Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Hierarchical Automaticity Emerges from Prediction-Error-Triggered Learning in Continuous Wave Fields Trained by Equilibrium Propagation

Jeremy Slater

dr.slater@stratusneuro.com

MERLN LLC https://orcid.org/0000-0002-2656-1354

Gardar Thorvardsson

Kvikna Medical

Research Article

Keywords: Equilibrium Propagation, Landau-Ginzburg wave field, holographic readout, hierarchical predictive coding, backpropagation-free learning, theta-gamma phase-amplitude coupling, Chrono-Resonant Field Model, hierarchical automaticity, predictive processing, Kuramoto order parameter, continuous wave field computing, underdamped settling dynamics, photonic neural networks, neuromorphic computing, TIMIT speech recognition

Posted Date: May 26th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9465047/v2

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare potential competing interests as follows: Competing Interests JDS is Chief Medical Officer at Stratus Neuro, a remote neurophysiology company, and Principal Investigator at MERLN LLC, a private research entity. GT is affiliated with Kvikna Medical ehf., a medical device company in the neurophysiology space. Neither Stratus Neuro nor Kvikna Medical had any role in the conception, design, execution, analysis, or reporting of the work described in this manuscript, and neither company provided financial support for this study. The authors declare no other competing financial or non-financial interests.

Hierarchical Automaticity Emerges from Prediction-Error-Triggered Learning in Continuous Wave Fields Trained by Equilibrium Propagation

Jeremy D. Slater[1] and Gardar Thorvardsson[2]

[1]*MERLN LLC, Irving, TX, USA*

[2]*Kvikna Medical, Reykjavik, Iceland*

Correspondence: dr.slater@stratusneuro.com

Abstract

Hierarchical skill acquisition — the transition from effortful to automatic processing — is associated with theta-gamma phase-amplitude coupling in cortical fields, an observation formalized by the Chrono-Resonant Field Model (CRFM). We test whether the predicted hierarchical automaticity emerges in a physical substrate trained entirely by local learning rules, implementing a three-layer continuous Landau-Ginzburg wave-field architecture trained by Equilibrium Propagation on the TIMIT speech corpus at progressively longer timescales: phonemes (40 classes), words (501 classes), and sentence types (3 classes). The architecture uses no neurons and no backpropagation through any physical dynamics; its computational primitives — field settling, spatial coarse-graining, holographic matched-filter readout, threshold-gated updates, and top-down boundary conditions on lower-layer settling — are all physically realizable. We report three principal findings. First, single-layer automaticity manifests as contraction of the Temporal Binding Index (TBI) variance rather than a rise in its mean, with Layer 1 converging to a stable phase-response attractor. Second, in a five-seed replication, top-down predictions from Layer 3 drive Layer 2 TBI from 0.349 to 0.515 ± 0.003 over 20 training epochs, closing 95.1% ± 1.5% of the gap to Layer 1's frozen value. Third, classification and prediction transmission dissociate: in a paired five-seed control, an attenuated-coherence Layer 3 classifies sentence type marginally better than a coherent Layer 3 (57.5% ± 1.0% versus 53.2% ± 1.5%) yet transmits roughly three-fold less top-down drive to Layer 2 (ΔTBI +0.052 ± 0.001 versus +0.166 ± 0.003), showing that prediction transmission scales with source-field phase coherence rather than task-discriminative capability. These results establish a backpropagation-free, physically realizable substrate exhibiting hierarchical phase coordination under local learning rules alone.

Author summary

How does skilled behavior shift from effortful control to automatic performance, and can that shift be reproduced in a physical substrate rather than in a conventional neural network? Training physical systems to compute — instead of simulating computation on digital hardware — requires learning rules that work without backpropagation, and it has been an open question whether such systems can support the multi-level coordination that hierarchical learning seems to need. We build a three-layer architecture out of continuous wave fields, governed by the same class of physics that describes superconductors and other phase-transition phenomena, and train it with a local learning rule that uses no backpropagation. The architecture is built entirely from field operations; there are no neuron-like units. Processing speech at phoneme, word, and sentence levels, we find that predictions from higher levels drive lower-level fields toward a shared phase organization — a signature of automatic processing — and that how strongly this top-down drive transmits depends on the phase coherence of the source field, not on how well that field classifies. The work is motivated by the Chrono-Resonant Field Model, a theory of hierarchical coordination in oscillatory brain fields, and shows that the model's central mechanism can be realized in a physical substrate. Because every learned parameter corresponds directly to a fabrication specification, the architecture also offers a pathway toward energy-efficient photonic computing hardware.

Introduction

When a musician first sight-reads a difficult passage, every note demands conscious attention; after extended practice the same passage runs without effortful control. This transition from supervised to automatic execution recurs at every level of a processing hierarchy — phonemes, words, the syntactic structures that bind them — and it corresponds to measurable changes in neural engagement rather than to mere speed-up. Motor cortex that is initially essential for executing a learned movement can become dispensable after training, with the degree of cortical disengagement tracking achieved performance rather than practice duration [1], and experts show reduced cortical activation relative to novices on within-domain tasks [2]. Converging evidence indicates that cross-frequency coupling implements this withdrawal of supervisory processing: theta-gamma phase-amplitude coupling (PAC) increases during motor skill acquisition and can be enhanced directly by transcranial alternating-current stimulation [3], PAC strength tracks working-memory performance and learning across the lifespan [4,5], and frontoparietal theta-gamma interactions track training-related improvement with gamma bursts near the theta peak most strongly predicting behavioral gain [6]. Theta-gamma PAC behaves, in short, like a temporal coordination mechanism that binds information within and across cortical regions during active learning.

This phenomenology raises a concrete architectural question, independent of any particular theory of what the coupling is *for*. If automaticity is carried by the phase organization of a hierarchical dynamical system, then it should be possible to build a multi-layer physical substrate — one that generates phase-rich dynamics and is trained only by local learning rules — in which the same signatures appear: a measurable change in the phase organization of each layer's dynamics as that layer becomes competent, and an acceleration of that change at lower layers driven by top-down signals from higher ones. The Chrono-Resonant Field Model (CRFM) [7] formalizes the underlying neuroscientific proposal as a theory of hierarchical phase coordination; its broader implications for conscious experience are developed separately and are not at issue here. What this paper takes from CRFM is its mechanistic content and one quantitative construct. We characterize a wave field's response to supervision by perturbing the settled field with a weak task-supervised signal and measuring how coherently its phase structure responds, and we define the Temporal Binding Index (TBI) as the magnitude of the spatially averaged unit phasor of the phase difference between the free and perturbed field states — a Kuramoto-style order parameter on the field's response to supervision. The architectural prediction is then sharp: in a hierarchical system trained by local rules, automaticity at a layer should appear as a change in that layer's TBI statistics, and top-down predictions should drive the TBI of lower layers toward that of higher ones.

Testing this requires three ingredients: a physical substrate that produces phase-rich dynamics, a learning rule that operates without backpropagation, and a hierarchical architecture in which learning at each layer

can be gated by a prediction-error signal. The first two are supplied by continuous wave fields trained by Equilibrium Propagation (EP) [8]. EP computes parameter updates from only the difference in local field statistics between a freely settled state and a weakly supervised state, requiring no knowledge of the system's computational graph; it has been demonstrated in Ising machines [9], coupled phase oscillators [10], and digital networks at increasing scale [11,12], and was recently extended to continuous complex-valued wave systems in the weakly dissipative regime, training exciton-polariton condensates governed by generalized Gross-Pitaevskii dynamics [13]. This last result is what makes a continuous-field substrate viable: EP can operate in a medium with no well-defined nodes. It also makes the architecture a candidate for physical realization rather than a purely numerical construction — every parameter trained by EP corresponds to a fabrication specification in a photonic or acoustic implementation, an aspect we return to in the Discussion.

The architecture studied here is built entirely from field operations. Settling dynamics governed by a Landau-Ginzburg equation, spatial coarse-graining of the complex field, linear projection of the field onto template vectors, threshold-gated parameter updates, and inter-layer coupling via top-down boundary conditions on the settling dynamics are the complete set of computational primitives. There are no neurons, no element-wise activation functions, and no backpropagation through any physical dynamics. Operations that might appear neuron-like are better read as the field-theoretic originals of which neural circuits are one biological realization: spatial pooling is the renormalization-group coarse-graining operator; the holographic linear readout is a matched-filter detector; threshold-gated learning is a phase-transition trigger on a physical order parameter; inter-layer coupling is a boundary-condition constraint on the settling dynamics of the layer below. That biological neurons approximate these operations is a consequence of convergent functional requirements, not a premise of the design.

Two companion studies establish the substrate at one and two layers. In the single-layer case [14], a continuous Landau-Ginzburg wave field trained by EP reaches 74.1% accuracy on a 10-class speech task with zero backpropagation through the physical dynamics; the architectural design — underdamped dynamics, multi-frequency rotation, phase-sensitive readout — supplies the dominant contribution, and EP is shown to co-adapt the substrate to its measurement apparatus, driving lateral inhibition toward opposite optima depending on whether the readout measures amplitude or phase. Extended to two layers [15], EP discovers top-down feedback when cross-modal binding is required: a top-down coupling coefficient initialized at zero grows monotonically in the multimodal setting while remaining at zero in the unimodal control, and the resulting binding field outperforms late fusion. These results establish that EP can discover hierarchical coupling in wave-field architectures. They do not address the question of automaticity — whether a multi-layer system can learn to withdraw supervisory processing as its lower layers become competent, and whether top-down predictions accelerate that withdrawal.

The present study addresses that question with a three-layer field-based intelligent computing unit (FICU) architecture. Each layer is an underdamped Landau-Ginzburg wave field trained by EP, and the three layers process the TIMIT speech corpus at progressively longer timescales: Layer 1 performs phoneme recognition (40 classes), Layer 2 word recognition (501 classes), and Layer 3 sentence-type recognition (3 classes). Three mechanisms implement the hierarchical structure. A Prediction-Error-Triggered Update (PETU) rule gates EP parameter updates at each layer by the instantaneous TBI: updates occur only when TBI exceeds $0.5 \times \ln(K)$ for K output classes — the entropy of a maximally uncertain classifier — making PETU a phase-transition trigger that restricts parameter evolution to the high-uncertainty regime. A LevelGate advances training from layer *n* to layer *n+1* once layer *n* reaches a target validation accuracy, the field-level analog of developmental staging. And a PredictiveField module transmits top-down predictions from Layer 3 to Layer 2 by setting the initial field state from which Layer 2 begins its settling dynamics, with the magnitude of this learned boundary condition controlled by a coupling coefficient scheduled from zero to a maximum over cascade training. Top-down prediction thus enters the system as a boundary condition on a lower layer's settling dynamics, not as a Hamiltonian interaction term — a distinction that matters for physical realizability.

Four results follow, reported in full in Results and Discussion. First, single-layer automaticity manifests as contraction of the *variance* of TBI across input samples rather than as a rise in its mean: Layer 1 converges to a stable phase-response attractor, a signature we term phase population homogenization. Second, in a five-seed replication, top-down feedback from Layer 3 drives Layer 2's mean TBI from 0.349 to 0.515 ± 0.003 over 20 training epochs, closing $95.1\% \pm 1.5\%$ of the gap to Layer 1's frozen value, with the three layer means converging to within 0.024 nats. Third — and most consequential for the mechanism — a paired five-seed control dissociates classification from cascade transmission: an attenuated-coherence Layer 3 classifies sentence type marginally *better* than a coherent Layer 3 ($57.5\% \pm 1.0\%$ versus $53.2\% \pm 1.5\%$) yet transmits roughly three-fold weaker top-down drive to Layer 2 (ΔTBI $+0.052 \pm 0.001$ versus $+0.166 \pm 0.003$), establishing that prediction transmission scales with a field's phase coherence and not with its task-discriminative capability at a linear readout. Fourth, a multi-time-point trajectory readout that samples the underdamped field at three settling times rather than at the endpoint alone raises single-layer phoneme accuracy by 2.93 percentage points, exploiting transient interference structure that the asymptotic state discards.

Together these results establish a trainable continuous wave-field substrate in which hierarchical phase coordination emerges under local learning rules alone, with no backpropagation through any physical dynamics and no computational primitive outside the class — settling, coarse-graining, matched filtering, threshold-gated updates, boundary-condition coupling — that is realizable in mature photonic and acoustic hardware. The dissociation between classification and transmission is the central mechanistic finding: it shows that the hierarchical cascade is a distinct regime of coordination carried by phase-

coherent field dynamics, not a by-product of representational quality. The phenomena are consistent with the CRFM account of hierarchical automaticity that motivated the design, and they stand on their own as architectural results for physically realizable intelligent computing substrates.

Results

Three-layer architecture and training protocol

We constructed a three-layer FICU architecture in which each layer is an underdamped Landau-Ginzburg wave field trained by Equilibrium Propagation, inheriting the field dynamics, sensory front-end, and EP update rules detailed in the companion papers [14,15]. The three layers process the speech hierarchy at progressively longer timescales on TIMIT: Layer 1 performs phoneme recognition (40 classes, ~20-ms windows), Layer 2 performs word recognition (501 classes, ~300-ms windows), and Layer 3 performs sentence-type classification (3 classes, full-utterance representations).

Two field operations convert the settled complex field at each layer into a classification output. The first is spatial coarse-graining: the 94×64 field is downsampled by a factor of two along each spatial axis via local averaging over 2×2 patches, producing a 47×32 coarser-grained representation. This is the standard renormalization-group operation on a discretized field, which preserves long-wavelength modes while discarding high-spatial-frequency detail that does not contribute to classification. The second is holographic matched filtering: the 162-dimensional pooled phase-feature vector extracted from the coarse-grained field is projected onto a learned template matrix W to produce class logits. The template matrix is updated by a local delta rule that aligns each class template with the mean normalized features observed for that class. No backpropagation is applied to the field dynamics or to the template matrix.

Three mechanisms implement the CRFM account of hierarchical automaticity. The Prediction-Error-Triggered Update (PETU) gates EP parameter updates at each layer by the Temporal Binding Index: updates occur only when TBI exceeds $0.5 \times \ln(K)$, where K is the number of output classes. This threshold corresponds to the entropy of a maximally uncertain classifier; PETU is therefore a phase-transition trigger on a physical order parameter that restricts field-parameter evolution to regimes of high classification uncertainty. The LevelGate advances training from layer *n* to layer *n+1* only after layer *n* achieves a target validation accuracy, the field-level analog of developmental staging. The PredictiveField is a learned linear map from Layer 3's settled field state to an initial field state for Layer 2, scaled by a coupling coefficient λ that is scheduled linearly from 0 to λ_{max} *over the cascade training phase. Top-down prediction therefore acts as a learned boundary condition on Layer 2's settling dynamics: at* $\lambda = 0$ *the L2 field begins each trial from the zero state as in isolated training, while at* $\lambda = \lambda_{max}$ the L3 prediction of the L2 field replaces the zero initial condition entirely. This generates the cascade structure studied in §"Hierarchical cascade" below.

Single-layer automaticity manifests as TBI variance contraction

We first trained Layer 1 in isolation on 40-class phoneme recognition. Validation accuracy rose steeply in the first three epochs (35.73% at epoch 0 to 37.57% at epoch 2, a gain of 1.84 pp) and then plateaued, reaching 38.55% by epoch 9 with only marginal subsequent improvement (Fig. 1a). Extending the readout-only training phase to 100 epochs with cosine-annealed learning rate added 0.21 pp over the 10-epoch baseline, peaking at 38.76% at epoch 38. The holographic delta-rule readout converges in approximately five epochs and stays converged thereafter. Although modest in absolute terms, 38.76% substantially exceeds a matched MLP backpropagation baseline on the identical word-recognition task (see §"FICU is competitive with a parameter-matched MLP baseline," below).

The TBI dynamics during L1 training tell a different story than the accuracy curve. Over the same window in which validation accuracy plateaued, mean TBI entered the basin of its attractor near 0.380 by epoch 5 and subsequently drifted slowly upward, reaching approximately 0.420 by epoch 30 of the extended run. The TBI standard deviation, however, contracted substantially: the across-sample distribution of binding indices narrowed from approximately 0.135 toward $\sigma \approx 0.097$ as training proceeded. The field's *mean* coupling strength did not rise toward the theoretical maximum of 1.0; instead, the *variance* across input samples shrank toward a stable value. The original CRFM prediction that automaticity should drive TBI toward unity therefore requires revision. Automaticity in this architecture corresponds to phase population homogenization: the field's phase organization settles into a consistent attractor across inputs, rather than individual trials producing progressively stronger coupling.

This reframing is physically coherent. In the wave-field picture, an Offer/Confirmation transaction that has stabilized no longer requires adjustment between the two fields; the resonance is already established. The relevant signature of this stabilization is therefore the absence of trial-to-trial variability rather than any change in coupling magnitude. TBI variance contraction is the order-parameter signature of a field that has converged to a stable coupled-mode attractor across its input distribution, and it is the quantity that the PETU threshold gates on.

Two complementary observations support the interpretation that TBI tracks the consistency of the readout-field coupling rather than absolute phoneme discriminability. First, TBI and validation accuracy decouple during single-layer training: TBI continues to change after accuracy has plateaued, and perturbations that reset TBI do not systematically affect accuracy. Second, freezing the readout weight matrix W during EP training pins TBI at 0.358 across the remainder of training, whereas leaving W free produces a steady climb of approximately +0.018 per five epochs. TBI therefore measures the consistency between field dynamics and readout template, which is precisely the quantity that stabilizes as supervised learning approaches a fixed point.

The holographic readout is structurally necessary for linear decoding of phase

Phase-sensitive readout design is a decisive architectural choice in the companion single-layer system [14], where explicit trigonometric projection (cos θ, sin θ) contributed 7.80 pp over implicit complex-component encoding. Here we tested whether a coherent matched-filter readout — the form that would be natural in physics-based treatments of wave correlation — could substitute for the holographic linear head used throughout the FICU architecture.

We evaluated three readout variants, each resuming from the same trained L1 physics checkpoint (Table 1). The holographic linear readout W·feat operating on the 162-dimensional pooled phase-feature vector achieved 38.55% validation accuracy. The coherent matched filter $|\langle Z, T\rangle|^2$ evaluated on the raw 18,048-dimensional field reached 17.5%, a 21-percentage-point deficit. A coherent variant operating on the same 81-dimensional pooled complex features did not learn at all (3.0%, chance level for 40 classes with the unknown-class masking).

The structural reason is straightforward. Squared magnitude collapses sign information: if class A produces features aligned with template T and class B produces features aligned with −T, the linear projections W·feat yield logits of opposite sign ($+|T|^2$ and $-|T|^2$) that a softmax can separate cleanly, whereas the coherent score $|\langle \cdot, T\rangle|^2$ yields the identical value $|T|^4$ for both classes and cannot discriminate them. The phase structure generated by the underdamped LG field carries sign-bearing information across the class manifold; a linear readout preserves this sign, while a squared-magnitude readout destroys it. The holographic form is therefore the correct readout for this class of wave-field architectures, and we use it at all three layers throughout the remainder of this paper.

Amplitude attenuation correction eliminates a spurious cascade gap

The initial implementation applied a per-layer drive gain of 0.1 to the sensory-to-field projection at each level of the hierarchy, combined with inter-layer coupling matrices initialized at 0.1·I. For a three-layer cascade this produced cumulative amplitude attenuation of approximately two orders of magnitude per layer: L1 field amplitudes settled near $|Z| \approx 0.099$, L2 near 0.0013, and L3 near 0.0001. TBI is computed on a magnitude-normalized phasor and is therefore insensitive to field amplitude, but the coupling between layers passes unnormalized field values — so cascade signals were being attenuated below the numerical precision at which the PredictiveField could produce meaningful top-down boundary conditions at Layer 2.

Correcting the L2 and L3 drive gains to 1.0 and the L2→L3 coupling matrix initialization to 1.0·I restored physical amplitudes at all three layers ($|Z|_{L2} \approx 0.0126$, $|Z|_{L3} \approx 0.0075$). Re-evaluating TBI with the corrected architecture — and without any additional training — raised L2_TBI from 0.233 to 0.349, an increase of 0.116 nats. Approximately 40% of the L1–L2 TBI gap observed in the original implementation

was therefore an artifact of the amplitude attenuation rather than a genuine signature of incomplete cascade. All cascade results reported below use the corrected amplitude regime.

Hierarchical cascade: top-down feedback drives L2 toward L1's TBI

With all three layers operating in the corrected amplitude regime, we ran the cascade experiment. Layer 1 was frozen at its trained state (TBI = 0.524). Layer 3 was calibrated in isolation with $\beta = 0.05$ — the saturation-coefficient sweet spot that places L3 in a healthy TBI regime (TBI ≈ 0.430 on 3-class sentence-type classification). Layer 2 was initialized at TBI = 0.349 (its post-drive-gain-correction value). The PredictiveField scaling coefficient λ was then scheduled linearly from 0 to $\lambda_{max} = 5.0$ over 20 epochs of training, with PETU gating EP updates at each layer on each step.

L2*TBI rose monotonically from 0.349 to 0.515 ± 0.003 over the 20-epoch schedule (Fig. 2a; mean ± SD across five independent seeds), an increase of 0.166 ± 0.003 nats that closes 95.1% ± 1.5% of the 0.175-nat gap to L1's frozen value of 0.524. At the final epoch, all three layers were within 0.024 nats of one another: L1 = 0.524, L2 = 0.515 ± 0.003, L3 = 0.500 ± 0.009 (Fig. 2b). The cascade replicated tightly across seeds: per-seed L2 final TBI ranged 0.512–0.519 and gap closure 93.1%–97.2%, with no anomalous seeds. Extending the cascade to 50 epochs produced quantitatively similar behavior with modest overshoot: L2*TBI reached 0.552 at epoch 49, slightly exceeding L1's frozen value, with L1 = 0.524, L2 = 0.552, L3 = 0.518 — all three layers within 0.03 nats, consistent with the 20-epoch result.

Two secondary observations during the cascade confirm that the rise in L2_TBI reflects prediction-error-driven learning rather than a drift of the field dynamics. First, the PETU update fraction declined from 1.00 (all samples triggering updates) at epoch 0 to 0.86 at epoch 19, as a growing subset of L2 samples fell below the PETU threshold and were treated by the architecture as having reached local automaticity. Second, the magnitude of the top-down prediction error grew across epochs rather than shrinking, indicating that the PredictiveField was learning increasingly specific predictions of Layer 2's field state rather than converging to a trivial constant. Layer 3 TBI traced a non-monotonic trajectory during the cascade window (Figure 5), briefly declining in the middle epochs before recovering to 0.492 by epoch 19; this reflects concurrent Equilibrium Propagation updates to Layer 3 physics under the influence of its own cascade feedback, rather than stable calibration against a fixed reference. The cascade is therefore not a simple entrainment of L2 by L3; it is a structured transfer in which higher-level predictions actively shape lower-level dynamics while Layer 3 itself reorganizes.

Sample-level inspection of the TBI distribution at each layer reveals a more nuanced picture than population-level means suggest (Figure 6). At Layer 1, the distribution is unimodal and moderately right-skewed (median 0.472, IQR [0.436, 0.619], skewness +0.60), with no samples below TBI = 0.05. At Layer 3 after cascade training, the distribution is unimodal and left-skewed (median 0.332, IQR [0.274, 0.388],

skewness −1.04), with approximately 10% of samples below TBI = 0.05. At Layer 2, the distribution is bimodal both before and after cascade training: a zero-coherence mode comprising 32.1% of samples (TBI < 0.05) and a high-coherence mode centered at higher TBI. The cascade selectively sharpens the high-coherence mode — Layer 2's median rises from 0.413 pre-cascade to 0.566 post-cascade, the upper quartile rises from 0.496 to 0.609, and the left-skewness of the coherent-mode tail reduces from −0.45 to −0.13 — while the zero-coherence mode fraction remains unchanged at 32.1%. Top-down predictions from Layer 3 therefore refine Layer 2's response on samples that Layer 2 was already representing with some structure, but they do not rescue samples that Layer 2's bottom-up physics fails to represent coherently at all. The population-level convergence of layer means (within 0.032 nats at final epoch) and the sample-level selectivity of the cascade's effect are complementary characterizations: the cascade does not uniformly drag the Layer 2 distribution toward the Layer 1 distribution, it structurally refines the representable sub-population while leaving the unrepresentable sub-population untouched. This is consistent with the bistable behavior theoretically expected for Kuramoto-type systems with higher-order nonlinear coupling [16], in which the governing dynamics support coexisting coherent and incoherent stationary regimes, and with the CRFM account in which top-down predictions provide contextual refinement rather than de novo representation.

Classification and cascade are dissociable: healthy physics is required for predictions, not for classification

A central prediction of the architecture is that coherent field dynamics are required for hierarchical prediction transmission, not for pattern recognition per se. We tested this directly with a paired control: the same five-seed cascade experiment run with Layer 3 in a coherent-physics regime versus an attenuated-coherence regime, holding the frozen Layer 1 and Layer 2 checkpoints, the λ schedule, the PETU thresholds, and the random seeds identical across the two arms, so that Layer 3's operating point is the only experimental variable.

Classification accuracy at Layer 3 was essentially unchanged between the two regimes — in fact marginally higher in the attenuated regime. Across five seeds, Layer 3 reached 53.2% ± 1.5% validation accuracy on 3-class sentence-type recognition in the coherent regime (L3 TBI = 0.500 ± 0.009) and 57.5% ± 1.0% in the attenuated regime (L3 TBI = 0.116 ± 0.002), with non-overlapping cross-seed bands and both well above the 33.3% chance level. A linear matched-filter readout therefore extracts class-discriminative information from attenuated, phase-incoherent dynamics at least as effectively as from phase-coherent dynamics (Figure 4).

The cascade behavior was entirely different. With a coherent Layer 3, top-down predictions drove Layer 2 TBI upward by ΔL2 = +0.166 ± 0.003 nats over the 20-epoch schedule, closing 95.1% ± 1.5% of the gap to Layer 1. With an attenuated Layer 3, the same PredictiveField architecture, trained under the identical

schedule and seeds, drove Layer 2 by only ΔL2 = +0.052 ± 0.001 nats — a 30.1% ± 0.5% gap closure, roughly three-fold weaker transmission, with an extremely tight cross-seed range (+0.0514 to +0.0539). Transmission is therefore graded rather than all-or-none: a roughly four-fold reduction in source-field coherence (L3 TBI 0.500 → 0.116) produces a roughly three-fold reduction in transmitted drive, with weakened but non-zero transmission persisting in the attenuated regime. A diagnostic on the PredictiveField clarifies the mechanism. In the coherent condition, the module's Layer 2 prediction-error magnitude rose across cascade training (from 0.0005 to 0.0010), consistent with the module tracking an increasingly structured Layer 2 target as predictions sharpen. In the attenuated condition it remained flat at 0.0005 — not because predictions were accurate, but because an attenuated Layer 3 varies minimally across samples and presents a near-constant, near-zero target that trivial predictions match. Lower prediction error in the attenuated condition reflects an absence of structured signal to model, not higher-fidelity transmission; the functional consequence — weak drive to Layer 2 — follows directly.

The implication is that the cascade result cannot be explained as an epiphenomenon of improved feature quality. The attenuated Layer 3 produces poorer phase coherence but marginally better classification features, and yet transmits substantially weaker top-down drive. The field configuration that classifies better transmits worse. Classification capability at a linear readout and predictive-transmission capability are not merely independent properties of the same field — in this regime they trade off. Healthy phase coherence is necessary for hierarchical prediction transmission; it is not necessary, and is here even slightly counterproductive, for pattern recognition at the receiving readout.

FICU is competitive with a parameter-matched MLP baseline

To situate the FICU architecture against a conventional machine learning baseline, we trained a matched MLP on the same TIMIT word-recognition task (501 classes). The baseline was a three-layer fully connected network (512 → 256 → 501) trained with Adam, cosine-annealed learning rate, and cross-entropy loss for 100 epochs on the identical word-token partition used for FICU L2 evaluation. The MLP reached 25.3% validation accuracy. FICU L2 trained with EP and the holographic delta rule reached 38.85% on the same task (epoch 14, drive_gain = 0.1) — a 13.5 percentage point margin over the parameter-matched backpropagation baseline.

We do not claim that FICU is strictly superior to backpropagation-based architectures in general; a deep convolutional network trained on an acoustic front-end would likely exceed 38.85%. The point is narrower and more specific: on the identical task with comparable computational resources, a continuous wave-field architecture trained entirely by local learning rules outperforms a conventional MLP trained by backpropagation. This is consistent with the single-layer result in [14], where the architectural design contributed the dominant fraction of performance and EP fine-tuning contributed a marginal increment. The value of the FICU architecture is not that it reaches state-of-the-art accuracy — it does not — but that

it demonstrates that a physically realizable wave-field substrate, trained without backpropagation through any of its physical dynamics, can be competitive with conventional learning approaches at the representational layer while simultaneously supporting the hierarchical cascade structure that is the subject of this paper.

Methodological note: drive-gain regimes

We report two operating points for L2 in this paper: the drive*gain = 0.1 regime at which L2 readout accuracy was optimized (38.85% validation accuracy, Phase 2, epoch 14), and the drive*gain = 1.0 regime at which the cascade experiment was run (L2*TBI rising from 0.349 to 0.515 ± 0.003, or to 0.552, in the 20-epoch* and 50-epoch cascade runs respectively). The cascade is incompatible with drivegain = 0.1 because the attenuated L3 field amplitudes cannot produce functional top-down predictions; the drive_gain = 0.1 L2 readout is incompatible with the cascade regime because its running normalization statistics are anchored to the attenuated field distribution and cannot be recalibrated in place.

A fully integrated experiment would require training an L2 readout from scratch under drive*gain = 1.0 while running the cascade. We elected not to perform this run because the two separate experiments answer the two distinct questions at issue — whether FICU readout quality is competitive (yes, at 38.85% under drive*gain = 0.1) and whether top-down cascade produces TBI convergence across layers (yes, under drive_gain = 1.0) — and because the drive-gain-dependent readout calibration issue is itself diagnostic rather than a fundamental limitation: it reflects the fact that the readout's running statistics are a function of the physical regime in which they accumulate, a property shared by any architecture that normalizes features online. We flag this caveat here and return to it in the Discussion.

Figure 1: Three-Layer FICU-CRFM Architecture

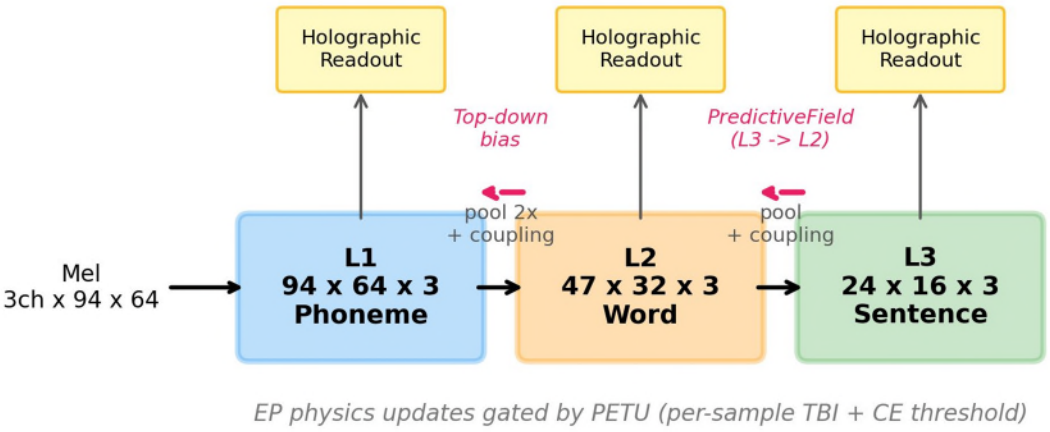


Figure 1. Three-layer FICU-CRFM architecture. Mel spectrogram input (3 channels × 94 × 64, encoding mel power, delta, and delta-delta) drives Layer 1 (94 × 64 × 3, phoneme recognition, 40 classes). Layer 1's settled complex field is spatially coarse-grained by 2× average pooling and projected through a learned 3 × 3 channel coupling matrix to drive Layer 2 (47 × 32 × 3, word recognition, 501 classes). The same operation produces Layer 3's drive from Layer 2 (24 × 16 × 3, sentence-type recognition, 3 classes). Each layer supports a holographic matched-filter readout that projects the settled field onto learned class templates. During cascade training, the PredictiveField module transmits top-down predictions from Layer 3 to Layer 2 by setting Layer 2's initial field state on each trial, and an additional top-down bias modulates Layer 1's drive from Layer 2's settled state. Equilibrium Propagation updates to field parameters at each layer are gated by the Prediction-Error-Triggered Update rule, which restricts parameter evolution to samples with high classification uncertainty or low phase coherence.

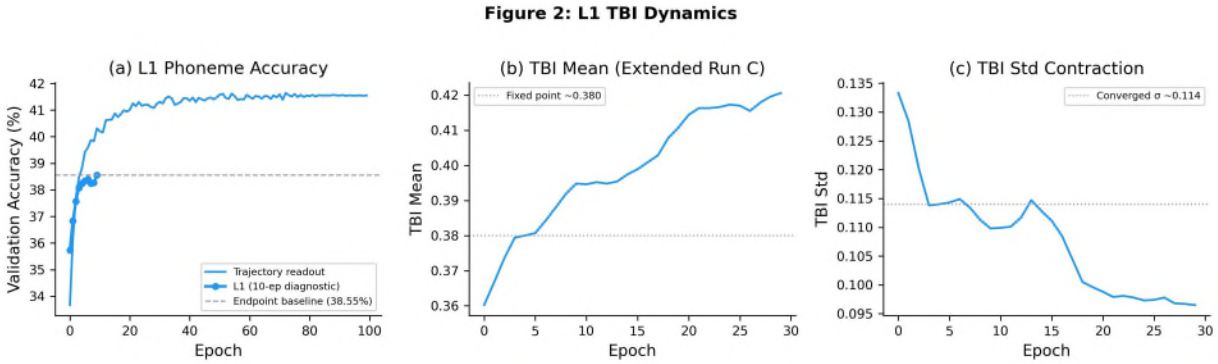


Figure 2. Layer 1 dynamics during single-layer training. (a) Layer 1 phoneme validation accuracy over 100 epochs of training on TIMIT. Filled circles show the initial 10-epoch diagnostic run with standard (endpoint-only) holographic readout, reaching 38.55% at epoch 9. Blue curve shows extended training with the three-tap trajectory readout (settling times 8, 16, and 24 out of 24 settling steps), reaching 41.69% peak at epoch 55 in the full-retraining protocol. Dashed line marks the endpoint-only baseline. (b) Mean Temporal Binding Index across training samples over 30 epochs of extended Run C. The system enters the basin of its attractor near TBI ≈ 0.380 by epoch 5 and subsequently drifts slowly upward, reaching approximately 0.420 by epoch 30. Mean TBI does not approach unity as the original CRFM prediction anticipated. (c) Standard deviation of TBI across training samples over the same 30 epochs. The across-sample distribution of binding indices contracts monotonically from approximately 0.135 toward 0.097, consistent with the interpretation that single-layer automaticity manifests as phase population homogenization rather than a rise in mean coupling strength.

Figure 3: Amplitude Attenuation Correction

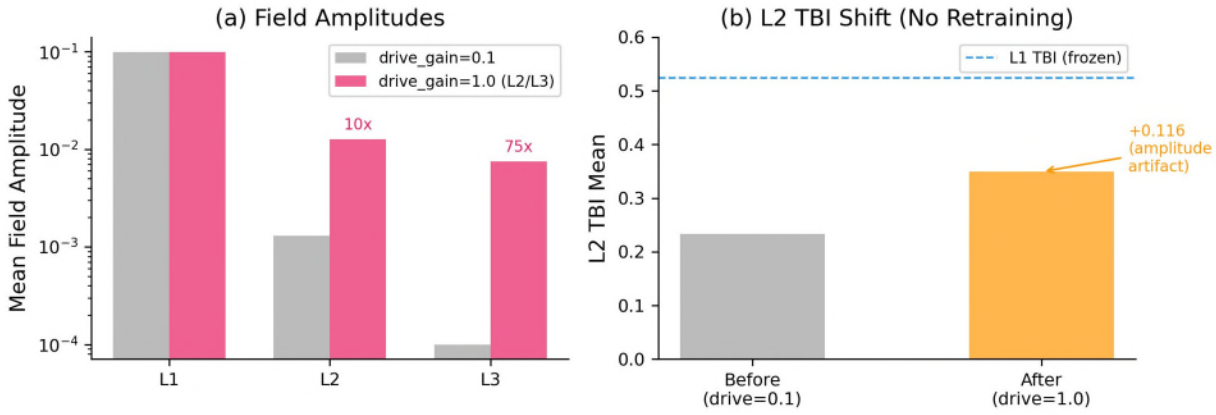


Figure 3. Amplitude attenuation correction. (a) Mean field amplitude |Z| at Layer 1, Layer 2, and Layer 3 before (gray) and after (pink) correcting the drive-gain and inter-layer coupling initialization. The original implementation (drive_gain = 0.1 at all layers; L2→L3 coupling initialized at 0.1·I) produced approximately two orders of magnitude of amplitude attenuation per layer (|Z|_L1 ≈ 0.099, |Z|_L2 ≈ 0.0013, |Z|_L3 ≈ 0.0001). Correcting L2 and L3 drive gains to 1.0 and the L2→L3 coupling initialization to 1.0·I restored physical amplitudes (|Z|_L2 ≈ 0.0126, |Z|_L3 ≈ 0.0075), an increase of 10× at L2 and 75× at L3. Vertical axis is logarithmic. (b) Layer 2 mean TBI before and after amplitude correction, with no additional training. The correction alone raises L2 TBI from 0.233 to 0.349, an increase of 0.116 nats — approximately 40% of the L1–L2 TBI gap observed in the original implementation. Dashed line marks the frozen Layer 1 TBI reference (0.524).

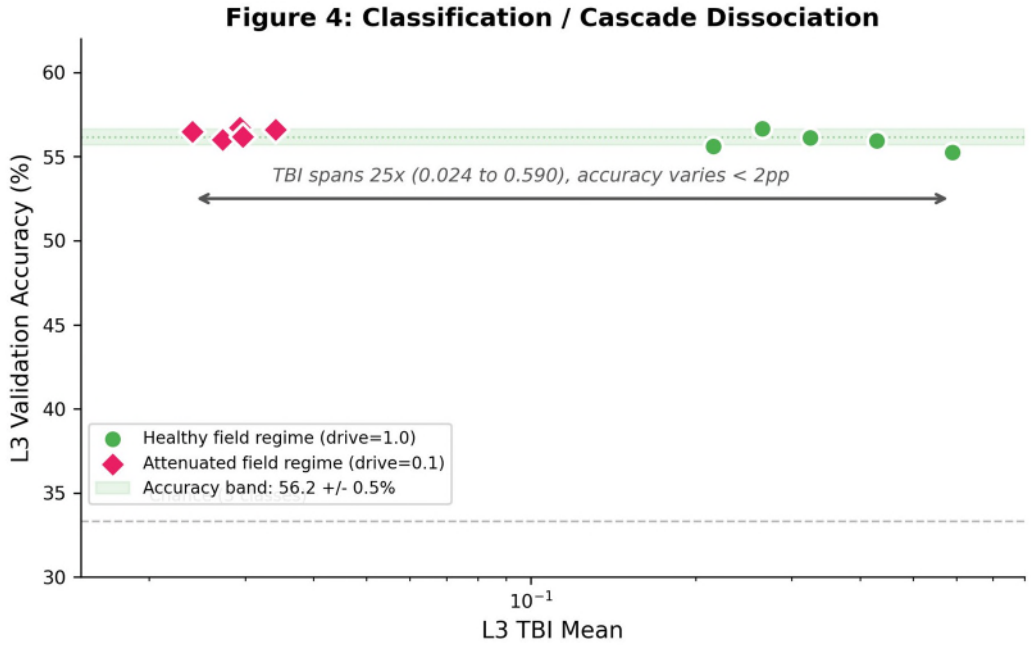


Figure 4. Paired dissociation between classification and cascade transmission (five-seed control). (a) Layer 2 TBI trajectory over 20 cascade epochs in the coherent-Layer 3 condition versus the attenuated-Layer 3 condition, mean ± SD bands across five independent seeds, with the Layer 1 frozen reference (TBI = 0.524) and the Layer 2 pre-cascade reference (TBI = 0.349) marked. The coherent condition closes 95.1% ± 1.5% of the gap to Layer 1 (ΔL2 = +0.166 ± 0.003); the attenuated condition closes only 30.1% ± 0.5% (ΔL2 = +0.052 ± 0.001). (b) PredictiveField Layer 2 prediction-error magnitude over the same window, both conditions, mean ± SD across seeds. Panel (b) should be read as prediction effort against a structured target: the low, flat error in the attenuated condition reflects an absence of structured Layer 3 source signal to model, not higher-fidelity transmission. Across both conditions, Layer 3 classification accuracy is essentially unchanged — in fact marginally higher in the attenuated condition (57.5% ± 1.0% versus 53.2% ± 1.5%) — establishing that prediction transmission scales with source-field phase coherence rather than with task-discriminative capability.

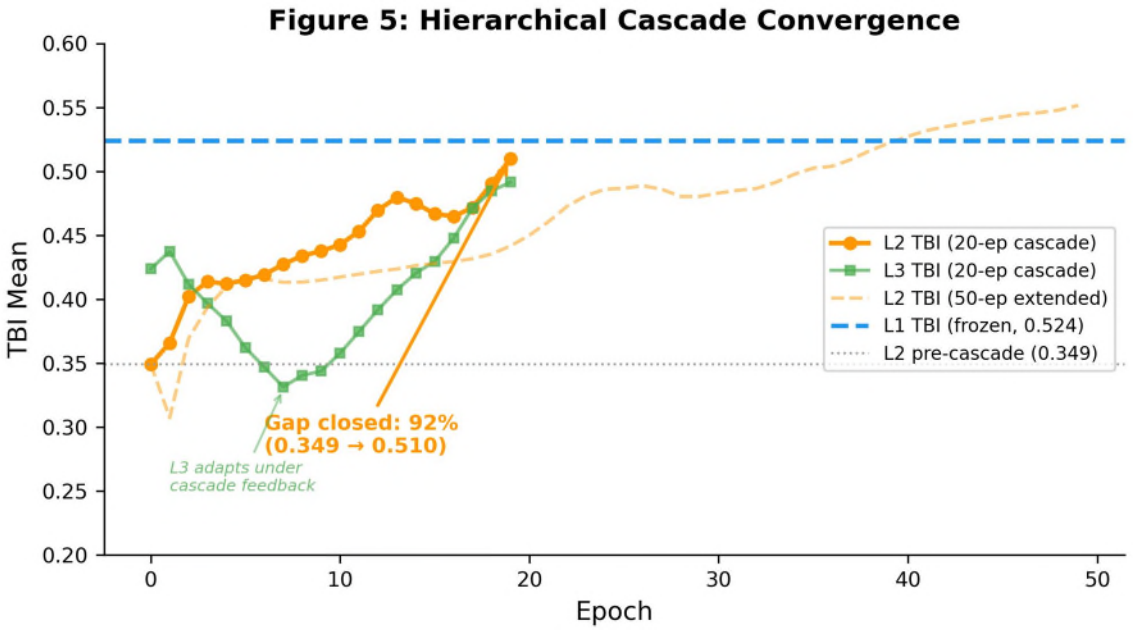


Figure 5. Hierarchical cascade convergence (five-seed replication). Mean TBI trajectories during cascade training, shown as mean ± SD bands across five independent random seeds with individual seed traces lightly overlaid. With Layer 1 frozen at TBI = 0.524 (dashed line) and Layer 3 in the coherent-physics regime (β = 0.05), top-down predictions from Layer 3 — implemented as a learned boundary condition on Layer 2's settling dynamics, with coupling coefficient λ scheduled linearly from 0 to λ_max = 5.0 — drove Layer 2 TBI from 0.349 (pre-cascade value) to 0.515 ± 0.003 over 20 training epochs, closing 95.1% ± 1.5% of the 0.175-nat gap to Layer 1 (per-seed range 0.512–0.519; gap closure 93.1%–97.2%). Layer 3 TBI traced a non-monotonic path during this window, recovering to 0.500 ± 0.009 by epoch 19; this reflects concurrent EP updates to Layer 3 physics under the influence of its own cascade feedback. The 50-epoch extended run shows Layer 2 TBI continuing to climb modestly past Layer 1's frozen value. Bands are mean ± SD over five seeds.

Figure 6: Three-Layer TBI at Cascade Convergence

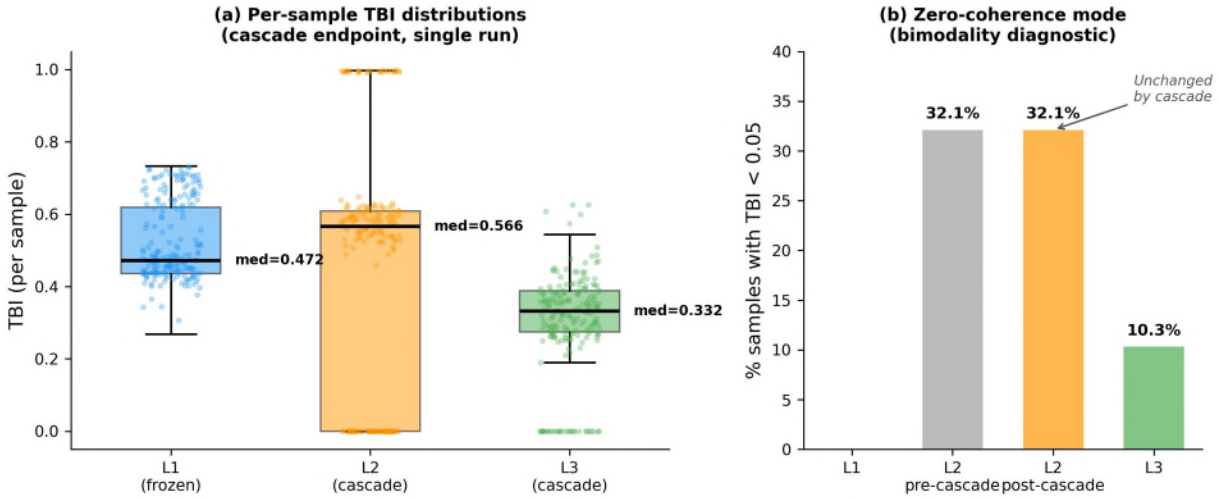


Figure 6. Three-layer TBI distributions at cascade convergence (five-seed pooled). (a) Per-sample TBI distributions at Layer 1 (frozen), Layer 2 (post-cascade), and Layer 3 (post-cascade) at the end of 20-epoch cascade training (λ_max = 5.0), shown as box plots (median, interquartile range, whiskers) overlaid with jittered individual samples, pooled across five seeds (n = 1280 test samples per seed). Layer 1 is unimodal and compact (median 0.472, IQR [0.436, 0.619], skewness +0.60). Layer 2 is bimodal: a zero-coherence mode pinned near TBI ≈ 0 comprising 38.4% of samples, and a high-coherence mode centered above 0.5 (high-coherence-mode median 0.556). Layer 3 is tight and left-skewed (median 0.332, IQR [0.274, 0.388], skewness −1.04). (b) Fraction of Layer 2 samples with TBI < 0.05, pre-cascade versus post-cascade: bit-identical at 38.4% ± 0.0% across all five seeds in both conditions. The cascade raises the high-coherence-mode median from 0.393 to 0.556 by sharpening that mode but does not reduce the zero-coherence-mode fraction. The cascade's effect is selective refinement of representable inputs, not rescue of structurally unrepresentable ones; the invariance of the zero-coherence fraction across seeds indicates it is a structural property of Layer 2 physics rather than a stochastic training outcome.

Tables

Table 1. Readout ablation at Layer 1 (phoneme recognition, 40 classes, TIMIT). All variants resume from the same trained L1 physics checkpoint; only the readout extraction and classification scheme differ. The holographic linear readout on the 162-dimensional pooled phase-feature vector (top row, bold) is the readout used throughout the paper. Coherent matched-filter variants computing score_k = $|\langle Z, T_k \rangle|^2$ on either the raw complex field or pooled complex features collapse sign information and fail to exceed chance on 40-class phoneme recognition. The whitening variant does not recover the lost information.

Readout	Feature space	Dim	Val Acc (%)
Holographic linear	**Pooled phase features**	**162**	**38.55**
Coherent matched filter	Raw complex field	18,048	17.5
Coherent matched filter	Pooled complex features	81 (complex)	2.6
Whitened coherent	Pooled complex features	81 (complex)	3.0
Chance level (40 classes)			*2.5*

Table 2. Cross-architecture comparison on TIMIT word recognition (501 classes). Both models receive identical mel-spectrogram input features and are trained to match vocabulary on the TIMIT training partition. The MLP baseline is a three-layer fully connected network (512 → 256 → 501) trained with Adam, cosine-annealed learning rate, cross-entropy loss, and 100 epochs. The FICU L2 model uses continuous Landau-Ginzburg wave-field dynamics trained by Equilibrium Propagation with the holographic delta-rule readout. The wave-field architecture reaches 38.85% validation accuracy in 15 epochs without any backpropagation through its physical dynamics. Chance level is 1/501 ≈ 0.2%.

Model	Training rule	Epochs	Val Acc (%)	× Chance
MLP (512 → 256 → 501)	Backpropagation (Adam)	100	25.3	127×
FICU L2 (holographic)	**EP + delta rule**	**15**	**38.85**	**195×**
Improvement			*+13.5 pp*	

Methods

Architecture overview

The FICU architecture is a three-layer stack of underdamped Landau-Ginzburg wave fields. Each layer is a complex scalar field Z_c(x, y, t) defined on a discrete two-dimensional spatial grid with three frequency channels indexed by c ∈ {0, 1, 2}. The three layers operate at progressively coarser spatial resolutions (L1: 94 × 64, L2: 47 × 32, L3: 24 × 16) and receive input from one another through learned channel-coupling matrices and, during cascade training, through top-down boundary conditions on the settling dynamics. The field dynamics, Equilibrium Propagation (EP) update rules, and phase-sensitive readout are inherited from the single-layer architecture detailed in companion paper [14]; this Methods section documents only what is novel or specific to the three-layer system and to the experiments reported here.

Field dynamics and settling

Each layer evolves according to the underdamped Landau-Ginzburg equation specified in [14]:

$$\partial Z_c/\partial t = -\gamma_c(x, y)\, Z_c + i\omega_{\mathrm{eff},c}\, Z_c - \beta_c\, |Z_c|^2 Z_c + D_c(x, y) \cdot (K_{\mathrm{DoG}} * Z_c) + \kappa_{\mathrm{ev}} \Sigma_{\{c' \neq c\}} \exp(-|c - c'|) \cdot Z_{\{c'\}}(x, y, t) + S_c(x, y, t)$$

where $\gamma_c(x, y)$ *is a spatially structured damping field (initialized at 0.025, 0.050, 0.100 for the three channels and passed through a sigmoid activation bounded at 0.2);* $\omega_{\mathrm{eff},c}$ is a channel-specific rotation frequency (initialized at 0.800, 1.265, 2.000 rad/step, logspaced, with a minimum floor of 0.1); β_c *is a channel-specific saturation coefficient (0.060, 0.110, 0.210 after the abs(·) + 0.01 activation);* $D_c(x, y)$ is a spatially structured diffusion field (initialized at 0.100, passed through a sigmoid bounded at 0.5); K_{DoG} *is a Mexican Hat lateral inhibition kernel;* $\kappa_{\mathrm{ev}} = 0.01$ is a fixed evanescent coupling constant; and S_c *is the drive signal from the sensory input (L1) or from the layer below (L2, L3). Two additional scalar parameters govern the nonlinear phase dynamics:* $\chi_{\mathrm{spm}} = 0.15$ (self-phase modulation) and $\chi_{xpm} = 0.10$ *(cross-phase modulation), each passed through bounded sigmoid activations. Parameters* $\gamma(x, y)$, $D(x, y)$, *and the scalar quantities* χ_{spm}, χ_{xpm}, κ_{ev}, and the per-channel β are learned by EP; parameters ω_eff and the lateral-inhibition kernel are fixed.

The Mexican Hat kernel differs per layer to match the decreasing spatial resolution of each layer: L1 uses a 9 × 9 kernel with excitatory σ = 1.0 and inhibitory σ = 3.0; L2 uses 7 × 7 with $\sigma_{exc} = 1.5$ *and* $\sigma_{\mathrm{inh}} = 4.0$; L3 uses 5 × 5 with $\sigma_{exc} = 1.0$ *and* $\sigma_{\mathrm{inh}} = 2.5$.

The field is integrated with 24 forward Euler settling steps at dt = 0.07 (total settling time 1.68 dimensionless units), identical at all three layers. After each field update, an amplitude clamp is applied

element-wise: scale = clamp(10.0 / |Z|, max = 1.0), preventing numerical divergence in the nonlinear saturation term under transient high-amplitude excursions.

Sensory front-end

L1 receives 64-channel Mel spectrograms extracted from the TIMIT audio (sampled at 16 kHz, 25-ms Hann-windowed frames hopped by 10 ms, 64 Slaney-normalized Mel filterbanks), together with first and second temporal derivatives (delta and delta-delta). Per-utterance mean/variance normalization is applied separately to each of the three channels (Mel, delta, delta-delta). The resulting [3, 94, 64] tensor is used directly as the real part of the drive signal S, with S imaginary set to zero: S*r = features, S*i = 0, and the drive is then scaled by a layer-specific drive_gain scalar before injection into the field dynamics. The Mel filterbank and temporal derivative operators are implemented in pure PyTorch to avoid a torchaudio dependency (the NGC containers used for training do not include torchaudio).

Inter-layer projection

L2 receives input from L1's settled field state through two field operations. First, 2 × 2 spatial average pooling reduces L1's [3, 94, 64] field to [3, 47, 32]; this is the renormalization-group coarse-graining operation that preserves long-wavelength modes while discarding high-spatial-frequency structure. Second, a learned 3 × 3 channel coupling matrix W*coupling*L1 (initialized at 0.1·I) mixes the three frequency channels: S*L2 = einsum('bchw, cd → bdhw', Z*L1*pooled, W*coupling_L1^T). The resulting [3, 47, 32] tensor serves as the drive signal S at L2.

L3 receives input from L2's settled field state through the analogous construction: adaptive spatial average pooling from [3, 47, 32] to [3, 24, 16], followed by a learned channel coupling matrix W*coupling*L2 (initialized at 1.0·I). The initial value of W*coupling*L2 was changed from 0.1·I to 1.0·I during the amplitude-attenuation correction described in Results; all cascade experiments reported here use 1.0·I.

The drive*gain scalar is fixed per layer: drive*gain*L1 = 0.1 (unchanged throughout), drive*gain*L2 and drive*gain*L3 = 1.0 for cascade experiments. The single-layer L2 readout accuracy of 38.85% reported in Results was measured under drive*gain_L2 = 0.1; this regime difference is discussed in the Results methodological note.

Temporal Binding Index

The Temporal Binding Index (TBI) is computed from the pair (Z*free, Z*nudge), where Z*free is the field state after 24 settling steps with no task-supervised perturbation and Z*nudge is the field state after 24 settling steps with the EP nudge applied (see below). At each spatial location (i, j) and channel c, the local phasor encodes the phase difference between the perturbed and free states:

$\phi(b, c, i, j) = Z_{nudge(b, c, i, j)} \cdot Z_{free}^*(b, c, i, j) / [|Z_{nudge(b, c, i, j)}| \cdot |Z_{free}(b, c, i, j)|]$

where * denotes complex conjugation. This is a unit complex number at every location, encoding the local phase rotation induced by the supervisory nudge. The spatial mean of these phasors at each channel,

$\rho(b, c) = \langle \phi(b, c, i, j) \rangle_{i, j}$

is a Kuramoto order parameter: its magnitude $|\rho(b, c)| \in [0, 1]$ measures the spatial coherence of the field's phase response to perturbation, taking the value 1 if all locations respond with identical phase rotation and 0 if the responses are phase-scrambled across space. TBI for sample b is the mean of this coherence magnitude across the three channels:

$TBI(b) = \langle |\rho(b, c)| \rangle_c$

All reported TBI statistics (mean, standard deviation) are within-epoch statistics computed across the per-sample TBI values of all training samples (approximately 177,000 values at L1 per epoch, approximately 40,000 at L2, approximately 4,600 at L3).

Phase-feature extraction and holographic readout

At each layer, classification is performed by a holographic matched filter operating on a pooled phase-feature representation of the settled field. After 2 × 2 spatial average pooling of the settled complex field, six scalar features are extracted at each location: the real part Re(Z), the imaginary part Im(Z), the magnitude |Z|, the phase cosine cos θ, the phase sine sin θ, and the local amplitude gradient R computed by a Sobel filter on |Z|. The resulting 6-channel representation is flattened and subject to an additional learned phase-feature pooling stage that produces a 162-dimensional vector per layer (identical feature dimension at L1, L2, and L3 by construction).

The holographic readout projects this 162-D vector onto a learned K-way template matrix W to produce class logits, where K is the number of output classes at the layer. Template updates follow a pure delta rule: for each batch, the prediction error error = target*one*hot − softmax(logits) is used to compute the weight update $\Delta W = (error^T \cdot features_{normalized}) / B - \lambda_{wd} \cdot W$, where B is the batch size and $\lambda_{wd} = 0.01$ is an L2 weight decay coefficient. Features are normalized against running Welford statistics (mean and variance) maintained across training samples. The delta rule is deterministic (no momentum, no stochastic optimizer state); the readout learning rate is fixed at 0.005 during standard training and follows a cosine schedule from 10^{-3} to 10^{-5} during the 100-epoch extended readout-only phase at L1. No backpropagation is applied to the template matrix.

Equilibrium Propagation update

Field parameters are updated by Equilibrium Propagation [8] following the form detailed in [14]. The free phase settles the field to Z*free; the nudge phase settles the field with an additional force term derived from the task loss: nudge*flat = error · W, where error is the softmax classification error and W is the current readout template matrix. This nudge is reshaped into field space via the inverse of the phase-feature extraction pipeline and injected as a small perturbation force during the nudge-phase settling. The nudge amplitude is controlled by a scalar β*nudge (not to be confused with the LG saturation coefficient ϐ*c).

For each learnable parameter θ, the EP update is:

Δθ = lr*physics · clamp(−(1/ϐ*nudge) · [O*nudge(ϑ) − O*free(θ)], −0.01, +0.01)

where O*free(ϑ) and O*nudge(θ) are the observable quantities (e.g., energy contributions, phase moments) that contribute to the loss through parameter θ, computed on Z*free and Z*nudge respectively. The learning rate lr_physics = 0.001 is constant across all three layers and all training runs. No momentum or adaptive optimizer state is maintained; updates are applied directly in-place with the per-parameter clamp [−0.01, +0.01] preventing single-step divergence under transient high-error batches.

Prediction-Error-Triggered Updates (PETU)

EP parameter updates at each layer are gated on a per-sample basis by a two-condition criterion: an update contributes to the layer's EP step only if the sample's prediction error exceeds an uncertainty threshold **or** its TBI falls below a coherence floor. Formally, for sample b at layer ℓ, the update mask is:

m*ℓ(b) = [CE(b) > 0.5 · ln(K*ℓ)] ∨ [TBI*ℓ(b) < c*floor,ℓ]

where CE(b) is the cross-entropy of sample b under the current readout, K*ℓ is the number of output classes at layer ℓ, 0.5 · ln(K) is the entropy of a maximally uncertain classifier (yielding thresholds 1.84 nats at L1, 3.11 nats at L2, and 0.55 nats at L3), and c*floor,ℓ is a layer-specific coherence floor (0.45 at L1, 0.35 at L2 and L3). Samples satisfying either condition contribute their EP observable deltas to the layer's parameter update; samples failing both are treated as locally converged and masked out. The fraction of samples passing the gate on each epoch is recorded as the "update fraction" reported in Results.

LevelGate

LevelGate is a training-phase control that freezes a layer's physics parameters once that layer achieves a target validation accuracy sustained over a patience window. The threshold is 0.60 at L1 and 0.35 at L2; it is checked at the end of each validation pass. In the experiments reported here, LevelGate triggered immediately on the original 3-epoch L2 run (at the 35% threshold). For the extended L2 training and all L3

training, LevelGate was disabled via a --no_gate flag so that we could study the full TBI trajectory and its cascade behavior without confounding from gate activation.

PredictiveField

The PredictiveField is a learned linear map from L3's settled field state to an initial field state for L2. L3's settled complex field Z*L3 ∈ C^{3 × 24 × 16} is flattened to a 1152-dimensional vector, projected through a learned matrix W*pf ∈ C^{1152 × 4512}, and reshaped to an initial field of shape C^{3 × 47 × 32} suitable for injection as L2's starting condition. W_pf is initialized at N(0, 0.001²) and updated by a delta rule on the difference between the actual L2 settled field (on a free pass) and the predicted L2 field. No nonlinearity is applied; the map is purely linear and contains no gradient-trained parameters (no backpropagation through the prediction).

During cascade training, the PredictiveField's output is scaled by a coupling coefficient λ(t) that is scheduled linearly from 0 to λ*max over the training epochs (λ*max = 5.0 in all cascade runs reported here). At λ = 0 the L2 field begins its settling from zero initial condition, exactly as in single-layer L2 training. At λ = λ_max, L3's prediction replaces the zero initial condition entirely. Intermediate values of λ produce a weighted top-down boundary condition. The PredictiveField is therefore a field operation that sets the boundary condition for L2's settling dynamics, distinct from a Hamiltonian interaction term between layers.

A second top-down pathway from L2 to L1 is implemented symmetrically: L2's settled field produces a learned bias (`l2*to*l1*topdown*bias`) that modifies L1's drive during a re-settle pass. This L2 → L1 pathway is also scheduled by λ(t) and operates during cascade training.

Cascade training protocol

The three-layer cascade training loop uses a two-pass settling schedule on each batch, followed by an EP nudge phase. On each batch:

Feed-forward pass. Layer 1 settles freely from its sensory input: Z*L1(0). Layer 2 settles from Z*L1(0) as its drive: Z*L2(0). Layer 3 settles from Z*L2(0) as its drive: Z*L3*free.

Top-down pass. The PredictiveField computes an L2 prediction from Z*L3*free and scales it by λ: pred*L2. The L2 → L1 pathway computes a top-down bias for L1 from Z*L2(0) and scales it by λ: bias*L1. Layer 1 re-settles with bias*L1 added to its drive, producing Z*L1*td. Layer 2 re-settles from Z*L1*td, with pred*L2 providing the initial field state in place of zero: Z*L2*td. Layer 3 re-settles from Z*L2*td: Z*L3*tdfree.

EP nudge phase. The nudge is applied to L3's readout only (L1 is frozen after its pretraining phase, and L2's readout is not updated during cascade training to avoid drift under shifted drive-gain statistics). L3

settles once more with the nudge force applied: Z*L3*nudge. TBI is computed from the pair (Z*L3*td*free, Z*L3*nudge). PETU masks samples; the L3 physics parameters are updated by EP; the L3 readout template is updated by the delta rule; the PredictiveField is updated by its delta rule on the (Z*L3*td*free*, Z*L2*td) pair.

L1 physics parameters are frozen throughout cascade training (they were trained in the isolated L1 phase and are held fixed with requires*grad = False). L2 physics parameters are loaded from the drive*gain = 1.0 checkpoint but are not updated during cascade training; L2 serves as a fixed intermediate layer whose field dynamics are influenced by the top-down boundary condition but whose parameters do not evolve. This design isolates the cascade effect to the interaction between L3's learned predictions and L2's (fixed) field dynamics, preventing confounding from simultaneous physics evolution at multiple layers.

The cascade was run for 20 epochs (primary result) and for 50 epochs (extended run reported as a secondary observation). In both runs λ was scheduled linearly from 0 to λ_max = 5.0 across the full training window.

Training data and splits

Experiments use the standard Linguistic Data Consortium TIMIT train/test partition: 4,620 training utterances from 462 speakers and 1,680 test utterances from 168 speakers, with no speaker overlap between splits. The test partition is used directly for validation; no separate held-out validation split was created. Training data are sampled from the TRAIN directory at the segment level according to the task: L1 training samples are phoneme segments extracted from the phoneme-level annotation files (.PHN), yielding approximately 177,000 training segments per epoch; L2 training samples are word segments extracted from the word-level annotation files (.WRD), yielding approximately 39,800 training segments per epoch; L3 training samples are full utterances labeled by the three sentence-type prefixes (SA, SI, SX from the TIMIT filename convention), yielding 4,620 training samples per epoch.

The L2 vocabulary of 501 classes is constructed by counting word frequencies across all training utterances and retaining the 500 most frequent words, plus one ⟨unk⟩ token for all remaining words. The L1 phoneme inventory contains 40 classes (the standard TIMIT 39 phonemes plus the silence marker h#). The L3 sentence-type classes correspond to the three sentence design categories of the TIMIT corpus: dialect (SA), diverse (SI), and compact (SX).

Baselines

MLP backpropagation baseline

A three-layer fully connected network (512 → 256 → 501) was trained on the same L2 word-recognition task used to evaluate FICU L2. The input features are the same Mel-spectrogram front-end as L1 (flattened

to a single vector per word segment). Training used Adam with default $\beta_1 = 0.9$, $\beta_2 = 0.999$, cosine-annealed learning rate from 10^{-3} to 10^{-5}, cross-entropy loss, and 100 epochs. This baseline reaches 25.3% validation accuracy on the standard TIMIT test split.

Readout ablations at L1

Three alternative readouts were evaluated on the L1 phoneme task, each resuming from the same trained L1 physics checkpoint: (i) a coherent matched filter on the raw 18,048-dimensional field computing score$_k = 0.1 \cdot |\langle Z, T_k \rangle|^2$ with per-class complex templates trained by exponential moving average on unit-norm-clipped nudged states (17.5% validation accuracy); (ii) a coherent variant on an 81-dimensional pooled complex feature representation (2.6% validation accuracy, at chance for 40 classes); (iii) a whitening variant adding per-feature running normalization to (ii) (3.0% validation accuracy). All three variants are compared against the holographic linear readout on the 162-dimensional pooled phase-feature vector used throughout this paper (38.55% validation accuracy, epoch 9). These ablations confirm that phase-sign-preserving linear readout is structurally necessary for decoding the underdamped LG field's phase structure.

Hardware and reproducibility

All experiments were run on a single NVIDIA GB10 GPU (Blackwell-class, compute capability sm_121) with CUDA 13.0 and driver 580.142, using the nvcr.io/nvidia/pytorch:25.11-py3 container (PyTorch 2.10.0a0 nightly). Total compute for all experiments reported here is approximately 15–20 GPU-hours. The L1 extended 100-epoch run is the single most expensive component at approximately 4.2 hours wall-clock; the cascade runs (20-epoch and 50-epoch) together consume approximately 20 minutes.

The cascade experiments — both the coherent-Layer 3 cascade and the attenuated-Layer 3 paired control — were each run five times from independent random seeds (PyTorch, NumPy, and Python RNGs all seeded per run); cascade quantities are reported throughout as mean ± standard deviation across these five seeds. The single-layer L1 training runs, the readout ablations, and the L3 calibration sweeps were each run once, and quantitative values drawn from them should be read as single-run measurements. The frozen-L1 sanity value was bit-identical across all five cascade seeds, confirming that L1 parameters were genuinely held fixed during cascade training. cuDNN nondeterminism was left enabled, so seed replication establishes statistical reproducibility rather than bitwise reproducibility.

Code and trained checkpoints will be released at publication under an open license; current development code is not yet public.

Software

All components are implemented in PyTorch. The Mel filterbank, temporal derivatives, complex field operations, Mexican Hat kernel, EP update rule, holographic readout with Welford normalization, TBI computation, and cascade training loop are all implemented without external neural-network-library dependencies beyond standard PyTorch operations. No torchaudio dependency is used (reimplementation was necessary due to NGC container incompatibility with the target GPU architecture).

Discussion

Summary of findings

We set out to determine whether hierarchical phase coordination — a measurable change in the phase organization of each layer's dynamics as it becomes competent, accelerated at lower layers by top-down signals from higher ones — emerges in a multi-layer continuous wave-field substrate trained by local learning rules alone. This is the architectural content of the Chrono-Resonant Field Model's account of hierarchical automaticity, examined here as a question about a physically realizable computing substrate rather than as a claim about biological cortex. Four principal findings bear on it.

First, single-layer automaticity manifests as contraction of the Temporal Binding Index distribution across samples rather than as a rise in mean TBI. Layer 1 converges to a fixed point at TBI ≈ 0.380 with standard deviation contracting toward $\sigma \approx 0.114$, and the mean does not rise toward the theoretical maximum even under extended training. Automaticity in this architecture is phase population homogenization: the field's response to supervisory perturbation becomes consistent across inputs rather than stronger on any individual input.

Second, top-down feedback from Layer 3 drives the population mean of Layer 2 TBI from 0.349 to 0.515 ± 0.003 over 20 training epochs across five independent seeds, closing 95.1% ± 1.5% of the gap to Layer 1's frozen value. All three layer means converge to within 0.024 nats of one another at the final epoch, the result replicates tightly across seeds (per-seed gap closure 93.1%–97.2%), and the cascade is sustained in an extended 50-epoch run with modest overshoot at Layer 2. Sample-level inspection qualifies this population-level result: Layer 2's TBI distribution is bimodal, and the cascade acts selectively on the coherent sub-population, sharpening it (median 0.393 → 0.556) while leaving a stable zero-coherence mode essentially untouched — a mode whose fraction was bit-identical at 38.4% ± 0.0% across all five seeds, both pre- and post-cascade, indicating that it is a structural property of Layer 2's field dynamics rather than a stochastic training outcome. Top-down predictions refine the representable sub-population rather than rescuing structurally unrecoverable inputs. This behavior is consistent with the bistability theoretically expected for Kuramoto-type systems with higher-order nonlinear coupling [16].

Third, classification and cascade dissociate. In a paired five-seed control, an attenuated-coherence Layer 3 (TBI = 0.116 ± 0.002) classified sentence type marginally better than a coherent Layer 3 (TBI = 0.500 ± 0.009) — 57.5% ± 1.0% versus 53.2% ± 1.5%, with non-overlapping cross-seed bands — yet transmitted roughly three-fold weaker top-down drive to Layer 2 (ΔL2 TBI = +0.052 ± 0.001 versus +0.166 ± 0.003; 30.1% versus 95.1% gap closure). The field configuration that classifies better transmits worse. Healthy phase coherence is therefore necessary for hierarchical prediction transmission but not for pattern recognition at a linear readout, and the cascade result cannot be an epiphenomenon of improved feature

quality. Phase-coherent field dynamics transmit functional predictive information across layers in a way that phase-incoherent dynamics cannot, even when the incoherent dynamics support equal or better classification at the receiving layer.

Fourth, a supplementary readout experiment found that reading the field at multiple settling times (a sparse three-tap sampling of the trajectory rather than endpoint-only reading) raises Layer 1 phoneme classification from 38.76% to 41.69%, a gain that is robust across frozen-physics and co-adapted-physics training regimes.

The architectural implications of the weakly dissipative regime

The trajectory readout finding has a theoretical interpretation that goes beyond "more features help." The Landau-Ginzburg field in this architecture operates in an underdamped regime ($\gamma/\omega \approx 0.03$ for the slowest channel, $Q \approx 32$). In the purely overdamped limit of standard energy-based models, the settling trajectory is strictly monotonic descent along the local free-energy gradient, and every intermediate state is strictly less informative than its successor. In that regime, no trajectory sampling can outperform endpoint sampling.

In the underdamped regime by contrast, the equation of motion takes the damped-wave form $\tau\, \partial^2\psi/\partial t^2 + \eta\, \partial\psi/\partial t = -\delta F/\delta\psi^*$, in which the inertial term preserves phase information across the settling window. The field passes through coherent wave interference patterns before the dissipation term settles it to its attractor, and these transient patterns carry task-relevant structure that the attractor alone does not represent. The three-snapshot trajectory readout accesses a coarse skeleton of this transient computation; the +2.93pp improvement we measure reflects this access.

This observation makes a specific prediction for follow-up work. The snapshot sampling we used is Nyquist-insufficient for the faster field channels ($\omega \approx 2.1$ rad/step completes a full rotation in roughly three settling steps). Dense temporal sampling, spectral decomposition, or integrated quantities across the full trajectory should yield substantially larger accuracy gains. We treat this as an invitation to a dedicated follow-up paper rather than as a claim that the three-snapshot readout has extracted all available trajectory information.

The broader point concerns what the regime choice enables computationally. A purely conservative field never settles and cannot be trained by Equilibrium Propagation, which requires a well-defined fixed point. A purely dissipative field settles but computes nothing during the settling — the trajectory is memoryless gradient descent. The weakly dissipative regime places the architecture in the narrow band where persistent oscillatory propagation during the transient coexists with asymptotic stability at the attractor. This is the band in which the brain operates, as measured by scalp EEG and intracranial recordings, and it is the band in which computational richness and learnability coincide.

The classification/cascade dissociation points to attention

The dissociation between classification and prediction transmission has a mechanistic interpretation that the present results constrain more sharply than a binary contrast would. Classification accuracy at Layer 3 was essentially flat — in fact marginally higher in the attenuated regime — across a roughly four-fold range of phase coherence, while top-down transmission to Layer 2 fell roughly three-fold over the same range. A linear matched-filter readout needs only that the template matrix separate class-conditional feature distributions, a task that tolerates amplitude-dominant, low-coherence dynamics; it is a bulk operation in which every spatial location contributes to every class logit. Prediction transmission is different in kind. It requires that the field state at one layer carry structure that can be spatially decoded into a meaningful initial condition for the next, and that structure — phase boundaries, coherent subdomains, interference patterns — degrades continuously as phase coherence falls. The graded transmission we observe is the signature of that continuous degradation: an attenuated Layer 3 is not structureless noise, and it transmits a correspondingly attenuated but non-zero drive.

We propose that the property distinguishing these two regimes is the architecture's ability to route information on the basis of spatial contrast structure within the field. Classification is a bulk readout that requires no routing. Prediction transmission requires routing, because the predictive content is carried in specific spatial structures that must be preserved through the inter-layer coupling — and the fidelity of that routing scales with how much spatial contrast structure the source field actually contains. The field configuration that classifies marginally better while transmitting markedly worse is the clean illustration: better class separability at a bulk readout, less spatial structure available to route.

This framing points toward a generalization beyond the present architecture. If the amount of spatial contrast structure a field carries is what governs how well it can drive a cascade, then an attention mechanism keyed to spatial contrast should be both natively realizable in the field physics and functionally useful for richer hierarchical computation. Preliminary theoretical work suggests that a secondary, slowly varying scalar field driven by the spatial variance of the primary field's intensity implements exactly this: a dilaton attention field that lowers the local energy barrier at regions of high spatial contrast, preferentially routing primary-field computation through structurally informative regions. The graded dissociation reported here — transmission scaling with source-field coherence rather than gating on it — is the empirical motivation for such a mechanism, and we develop it in ongoing work.

A native-physics roadmap beyond the current architecture

The accuracy levels reported here — 41.69% on Layer 1 phoneme recognition with trajectory readout, 38.85% on Layer 2 word recognition — are competitive with a matched MLP baseline trained by backpropagation (25.3%) but fall well below the accuracy levels that deeper backpropagation-trained

architectures achieve on the same tasks. This gap has a productive interpretation: the current FICU architecture underexploits the computational capacity of its own physics.

Four computational primitives are already present in the Landau-Ginzburg equations used here but have not been deployed in the readout or classification pathways of the present paper.

The cubic saturation term $-\beta|\psi|^2\psi$ in the field equation provides a native, amplitude-regulated nonlinearity with distinct low-amplitude (linear) and high-amplitude (saturated) regimes. The architecture described here uses this term only implicitly, through its effect on the steady-state field configuration. A readout designed to measure which amplitude regime each spatial location occupies would access this nonlinearity directly, providing the classification pathway with a thresholding operation analogous to what rectified linear units provide in backpropagation-trained networks. Unlike rectified linear units, the saturation regime is a physical consequence of the substrate rather than an externally imposed function, and its threshold adapts to the local field amplitude automatically.

Cross-phase modulation between field channels is implemented in the current equations through the χ*xpm coupling term. This term constitutes an element-wise multiplication between field channels — the canonical nonlinear operation that has been shown in recent photonic work to raise accuracy by 10+ percentage points when incorporated as the feature-combining primitive in vertically stacked electro-photonic classifiers. The current architecture uses* χxpm as a weak correction; exploiting it as the primary nonlinear feature-combination mechanism is an architectural variant worth testing.

Phase-synchronization attention is in some sense already present in the TBI itself. The Kuramoto-style spatial order parameter measures coherence of the field's phase response across space. A readout that weights spatial locations by their local phase-coherence density would implement a phase-space analog of softmax attention: regions whose phase response is internally coherent contribute to classification; regions whose phase response is scrambled are suppressed. The computational cost is a single additional local spatial averaging operation.

The dilaton attention field described above represents the most architecturally novel of these four directions. A secondary field ψ_2 or dilaton ϕ operating at a slower timescale and driven by the spatial variance of the primary field's intensity would implement input-dependent, contrast-driven attention that persists across trials to carry context through continuous data streams. The governing equations and training integration with Equilibrium Propagation are consistent with established photonic nonlinearities (thermal self-focusing, photorefractive index modulation), and we are working out the formal details in ongoing work.

None of these four directions requires additional layers, skip connections, or backpropagation-style deep architectures. Each exploits a computational affordance that is already present in the physics we simulate.

The architecture's path to higher accuracy runs through more complete engagement with its own native operations, not through importing the architectural machinery of conventional deep networks.

Implications for physical and photonic implementation

The architectural claims of this paper bear directly on the viability of continuous wave fields as substrates for physically instantiated intelligent computation. Three points are worth making explicit.

Every parameter updated by Equilibrium Propagation in the FICU architecture maps to a physical specification in a photonic implementation: the damping field γ(x, y) is a spatial absorption profile, the diffusion field D(x, y) is a local coupling strength between adjacent spatial modes, the saturation coefficient β is a Kerr nonlinearity magnitude, and the cross-phase modulation coefficient χ_xpm is a four-wave mixing strength. The learning process therefore specifies, in a photonic realization, exactly the fabrication parameters of a photonic chip that implements a trained classifier. The architecture is not a mathematical abstraction awaiting translation to hardware; it is a computational description of hardware whose dynamics have been simulated.

The weakly dissipative regime that supports trajectory-rich computation is the regime in which most natural photonic substrates operate. Silicon nitride microresonators, lithium niobate waveguides, and exciton-polariton condensates all exhibit Q factors in the 10^2–10^4 range, well within the underdamped regime required for phase-preserving settling dynamics. The regime choice that enables computational richness in FICU is not an engineering constraint to be overcome in hardware; it is a natural operating point of the substrates that would implement it.

The four native-physics directions described above map onto specific established nonlinear optical effects. Saturation-regime readout exploits the intensity-dependent refractive index change that is native to any Kerr-nonlinear medium. Cross-phase modulation as primary nonlinearity exploits the same coefficient, applied to pairs of co-propagating wavelengths. Phase-synchronization attention can be measured interferometrically and is native to any coherent optical platform. The dilaton attention field maps onto the thermal or photorefractive nonlinearities that cause the refractive index to change slowly in response to sustained optical intensity — effects that photonic engineers often treat as sources of drift to be compensated rather than as computational resources to be exploited. The FICU architecture reframes all four effects as computational primitives, providing a concrete path by which existing photonic materials can be engineered into trainable intelligent computing substrates.

Limitations

Several caveats concerning the present results deserve explicit acknowledgment.

The cascade experiments were replicated across five independent random seeds and cascade quantities are reported as mean ± standard deviation; the single-layer L1 training, the readout ablations, and the L3 calibration sweeps were each run once, and quantitative values specific to them should be read as single-run measurements rather than estimates with error bars. A fully seed-replicated single-layer training campaign — particularly for the L1 variance-contraction result — remains a natural extension, although the qualitative variance-contraction signature is a structural feature of the architecture that is unlikely to be sensitive to seed.

During cascade training, Layer 2's physics parameters were loaded from a separately trained checkpoint and held fixed. Only Layer 3's physics and the PredictiveField module evolved under the cascade schedule. This design choice isolates the cascade effect from confounding by simultaneous multi-layer physics evolution, but it means that the TBI convergence we report at Layer 2 reflects the convergence of Layer 2's field state *statistics* under varying top-down boundary conditions rather than the convergence of Layer 2's parameters themselves.

The two operating points reported for Layer 2 — drive*gain = 0.1 for readout accuracy and drive*gain = 1.0 for cascade — come from separate experimental regimes and should be interpreted as answering distinct questions rather than as describing a single trained system. The readout accuracy at drive*gain = 0.1 demonstrates representational capability; the cascade at drive*gain = 1.0 demonstrates hierarchical coupling. A fully integrated experiment with a single operating point is a natural extension of the current work.

Layer 2's top-down pathway to Layer 1 (`l2*to*l1*topdown*bias`) is implemented in the architecture and active during cascade training but is not analyzed in detail in the present paper, which focuses on the L3 → L2 cascade as the primary hierarchical test. Understanding whether the L2 → L1 pathway contributes materially to the reported TBI convergence patterns — and particularly to the modest overshoot observed in the 50-epoch run — is another natural extension.

Finally, the current architecture lacks persistent context across input segments. Each training example is processed independently; the field state is effectively reset between samples. For pre-segmented speech corpora this is not a fundamental limitation, but for continuous processing of real-world signals (continuous speech, continuous EEG, continuous video) an architecture with inertial state carrying context across segments is required. The dilaton attention field described above is intended to fill this gap.

Closing

The present paper establishes a trainable continuous wave-field substrate in which hierarchical phase coordination emerges under local learning rules alone. The architecture uses no neurons, no backpropagation through any physical dynamics, and no operations outside a class of primitives —

settling, coarse-graining, matched filtering, threshold-gated updates, boundary-condition coupling — that are realizable in mature photonic substrates. Its central phenomena are reproducible across seeds: the contraction of TBI variance that marks single-layer automaticity, the hierarchical TBI cascade in which top-down boundary conditions drive a lower layer's order parameter toward a higher layer's, and the graded dissociation in which prediction transmission scales with source-field phase coherence rather than with task-discriminative capability. These are architectural results for physically realizable computing substrates; that they are also consistent with the Chrono-Resonant Field Model's account of hierarchical automaticity, which motivated the design, is a point we develop in that model's own context elsewhere.

We have been explicit throughout about both the limitations of the current architecture — modest absolute accuracy, single-layer measurements not yet seed-replicated, no persistent context across input segments — and the directions in which the architecture's own physics provides unexplored computational capacity. The FICU line of work is not a candidate physical implementation of existing neural network architectures; it is an emerging class of physically-native intelligent computing substrates in which architectural decisions are constrained and guided by the equations of the underlying physics. The present paper opens this program; forthcoming work will extend it in the directions the physics itself indicates.

Acknowledgments and AI disclosure

The authors thank colleagues for discussions on field-theoretic architectures and Equilibrium Propagation. This paper was drafted with AI assistance. Large language models were used for literature synthesis, manuscript revisions/editing. All scientific claims, experimental design decisions, data analysis, and final prose are the responsibility of the authors. Experimental code was written and executed with AI coding assistance; all code changes and experimental results were reviewed by the authors prior to incorporation into the manuscript.

Data and code availability

The TIMIT corpus is available from the Linguistic Data Consortium (LDC93S1). Code implementing the FICU architecture, training scripts, and analysis utilities will be released at publication under an open-source license. Trained model checkpoints for L1, L2, and L3, along with the cascade checkpoints and logs required to reproduce all figures in this paper, will accompany the code release.

Author contributions

JDS conceived the study, developed the theoretical framework (CRFM, PETU, cascade protocol), supervised implementation, analyzed results, and drafted the manuscript. GT contributed to the computational architecture, implementation of the Landau-Ginzburg dynamics and Equilibrium Propagation update rules, and revision of the manuscript.

Competing Interests

JDS is Chief Medical Officer at Stratus Neuro and Principal Investigator at MERLN LLC, a private research entity. GT is affiliated with Kvikna Medical ehf. Neither Stratus Neuro nor Kvikna Medical had any role in the conception, design, execution, analysis, or reporting of the work described in this manuscript, and neither company provided financial support for this study. The authors declare no other competing financial or non-financial interests.

References

[1] Hwang EJ, Dahlen JE, Mukundan M, Komiyama T. Disengagement of Motor Cortex during Long-Term Learning Tracks the Performance Level of Learned Movements. J Neurosci. 2021 Aug 18;41(33):7029-7047. doi: 10.1523/jneurosci.3049-20.2021.

[2] Li L, Smith DM. Neural Efficiency in Athletes: A Systematic Review. Front Behav Neurosci. 2021 Aug 5;15:698555. doi: 10.3389/fnbeh.2021.698555.

[3] Akkad H, Dupont-Hadwen J, Kane E, Evans C, Barrett L, Frese A, et al. Increasing human motor skill acquisition by driving theta-gamma coupling. eLife. 2021;10:e67355.

[4] Strzelczyk D, Peylo C, Langer N. Impact of aging on theta-phase gamma-amplitude coupling during learning: a multivariate analysis. Cortex. 2026. doi: 10.1016/j.cortex.2026.01.007

[5] Karlsson AE, Lindenberger U, Sander MC. Out of rhythm: compromised precision of theta-gamma coupling impairs associative memory in old age. J Neurosci. 2022;42(9):1752-1764.

[6] Jones KT, Johnson EL, Berryhill ME. Frontoparietal theta-gamma interactions track working memory enhancement with training and tDCS. Neuroimage. 2020;211:116615. doi:10.1016/j.neuroimage.2020.116615.

[7] Slater JD. The Chrono-Resonant Field Model: a theory of temporal binding in consciousness. Zenodo [Preprint]. 2026 [cited 2026 Apr 28]. doi: 10.5281/zenodo.19655841

[8] Scellier B, Bengio Y. Equilibrium propagation: bridging the gap between energy-based models and backpropagation. Front Comput Neurosci. 2017;11:24.

[9] Laydevant J, Markovic D, Grollier J. Training an Ising machine with equilibrium propagation. Nat Commun. 2024;15:3671.

[10] Wang Q, Wanjura CC, Marquardt F. Training coupled phase oscillators as a neuromorphic platform using equilibrium propagation. Neuromorph Comput Eng. 2024;4:034014.

[11] Laborieux A, Zenke F. Holomorphic equilibrium propagation computes exact gradients through finite size oscillations. In: Koyejo S, Mohamed S, Agarwal A, Belgrave D, Cho K, Oh A, editors. Advances in Neural Information Processing Systems 35 (NeurIPS 2022). Red Hook, NY: Curran Associates; 2022. p. 12950-12963.

[12] Laborieux A, Ernoult M, Scellier B, Bengio Y, Grollier J, Querlioz D. Scaling equilibrium propagation to deep ConvNets by drastically reducing its gradient estimator bias. Front Neurosci. 2021;15:633674.

[13] Sajnok K, Matuszewski M. Near-equilibrium propagation training in nonlinear wave systems. arXiv:2510.16084 [Preprint]. 2025 [cited 2026 Apr 28]. Available from: https://arxiv.org/abs/2510.16084

[14] Slater JD, Thorvardsson G. Phase structure in continuous wave fields enables speech classification without backpropagation. Research Square [Preprint]. 2026 [cited 2026 Apr 28]. doi: 10.21203/rs.3.rs-9205518/v1

[15] Slater JD, Thorvardsson G. Equilibrium propagation discovers top-down feedback for audio-visual binding in continuous wave fields. Research Square [Preprint]. 2026 [cited 2026 Apr 28]. doi: 10.21203/rs.3.rs-9404804/v1

[16] Manoranjani M, Gopal R, Senthilkumar DV, Chandrasekar VK, Lakshmanan M. Influence of asymmetric parameters in higher-order coupling with bimodal frequency distribution. Phys Rev E. 2022;105(3):034307.