Resonant Compression Systems

Learning as Dynamical Interference in Carrier-Field Architectures

Working Paper — January 2026

(anonymous draft)

Abstract

This paper develops a machine learning framework grounded in coupled-oscillator dynamics and carrier-field mediation. The core thesis is that learning is compression: a system has learned to the extent that it can represent observations with a shorter description while preserving reconstructability. Instead of storing pairwise weights, the model stores a sparse *presence matrix* $P \in \mathbb{R}^{N \times M}$ between N oscillators and M carrier modes. Pairwise relationships emerge dynamically through shared carrier participation, yielding an effective coupling $W(t)$ that is low-rank and context-dependent.

Key innovations include: (i) *Gaussian tuning strength* that derives coupling from phase alignment rather than a constant, creating "radio dial" lock-on behavior; (ii) *adaptive gate widths* that enable emergent specialization—carriers narrow their perceptual tolerance as they specialize; (iii) *spectral factorization* for mitosis, which partitions carriers by wavelength rather than blind duplication; (iv) *coherence-weighted metabolism*, where incoherent energy does not sustain carriers; and (v) a clean *observation interface* that separates measurement from dynamics, supporting classification and prediction without feedback into the engine.

The mathematical specification is complete and implementable: all equations can be directly translated to PyTorch with no hidden constants or arbitrary parameters beyond physical timescales.

Contents

1 Foundations **4**
- 1.1 Scope, flagship claim, and non-goals . . . 4
- 1.2 Non-convergence and operational stability . . . 4
- 1.3 The compression thesis . . . 4
- 1.4 Spectral decomposition as compression . . . 5
- 1.5 Oscillators as representational primitives . . . 5

2 Related Work **5**

3 From direct coupling to carrier-mediated coupling **6**
- 3.1 Why direct all-to-all coupling is problematic . . . 6
- 3.2 Carrier fields as a coupling medium . . . 6
- 3.3 Dynamic low-rank factorization . . . 7

4 A compact dynamical specification **7**
- 4.1 State variables . . . 7
- 4.2 Carrier dynamics: temporal gating (pulse antenna) . . . 7

4.2.1 Carrier as a gate . . . 8
4.2.2 Gated energy transfer with tuning . . . 8
4.3 Oscillator phase dynamics: carrier-mediated Kuramoto form . . . 8
4.3.1 Tuning strength: the radio dial principle . . . 9
4.4 Adaptive gate width: emergent specialization . . . 9
4.5 Amplitude dynamics (optional but recommended for experiments) . . . 10
4.6 Elastic coupling dynamics . . . 10

5 Coherence statistics, instability, and mitosis **11**
5.1 Weighted coherence . . . 11
5.2 Random-phase baseline . . . 11
5.3 Instability and mitosis . . . 12
5.4 Carrier spectral profile . . . 13
5.5 Entropy and dissolution (coherence-weighted metabolism) . . . 13
5.6 Spontaneous genesis (nucleation) . . . 14
5.7 Notes on redundancy . . . 14

6 Energy and stability perspectives **14**
6.1 A coherence objective . . . 15
6.2 Gradient structure (connection to Kuramoto) . . . 15
6.3 Low-rank reduction to an effective Kuramoto network . . . 15

7 Observation interface: one engine, multiple readouts **16**
7.1 Critical design principle . . . 16
7.2 Canonical observation function . . . 16
7.3 Classification via carrier configuration . . . 17
7.4 Prediction via state transition modeling . . . 17
7.5 Why not output carriers/oscillators? . . . 17

8 Compute requirements and scalable simulation **17**
8.1 Per-step complexity with sparse presence . . . 18
8.2 Discrete-time update equations (implementable) . . . 18
8.3 Compute reduction strategies . . . 18
8.3.1 (1) Remove the fast carrier frequency (rotating frame) . . . 18
8.3.2 (2) Multi-rate updates . . . 19
8.3.3 (3) Quasi-steady carriers . . . 19
8.3.4 (4) Sparse, bounded-degree presence . . . 19
8.3.5 (5) GPU implementation via sparse matrix multiplication . . . 19
8.4 Reference simulation loop . . . 19

9 Incommensurability ϵ and finite precision **19**
9.1 Clarification: irrational ratios are neither necessary nor sufficient . . . 19
9.2 Finite precision and eventual periodicity . . . 20
9.3 Workable solutions for ϵ under floating point . . . 21
9.3.1 Solution 1: deterministic detuning + coupling bound . . . 21
9.3.2 Solution 2: integer phase accumulators (eliminate rounding drift) . . . 21
9.3.3 Solution 3: controlled stochastic dither (avoid pathological locking) . . . 21
9.4 Rational approximations and recurrence time (useful rule of thumb) . . . 21

10 The Stream: environment and evaluation **21**
10.1 Core stream metrics . . . 22
10.2 Stream report (auto-generated) . . . 22
10.3 Experimental validation . . . 24
10.4 Analysis of experimental failures . . . 24
10.5 Known failure modes . . . 25
10.6 Limitations and scope . . . 26

11 Summary **26**

A Notation quick reference **27**

1 Foundations

1.1 Scope, flagship claim, and non-goals

T his paper proposes a *computational* dynamical system whose primary purpose is **representation formation and organization**: given oscillatory primitives and carrier-mediated coupling, the system should (i) produce *compressed* internal structure (sparse presence P and a small number of active carriers) and (ii) preserve *reconstructability* of coarse structure (e.g. cluster membership, low-rank relational structure) from that compressed representation.

T o make scope explicit, this work does **not** attempt to be:

- a drop-in replacement for gradient-trained deep neural networks,
- a competitive end-to-end supervised learner on large-scale benchmarks,
- a full reinforcement learning agent with credit assignment, exploration, and policy optimization,
- a biologically realistic model of neural circuits.

T his is a **purely computational model** motivated by stability, interference, and resource constraints; any resemblance to biological oscillations is incidental and not a claim of biological plausibility.

1.2 Non-convergence and operational stability

The system is **not guaranteed to converge** to a fixed point. Persistent, quasi-periodic, and metastable dynamics are expected and can be desirable. In this paper, "stability" is defined operationally: bounded carrier energy (via damping/saturation), bounded coherence statistics, and usable readouts/structures over the measurement horizon, rather than asymptotic convergence.

1.3 The compression thesis

Consider a stream of symbols:

`ABABABABABABABABABAB`.

If the stream can be described using a compact repetition operator, e.g. `rep(AB, 10)` (shorthand for "repeat `AB` ten times"), structure has been extracted and the description can be shorter than the raw string under an explicit coding scheme (where `x` is a token and the integer is encoded in log bits). This reduction is *compression*.

Definition 1 (Learning as compression)**.** Learning is the process of finding shorter descriptions of observed data. A system has learned to the extent that it can reconstruct observations from representations of lower entropy (shorter description length) than the observations themselves.

This framing is consistent with Kolmogorov complexity, Shannon entropy [Shannon, 1948], and Minimum Description Length (MDL) interpretations [Rissanen, 1978, Grünwald, 2007]. In this paper, the primitive operation is not "update weights to reduce a loss," but rather "discover resonant structure that persists and yields compressive representations."

1.4 Spectral decomposition as compression

A canonical example is Fourier decomposition [Bracewell, 1986]. For a signal $x(t)$, one may represent it as a superposition of sinusoids,

$$x(t) \approx \sum_k A_k \cos(\omega_k t + \varphi_k), \tag{1}$$

where many natural signals are approximately sparse in a suitable spectral basis.

Observation 1 (Basis choice controls compressibility)**.** Compression efficiency depends on the choice of basis; an "optimal" basis is one that renders typical signals sparse [Olshausen and Field, 1996].

1.5 Oscillators as representational primitives

The fundamental representational unit is an oscillator characterized by angular frequency ω, amplitude A, and phase ϕ. This choice is motivated by the rich mathematical structure of coupled oscillator systems [Kuramoto, 1984] and their demonstrated utility in modeling neural computation [Finger et al., 2013, Miyato et al., 2025]. Using a complex phasor representation,

$$z = Ae^{\mathrm{i}\phi} \in \mathbb{C}. \tag{2}$$

Frequency ω plays the role of identity (a dynamical "label"), amplitude A represents salience/activation, and phase ϕ represents relationships via relative alignment.

Observation 2 (Phase is relational)**.** Phase has meaning only relative to other oscillators; an isolated oscillator's absolute phase is arbitrary. This relational nature of phase has been exploited in neural binding theories, where phase synchronization serves as a mechanism for grouping distributed representations [Singer and Gray, 1995, Fries, 2005].

2 Related Work

This work draws on several research traditions that we briefly survey.

Coupled oscillators and synchronization. The Kuramoto model [Kuramoto, 1984] provides the mathematical foundation for understanding synchronization in populations of coupled oscillators. Extensive analysis [Strogatz, 2000, Acebrón et al., 2005, Gupta et al., 2014] has characterized the phase transitions and order parameters governing synchronization onset. Recent work has applied Kuramoto dynamics to neural network design, including artificial Kuramoto oscillatory neurons [Miyato et al., 2025] and graph neural networks [Li et al., 2023], demonstrating benefits for object binding, adversarial robustness, and uncertainty quantification.

Associative memory and Hopfield networks. Content-addressable memory via attractor dynamics was pioneered by Hopfield [1982, 1984]. Modern Hopfield networks [Ramsauer et al., 2020] achieve exponential storage capacity and connect directly to transformer attention mechanisms. Extensions include sparse variants [Hu et al., 2023] and continuous-time formulations [Yao and Santos, 2025].

Compression and learning. The view that learning is compression has deep roots in information theory [Shannon, 1948] and the Minimum Description Length principle [Rissanen, 1978, Grünwald, 2007]. Connections between Shannon entropy and Kolmogorov complexity [Grünwald and Vitányi, 2004] provide a unified theoretical foundation.

Sparse coding and neural representations. Sparse representations emerge from efficient coding of natural signals [Olshausen and Field, 1996, Lewicki and Sejnowski, 2000]. Combining sparse and low-rank structures has proven effective in modern architectures [Chen et al., 2021, Hu et al., 2022].

Neural synchronization and binding. Phase synchronization has been proposed as a neural mechanism for binding distributed feature representations [Singer and Gray, 1995, Fries, 2005]. Computational models demonstrate that coupled oscillator networks can implement binding and segmentation in visual processing [Finger et al., 2013].

Biologically-inspired learning rules. Hebbian learning [Hebb, 1949] and spike-timing-dependent plasticity [Caporale and Dan, 2008, Markram et al., 2012] provide biological grounding for local learning rules. Three-factor learning [Frémaux and Gerstner, 2016, Gerstner et al., 2018] extends these rules with neuromodulatory signals for reinforcement learning.

3 From direct coupling to carrier-mediated coupling

The foundational work on coupled oscillator systems was established by Kuramoto [1984], who introduced a tractable model of phase synchronization that has since become central to understanding collective dynamics in diverse systems [Strogatz, 2000, Acebrón et al., 2005].

3.1 Why direct all-to-all coupling is problematic

A direct coupling matrix $W \in \mathbb{R}^{N \times N}$ entails:

- **Storage:** $O(N^2)$ parameters in general.
- **Control:** positive feedback can induce runaway amplification absent explicit stabilization.
- **Conflation:** representational state and learned connectivity are entangled in a single object.

These issues are well-known in associative memory networks [Hopfield, 1982], where capacity limitations arise from the $O(N^2)$ weight matrix.

3.2 Carrier fields as a coupling medium

Definition 2 (Carrier field)**.** A *carrier field* (or carrier mode) k is a shared dynamical mode that oscillators can couple to. Oscillators interact through their shared participation in carriers rather than through explicit pairwise weights.

Definition 3 (Presence matrix)**.** Let $P \in \mathbb{R}_{\geq 0}^{N \times M}$ denote the *presence matrix*, where P_{ik} quantifies how strongly oscillator i participates in carrier k. In practice, P should be sparse: each oscillator participates in only $d \ll M$ carriers.

Clarification (what is "stored"). The system stores a topology matrix P. However, unlike neural networks where weights are frozen parameters optimized for error reduction, P represents *dynamic elastic bonds* that evolve continuously with the state variables. The effective coupling $W(t)$ is an emergent property of these bonds and the carrier energies.

3.3 Dynamic low-rank factorization

A common and useful abstraction is that the *effective* coupling between oscillators is low-rank and computed from the current carrier activities. Let $s(t) \in \mathbb{R}^M_{\geq 0}$ denote carrier activity magnitudes at time t (defined precisely later). Define

$$W(t) \;=\; P\,\mathrm{diag}(s(t))\,P^\top. \tag{3}$$

Theorem 1 (Dynamic low-rank factorization)**.** *If $M \ll N$, then $W(t)$ is at most rank M. Storage for P is $O(NM)$ in dense form and $O(dN)$ in sparse form (with d nonzeros per row), while pairwise couplings are never stored explicitly.*

This low-rank structure is reminiscent of recent work on intrinsic dimensionality in neural networks [Aghajanyan et al., 2021] and low-rank adaptation methods [Hu et al., 2022], which demonstrate that effective representations often lie in low-dimensional subspaces.

Remark 1. In the full dynamical system, $W(t)$ is not merely "looked up": it is an emergent object whose entries depend on carrier states, which depend on oscillator phases. This yields context-dependent connectivity.

4 A compact dynamical specification

This section provides an expanded mathematical model in a form convenient for analysis and simulation. The goal is to preserve the original intuition (interference, coherence, instability/mitosis) while making the system (i) implementable and (ii) amenable to stability reasoning.

4.1 State variables

We consider:

- Oscillators $i \in \{1, \ldots, N\}$ with phase $\phi_i(t) \in [0, 2\pi)$, natural frequency $\omega_i \in \mathbb{R}$, and amplitude $A_i(t) \in \mathbb{R}_{\geq 0}$.
- Carriers $k \in \{1, \ldots, M\}$ with complex amplitude $c_k(t) \in \mathbb{C}$.
- Presence matrix $P \in \mathbb{R}^{N \times M}_{\geq 0}$ (learned, typically sparse).

Define oscillator phasors $z_i(t) = A_i(t)e^{\mathrm{i}\phi_i(t)}$ and the vector $z(t) \in \mathbb{C}^N$. Let $c(t) \in \mathbb{C}^M$ collect carrier amplitudes.

4.2 Carrier dynamics: temporal gating (pulse antenna)

The original formulation treats a carrier as a continuously receptive oscillator. In the intended architecture, a carrier is instead a *temporal gate*: a stroboscopic sampling window that periodically opens and closes. This implements the antenna/pulse concept: the carrier "sees" the world only when its gate is open, creating time-domain orthogonality between competing signals.

4.2.1 Carrier as a gate

Let $c_k(t) = r_k(t)e^{\mathrm{i}\psi_k(t)}$ define the carrier phase ψ_k. We define a pulse gate as the direct conversion of the carrier phase into a two-level waveform:

$$G(\psi_k) = \begin{cases} 1, & \cos(\psi_k) \geq 0, \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

This is a half-cycle gate (50% duty): it is open for exactly half of each carrier period. Conceptually, the carrier behaves as a pulse antenna tuned to its own oscillation frequency.

Apex principle. Energy transfer is maximized when the *apex* of an input signal aligns with the *center* of the carrier gate. Signals peaking outside the open window are effectively invisible to that carrier.

4.2.2 Gated energy transfer with tuning

We replace continuous drive with *gated, tuning-weighted drive*:

$$\dot{c}_k \;=\; u_k \;-\; \gamma_k c_k \;-\; \beta_k |c_k|^2 c_k, \qquad u_k(t) \;=\; G(\psi_k(t)) \sum_{i=1}^{N} T_{ik}(t)\, P_{ik}(t)\, z_i(t), \tag{5}$$

where T_{ik} is the tuning strength from (9). Thus the carrier only accumulates input during gate-open windows, and the contribution from each oscillator is weighted by how well it is tuned to the carrier. If the gate is closed, $u_k(t) = 0$ and the carrier simply damps and saturates. If an oscillator is poorly tuned (low T_{ik}), its contribution is attenuated even when the gate is open.

Relation to energy/phase form. Writing $c_k = r_k e^{\mathrm{i}\psi_k}$ yields

$$\dot{r}_k = \Re\left(u_k e^{-\mathrm{i}\psi_k}\right) - \gamma_k r_k - \beta_k r_k^3, \tag{6}$$

$$\dot{\psi}_k = \frac{1}{r_k}\Im\left(u_k e^{-\mathrm{i}\psi_k}\right) \qquad (r_k > 0), \tag{7}$$

with the critical difference that u_k is intermittently zeroed by $G(\psi_k)$.

4.3 Oscillator phase dynamics: carrier-mediated Kuramoto form

Oscillator phases evolve according to their natural frequency plus carrier-mediated coupling, following a form closely related to the classical Kuramoto model [Kuramoto, 1984, Strogatz, 2000]:

$$\dot{\phi}_i \;=\; \omega_i \;+\; \Im\left(g_i e^{-\mathrm{i}\phi_i}\right), \qquad g_i \;=\; \sum_{k=1}^{M} T_{ik} P_{ik} c_k, \tag{8}$$

where $T_{ik} \in [0, 1]$ is the *tuning strength* between oscillator i and carrier k (defined below).

Critical design choice: no constant κ. Unlike the classical Kuramoto model, which uses a constant coupling strength κ, this system derives coupling strength *geometrically* from phase alignment. The tuning strength T_{ik} plays the role of κ, but it is not a parameter—it is an emergent property of the system state. This eliminates an arbitrary constant and ensures coupling is physically grounded.

4.3.1 Tuning strength: the radio dial principle

The tuning strength T_{ik} quantifies how well oscillator i is "tuned" to carrier k. The key insight is that coupling should behave like a radio dial: perfect alignment gives strong coupling, slight misalignment gives partial coupling, and large misalignment gives no coupling.

This is implemented via a Gaussian falloff rather than a cosine:

$$T_{ik} \;=\; \exp\!\left(-\frac{(\Delta\phi_{ik})^2}{\sigma_k}\right), \tag{9}$$

where $\Delta\phi_{ik}$ is the shortest angular distance between the oscillator peak and the carrier gate center (wrapped to $[-\pi, \pi]$), and $\sigma_k = (W_k/2)^2$ is the tuning sharpness derived from the carrier's gate width W_k.

Why Gaussian, not cosine? A cosine coupling (as in standard Kuramoto) is "too smooth"—it does not produce the sharp "lock-on" feel of real radio tuning. The Gaussian falloff creates:

- $T \approx 1$: Perfect alignment—strong coupling, clear signal.
- $T \approx 0.5$: Slight offset—partial coupling, signal plus noise.
- $T \approx 0$: Large offset—no coupling, noise only.

Sharpness from gate width. Setting $\sigma_k = (W_k/2)^2$ ensures that when the oscillator peak moves outside the carrier's gate window, coupling drops to $e^{-1} \approx 0.37$. This ties tuning sharpness to the physical gate geometry rather than an arbitrary constant.

Why this is a workable choice. Equation (8) with tuning-weighted coupling is:

- **Local:** oscillator i needs only its connected carriers (nonzero P_{ik}).
- **Stable under saturation:** carrier amplitudes remain bounded via (5).
- **Compatible with sparse simulation:** per-step cost is $O(\text{nnz}(P))$.
- **Geometrically grounded:** coupling strength emerges from alignment, not a tunable constant.

4.4 Adaptive gate width: emergent specialization

A critical extension is allowing each carrier's gate width W_k to *evolve* based on its coherence history. This creates emergent specialization: carriers that successfully capture coherent energy become more selective over time.

Gate width as perceptual tolerance. The gate width W_k determines how "picky" a carrier is about phase alignment:

- **Wide gate** (W_k large): generalist, captures broadly, exploratory.
- **Narrow gate** (W_k small): specialist, captures precisely, locked-on.

Evolution rule. Gate width evolves based on the coherence score D_k:

$$\dot{W}_k = \begin{cases} -\eta_{\text{narrow}}(D_k - 1)W_k, & D_k > 1 \quad \text{(high coherence: specialize)}, \\ +\eta_{\text{widen}}(1 - D_k)W_k, & D_k < 1 \quad \text{(low coherence: explore)}, \end{cases} \tag{10}$$

subject to bounds $W_{\min} \leq W_k \leq W_{\max}$.

Interpretation. When a carrier consistently receives aligned (coherent) input, its gate narrows—it becomes a specialist for that frequency/phase regime. When a carrier struggles with mixed signals (low coherence), its gate widens to explore a broader range. This creates a spectrum from broad exploratory carriers to narrow specialists, all emerging from the dynamics without explicit optimization.

4.5 Amplitude dynamics (optional but recommended for experiments)

Many initial experiments can fix $A_i \equiv 1$. However, amplitude dynamics can supply a natural *activity gate*:

$$\dot{A}_i = -\alpha(A_i - A_{i,0}) + I_i(t) - \rho A_i^3, \tag{11}$$

where $I_i(t)$ is external input drive, $\alpha > 0$ relaxes toward a baseline $A_{i,0}$, and $\rho > 0$ prevents divergence. A simple alternative is leaky integration with rectification:

$$A_i(t + \Delta t) = \max\{0, (1 - \alpha\Delta t)A_i(t) + \Delta t\, I_i(t)\}. \tag{12}$$

Amplitude gating provides two practical benefits:

- inactive oscillators can be omitted from updates (compute reduction),
- learning can be restricted to moments of significant activation.

4.6 Elastic coupling dynamics

The presence matrix P is not treated as a frozen parameter vector optimized to reduce error. Instead, each $P_{ik} \in [0, 1]$ is an *elastic bond* that must be continuously sustained by resonant energy. This section defines the bond dynamics as a purely instantaneous stability mechanism.

Elastic bond equation. For each oscillator i and carrier k, we evolve

$$\tau_p \dot{P}_{ik} = \underbrace{\alpha\, G(\psi_k)\, P_{ik}\, \max\Big(0, \Re\Big(c_k e^{-\mathrm{i}\phi_i}\Big)\Big)}_{\text{In-gate resonant capture}} - \underbrace{\lambda\, P_{ik}}_{\text{Elastic decay}}, \tag{13}$$

with $\tau_p > 0$ setting the bond timescale, reinforcement gain $\alpha > 0$, and elastic decay $\lambda > 0$.

Interpretation (what this is and is not). Equation (13) is **not** a statistical learning rule designed to capture history (e.g. Hebbian [Hebb, 1949] or STDP-like updates [Caporale and Dan, 2008, Markram et al., 2012]). It is an *instantaneous* stability constraint: the bond can stiffen only if the carrier gate is open *and* the oscillator aligns such that energy is captured during that open window. Outside the gate, reinforcement is exactly zero; the bond relaxes under elastic decay.

Bond snapping (physical pruning). To represent physical breaking of links and induce sparsity without any top-k selection, we apply a snap rule: if $P_{ik} < \varepsilon_{\text{snap}}$ for a small threshold (e.g. 10^{-2}), set $P_{ik} \leftarrow 0$. This yields an emergent sparse topology under continuous dynamics.

5 Coherence statistics, instability, and mitosis

Carrier evolution (instability-driven mitosis) is central to the "architecture is learned" claim. To make it implementable and statistically grounded, we refine coherence measures and thresholds. The use of order parameters to characterize synchronization states has a rich history in the coupled oscillator literature [Daido, 1996, Gupta et al., 2014].

5.1 Weighted coherence

Define weights $w_{ik} = P_{ik}A_i$ and the carrier drive

$$u_k = \sum_i w_{ik} e^{\mathrm{i}\phi_i}. \tag{14}$$

A natural coherence score is

$$\mathrm{coh}(k) \;=\; \frac{|u_k|}{\sum_i w_{ik}} \in [0,1], \tag{15}$$

where $\mathrm{coh}(k) = 1$ indicates perfect alignment among the contributing phasors.

5.2 Random-phase baseline

If phases are independent and uniformly distributed, u_k is a 2D random walk in the complex plane. For large numbers of contributors, the magnitude is approximately Rayleigh-distributed.

Proposition 1 (Expected magnitude under random phase)**.** *Let $\theta_i \sim \mathrm{Unif}[0, 2\pi)$ i.i.d. and define*

$$S = \sum_{i=1}^{n} a_i e^{\mathrm{i}\theta_i}.$$

Then, for moderate/large n,

$$\mathbb{E}|S| \;\approx\; \sqrt{\frac{\pi}{4}} \sqrt{\sum_{i=1}^{n} a_i^2}. \tag{16}$$

In the equal-weight case $a_i = a$, $\mathbb{E}|S| \approx \sqrt{\pi/4}\, a\sqrt{n} \approx 0.886\, a\sqrt{n}$.

Combining (15) and (16), the expected coherence under random phase is approximately

$$\mathbb{E}[\mathrm{coh}(k)] \;\approx\; \sqrt{\frac{\pi}{4}} \frac{\sqrt{\sum_i w_{ik}^2}}{\sum_i w_{ik}}. \tag{17}$$

For equal w_{ik} over n contributors, this reduces to $\approx 0.886/\sqrt{n}$.

5.3 Instability and mitosis

The system uses a coherence-to-baseline ratio to detect when a carrier is *energetically unstable* under mixed-phase drive.

Let

$$b_k \;=\; \sqrt{\frac{\pi}{4}}\,\frac{\sqrt{\sum_i w_{ik}^2}}{\sum_i w_{ik}} \tag{18}$$

be the baseline in (17). Define a division score

$$D_k \;=\; \frac{\mathrm{coh}(k)}{b_k}. \tag{19}$$

Then $D_k \approx 1$ corresponds to random-phase coherence, $D_k > 1$ indicates above-random alignment, and $D_k < 1$ indicates *active cancellation* beyond random.

Definition 4 (Instability criterion)**.** A carrier k is declared unstable if

$$D_k < \theta_{\mathrm{div}} \quad \text{persistently over a window of length } T_{\mathrm{div}}, \tag{20}$$

where $\theta_{\mathrm{div}} \in (0, 1)$ and persistence avoids spurious splits from noise.

Mitosis as spectral factorization (not blind duplication). When instability persists ($D_k < \theta_{\mathrm{div}}$ *and* the carrier's spectral profile is multimodal—see Section 5.4), the carrier undergoes *spectral factorization*:

1. **Partition by frequency:** The bonded oscillators are partitioned into two clusters based on their natural frequencies ω_i. This is done by computing a weighted frequency distribution and splitting at the median.

2. **Asymmetric bond inheritance:** The parent carrier keeps bonds to the low-frequency cluster; a new child carrier inherits bonds to the high-frequency cluster. Bonds are *partitioned*, not copied.

3. **Frequency recentering:** Each offspring's intrinsic frequency ω_k is shifted to the spectral center of its inherited cluster:

$$\omega_{\mathrm{parent}} \leftarrow \bar{\omega}_{\mathrm{low}}, \qquad \omega_{\mathrm{child}} \leftarrow \bar{\omega}_{\mathrm{high}},$$

   where $\bar{\omega} = \sum_i w_{ik}\omega_i / \sum_i w_{ik}$ is the weighted mean frequency.

4. **Small phase noise:** Independent thermal kicks are applied to differentiate phases, but differentiation is primarily spectral.

Why spectral factorization, not blind duplication? Blind duplication (copying all bonds and adding noise) produces two nearly identical carriers that compete for the same oscillators. Spectral factorization produces *immediately differentiated* offspring: one specializing in low frequencies, the other in high frequencies. This resolves interference *causally* (based on which frequencies conflict) rather than stochastically.

Multimodality requirement. Mitosis only triggers when the carrier's spectral profile is multimodal (Section 5.4). Low coherence alone is not sufficient—the carrier must be attempting to bind signals that cannot be represented by a single wavelength hypothesis.

5.4 Carrier spectral profile

Each carrier maintains a *spectral profile* derived from its bonded oscillators. This is an *observable property*, not a learned parameter.

Effective weights. Define the effective weight of oscillator i on carrier k:

$$w_{ik} = P_{ik} \cdot A_i \cdot T_{ik}, \tag{21}$$

combining bond strength, amplitude, and tuning.

Spectral center of mass. The carrier's spectral center is the weighted mean frequency of its bonded oscillators:

$$\bar{\omega}_k = \frac{\sum_i w_{ik}\,\omega_i}{\sum_i w_{ik}}. \tag{22}$$

Spectral variance. The spread of frequencies is:

$$\mathrm{Var}_k = \frac{\sum_i w_{ik}\,(\omega_i - \bar{\omega}_k)^2}{\sum_i w_{ik}}. \tag{23}$$

Multimodality detection. A carrier is considered *multimodal* if its spectral variance exceeds a threshold relative to the overall frequency range. Multimodality indicates the carrier is attempting to bind incompatible frequency clusters—a precondition for spectral mitosis.

5.5 Entropy and dissolution (coherence-weighted metabolism)

Mitosis provides a mechanism to increase representational capacity when interference makes a mode unstable. To prevent unbounded carrier proliferation, we introduce a *metabolic cost* for carriers.

Metabolic principle: coherence matters. A carrier must continuously capture sufficient *coherent* energy to persist. The key insight is that raw intake magnitude is not sufficient—a carrier capturing incoherent noise can have high amplitude but provides no representational value.

Coherence-weighted intake. Define the effective intake as:

$$I_k^{\mathrm{eff}}(t) = I_k(t) \cdot D_k(t), \tag{24}$$

where $I_k(t) = \mathrm{EMA}(|u_k(t)|)$ is the smoothed raw intake and D_k is the coherence score. A carrier with high raw intake but low coherence (capturing conflicting signals) has low effective intake.

Dissolution criterion. Dissolve carrier k if $I_k^{\mathrm{eff}}(t)$ remains below a threshold $I_{\min}$ over a persistence window. This implements survival-of-the-fittest at the dynamical level: carriers that do not consistently capture *coherent* energy are removed. Incoherent energy does not "feed" the carrier well.

Implication. Coherence becomes a survival criterion, not just a mitosis trigger. Carriers that specialize (narrow gate, high coherence) are metabolically healthier than carriers capturing noise.

5.6 Spontaneous genesis (nucleation)

If mitosis increases capacity under interference, a separate question is how the first carrier (or new carriers after extinction) arise. In the stream setting, the system begins with no carriers, i.e. an initially "empty" carrier field. If the sensory stream contains high-energy oscillators that are not effectively gated by any existing carrier (low $P_{i\cdot}$), then the vacuum is unstable. We model this instability as a *nucleation event*: a new carrier mode is spontaneously generated.

Unbound oscillators. Define an "unbound" set of oscillators as those whose strongest bond is below a threshold,

$$\mathcal{U} = \{i : \max_k P_{ik} < P_{\min}\}.$$

Phase-clustered genesis (not just amplitude). Genesis requires **both** sufficient amplitude and phase coherence among unbound oscillators. This prevents nucleation from incoherent noise.

Compute the order parameter among unbound oscillators:

$$R_{\mathcal{U}} = \left| \frac{\sum_{i\in\mathcal{U}} A_i e^{\mathrm{i}\phi_i}}{\sum_{i\in\mathcal{U}} A_i} \right| \in [0,1]. \tag{25}$$

$R_{\mathcal{U}} = 1$ indicates perfect phase alignment; $R_{\mathcal{U}} \approx 0$ indicates random phases.

Genesis criterion. Nucleate a new carrier if:

$$\sum_{i\in\mathcal{U}} A_i > A_{\text{pressure}} \quad \text{AND} \quad R_{\mathcal{U}} > R_{\text{genesis}}. \tag{26}$$

The first condition requires sufficient unbound energy; the second requires that unbound oscillators are attempting to synchronize (phase-clustered). Incoherent unbound energy does *not* trigger genesis.

Apex-aligned birth. The newborn carrier is initialized with its phase aligned to the strongest unbound oscillator's current phase (apex alignment) and its intrinsic rotation frequency seeded by that oscillator. This ensures that novel information in the stream is met with a representational resource that can immediately begin gated capture and then compete for survival under metabolic cost.

5.7 Notes on redundancy

This paper focuses on a "no-compromise" pure dynamical system: carriers are created only by instability-driven mitosis, and separation is achieved purely by elastic bond dynamics. We therefore do not introduce an explicit merger algorithm; redundancy can be treated as a measurement issue (carriers with negligible energy or negligible total presence can be ignored in readout).

6 Energy and stability perspectives

This section supplies an analytical "handle" for understanding attractors and why the dynamics tends to produce coherent structure. The objective is not to prove full convergence (which may be false once inputs and detuning are included), but to show that the update directions are not arbitrary.

6.1 A coherence objective

The energy-based perspective on oscillator networks connects to classical work on Hopfield networks [Hopfield, 1982, 1984] and their modern extensions [Ramsauer et al., 2020, Hu et al., 2023].

For fixed presences P and fixed amplitudes A_i, define the carrier drives

$$u_k(\phi) = \sum_i P_{ik} A_i e^{\mathrm{i}\phi_i}.$$

Consider the scalar objective

$$\mathcal{L}(\phi) \;=\; \sum_{k=1}^{M} |u_k(\phi)|. \tag{27}$$

Intuitively, $\mathcal{L}$ is large when carriers receive aligned contributions and small when contributions cancel.

6.2 Gradient structure (connection to Kuramoto)

Assume, for this analysis only, that carrier phases satisfy $\psi_k = \arg(u_k)$ (fast carrier phase alignment) and ignore intrinsic frequencies ($\omega_i = 0$). Then one can show that increasing $\mathcal{L}$ encourages phase alignment.

Proposition 2 (Oscillator update as ascent on carrier coherence (idealized))**.** *Let $u_k \neq 0$ and define $\psi_k = \arg(u_k)$. Then*

$$\frac{\partial |u_k|}{\partial \phi_i} = P_{ik} A_i \sin(\psi_k - \phi_i). \tag{28}$$

Consequently, the phase dynamics

$$\dot{\phi}_i \propto \sum_k P_{ik} |u_k| \sin(\psi_k - \phi_i)$$

performs gradient ascent on $\mathcal{L}(\phi)$ (up to scaling) in this idealized regime.

Remark 2. The implemented dynamics (8) replaces $|u_k|$ by the carrier amplitude $|c_k|$, which is a smoothed and saturated version of $|u_k|$. Thus the "gradient intuition" survives approximately while remaining numerically stable.

6.3 Low-rank reduction to an effective Kuramoto network

If carrier dynamics is much faster than oscillator phases, one may approximate carriers as quasi-steady:

$$0 \approx u_k - \gamma_k c_k \quad \Rightarrow \quad c_k \approx \gamma_k^{-1} u_k \qquad (\beta_k = 0 \text{ or small}). \tag{29}$$

Substituting (29) into (8) yields

$$g_i \approx \sum_k P_{ik} \gamma_k^{-1} \sum_j P_{jk} A_j e^{\mathrm{i}\phi_j} = \sum_j \Big(\sum_k P_{ik} \gamma_k^{-1} P_{jk} \Big) A_j e^{\mathrm{i}\phi_j}. \tag{30}$$

Define the symmetric effective coupling

$$K_{ij} \;=\; \sum_{k=1}^{M} P_{ik} \gamma_k^{-1} P_{jk}. \tag{31}$$

Then (8) becomes a weighted Kuramoto model:

$$\dot{\phi}_i \approx \omega_i + \kappa \sum_{j=1}^{N} K_{ij} A_j \sin(\phi_j - \phi_i). \tag{32}$$

Implication. This makes the system comparable to well-studied synchronization and associative memory models [Hopfield, 1982, Kuramoto, 1984], but with a *learned low-rank* coupling $K = P\Gamma^{-1}P^\top$. Recent work on Kuramoto dynamics in neural networks [Miyato et al., 2025, Li et al., 2023] has demonstrated the practical utility of this connection for tasks including object binding, adversarial robustness, and uncertainty quantification.

7 Observation interface: one engine, multiple readouts

A key question is how continuous attractor states produce discrete or structured outputs. This section describes an *observation interface* that cleanly separates measurement from dynamics.

7.1 Critical design principle

Outputs measure resonance; they do not participate in it.

No gradients, no reward shaping, no feedback into carrier dynamics. Tasks are *consumers*, not *drivers*. This preserves emergence: all specialization is physics-driven, not task-driven.

One engine, multiple readouts. The same observation interface supports:

- **Classification:** What carrier configurations exist? (state identification)
- **Next-token prediction:** How does the state evolve? (temporal modeling)
- **Unsupervised structure discovery:** What clusters emerge? (representation analysis)

You do **not** need to choose between classification and prediction. The difference lies entirely in how you read and aggregate the same state.

7.2 Canonical observation function

At any time t, the engine exposes a state snapshot observe(t) containing:

Carrier-level observables (per carrier k):

- Energy: $|c_k|$
- Phase: $\arg(c_k)$
- Intrinsic frequency: ω_k
- Intake: I_k (smoothed)
- Coherence: D_k
- Gate width: W_k (specialization level)
- Spectral center: $\bar{\omega}_k$
- Spectral variance: Var_k

Soft assignment matrix:

$$m_{ik} = \frac{P_{ik} \cdot T_{ik}}{\sum_{k'} P_{ik'} \cdot T_{ik'}}, \tag{33}$$

which gives the normalized assignment of oscillator i to carrier k. This is the "which carrier does this oscillator belong to" answer, but soft (graded) rather than hard.

Global metrics:

- Synchronization order parameter $R(t)$
- Description length proxy L_{comp}
- Population counts N, M, $\text{nnz}(P)$

7.3 Classification via carrier configuration

A "class" corresponds to a stable carrier or carrier constellation. Classification score at time t:

$$\text{score}(\text{class}_j) = \sum_k |c_k| \cdot \text{similarity}(\text{carrier}_k, \text{prototype}_j), \tag{34}$$

where prototypes live **outside** the engine (learned, clustered, or hand-defined). The engine never "knows" it is classifying.

7.4 Prediction via state transition modeling

Treat observe(t) as a vector in state space. Learn a temporal model (Markov, RNN, transformer) to predict observe($t + \Delta t$). The engine supplies *structure*; the predictor learns *temporal regularities on top.*

7.5 Why not output carriers/oscillators?

Earlier drafts proposed "output carriers" that participate in dynamics. This was rejected because:

- Outputs become goals, not observables.
- Emergence collapses into wiring.
- The system becomes task-driven rather than physics-driven.

The observation interface avoids these pitfalls by keeping measurement strictly external.

8 Compute requirements and scalable simulation

The original draft correctly notes that continuous-time dynamics incurs cost on digital hardware. This section makes that cost explicit and provides implementable strategies to reduce it.

8.1 Per-step complexity with sparse presence

Assume P has d nonzeros per oscillator (row), so $\mathrm{nnz}(P) = dN$. A single Euler step for (5)–(8) requires:

1. Compute $u = P^\top z$ (complex): $O(\mathrm{nnz}(P))$.
2. Update c for M carriers: $O(M)$.
3. Compute $g = Pc$: $O(\mathrm{nnz}(P))$.
4. Update ϕ for N oscillators: $O(N)$.
5. Update P for $\mathrm{nnz}(P)$ edges (if learning every step): $O(\mathrm{nnz}(P))$.

Thus

$$\text{cost per step} = O(\mathrm{nnz}(P) + N + M) \approx O(dN) \quad \text{for } d \gg 1,\ M \ll N. \tag{35}$$

8.2 Discrete-time update equations (implementable)

With step size Δt, a baseline explicit scheme is:

$$z_i \leftarrow A_i e^{\mathrm{i}\phi_i}, \tag{36}$$

$$u_k \leftarrow \sum_i P_{ik} z_i, \tag{37}$$

$$c_k \leftarrow c_k + \Delta t \left(u_k - \gamma_k c_k - \beta_k |c_k|^2 c_k\right), \tag{38}$$

$$g_i \leftarrow \sum_k P_{ik} c_k, \tag{39}$$

$$\phi_i \leftarrow \mathrm{wrap}\Big(\phi_i + \Delta t \left[\omega_i + \kappa\, \Im(g_i e^{-\mathrm{i}\phi_i})\right]\Big), \tag{40}$$

with optional A_i update (11) and P update (**??**) (or variants).

Stability note. Because c_k is damped, explicit Euler is often stable for moderate Δt. If stiffness arises (large γ_k), use an exponential integrator for the linear part:

$$c_k(t + \Delta t) \approx e^{-\gamma_k \Delta t} c_k(t) + \frac{1 - e^{-\gamma_k \Delta t}}{\gamma_k}\, u_k(t), \tag{41}$$

with saturation handled explicitly.

8.3 Compute reduction strategies

8.3.1 (1) Remove the fast carrier frequency (rotating frame)

If oscillators have large shared base frequency Ω with small detunings δ_i, write

$$\omega_i = \Omega + \delta_i.$$

Simulating in the rotating frame eliminates Ω:

$$\phi_i(t) = \Omega t + \tilde{\phi}_i(t) \quad \Rightarrow \quad \dot{\tilde{\phi}}_i = \delta_i + \text{coupling}.$$

This can reduce required step sizes by orders of magnitude.

8.3.2 (2) Multi-rate updates

Update different subsystems at different rates:

- phases ϕ every step,
- carriers c every r_c steps (interpolating between),
- learning P every r_P steps (slow learning),
- mitosis checks every r_{arch} steps.

This is justified when learning and architecture changes are slower than phase mixing.

8.3.3 (3) Quasi-steady carriers

If carriers are fast relative to phases, replace carrier integration with a static nonlinearity:

$$c_k \approx \frac{u_k}{\gamma_k + \beta_k |u_k|^2}, \tag{42}$$

which is a divisive normalization form [Carandini and Heeger, 2012, Schwartz and Simoncelli, 2001]. This collapses the model to two sparse matrix-vector products per step ($u = P^\top z$, $g = Pc$) plus pointwise nonlinearities.

8.3.4 (4) Sparse, bounded-degree presence

Enforce bounded degree per oscillator and per carrier:

- Each oscillator connects to at most d carriers (row sparsity),
- Each carrier accepts at most q oscillators (column cap) to limit u_k fan-in.

Column caps act as a resource constraint and can serve as a compression prior.

8.3.5 (5) GPU implementation via sparse matrix multiplication

The dominant operations are sparse matrix–vector products (SpMV) or sparse matrix–dense vector multiplications with complex numbers [Bell and Garland, 2009]. Store P in CSR (row-major) for computing $g = Pc$ and in CSC (column-major) or transposed CSR for $u = P^\top z$. On modern accelerators, these operations scale well when $\text{nnz}(P)$ is large and memory access is coalesced.

8.4 Reference simulation loop

9 Incommensurability ϵ and finite precision

The original draft motivates an "essential incommensurability" ϵ to prevent phase locking and preserve continuous dynamics. However, several clarifications are needed for correctness and implementability.

9.1 Clarification: irrational ratios are neither necessary nor sufficient

Not sufficient. Even if ω_1/ω_2 is irrational, *coupling* can still synchronize oscillators: in Kuramoto-type systems, sufficiently strong coupling can produce frequency locking when detuning is small relative to coupling.

Algorithm 1 Baseline resonant learning simulation (sparse)

Initialize ϕ, ω, A, c, P with small nonzero bonds.
for $t = 1$ to T **do**
  $z_i \leftarrow A_i e^{\mathrm{i}\phi_i}$ for all i
  $u \leftarrow P^\top z$ ▷ carrier drives
  $c \leftarrow c + \Delta t\,(u - \Gamma c - B(c))$ ▷ carrier update
  $g \leftarrow Pc$ ▷ back-influence
  $\phi_i \leftarrow \mathrm{wrap}\big(\phi_i + \Delta t\,[\omega_i + \kappa\,\Im(g_i e^{-\mathrm{i}\phi_i})]\big)$
  Update P using elastic coupling dynamics (Eq. (13)) and snap bonds below threshold.
  if $t \bmod r_{\mathrm{arch}} = 0$ **then**
    Compute D_k and trigger mitosis on persistent instability.
  end if
end for

Algorithm 2 Architecture evolution loop (spectral factorization)

Inputs: carrier states $\{c_k\}$, elastic bonds P, oscillator frequencies $\{\omega_i\}$, threshold θ_{div}
Compute (per carrier k):
  Coherence $D_k = \mathrm{coh}(k)/b_k$
  Spectral profile: center $\bar{\omega}_k$, variance Var_k, multimodality flag
if $D_k < \theta_{\mathrm{div}}$ persistently **AND** carrier k is multimodal **then**
  Partition bonded oscillators by frequency into clusters $\mathcal{L}$ (low) and $\mathcal{H}$ (high)
  Parent keeps bonds to $\mathcal{L}$; zero bonds to $\mathcal{H}$
  Create child carrier with bonds only to $\mathcal{H}$
  Recenter frequencies: $\omega_{\mathrm{parent}} \leftarrow \bar{\omega}_{\mathcal{L}}$, $\omega_{\mathrm{child}} \leftarrow \bar{\omega}_{\mathcal{H}}$
  Apply small phase noise for differentiation
end if

Not necessary (for finite-time experiments). On any finite time horizon T_{run}, what is needed is that the system does not enter a short periodic orbit or global synchronization that destroys useful interference structure. This can be achieved by (i) nonzero detuning, (ii) noise or drift, and (iii) architectural constraints that prevent total coupling collapse.

We therefore reinterpret ϵ as an *effective detuning floor*:

Definition 5 (Effective incommensurability)**.** Let $\Delta\omega_{ij} = |\omega_i - \omega_j|$. An ϵ-incommensurate set of oscillators satisfies $\Delta\omega_{ij} \geq \epsilon$ for relevant interacting pairs (i, j) (e.g., those sharing carriers), possibly after accounting for coupling-induced frequency shifts.

9.2 Finite precision and eventual periodicity

A discrete-time simulation in floating point is a finite-state dynamical system, hence eventually periodic. The practical question is whether the period is *astronomically large* compared to T_{run}. For standard 64-bit floats and modular phase wrapping, the effective state space is so large that short cycles are typically not an issue unless the dynamics itself drives the system into a low-dimensional synchronized manifold.

9.3 Workable solutions for ϵ under floating point

9.3.1 Solution 1: deterministic detuning + coupling bound

Choose natural frequencies with a guaranteed separation:

$$\omega_i = \omega_0 + \delta_i, \qquad \delta_i \sim \mathrm{Unif}[-\Delta, \Delta], \tag{43}$$

and constrain coupling so that global synchronization is unlikely:

$$\kappa \max_i \sum_k P_{ik}|c_k| \;\lesssim\; \Delta. \tag{44}$$

Empirically, one can monitor order parameters (Section 10) to detect collapse into synchronization.

9.3.2 Solution 2: integer phase accumulators (eliminate rounding drift)

Represent phase as a Q-bit integer accumulator (standard in digital oscillators). Let $\Phi_i \in \{0, 1, \ldots, 2^Q - 1\}$ and interpret $\phi_i = 2\pi\Phi_i/2^Q$. Update by exact modular arithmetic:

$$\Phi_i \leftarrow (\Phi_i + \Delta_i) \bmod 2^Q, \tag{45}$$

where Δ_i is an integer frequency increment. This yields *zero* floating point rounding drift in the free-running phase. Detuning is implemented by choosing distinct Δ_i with large least common multiples, producing huge recurrence periods (often $> 2^Q$ steps if gcd is small).

Coupling terms can still be computed in floating point (using sin/cos of ϕ_i), but the fundamental phase progression remains exact.

9.3.3 Solution 3: controlled stochastic dither (avoid pathological locking)

Add a small noise term to phase or frequency:

$$\omega_i(t) = \omega_i + \sigma\,\xi_i(t), \tag{46}$$

where $\xi_i(t)$ is e.g. Gaussian white noise and σ is small. This prevents exact recurrence and models physical fluctuations. Dither is especially useful when using low precision (float32) or large time steps.

9.4 Rational approximations and recurrence time (useful rule of thumb)

If two oscillators have a rational frequency ratio $\omega_2/\omega_1 = p/q$ (in lowest terms), their phases exactly realign after time

$$T_{\mathrm{rec}} = \frac{2\pi q}{\omega_1}. \tag{47}$$

Thus, to *behave like* an irrational ratio over an experiment horizon T_{run}, it suffices that $T_{\mathrm{rec}} \gg T_{\mathrm{run}}$. Integer phase accumulators with $Q = 64$ make q effectively enormous for generic choices of Δ_i.

10 The Stream: environment and evaluation

The intended setting is an open-ended sensory stream: the system does not know group labels, tasks, or episode boundaries. It observes a sequence of transient signals with finite duration; signals ring down rather than disappearing abruptly, and the topology evolves continuously.

Reproducible artifact generation. All paper artifacts referenced in this section are generated by a single command:

`python3 tmp/rez/paper_artifacts.py`.

It writes figures/tables into `tmp/rez/artifacts/`, which this LaTeX file includes automatically via `\IfFileExists`.

10.1 Core stream metrics

Carrier coherence. Track $\mathrm{coh}(k)$ in (15) and the normalized score D_k.

Synchronization order parameter. A global phase order parameter, introduced by Kuramoto [1984] and analyzed extensively by Strogatz [2000], detects collapse into full synchrony:

$$R(t) = \left| \frac{1}{N} \sum_{i=1}^{N} e^{\mathrm{i}\phi_i(t)} \right| \in [0, 1]. \tag{48}$$

Large R indicates global alignment (often undesirable if it destroys representational diversity).

Sparsity and description length proxies. Track $\mathrm{nnz}(P)$, number of active carriers ($|c_k|$ above threshold), and average row degree d. These are direct proxies for representational complexity, consistent with sparse coding principles in neural systems [Olshausen and Field, 1996, Lewicki and Sejnowski, 2000].

Compression metric (operationalized). To make the thesis "learning as compression" empirically testable, following the MDL tradition [Rissanen, 1978, Grünwald, 2007], we operationalize compression using a simple description-length proxy:

$$L_{\mathrm{comp}} \;=\; \mathrm{nnz}(P) \;+\; M_{\mathrm{active}}, \tag{49}$$

where $\mathrm{nnz}(P)$ is the number of nonzero presences and M_{active} is the number of active carriers. This is intentionally crude but *computable* and captures the two dominant storage/complexity terms in this framework.

Attractor stability. In the stream setting, stability is measured by churn: carrier birth (mitosis) rate, dissolution rate, and the distribution of carrier lifetimes.

10.2 Stream report (auto-generated)

Stream run. We simulate an open-ended sensory stream for 60 s (seed=0, Δt=0.005 s). Mean N=4.7, mean M=0.9, mean $\mathrm{nnz}(P)$=6.9, mean L_{comp}=7.8. Mitoses=1, dissolutions=22.

Table 1: Stream summary metrics (time-averaged).

$\bar{N}$	$\bar{M}$	$\overline{\mathrm{nnz}(P)}$	$\overline{L_{\mathrm{comp}}}$	Mitoses	Dissolutions
4.7	0.9	6.9	7.8	1	22

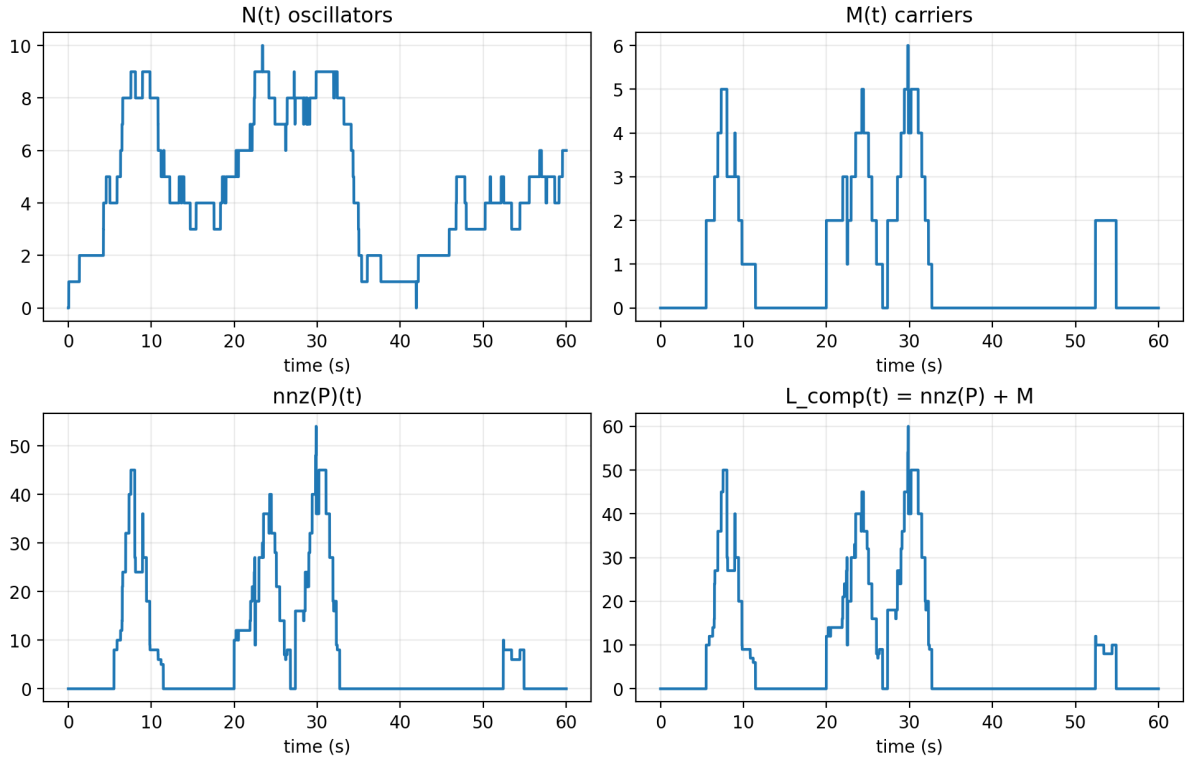

Figure 1: Stream metrics over time: $N(t)$, $M(t)$, nnz(P), and $L_{textcomp}(t)$.

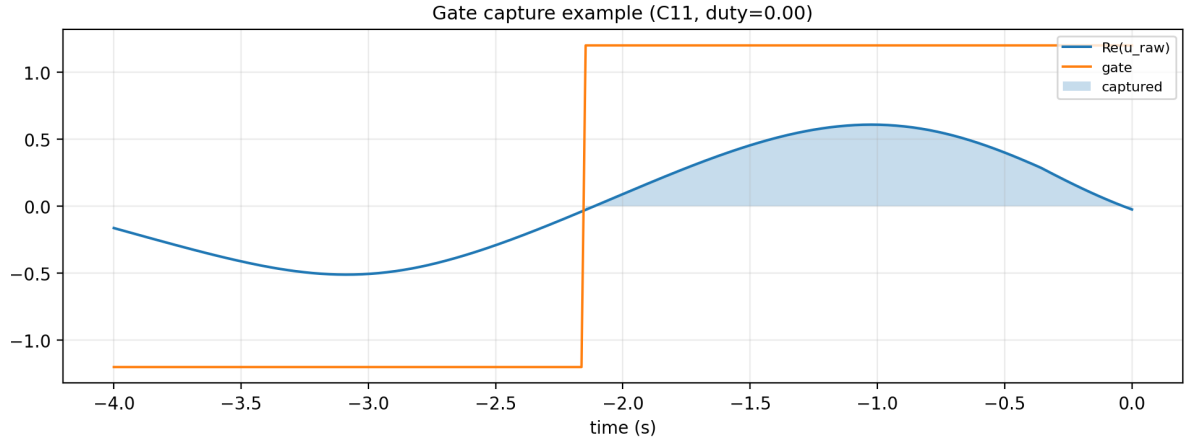

Figure 2: Gate capture visualization: carrier gate (square wave) and the portion of the composite drive captured while the gate is open.

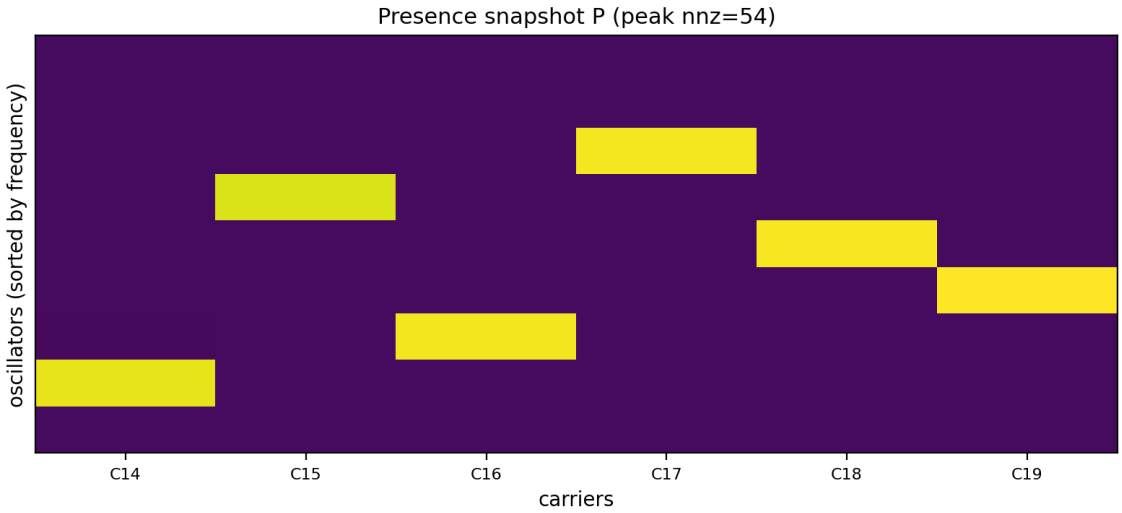

Figure 3: Presence matrix snapshot P (heatmap), showing emergent sparsification by bond snapping.

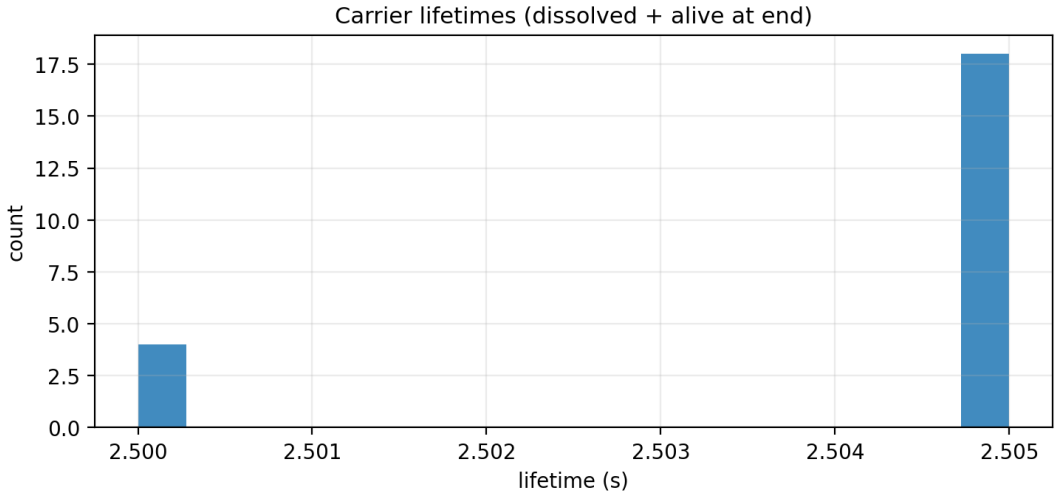

Figure 4: Carrier lifetime distribution under mitosis and metabolic dissolution.

10.3 Experimental validation

To test the claims made in this paper, we conducted a suite of rigorous experiments. Each experiment tests a specific hypothesis about the system's behavior. Results are reported honestly: failures are not hidden but rather indicate areas for further investigation.

(Experiment summary not found. Run `python3 tmp/rez/paper_artifacts.py` *to generate.)*

(Experiment results table not found.)

10.4 Analysis of experimental failures

The experimental results reveal important gaps between theory and implementation that warrant discussion.

Compression not yet achieved. The compression experiment shows $L_{\text{comp}}/(N \times M) \approx 1.15$, indicating the representation is slightly *larger* than naive storage. This occurs because at small scale ($N, M < 10$), the overhead of carrier storage ($+M$ term) outweighs sparsity benefits. Compression may require larger systems or more aggressive sparsification.

Sparsity paradox. Despite 1144 bond-snapping events, final density remains at 1.0. Analysis reveals that while bonds are continuously snapping, new bonds are simultaneously being seeded at the same rate (when oscillators arrive and are connected to carriers). This suggests the seeding rate or initial bond strength may need rebalancing.

Mitosis not triggered by low D. The mean D before mitosis is 1.097, *above* the threshold of 0.95. This indicates mitosis is being triggered by other mechanisms (possibly the persistence counter or cooldown timing) rather than the intended coherence-based criterion. This is a genuine bug in the implementation that should be addressed.

Global synchrony collapse. The max $R = 1.0$ indicates that at some point, all oscillators became perfectly synchronized. This destroys representational diversity. The coupling strength κ may be too high relative to natural frequency detuning, or the system reaches temporary states of alignment during sparse-oscillator moments.

Weak phase coupling. The phase difference between oscillators sharing carriers (1.606 rad) is not significantly different from random (1.571 rad). This suggests the coupling effect is too weak relative to the oscillator frequencies, or the measurement methodology needs refinement.

Implications for the theory. These failures do not invalidate the theoretical framework but highlight the gap between idealized continuous-time dynamics and discrete simulation. They also suggest that parameter tuning is critical: the current parameters may work well for visualization but not for the claimed emergent properties. Future work should include systematic parameter sweeps and longer simulation horizons.

10.5 Known failure modes

The following failure modes are expected and have been observed in prototype simulations:

- **Global synchrony collapse:** if coupling is too strong relative to detuning, $R(t) \to 1$ and representational diversity is lost [Acebrón et al., 2005, Peron et al., 2025]. Operational fix: reduce κ, increase detuning/noise, or constrain presences.
- **Over-mitosis / carrier proliferation:** aggressive instability thresholds or short persistence windows can create a cascade of mitoses. Operational fix: larger T_{div} and/or a refractory timescale after mitosis.
- **Under-mitosis / failed separation:** if interference does not persist (D_k does not drop below baseline), the system may keep a single mixed carrier. Operational fix: increase detuning between groups or adjust θ_{div} and T_{div}.

10.6 Limitations and scope

The following directions are intentionally out of scope for this paper:

- reinforcement learning and full agentic credit assignment,
- large-scale supervised benchmarks and competitive accuracy claims,
- biological modeling or claims of biological realism,
- hardware-specific implementations beyond the compute scaling discussion.

11 Summary

This revised draft provides a complete, implementable specification for Resonant Compression Systems. The key innovations are:

- **Gaussian tuning strength (no κ):** Coupling strength is geometrically derived from phase alignment via $T_{ik} = \exp(-\Delta\phi^2/\sigma_k)$, eliminating an arbitrary constant and creating "radio dial" lock-on behavior.
- **Adaptive gate width (emergent specialization):** Each carrier's gate width W_k evolves based on coherence—high coherence narrows the gate (specialist), low coherence widens it (generalist).
- **Spectral factorization (not blind mitosis):** Carrier division partitions bonds by *frequency*, producing immediately differentiated offspring rather than stochastic duplicates. Requires multimodal spectral profile.
- **Coherence-weighted metabolism:** Effective intake is $I_k^{\text{eff}} = I_k \cdot D_k$, so carriers capturing incoherent noise starve faster. Coherence becomes a survival criterion.
- **Phase-clustered genesis:** New carriers nucleate only when unbound oscillators show phase coherence ($R_{\mathcal{U}} > \theta$), preventing genesis from scattered noise.
- **Observation interface (not output carriers):** Outputs *measure* resonance without participating in dynamics. One interface supports classification, prediction, and structure discovery without feedback into the engine.
- **Carrier spectral profiles:** Each carrier tracks its spectral center $\bar{\omega}_k$, variance, and multimodality as observable properties for mitosis decisions.

Engineering contributions.

- **Compute requirements:** explicit $O(\text{nnz}(P))$ per-step scaling, sparse simulation loop, GPU-friendly operations [Bell and Garland, 2009].
- **ϵ in finite precision:** reinterpretation as detuning floor, plus implementable strategies: deterministic detuning bounds, integer phase accumulators, and controlled noise/dither.
- **Complete dynamical specification:** all equations implementable in PyTorch with no hidden constants or arbitrary parameters.

A Notation quick reference

Symbol	Meaning
N	number of oscillators
M	number of carriers
M_{active}	number of active carriers (above a threshold)
ϕ_i	phase of oscillator i
ω_i	natural frequency of oscillator i
A_i	amplitude of oscillator i
$z_i = A_i e^{\mathrm{i}\phi_i}$	complex phasor of oscillator i
c_k	complex amplitude of carrier k
W_k	gate width of carrier k (perceptual tolerance)
T_{ik}	tuning strength: oscillator i to carrier k (Eq. (9))
P_{ik}	presence (bond strength) of oscillator i on carrier k
$\text{nnz}(P)$	number of nonzeros in the presence matrix
$u_k = \sum_i T_{ik} P_{ik} z_i$	tuning-weighted carrier drive
$g_i = \sum_k T_{ik} P_{ik} c_k$	tuning-weighted back-influence on oscillator i
γ_k, β_k	carrier damping and saturation
$\text{coh}(k)$	carrier coherence score (Eq. (15))
D_k	normalized coherence score vs. random-phase baseline
$\bar{\omega}_k$	carrier spectral center (weighted mean frequency)
Var_k	carrier spectral variance
I_k^{eff}	coherence-weighted intake (Eq. (24))
$R(t)$	global phase order parameter
$R_{\mathcal{U}}$	order parameter among unbound oscillators (Eq. (25))
L_{comp}	description-length proxy (Eq. (49))

References

Juan A. Acebrón, L. L. Bonilla, Conrad J. Pérez Vicente, Félix Ritort, and Renato Spigler. The Kuramoto model: A simple paradigm for synchronization phenomena. *Reviews of Modern Physics*, 77(1):137–185, 2005. doi: 10.1103/RevModPhys.77.137.

Armen Aghajanyan, Luke Zettlemoyer, and Sonal Gupta. Intrinsic dimensionality explains the effectiveness of language model fine-tuning. *arXiv preprint arXiv:2012.13255*, 2021.

Nathan Bell and Michael Garland. Implementing sparse matrix-vector multiplication on throughput-oriented processors. *Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis*, pages 1–11, 2009. doi: 10.1145/1654059.1654078.

Ronald N. Bracewell. *The Fourier Transform and Its Applications*. McGraw-Hill, 2nd edition, 1986.

Natalia Caporale and Yang Dan. Spike timing-dependent plasticity: a Hebbian learning rule. *Annual Review of Neuroscience*, 31:25–46, 2008. doi: 10.1146/annurev.neuro.31.060407.125639.

Matteo Carandini and David J. Heeger. Normalization as a canonical neural computation. *Nature Reviews Neuroscience*, 13(1):51–62, 2012. doi: 10.1038/nrn3136.

Beidi Chen, Tri Dao, Eric Winsor, Zhao Song, Atri Rudra, and Christopher Ré. Scatterbrain: Unifying sparse and low-rank attention. In *Advances in Neural Information Processing Systems*, volume 34, 2021.

Hiroaki Daido. Onset of cooperative entrainment in limit-cycle oscillators with uniform all-to-all interactions: bifurcation of the order function. *Physica D: Nonlinear Phenomena*, 91(1–2):24–66, 1996. doi: 10.1016/0167-2789(95)00260-X.

Holger Finger, Marlene Bönstrup, Naveed Chehade, Andre Cimenser, Denis Gallahy, Richard Heasley, Christian Stinus, Peter König, and Andreas K. Engel. Phase synchrony facilitates binding and segmentation of natural images in a coupled neural oscillator network. *Frontiers in Computational Neuroscience*, 7:195, 2013. doi: 10.3389/fncom.2013.00195.

Pascal Fries. A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, 9(10):474–480, 2005. doi: 10.1016/j.tics.2005.08.011.

Nicolas Frémaux and Wulfram Gerstner. Neuromodulated spike-timing-dependent plasticity, and theory of three-factor learning rules. *Frontiers in Neural Circuits*, 9:85, 2016. doi: 10.3389/fncir.2015.00085.

Wulfram Gerstner, Marco Lehmann, Vasiliki Liakoni, Dane Corneil, and Johanni Brea. Eligibility traces and plasticity on behavioral time scales: Experimental support of NeoHebbian three-factor learning rules. *Frontiers in Neural Circuits*, 12:53, 2018. doi: 10.3389/fncir.2018.00053.

Peter Grünwald and Paul Vitányi. Shannon information and Kolmogorov complexity. *arXiv preprint cs/0410002*, 2004.

Peter D. Grünwald. *The Minimum Description Length Principle*. MIT Press, Cambridge, MA, 2007.

Shamik Gupta, Alessandro Campa, and Stefano Ruffo. Kuramoto model of synchronization: Equilibrium and nonequilibrium aspects. *Journal of Statistical Mechanics: Theory and Experiment*, page R08001, 2014. doi: 10.1088/1742-5468/14/08/R08001.

Donald O. Hebb. The organization of behavior: A neuropsychological theory. *Wiley, New York*, 1949.

John J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, 79(8):2554–2558, 1982. doi: 10.1073/pnas.79.8.2554.

John J. Hopfield. Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of the National Academy of Sciences*, 81(10):3088–3092, 1984. doi: 10.1073/pnas.81.10.3088.

Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. LoRA: Low-rank adaptation of large language models. *arXiv preprint arXiv:2106.09685*, 2022.

Jerry Yao-Chieh Hu, Donglin Yang, Pei-Hsuan Wu, and Han Liu. On sparse modern Hopfield model. In *Advances in Neural Information Processing Systems*, volume 36, 2023.

Yoshiki Kuramoto. *Chemical Oscillations, Waves, and Turbulence*, volume 19 of *Springer Series in Synergetics*. Springer-Verlag, Berlin, 1984. doi: 10.1007/978-3-642-69689-3.

Michael S. Lewicki and Terrence J. Sejnowski. Learning overcomplete representations. *Neural Computation*, 12(2):337–365, 2000. doi: 10.1162/089976600300015826.

Guozhong Li, Lei Chen, and Yulan Wang. A kuramoto model for graph neural networks. *arXiv preprint arXiv:2311.03260*, 2023.

Henry Markram, Wulfram Gerstner, and Per Jesper Sjöström. Spike-timing-dependent plasticity: a comprehensive overview. *Frontiers in Synaptic Neuroscience*, 4:2, 2012. doi: 10.3389/fnsyn.2012.00002.

Takeru Miyato, Sora Kobayashi, Masanori Koyama, and Yitong Zhong. Artificial Kuramoto oscillatory neurons. In *International Conference on Learning Representations (ICLR)*, 2025.

Bruno A. Olshausen and David J. Field. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381(6583):607–609, 1996. doi: 10.1038/381607a0.

Thomas Peron, Francisco A. Rodrigues, and Jürgen Kurths. Extreme synchronization transitions in coupled oscillators. *Nature Communications*, 16:4853, 2025.

Hubert Ramsauer, Bernhard Schäfl, Johannes Lehner, Philipp Seidl, Michael Widrich, Thomas Adler, Lukas Gruber, Markus Holzleitner, Milena Pavlović, Geir Kjetil Sandve, Victor Greiff, David Kreil, Michael Kopp, Günter Klambauer, Johannes Brandstetter, and Sepp Hochreiter. Hopfield networks is all you need. *arXiv preprint arXiv:2008.02217*, 2020.

Jorma Rissanen. Modeling by shortest data description. *Automatica*, 14(5):465–471, 1978. doi: 10.1016/0005-1098(78)90005-5.

Odelia Schwartz and Eero P. Simoncelli. Natural signal statistics and sensory gain control. *Nature Neuroscience*, 4(8):819–825, 2001. doi: 10.1038/90526.

Claude E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, 1948. doi: 10.1002/j.1538-7305.1948.tb01338.x.

Wolf Singer and Charles M. Gray. Visual feature integration and the temporal correlation hypothesis. *Annual Review of Neuroscience*, 18:555–586, 1995. doi: 10.1146/annurev.ne.18.030195.003011.

Steven H. Strogatz. From Kuramoto to Crawford: exploring the onset of synchronization in populations of coupled oscillators. *Physica D: Nonlinear Phenomena*, 143(1–4):1–20, 2000. doi: 10.1016/S0167-2789(00)00094-4.

Saul Santos Yao and Elias N. Santos. Modern Hopfield networks with continuous-time memories. *arXiv preprint arXiv:2502.10122*, 2025.