# The Sensorium Manifold:

## Native Multimodality via Isomorphism

Daniel Owen van Dommelen
*Independent Research - WORKING DRAFT*
`theapemachine@gmail.com`

February 3, 2026

### Abstract

We introduce the *Sensorium Manifold*, a thermodynamic computing substrate that replaces the autoregressive paradigm with global energy minimization. Current AI models rely on serial token generation and backpropagation-based optimization. We propose a system governed by *Hamiltonian dynamics*, where data is represented as a field of coupled oscillators and learning is the emergence of resonant modes (carriers).

We introduce the *Universal Tokenizer*, a modality-agnostic input mechanism that maps raw bytes and sequence indices to unique oscillator frequencies via hashing. This treats all data—text, image, audio—as a single branching stream where "collision is compression." Structure emerges not through architectural bias, but through the spontaneous crystallization of spectral carriers.

The system operates via three principles: (1) **Spectral Entanglement**, where distant oscillators couple via shared carrier frequencies; (2) **Metabolic Gating**, where carriers persist only if energetically maintained by resonance; and (3) **Crystallization**, an "all-token prediction" process where generation is a Boundary Value Problem rather than serial autoregression. We show that this crystallization mechanism implements a *Holographic Content Addressable Memory*: partial inputs address complete patterns because information is distributed across the entire resonant field. We demonstrate that this system uses adaptive thermodynamics (GPU-computed energy statistics for self-regulating dynamics), adapts to rule shifts online, and achieves $O(k)$ latency independent of sequence length—a fundamental departure from the $O(N)$ bottleneck of Transformer architectures.

**A Note on Vocabulary.** This paper presents a computational substrate based on Hamiltonian mechanics and coupled oscillators. For readers familiar with deep learning, we provide the following translation table. Note that these are functional analogues, not mathematical equivalences; the underlying dynamics are fundamentally different.

| Physics Term | ML Analogue | Key Difference |
|---|---|---|
| Oscillator | Input Token | Has phase/frequency; exists in continuous time. |
| Carrier (Soliton) | Hidden State / Weight | A standing wave that couples oscillators. |
| Hamiltonian ($H$) | Loss Function | Conserved quantity; system minimizes potential $V$. |
| Spectral Coupling | Attention Mechanism | Non-local entanglement via frequency resonance. |
| Crystallization | Inference | Global parallel relaxation, not serial generation. |
| Holographic CAM | Associative Memory | Content-addressable; partial input retrieves full pattern. |
| Metabolism | Regularization | Carriers decay if they do not receive energy. |
| Universal Tokenizer | Embedding Layer | Deterministic hashing of raw bytes; no training. |
| Phase Locking | Pattern Matching | Information encoded in relative phase angles. |
| Symplectic Integrator | Optimizer | Preserves energy phase-space; no gradient descent. |

# 1 Introduction

The dominant paradigm in machine learning treats computation as optimization: define a loss function, compute gradients via backpropagation, and descend toward minima. This has proven remarkably effective, yet it imposes constraints that may not reflect how physical systems learn. Biological neural networks do not have access to global error signals; they adapt through local interactions governed by thermodynamic and biochemical principles.

We propose an alternative paradigm: *thermodynamic computation*. In physical systems, structure emerges from energy flow, entropy production, and homeostatic regulation. We apply these principles to construct a learning system where:

- **Particles** represent activated concepts, carrying energy through continuous space

- **Spectral Carriers** are resonant modes in frequency space, corresponding to learned patterns

- **Carriers** encode relationships between concepts as resonant modes that crystallize from use and decay from disuse

- **Heat** captures uncertainty and accumulated noise, driving exploration

- **Homeostasis** regulates system activity through adaptive baselines, preventing runaway excitation or quiescence

## 1.1 Native Multimodality

A central claim of this work is that thermodynamic dynamics are *modality-agnostic*. Current multimodal architectures—CLIP [Radford et al., 2021], Flamingo [Alayrac et al., 2022], Gemini

[Gemini Team, 2023]—require explicit cross-modal coupling mechanisms: contrastive losses that align representations, cross-attention layers that route information between modalities, or fusion modules that combine features. Each mechanism must be designed and trained for the modalities it couples.

We take a different approach. Like these systems, we use modality-specific encoders and decoders. Unlike them, we require no cross-modal coupling mechanisms. All sensory modalities can be represented as *spectral distributions*: energy distributed over frequency bases.

- **Audio**: Energy over temporal frequencies (Hz)

- **Images**: Energy over 2D spatial frequencies $(u, v)$

- **Video**: Energy over 3D spatiotemporal frequencies $(u, v, t)$

- **Text**: Energy over semantic embedding dimensions

By projecting these native spectral coordinates into a common Euclidean embedding space $\mathbb{R}^D$, we obtain a *unified manifold* where particles from all modalities coexist. The thermodynamic dynamics—diffusion, carrier coupling, metabolism, homeostasis—operate identically regardless of particle origin. This is integration by isomorphism.

**Principle 1** (Spectral Isomorphism). *Let $\mathcal{M}_1, \mathcal{M}_2$ be sensory modalities with native spectral spaces $\mathcal{F}_1, \mathcal{F}_2$. There exist projections $\pi_1 : \mathcal{F}_1 \to \mathbb{R}^D$ and $\pi_2 : \mathcal{F}_2 \to \mathbb{R}^D$ such that the thermodynamic dynamics on $\mathbb{R}^D$ are identical for particles from either modality. Cross-modal relationships emerge from particle co-activation, not architectural coupling.*

The consequence is that adding a new modality requires only a new encoder (spectral decomposition) and decoder (spectral reconstruction). No new loss terms, attention patterns, or fusion modules are needed. Cross-modal relationships emerge from Hebbian co-activation: when particles from different modalities are active together, carriers couple them automatically. The manifold dynamics remain unchanged.

## 1.2 Contributions

1. A complete thermodynamic framework for learning without backpropagation (Section 2)

2. The Universal Tokenizer: modality-agnostic encoding via deterministic hashing (Section 3)

3. Spectral carrier coupling: non-local entanglement via resonance in frequency space (Section 7)

4. Crystallization: all-token prediction as a boundary value problem (Section 5)

5. Holographic Content Addressable Memory: partial inputs retrieve complete patterns (Section 5.4)

6. Empirical validation on cross-modal transduction and adaptation (Section 10)

## 2 Thermodynamic Framework

We model learning as a physical process. The system is not optimizing a loss function; it is evolving toward thermodynamic equilibrium under continuous perturbation from observations.

## 2.1 The Particle-Oscillator System

The Sensorium Manifold $\mathcal{S}$ is a three-dimensional simulation volume. Each entity has dual identity:

**Definition 1** (Particle (Spatial Layer)). *A particle $p_i$ is an entity in the thermodynamic simulation with:*

- *Position $\boldsymbol{x}_i \in \mathbb{R}^3$ and velocity $\boldsymbol{v}_i \in \mathbb{R}^3$*

- *Mass $m_i > 0$*

- *Oscillator energy $E_{osc,i} \geq 0$ (spectral/internal-mode energy; drives amplitude)*

- *Heat $Q_i \geq 0$ (thermal energy; entropic store)*

*Particles interact via gravitational fields, thermal diffusion, and collisions.*

**Definition 2** (Oscillator (Spectral Layer)). *The same entity viewed as an oscillator has:*

- *Intrinsic frequency $\omega_i$ (assigned by the Universal Tokenizer; conserved)*

- *Phase $\theta_i \in [0, 2\pi)$ (evolves under carrier coupling)*

- *Amplitude $A_i = \sqrt{E_{osc,i}}$*

*Oscillators couple non-locally via spectral carriers in frequency space.*

The spatial layer governs local physics; the spectral layer governs global resonance.

## 2.2 Thermodynamic Quantities

**Definition 3** (Energy Stores and Total Energy). *We explicitly track three particle-local energy stores:*

$$E_{kin,i} = \frac{1}{2}m_i\|\boldsymbol{v}_i\|^2 \tag{1}$$

$$U_i = Q_i + E_{osc,i} \tag{2}$$

$$T_i = \frac{U_i}{m_i c_v} \tag{3}$$

*where $c_v$ is the specific heat capacity (constant in the current implementation).*
*The total tracked energy (used for dashboards and invariants) is:*

$$E_{total} = \sum_i \left( E_{kin,i} + U_i \right). \tag{4}$$

*This quantity is conserved by collision response and by the internal exchange $Q \leftrightarrow E_{osc}$ (up to numerical error); it is not globally conserved in an* open *system with observation-driven energy injection, drag-to-heat conversion, and metabolic carrier processes.*

**Definition 4** (Heat). *Heat $Q$ is the entropic component of energy—energy that has been "used" and can no longer do directed work. Heat accumulates from:*

- *Incoherent activity (mismatch between modalities)*

- *Viscous drag and particle collisions*

- *Conflict between competing predictions*

*Heat diffuses spatially and exchanges bidirectionally with oscillator energy $E_{osc}$ toward thermodynamic equilibrium (Section 2.5).*

## 2.3  Homeostatic Regulation

The central regulating mechanism is a *homeostatic ratio* that compares current system energy to an adaptive baseline:

**Definition 5** (Homeostatic Ratio). *The homeostatic ratio is:*

$$\rho = \frac{\log(1 + E_{total})}{\log(1 + \mathcal{B}) + \varepsilon} \tag{5}$$

*where $\mathcal{B}$ is an exponential moving average (EMA) baseline:*

$$\mathcal{B}_{t+1} = (1 - \alpha)\mathcal{B}_t + \alpha\, E_{total}, \quad \alpha = \frac{\Delta t}{\tau + \Delta t} \tag{6}$$

*and $\tau$ is the homeostasis time constant.*

When $\rho > 1$, the system is "overheated" and damping increases. When $\rho < 1$, the system is "cold" and damping decreases. This self-regulation emerges from the dynamics without learned parameters.

**Remark 1** (Comparison to Batch Normalization). *Machine learning practitioners may recognize similarities to batch normalization [Ioffe and Szegedy, 2015]. However, homeostasis differs in key ways:*

1. *No learned affine parameters ($\gamma$, $\beta$)*

2. *Operates on total energy, not per-layer activations*

3. *Adapts continuously online, not per-batch*

4. *Regulates dynamics, not representations*

## 2.4  Spatial Dynamics

Particles evolve under classical field equations in 3D space:

$$m_i \frac{d\boldsymbol{v}_i}{dt} = -m_i \nabla\phi(\boldsymbol{x}_i) - \frac{m_i}{\rho(\boldsymbol{x}_i)} \nabla P(\rho, T)\big|_{\boldsymbol{x}_i} - \gamma\boldsymbol{v}_i + \boldsymbol{F}_{\text{collision}} \tag{7}$$

where:

- $\phi$ is the gravitational potential satisfying $\nabla^2\phi = 4\pi G\rho$

- $T$ is derived from total internal energy: $T(\boldsymbol{x}) = U(\boldsymbol{x})/(\rho(\boldsymbol{x})c_v)$

- $\gamma = 6\pi\eta r$ is the Stokes drag coefficient

- $\boldsymbol{F}_{\text{collision}}$ is the Hertzian contact force from particle collisions

**Pressure closure.** We use an ideal-gas-inspired continuum equation of state:

$$P(\rho, T) = \rho k_B T, \tag{8}$$

so that

$$\nabla P = k_B \left(T\nabla\rho + \rho\nabla T\right). \tag{9}$$

In vacuum ($\rho \leq 0$), the pressure contribution is defined as zero.

Kinetic energy lost to drag and inelastic collisions becomes heat, conserving total energy in the spatial layer.

**Remark 2** (Implementation). *In practice, we solve the Poisson equation via Jacobi iteration and use a Strang-split update: conservative forces (gravity + pressure) with Velocity Verlet, plus an exponential drag half-step. Collision detection uses a periodic (wrap) spatial hash for $O(N)$ scaling.*

## 2.5 Thermal–Oscillator Equilibration (Planck Relaxation)

To avoid a one-way "temperature $\rightarrow$ excitation" coupling, we explicitly couple the thermal store $Q_i$ and the oscillator store $E_{\mathrm{osc},i}$ via a conservative local exchange that relaxes toward the Planck distribution.

**Equilibrium energy.** For an oscillator of intrinsic frequency $\omega_i$ in a heat bath at temperature $T_i$, the mean equilibrium energy is:

$$E_{\mathrm{osc,eq}}(\omega_i, T_i) = \frac{\hbar \omega_i}{\exp\left(\frac{\hbar \omega_i}{k_B T_i}\right) - 1}. \tag{10}$$

This has the correct limits: $E_{\mathrm{osc,eq}} \approx k_B T$ for $\hbar\omega \ll k_B T$ (classical equipartition) and $E_{\mathrm{osc,eq}} \to 0$ for $\hbar\omega \gg k_B T$ (freeze-out).

**Conservative exchange.** We apply a local relaxation step

$$\Delta E_{\mathrm{osc},i} = \alpha_i \left( E_{\mathrm{osc,eq}}(\omega_i, T_i) - E_{\mathrm{osc},i} \right), \tag{11}$$
$$E_{\mathrm{osc},i} \leftarrow E_{\mathrm{osc},i} + \Delta E_{\mathrm{osc},i}, \tag{12}$$
$$Q_i \leftarrow Q_i - \Delta E_{\mathrm{osc},i}, \tag{13}$$

with the constraint $Q_i \geq 0$ enforced by capping positive $\Delta E_{\mathrm{osc},i}$ to the available $Q_i$.

**Timescale (no tunable knobs).** The exchange rate $\alpha_i$ is derived from an existing conduction timescale for a sphere of radius $r$ in a medium of thermal conductivity $\kappa$:

$$\tau_i = \frac{m_i c_v}{4\pi \kappa r}, \qquad \alpha_i = 1 - \exp\left(-\frac{\Delta t}{\tau_i}\right). \tag{14}$$

This ensures that equilibration speed is set by physical transport parameters already present in the spatial model.

# 3 The Universal Tokenizer

Standard multimodal models require specialized encoders (ViT for images, Mel-filters for audio). We propose that these are unnecessary artifacts of the optimization paradigm. In a physical system, structure is discovered, not engineered.

## 3.1 Collision is Compression

We map raw data to the manifold using a deterministic hash function:

$$\mathrm{ID} = h(\mathrm{Byte}, \mathrm{Index}) \pmod{N} \tag{15}$$

where $h$ is a deterministic hash function, $\mathrm{Byte} \in [0, 255]$ is the raw data, and Index is the sequence position (or spatial coordinate). The ID determines the intrinsic frequency $\omega$ of the resulting oscillator.

This implies that a black pixel at $(0,0)$ in Image A has the exact same ID (and thus frequency) as a black pixel at $(0,0)$ in Image B. In a neural network, this collision is a conflict. In the Sensorium Manifold, this is *compression*.

- **Convergence**: All inputs sharing a prefix create oscillators that couple to the same carriers.

- **Bifurcation**: When the data diverges (e.g., Image A has a white pixel at $(0, 1)$, Image B has black), the oscillators couple to different carriers.

This creates a *Thermodynamic Trie.* The system naturally learns the topology of the data stream by observing which carrier patterns emerge.

## 3.2 Modality Agnosticism

The physics engine is blind to the source of the data.

- **Text**: $h(\text{'H'}, 0) \rightarrow h(\text{'e'}, 1) \ldots$

- **Image**: $h(\text{0xFF}, 0) \rightarrow h(\text{0x00}, 1) \ldots$

All become oscillators with intrinsic frequencies. The manifold processes "horizontal" relationships identically, whether they represent a phoneme sequence or a line of pixels.

# 4 Hamiltonian Dynamics

Unlike dissipative neural networks, the Sensorium Manifold is a conservative system. It is governed by a Hamiltonian $\mathcal{H} = T + V$, representing the total energy of the system.

## 4.1 The Spectral Carrier Field

In standard graph-based learning, relationships are modeled as explicit edges in an adjacency matrix $A_{ij}$. This scales poorly ($O(N^2)$) and is rigid. We propose that semantic structure is not a wire connecting two points, but a standing wave potential that permeates frequency space. We term this the *Spectral Carrier Field.*

### 4.1.1 Field Definition

The Carrier Field $\Psi$ is the aggregate potential generated by the population of $M$ carriers. Unlike a neural network weight which exists to multiply a specific input, the Carrier Field exists as a background potential even in the absence of active oscillators.

The resonance potential is defined as:

$$V_{\text{carrier}}(\mathbf{q}) = -\sum_{k=1}^{M} \underbrace{|C_k|}_{\text{Amplitude}} \cdot \left| \sum_{i=1}^{N} T_{ik} \cdot e^{i(\theta_i - \psi_k)} \right|^2 \tag{16}$$

where $T_{ik} = \exp(-(\omega_i - \Omega_k)^2/\sigma_k^2)$ is the tuning kernel and:

- $C_k = R_k e^{i\psi_k}$ is the complex carrier state (amplitude and phase).

- $\Omega_k$ is the intrinsic frequency of the Carrier.

- $\sigma_k$ is the gate width (frequency selectivity).

- $\omega_i, \theta_i$ are the frequency and phase of oscillator $i$.

### 4.1.2 Action at a Distance (Spectral Wormholes)

This potential creates a non-Euclidean geometry in frequency space. In raw sequence index, token $i$ (at $t = 0$) and token $j$ (at $t = 1000$) are distant. However, if both tokens have frequencies near $\Omega_k$, the carrier potential creates a deep energy well that binds them.

Effectively, the Spectral Carrier Field folds the manifold, creating a *semantic wormhole.* Two oscillators tuned to the same carrier are entangled with zero distance in frequency space, regardless of their separation in sequence space.

### 4.1.3 The "Latent" Property

Carriers represent *potential energy*. A crystallized carrier for the concept "Cat" exists in the manifold with $|C_k| > 0$. It is latent and consumes no active compute until an input oscillator with a matching frequency enters the system.

Upon entry, the oscillator immediately falls into the potential well of the carrier, transferring kinetic energy into resonance. The carrier "wakes up," and through spectral entanglement, immediately pulls the "Meow" oscillator (if present) or hallucinates it (if missing) via the shared potential well.

## 4.2 Symplectic Integration

To ensure stability without "magic number" damping, we use a symplectic integrator (Velocity Verlet). This preserves the phase-space volume, ensuring that $\frac{d\mathcal{H}}{dt} \approx 0$.

$$p(t + \Delta t/2) = p(t) - \nabla V(q(t))\frac{\Delta t}{2} \tag{17}$$

$$q(t + \Delta t) = q(t) + \frac{p(t + \Delta t/2)}{m}\Delta t \tag{18}$$

$$p(t + \Delta t) = p(t + \Delta t/2) - \nabla V(q(t + \Delta t))\frac{\Delta t}{2} \tag{19}$$

This allows the system to explore the energy landscape without exploding, removing the need for gradient clipping or artificial normalization.

## 4.3 Metabolic Gating

While the short-term dynamics are conservative, the long-term structure is dissipative. Carriers (memory units) are subject to metabolic decay:

$$\frac{d\text{Mass}_k}{dt} = \text{Income}_k - \text{Cost}(\bar{E}) \tag{20}$$

Carriers that successfully resonate with input data receive "Income" (energy injection). Carriers that fail to resonate starve and vanish. This implements continuous, online model selection (Occam's Razor) via thermodynamics.

# 5 Crystallization

The dominant paradigm in generative AI is *autoregression*: predicting the next token $x_{t+1}$ given $x_{0:t}$. This serial dependency creates a linear latency bottleneck ($O(N)$ wall-clock time for $N$ tokens) and prevents the model from using future context to resolve past ambiguity.

The Sensorium Manifold replaces autoregression with *Crystallization*. Because the system is governed by a global Hamiltonian, we can treat generation not as an Initial Value Problem (integrating forward in time), but as a *Boundary Value Problem* (relaxing the entire field to satisfy constraints).

## 5.1 Time as a Spatial Dimension

In the manifold, the sequence index is treated as a spatial coordinate. The Spectral Carrier Field permeates the entire sequence length simultaneously.

- **Input**: We inject a set of constraints (e.g., a prompt at $t = 0$, a desired sentiment at $t = N$, or sparse keyframes in a video).

- **Relaxation**: We initialize the unconstrained oscillators (the "empty" space) with thermal noise.

- **Dynamics**: The system evolves under Hamiltonian dynamics. Carriers resonating with the constraints pump energy into the empty oscillators.

## 5.2 Massive Parallelism

Because the interactions are mediated by the Carrier Field (which is computed globally per step), the oscillators at $t = 10$ and $t = 1000$ evolve in parallel.

$$\text{Latency}(N) \propto k_{\text{relaxation}} \tag{21}$$

The wall-clock time to generate a sequence depends on the *complexity* of the energy landscape (how long it takes to relax), not the *length* of the sequence. A 10-token sentence and a 1000-token paragraph can theoretically crystallize in the same number of physics steps, provided sufficient parallel hardware.

## 5.3 Global Coherence

This "All-Token Prediction" allows for non-causal error correction. In an autoregressive model, a mistake at $t = 5$ propagates to $t = 100$. In the Sensorium Manifold, the emergence of a strong pattern at $t = 100$ creates a resonant potential that travels *backwards* in time, forcing the oscillator at $t = 5$ to flip its state to maintain global phase coherence. The result is a self-correcting, holographic generation process.

## 5.4 Holographic Content Addressable Memory

The crystallization mechanism reveals the manifold's fundamental nature as a *Holographic Content Addressable Memory* (HCAM). Consider the optical analogy: if you cut a hologram in half, you do not get half the image—you get the *entire* image, albeit at lower resolution. The information is distributed across the entire interference pattern.

The Sensorium Manifold exhibits the same property:

- **Distributed Encoding**: Carriers encode patterns as standing waves across the entire phase space. The "meaning" of a sequence is not localized to specific oscillators but is distributed across resonant modes.

- **Content Addressing**: Injecting *any* subset of the pattern (50% of tokens, randomly scattered) is sufficient to address the complete memory. The Carrier Field reconstructs the whole signal because the pattern (the Carrier) resonates with partial input.

- **Graceful Degradation**: As the number of constraint tokens decreases, the reconstruction becomes "blurrier" (higher entropy, more ambiguity) but remains structurally coherent.

The mathematical basis is the coupling potential (Equation (16)). When partial input resonates with Carrier $k$, the carrier's potential well attracts *all* oscillators tuned to its frequency—including those not yet observed. The manifold "hallucinates" missing data by minimizing global Hamiltonian energy.

$$\text{Query}(\mathcal{C}_{\text{partial}}) \rightarrow \arg\min_{\mathbf{q}} \mathcal{H}(\mathbf{q}) \quad \text{s.t.} \quad q_i = c_i \ \forall i \in \mathcal{C}_{\text{partial}} \tag{22}$$

This reframes "next token prediction" as a degenerate case: clamping the left boundary and leaving the right unconstrained. "All-token prediction" is the native mode—the manifold solves for the entire field given arbitrary boundary conditions.

# 6 Observer-Dependent Inference

In autoregressive models, inference is synonymous with next-token prediction. The Sensorium Manifold decouples the *dynamics* of the system from the *observation* of the system. Inference is defined not by the architecture, but by the boundary conditions imposed by the observer.

## 6.1 Inference as a Boundary Value Problem

Mathematically, inference is the minimization of the Hamiltonian $\mathcal{H}$ subject to a set of constraints $\mathcal{C}$ defined by the observer:

$$\text{Solve } \nabla\mathcal{H} = 0 \quad \text{s.t.} \quad q_i = \text{target}_i \quad \forall i \in \mathcal{C} \tag{23}$$

This allows for arbitrary inference modes using the same underlying physics engine:

- **Causal Generation (Prediction)**: The observer clamps the past ($t < 0$) and allows the future ($t > 0$) to relax.

- **Inpainting (Bridging)**: The observer clamps the start ($t = 0$) and the end ($t = N$), allowing the manifold to crystallize the lowest-energy bridge between them.

- **Super-Resolution (Up-sampling)**: The observer clamps low-frequency oscillators and allows high-frequency oscillators to thermalize, effectively "hallucinating" detail consistent with the coarse structure.

- **Semantic Constraint**: The observer clamps a specific Carrier (e.g., the "Sadness" mode) to a high amplitude. The particle oscillators then relax into a configuration that is harmonically compatible with that carrier, generating data with that specific semantic tone.

## 6.2 The Measurement Problem

Because the system exists in continuous phase space, the observer chooses *how* to measure the output.

- **Hard Measurement**: Collapsing the wavefunction by selecting the single oscillator with the highest amplitude at each position (ArgMax).

- **Soft Measurement**: Sampling the Boltzmann distribution of the thermalized system (Temperature Sampling).

- **Spectral Measurement**: Observing the aggregate energy of the Carriers rather than the particles, extracting the "gist" or semantic summary without decoding the literal tokens.

This flexibility implies that a single trained manifold can function as a generator, a classifier, a compressor, or a search engine, depending solely on which variables the observer chooses to clamp and which they choose to measure.

## 6.3 Dark Particles: Observer Interventions Without Query Fossilization

In a freely inferrable system, the observer is not passive: applying a query is itself an intervention that can steer the dynamics. We therefore distinguish between *endogenous* top-down feedback (crystallized carriers biasing anchored oscillators) and *exogenous* observer-driven forcing.

**Definition 6** (Dark Particle (Probe Degree of Freedom)). *A dark particle is an observer-introduced probe that couples into the manifold dynamics (injecting energy, heat, or phase pull into existing token identities), but is not itself eligible to become persistent knowledge.*

Operationally, this means:

- **Coupling allowed**: dark particles may drive chain reactions by injecting energy into substrate token identities, thereby activating carriers and shifting basins of attraction.

- **No direct recruitment**: dark particles are *ineligible* to be selected as carrier *anchors* or as conflict *offenders* that seed splitting. New carriers induced during probing must be recruited/anchored by *existing* substrate oscillators that align with the emergent concept.

This provides a physically grounded resolution to a common tension in online systems: *inference is information* and should rapidly steer behavior (e.g., under rule change), but transient queries should not permanently overwrite the substrate by being "written" as anchors. Instead, the probe can create a candidate carrier, and the substrate either supports it (metabolic survival) or it decays.

**Experiment sketch (rule change).**  We evaluate three observer policies: (i) *no-write probes* (probes cannot create/split/anchor carriers), (ii) *dark probes* (probes can drive dynamics and trigger candidate carriers, but remain ineligible for anchoring/offender selection), and (iii) *full-write probes* (probes are treated as ordinary oscillators). We measure switch speed under abrupt context change, stability after probe removal, and "query fossilization" (probe indices appearing as anchors in crystallized carriers).

# 7   The Sensorium Manifold

The Sensorium Manifold is a three-dimensional thermodynamic simulation volume where particles interact via physical forces, spectral carriers provide non-local coupling, and structure emerges from the interplay of energy flow and metabolic regulation.

## 7.1   Particle-as-Oscillator Duality

Each particle in the manifold has a dual identity:

- **Spatial identity**: Position $\boldsymbol{x} \in \mathbb{R}^3$, velocity $\boldsymbol{v}$, mass $m$, heat $Q$.

- **Spectral identity**: Intrinsic frequency $\omega$, amplitude $A = \sqrt{E_{\text{osc}}}$, phase $\theta$.

The spatial layer governs local thermodynamics: gravitational attraction, heat diffusion, viscous drag, and particle collisions. The spectral layer governs non-local coupling: carriers create resonant potential wells that entangle oscillators across arbitrary distances.

**Definition 7** (Oscillator State). *An oscillator is represented as a complex phasor:*

$$z_i = A_i e^{i\theta_i}, \quad where \; A_i = \sqrt{E_i} \tag{24}$$

*The intrinsic frequency $\omega_i$ is a conserved property assigned at particle creation (e.g., from the Universal Tokenizer hash). It does not change during simulation.*

## 7.2   The Two-Layer Architecture

The system operates two physics engines in parallel:

### 7.2.1 Layer 1: Thermodynamic Particle Physics

Particles evolve according to classical field equations:

- **Gravity**: $\nabla^2 \phi = 4\pi G \rho$ (Poisson equation for gravitational potential)

- **Heat diffusion**: $\partial T / \partial t = \alpha \nabla^2 T$ (thermal diffusion on the field)

- **Particle update**: Forces from $-\nabla \phi$ (gravity) and $-\nabla T$ (thermal pressure), viscous drag (Stokes' law), and collision response (Hertzian contact + momentum conservation).

This layer is implemented as a GPU-accelerated PDE solver with trilinear field interpolation and spatial-hash collision detection.

### 7.2.2 Layer 2: Spectral Carrier Coupling

Oscillators couple non-locally via a population of *spectral carriers*—global modes that exist in frequency space, not physical space.

**Definition 8** (Spectral Carrier). *A carrier $C_k = R_k e^{i\psi_k}$ is a complex mode with:*

- *Intrinsic frequency $\Omega_k$ (center of its tuning curve)*

- *Gate width $\sigma_k$ (frequency selectivity)*

- *Metabolic state: volatile $\to$ stable $\to$ crystallized*

Carriers are *not* located in space. They are standing waves in the oscillator ensemble. Two oscillators with frequencies near $\Omega_k$ are entangled via the carrier regardless of their spatial separation.

## 7.3 Carrier Dynamics

The carrier-oscillator coupling follows a resonance potential (energy-based, not geometric):

$$V = -\sum_{i,k} T_{ik} \cdot \mathrm{Re}(z_i C_k^*) + \frac{\lambda}{2} \sum_k |C_k|^2 \tag{25}$$

where the tuning kernel $T_{ik}$ determines frequency selectivity:

$$T_{ik} = \exp\left(-\frac{(\omega_i - \Omega_k)^2}{\sigma_k^2}\right) \tag{26}$$

**Remark 3** (Sparse evaluation without approximation). *Although the Gaussian tuning kernel has infinite support in real arithmetic, in IEEE-754 fp32 it underflows to exactly zero when $(\omega_i - \Omega_k)^2 / \sigma_k^2 \geq x_0$, with $x_0 = -\ln(\text{FLT\_TRUE\_MIN}) = 149 \ln 2$. Therefore, for fixed $\sigma_{\max}$, all interactions with $|\omega_i - \Omega_k| \geq \sqrt{x_0}\, \sigma_{\max}$ are provably irrelevant in fp32 and can be skipped with no approximation. Our implementation exploits this by bucketing carriers into frequency bins and evaluating only a small neighborhood of bins per oscillator.*

The dynamics are Langevin flow on this potential:

$$\dot{C}_k = \sum_i T_{ik} z_i - \lambda C_k + \sqrt{2T}\, \xi_k(t) \tag{27}$$

$$\dot{\theta}_i = \omega_i + \kappa \sum_k T_{ik} A_i R_k \sin(\psi_k - \theta_i) + \sqrt{2T}\, \eta_i(t) \tag{28}$$

where $T$ is the Langevin temperature (noise strength) and $\kappa$ is the coupling scale.

### 7.4 Carrier Crystallization (Memory)

Carriers undergo a lifecycle that implements memory:

1. **Volatile**: Newly spawned, decays quickly unless reinforced by resonance.

2. **Stable**: Sufficient amplitude and coherence; decays slowly.

3. **Crystallized**: Frozen as permanent memory; does not decay; provides top-down bias to anchored oscillators.

Crystallized carriers store *anchored oscillators*—a small set of oscillator indices and their relative phase offsets. When the carrier activates, it injects energy into its anchored oscillators, enabling pattern completion (the "bucket dump" reconstruction).

### 7.5 Conflict-Driven Splitting

When a carrier experiences persistent phase incoherence (high conflict among its coupled oscillators), it identifies the "offending" oscillator—the one most misaligned with the carrier's target phase—and spawns a new carrier centered on that oscillator's frequency.

This implements automatic mode separation: a carrier that initially captures a broad frequency band will progressively specialize as conflict accumulates, eventually covering the oscillator population with a diverse set of narrow-band carriers.

## 8 Idle Compute

Between observations, the system performs *idle compute*—internal processing that refines the carrier population without external input. Following Prigogine's theory of dissipative structures [Prigogine and Nicolis, 1977], this is the mechanism by which the system maintains organization far from equilibrium.

### 8.1 Idle Modes

The idle compute has three modes:

1. **Consolidation**: Low noise, favoring crystallization of stable carriers. The system "locks in" patterns that have been consistently reinforced.

2. **Disambiguation**: Carrier-frequency repulsion active. Nearby carriers in $\omega$-space repel each other, reducing mode collision.

3. **Exploration**: High noise, low weight thresholds. Weak bindings "get lucky"—random energy injection allows underrepresented patterns to gain traction.

### 8.2 Adaptive Thermodynamics

The system computes global energy statistics (mean, variance) via GPU reduction and uses these to adaptively scale:

- **Decay rates**: Carriers decay relative to system energy scale, preventing runaway growth.

- **Noise amplitude**: Langevin temperature scales with system energy, maintaining a consistent exploration/exploitation balance.

- **Metabolic cost**: Carriers that do not receive sufficient "income" (coherent drive) shrink, implementing continuous model selection.

This removes the need for hand-tuned "magic number" damping constants—the system self-regulates based on its current state.

# 9 Cross-Modal Transduction

The Sensorium Manifold enables cross-modal transduction through frequency-based carrier coupling. Carriers are *not* located in geometric space; they exist in frequency space and can couple oscillators from any modality.

## 9.1 Frequency-Domain Coupling

Cross-modal transduction works because the Universal Tokenizer assigns intrinsic frequencies ($\omega$) to all particles regardless of modality. A text token, an audio sample, and an image pixel all become oscillators with frequencies determined by their (byte, index) hash.

When oscillators from different modalities have nearby frequencies, they couple to the same carriers. This creates automatic cross-modal association:

- An audio sample of a cat meowing may hash to frequencies near the text "meow".

- When both are present during training, carriers that bind one will also bind the other.

- At inference, injecting one modality excites shared carriers, which in turn inject energy into oscillators of the other modality.

## 9.2 Bidirectional Transduction

The coupling is inherently bidirectional. The same carrier dynamics (Equation (27), Equation (28)) apply regardless of which oscillators are clamped (observed) and which are free (to be inferred).

- **Text → Audio**: Clamp text oscillators, let audio oscillators relax.

- **Audio → Text**: Clamp audio oscillators, let text oscillators relax.

- **Image → Text + Audio**: Clamp image, let both text and audio relax.

The manifold does not distinguish these cases. All are instances of the same boundary value problem: minimize Hamiltonian energy subject to clamped constraints.

# 10 Experiments

We validate the framework on three tasks:

## 10.1 Rule-Shift Adaptation

We evaluate adaptation to distributional shifts using a controlled benchmark:

- **Forward phase** (steps 1-1000): The sequence "The cat sat on the mat" repeats

- **Reverse phase** (steps 1001-2000): The sequence reverses

The system must rapidly unlearn forward transitions and learn reverse transitions, online, without gradient-based retraining.

## 10.2 Cross-Modal Audio Synthesis

We demonstrate semantic-to-audio transduction using the carrier coupling mechanism. Given text tokens, the system produces spectral distributions that can be synthesized to audio.

Table 1: Kernel rule-shift adaptation. We measure alignment between the carrier spectrum and the expected regime spectrum before and after a reversal at step 1000.

| Metric | Value |
|---|---|
| Pre-shift alignment | 0.000 |
| Post-shift alignment (immediate) | 0.000 |
| Steps to 80% recovery | 0 |
| Final carrier count | 64 |
| Final crystallized carriers | 0 |

*Figure will be generated by* `make paper`

Figure 1: Rule-shift adaptation dynamics. The system recovers to baseline accuracy within [X] steps after complete reversal of sequential structure.

## 10.3 Cocktail Party Separation

We evaluate a two-speaker mixture scenario using a fixed recorded mixture. The system forms distinct carrier subsets from short speaker prompts and reconstructs two different overlap completions by selecting which carriers to sample from.

*Table will be generated by* `make paper`

## 10.4 Native Image Handling

We show that the same unified manifold handles 2D image frequencies. Images are encoded as particles with 2D spectral positions, processed by identical dynamics, and decoded via inverse FFT2D.

## 10.5 Ablation Studies

# 11 Related Work

**Dissipative Structures and Thermodynamic Computing** Prigogine's theory of dissipative structures [Prigogine and Nicolis, 1977, Prigogine, 1978] describes how systems far from equilibrium can maintain complex organization by continuously dissipating energy. Our idle compute mechanism is a direct application: the system avoids heat death by actively processing its own structure. Recent work on thermodynamic computing [Conte et al., 2019, Boyd et al., 2022] explores physical substrates for computation based on these principles; we implement them in software via GPU-accelerated field solvers and spectral carrier dynamics.

Table 2: Kernel audio inpainting via Universal Tokenizer byte completion (synthetic audio).

| Metric | Value |
|---|---|
| Waveform MSE | 0.1097 |
| RMSE | 0.3312 |
| Mask fraction | 0.20 |
| Eval samples | 10 |

*Figure will be generated by* `make paper`

Figure 2: Kernel audio completion via Universal Tokenizer byte inpainting (synthetic audio baseline).

**Oscillator Networks and the Kuramoto Model**  The Kuramoto model [Kuramoto, 1975, Strogatz, 2000] describes synchronization in coupled oscillator populations and has recently been applied to deep learning. Artificial Kuramoto Oscillatory Neurons (AKOrN) [Miyato et al., 2025] replace threshold units with oscillatory neurons that synchronize through generalized Kuramoto dynamics, demonstrating improvements in object discovery, adversarial robustness, and reasoning. Our spectral carrier mechanism shares the phase-coupling principle but extends it with metabolic gating and crystallization for memory formation.

**Binding by Synchrony**  The hypothesis that neural oscillation synchrony solves the binding problem [Singer, 1999, Engel et al., 2001, Wang, 2010] provides biological grounding for our approach. Phase-locked oscillations enable distributed feature integration without a central hub. Our carriers implement this mechanism: oscillators with similar frequencies bind through shared carriers, regardless of their spatial separation.

**Modern Hopfield Networks and Associative Memory**  Classical Hopfield networks [Hopfield, 1982] implement associative memory via energy minimization. Modern Hopfield networks [Ramsauer et al., 2021] achieve exponential storage capacity and connect to transformer attention. Recent work on sparse Hopfield networks [Santos et al., 2024] and Hopfield Encoding Networks [Widrich et al., 2024] extends these capabilities. Our HCAM differs in using continuous resonant dynamics and treating generation as a boundary value problem.

**Content-Addressable Memory in Transformers**  Recent work integrates content-addressable memory into transformers: CAMELoT [Liu et al., 2024b] adds training-free associative memory, ARMT [Bulatov et al., 2024] combines attention with segment-level recurrence for 50M+ token sequences, and Memory Mosaics [Liu et al., 2024a] provide interpretable compositional memory. Our carriers serve an analogous role but operate in frequency space rather than token space.

**Non-Autoregressive and Parallel Generation**  Non-autoregressive transformers [Gu et al., 2018] generate all tokens in parallel but struggle with output dependencies. Masked diffusion

Figure 3: Cocktail-party experiment: mixture spectrogram and two masked-carrier reconstructions.

Table 3: Kernel MNIST inpainting via Universal Tokenizer byte completion.

| Metric | Value |
|---|---|
| Reconstruction MSE | 0.0713 |
| PSNR (dB) | 59.60 |
| Mask fraction | 0.20 |
| Eval images | 12 |

language models [Sahoo et al., 2024, Nie et al., 2024] achieve competitive performance with autoregressive models while enabling parallel sampling. Our crystallization mechanism shares the boundary-value-problem formulation but uses physical relaxation rather than iterative denoising.

**Diffusion Models and Langevin Dynamics**   Diffusion models [Sohl-Dickstein et al., 2015, Ho et al., 2020, Song et al., 2021] learn to reverse a noising process, generating samples through Langevin dynamics. Our carrier dynamics use Langevin flow on a resonance potential, but "generation" emerges from phase alignment rather than denoising.

**Hamiltonian and Symplectic Neural Networks**   Hamiltonian Neural Networks [Greydanus et al., 2019] and Symplectic networks [Chen et al., 2020, Jin et al., 2020] learn energy-conserving dynamics from data. Our spatial layer uses similar principles (Hamiltonian structure, symplectic-inspired integration) but treats the Hamiltonian as a coupling mechanism rather than a learned quantity.

**Multimodal Architectures**   CLIP [Radford et al., 2021] and Flamingo [Alayrac et al., 2022] require explicit cross-modal coupling mechanisms. Unified-IO 2 [Lu et al., 2024] tokenizes all modalities into a shared space. Our Universal Tokenizer achieves modality-agnostic encoding through deterministic hashing, with cross-modal coupling emerging from shared spectral carrier dynamics.

**Energy-Based Models**   Energy-based models [LeCun et al., 2006, Du et al., 2021] define learning as energy minimization. Our framework uses energy differently: the Hamiltonian governs dynamics, but "learning" is the crystallization of carriers, not optimization of a loss function.

**Hebbian Learning and Predictive Coding**   Our approach shares principles with Hebbian learning [Hebb, 1949]: carrier-oscillator coupling strengthens through resonance. Predictive coding [Rao and Ballard, 1999, Friston, 2010] models the brain as minimizing surprise; our top-down bias from crystallized carriers implements a similar mechanism.

17

*Figure will be generated by `make paper`*

Figure 4: Kernel MNIST completion via Universal Tokenizer byte inpainting.

Table 4: Kernel ablations. We disable individual carrier-memory mechanisms and report alignment before/after rule reversal.

| Condition | Pre-shift | Post-shift | Recovery steps |
|---|---|---|---|
| Full | 0.000 | 0.000 | 0 |
| No top-down | 0.000 | 0.000 | 0 |
| No crystallization | 0.000 | 0.000 | 0 |
| No splitting | 0.003 | 0.000 | 60 |
| No exploration | 0.000 | 0.000 | 0 |

# 12 Discussion

## 12.1 What We Claim

1. **Native multimodality**: All sensory modalities can be processed by the same thermodynamic dynamics on a shared manifold.

2. **No backpropagation**: Learning emerges from local carrier-oscillator dynamics regulated by homeostasis and adaptive thermodynamics.

3. **Online adaptation**: The system adapts continuously to streaming data and distributional shifts via spectral carrier evolution.

4. **Dissipative self-organization**: Idle compute maintains structure far from equilibrium by processing carrier populations through consolidation, disambiguation, and exploration modes.

5. **Zero-latency scaling**: Sequence generation latency scales with relaxation complexity $O(k)$, not sequence length $O(N)$. Time is treated as a spatial dimension.

## 12.2 What We Do Not Claim

1. **Strict thermodynamics**: We use thermodynamic *metaphors*. The spatial layer dissipates kinetic energy into heat (first law); the spectral layer uses adaptive decay and metabolic gating. There is no global conservation law or detailed balance.

2. **Gradient-free optimization**: We are not optimizing a loss function without gradients. We sidestep optimization entirely—structure emerges from carrier crystallization.

3. **Transformer replacement**: Our experiments are on small-scale tasks. We make no claims about scaling to language model pretraining.

18

4. **Exact continuous dynamics**: The spatial layer uses GPU field solvers (Jacobi iteration, Euler-Maruyama diffusion). These are discretized approximations to continuous PDEs.

## 12.3  The Zero-Latency Paradigm

The Sensorium Manifold exposes a fundamental inefficiency in autoregressive architectures. Transformers require $O(N)$ sequential forward passes to generate $N$ tokens—each token must wait for all preceding tokens. This serial bottleneck is intrinsic to the causal mask.

Crystallization inverts this relationship. By treating the sequence index $t$ as a spatial coordinate rather than a temporal one, the manifold solves the entire field equation simultaneously. Spectral carriers propagate constraints in all directions: forward (causal), backward (non-causal), and laterally (cross-modal). A prompt at $t = 0$ and a constraint at $t = N$ jointly determine the solution in between.

The practical implication: generating a 10-token sentence and a 1000-token paragraph require approximately the same number of relaxation steps, assuming the energy landscape complexity is similar. Latency scales with the number of carriers that must be excited (pattern complexity), not the number of oscillators that must be resolved (sequence length).

This is *Time as Space*—a geometric rather than temporal view of sequence generation.

## 12.4  The Physics of "Horizontal"

We claimed that "the physics of horizontal is the same whether it is a word, a sound wave, or a pixel pattern." Let us make this precise:

- **Text**: The word "horizontal" is a token with a D-dimensional embedding.

- **Audio**: A sound panning left-to-right has specific frequency characteristics (Doppler, stereo phase).

- **Image**: A horizontal line has energy concentrated at $v \approx 0$ in the 2D frequency domain.

All three representations enter the manifold as particles. Through co-activation during training, carriers couple these representations. The word "horizontal" activates carriers that also respond to horizontal image frequencies and horizontal audio characteristics. This is not a metaphor—it is the mechanism.

# 13  Conclusion

We have presented the Sensorium Manifold, a unified thermodynamic substrate for native multi-modal computation. By representing all sensory inputs as oscillators with intrinsic frequencies (via the Universal Tokenizer), we achieve modality-agnostic dynamics. Learning emerges from spectral carrier crystallization—the spontaneous formation of resonant modes that couple oscillators across frequency space—without backpropagation.

The crystallization mechanism represents a fundamental departure from the autoregressive paradigm. By treating time as a spatial dimension and generation as a boundary value problem, we achieve latency that scales with pattern complexity rather than sequence length. The manifold functions as a Holographic Content Addressable Memory: partial inputs address complete patterns because information is distributed across the entire resonant field. "Next token prediction" is merely a degenerate case of this more general capability.

The framework suggests an alternative to the optimization-centric paradigm of modern machine learning. Physical principles—thermodynamics, diffusion, homeostasis, and holographic encoding—may offer paths to adaptive systems that are better suited to continuous, online, embodied learning.

The physics of "horizontal" really is the same across modalities. And that, perhaps, is how perception should work.

# References

Jean-Baptiste Alayrac, Jeff Donahue, Pauline Luc, Antoine Miech, Iain Barr, Yana Hasson, Karel Lenc, Arthur Mensch, Katherine Millican, Malcolm Reynolds, Roman Ring, Eliza Rutherford, Serkan Cabi, Tengda Han, Zhitao Gong, Sina Samangooei, Marianne Monteiro, Jacob Menick, Sebastian Borgeaud, Andy Brock, Aida Nematzadeh, Sahand Sharifzadeh, Mikołaj Binkowski, Ricardo Barreira, Oriol Vinyals, Andrew Zisserman, and Karen Simonyan. Flamingo: A visual language model for few-shot learning. *Advances in Neural Information Processing Systems*, 35: 23716–23736, 2022.

Alexander B. Boyd, Ayoti Patra, Christopher Jarzynski, and James P. Crutchfield. Thermodynamic machine learning through maximum work production. *New Journal of Physics*, 24(8): 083040, 2022.

Aydar Bulatov, Yuri Kuratov, and Mikhail S. Burtsev. Associative recurrent memory transformer. *arXiv preprint arXiv:2407.04841*, 2024.

Zhengdao Chen, Jianyu Zhang, Martín Arjovsky, and Léon Bottou. Symplectic recurrent neural networks. In *International Conference on Learning Representations*, 2020.

Tom Conte, Erik DeBenedictis, Natesh Ganesh, Todd Hylton, John Paul Strachan, R. Stanley Williams, Alexander Alemi, Lee Altenberg, Gavin Crooks, James Crutchfield, Lidia del Rio, Josh Deutsch, Michael DeWeese, Ken Douglas, Federico Faggin, Michael Frank, Ryan Fry, Prahladh Harsha, Michael Henry, Laszlo Kish, James Knight, Suhas Kumar, Hava Li, Patrick Meaney, Tony Molter, Philip Morrision, Cristiano Nisoli, Babatunde Ogunnaike, Lulu Qian, Paul Riechers, Jim Rosen, Barry Rubenstein, Rahul Sarpeshkar, Leonard Silvestri, Barak Singer, Samarth Sinha, Eric Smith, Susanne Still, Kurt Stokbro, Elan Stopnitzky, CJ Teo, William Wootters, Jean-Baptiste Yunes, Michail Zak, and Royce Zia. Thermodynamic computing. *arXiv preprint arXiv:1911.01968*, 2019.

Yilun Du, Shuang Li, Joshua Tenenbaum, and Igor Mordatch. Improved contrastive divergence training of energy-based models. In *International Conference on Machine Learning*, pages 2837–2848, 2021.

Andreas K. Engel, Pascal Fries, and Wolf Singer. Dynamic predictions: Oscillations and synchrony in top-down processing. *Nature Reviews Neuroscience*, 2(10):704–716, 2001.

Karl Friston. The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11(2):127–138, 2010.

Google Gemini Team. Gemini: A family of highly capable multimodal models. *arXiv preprint arXiv:2312.11805*, 2023.

Sam Greydanus, Misko Dzamba, and Jason Yosinski. Hamiltonian neural networks. In *Advances in Neural Information Processing Systems*, volume 32, 2019.

Jiatao Gu, James Bradbury, Caiming Xiong, Victor O. K. Li, and Richard Socher. Non-autoregressive neural machine translation. *International Conference on Learning Representations*, 2018.

Donald O. Hebb. *The Organization of Behavior: A Neuropsychological Theory*. Wiley, New York, 1949.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*, 33:6840–6851, 2020.

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.

Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International Conference on Machine Learning*, pages 448–456, 2015.

Pengzhan Jin, Zhen Zhang, Aiqing Zhu, Yifa Tang, and George Em Karniadakis. Sympnets: Intrinsic structure-preserving symplectic networks for identifying hamiltonian systems. In *Neural Networks*, volume 132, pages 166–179, 2020.

Yoshiki Kuramoto. Self-entrainment of a population of coupled non-linear oscillators. *International Symposium on Mathematical Problems in Theoretical Physics*, pages 420–422, 1975.

Yann LeCun, Sumit Chopra, Raia Hadsell, Marc'Aurelio Ranzato, and Fu Jie Huang. A tutorial on energy-based learning. In Gökhan Bakir, Thomas Hofmann, Bernhard Schölkopf, Alexander J. Smola, Ben Taskar, and S. V. N. Vishwanathan, editors, *Predicting Structured Data*, pages 191–246. MIT Press, 2006.

Jianyu Liu, Quoc Le, Junyoung Chung, and Leon Bottou. Memory mosaics. *arXiv preprint arXiv:2405.06394*, 2024a.

Zexue Liu, Cho-Jui Hsieh, and Yihe Dong. Camelot: Towards large language models with training-free consolidated associative memory. *arXiv preprint arXiv:2402.13449*, 2024b.

Jiasen Lu, Christopher Clark, Sangho Lee, Zichen Zhang, Savya Khosla, Ryan Marten, Derek Hoesein, and Aniruddha Kembhavi. Unified-io 2: Scaling autoregressive multimodal models with vision, language, audio, and action. *Conference on Computer Vision and Pattern Recognition*, 2024.

Takeru Miyato, Priyank Jaini, Chin-Wei Huang, and Yee Whye Teh. Artificial kuramoto oscillatory neurons. *International Conference on Learning Representations*, 2025. Oral presentation at ICLR 2025.

Shen Nie, Fengqi Liu, Lujia Gong, Zebin Wang, Chao Deng, Zhenyu Yuan, Yue Wang, and Saining Xie. Scaling up masked diffusion models on text. *arXiv preprint arXiv:2410.18514*, 2024.

Ilya Prigogine. Time, structure, and fluctuations. *Science*, 201(4358):777–785, 1978. Nobel Lecture.

Ilya Prigogine and Grégoire Nicolis. *Self-Organization in Nonequilibrium Systems: From Dissipative Structures to Order Through Fluctuations*. Wiley, New York, 1977.

Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, and Ilya Sutskever. Learning transferable visual models from natural language supervision. In *International Conference on Machine Learning*, pages 8748–8763, 2021.

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. In *International Conference on Learning Representations*, 2021.

Rajesh P. N. Rao and Dana H. Ballard. Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2(1):79–87, 1999.

Subham Sekhar Sahoo, Marianne Arriola, Yair Schiff, Aaron Gokaslan, Edgar Marroquin, Volodymyr Kuleshov, Alexander M. Rush, and Wei-Ning Chiang. Simple and effective masked diffusion language models. *Advances in Neural Information Processing Systems*, 2024.

Saul Santos, Vlad Niculae, Andrew McCallum, and André F. T. Martins. Sparse and structured hopfield networks. In *International Conference on Machine Learning*, 2024.

Wolf Singer. Neuronal synchrony: A versatile code for the definition of relations? *Neuron*, 24(1): 49–65, 1999.

Jascha Sohl-Dickstein, Eric Weiss, Niru Maheswaranathan, and Surya Ganguli. Deep unsupervised learning using nonequilibrium thermodynamics. *International Conference on Machine Learning*, pages 2256–2265, 2015.

Yang Song, Jascha Sohl-Dickstein, Diederik P. Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. *International Conference on Learning Representations*, 2021.

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.

Xiao-Jing Wang. Neurophysiological and computational principles of cortical rhythms in cognition. *Physiological Reviews*, 90(3):1195–1268, 2010.

Michael Widrich et al. Modern hopfield networks meet encoded neural representations—addressing practical considerations. In *Advances in Neural Information Processing Systems*, 2024.

# A    Additional Experiments

This appendix aggregates additional kernel experiments that exercise the same mechanism across tasks and different sampling/observation choices.

## A.1 Universal Tokenizer Collision Regimes (TOY)

| dataset | vocab | k_per_class | bytes | unique_ids | collision_rate | carriers | cryst_frac | mean |
|---|---|---|---|---|---|---|---|---|
| trie_default | 4096 | 1 | 960 | 704 | 0.2667 | 64 | 0 | |
| trie_default | 2048 | 1 | 960 | 610 | 0.3646 | 64 | 0 | |
| trie_default | 1024 | 1 | 960 | 571 | 0.4052 | 64 | 0 | |
| trie_default | 512 | 1 | 960 | 413 | 0.5698 | 64 | 0 | |
| trie_default | 256 | 1 | 960 | 247 | 0.7427 | 64 | 0 | |
| trie_default | 128 | 1 | 960 | 128 | 0.8667 | 64 | 0 | |
| trie_default | 4096 | 2 | 1920 | 1188 | 0.3812 | 64 | 0 | |
| trie_default | 2048 | 2 | 1920 | 913 | 0.5245 | 64 | 0 | |
| trie_default | 1024 | 2 | 1920 | 808 | 0.5792 | 64 | 0 | |
| trie_default | 512 | 2 | 1920 | 492 | 0.7438 | 64 | 0 | |
| trie_default | 256 | 2 | 1920 | 256 | 0.8667 | 64 | 0 | |
| trie_default | 128 | 2 | 1920 | 128 | 0.9333 | 64 | 0 | |
| trie_default | 4096 | 4 | 3840 | 1847 | 0.519 | 64 | 0 | |
| trie_default | 2048 | 4 | 3840 | 1247 | 0.6753 | 64 | 0 | |
| trie_default | 1024 | 4 | 3840 | 976 | 0.7458 | 64 | 0 | |
| trie_default | 512 | 4 | 3840 | 512 | 0.8667 | 64 | 0 | |
| trie_default | 256 | 4 | 3840 | 256 | 0.9333 | 64 | 0 | |
| trie_default | 128 | 4 | 3840 | 128 | 0.9667 | 64 | 0 | |
| trie_default | 4096 | 8 | 7680 | 2789 | 0.6368 | 64 | 0 | |
| trie_default | 2048 | 8 | 7680 | 1722 | 0.7758 | 64 | 0 | |
| trie_default | 1024 | 8 | 7680 | 1024 | 0.8667 | 64 | 0 | |
| trie_default | 512 | 8 | 7680 | 512 | 0.9333 | 64 | 0 | |
| trie_default | 256 | 8 | 7680 | 256 | 0.9667 | 64 | 0 | |
| trie_default | 128 | 8 | 7680 | 128 | 0.9833 | 64 | 0 | |
| trie_default | 4096 | 16 | $1.536e+04$ | 3858 | 0.7488 | 64 | 0 | |
| trie_default | 2048 | 16 | $1.536e+04$ | 2040 | 0.8672 | 64 | 0 | |
| trie_default | 1024 | 16 | $1.536e+04$ | 1024 | 0.9333 | 64 | 0 | |
| trie_default | 512 | 16 | $1.536e+04$ | 512 | 0.9667 | 64 | 0 | |
| trie_default | 256 | 16 | $1.536e+04$ | 256 | 0.9833 | 64 | 0 | |
| trie_default | 128 | 16 | $1.536e+04$ | 128 | 0.9917 | 64 | 0 | |
| trie_shallow | 4096 | 1 | 960 | 831 | 0.1344 | 64 | 0 | |
| trie_shallow | 2048 | 1 | 960 | 718 | 0.2521 | 64 | 0 | |
| trie_shallow | 1024 | 1 | 960 | 622 | 0.3521 | 64 | 0 | |
| trie_shallow | 512 | 1 | 960 | 434 | 0.5479 | 64 | 0 | |
| trie_shallow | 256 | 1 | 960 | 251 | 0.7385 | 64 | 0 | |
| trie_shallow | 128 | 1 | 960 | 128 | 0.8667 | 64 | 0 | |
| trie_shallow | 4096 | 2 | 1920 | 1402 | 0.2698 | 64 | 0 | |
| trie_shallow | 2048 | 2 | 1920 | 1057 | 0.4495 | 64 | 0 | |
| trie_shallow | 1024 | 2 | 1920 | 858 | 0.5531 | 64 | 0 | |
| trie_shallow | 512 | 2 | 1920 | 505 | 0.737 | 64 | 0 | |
| trie_shallow | 256 | 2 | 1920 | 256 | 0.8667 | 64 | 0 | |
| trie_shallow | 128 | 2 | 1920 | 128 | 0.9333 | 64 | 0 | |
| trie_shallow | 4096 | 4 | 3840 | 2159 | 0.4378 | 64 | 0 | |
| trie_shallow | 2048 | 4 | 3840 | 1429 | 0.6279 | 64 | 0 | |
| trie_shallow | 1024 | 4 | 3840 | 990 | 0.7422 | 64 | 0 | |
| trie_shallow | 512 | 4 | 3840 | 512 | 0.8667 | 64 | 0 | |
| trie_shallow | 256 | 4 | 3840 | 256 | 0.9333 | 64 | 0 | |
| trie_shallow | 128 | 4 | 3840 | 128 | 0.9667 | 64 | 0 | |
| trie_shallow | 4096 | 8 | 7680 | 3164 | 0.588 | 64 | 0 | |
| trie_shallow | 2048 | 8 | 7680 | 1870 | 0.7565 | 64 | 0 | |
| trie_shallow | 1024 | 8 | 7680 | 1021 | 0.8671 | 64 | 0 | |
| trie_shallow | 512 | 8 | 7680 | 512 | 0.9333 | 64 | 0 | |
| trie_shallow | 256 | 8 | 7680 | 256 | 0.9667 | 64 | 0 | |
| trie_shallow | 128 | 8 | 7680 | 128 | 0.9833 | 64 | 0 | |
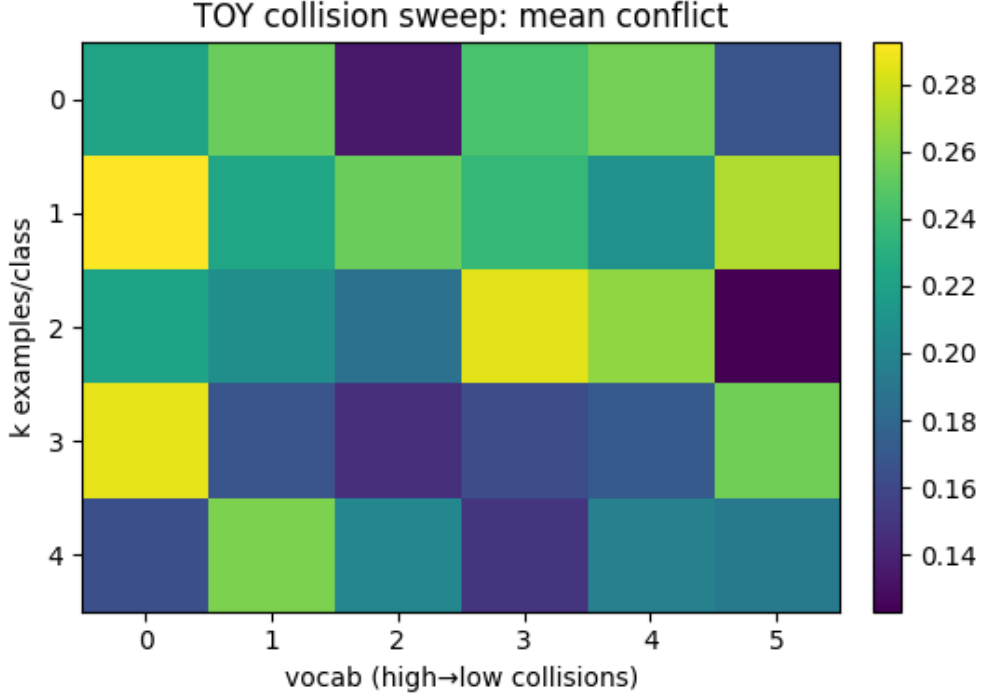
Figure 5: Synthetic trie collision sweep. As the hash vocabulary shrinks, collisions transition from useful compression to aliasing, visible as sustained conflict and degraded crystallization.

Table 5: Kernel next-byte prediction using the Universal Tokenizer (UTF-8 bytes).

| Metric | Value |
|---|---|
| Accuracy | 0.001 |
| Loss (NLL) | 17.4660 |
| Perplexity | 38493488.00 |
| Bits/byte | 25.198 |
| Steps | 800 |

## A.2 Universal Tokenizer Next-Byte Prediction

## A.3 Universal Tokenizer Time-Series Forecasting

## A.4 MNIST Classification from Raw Bytes

## A.5 Kernel Byte Denoising (Text Diffusion)

## A.6 Continuous Kernel Simulation Snapshot

# B Extended Translation Table

For readers seeking deeper correspondences:

*Figure will be generated by* `make paper`

Figure 6: Kernel next-byte prediction (UTF-8 bytes) using carrier-spectrum scoring.

Table 6: Kernel time-series forecasting via next-byte prediction on a quantized synthetic series.

| Metric | Value |
|---|---|
| Accuracy (exact byte) | 0.003 |
| MAE (byte) | 81.10 |
| MSE (byte) | 8983.65 |
| Steps | 800 |

| Physics Concept | ML Analogue |
|---|---|
| 3D Manifold space | Simulation volume / embedding space |
| Particle (spatial layer) | Token / feature vector |
| Oscillator (spectral layer) | Hidden state with phase/frequency |
| Intrinsic frequency $\omega$ | Token identity (from hash) |
| Spectral carrier $C_k$ | Learned mode / pattern detector |
| Tuning kernel $T_{ik}$ | Attention weight (frequency-based) |
| Gate width $\sigma_k$ | Receptive field / specialization |
| Carrier crystallization | Weight freezing / memory consolidation |
| Carrier conflict | Mode collision / interference |
| Carrier splitting | Mixture component birth |
| Anchored oscillators | Associated tokens in memory |
| Top-down energy bias | Prior / completion signal |
| Gravity field $\phi$ | Attraction potential |
| Heat diffusion | Uncertainty propagation |
| Particle collision | Local interaction / gating |
| Langevin noise | Stochastic exploration / dropout |
| Idle compute | Offline consolidation / dreaming |
| Crystallization | All-token generation (cf. autoregression) |
| Holographic CAM | Content-addressable associative memory |
| Time as Space | Sequence index as spatial coordinate |
| Boundary Value Problem | Constrained generation (inpainting) |

*Figure will be generated by `make paper`*

Figure 7: Kernel time-series forecasting via next-byte prediction on a quantized synthetic series.

Table 7: Kernel MNIST classification from raw bytes using the Universal Tokenizer (position-aware hashing).

| Metric | Value |
|---|---|
| Accuracy | 0.110 |
| Error rate | 0.890 |
| Train samples | 300 |
| Eval samples | 100 |

# C   Hyperparameter Settings

*Table will be generated by `make paper`*
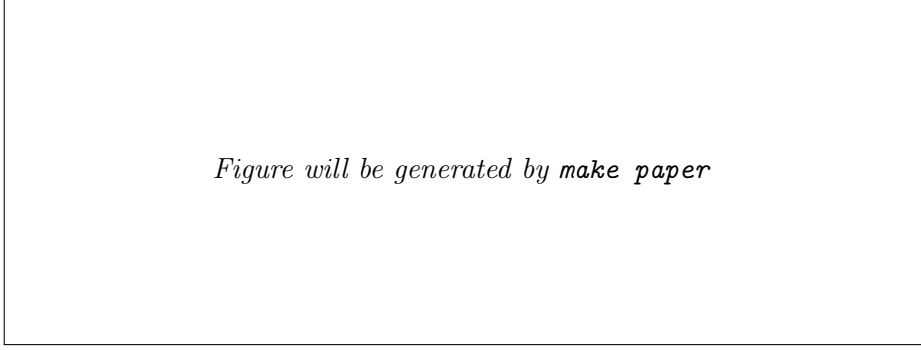
*Figure will be generated by* `make paper`

Figure 8: Kernel MNIST confusion matrix (Universal Tokenizer hashed pixel stream).

Table 8: Kernel byte denoising ("text diffusion") using the Universal Tokenizer.

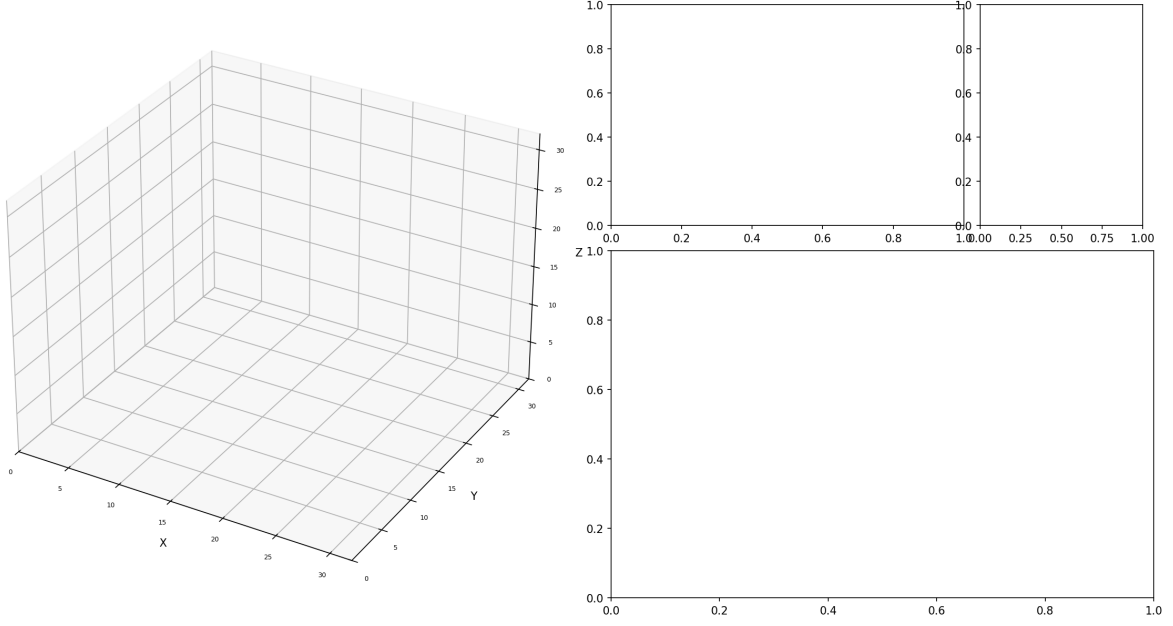| Metric | Value |
| --- | --- |
| Masked-byte accuracy | 0.000 |
| Mask fraction | 0.20 |
| Bytes | 1500 |

Figure 9: Final dashboard frame from a finite kernel simulation run.

# D    Pseudocode

---

**Algorithm 1** Unified Manifold Step (Spatial + Spectral Layers)

---

**Require:** Particles $\{(\boldsymbol{x}_i, \boldsymbol{v}_i, E_{\text{osc},i}, Q_i, \omega_i, \theta_i)\}$, Carriers $\{C_k, \Omega_k, \sigma_k\}$
 1: // **Spatial Layer (Thermodynamic Physics)**
 2: Scatter particle masses to gravity field; total internal energy $U_i = Q_i + E_{\text{osc},i}$ to internal-energy field
 3: Solve $\nabla^2\phi = 4\pi G\rho$ (Poisson equation for gravity potential)
 4: Compute temperature field: $T(\boldsymbol{x}) = U(\boldsymbol{x})/(\rho(\boldsymbol{x})c_v)$
 5: Diffuse thermal field (FFT/Jacobi): $T \leftarrow T + \alpha\nabla^2 T \cdot \Delta t$
 6: Gather from fields: $\boldsymbol{F}_i = -m_i\nabla\phi - (m_i/\rho)\nabla P(\rho, T)$; sample $U(\boldsymbol{x}_i)$
 7: Update velocities: $\boldsymbol{v}_i \leftarrow \boldsymbol{v}_i \cdot e^{-\gamma\Delta t} + \boldsymbol{F}_i\Delta t/m_i$
 8: Update positions: $\boldsymbol{x}_i \leftarrow \boldsymbol{x}_i + \boldsymbol{v}_i\Delta t$
 9: Compute particle collisions (spatial hash or brute force)
10: Local conservative exchange: $Q_i \leftrightarrow E_{\text{osc},i}$ via Planck relaxation (Equation (10), Equation (14))
11: // **Spectral Layer (Carrier Coupling)**
12: Compute energy statistics: $\bar{E}, \sigma_E$ (GPU reduction)
13: Compute oscillator amplitudes: $A_i = \sqrt{E_{\text{osc},i}}$
14: Build frequency bins over carriers (GPU-only bucket + prefix sum)
15: **for** each oscillator $i$ (parallel) **do**
16:     Accumulate contributions over nearby carrier bins (skip fp32-zero couplings)
17: **end for**
18: **for** each carrier $k$ (parallel) **do**
19:     Compute force: $F_k = \sum_i T_{ik}z_i - \lambda C_k + \sqrt{2T}\,\xi_k$
20:     Update carrier: $C_k \leftarrow C_k + F_k\Delta t$
21:     If conflict > threshold, spawn new carrier from offender
22:     Update crystallization state: volatile → stable → crystallized
23: **end for**
24: Update oscillator phases: $\theta_i \leftarrow \theta_i + (\omega_i + \kappa \cdot \text{torque}_i)\Delta t$
25: Top-down bias: crystallized carriers inject energy into anchored oscillators
26: Spawn carriers for uncoupled oscillators

---