The Dimensional Hierarchy of Cortical Oscillations: From Analog Substrate to Symbolic Codes

Ian Todd
Sydney Medical School, University of Sydney
itod2305@uni.sydney.edu.au

November 29, 2025

Abstract

We propose that cortical oscillations implement a *dimensional hierarchy*: a cascade of progressively tighter information bottlenecks from slow to fast frequencies. Using graph Laplacian analysis on modular networks approximating cortical column structure, we show that slow eigenmodes engage substantially more oscillators than fast modes ($r = -0.75$), establishing the high-dimensional geometric substrate. Using encoder-decoder networks, we show that discrete symbolic codes emerge at a critical bottleneck width of $k = 2$, while $k = 3$ preserves continuous "compliant" dynamics—analogous to compliant mechanisms in engineering, which gain functionality through flexible deformation rather than rigid joints—capable of representing self-referential structures without trajectory collision. We propose that this hierarchy maps onto frequency bands: slow oscillations maintain the volumetric analog context; beta ($k \approx 3$) supports manipulation and meta-cognition; gamma ($k \approx 2$) forces categorical commitment. The capacity to sustain $k \geq 3$ dynamics—to hold contradictions without collapsing them—may be a geometric signature of cognitive and emotional maturity. Long-wavelength oscillations provide the temporal stability and geometric depth required for collision-free representations; their disruption biases the system toward premature discretisation.

Keywords: cortical oscillations, information bottleneck, participation ratio, graph Laplacian, dimensional hierarchy, neural coding

1 Introduction

1.1 The Dimensionality Question

Influential work on oscillatory dynamics in prefrontal cortex suggests that different frequency bands serve distinct computational roles (Miller et al., 2018; Lundqvist et al., 2016). Low-frequency oscillations are often characterised as "low-dimensional" coordinating signals, while gamma activity is associated with "high-dimensional" information processing (Bastos et al., 2015). Meanwhile, large-scale neural recordings reveal that population activity can occupy surprisingly high-dimensional spaces (Cunningham and Yu, 2014; Stringer et al., 2019).

However, this framing conflates two distinct notions of dimensionality:

1. **Temporal complexity**: How many independent time-varying components describe the signal at a single site.

2. **Geometric dimensionality**: How many degrees of freedom participate coherently across space.

A slow wave sweeping across cortex may appear "simple" at a single electrode but coordinates thousands of oscillators into coherent phase relationships (high geometric dimensionality). Conversely, a gamma burst may exhibit complex temporal structure but engage only a small cortical population (low geometric dimensionality).

1.2 The Hierarchy Hypothesis

We propose that the frequency spectrum implements a *dimensional hierarchy*—a cascade of information bottlenecks characterised by spatial participation (how many oscillators engage coherently), not latent state-space dimensionality. As a working hypothesis, we suggest the following mapping between bands and effective bottleneck widths:

Band	**Bottleneck**	**Topology**	**Function**
Delta/Theta	$k \gg 3$	Volumetric	Raw substrate
Beta	$k \approx 3$	Compliant manifold	Manipulation, meta-cognition
Gamma	$k \approx 2$	Discrete clusters	Symbols, decisions

The key insight is that **different bottleneck widths support qualitatively different computations**. At $k = 2$, the system is forced to discretise—continuous manifolds collapse into distinct attractor basins ("symbols"). At $k \geq 3$, the system retains enough dimensionality to represent continuous processes, including self-referential structures that would produce trajectory collisions in lower dimensions.

1.3 Hypothesis: Maturity as Dimensional Capacity

We further hypothesise that the capacity to sustain $k \geq 3$ dynamics—to hold contradictions, ambiguity, and nuance without forcing premature resolution—constitutes a geometric signature of cognitive and emotional maturity. Under this hypothesis, the immature or stressed mind collapses to $k = 2$: black/white, good/bad, us/them. The mature mind can inhabit the "compliant" space where paradoxes coexist.

Long-wavelength (slow) oscillations may provide the *temporal stability* required for collision-free representations. When slow-wave power is reduced (stress, sleep deprivation, developmental immaturity), the system may lose the substrate needed to maintain $k \geq 3$ dynamics and default to rigid categorical processing.

2 Methods

2.1 Laplacian Eigenmodes and Participation Ratio

Graph Laplacian eigenmodes provide a natural basis for analysing spatially extended oscillatory patterns (Atasoy et al., 2016). Lower eigenvalues correspond to smoother, longer-wavelength modes; higher eigenvalues correspond to more localised, shorter-wavelength modes. Under a diffusive wave equation on the graph, eigenvalue λ relates to characteristic frequency as $\omega \propto \sqrt{\lambda}$, so we interpret normalised eigenvalue as approximately proportional to frequency squared. We analyse these *structural* eigenmodes as defining the geometric constraints within which dynamics unfold; while non-linear neural dynamics can diverge from this basis, the Laplacian defines the manifold capacity available to the system.

2.1.1 Network Construction

We model cortical connectivity using a modular (stochastic block) network with $N = 2500$ nodes organised into 25 modules of 100 nodes each. Connection probability within modules was $p_{\text{within}} = 0.3$; between modules, $p_{\text{between}} = 0.01$.

The unnormalised graph Laplacian L has diagonal entries equal to node degree and off-diagonal entries of -1 for each edge:

$$L_{ij} = \begin{cases} \deg(i) & \text{if } i = j \\ -1 & \text{if } i \sim j \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

2.1.2 Participation Ratio

The participation ratio (PR), originally developed to quantify localisation in disordered solids (Bell and Dean, 1970), measures how many oscillators contribute significantly to a given mode. For a normalised eigenmode ψ across N oscillators:

$$\text{PR}(\psi) = \frac{1}{\sum_i |\psi_i|^4}, \quad \text{with } \sum_i |\psi_i|^2 = 1. \tag{2}$$

PR $\approx N$ when all oscillators contribute equally; PR ≈ 1 when activity is confined to a single oscillator. We computed the 150 smallest eigenpairs using ARPACK. The negative correlation between eigenvalue and PR is strongest for modular networks ($r \approx -0.85$) and weaker for simple lattices ($r \approx -0.22$); small-world networks show the opposite pattern ($r \approx +0.73$; see Appendix, Figure 6). We report results for the modular network as it better captures cortical column structure—the effect depends on having distinct functional modules with sparse inter-module connections.

Note on dimensionality terminology. PR quantifies *spatial participation*—how many oscillators contribute to a given mode—not the intrinsic dimensionality of the state-space manifold (as typically measured by PCA, factor analysis, or manifold learning). When we say "high-dimensional" in this paper, we mean "many units participate coherently" unless otherwise specified. Our claim is about the geometric substrate: slower modes engage more degrees of freedom, providing a higher-capacity "canvas."

2.1.3 Synthetic Time-Series Validation

To verify that the participation ratio metric correctly distinguishes global from local activity in time-series data, we generated synthetic multi-channel signals. Slow activity (2–8 Hz) was simulated as a global oscillation with high inter-channel correlation ($r = 0.8$); fast activity (30–50 Hz) was simulated as sparse, uncorrelated bursts. The combined signal was bandpass filtered and PR was computed on amplitude envelopes over 200 ms sliding windows. This establishes that PR correctly identifies global versus local activity when the spatial structure is uncontaminated by volume conduction.

2.2 Bottleneck Compression

The information bottleneck principle (Tishby et al., 2000) formalises the trade-off between compression and preservation of task-relevant information. We use encoder-decoder networks to explore how bottleneck dimensionality affects code formation.

2.2.1 Category Construction

We defined six categories as smooth phase gradient patterns across 256 oscillators:

$$\phi_c(i) = 0.3\cos(x_i + \theta_c) + 0.2\sin(0.5x_i + \theta_c) \tag{3}$$

where $x_i \in [0, 4\pi]$ and $\theta_c = \pi c/6$. Samples were generated by adding Gaussian noise ($\sigma = 0.5$), yielding 200 samples per category. We chose phase-gradient inputs not to force circular topology arbitrarily, but because grid cells, head direction cells, and theta phase precession demonstrate that the brain fundamentally encodes continuous variables using cyclic phase codes (S^1 topology).

2.2.2 Network Architecture

The encoder-decoder compresses 256-D input through bottlenecks of width $k \in \{1, 2, 3, 4, 8, 16, 32\}$:

- **Encoder**: $256 \to 128 \to 128 \to k$ (ReLU)
- **Bottleneck**: k dimensions with additive Gaussian noise ($\sigma = 0.5$)
- **Decoder**: $k \to 128 \to 128 \to 256$ (ReLU)

Training used Adam (lr $= 10^{-3}$) for 150 epochs, minimising MSE reconstruction error. Code formation was measured by Adjusted Rand Index (ARI) between k-means clustering of bottleneck codes and true labels.

2.3 Paradox Topology Simulation

To demonstrate collision dynamics, we simulated a self-referential process (the Liar's Paradox: TRUE $\to$ FALSE $\to$ TRUE $\to$ ...) as a continuous trajectory and embedded it in 2D versus 3D spaces:

- **2D embedding**: The cyclic trajectory must trace a closed loop, producing self-intersections ("collisions") at every cycle.
- **3D embedding**: The trajectory can spiral upward (helix), with time as the third dimension, avoiding self-intersection.

Self-intersections were counted as pairs of trajectory points that are spatially close (< 0.15 units) but temporally distant (> 10 steps apart). We also trained a linear autoencoder to compress and reconstruct the 3D helix through $k = 2$ versus $k = 3$ bottlenecks, quantifying the irreducible reconstruction error when the time dimension is lost.

3 Results

We present three linked results that together demonstrate the dimensional hierarchy. First, we show that slow modes are geometrically high-dimensional—they engage many oscillators coherently (Section 3.1). Second, we show that a narrow bottleneck acting on this substrate naturally produces discrete codes at $k \approx 2$–3 (Section 3.2). Third, we demonstrate that $k \geq 3$ is required to represent self-referential dynamics without trajectory collision (Section 3.3). Finally, we show that increased noise shifts the optimal bottleneck width downward, providing a mechanism for stress-induced categorical collapse (Section 3.4).

3.1 Slow Modes Have Higher Participation

Figure 1 shows participation ratio versus normalised eigenvalue. There is a strong negative correlation ($r = -0.75$, $p < 0.001$): slower modes engage substantially more oscillators. The slowest 15 modes have mean PR $\approx$ 560 (23% of nodes); the fastest 15 modes have mean PR $\approx$ 175 (7% of nodes)—a 3-fold difference. This effect is largest in modular networks; lattice graphs show weaker correlations and small-world graphs can show reversed patterns (Appendix, Figure 6), predicting that the slow>fast PR difference should be strongest in association cortex with pronounced modularity.

Synthetic time-series validation (Figure 2) confirms that PR correctly distinguishes global from local activity. Slow-band activity (2–8 Hz) engaged 99% of channels (PR = 63.2/64); fast-band activity (30–50 Hz) engaged only 42% (PR = 26.6/64)—a 2.4-fold difference matching the Laplacian prediction. (Preliminary application to 64-channel scalp EEG yielded null results, consistent with the spatial-blurring limitations discussed in Section 4.7.)

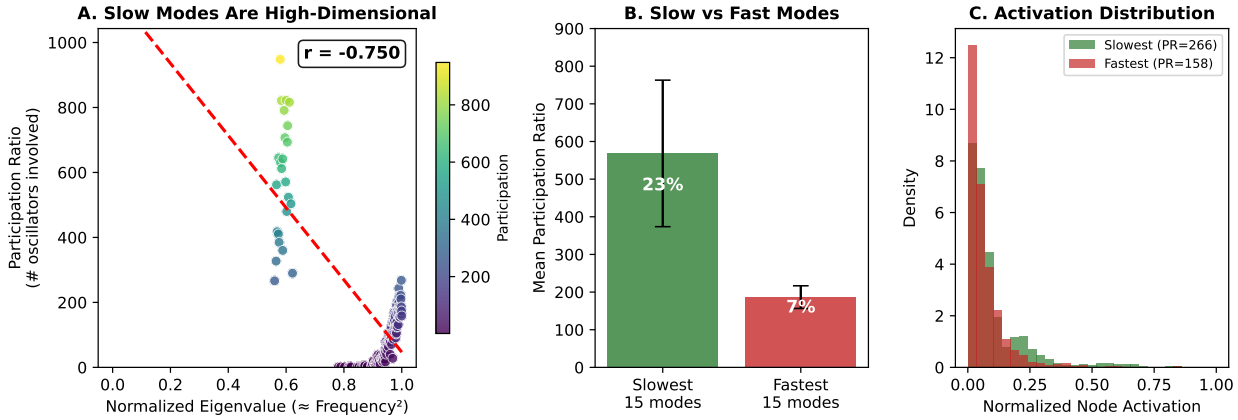

Figure 1: **Slow modes are geometrically high-dimensional.** (A) Participation ratio decreases with eigenvalue ($r = -0.75$). (B) Mean participation for slowest vs fastest modes shows 3-fold difference. (C) Activation distributions: slow modes spread activity broadly.

3.2 Discrete Codes Emerge at $k = 2$

Figure 3 shows that code formation peaks at the critical bottleneck width $k \approx 2$–3 (ARI = 0.88; representative run with fixed seed, qualitatively stable across seeds). At $k = 1$, information is lost (ARI = 0.64). At $k \geq 4$, discretisation pressure diminishes and codes become more distributed (ARI = 0.82–0.87).

The critical $k \approx 2$ reflects a *static* topological constraint: the input categories are distinguished by phase relationships with S^1 (circular) topology. In our construction, they are embedded as a circle in $\mathbb{R}^2$, so a 1-D bottleneck cannot separate them, whereas $k \geq 2$ is sufficient for static classification. However, as Section 3.3 demonstrates, when *dynamical* trajectories must evolve over time without self-intersection, an additional dimension ($k \geq 3$) becomes functionally important—the distinction between "enough dimensions to separate categories" and "enough dimensions to avoid collision during temporal evolution."

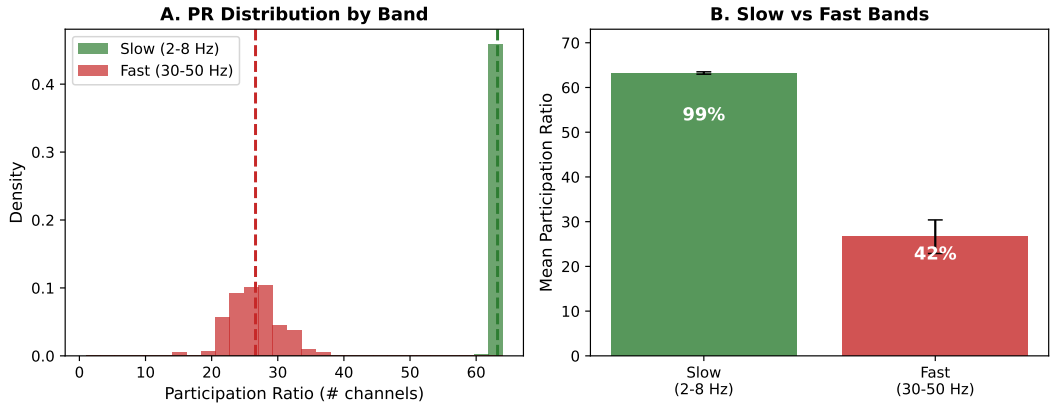

Figure 2: **Synthetic validation confirms PR metric.** (A) PR distributions for slow (2–8 Hz) versus fast (30–50 Hz) bands in synthetic multi-channel data. (B) Mean PR: slow band engages 99% of channels; fast band engages 42%. This 2.4-fold difference confirms that the metric correctly identifies global versus local activity when spatial structure is preserved.

3.3 Self-Reference Benefits from $k \geq 3$

Figure 4 demonstrates both geometric and computational evidence for the $k \geq 3$ requirement. The Liar's Paradox trajectory produces 1511 self-intersections when confined to 2D (panel A); in 3D, it forms a collision-free helix (panel B). The computational proof is equally stark: a linear autoencoder trained to compress and reconstruct the 3D helix through a $k = 2$ bottleneck fails catastrophically on the time dimension (MSE $= 0.108$), while $k = 3$ achieves perfect reconstruction (MSE $< 10^{-6}$; panels C–D). The high error at $k = 2$ is *irreducible*: it represents the necessary loss of temporal information to satisfy the planar constraint. This is a *topological* constraint, not merely a limitation of linear autoencoders: in 2D, any continuous trajectory that returns to the same logical state while keeping process states distinct must self-intersect. Crucially, because the paradox is a logical *cycle* (it returns to the start state), "unrolling" the helix would require cutting the loop, destroying the continuous temporal dynamics required for stability.

3.4 Noise Shifts the Optimal Bottleneck

Figure 5 shows that the optimal bottleneck width depends on channel noise. Quantitatively, the argmax over k shifts from $k \geq 3$ at $\sigma \leq 0.3$ to $k = 2$ at $\sigma \geq 0.7$. High-dimensional "helical" representations are fragile—they require precise coordination across many dimensions. Binary contrast ($k = 2$) is robust: it survives noise because it only needs to distinguish two categories. This provides a computational mechanism for the stress–rigidity link: degraded signal-to-noise ratio (from stress, fatigue, or developmental immaturity) drives the system toward categorical processing not by choice but by information-theoretic necessity. The peak at $k \approx 2$–3 generalises across different numbers of input categories (Appendix, Figure 7).

4 Discussion

4.1 The Frequency Hierarchy

These results suggest that cortical frequency bands implement a cascade of dimensional bottlenecks:

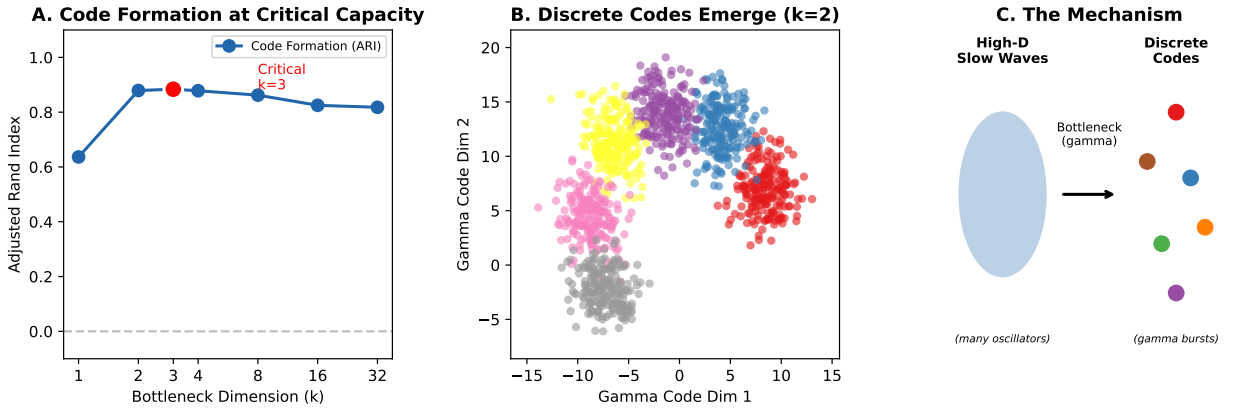

Figure 3: **Discrete codes emerge at critical capacity.** (A) Code formation (ARI) peaks at $k \approx 2$–3, reflecting the S^1 topology of the phase-defined categories. (B) Bottleneck codes at $k = 2$ form six distinct clusters, each interpretable as a proto-symbol. (C) Schematic: high-dimensional slow-wave patterns compress through a noisy bottleneck, forcing discretisation when channel capacity is limited.

- **Slow oscillations (delta, theta)**: Maintain the high-dimensional volumetric substrate. Many oscillators participate coherently, providing the "canvas" for computation.
- **Beta oscillations ($k \approx 3$)**: Intermediate compression. Beta rhythms are associated with maintenance of the current cognitive state (Engel and Fries, 2010). The manifold is constrained but remains "compliant"—continuous enough for analog manipulation, mental rotation, holding ambiguity.
- **Gamma oscillations ($k \approx 2$)**: Tight compression. Gamma rhythms support selective information transmission through coherence (Fries, 2015). Forces categorical commitment—symbols, decisions, assertions. Optimal for transmission but sacrifices nuance.

The brain may dynamically adjust bottleneck width via frequency shifts: increasing beta power to "loosen" the constraint when flexibility is needed; increasing gamma power to "tighten" when commitment is required.

4.2 Collision Dynamics and Meta-Cognition

The paradox simulation reveals a fundamental distinction:

- $k = 2$ **(Gamma)**: Self-referential structures produce collisions. The system cannot stably represent "This statement is false." Attempting to do so produces oscillation, confusion, or collapse.
- $k \geq 3$ **(Beta)**: The extra dimension allows trajectories to "lift" over themselves. The system can represent the *process* of the paradox—understand it—without being forced to resolve it.

This offers a geometric interpretation of why we can *think about* paradoxes (beta, meta-cognition) but cannot *decide* them (gamma, assertion). Other encoding schemes (e.g., discrete state machines) can represent the paradox without an explicit extra spatial dimension; our point is that *continuous geometric* representations of self-reference naturally demand $k \geq 3$.

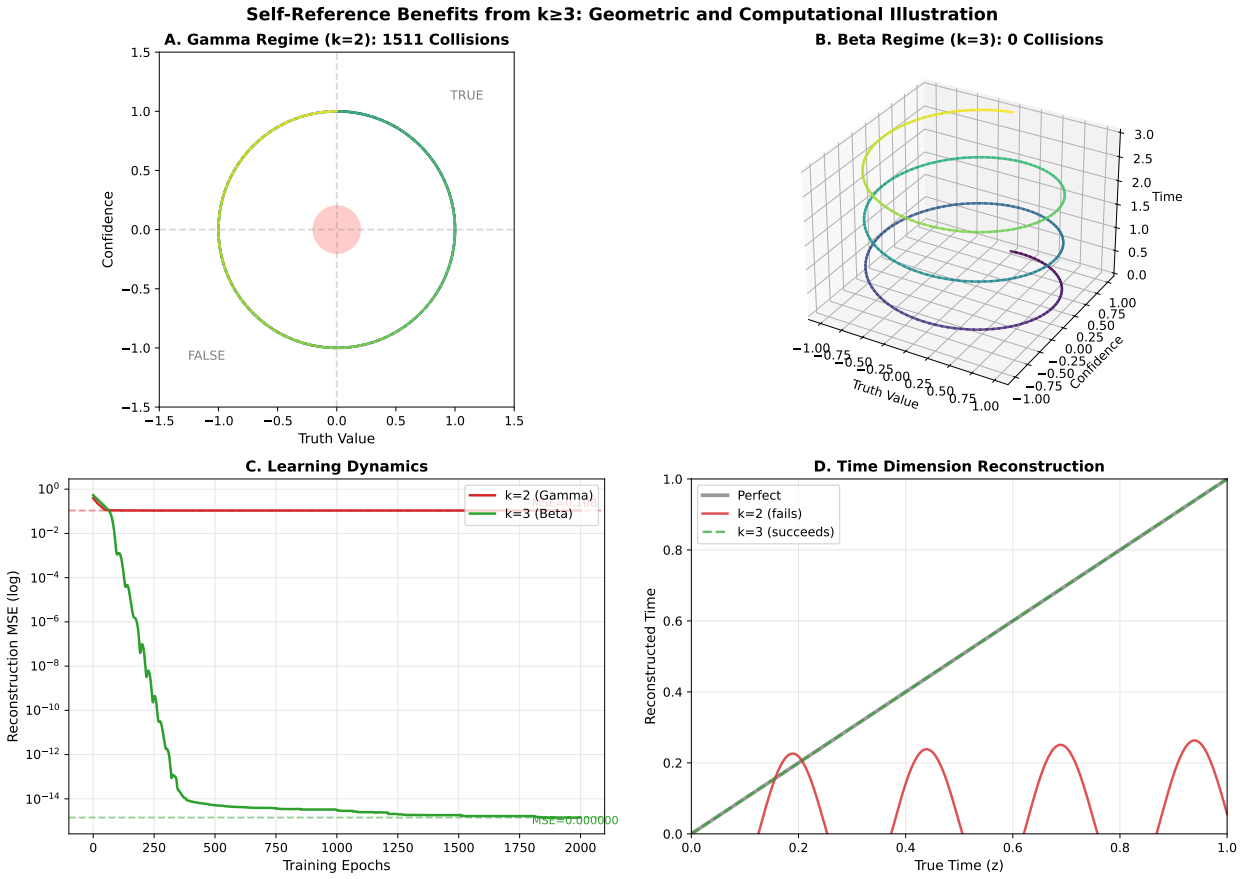

Figure 4: **Geometric and computational proof that self-reference benefits from $k \geq 3$.** (A) In 2D, the paradox trajectory collides with itself (1511 intersections; defined as point pairs < 0.15 units apart but > 10 steps distant). Colour encodes time. (B) In 3D, the trajectory spirals upward (helix), avoiding collision. (C) Learning dynamics: $k = 2$ plateaus at high MSE; $k = 3$ converges to zero. (D) Time reconstruction: $k = 2$ cannot recover the temporal dimension (red); $k = 3$ achieves perfect reconstruction (green).

4.3 Hypothesis: Maturity and Long-Wavelength Stability

We hypothesise that the capacity to sustain $k \geq 3$ dynamics constitutes a geometric signature of cognitive and emotional maturity:

- **Immature/stressed cognition**: Tends toward an effectively $k \approx 2$ regime, favouring binary categorisation. Contradictions may be difficult to tolerate; ambiguity may provoke discomfort.
- **Mature cognition**: Can sustain $k \geq 3$ dynamics. Holds contradictions without forcing resolution. Tolerates ambiguity. Maintains multiple perspectives simultaneously.

This geometric capacity likely tracks the biological maturation of the prefrontal cortex (PFC). The PFC is the primary generator of top-down beta oscillations (Miller et al., 2018; Lara and Wallis, 2015) and is the last cortical region to fully myelinate, a process that continues into the mid-20s (Gogtay et al., 2004). The ability to sustain the "helical" $k \geq 3$ dynamics required for collision-free logic may be physically rate-limited by this development. The immature brain,

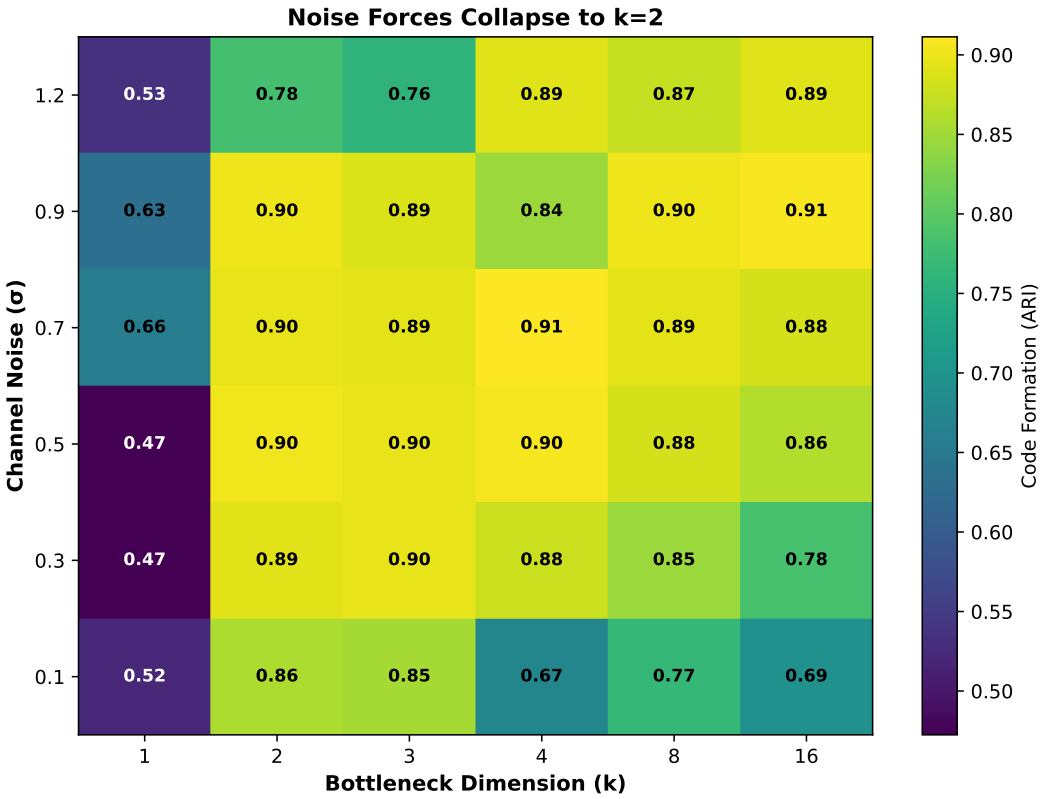

Figure 5: **Noise forces topological collapse.** Code formation (ARI) heatmap across bottleneck width (k) and channel noise (σ). At low noise (bottom), the system supports higher-dimensional nuance ($k \geq 3$). At high noise (top), optimal coding collapses to $k = 2$: binary categorisation becomes the only robust strategy. This provides a computational mechanism for stress-induced rigid thinking.

lacking the myelinated bandwidth to sustain stable high-dimensional beta loops, falls back on earlier-developing, sensory-driven gamma circuitry ($k = 2$). Consequently, it defaults to a "planar" topology where paradoxes force immediate, binary resolution (impulsivity) rather than being held in suspension (cognitive control).

Long-wavelength oscillations provide the *temporal stability* needed for this capacity. Slow waves have long autocorrelation times—they change gradually, providing a stable backdrop against which complex, collision-free trajectories can unfold. Slow-wave sleep, in particular, is critical for memory consolidation (Diekelmann and Born, 2010); its disruption impairs not only declarative memory but also the cognitive flexibility that depends on integrated representations. When slow-wave power is compromised (stress, sleep deprivation, developmental immaturity, certain psychiatric conditions), the system loses this stability and collapses to categorical processing. The noise simulation (Figure 5) formalises this intuition: as channel signal-to-noise ratio degrades, the optimal bottleneck width shifts from $k \geq 3$ (nuance) to $k = 2$ (binary contrast). This is not a failure of the system but an adaptive response—binary codes are simply more robust to noise. While we model this mathematically as additive noise, biologically it corresponds to any reduction in effective channel capacity: synaptic fatigue, inhibitory dysregulation, or information overload where input complexity exceeds processing bandwidth. These links to psychiatric conditions are speculative and would require careful, dedicated clinical studies to test.

4.4 Cross-Frequency Coupling

The well-documented coupling between slow-wave phase and gamma amplitude (Canolty and Knight, 2010; Lisman and Jensen, 2013) may implement exactly this hierarchy. The slow oscillation defines the current position on the high-dimensional substrate; the gamma burst transmits a discrete "snapshot" of that position. The number of distinct gamma codes per slow cycle would be limited by $k = 2$ capacity, potentially explaining the 7 ± 2 limit on working memory items (Miller, 1956; Cowan, 2001; Lisman and Jensen, 2013).

4.5 Laminar and Bioelectric Connections

This architecture maps onto cortical laminar structure: deep layers (L5/6) with extensive horizontal connectivity support the slow, high-dimensional substrate; superficial layers (L2/3) with gamma-dominant activity implement the bottleneck (van Kerkoerle et al., 2014). The laminar segregation of alpha/beta (deep, feedback) versus gamma (superficial, feedforward) aligns precisely with the proposed hierarchy.

More broadly, the motif of electric fields interacting with 2D surfaces recurs across biology: cell membranes, bioelectric gradients in morphogenesis (Levin, 2021), and cortical sheets all share this geometric configuration. Ephaptic coupling—where neurons influence each other through extracellular fields rather than synapses (Anastassiou et al., 2011)—represents the same principle operating at faster timescales; Levin's bioelectric code and neural ephaptic fields may be two expressions of a common computational motif. The 2D surface may be a broadly useful architecture for transforming continuous volumetric dynamics into discrete signals.

4.6 Testable Predictions

The framework generates several concrete predictions for future empirical work:

1. **PR across frequency bands**: In dense intracranial recordings (Utah arrays, high-density ECoG), beta-band activity should engage more channels (higher PR) than gamma-band activity during flexible cognition tasks, while both should engage more than rest.

2. **State-space trajectories**: During tasks requiring cognitive flexibility (e.g., set-shifting), decoded neural trajectories should exhibit more continuous, less clustered structure than during categorical decision tasks.

3. **Cross-frequency structure**: The number of distinguishable gamma "packets" per slow-wave cycle should be constrained to approximately $k = 2$ worth of information, potentially explaining working memory capacity limits.

4. **Developmental trajectory**: PR measures for beta-band activity in frontal regions should increase with age through adolescence, paralleling PFC myelination.

4.7 Limitations and Measurement Requirements

While synthetic validation confirms the PR metric's utility (Figure 2), preliminary application to scalp EEG (64-channel recordings) yielded null results. This is expected and consistent with the spatial Nyquist theorem: capturing the geometric difference between global and local modes requires sampling density higher than the spatial frequency of the mode. Scalp EEG (inter-electrode distance $\sim$2–3 cm) spatially aliases local gamma structure, making it indistinguishable from noise.

The slow-wave versus fast-wave PR difference requires electrode spacing finer than the spatial scale of functional oscillator units. Cortical columns span ∼300–500 μm; resolving participation at this scale requires Utah-array density (∼400 μm pitch) or high-density ECoG. Standard ECoG grids (1 cm spacing) may still average over too many columns. This represents a clear experimental target: PR analysis on dense intracranial recordings during tasks requiring flexibility (beta-dominant) versus commitment (gamma-dominant).

The models presented here isolate geometric and information-theoretic aspects; they do not capture excitatory-inhibitory dynamics, conduction delays, or synaptic nonlinearities. Additionally, Laplacian eigenmodes represent standing wave constraints based on diffusion; while axonal delays enable traveling waves that deviate from these standing modes, the Laplacian defines the structural manifold capacity—the baseline geometric availability of the network before temporal dynamics are applied. Future work should test these predictions in biologically realistic spiking networks and empirical data with appropriate spatial resolution.

5 Conclusion

Cortical oscillations implement a dimensional hierarchy: slow waves maintain the high-dimensional analog substrate; beta provides intermediate "compliant" compression for manipulation; gamma enforces discrete symbol formation. The capacity to sustain $k \geq 3$ dynamics—to hold paradoxes without collision—may be a geometric signature of maturity. Long-wavelength stability provides the temporal substrate for collision-free cognition; its disruption forces the mind into rigid categorical processing.

Intelligence emerges from the controlled collapse of analog into digital. The canvas is slow; the brushstrokes are fast.

Acknowledgements

The author thanks the Fulcher Lab, University of Sydney, for feedback on an early version of this work, and Prof. Michael Levin (Tufts University) for generously answering questions at office hours—particularly the insight that the Liar's Paradox can be understood as an oscillation in phase space.

Statements and Declarations

Funding. The author did not receive support from any organisation for the submitted work.
Competing Interests. The author has no relevant financial or non-financial interests to disclose.
Code Availability. All simulation code and analysis scripts are available at: `https://github.com/todd866/brainwavedimensionality`
AI Assistance. Large language models (Claude, GPT) were used for AI-assisted copy editing and iterative manuscript refinement. The author takes full responsibility for the content.

A Robustness Analyses

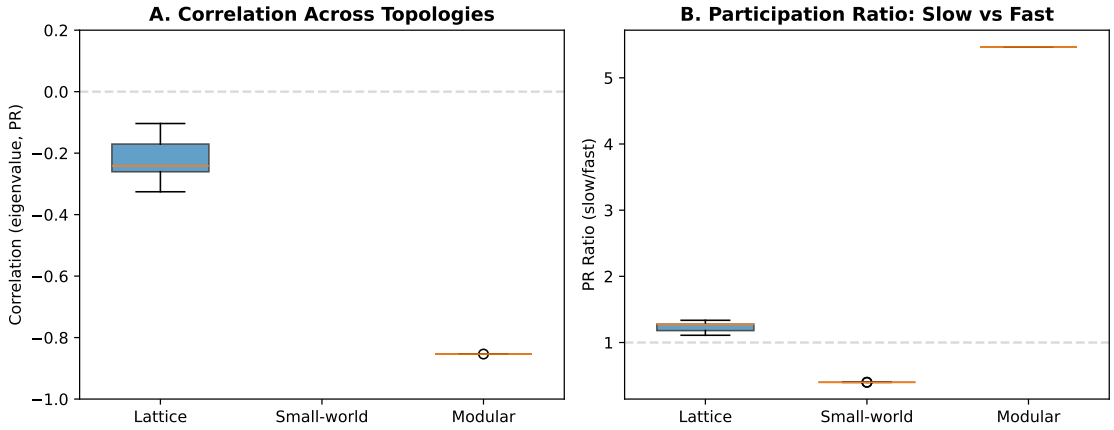

Figure 6: **Participation ratio effect depends on network topology.** (A) Correlation between eigenvalue and PR across network types. The negative correlation is strongest for modular networks ($r \approx -0.85$), weaker for lattices ($r \approx -0.22$), and reverses for small-world networks ($r \approx +0.73$). (B) PR ratio (slowest 10 modes / fastest 10 modes) confirms modular networks show the largest slow/fast difference. The effect requires distinct functional modules with sparse inter-module connections—consistent with cortical column architecture.

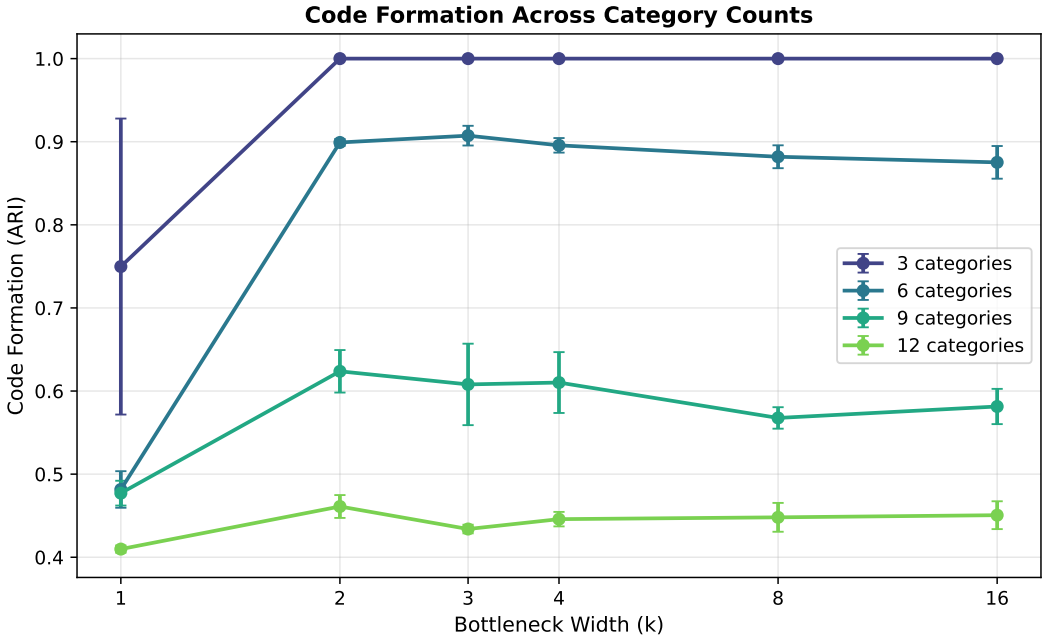

Figure 7: **Code formation peaks at $k \approx 2$–3 across category counts.** ARI versus bottleneck width for different numbers of input categories (3, 6, 9, 12). Regardless of category count, peak code formation occurs at $k = 2$–3, confirming this is a robust feature of the bottleneck architecture rather than an artefact of the specific 6-category construction used in the main text.

References

Anastassiou, C. A., Perin, R., Markram, H., and Koch, C. (2011). Ephaptic coupling of cortical neurons. *Nature Neuroscience*, 14(2):217–223.

Atasoy, S., Donnelly, I., and Pearson, J. (2016). Human brain networks function in connectome-specific harmonic waves. *Nature Communications*, 7:10340.

Bastos, A. M., Vezoli, J., Bosman, C. A., Schoffelen, J.-M., Oostenveld, R., Dowdall, J. R., De Weerd, P., Kennedy, H., and Fries, P. (2015). Visual areas exert feedforward and feedback influences through distinct frequency channels. *Neuron*, 85(2):390–401.

Bell, R. J. and Dean, P. (1970). Atomic vibrations in vitreous silica. *Discussions of the Faraday Society*, 50:55–61.

Canolty, R. T. and Knight, R. T. (2010). The functional role of cross-frequency coupling. *Trends in cognitive sciences*, 14(11):506–515.

Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24(1):87–114.

Cunningham, J. P. and Yu, B. M. (2014). Dimensionality reduction for large-scale neural recordings. *Nature Neuroscience*, 17(11):1500–1509.

Diekelmann, S. and Born, J. (2010). The memory function of sleep. *Nature Reviews Neuroscience*, 11(2):114–126.

Engel, A. K. and Fries, P. (2010). Beta-band oscillations—signalling the status quo? *Current Opinion in Neurobiology*, 20(2):156–165.

Fries, P. (2015). Rhythms for cognition: communication through coherence. *Neuron*, 88(1):220–235.

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., Nugent, T. F., Herman, D. H., Clasen, L. S., Toga, A. W., Rapoport, J. L., and Thompson, P. M. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences*, 101(21):8174–8179.

Lara, A. H. and Wallis, J. D. (2015). The role of prefrontal cortex in working memory: a mini review. *Frontiers in Systems Neuroscience*, 9:173.

Levin, M. (2021). Bioelectric signaling: Reprogrammable circuits underlying embryogenesis, regeneration, and cancer. *Cell*, 184(6):1971–1989.

Lisman, J. E. and Jensen, O. (2013). The theta-gamma neural code. *Neuron*, 77(6):1002–1016.

Lundqvist, M., Rose, J., Herman, P., Brincat, S. L., Buschman, T. J., and Miller, E. K. (2016). Gamma and beta bursts underlie working memory. *Neuron*, 90(1):152–164.

Miller, E. K., Lundqvist, M., and Bastos, A. M. (2018). Working memory 2.0. *Neuron*, 100(2):463–475.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63(2):81–97.

Stringer, C., Pachitariu, M., Steinmetz, N., Reddy, C. B., Carandini, M., and Harris, K. D. (2019). Spontaneous behaviors drive multidimensional, brainwide activity. *Science*, 364(6437):eaav7893.

Tishby, N., Pereira, F. C., and Bialek, W. (2000). The information bottleneck method. In *Proceedings of the 37th Allerton Conference on Communication, Control, and Computing*, pages 368–377.

van Kerkoerle, T., Self, M. W., Dagnino, B., Gariel-Mathis, M.-A., Poort, J., van der Togt, C., and Roelfsema, P. R. (2014). Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. *Proceedings of the National Academy of Sciences*, 111(40):14332–14341.