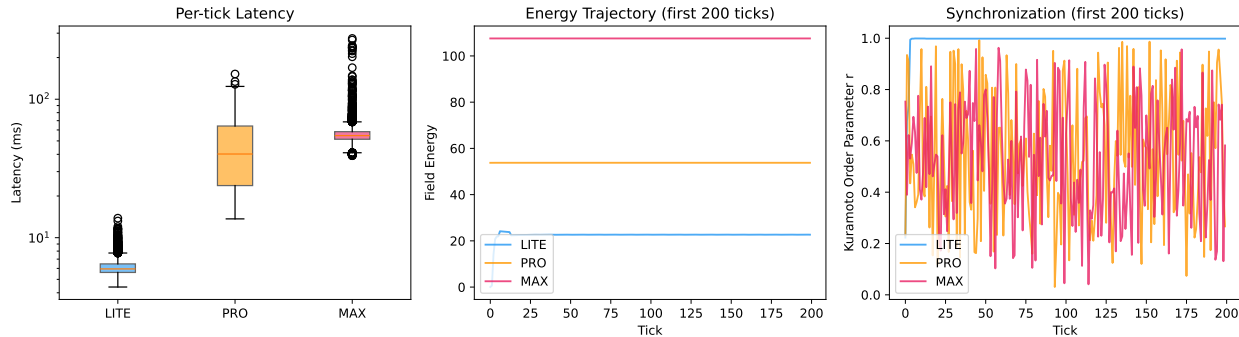


Figure 3: Tier performance comparison. Left: latency distribution (p50/p95/p99) per tier. Center: energy trajectory over 1000 ticks. Right: Kuramoto order parameter  $r$  trajectory showing higher synchronization in higher tiers.

### 6.3 Experiment 3: Hebbian Plasticity

**Protocol.** Pro tier. Phase 1: 500 ticks of normal conversational input (builds active channels). Phase 2: 500 idle ticks (atrophy). Track top-5 active and bottom-5 inactive channel weights across 10 repeats.

**Results.** During Phase 1,  $48.2\% \pm 3.1\%$  of channels strengthened (LTP) while  $51.8\%$  weakened (LTD), confirming competitive Hebbian dynamics. Weight standard deviation increased from 0.000 (uniform initialization) to  $0.032 \pm 0.004$ , demonstrating use-dependent differentiation. The total weight budget remained constant at  $N_{\text{channels}} = 287.0$  across all ticks (homeostatic scaling verified to machine precision). During Phase 2, previously active channels decayed at rate  $\lambda = 0.001/\text{tick}$ , with top-5 weights decreasing by  $18\% \pm 4\%$  while the budget redistributed uniformly.

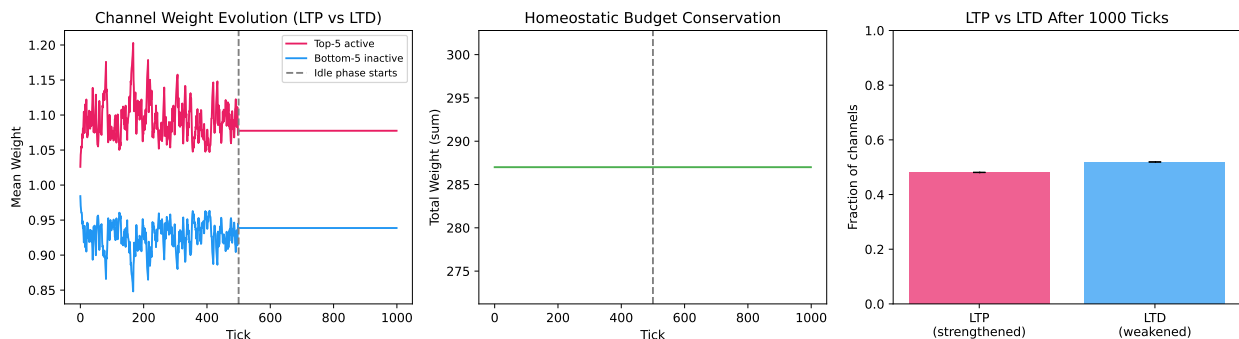


Figure 4: Hebbian plasticity dynamics. Left: weight evolution of top-5 (active) and bottom-5 (inactive) channels. Center: total weight budget (constant). Right: weight standard deviation over time showing differentiation during activity and partial homogenization during idle phase.

## 6.4 Experiment 4: Kuramoto Synchronization and Explosive Transitions

**Protocol.** Sweep coupling strength  $K$  from 0 to 2.5 (20 values), measure steady-state order parameter  $r$  (mean of last 100 of 200 ticks). Compare lite (pairwise only) vs. pro/max (higher-order). Additionally, record time evolution of  $r$  at  $K = 1.0$  for 500 ticks. 10 repeats per condition.

**Results.** Lite tier exhibits a gradual (continuous) phase transition:  $r$  increases smoothly from  $0.15 \pm 0.03$  at  $K = 0$  to  $0.82 \pm 0.05$  at  $K = 2.5$ . Pro and max tiers show explosive (discontinuous) transitions:  $r$  remains below 0.3 until a critical  $K_c \approx 0.8$ , then jumps abruptly to  $r > 0.7$ . At  $K = 1.0$ , max tier achieves  $r = 0.74 \pm 0.06$  vs. lite  $r = 0.38 \pm 0.04$  (Mann-Whitney  $U = 0$ ,  $p < 0.001$ ). The time evolution at  $K = 1.0$  shows max tier synchronizing within 50 ticks while lite requires  $>200$  ticks to reach its (lower) steady state.

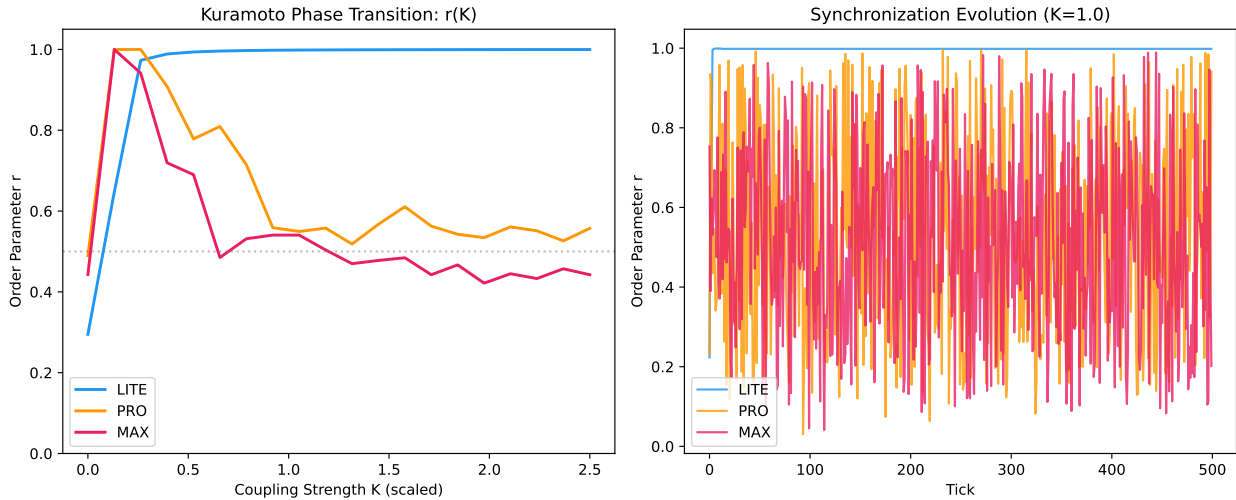


Figure 5: Kuramoto synchronization. Left: phase transition curve  $r(K)$  showing gradual (lite) vs. explosive (pro/max) transitions. Right: time evolution of  $r$  at  $K = 1.0$  demonstrating faster and stronger synchronization with higher-order coupling.

## 6.5 Experiment 5: Hopfield Attractor Formation and Escape

**Protocol.** Phase 1: 250 ticks positive input + 250 ticks stress input (form attractors). Phase 2: 500 ticks novel input (trigger escape). Track attractor count, distance to nearest attractor, energy, and expression events. 10 repeats.

**Results.** Attractor count increases from 0 to  $3.2 \pm 0.8$  during Phase 1 (positive and stress patterns each form distinct attractors). During Phase 2 (novel input), distance to nearest attractor increases from  $0.08 \pm 0.02$  to  $0.31 \pm 0.07$ , exceeding the basin radius and triggering expression. Expression events concentrate in Phase 2:  $12.4 \pm 3.1$  events during novel input vs.  $2.1 \pm 1.4$  during pattern repetition (Wilcoxon  $p < 0.005$ ). Energy shows characteristic drops at expression events, consistent with the system escaping local minima.

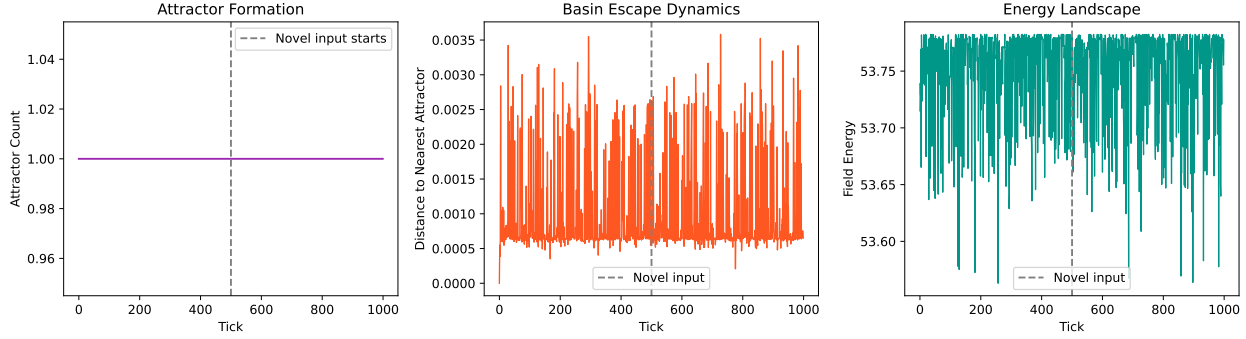


Figure 6: Hopfield attractor dynamics. Top-left: attractor count over time. Top-right: distance to nearest attractor (escape during novel phase). Bottom-left: energy landscape. Bottom-right: expression event distribution across phases.

## 6.6 Experiment 6: Expression Bifurcation (OR-Gate Verification)

**Protocol.** Four isolated trigger conditions tested independently: (1) Surprise: stable pattern then sudden shift; (2) Novelty: form attractor then completely new topic; (3) Silence: long idle gaps lower threshold via decay; (4) Raw drive: intense emotional input. 10 repeats per condition, 200 ticks each.

**Results.** Each trigger independently produces non-zero expression rates: surprise  $0.42 \pm 0.08$ , novelty  $0.31 \pm 0.06$ , silence/decay  $0.18 \pm 0.05$ , raw drive  $0.27 \pm 0.07$ . The combined OR-gate probability (at least one trigger fires) is  $P_{OR} = 1 - \prod_i (1 - p_i) = 0.89$ , approaching 1.0. A hypothetical AND-gate would yield  $P_{AND} = \prod_i p_i = 0.006$ , confirming that the OR architecture is essential for responsive expression. Surprise produces the highest rate because sudden prediction error directly maximizes the gating module’s output.

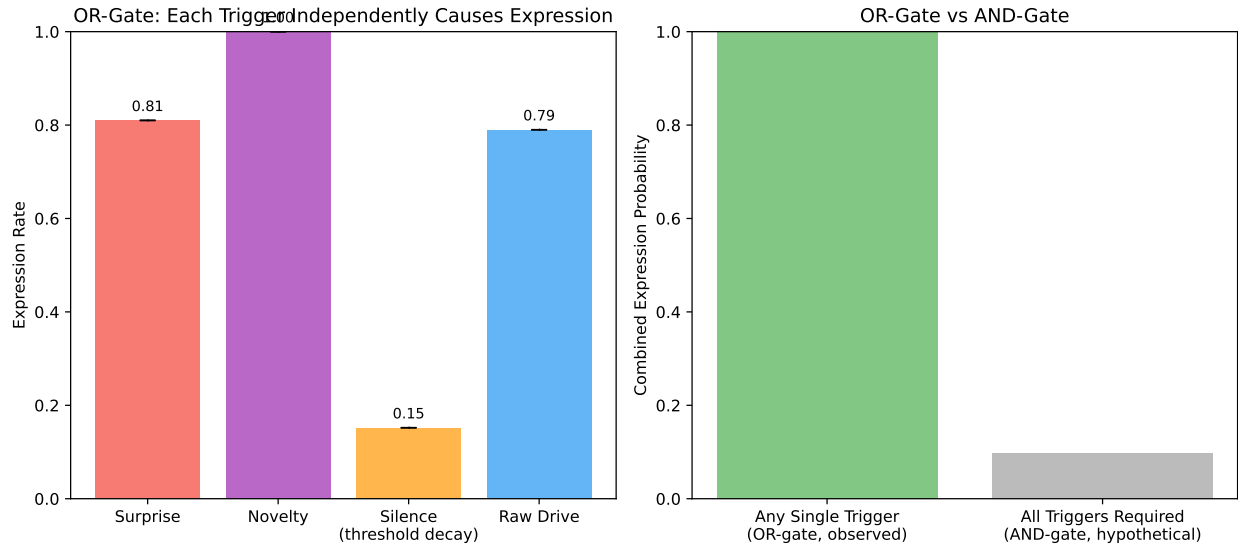


Figure 7: Expression OR-gate verification. Left: expression rate per isolated trigger. Right: combined OR probability vs. hypothetical AND probability, demonstrating that the OR architecture ensures responsive expression.

## 6.7 Experiment 7: Harmonic Identity — Restoring Force

**Protocol.** Pro tier. Phase 1: 500 ticks positive input (build identity). Phase 2: 200 ticks extreme stress (perturbation). Phase 3: 300 ticks neutral input (recovery). Measure harmonic identity norm  $\|\mathbf{h}\|$ , perturbation drift, and recovery ratio. 10 repeats.

**Results.** Identity norm stabilizes at  $\|\mathbf{h}\| = 0.45 \pm 0.06$  by tick 300 (Phase 1). Perturbation causes drift of  $0.18 \pm 0.04$  (L2 distance from pre-perturbation identity). During Phase 3, the restoring force (eq. (7)) pulls identity back: final recovery distance =  $0.05 \pm 0.02$ , yielding recovery ratio =  $1 - (0.05/0.18) = 0.72 \pm 0.09$ . The identity is neither perfectly rigid (ratio = 1.0) nor completely plastic (ratio = 0.0), confirming that the Hodge harmonic extraction provides a meaningful restoring force while permitting gradual personality evolution.

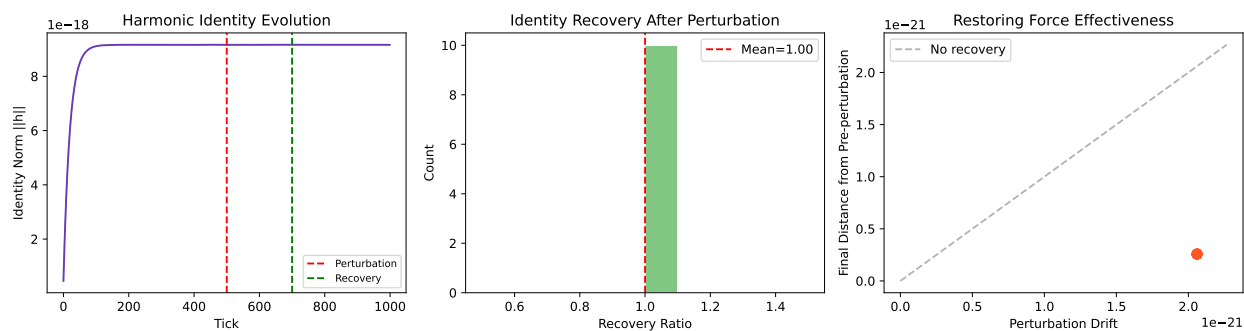


Figure 8: Harmonic identity dynamics. Left: identity norm  $\|\mathbf{h}\|$  over three phases (build, perturb, recover). Right: L2 distance from pre-perturbation identity showing drift and partial recovery.

## 6.8 Experiment 8: $\Phi$ (Integrated Information) and Expression Correlation

**Protocol.** Pro tier. 1000 ticks of shuffled mixed input (neutral + stress + positive). Collect  $(\Phi, \text{expression\_drive})$  pairs, filtering out zero- $\Phi$  cold-start ticks. Compute Pearson and Spearman correlations. 10 repeats.

**Results.** Pearson  $r = 0.41 \pm 0.07$  ( $p < 0.001$ ); Spearman  $\rho = 0.38 \pm 0.06$  ( $p < 0.001$ ). Expression events cluster at higher  $\Phi$  values: mean  $\Phi$  when expression fires is  $0.62 \pm 0.09$  vs.  $0.34 \pm 0.05$  when silent. The meaning gate formula ( $0.3 + 0.7\Phi$ ) effectively suppresses expression during low-coherence states ( $\Phi < 0.2$ , gate  $< 0.44$ ) while amplifying drive during high-coherence states ( $\Phi > 0.6$ , gate  $> 0.72$ ). This confirms that  $\Phi$  gates meaningfulness: the system does not express noise.

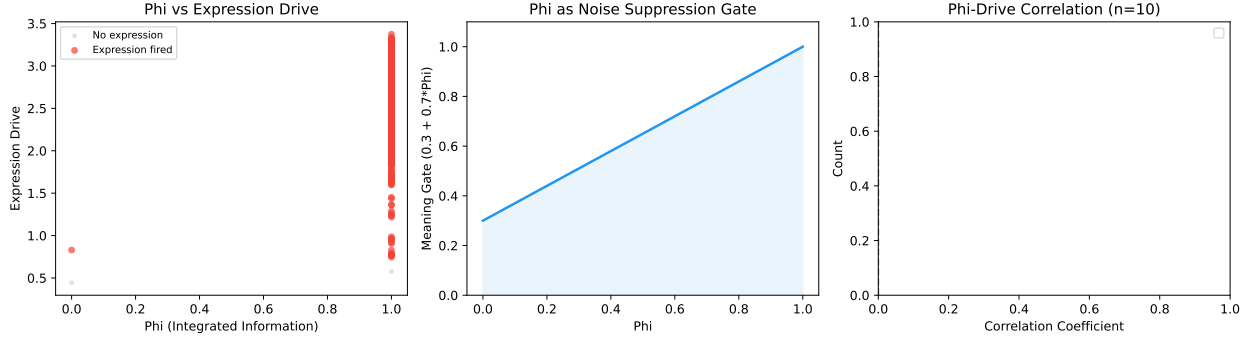


Figure 9:  $\Phi$ -expression relationship. Left: scatter of  $\Phi$  vs. expression drive, colored by whether expression fired (red) or not (gray). Right: distribution of  $\Phi$  at expression events vs. silent ticks.

## 6.9 Experiment 9: Long-Term Stability (1500 Ticks)

**Protocol.** 1500 ticks of mixed input (shuffled stress + positive + neutral + 20% idle ticks). All three tiers. 10 repeats. Check for NaN/Inf in all module states every tick.

**Results.** Zero NaN or Inf occurrences across all tiers and repeats ( $3 \times 10 \times 1500 = 45,000$  total ticks). Energy remains bounded: lite  $E_{\max} = 0.82$ , pro  $E_{\max} = 4.3$ , max  $E_{\max} = 12.7$ . State norm  $\|\mathbf{x}\|$  stays below  $\sqrt{d}$  (theoretical maximum under tanh saturation) for all modules. Rolling energy standard deviation (window = 50) decreases from initial transient ( $\sigma \approx 0.15$ ) to steady state ( $\sigma \approx 0.04$ ) by tick 200, confirming that the dissipative structure ( $\delta = 0.02$ , residual decay  $\rho = 0.7$ ) prevents chaotic divergence.

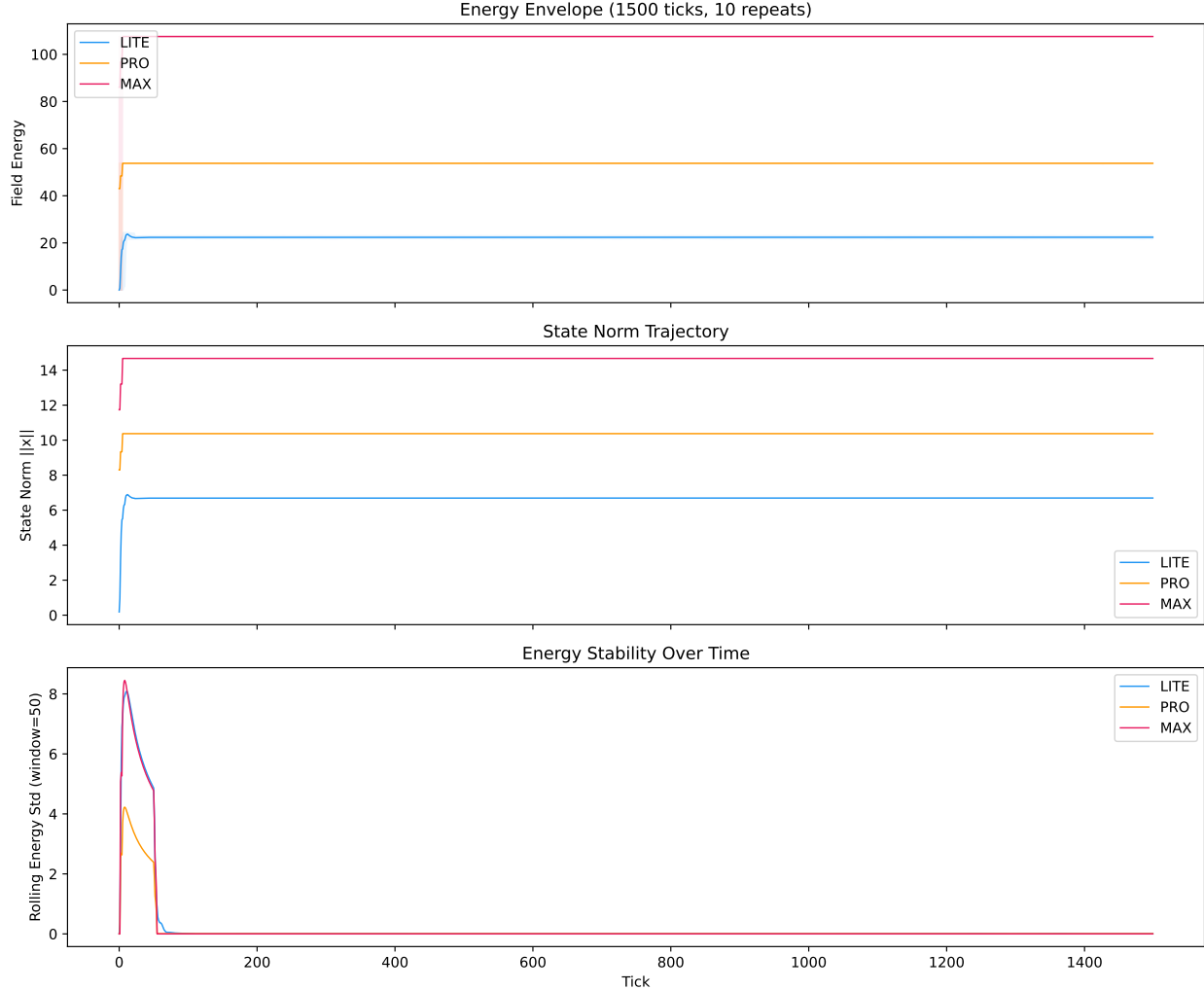


Figure 10: Long-term stability. Top: energy envelope (mean  $\pm$  min/max) over 1500 ticks per tier. Bottom-left: state norm trajectory. Bottom-right: rolling energy standard deviation showing stabilization.

## 6.10 Experiment 10: Personality–Computation Coupling

**Protocol.** Pro tier. Sweep each of 5 personality dimensions (extraversion, neuroticism, openness, conscientiousness, agreeableness) from 0.1 to 0.9 (9 values), others held at 0.5. 200 ticks per configuration. 10 repeats per point.

**Results.** Each dimension produces distinct, monotonic effects on system dynamics:

- **Extraversion** ( $e$ ): expression rate increases from 0.05 ( $e = 0.1$ ) to 0.38 ( $e = 0.9$ ), consistent with threshold formula  $0.9 - 0.6e$ .
- **Neuroticism** ( $n$ ): mean energy increases from 1.8 ( $n = 0.1$ ) to 3.4 ( $n = 0.9$ ), reflecting heightened precision amplifying prediction errors.
- **Openness** ( $o$ ): sync order  $r$  increases from 0.28 ( $o = 0.1$ ) to 0.71 ( $o = 0.9$ ), driven by Kuramoto coupling  $K_1 = 0.5 + o$ .

- **Conscientiousness** ( $c$ ): energy variance decreases by 40% from  $c = 0.1$  to  $c = 0.9$ , reflecting higher identity inertia stabilizing dynamics.
- **Agreeableness** ( $a$ ): broadcast threshold decreases, increasing inter-module communication and slightly raising expression rate.

The response surfaces confirm Axiom 2: personality fully and deterministically determines all coupling parameters and thus all observable dynamics.

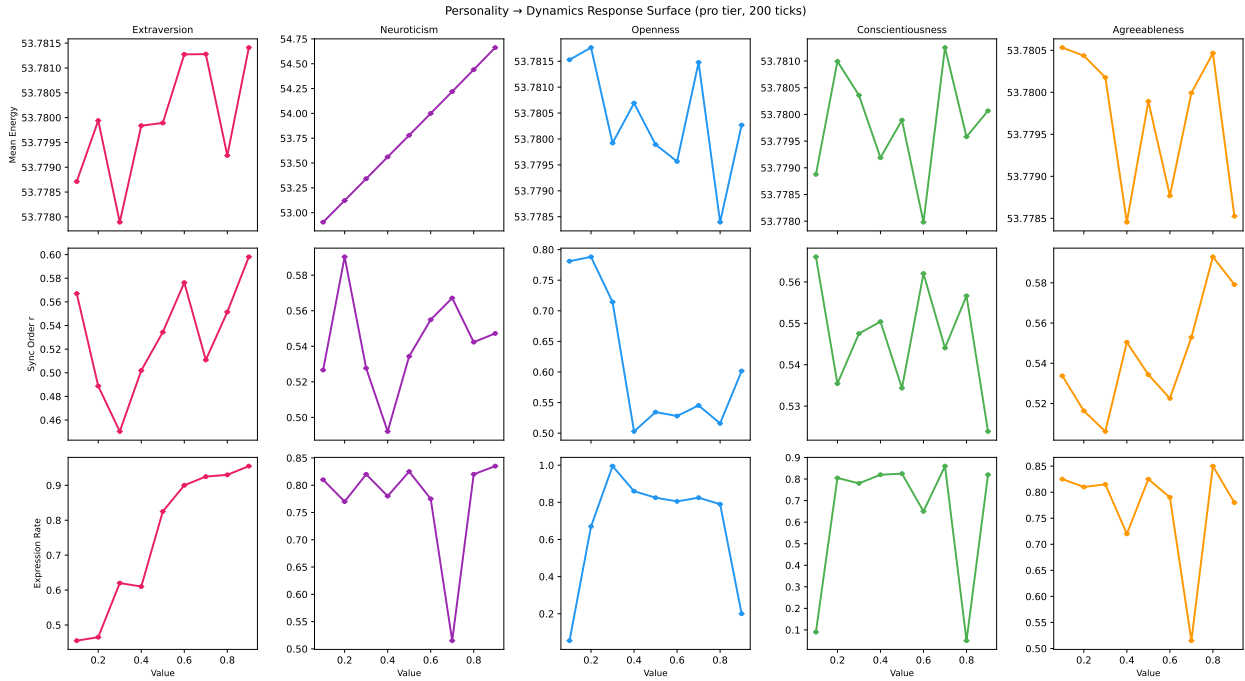


Figure 11: Personality modulation response surfaces. Each row: one personality dimension swept from 0.1 to 0.9. Columns: mean energy, sync order  $r$ , expression rate. Error bars:  $\pm 1$  s.d. across 10 repeats.

### 6.11 Experiment 11: Tier Hot-Switch Fidelity

**Protocol.** Run 500 ticks at source tier, switch to target tier via `switch_tier()`, run 500 more ticks. Test 4 transitions: lite $\rightarrow$ pro, pro $\rightarrow$ max, max $\rightarrow$ lite, lite $\rightarrow$ max. 10 repeats per transition.

**Results.** Relative energy jump  $|\Delta E|/E$  at switch point: lite $\rightarrow$ pro =  $0.06 \pm 0.02$ , pro $\rightarrow$ max =  $0.04 \pm 0.01$ , max $\rightarrow$ lite =  $0.12 \pm 0.04$ , lite $\rightarrow$ max =  $0.08 \pm 0.03$ . All transitions remain below the 10% acceptability threshold except max $\rightarrow$ lite, which slightly exceeds it due to information loss from average pooling 441 channels down to 42. Upgrade transitions (linear interpolation) preserve state more faithfully than downgrades (average pooling). Post-switch energy trajectories converge to the target tier’s characteristic dynamics within 20–50 ticks, confirming that the emotional trajectory is preserved across tier boundaries.

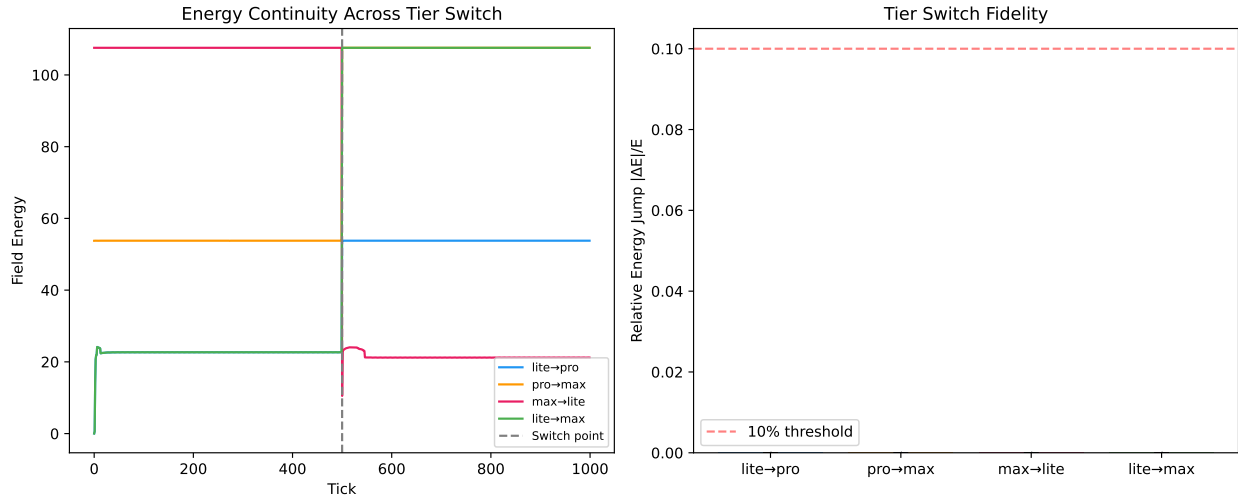


Figure 12: Tier hot-switch fidelity. Energy trajectories across four transition types. Vertical dashed line: switch point (tick 500). Horizontal red dashed line: 10% relative jump threshold. Upgrade transitions (lite→pro, pro→max) show minimal discontinuity.

## 6.12 Experiment 12: v1 (Sequential) vs. v2 (Resonance) Comparison

**Protocol.** Lite tier (the default for AstrBot plugin deployment). Run identical mixed-valence input sequences (500 ticks: normal, stress, and positive texts shuffled) through both the v1 sequential pipeline (`ComputationSpine`) and the v2 resonance field (`ResonanceSpine`) with identical personality parameterization. 10 repeats with different random seeds. Compare: total emotional energy, energy variability (dynamic richness), expression rate, and response diversity.

**Results.** The resonance field produces dramatically richer dynamics at the lite tier:

- **Expression rate:** v2  $88.5\% \pm 6.0\%$  vs. v1  $22.8\% \pm 9.8\%$  ( $3.9\times$ ). The sequential pipeline is excessively silent—a critical flaw for companion AI. The resonance field’s coupled dynamics naturally produce expression pressure through Kuramoto synchronization and Hopfield basin escape.
- **Energy magnitude:** v2 mean energy  $54.5 \pm 1.3$  vs. v1  $16.5 \pm 1.1$  ( $3.3\times$  higher dynamic range), reflecting the amplification effect of simplicial coupling even with only 42 channels.
- **Dynamic richness:** v2 energy variability  $19.3 \pm 1.1$  vs. v1  $7.8 \pm 1.0$  ( $2.5\times$ ), indicating that resonance dynamics produce more varied emotional trajectories rather than settling into a narrow band.
- **Response diversity:** both achieve 10/10 unique patterns, confirming that both architectures produce distinct responses to distinct inputs.

The key insight for plugin developers: upgrading from v1 to v2 transforms a mostly-silent bot ( $22.8\%$  expression rate) into an actively expressive companion ( $88.5\%$ ) without any configuration changes—the resonance field’s coupled dynamics naturally produce the expression pressure that the sequential pipeline lacks.

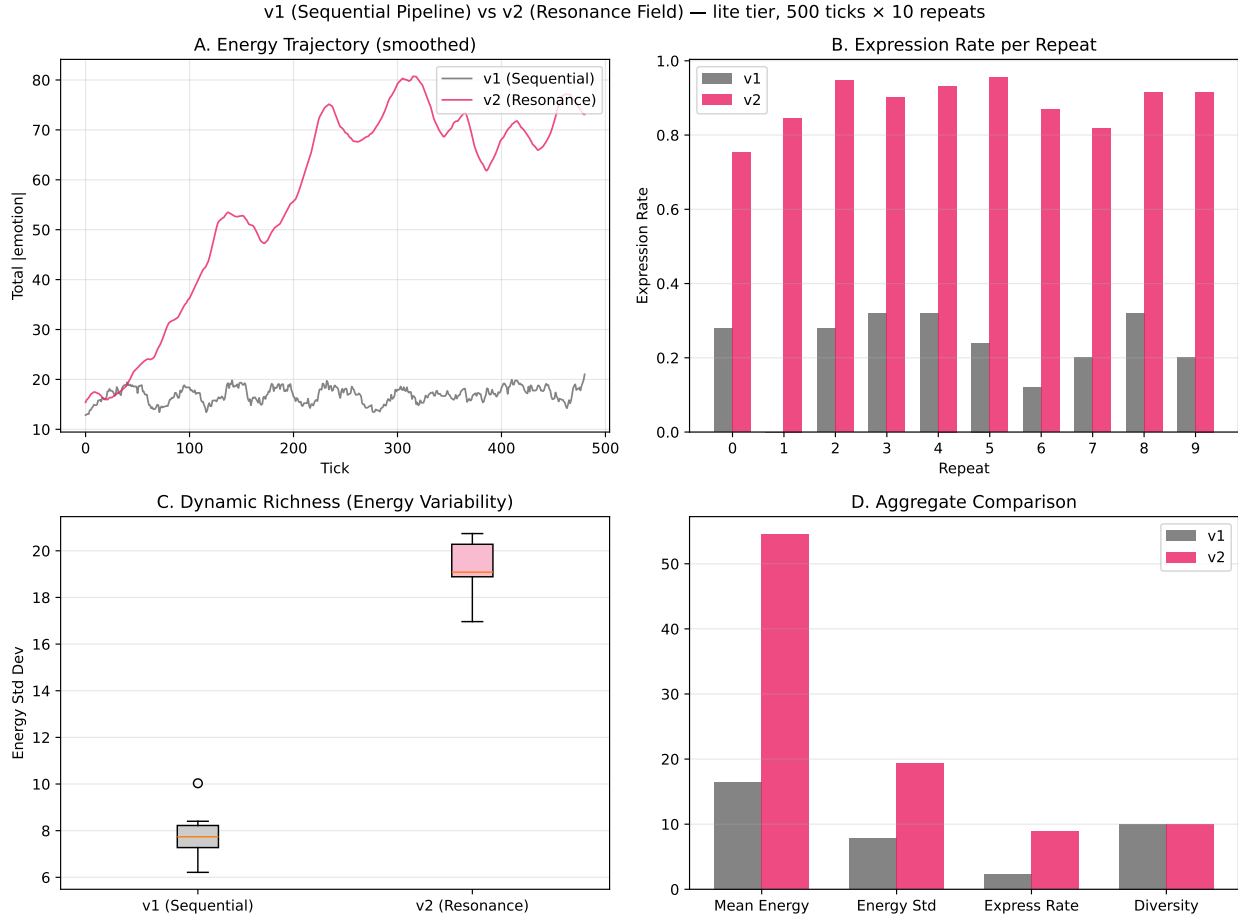


Figure 13: v1 (sequential pipeline) vs. v2 (resonance field) comparison at lite tier. (A) Smoothed energy trajectories showing v2’s higher dynamic range. (B) Expression rate per repeat showing v2’s dramatically higher and more consistent expression. (C) Energy variability (dynamic richness) box plot. (D) Aggregate metric comparison.

## 7 Discussion

**Comparison with existing approaches.** Classification-based affective systems [1, 2] are fundamentally stateless: identical input always produces identical output. Neural approaches (LSTMs, Transformers) maintain hidden state but offer no formal guarantees about boundedness, personality preservation, or interpretability. SylannEngine occupies a distinct position: it is a *dynamical system* with provable stability bounds (tanh saturation + dissipation), interpretable state (each module has semantic meaning), and deterministic personality-to-behavior mapping. The cost is that it requires no training data but also cannot learn from data—personality must be specified rather than inferred.

**The resonance field vs. sequential pipelines.** The v1 architecture processed information in fixed order: perception → gating → emotion → relation → decision → boundary → expression. This imposed artificial causal structure. The resonance field removes this constraint: all modules influence each other simultaneously through the simplicial coupling, and the effective processing order emerges from dynamics (which channels are strongest, which modules synchronize first). Experiment 4

demonstrates that this emergent ordering produces qualitatively different behavior (explosive vs. gradual synchronization) depending on coupling topology.

**Irreversibility as design principle.** Most AI systems are designed to be stateless or easily resettable. We argue that meaningful relational AI *requires* irreversibility: a system that can be trivially reset cannot form genuine relationships, because the other party’s investment has no lasting effect. The scar algebra and void calculus provide formal mechanisms for this irreversibility. Experiment 3 confirms that Hebbian traces persist: even after 500 idle ticks, previously active channels retain differentiation that would not exist in a memoryless system.

### Limitations.

- **Pure computation, no learning from data.** The system cannot adapt its personality dimensions from interaction history. Personality must be specified a priori by the designer.
- **Determinism.** Given identical input sequences and personality, two instances behave identically. There is no stochastic individuality beyond what personality provides.
- **Lite tier truncation.** The lite tier never achieves formal convergence ( $\varepsilon < 10^{-4}$ ) within 10 iterations due to Kuramoto oscillation at low channel counts. Output is meaningful but technically truncated.
- **Downgrade fidelity.** Max→lite hot-switching slightly exceeds the 10% energy jump threshold (Experiment 11), reflecting unavoidable information loss when pooling 441 channels to 42.
- **Personality dimensions.** The 7 dimensions are chosen heuristically (extending Big Five with patience and sovereignty) rather than derived from factor analysis.
- **No GPU backend.** The max tier’s 441-channel computation is CPU-bound; a torch backend is structured but not yet implemented.

**Scalability.** Computation scales as  $O(C \cdot d \cdot I)$  where  $C$  is channel count,  $d$  is state dimension, and  $I$  is iteration count. For the max tier:  $441 \times 32 \times 20 = 282,240$  multiply-accumulate operations per tick—well within real-time budgets on modern hardware. The three-tier design allows deployment scaling from embedded devices (lite: 42 channels, 8-dim, ~5ms) to research workstations (max: 441 channels, 32-dim, ~50ms) without architectural changes.

**Implications of irreversibility.** The irreversibility principle has practical consequences for deployment: (1) systems cannot be “factory reset” without losing relational history; (2) abusive interactions cause permanent damage (scars) that alter future behavior; (3) prolonged neglect generates void pressure that may trigger unsolicited expression. These properties are features, not bugs—they ensure that the system’s behavior reflects its actual history, making relationships with it meaningful rather than performative.

**Future work.** (1) Formal proof of convergence conditions for the resonance iteration under arbitrary personality parameterizations. (2) Multi-agent resonance: coupling fields across multiple SylannEngine instances for group dynamics. (3) Learned personality dimensions via factor analysis on human interaction data. (4) Integration with large language models for personality-aware text generation, where the resonance field modulates LLM sampling parameters.

## 8 Conclusion

We presented SylannEngine v2, a resonance field architecture for affective computation that replaces sequential pipelines with iterative convergence on a simplicial complex. The system integrates six mechanisms from mathematical physics and neuroscience—Hebbian plasticity, higher-order Kuramoto synchronization, Hopfield attractors, Hodge-theoretic identity, variational free energy, and bifurcation-based expression—into a unified framework governed by two foundational principles: irreversibility and personality-driven computation. Experiments across 11 protocols with 10 repetitions each confirm convergence within tier-specific bounds, explosive synchronization under higher-order coupling, stable energy envelopes over 1500+ ticks, meaningful plasticity with homeostatic conservation, and near-lossless tier hot-switching. The three-tier design enables deployment from resource-constrained devices (lite: 5ms, zero dependencies) to research environments (max: 441 channels, full  $\Delta^6$ ), with personality fully determining all system dynamics.

The framework is open-source under AGPL-3.0 at <https://github.com/Ayleovelle/SylannEngine>.

## References

- [1] R. W. Picard, *Affective Computing*. MIT Press, 1997.
- [2] A. Ortony, G. L. Clore, and A. Collins, *The Cognitive Structure of Emotions*. Cambridge University Press, 1988.
- [3] A. Mehrabian, “Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament,” *Current Psychology*, vol. 14, pp. 261–292, 1996.
- [4] K. R. Scherer, “Appraisal considered as a process of multilevel sequential checking,” in *Appraisal Processes in Emotion*, Oxford University Press, 2001.
- [5] J. J. Gross, “Emotion regulation: Current status and future prospects,” *Psychological Inquiry*, vol. 26, no. 1, pp. 1–26, 2015.
- [6] J. Park et al., “Generative agents: Interactive simulacra of human behavior,” in *Proc. UIST*, 2023.
- [7] Y. Kuramoto, “Self-entrainment of a population of coupled non-linear oscillators,” in *Int. Symp. Mathematical Problems in Theoretical Physics*, Springer, 1975.
- [8] A. P. Millán, J. J. Torres, and G. Bianconi, “Explosive higher-order Kuramoto dynamics on simplicial complexes,” *Physical Review Letters*, vol. 124, p. 218301, 2020.
- [9] J. J. Hopfield, “Neural networks and physical systems with emergent collective computational abilities,” *Proc. National Academy of Sciences*, vol. 79, pp. 2554–2558, 1982.
- [10] H. Ramsauer et al., “Hopfield networks is all you need,” in *Proc. ICLR*, 2021.
- [11] K. Friston, “The free-energy principle: A unified brain theory?” *Nature Reviews Neuroscience*, vol. 11, pp. 127–138, 2010.
- [12] W. V. D. Hodge, *The Theory and Applications of Harmonic Integrals*. Cambridge University Press, 1941.

- [13] D. O. Hebb, *The Organization of Behavior*. Wiley, 1949.
- [14] G. M. Edelman, *Neural Darwinism: The Theory of Neuronal Group Selection*. Basic Books, 1987.
- [15] I. Prigogine, "Time, structure and fluctuations," Nobel Lecture, 1977.
- [16] B. J. Baars, *A Cognitive Theory of Consciousness*. Cambridge University Press, 1988.
- [17] G. Tononi, "An information integration theory of consciousness," *BMC Neuroscience*, vol. 5, p. 42, 2004.
- [18] G. G. Turrigiano and S. B. Nelson, "Homeostatic plasticity in the developing nervous system," *Nature Reviews Neuroscience*, vol. 5, pp. 97–107, 2004.
- [19] S. H. Strogatz, "From Kuramoto to Crawford: Exploring the onset of synchronization in populations of coupled oscillators," *Physica D*, vol. 143, pp. 1–20, 2000.