

Information Geometry of Nested Frequency Hierarchies: Why Biological Systems Exhibit High-Dimensional Coherence

Ian Todd

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

Abstract

Biological systems from neural circuits to cellular networks exhibit nested frequency hierarchies: oscillations at multiple timescales with systematic phase-amplitude coupling between scales. We provide an information-geometric framework showing how this architecture creates high-dimensional statistical structure. Under a generative model where slow oscillatory phases modulate fast dynamics, we prove that: (1) K frequency bands induce a $(K-1)$ -dimensional parameter space of relative phases after quotienting global phase; (2) phase-amplitude coupling generically produces off-diagonal Fisher information, enabling identification of slow phase from fast-band observations that would otherwise be unidentifiable; and (3) global clock synchronization collapses the accessible parameter family to at most one dimension. A worked example with theta-gamma coupling demonstrates how coupling activates identifiability: without modulation, slow phase drops out of the likelihood; with modulation, it becomes estimable. These results clarify the geometric basis for why biological substrates—which universally employ nested frequency architectures—support richer statistical structure than globally clocked systems.

Keywords: Fisher information, product manifolds, phase-amplitude coupling, biological oscillations, identifiability

1 Introduction

Biological systems operate through nested oscillatory hierarchies. In neural systems, gamma oscillations (30–80 Hz) are modulated by theta rhythms (4–8 Hz), which are in turn modulated by slower delta and infra-slow fluctuations [2]. In cellular biology, circadian rhythms organize ultradian metabolic cycles which contain faster enzymatic oscillations [7]. This hierarchical frequency organization appears ubiquitous across living systems [4].

Recent theoretical work proposes that intelligence arises from maintaining high-dimensional coherent dynamics [9]—systems that sustain many coupled degrees of freedom. But where does this dimensionality come from, and what makes it statistically meaningful?

We argue that **nested frequency hierarchies create high-dimensional statistical structure** through a specific mechanism: phase-amplitude coupling between timescales makes slow contextual variables identifiable from fast observables. This is not merely a matter of having multiple oscillators (which trivially yields a product space), but of how cross-scale coupling creates *informative dependencies* that would not exist under independence.

1.1 Main Contributions

We develop an information-geometric framework with an explicit observation model, proving three results:

1. **Phase space structure** (Lemma 2): K oscillatory bands with unobservable global phase yield a $(K - 1)$ -dimensional parameter space \mathbb{T}^{K-1} of relative phases.
2. **Coupling activates identifiability** (Theorem 7): Under phase-amplitude coupling, the Fisher information matrix generically has nonzero off-diagonal entries, and parameters that are unidentifiable under independence become identifiable. We demonstrate this explicitly with a theta-gamma example.
3. **Clock collapse** (Theorem 10): Global clock synchronization reduces the relative phase parameters to deterministic functions of a single time-offset parameter, collapsing identifiable dimension to at most 1.

The key insight is that *nesting*—where slow phases modulate fast dynamics—creates statistical dependencies that make the slow phase observable through the fast signal. This is distinct from merely having multiple independent oscillators.

1.2 Notation

We use the following notation throughout:

- $\phi_k \in S^1$: phase of the k -th oscillatory band
- $\boldsymbol{\phi} = (\phi_1, \dots, \phi_K) \in \mathbb{T}^K$: full phase vector
- $\psi_j = \phi_j - \phi_1$ for $j \geq 2$: relative phases (parameters)
- $\boldsymbol{\psi} = (\psi_2, \dots, \psi_K) \in \mathbb{T}^{K-1}$: relative phase vector
- y : observed signal
- $p(y | \boldsymbol{\psi})$: likelihood of observation given parameters
- $\mathcal{I}(\boldsymbol{\psi})$: Fisher information matrix at $\boldsymbol{\psi}$

2 Preliminaries

2.1 Statistical Manifolds and Fisher Information

A *statistical manifold* is a smooth family of probability distributions $\{p(y \mid \theta) : \theta \in \Theta\}$ parameterized by $\theta \in \Theta \subseteq \mathbb{R}^n$. The Fisher information matrix at θ is:

$$\mathcal{I}_{ij}(\theta) = \mathbb{E}_{y \sim p(\cdot \mid \theta)} \left[\frac{\partial \log p(y \mid \theta)}{\partial \theta^i} \frac{\partial \log p(y \mid \theta)}{\partial \theta^j} \right] \quad (1)$$

where the expectation is over observations y drawn from $p(y \mid \theta)$.

The *Fisher rank* at θ is $\text{rank}(\mathcal{I}(\theta))$: the number of locally identifiable parameter directions. If \mathcal{I} has a zero eigenvalue in direction v , then parameters cannot be distinguished along v from observations—the model is *unidentifiable* in that direction.

Remark 1. *Throughout this paper, Fisher information is always computed with respect to parameters θ (here, relative phases ψ) in a likelihood $p(y \mid \theta)$, not with respect to coordinates of a distribution on phase space. This distinction is essential.*

2.2 Phase Variables and the Torus

A system with K oscillators has phase variables $\phi = (\phi_1, \dots, \phi_K) \in \mathbb{T}^K$, where $\mathbb{T}^K = (S^1)^K$ is the K -torus. In many settings, only *relative* phases are observable—the global phase $\Phi = \frac{1}{K} \sum_i \phi_i$ cannot be determined from observations invariant under simultaneous phase shifts.

Lemma 2 (Phase space reduction). *If observations are invariant under the diagonal action $\phi_k \mapsto \phi_k + \alpha$ for all k , then the effective parameter space is the quotient:*

$$\mathbb{T}^K / S^1 \cong \mathbb{T}^{K-1} \quad (2)$$

with coordinates $\psi_j = \phi_j - \phi_1$ for $j = 2, \dots, K$.

Proof. The diagonal S^1 action on \mathbb{T}^K has orbits $\{(\phi_1 + \alpha, \dots, \phi_K + \alpha) : \alpha \in S^1\}$. Invariant functions depend only on phase differences. The map $(\phi_1, \dots, \phi_K) \mapsto (\phi_2 - \phi_1, \dots, \phi_K - \phi_1)$ provides coordinates on the quotient, which is diffeomorphic to \mathbb{T}^{K-1} . \square

This lemma is a standard fact about product tori; the key question is what statistical structure exists on \mathbb{T}^{K-1} under different observation models.

2.3 Observation Model: Phase-Amplitude Coupling

We now specify the generative model that makes “nested frequency hierarchy” statistically meaningful.

Definition 3 (Nested frequency observation model). *A nested frequency observation model consists of:*

1. Latent phases $\phi = (\phi_1, \dots, \phi_K)$ with frequencies $\omega_1 < \omega_2 < \dots < \omega_K$ (slow to fast).

2. A signal $y(t)$ of the form:

$$y(t) = A(\phi_1, \dots, \phi_{K-1}) \cdot g(\phi_K) + \epsilon(t) \quad (3)$$

where $A(\cdot)$ is the amplitude modulation by slower phases, $g(\cdot)$ is the fast carrier (e.g., $\cos(\phi_K)$), and $\epsilon(t)$ is observation noise.

3. Parameters $\psi = (\psi_2, \dots, \psi_K)$ are relative phases.

The critical feature is that **slow phases modulate the amplitude of fast oscillations**. This is the defining characteristic of nested frequency hierarchies, distinguishing them from independent oscillators.

Example 4 (Theta-gamma coupling). *In hippocampal recordings, gamma oscillations (~ 40 Hz) have amplitude modulated by theta phase (~ 8 Hz) [6]. The signal model is:*

$$y(t) = [1 + \alpha \cos(\phi_\theta)] \cos(\phi_\gamma) + \epsilon(t) \quad (4)$$

where $\alpha \in [0, 1]$ is the modulation depth. When $\alpha = 0$, the phases are statistically independent in the observation; when $\alpha > 0$, theta phase affects the observable signal.

3 Fisher Information Under Independence

We first establish a baseline: what is the Fisher information structure when oscillatory bands are statistically independent in the observation model?

Proposition 5 (Block-diagonal Fisher under independence). *Suppose the likelihood factorizes as $p(y|\psi) = \prod_{j=2}^K p_j(y_j|\psi_j)$ where y_j are independent observations of each relative phase. Then the Fisher information matrix is block-diagonal:*

$$\mathcal{I}(\psi) = \text{diag}(\mathcal{I}_2(\psi_2), \dots, \mathcal{I}_K(\psi_K)) \quad (5)$$

where $\mathcal{I}_j(\psi_j) = \mathbb{E}[(\partial_{\psi_j} \log p_j)^2]$.

Proof. Under independence, $\log p(y|\psi) = \sum_j \log p_j(y_j|\psi_j)$. The score for ψ_j is $\partial_{\psi_j} \log p = \partial_{\psi_j} \log p_j$, which depends only on y_j . For $j \neq k$:

$$\mathcal{I}_{jk} = \mathbb{E}[\partial_{\psi_j} \log p \cdot \partial_{\psi_k} \log p] = \mathbb{E}[\partial_{\psi_j} \log p_j] \cdot \mathbb{E}[\partial_{\psi_k} \log p_k] = 0 \quad (6)$$

since $\mathbb{E}[\partial_\psi \log p] = 0$ for any likelihood. \square

Remark 6 (Unidentifiable phases under partial observation). *Under the independent model, if we only observe the fast band (e.g., gamma), then slow phases (e.g., theta) do not appear in the likelihood at all. The corresponding rows and columns of \mathcal{I} are zero—these parameters are completely unidentifiable.*

This sets up the key question: how does phase-amplitude coupling change this picture?

4 Coupling Activates Identifiability

We now prove that phase-amplitude coupling generically produces off-diagonal Fisher information and, crucially, can make previously unidentifiable parameters identifiable.

4.1 General Result

Theorem 7 (Coupling activates identifiability). *Consider the observation model (3) with Gaussian noise $\epsilon \sim \mathcal{N}(0, \sigma^2)$. Let $A(\phi_1, \dots, \phi_{K-1}) = A_0 + \sum_{j=1}^{K-1} \alpha_j f_j(\phi_j)$ where f_j are non-constant functions and α_j are modulation depths.*

- (i) *If all $\alpha_j = 0$ (no modulation), then slow phases $\phi_1, \dots, \phi_{K-1}$ are unidentifiable from observations of y : the Fisher information for these parameters is zero.*
- (ii) *If $\alpha_j \neq 0$ for some j , then ϕ_j becomes identifiable: $\mathcal{I}_{jj} > 0$. Moreover, if multiple $\alpha_j \neq 0$, the Fisher matrix generically has nonzero off-diagonal entries.*

Proof. For Gaussian noise, the log-likelihood is:

$$\log p(y|\phi) = -\frac{1}{2\sigma^2} [y - A(\phi_1, \dots, \phi_{K-1})g(\phi_K)]^2 + \text{const} \quad (7)$$

The score for ϕ_j (with $j < K$) is:

$$\frac{\partial \log p}{\partial \phi_j} = \frac{1}{\sigma^2} [y - Ag(\phi_K)] \cdot \frac{\partial A}{\partial \phi_j} \cdot g(\phi_K) \quad (8)$$

- (i) If $\alpha_j = 0$, then $\partial A / \partial \phi_j = 0$, so the score is identically zero and $\mathcal{I}_{jj} = 0$.
- (ii) If $\alpha_j \neq 0$, then $\partial A / \partial \phi_j = \alpha_j f'_j(\phi_j) \neq 0$ generically. The Fisher information is:

$$\mathcal{I}_{jj} = \mathbb{E} \left[\left(\frac{\partial \log p}{\partial \phi_j} \right)^2 \right] \quad (9)$$

$$= \frac{1}{\sigma^4} \mathbb{E}[(y - Ag)^2] \cdot \left(\frac{\partial A}{\partial \phi_j} \right)^2 g(\phi_K)^2 \quad (10)$$

$$= \frac{1}{\sigma^2} (\alpha_j f'_j(\phi_j))^2 g(\phi_K)^2 > 0 \quad (11)$$

For the off-diagonal terms with $j \neq k$ (both $< K$):

$$\mathcal{I}_{jk} = \frac{1}{\sigma^2} \frac{\partial A}{\partial \phi_j} \frac{\partial A}{\partial \phi_k} g(\phi_K)^2 \quad (12)$$

which is generically nonzero when both $\alpha_j, \alpha_k \neq 0$. \square

Interpretation. The theorem formalizes a simple but important point: *phase-amplitude coupling makes slow phases visible in fast observations*. Without coupling, observing gamma tells you nothing about theta. With coupling, theta phase modulates gamma amplitude, so theta becomes estimable from gamma recordings.

4.2 Worked Example: Theta-Gamma System

We now compute the Fisher information matrix explicitly for the canonical two-band case.

Example 8 (Fisher information for theta-gamma coupling). *Consider the model (4):*

$$y = [1 + \alpha \cos(\psi)] \cos(\phi_\gamma) + \epsilon, \quad \epsilon \sim \mathcal{N}(0, \sigma^2) \quad (13)$$

where $\psi = \phi_\theta - \phi_\gamma$ is the relative phase (the parameter of interest, since global phase is unobservable) and we condition on ϕ_γ being known from the fast carrier.

The log-likelihood is:

$$\log p(y|\psi) = -\frac{1}{2\sigma^2} [y - (1 + \alpha \cos \psi) \cos \phi_\gamma]^2 + \text{const} \quad (14)$$

The score is:

$$\frac{\partial \log p}{\partial \psi} = \frac{\alpha \sin \psi \cos \phi_\gamma}{\sigma^2} [y - (1 + \alpha \cos \psi) \cos \phi_\gamma] \quad (15)$$

The Fisher information is:

$$\mathcal{I}(\psi) = \mathbb{E} \left[\left(\frac{\partial \log p}{\partial \psi} \right)^2 \right] \quad (16)$$

$$= \frac{\alpha^2 \sin^2 \psi \cos^2 \phi_\gamma}{\sigma^4} \cdot \mathbb{E}[(y - (1 + \alpha \cos \psi) \cos \phi_\gamma)^2] \quad (17)$$

$$= \frac{\alpha^2 \sin^2 \psi \cos^2 \phi_\gamma}{\sigma^2} \quad (18)$$

Key observations:

1. When $\alpha = 0$ (no coupling): $\mathcal{I}(\psi) = 0$. Theta phase is completely unidentifiable from gamma observations.
2. When $\alpha > 0$ (coupling present): $\mathcal{I}(\psi) > 0$ for generic ψ (except at $\psi = 0, \pi$ where modulation effect vanishes). Theta phase becomes identifiable.
3. Fisher information scales as α^2 : stronger coupling yields more precise estimation.
4. Fisher information depends on ϕ_γ : observations at $\phi_\gamma = \pi/2$ (zero-crossings of carrier) are uninformative.

This example illustrates the core mechanism: nested frequency structure, through phase-amplitude coupling, creates statistical dependencies that “lift” hidden slow variables into the observable fast signal.

4.3 What Coupling Does *Not* Do

To prevent misunderstanding, we clarify what Theorem 7 does *not* claim:

1. **Coupling does not increase dimension beyond the parameter space.** If you have $K - 1$ relative phases, the Fisher matrix is at most $(K - 1) \times (K - 1)$ with rank at most $K - 1$. Coupling cannot create dimensions that don't exist.
2. **Strong coupling can reduce effective dimension.** Phase-locking (complete synchronization) imposes constraints that *reduce* the number of free parameters. Our "activation" result concerns the regime where coupling creates informative dependencies, not hard constraints.
3. **Identifiability requires appropriate observations.** If your measurement is insensitive to amplitude modulation, coupling won't help. The result is model-dependent.

5 Clock Synchronization Collapses Dimension

We now characterize what happens when all oscillators are locked to a global clock.

Definition 9 (Clock synchronization). *A system of K oscillators is clock-synchronized if all phases are deterministic functions of a global time reference:*

$$\phi_k(t) = \omega_k t + \phi_0 \quad \text{for all } k \tag{19}$$

where ϕ_0 is the initial clock phase and $\omega_k = n_k \omega_0$ are integer multiples of a base frequency ω_0 .

Under clock synchronization, the relative phases become:

$$\psi_j(t) = \phi_j(t) - \phi_1(t) = (n_j - n_1)\omega_0 t \tag{20}$$

These are deterministic functions of time, not free parameters.

Theorem 10 (Clock collapse). *Under clock synchronization:*

- (i) *The relative phases $\psi_j(t)$ are deterministic functions of the single parameter t (or equivalently, ϕ_0).*
- (ii) *The effective parameter dimension is at most 1: only the time offset ϕ_0 (equivalently, the clock phase) is a free parameter.*
- (iii) *The trajectory $\psi(t)$ traces a 1-dimensional curve in \mathbb{T}^{K-1} . At any fixed time, the system occupies a single point, not a distribution over phases.*

Proof. (i) By definition, $\psi_j(t) = (n_j - n_1)\omega_0 t$ is determined by t .

(ii) Any observation $y(t)$ depends on $\phi(t)$, which depends only on t and ϕ_0 . The likelihood $p(y|\phi_0)$ is parameterized by a single parameter, so Fisher rank ≤ 1 .

(iii) The map $t \mapsto \psi(t)$ is a curve in \mathbb{T}^{K-1} . If frequencies are commensurate (all n_j integers), this is a closed orbit; if incommensurate, it winds densely. Either way, it is 1-dimensional. \square

Interpretation. A global clock destroys the product structure of independent phases. Instead of $K - 1$ free relative phases, you have one clock phase that determines everything. The system cannot maintain high-dimensional phase relationships because all timing is slaved to the clock.

5.1 Application to Digital Systems

Digital computers with a global clock operate in this regime: all state transitions are synchronized to clock edges, and relative timing between operations is fixed to integer clock cycles.

Remark 11 (Nuances for digital systems). *The clock-collapse result applies to ideal globally-clocked systems. Real digital systems have complications:*

1. **Multiple clock domains:** Modern processors have multiple asynchronous clock domains, partially recovering independent phase relationships.
2. **Clock jitter:** Thermal and quantum noise create phase uncertainty, though typically small.
3. **Asynchronous circuits:** Clockless designs exist and do not suffer clock collapse.

The contrast we emphasize is between systems designed around a global timing reference versus systems with intrinsically asynchronous, hierarchically coupled timing. The former collapse relative phase degrees of freedom; the latter preserve them.

Remark 12 (Dimension in digital systems). *Digital systems have enormous discrete state spaces (exponential in bit count). Theorem 10 concerns continuous relative phase degrees of freedom, not discrete state dimension. A clocked digital computer gains combinatorial complexity in bits but loses the kind of continuous multiscale phase manifold that characterizes biological oscillatory systems.*

6 Summary of Statistical Structure

We summarize the Fisher information structure under different architectures:

Architecture	Parameter space	Fisher structure	Identifiable dim.
K independent bands	\mathbb{T}^{K-1}	Diagonal (or zero)	$\leq K - 1$
K nested bands (coupled)	\mathbb{T}^{K-1}	Dense, off-diagonal	$\leq K - 1^*$
Clock-synchronized	1D (time offset)	Scalar	≤ 1

Table 1: Fisher information structure under different oscillatory architectures. *Coupling activates identifiability of phases that are unidentifiable under independence; see Example 8.

Corollary 13 (Identifiability gap). *Under the nested frequency observation model (Definition 3):*

1. *With K coupled bands and generic modulation depths, all $K - 1$ relative phases are identifiable from fast-band observations.*
2. *Under global clock synchronization, at most 1 parameter (time offset) is identifiable.*

The identifiability gap between nested biological systems and clock-synchronized systems is $K - 2$ dimensions.

For neural systems with 4–6 canonical frequency bands (delta, theta, alpha, beta, gamma, high-gamma), this represents 3–5 identifiable relative phases under coupling, versus at most 1 under global clock synchronization.

7 Discussion

7.1 What Nesting Contributes Beyond Multiple Oscillators

A reviewer might ask: “Any K oscillators give a K -torus; what is special about nesting?”

The answer is in the observation model. *Nesting* refers to the hierarchical structure where slow phases modulate fast dynamics:

1. **Slow phases are not directly observed**—in typical recordings, you measure fast activity (spikes, gamma, etc.), not the slow envelope directly.
2. **Phase-amplitude coupling makes slow phases observable through fast signals**—this is the “activation” of Theorem 7.
3. **The hierarchy provides natural multiscale coordinates**—slow phases set context, fast phases carry content, and the nesting structure determines which variables can be inferred from which observations.

Without the nesting structure (phase-amplitude coupling), independent oscillators would give a product space with diagonal (or zero) Fisher information—you could estimate phases you directly observe, but not others. Nesting creates the statistical dependencies that make the full phase structure accessible.

7.2 Why Biology Uses Nested Frequencies

The ubiquity of nested frequency hierarchies in biological systems [2, 6] is often attributed to functional requirements: temporal segregation, multiplexing, context-dependent processing.

Our analysis suggests a complementary geometric explanation: **nested frequencies are how biological systems create identifiable multiscale structure**. The slow contextual variables (mood, task state, circadian phase) become estimable from fast observables (neural spikes, gene expression bursts) precisely because of cross-scale coupling.

This may explain why biological systems consistently employ nested architectures rather than independent oscillators: the nesting structure maximizes what can be inferred about system state from available observations.

7.3 Implications for Artificial Systems

If high-dimensional identifiable structure requires nested frequency architectures with phase-amplitude coupling, then achieving similar capabilities in artificial systems may require:

1. Multiple oscillatory timescales (not a single global clock)
2. Asynchronous phase relationships between scales
3. Coupling mechanisms that make slow state visible in fast dynamics

Neuromorphic computing architectures that incorporate these features [5, 8] may achieve richer statistical structure than globally clocked systems. The key is not merely parallelism or speed, but the geometric structure of timing relationships.

7.4 Limitations

This analysis makes several idealizations:

1. **Idealized oscillator models:** Real biological oscillations have amplitude dynamics, non-sinusoidal waveforms, and complex coupling functions.
2. **Gaussian noise:** The explicit calculations assume Gaussian observation noise; non-Gaussian cases may differ quantitatively.
3. **Known carrier phase:** The theta-gamma example conditions on known ϕ_γ ; in practice, both phases must be estimated jointly.
4. **Single-sample Fisher information:** We compute Fisher information for single observations; temporal sequences provide additional structure.

Despite these limitations, the core insight—that cross-scale coupling activates identifiability of slow variables from fast observations—should be robust across more realistic models.

8 Conclusion

Nested frequency hierarchies are not merely a convenient encoding scheme. They are a mechanism for creating *identifiable* high-dimensional statistical structure. Each frequency band contributes a potential phase parameter; cross-scale coupling (phase-amplitude modulation) makes these parameters jointly estimable from partial observations; and the resulting Fisher information structure has dimension scaling with the number of nested bands.

Global clock synchronization destroys this structure. By making all relative phases deterministic functions of a single clock, synchronization collapses the identifiable parameter family to one dimension. This is a geometric fact, not an implementation detail.

The identifiability gap between nested biological systems and clock-synchronized artificial systems provides one explanation for persistent differences in their capabilities. Whether this gap can be bridged by alternative computing architectures—those that preserve multiscale asynchronous phase structure—remains an important open question.

Acknowledgements

The author thanks anonymous reviewers for detailed feedback that substantially improved the manuscript.

Declarations

Funding. The author received no specific funding for this work.

Conflicts of interest. The author has no conflicts of interest to declare.

Data availability. No datasets were generated or analyzed.

AI assistance. Claude Code with Opus 4.5 (Anthropic) was used for drafting. The author reviewed all content and takes full responsibility.

References

- [1] Amari, S.-I. (2016). *Information Geometry and Its Applications*. Springer.
- [2] Buzsáki, G. (2006). *Rhythms of the Brain*. Oxford University Press.
- [3] Canolty, R. T., Knight, R. T. (2010). The functional role of cross-frequency coupling. *Trends in Cognitive Sciences*, 14(11), 506–515.
- [4] Helm, B., et al. (2017). Annual rhythms that underlie phenology: biological time-keeping meets environmental change. *Proceedings of the Royal Society B*, 284, 20170046.
- [5] Indiveri, G., et al. (2011). Neuromorphic silicon neuron circuits. *Frontiers in Neuroscience*, 5, 73.
- [6] Lisman, J. E., Jensen, O. (2013). The theta-gamma neural code. *Neuron*, 77, 1002–1016.
- [7] Lloyd, D., Rossi, E. L. (2006). *Ultradian Rhythms from Molecules to Mind*. Springer.
- [8] Schuman, C. D., et al. (2017). A survey of neuromorphic computing and neural networks in hardware. *arXiv:1705.06963*.
- [9] Todd, I. (2026). Intelligence as high-dimensional coherence: The observable dimensionality bound and computational tractability. *BioSystems*, 258, 105704.