Alignment Probabilities on Product Statistical Manifolds: Fisher Information and Coordination Depth

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

Abstract

We study the probability of simultaneous alignment events on product statistical manifolds, with applications to synchronization in coupled oscillator systems. For M random variables drawn independently from a common exponential family on the circle (von Mises distributions with concentration κ), the probability that all variables fall within tolerance ε of each other scales as $p_1(\varepsilon, \kappa)^{M-1}$, where p_1 is the single-variable window probability. The exponent $(M-1)$ rather than M arises from quotient geometry: alignment is invariant under global rotation, so the constraint has *codimension* $M-1$ in the M-torus after quotienting by the diagonal S^1 action. We prove this product formula, characterize the single-variable probability p_1 in terms of Fisher information on the circle, and derive the resulting scaling law for expected hitting times: $\tau \propto p_1^{-(M-1)}$. The framework provides a coordinate-invariant foundation for understanding how "coordination depth" (the number of components requiring simultaneous alignment) determines waiting times in stochastic systems. Applications include neural synchronization, molecular coordination, and distributed consensus.

Keywords: Fisher information; von Mises distribution; product manifolds; hitting times; synchronization; coordination depth

1 Introduction

Many physical and biological systems require *simultaneous alignment* of multiple components before a collective event can occur. Neurons must achieve phase coherence for perceptual binding; molecular machines require conformational alignment for catalysis; distributed agents must reach consensus for coordinated action. A natural question arises: how does the probability of alignment—and hence the expected waiting time—scale with the number of components?

This paper provides an information-geometric answer. We consider M random variables drawn independently from a common distribution on the circle (specifically, the von Mises exponential family). The *alignment event* is the simultaneous occurrence of all variables within a tolerance window ε of each other. We prove:

1. The alignment probability factorizes as $P(\text{align}) = p_1(\varepsilon, \kappa)^{M-1}$, where p_1 is the single-variable window probability and κ is the concentration parameter.

2. The exponent is $(M-1)$ rather than M because alignment is invariant under global rotation: we quotient by the S^1 action, reducing the constraint codimension by one.

3. The single-variable probability $p_1(\varepsilon, \kappa)$ is determined by the Fisher information geometry of the von Mises family, connecting alignment rates to statistical distinguishability.

4. Expected hitting times scale as $\tau \propto p_1^{-(M-1)}$, exponential in coordination depth M.

The coordinate-invariant nature of this result—expressed in terms of Fisher information and manifold dimension—makes it applicable across substrates: neural oscillators, molecular rotors, or abstract phase variables.

1.1 Related Work

The von Mises distribution and its information geometry are classical [**?**, **?**]. Synchronization in coupled oscillators has been extensively studied via the Kuramoto model [**?**, **?**], but typically focused on order parameter dynamics rather than hitting times. Our contribution is connecting the alignment probability to product manifold geometry and Fisher information, yielding a principled derivation of the $(M-1)$ scaling exponent.

2 The Von Mises Family and Fisher Information on the Circle

2.1 The Von Mises Distribution

The von Mises distribution is the maximum-entropy distribution on S^1 for a given mean direction and circular concentration [**?**]. With mean $\mu = 0$ (without loss of generality by rotational symmetry), the density is:

$$p(\theta; \kappa) = \frac{e^{\kappa \cos\theta}}{2\pi I_0(\kappa)}, \quad \theta \in [-\pi, \pi) \tag{1}$$

where $\kappa \geq 0$ is the concentration parameter and $I_0(\kappa)$ is the modified Bessel function of the first kind.

The von Mises family is an exponential family with natural parameter κ and sufficient statistic $\cos\theta$. Key properties:

- $\kappa = 0$: uniform distribution on S^1

- $\kappa \to \infty$: distribution concentrates at $\theta = 0$
- Mean resultant length: $r = \mathbb{E}[\cos\theta] = I_1(\kappa)/I_0(\kappa)$

The relationship $r = I_1(\kappa)/I_0(\kappa)$ provides a bijection between the natural parameter $\kappa \in [0, \infty)$ and the mean resultant $r \in [0, 1)$. For moderate κ, the approximation $\kappa \approx r(2 - r^2)/(1 - r^2)$ is useful [**?**].

2.2 Fisher Information on the Circle

The Fisher information for the von Mises family with respect to κ is the variance of the score:

$$I(\kappa) = \mathbb{E}\left[\left(\frac{\partial \log p}{\partial \kappa}\right)^2\right] = \mathrm{Var}(\cos\theta) = \mathbb{E}[\cos^2\theta] - r^2 \tag{2}$$

Using $\cos^2\theta = (1 + \cos 2\theta)/2$ and $\mathbb{E}[\cos n\theta] = I_n(\kappa)/I_0(\kappa)$:

$$I(\kappa) = \frac{1}{2}\left(1 + \frac{I_2(\kappa)}{I_0(\kappa)}\right) - r^2 \tag{3}$$

Applying the Bessel recurrence $I_2(\kappa) = I_0(\kappa) - (2/\kappa)I_1(\kappa)$:

$$I(\kappa) = 1 - \frac{r}{\kappa} - r^2 \tag{4}$$

For small κ (near uniform), $r \to 0$ and $I(\kappa) \to 1$: samples are maximally informative about κ. For large κ (high concentration), $r \to 1$ and $I(\kappa) \to 0$: the distribution is nearly deterministic, so samples provide little information about small changes in κ.

Remark 1 (Fisher information and circular variance)**.** *The circular variance is $V = 1 - r$. High Fisher information (near 1) corresponds to high circular variance (diffuse distribution); low Fisher information corresponds to low circular variance (concentrated distribution). This inverse relationship connects statistical distinguishability to phase dispersion.*

2.3 Window Probability

Define the *window probability* as the probability mass in an interval of width ε centered at the mode:

$$p_1(\varepsilon, \kappa) = \int_{-\varepsilon/2}^{\varepsilon/2} \frac{e^{\kappa\cos\theta}}{2\pi I_0(\kappa)} d\theta \tag{5}$$

For the uniform distribution ($\kappa = 0$): $p_1(\varepsilon, 0) = \varepsilon/(2\pi)$.

For large κ (Gaussian approximation): $p_1(\varepsilon, \kappa) \approx \mathrm{erf}(\varepsilon\sqrt{\kappa}/2)$, approaching 1 as $\kappa \to \infty$.

The ratio $p_1(\varepsilon, \kappa)/p_1(\varepsilon, 0) = 2\pi p_1(\varepsilon, \kappa)/\varepsilon$ measures how much concentration enhances alignment probability relative to the geometric baseline.

2.4 Explicit Computation of p_1

While no closed form exists for $p_1(\varepsilon, \kappa)$ in general, we derive asymptotic expressions in the key regimes.

Small window ($\varepsilon \to 0$). For ε small relative to the distribution width:

$$p_1(\varepsilon, \kappa) \approx \varepsilon \cdot p(0; \kappa) = \varepsilon \cdot \frac{e^\kappa}{2\pi I_0(\kappa)} \tag{6}$$

The window probability is proportional to the density at the mode times window width.

Uniform distribution ($\kappa = 0$). When $\kappa = 0$, $p(\theta) = 1/(2\pi)$ uniformly, giving:

$$p_1(\varepsilon, 0) = \frac{\varepsilon}{2\pi} \tag{7}$$

exactly, for any $\varepsilon \leq 2\pi$.

High concentration ($\kappa \gg 1$). For large κ, the von Mises distribution approaches a wrapped Gaussian with variance $1/\kappa$ (standard deviation $\sigma = 1/\sqrt{\kappa}$). Integrating from $-\varepsilon/2$ to $\varepsilon/2$:

$$p_1(\varepsilon, \kappa) \approx \operatorname{erf}\left(\frac{\varepsilon\sqrt{\kappa}}{2\sqrt{2}}\right) \tag{8}$$

For $\varepsilon\sqrt{\kappa} \gg 1$, this saturates: $p_1 \to 1$.

Intermediate regime. For moderate κ and ε, numerical integration is required. The key observation is that $p_1(\varepsilon, \kappa)$ interpolates monotonically between the uniform baseline $\varepsilon/(2\pi)$ and unity as κ increases.

Lemma 2 (Monotonicity in κ)**.** *For fixed $\varepsilon > 0$, $p_1(\varepsilon, \kappa)$ is strictly increasing in κ.*

Proof. The von Mises family has monotone likelihood ratio: for $\kappa_2 > \kappa_1$, the ratio $p(\theta; \kappa_2)/p(\theta; \kappa_1)$ is increasing in $\cos\theta$. This implies that probability mass shifts toward $\theta = 0$ as κ increases, strictly increasing $p_1(\varepsilon, \kappa)$ for any centered window. □

Series expansion. Using the Fourier expansion of the von Mises density:

$$p(\theta; \kappa) = \frac{1}{2\pi}\left[1 + 2\sum_{n=1}^{\infty} \frac{I_n(\kappa)}{I_0(\kappa)} \cos(n\theta)\right] \tag{9}$$

we obtain:

$$p_1(\varepsilon, \kappa) = \frac{\varepsilon}{2\pi} + \frac{2}{\pi}\sum_{n=1}^{\infty} \frac{I_n(\kappa)}{I_0(\kappa)} \cdot \frac{\sin(n\varepsilon/2)}{n} \tag{10}$$

The first term is the uniform contribution; subsequent terms capture concentration-induced enhancement.

3 Product Manifolds and Alignment Geometry

3.1 The M-Torus as Product Manifold

Consider M independent random variables $\theta_1, \ldots, \theta_M$, each distributed as von Mises with common concentration κ. The joint distribution lives on the M-torus $T^M = S^1 \times \cdots \times S^1$:

$$p(\theta_1, \ldots, \theta_M; \kappa) = \prod_{i=1}^{M} \frac{e^{\kappa \cos \theta_i}}{2\pi I_0(\kappa)} \tag{11}$$

The statistical manifold is the product $\mathcal{P}^M = \mathcal{P}_1 \times \cdots \times \mathcal{P}_M$, where each $\mathcal{P}_i$ is the von Mises family on the i-th circle. The Fisher metric on $\mathcal{P}^M$ is block-diagonal, with each block equal to $I(\kappa)$.

3.2 The Alignment Region

Define the *alignment region* $A_\varepsilon \subset T^M$ as:

$$A_\varepsilon = \{(\theta_1, \ldots, \theta_M) : |\theta_i - \theta_1| < \varepsilon/2 \text{ for } i = 2, \ldots, M\} \tag{12}$$

This is the set where all phases lie within $\varepsilon/2$ of the reference phase θ_1—equivalently, all phases lie within a window of width ε centered on the reference. Throughout, ε denotes the **full window width**.

Remark 3 (Rotational invariance)**.** *The alignment region A_ε is invariant under the diagonal S^1 action: $(\theta_1, \ldots, \theta_M) \mapsto (\theta_1 + \phi, \ldots, \theta_M + \phi)$ for any $\phi \in S^1$. This symmetry is crucial for the $(M-1)$ exponent.*

3.3 Quotient Geometry

Let $\bar{\theta} = (1/M)\sum_i \theta_i$ be the mean phase. The alignment constraint depends only on the *relative phases* $\phi_i = \theta_i - \bar{\theta}$, which satisfy $\sum_i \phi_i = 0$. The constraint $|\theta_i - \theta_j| < \varepsilon$ for all pairs is equivalent to all ϕ_i lying in an interval of width ε.

The quotient T^M/S^1 (identifying configurations that differ by global rotation) is an $(M-1)$-dimensional space parameterized by the relative phases. The alignment region A_ε projects to a region in this quotient of dimension $(M-1)$.

3.3.1 Formal Structure

Consider the M-torus $T^M = S^1 \times \cdots \times S^1$ with the product topology and smooth structure. The diagonal S^1 action is:

$$\rho_\phi : (\theta_1, \ldots, \theta_M) \mapsto (\theta_1 + \phi, \ldots, \theta_M + \phi) \tag{13}$$

for $\phi \in S^1$. This action is free (no fixed points) and proper, so the quotient T^M/S^1 is a smooth manifold.

Proposition 4 (Quotient structure)**.** *The quotient T^M/S^1 is diffeomorphic to T^{M-1}.*

Proof. Define the projection $\pi : T^M \to T^{M-1}$ by:

$$\pi(\theta_1, \ldots, \theta_M) = (\theta_2 - \theta_1, \theta_3 - \theta_1, \ldots, \theta_M - \theta_1) \tag{14}$$

This map is well-defined modulo 2π, smooth, and surjective. The fibers $\pi^{-1}(y)$ are exactly the S^1-orbits. By the slice theorem, π descends to a diffeomorphism $T^M/S^1 \cong T^{M-1}$. □

3.3.2 Coordinates on the Quotient

The coordinates $\psi_i = \theta_{i+1} - \theta_1$ for $i = 1, \ldots, M-1$ provide a global chart on T^{M-1}. In these coordinates:

- The alignment region becomes $\{|\psi_i| < \varepsilon/2 : i = 1, \ldots, M-1\}$
- The product measure on T^M descends to a measure on T^{M-1} (after integrating out the global phase)
- The alignment probability in quotient coordinates is exactly $p_1(\varepsilon, \kappa)^{M-1}$

3.3.3 The Codimension Argument

The alignment region $A_\varepsilon \subset T^M$ has dimension M (it is an open set). However, it is invariant under the S^1 action, so its image in T^{M-1} has dimension $M-1$.

Alternatively: the alignment constraint imposes $M-1$ independent conditions ($|\theta_i - \theta_1| < \varepsilon/2$ for $i = 2, \ldots, M$, or equivalently $|\psi_i| < \varepsilon/2$ for $i = 1, \ldots, M-1$). This gives codimension $M-1$ in T^M, or equivalently, dimension 1 (the free global phase).

Remark 5 (Why $(M-1)$, not $\binom{M}{2}$)**.** *One might expect $\binom{M}{2}$ constraints (all pairwise differences small), but these are not independent. Using θ_1 as reference, the $M-1$ constraints $|\theta_i - \theta_1| < \varepsilon/2$ imply $|\theta_i - \theta_j| < \varepsilon$ by the triangle inequality. The $M-1$ reference-based constraints fully determine the alignment event.*

4 Main Result: The $(M-1)$ Exponent

Theorem 6 (Alignment probability on product manifolds)**.** *Let $\theta_1, \ldots, \theta_M$ be independent von Mises random variables with common concentration κ. The probability that all phases lie within ε of each other is:*

$$P(A_\varepsilon) = p_1(\varepsilon, \kappa)^{M-1} \tag{15}$$

where $p_1(\varepsilon, \kappa)$ is the single-variable window probability.

Proof. By rotational invariance, we may condition on $\theta_1 = 0$ (or equivalently, use θ_1 as the reference phase). The alignment event A_ε then requires each of $\theta_2, \ldots, \theta_M$ to lie in the window $(-\varepsilon/2, \varepsilon/2)$.

Since the θ_i are independent and identically distributed:

$$P(A_\varepsilon) = P(\theta_2 \in (-\varepsilon/2, \varepsilon/2), \ldots, \theta_M \in (-\varepsilon/2, \varepsilon/2) \,|\, \theta_1 = 0) \tag{16}$$

$$= \prod_{i=2}^{M} P(\theta_i \in (-\varepsilon/2, \varepsilon/2)) \tag{17}$$

$$= p_1(\varepsilon, \kappa)^{M-1} \tag{18}$$

The exponent is $(M-1)$ because θ_1 serves as the reference; only the remaining $M-1$ phases must independently fall within the tolerance window. □

Corollary 7 (Hitting time scaling)**.** *If phases explore the circle at rate $\Delta\omega$ (e.g., via diffusion or frequency mismatch), the expected hitting time to alignment satisfies:*

$$\mathbb{E}[\tau] \propto \frac{1}{\Delta\omega} \cdot p_1(\varepsilon, \kappa)^{-(M-1)} \tag{19}$$

Equivalently, $\log \mathbb{E}[\tau] \propto (M-1)\log(1/p_1)$*: logarithmic hitting time is linear in coordination depth.*

Proof. For a renewal process with success probability p per attempt and attempt rate $\Delta\omega$, the expected waiting time is $\mathbb{E}[\tau] = 1/(\Delta\omega \cdot p)$. Substituting $p = P(A_\varepsilon) = p_1^{M-1}$ yields the result. □

Remark 8 (Geometric interpretation)**.** *The $(M-1)$ exponent has a clean geometric meaning: the alignment constraint removes $M-1$ degrees of freedom from the M-torus. Equivalently, after quotienting by global rotation, the relative-phase space is $(M-1)$-dimensional, and alignment requires all $M-1$ relative phases to independently satisfy a window constraint.*

5 Fisher Information Characterization

5.1 Window Probability and Concentration

The window probability $p_1(\varepsilon, \kappa)$ interpolates between two limits:

- **Uniform limit** ($\kappa = 0$): $p_1 = \varepsilon/(2\pi)$
- **Concentrated limit** ($\kappa \to \infty$): $p_1 \to 1$

For intermediate κ, we can write:

$$p_1(\varepsilon, \kappa) = \frac{\varepsilon}{2\pi} \cdot g(\kappa) \tag{20}$$

where $g(\kappa) = 2\pi p_1(\varepsilon, \kappa)/\varepsilon$ is the *concentration enhancement factor*, satisfying $g(0) = 1$ and $g(\kappa) \to 2\pi/\varepsilon$ as $\kappa \to \infty$.

5.2 Connection to Fisher Information

The concentration κ and Fisher information $I(\kappa)$ are inversely related: high κ means concentrated distribution (low Fisher information), enabling fast alignment. We can express the hitting time scaling in terms of Fisher information:

Proposition 9 (Fisher-alignment duality)**.** *Let $V(\kappa) = 1 - r(\kappa)$ be the circular variance. Then:*

$$\log p_1(\varepsilon, \kappa) \approx \log(\varepsilon/2\pi) + \alpha(\varepsilon) \cdot (1 - V(\kappa)) \tag{21}$$

for a function α depending on ε. Consequently:

$$\log \mathbb{E}[\tau] \propto (M-1) \cdot V(\kappa) \cdot |\log(\varepsilon/2\pi)| \tag{22}$$

Higher circular variance (more phase dispersion, higher Fisher information) leads to longer hitting times.

For the von Mises family, this formalizes the intuition: concentrated distributions (low V, low I) align quickly; dispersed distributions (high V, high I) align slowly.

6 Beyond Independence: Weak Coupling and the α Parameter

The product formula $P(A_\varepsilon) = p_1^{M-1}$ assumes independence across components. In physical systems—coupled oscillators, interacting molecules, communicating agents—components are correlated. Here we derive how weak coupling modifies the alignment probability, yielding the topology-dependent parameter α.

6.1 The Kuramoto Model as Statistical Model

Consider N coupled phase oscillators with dynamics:

$$d\theta_i = \omega_i \, dt + \frac{K}{N} \sum_{j=1}^{N} A_{ij} \sin(\theta_j - \theta_i) \, dt + \sigma \, dW_i \tag{23}$$

where ω_i are natural frequencies (drawn from $g(\omega)$), K is coupling strength, A_{ij} encodes network topology, and σ provides phase noise.

For $K > 0$, the stationary distribution is *not* a product of independent von Mises distributions. However, for weak coupling or modular topologies, we can characterize the deviation from independence.

6.2 Modular Decomposition

Suppose the network has M modules, each internally strongly coupled but weakly coupled to other modules. Let K_{intra} be intra-module coupling and $K_{\text{inter}} \ll K_{\text{intra}}$ be inter-module coupling.

Within each module m, the phases synchronize to a common module phase Φ_m with small fluctuations. The module phases $(\Phi_1, \ldots, \Phi_M)$ evolve on a slower timescale, governed by the weak inter-module coupling.

Definition 10 (Module order parameter)**.** *For module m containing oscillators $\{i : i \in m\}$, define:*

$$r_m e^{i\Phi_m} = \frac{1}{|m|} \sum_{i \in m} e^{i\theta_i} \tag{24}$$

where r_m is the intra-module coherence and Φ_m is the module phase.

In the modular regime, $r_m \approx 1$ (strong intra-module synchronization) while the module phases $\Phi_1, \ldots, \Phi_M$ remain weakly correlated.

6.3 Effective Distribution on Module Phases

The stationary distribution of module phases can be approximated as:

$$p(\Phi_1, \ldots, \Phi_M) \approx \frac{1}{Z} \exp\left[\kappa_0 \sum_m \cos \Phi_m + \kappa_{\text{int}} \sum_{m<n} \cos(\Phi_m - \Phi_n)\right] \tag{25}$$

where:

- κ_0 captures the marginal concentration (from natural frequency spread and noise)
- $\kappa_{\text{int}} \propto K_{\text{inter}}$ captures inter-module coupling
- Z is the normalization constant

For $\kappa_{\text{int}} = 0$ (no inter-module coupling), this reduces to a product of independent von Mises distributions, recovering the $(M-1)$ formula exactly.

6.4 Perturbative Expansion

For small κ_{int}, expand the alignment probability:

$$P(A_\varepsilon) = P_0(A_\varepsilon) + \kappa_{\text{int}} \cdot P_1(A_\varepsilon) + O(\kappa_{\text{int}}^2) \tag{26}$$

where $P_0 = p_1(\varepsilon, \kappa_0)^{M-1}$ is the independent baseline.

The first-order correction P_1 depends on the topology of inter-module connections. For pairwise coupling:

$$P_1(A_\varepsilon) = \sum_{m<n} C_{mn}(\varepsilon, \kappa_0) \cdot p_1^{M-3} \tag{27}$$

where C_{mn} is a correlation integral over the alignment region.

Proposition 11 (Effective exponent under weak coupling)**.** *For weak inter-module coupling $\kappa_{\text{int}} \ll 1$, the alignment probability can be* approximated *as:*

$$P(A_\varepsilon) \approx p_1(\varepsilon, \kappa_{\text{eff}})^{\alpha(M-1)} \tag{28}$$

where:

- κ_{eff} *is an effective concentration incorporating both* κ_0 *and coupling effects*
- $\alpha = 1 - \beta(\textit{topology}, \kappa_{\text{int}})$ *with* $\beta \geq 0$ *a topology-dependent correction*

The correction β *is largest for densely connected topologies (all-to-all) and smallest for sparse topologies.*

Proof sketch. Write $\log P(A_\varepsilon) = (M-1)\log p_1 + \Delta(\kappa_{\text{int}})$ where Δ captures the coupling correction. For small κ_{int}:

$$\Delta \approx \kappa_{\text{int}} \cdot |E| \cdot f(\varepsilon, \kappa_0) \tag{29}$$

where $|E|$ is the number of inter-module edges and f is a positive function (coupling facilitates alignment). Factoring:

$$\log P(A_\varepsilon) = (M-1)\log p_1 \left[1 + \frac{\kappa_{\text{int}}|E|f}{(M-1)\log p_1}\right] \tag{30}$$

Defining $\alpha = 1 - \kappa_{\text{int}}|E|f/[(M-1)|\log p_1|]$ yields the stated form. □

6.5 Topology Dependence of α

The correction $\beta = 1 - \alpha$ depends on:

1. **Edge density**: More inter-module connections ⇒ stronger correlation ⇒ larger β
2. **Coupling strength**: Larger $K_{\text{inter}}/K_{\text{intra}}$ ⇒ larger β
3. **Module coherence**: Higher intra-module r_m ⇒ cleaner separation ⇒ smaller β

Table 1: Qualitative α regimes by inter-module coupling (schematic)

Regime	Inter-module coupling	Predicted α
Independent modules	None	$\alpha = 1$ (exact)
Weak inter-module	Few connections, weak	$\alpha \approx 0.7$–0.9
Moderate inter-module	Sparse, moderate strength	$\alpha \approx 0.4$–0.6
Strong inter-module	Dense or strong	$\alpha \approx 0.2$–0.4

These are order-of-magnitude expectations based on the perturbative analysis, not formal bounds; exact values depend on specific coupling parameters. The key principle: stronger inter-module coupling ⇒ stronger correlation ⇒ lower effective exponent α.

This explains the empirically observed variation in α across topologies: the parameter captures how much inter-module correlation deviates from the independent-product baseline.

6.6 Information-Geometric Interpretation

The α parameter has a clean information-geometric meaning:

Proposition 12 (as correlation dimension)**.** *Define the* correlation dimension *of the module-phase distribution as:*

$$d_{\text{corr}} = \frac{\textit{(effective degrees of freedom)}}{\textit{(nominal degrees of freedom)}} = \frac{d_{\text{eff}}}{M-1} \tag{31}$$

where d_{eff} is the participation ratio of the Fisher information matrix eigenvalues. Then $\alpha \approx d_{\text{corr}}$.

For independent modules, the Fisher information matrix is diagonal with equal entries, giving $d_{\text{eff}} = M-1$ and $\alpha = 1$. Coupling introduces off-diagonal structure, reducing effective dimension and hence $\alpha < 1$.

This connects α to the information-geometric concept of *effective model dimension*: coupling reduces the number of independent directions in parameter space, slowing the rate at which alignment probability decays with M.

7 Numerical Validation

To validate the theoretical predictions, we simulated modular Kuramoto networks with varying inter-module coupling.

7.1 Setup

- $N = 100$ oscillators partitioned into $M = 3$ to 10 modules
- Intra-module coupling $K_{\text{intra}} = 1.0$
- Inter-module coupling $K_{\text{inter}} \in \{0, 0.05, 0.15, 0.30\}$
- Natural frequencies $\omega_i \sim \mathcal{N}(0, 0.3)$
- Phase noise $\sigma = 0.1$
- Alignment tolerance $\varepsilon = 2.0$ rad
- 20 trials per (M, K_{inter}) combination

7.2 Results

Fitting $\log \tau = \hat{\alpha}(1-\bar{r})(M-1)\log(2\pi/\varepsilon) + c$:

Key observations:

1. At $K_{\text{inter}} = 0$, $\hat{\alpha} \approx 1$ as predicted by the product formula
2. Increasing inter-module coupling decreases α, as predicted by Proposition **??**
3. The R^2 decreases with coupling strength, reflecting breakdown of the weak-coupling approximation

Table 2: Fitted $\hat{\alpha}$ vs inter-module coupling strength

K_{inter}	$\bar{r}$	$\hat{\alpha}$	R^2
0 (independent)	0.91	0.98	0.94
0.05 (weak)	0.93	0.52	0.81
0.15 (moderate)	0.94	0.35	0.71
0.30 (strong)	0.96	0.21	0.58

8 Worked Examples

We illustrate the theoretical framework with concrete calculations across three parameter regimes.

8.1 Example 1: Uniform Phases (Diffusive Limit)

Consider $M = 5$ independent phases uniformly distributed on S^1 ($\kappa = 0$). For alignment tolerance $\varepsilon = \pi/6$ (30°):

Single-window probability:

$$p_1(\pi/6, 0) = \frac{\pi/6}{2\pi} = \frac{1}{12} \approx 0.083 \tag{32}$$

Alignment probability:

$$P(A_\varepsilon) = p_1^{M-1} = \left(\frac{1}{12}\right)^4 = \frac{1}{20736} \approx 4.8 \times 10^{-5} \tag{33}$$

Expected hitting time: If phases explore at rate $\Delta\omega = 1$ rad/s (one cycle per 2π seconds):

$$\mathbb{E}[\tau] \sim \frac{1}{\Delta\omega \cdot P(A_\varepsilon)} \approx 2.1 \times 10^4 \text{ s} \approx 5.8 \text{ hours} \tag{34}$$

Adding a sixth component increases the exponent by 1:

$$P(A_\varepsilon; M = 6) = \left(\frac{1}{12}\right)^5 \approx 4.0 \times 10^{-6} \tag{35}$$

a 12-fold reduction, demonstrating the exponential penalty.

8.2 Example 2: Moderate Concentration (Neural Oscillators)

Consider neural gamma oscillators (∼40 Hz) with concentration $\kappa = 3$ (typical for cortical phase distributions). For $M = 4$ modules and $\varepsilon = \pi/4$ (45°):

Single-window probability: Using the Fourier series with $I_1(3)/I_0(3) \approx 0.787$:

$$p_1(\pi/4, 3) \approx \frac{\pi/4}{2\pi}\left[1 + 2 \times 0.787 \times \frac{\sin(\pi/8)}{\pi/8}\right] \approx 0.125 \times 2.52 \approx 0.32 \tag{36}$$

Alignment probability:

$$P(A_\varepsilon) = (0.32)^3 \approx 0.033 \tag{37}$$

Comparison to uniform: Concentration increases alignment probability by a factor:

$$\frac{p_1(3)^3}{p_1(0)^3} = \left(\frac{0.32}{0.125}\right)^3 \approx 17 \tag{38}$$

Concentration reduces waiting time 17-fold.

8.3 Example 3: High Concentration with Coupling Correction

Consider strongly synchronized modules ($\kappa = 10$, near-Gaussian) with weak inter-module coupling giving $\alpha = 0.4$. For $M = 8$ modules and $\varepsilon = \pi/3$ (60°):

Single-window probability (Gaussian approximation):

$$p_1(\pi/3, 10) \approx \text{erf}\left(\frac{(\pi/3)\sqrt{10}}{2\sqrt{2}}\right) = \text{erf}(1.17) \approx 0.90 \tag{39}$$

Alignment probability without coupling ($\alpha = 1$):

$$P_{\alpha=1} = (0.90)^7 \approx 0.48 \tag{40}$$

Alignment probability with coupling ($\alpha = 0.4$):

$$P_{\alpha=0.4} = (0.90)^{0.4\times 7} = (0.90)^{2.8} \approx 0.75 \tag{41}$$

The coupling correction increases alignment probability from 48% to 75%—the reduction in effective exponent from 7 to 2.8 substantially eases alignment. The relative change in $\log P$:

$$\frac{\log P_{\alpha=0.4}}{\log P_{\alpha=1}} = \frac{0.4 \times 7 \times \log(0.90)}{7 \times \log(0.90)} = 0.4 \tag{42}$$

For lower p_1 (harder alignment), the coupling correction has even larger absolute effect.

8.4 Scaling Summary

The table below summarizes alignment probabilities across parameter combinations:

The dominant effect is the exponential penalty in M. Increasing concentration κ provides polynomial mitigation but cannot overcome large M.

9 Discussion

9.1 Summary of Contributions

This paper provides:

Table 3: Alignment probabilities $P(A_\varepsilon)$ for various (M, κ, ε)

M	κ	ε	p_1	$P(A_\varepsilon) = p_1^{M-1}$
3	0	$\pi/6$	0.083	0.0069
3	3	$\pi/6$	0.25	0.063
5	0	$\pi/6$	0.083	4.8×10^{-5}
5	3	$\pi/6$	0.25	0.0039
5	10	$\pi/6$	0.59	0.12
8	3	$\pi/4$	0.32	0.0011
10	0	$\pi/4$	0.125	1.9×10^{-9}

1. A rigorous derivation of the $(M-1)$ exponent for alignment probabilities on product manifolds (Theorem **??**)

2. A geometric explanation: the exponent equals the codimension of the alignment constraint after quotienting by global rotation

3. Connection to Fisher information: alignment rates are determined by the statistical geometry of the component distributions

4. Hitting time scaling (Corollary **??**): $\tau \propto p_1^{-(M-1)}$, exponential in coordination depth

9.2 Applications

The framework applies to any system where multiple stochastic components must achieve simultaneous alignment:

- **Neural synchronization**: Phase alignment across cortical modules for perceptual binding
- **Molecular machines**: Conformational alignment of subunits for catalysis
- **Distributed consensus**: Agreement among agents with noisy communication
- **Coupled oscillators**: Kuramoto-type synchronization with finite tolerance

In each case, the $(M-1)$ scaling provides a principled prediction for how coordination depth affects waiting times.

9.3 Relation to Companion Work

This paper provides the mathematical foundation for a companion paper on neural coherence time scaling (Todd, under review at BioSystems), which applies the $(M-1)$ framework to cortical gamma oscillations.

The division of labor:

- **This paper** (Information Geometry): Derives the $(M-1)$ exponent from first principles, proves the product formula, characterizes p_1 via Fisher information, and shows how coupling yields $\alpha < 1$.

- **Companion paper** (BioSystems): Applies the framework to empirical neural data, estimates α from multi-electrode recordings, tests the scaling law across species and brain regions, and interprets coherence time bounds in terms of cognitive function.

The key empirical finding is that neural oscillators exhibit $\alpha \approx 0.3$–0.5, consistent with the moderate inter-module coupling predicted by Proposition **??**. This sub-unity α relaxes the exponential penalty somewhat, but the $(M-1)$ scaling still imposes strong constraints on multi-region coordination.

Why two papers? The mathematical content (product manifold geometry, Fisher information characterization, quotient space analysis) is independent of the neural application and merits separate treatment in an information-geometry venue. The companion paper focuses on empirical validation, biological interpretation, and cognitive implications, appropriate for a systems biology journal.

9.4 Limitations and Extensions

Independence assumption. The product structure assumes independence across components. For coupled systems (e.g., Kuramoto oscillators with nonzero coupling), the phases are correlated and the product formula is approximate. Section 6 provides a perturbative analysis showing how coupling modifies the exponent, but strongly coupled systems may require different techniques.

Stationary distribution. We assume components have reached their stationary (von Mises) distribution. Transient dynamics may exhibit different scaling; in particular, the hitting time formula assumes renewal-like behavior which may not hold during transient synchronization.

Beyond von Mises. The framework extends to other exponential families on compact spaces, with the window probability determined by the appropriate Fisher information. For non-circular state spaces (e.g., tori, spheres, or Lie groups), the quotient geometry argument generalizes but the specific form of p_1 changes.

Heterogeneous concentrations. We assumed common κ across all components. For heterogeneous concentrations $\kappa_1, \ldots, \kappa_M$, the product formula becomes:

$$P(A_\varepsilon) = \prod_{i=2}^{M} p_1(\varepsilon, \kappa_i) \tag{43}$$

(taking θ_1 as reference). The exponent remains $(M-1)$ but the effective base varies.

Finite-time corrections. The hitting time scaling $\tau \propto p_1^{-(M-1)}$ assumes many independent alignment attempts. For fast systems or tight tolerances, finite-time correlations between attempts may modify the prefactor.

10 Conclusion

The probability of simultaneous alignment across M independent circular random variables scales as p_1^{M-1}, where p_1 is the single-component window probability. The exponent $(M-1)$—not M—arises from the quotient geometry: one component serves as reference, leaving $M-1$ independent constraints. This provides a coordinate-invariant, information-geometric foundation for understanding how coordination depth determines waiting times in stochastic systems.

The key insight is that "coordination depth" is a geometric quantity: it counts the codimension of the alignment constraint on the product manifold. Fisher information enters through the single-component window probability, connecting statistical distinguishability to alignment rates.

Code availability: Supplementary code is available at `https://github.com/todd866/alignment-geometry`

References

[1] K. V. Mardia and P. E. Jupp. *Directional Statistics*. Wiley, 2000.

[2] S.-I. Amari. *Information Geometry and Its Applications*. Springer, 2016.

[3] S. H. Strogatz. From Kuramoto to Crawford: exploring the onset of synchronization in populations of coupled oscillators. *Physica D* 143 (2000), 1–20.

[4] J. A. Acebrón, L. L. Bonilla, C. J. Pérez Vicente, F. Ritort, and R. Spigler. The Kuramoto model: A simple paradigm for synchronization phenomena. *Rev. Mod. Phys.* 77 (2005), 137–185.