

Communication Beyond Information: Manifold Expansion via High-Dimensional Coupling

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

Abstract

Many information-theoretic bounds—in learning, estimation, and communication—assume a fixed statistical model class. We identify a regime where this assumption can fail: when high-dimensional systems couple, the *identifiable* parameter set can change because the image rank of the dynamics-to-distribution map changes. (We do not claim violations of information inequalities; we track geometric changes in the accessible family.) We formalize this as *manifold expansion*: a coupling-induced increase in the Fisher rank of the accessible statistical family. We show that coupling can produce *superadditive* growth in accessible complexity: the coupled system exhibits identifiable degrees of freedom not attributable to any subsystem alone, so that the whole exceeds the sum of its parts ($r(\kappa) > \sum_i r_i$). Our main theorem provides two checkable criteria for rank transitions: the *transversality criterion* (coupling moves the accessible family off a constraint submanifold) and the *symmetry-breaking criterion* (coupling breaks a group invariance, making orbit directions identifiable). We prove that Fisher rank equals Jacobian rank of the dynamics-to-distribution map, so rank transitions are coordinate-invariant. Two worked examples demonstrate the criteria: coupled Ornstein–Uhlenbeck processes (transversality, $\kappa_c = 0$) and Kuramoto oscillators (symmetry breaking, $\kappa_c > 0$), both exhibiting strict superadditivity. The framework yields testable predictions: Fisher eigenvalues should emerge from zero as coupling crosses critical thresholds.

Keywords: Fisher information metric; parameter identifiability; superadditivity; coupled dynamical systems; statistical manifolds; synchronization

1 Introduction

Information theory provides powerful bounds on communication: channel capacity limits transmission rates, mutual information bounds learning, and data processing inequalities

constrain inference [7]. Many commonly used bounds—PAC generalization, channel coding theorems, Cramér–Rao bounds—assume a fixed model class, channel law, or hypothesis space.

This assumption is usually invisible. When two agents exchange messages, we take for granted that the receiver’s model class is fixed: decoding maps incoming symbols to a pre-existing internal space. Complexity growth is then bounded by how much information the channel can carry.

But what happens when the communicating systems are themselves high-dimensional dynamical objects? When a protocell network coordinates, or when coupled oscillators synchronize, or when interacting agents learn together—the systems don’t merely exchange symbols. They *reshape each other’s dynamics*.

This paper identifies a regime where the fixed-manifold assumption need not hold, so that bounds derived for fixed model classes may underpredict complexity growth:

Core claim: When high-dimensional coherent systems couple, they exchange constraints rather than tokens. This can create new collective coordinates, increasing effective dimensionality beyond the sum of parts.

We call this *manifold expansion*: the accessible statistical manifold grows under coupling. This is not a violation of information theory—it is a regime where information theory’s foundational assumption (fixed model class) need not hold.

The implications are significant:

- Model complexity (identifiable dimensions) can grow faster than fixed-class bounds predict—because coupling activates new coordinates, not because information bounds are violated
- Communication becomes manifold deformation, not message passing
- The mechanism is scale-free: it applies wherever high-D systems interact

1.1 The Fixed-Manifold Assumption in Practice

To be precise about what we are *not* claiming: we do not argue that data-processing inequalities are violated, or that channel capacity theorems fail. These results hold within their stated assumptions.

What we identify is a regime where the *model class itself* changes under coupling, so that bounds derived for a fixed class can underpredict emergent complexity. The fixed-manifold assumption appears in several standard settings:

1. **Generalization bounds:** PAC-learning and VC-dimension arguments assume a fixed hypothesis class \mathcal{H} . Sample complexity bounds scale with $\dim(\mathcal{H})$ [7].
2. **Channel capacity:** Shannon’s theorem assumes a fixed channel law $p(y|x)$. Capacity is computed over this fixed conditional distribution.

3. **Data processing inequalities:** DPI holds for fixed processing maps. If the map itself changes under interaction, DPI applies to each fixed map but not to the trajectory across maps.
4. **Fisher information bounds:** Cramér-Rao and related bounds assume a fixed parametric family. If coupling changes which parameters are identifiable, the bounds shift.

Our claim: *coupling between high-dimensional systems can change the identifiable parameter set, so that applying bounds derived for a fixed family may underpredict the achievable complexity.*

To be explicit: we are tracking changes in the *image dimension* of the parameter-to-distribution map, not claiming violations of information inequalities. The inequalities hold; what changes is the dimension of the family to which they apply.

1.2 Coherence vs. Information: The Key Distinction

We distinguish two conceptual modes of inter-system interaction that motivate our geometric framework:

Information transfer: System A transmits discrete, addressable, copyable tokens to system B . The receiver decodes tokens into a pre-existing internal representation. Complexity growth is bounded by channel capacity. This is the regime where standard information-theoretic bounds apply.

Constraint exchange: System A couples to system B via shared dynamical modes. The coupling reshapes both systems' accessible state spaces. Complexity growth depends on geometric properties of the coupling, not channel capacity. This is the regime we study.

High-dimensional coherent systems (our focus) are those with: (i) many coupled internal degrees of freedom, (ii) long-lived collective modes or metastable states, and (iii) dynamics that maintain internal correlations over timescales longer than the coupling timescale. Examples include oscillator networks, neural populations, and coupled reaction systems near bifurcations.

The mathematical content of this paper concerns the constraint-exchange regime. We prove that coupling can increase the Fisher rank of an accessible family—a phenomenon invisible to channel-capacity arguments that assume a fixed model class.

1.3 Paper Outline

This paper provides:

1. **Section 2:** Background on information geometry and the fixed-manifold assumption
2. **Section 3:** Formal setup—accessible families, collective coordinates, and the main theorem on coupling-induced rank transitions (Theorem 5)
3. **Section 4:** Consequences for complexity growth, including the complexity acceleration conjecture (Conjecture 20)
4. **Section 5:** Discussion, testable predictions, and scope

2 Background: Information Geometry

Information geometry studies the differential geometry of statistical manifolds [1, 4, 5]. A statistical manifold \mathcal{M} is a space of probability distributions, equipped with the Fisher information metric:

$$g_{ij}(\theta) = \mathbb{E} \left[\frac{\partial \log p(x|\theta)}{\partial \theta_i} \frac{\partial \log p(x|\theta)}{\partial \theta_j} \right] \quad (1)$$

The Fisher metric induces a Riemannian structure on \mathcal{M} . A key feature of IG is the existence of *dual affine connections* ($\nabla^{(e)}$ and $\nabla^{(m)}$) corresponding to exponential and mixture geodesics. In dually flat manifolds (including exponential families), KL divergence is a canonical divergence whose second-order expansion recovers the Fisher metric, and information projections satisfy a Pythagorean theorem [1].

Communication and learning are naturally described as motion on this manifold:

- Parameter estimation moves along geodesics (e- or m-geodesics depending on the estimation criterion)
- KL divergence locally approximates squared Fisher–Rao distance to second order
- Natural gradient descent follows the manifold’s intrinsic geometry
- Information projections onto submanifolds minimize divergence

The hidden assumption: The manifold \mathcal{M} is fixed. The parameterization θ may change, but the model class—the set of distributions considered—does not.

This assumption is reasonable when:

- Systems have fixed, known structure
- Communication channels are well-defined
- Complexity is measured by localization on a fixed space

It can be inapplicable when:

- Interacting systems are high-dimensional and adaptive
- Coupling creates new collective modes
- The model class itself is shaped by interaction

2.1 Relation to Hierarchical Information Geometry

Amari’s hierarchical decomposition of probability distributions provides the natural IG framework for our analysis [2]. For a joint distribution $p(x_1, x_2)$, the log-linear decomposition separates marginal and interaction terms:

$$\log p(x_1, x_2) = \theta_1(x_1) + \theta_2(x_2) + \eta(x_1, x_2) + \psi \quad (2)$$

where θ_1, θ_2 are marginal parameters and η captures interactions. The *independence submanifold* \mathcal{M}_{ind} is the e-flat submanifold where $\eta = 0$.

Amari showed that interaction coordinates η are always *structurally present* in the ambient model—they parameterize directions transverse to independence. Our contribution is orthogonal: we study when these coordinates become *dynamically accessible*.

The distinction is crucial. In standard IG, one chooses a model class \mathcal{M} and studies its geometry. The interaction coordinates either belong to \mathcal{M} or they don't—this is a modeling choice. We consider a different situation: the dynamics determine which distributions are reachable, and coupling can change the reachable set. The “manifold expansion” we study is not a change in the chosen model, but a change in what the dynamics can produce.

This connects to recent work on sufficient statistics and statistical morphisms [3]. The map from physical parameters to observable distributions is a statistical transformation; our rank-change theorems characterize when this transformation becomes more “expressive” under coupling.

Relation to singular models. Our focus on Fisher rank degeneracy connects to singular learning theory, where the Fisher information matrix is degenerate at certain parameter values [6]. In that literature, singularities are typically fixed features of a chosen model class. Here, the novelty is that degeneracy can be *lifted by coupling*: parameter directions that are unidentifiable (in $\ker(dF_0)$) at zero coupling become identifiable (leave the kernel) at positive coupling. This “identifiability activation” is a dynamic phenomenon, not a property of a fixed model.

3 Manifold Expansion Under Coupling

3.1 Setup: Three Distinct Objects

Consider two dynamical systems with state spaces $X_1 \subset \mathbb{R}^{n_1}$ and $X_2 \subset \mathbb{R}^{n_2}$. We distinguish three objects that are often conflated:

1. **Ambient manifold** $\mathcal{P}(X_1 \times X_2)$: The space of all probability distributions over the joint state space. This is fixed.
2. **Chosen model class** $\mathcal{M} \subset \mathcal{P}$: A parametric family the analyst uses for inference (e.g., “all bivariate Gaussians”). Standard bounds assume this is fixed.
3. **Accessible family** $\mathcal{M}_{\text{acc}}(\kappa)$: The distributions that the *dynamics* can actually produce at coupling strength κ . This may change with κ .

The independence submanifold: A key example of a submodel is $\mathcal{M}_{\text{ind}} = \{p_1 \otimes p_2 : p_i \in \mathcal{M}_i\}$ —the product distributions with zero interaction. In IG terminology, this is an *e-flat* submanifold corresponding to zero interaction coordinates in the hierarchical log-linear decomposition [1].

The fixed-model assumption: Many learning and communication bounds assume the accessible family is contained in a fixed model class. We do *not* claim this assumption is wrong in general—only that it need not hold in the regime we study.

Formal setup: Let $\phi_\kappa : X_1 \times X_2 \rightarrow X_1 \times X_2$ denote coupled dynamics with coupling strength $\kappa \geq 0$, and let $h : X_1 \times X_2 \rightarrow \mathbb{R}^m$ be an observation map. Let \mathcal{B} be a smooth d -dimensional manifold of *physical parameters* β (e.g., damping rates, noise intensities, frequencies).

Key convention. The coupling strength κ is an *index* selecting a dynamical system, not a coordinate in \mathcal{B} . This prevents trivial rank increases from simply adding κ as a new parameter. Rank changes must arise from how coupling restructures the parameter→distribution map, not from expanding the parameter space.

For each fixed κ , we have a map:

$$F_\kappa : \mathcal{B} \rightarrow \mathcal{P}(\mathbb{R}^m), \quad \beta \mapsto Q_{(\kappa, \beta)} = (h)_\# P_{(\kappa, \beta)} \quad (3)$$

where $P_{(\kappa, \beta)}$ is the stationary distribution of ϕ_κ at parameters β , and $(h)_\#$ denotes the pushforward through h . The *accessible family* at coupling κ is the image:

$$\mathcal{M}_{\text{acc}}(\kappa) = \text{Image}(F_\kappa) \subset \mathcal{P}(\mathbb{R}^m)$$

Role of the observation map. The observation map $h : X_1 \times X_2 \rightarrow \mathbb{R}^m$ determines which features of the joint state are statistically accessible. When h is the identity (full state observation), identifiability depends only on dynamics. When h is a low-dimensional projection (partial observation), some coordinates may remain unidentifiable even if the dynamics distinguish them. Our rank-change theorems concern the map $F_\kappa = (h)_\# \circ P_{(\kappa, \cdot)}$; both dynamics and observation contribute to the accessible family. In the examples, we specify h explicitly and note when identifiability depends on observation richness.

Standing assumptions. Throughout, we assume:

- (A1) For each (κ, β) , the dynamics ϕ_κ admit a (locally) unique stationary distribution $P_{(\kappa, \beta)}$.
- (A2) The map $\beta \mapsto P_{(\kappa, \beta)}$ (and hence F_κ) is smooth.
- (A3) The observation map h is smooth with sufficient regularity that $Q = (h)_\# P$ has a smooth density on its support.

These assumptions exclude bifurcation points where stationary distributions split or vanish. Our results apply to parameter regions where (A1)–(A3) hold; analysis at bifurcation points requires additional care.

Symmetry-breaking and (A1). At symmetry-breaking transitions (as in the Kuramoto example), assumption (A1) can appear to fail: above threshold, there may be an orbit of stationary solutions rather than a unique one. We handle this in two equivalent ways: (i) treat orbit coordinates as parameters conditional on an external reference (making the distribution unique given the orbit coordinate), or (ii) introduce a small symmetry-breaking perturbation to restore uniqueness and take the limit. In the Kuramoto example (Section 3.6), we adopt approach (i): the orbit coordinate Ψ becomes part of \mathcal{B} , and $F_K(\Psi)$ is unique for each Ψ .

Our claim: For high-dimensional coherent systems, the rank of dF_κ can increase as κ crosses a threshold—the image manifold gains dimension.

Definition 1 (Manifold expansion). Define the generic rank of F_κ as:

$$r(\kappa) := \sup_{\beta \in \mathcal{B}} \text{rank } dF_\kappa(\beta) \quad (4)$$

By lower semicontinuity of rank (the sets $\{\beta : \text{rank } dF_\kappa(\beta) \geq r\}$ are open), the set where rank equals the supremum is open. We call $r(\kappa)$ the “generic rank” when this open set is nonempty; in applications we restrict to parameter regions where the supremum is attained. (For real-analytic maps on connected domains, the maximum-rank set is open and dense; for smooth maps, it is open but density requires additional assumptions.)

Manifold expansion occurs at κ_c if the generic rank increases as κ crosses κ_c :

$$r(\kappa_c^+) > r(\kappa_c^-) \quad (5)$$

Equivalently, the dimension of $\mathcal{M}_{\text{acc}}(\kappa)$ increases: a parameter direction in \mathcal{B} that was generically in the kernel of dF_κ becomes generically non-degenerate under coupling.

In Fisher-metric terms: manifold expansion corresponds to an increase in Fisher rank—the number of generically linearly independent score directions.

This definition is coordinate-invariant: rank is preserved under reparameterization. It captures “new identifiable directions appear” in an intrinsic way.

Remark 2 (Manifold structure of the accessible family). The accessible family $\mathcal{M}_{\text{acc}}(\kappa) = \text{Image}(F_\kappa)$ is the image of a smooth map. Locally around points where rank is constant, the constant rank theorem guarantees that the image is an immersed submanifold of dimension $r(\kappa)$. At points where rank changes, the image may have stratified structure. Rank transitions as in Theorem 5 correspond to changes in local manifold dimension (the top stratum gains dimension).

The novelty is not that interaction parameters exist in the full parametric family—Amari’s hierarchical decomposition already accounts for this [1]. The novelty is that **dynamics and observation determine which interaction coordinates are identifiable**. Coupling can activate coordinates that were structurally present but dynamically inaccessible—a phenomenon we term **coupling-induced identifiability activation**.

3.2 Mechanism: Collective Coordinates

Let $\phi : X_1 \times X_2 \rightarrow X_1 \times X_2$ be the coupled dynamics. Under coupling, new *collective coordinates* can become accessible:

Definition 3 (Collective coordinate). A collective coordinate is a function $\psi : X_1 \times X_2 \rightarrow \mathbb{R}$ that is:

1. Not reducible to functions of X_1 or X_2 alone
2. Dynamically stable: slow manifold, metastable basin, or conserved quantity
3. Statistically identifiable: non-degenerate Fisher information along ψ under observation h

Remark 4. *Stability is a dynamical property; identifiability is a statistical property. Coupling can create one without the other. A collective mode may be dynamically stable but statistically invisible (unobservable through h), or statistically identifiable but dynamically unstable (transient). The interesting regime is when both hold.*

Examples:

- Phase difference between coupled oscillators [12, 14]
- Synchronization manifold coordinates [13]
- Order parameters of collective states (e.g., chimera states) [15]
- Interface modes at boundaries between systems

These coordinates represent genuinely new degrees of freedom: they were not accessible to either system in isolation.

3.3 Main Theorem: Coupling-Induced Rank Transitions

We now state and prove our main result, which provides checkable criteria for when coupling induces Fisher-rank transitions.

Structure and novelty. The theorem has three parts. Part I recalls the standard pullback characterization of Fisher information (see, e.g., Amari [1], Ch. 2); we include it for completeness and to fix notation. **The novel content is Parts II and III**, which provide checkable criteria—in terms of kernel activation—for when coupling *increases* the Fisher rank. These criteria, together with the superadditivity formalization (Section 3.4) and the “accessible family” dynamical framing, constitute the paper’s contribution.

Manifold setting. Throughout, we assume \mathcal{P} is a finite-dimensional statistical manifold—a smooth manifold of probability distributions equipped with the Fisher information metric g . Concretely, \mathcal{P} may be an exponential family, a mixture family, or any smooth finite-dimensional submanifold of the space of densities on observation space Y . The Fisher metric is positive-definite on tangent spaces to \mathcal{P} . (For infinite-dimensional \mathcal{P} , one can use the weak Fisher–Rao metric on smooth positive densities; our results extend with appropriate regularity assumptions, but we state them in the finite-dimensional setting for clarity.)

Theorem 5 (Coupling-induced rank transitions). *Let \mathcal{B} be a smooth d -dimensional parameter manifold, let \mathcal{P} be a finite-dimensional statistical manifold on observation space Y equipped with the Fisher information metric g , and let $F_\kappa : \mathcal{B} \rightarrow \mathcal{P}$ be a smooth family of maps indexed by coupling strength $\kappa \geq 0$, where $F_\kappa(\beta) = Q_{(\kappa, \beta)}$ is the distribution of observations under dynamics with parameters (κ, β) .*

Part I (Pullback identity—standard). *The Fisher information matrix on \mathcal{B} induced by the accessible family $\mathcal{M}_{\text{acc}}(\kappa) = \text{Image}(F_\kappa)$ is the pullback of the Fisher metric:*

$$I_{\mathcal{B}}(\kappa) = F_\kappa^* g \tag{6}$$

In coordinates, if $\beta = (\beta^1, \dots, \beta^d)$ and $Q_{(\kappa, \beta)}$ has density $q(y; \kappa, \beta)$, then:

$$(I_B(\kappa))_{ij} = \mathbb{E}_{Q_{(\kappa, \beta)}} \left[\frac{\partial \log q}{\partial \beta^i} \frac{\partial \log q}{\partial \beta^j} \right] \quad (7)$$

Consequently:

$$\text{rank } I_B(\kappa) = \text{rank } dF_\kappa \quad (8)$$

Part II (Constraint-release criterion). Let $\mathcal{M}_0 \subset \mathcal{P}$ be a smooth submanifold of codimension c (e.g., the independence submanifold). Choose local coordinates on \mathcal{P} near \mathcal{M}_0 so that $\mathcal{M}_0 = \{\eta = 0\}$ locally, and write $F_\kappa(\beta) = (f_\kappa(\beta), \eta_\kappa(\beta))$ where f are coordinates along \mathcal{M}_0 and η are transverse coordinates. Let $\beta_0 \in \mathcal{B}$ achieve the maximal rank $r(0) = \text{rank } dF_0(\beta_0)$. Suppose:

- (C1) $F_0(\mathcal{B}) \subseteq \mathcal{M}_0$, i.e., $\eta_0(\beta) \equiv 0$ (dynamics constrain to submanifold at $\kappa = 0$)
- (C2) For $\kappa > \kappa_c$, there exists β such that $\eta_\kappa(\beta) \neq 0$ (constraint released)
- (C3) **Kernel activation:** There exists $v \in \ker(dF_0(\beta_0))$ such that $d\eta_\kappa(v) \neq 0$ for $\kappa > \kappa_c$

Then for $\kappa > \kappa_c$:

$$r(\kappa) \geq r(0) + 1 \quad (9)$$

where $r(\kappa)$ is the maximal rank of F_κ . More generally, if there exist k linearly independent vectors $v_1, \dots, v_k \in \ker(dF_0(\beta_0))$ such that $d\eta_\kappa(v_1), \dots, d\eta_\kappa(v_k)$ are linearly independent for $\kappa > \kappa_c$, then $r(\kappa) \geq r(0) + k$.

Remark 6 (Mechanism of rank increase). The key condition is (C3): transverse variation must occur along parameter directions that were previously unidentifiable (in $\ker(dF_0)$). Without this, coupling could merely “tilt” the existing $r(0)$ -dimensional image out of $T\mathcal{M}_0$ without adding new independent directions. Condition (C3) guarantees that previously-kernel directions become transverse-accessible, yielding a genuine rank increase.

Part III (Symmetry-breaking criterion). Let G be a k -dimensional Lie group acting smoothly on \mathcal{P} , and let $\text{Fix}(G) = \{p \in \mathcal{P} : \gamma \cdot p = p \text{ for all } \gamma \in G\}$ be the fixed-point set. Let $\beta_0 \in \mathcal{B}$ achieve the maximal rank $r(0) = \text{rank } dF_0(\beta_0)$. Suppose:

- (S1) $F_0(\mathcal{B}) \subseteq \text{Fix}(G)$ (distributions are G -invariant at $\kappa = 0$)
- (S2) For $\kappa > \kappa_c$, there exists β such that $F_\kappa(\beta) \notin \text{Fix}(G)$ (symmetry broken)
- (S3) The orbit $G \cdot F_\kappa(\beta)$ has dimension k (free action locally)
- (S4) **Kernel-to-orbit activation:** There exist k linearly independent vectors $v_1, \dots, v_k \in \ker(dF_0(\beta_0))$ such that $dF_\kappa(v_1), \dots, dF_\kappa(v_k)$ project to linearly independent vectors in $T_p(G \cdot p)$ for $\kappa > \kappa_c$

Then for $\kappa > \kappa_c$:

$$r(\kappa) \geq r(0) + k \quad (10)$$

where $r(\kappa)$ is the maximal rank of F_κ .

Remark 7 (Mechanism of symmetry-breaking rank increase). *Condition (S4) is the key: orbit directions must be accessed by parameter directions that were previously in the kernel of dF_0 . This prevents “trading” old identifiable directions for orbit directions (which would leave rank unchanged). The condition guarantees that symmetry breaking activates previously-unidentifiable parameters, adding k new independent directions to the image.*

Proof. **Part I.** Let $\beta(t)$ be a smooth curve in \mathcal{B} with $\beta(0) = \beta_0$ and $\dot{\beta}(0) = v \in T_{\beta_0}\mathcal{B}$. The image curve $Q(t) = F_\kappa(\beta(t))$ has tangent vector $dF_\kappa(v) \in T_{Q_0}\mathcal{P}$.

The Fisher metric on \mathcal{P} evaluates tangent vectors via score functions. If $Q(t)$ has density $q(y; t)$, then the tangent vector corresponds to the score $s(y) = \frac{d}{dt}|_{t=0} \log q(y; t)$, and the Fisher metric is $g(s, s) = \mathbb{E}_Q[s^2]$.

By the chain rule:

$$\frac{\partial \log q(y; \kappa, \beta)}{\partial \beta^i} = \sum_{\alpha} \frac{\partial \log q}{\partial \theta^\alpha} \frac{\partial \theta^\alpha}{\partial \beta^i} \quad (11)$$

where θ are coordinates on \mathcal{P} . The pullback metric is:

$$(F_\kappa^* g)_{ij} = g \left(\frac{\partial F_\kappa}{\partial \beta^i}, \frac{\partial F_\kappa}{\partial \beta^j} \right) \quad (12)$$

$$= \mathbb{E}_{Q_{(\kappa, \beta)}} \left[\frac{\partial \log q}{\partial \beta^i} \frac{\partial \log q}{\partial \beta^j} \right] \quad (13)$$

$$= (I_{\mathcal{B}}(\kappa))_{ij} \quad (14)$$

This establishes the pullback identity.

For the rank equality: the pullback metric $F_\kappa^* g$ is degenerate precisely along the kernel of dF_κ . A vector $v \in T_{\beta}\mathcal{B}$ satisfies $(F_\kappa^* g)(v, v) = 0$ if and only if $dF_\kappa(v) = 0$ (since g is positive-definite on $T_Q\mathcal{P}$). Thus:

$$\ker(I_{\mathcal{B}}(\kappa)) = \ker(dF_\kappa) \quad (15)$$

and $\text{rank}(I_{\mathcal{B}}(\kappa)) = d - \dim \ker(dF_\kappa) = \text{rank } dF_\kappa$.

Part II. In local coordinates with $\mathcal{M}_0 = \{\eta = 0\}$, write $F_\kappa(\beta) = (f_\kappa(\beta), \eta_\kappa(\beta))$. At $\kappa = 0$, condition (C1) gives $\eta_0 \equiv 0$, so the image of dF_0 lies entirely in the f -directions:

$$\text{Image}(dF_0(\beta_0)) = \text{Image}(df_0(\beta_0)) \oplus \{0\}$$

with $\dim(\text{Image}(dF_0(\beta_0))) = r(0)$.

Let $v \in \ker(df_0(\beta_0))$ be as in (C3). At $\kappa = 0$, $dF_0(v) = (df_0(v), d\eta_0(v)) = (0, 0)$, so $v \in \ker(dF_0(\beta_0))$. By (C3), for $\kappa > \kappa_c$, $d\eta_\kappa(v) \neq 0$, so $dF_\kappa(v) = (df_\kappa(v), d\eta_\kappa(v))$ has a nonzero transverse component.

The key observation is that $T_p\mathcal{P} \cong T_p\mathcal{M}_0 \oplus N_p\mathcal{M}_0$ (tangent and normal components are complementary subspaces). At $\kappa = 0$, the image of dF_0 lies entirely in $T\mathcal{M}_0$ (the f -directions). The vector $dF_\kappa(v)$ has a nonzero component in $N\mathcal{M}_0$ (the η -direction), so it cannot lie in the span of the old $r(0)$ tangential directions.

Persistence of old directions: Since $(dF_\kappa)_{\kappa \geq 0}$ is a smooth family of linear maps (by smoothness of F_κ in κ and β), the $r(0)$ linearly independent vectors in $\text{Image}(dF_0(\beta_0))$

remain linearly independent for κ in a neighborhood of 0. More precisely, if $w_1, \dots, w_{r(0)}$ span the image at $\kappa = 0$, then $\det(G)$ of the Gram matrix $G_{ij} = \langle dF_\kappa(e_i), dF_\kappa(e_j) \rangle$ is nonzero at $\kappa = 0$, hence nonzero in a neighborhood by continuity. Combined with the new transverse direction, we have:

$$r(\kappa) \geq r(0) + 1 \quad \text{for } \kappa \text{ in a neighborhood of } \kappa_c^+$$

Part III. At $\kappa = 0$, (S1) implies $F_0(\mathcal{B}) \subseteq \text{Fix}(G)$, so $\text{Image}(dF_0) \subseteq T(\text{Fix}(G))$ (no orbit directions). At β_0 achieving maximal rank, $\text{rank}(dF_0(\beta_0)) = r(0)$.

Let $v_1, \dots, v_k \in \ker(dF_0(\beta_0))$ be as in (S4). At $\kappa = 0$, $dF_0(v_i) = 0$ for all i . For $\kappa > \kappa_c$, condition (S4) states that $dF_\kappa(v_1), \dots, dF_\kappa(v_k)$ project to k linearly independent vectors in the orbit tangent space $T_p(G \cdot p)$.

The key observation is that at points $p \notin \text{Fix}(G)$, the tangent space decomposes as $T_p\mathcal{P} \cong T_p(G \cdot p) \oplus (T_p(G \cdot p))^\perp$, where orbit and non-orbit directions are complementary. At $\kappa = 0$, the image of dF_0 lies entirely in $T(\text{Fix}(G))$, which contains no orbit directions. For $\kappa > \kappa_c$, the vectors $dF_\kappa(v_i)$ have nonzero projections onto $T_p(G \cdot p)$, so they cannot lie in the span of the old $r(0)$ directions (which had no orbit component). As in Part II, the old $r(0)$ directions persist by continuity of the Gram determinant. Therefore:

$$r(\kappa) \geq r(0) + k \quad \text{for } \kappa \text{ in a neighborhood of } \kappa_c^+$$

□

Remark 8 (Effective dimensionality). *For systems where Fisher rank may be ambiguous (e.g., eigenvalues near zero), one can use the participation ratio as a soft measure of effective dimensionality [9, 11]:*

$$d_{\text{eff}} = \frac{(\sum_i \lambda_i)^2}{\sum_i \lambda_i^2} \tag{16}$$

where λ_i are Fisher eigenvalues. Fisher eigenvalues are coordinate-dependent: under reparameterization $\beta \mapsto \tilde{\beta}(\beta)$, the FIM transforms by congruence $\tilde{I} = J^T I J$, changing eigenvalues. Participation ratio is therefore a heuristic for fixed parameterizations. In contrast, the rank of the pullback metric $F_\kappa^* g$ equals the rank of the differential dF_κ , which is coordinate-invariant.

Remark 9 (Connection to effective nonergodicity). *High-dimensional systems are often effectively nonergodic over finite timescales: measure concentration and metastability can confine trajectories to small regions of phase space for times much longer than observation windows. In such regimes, Theorem 5 describes what happens when two effectively nonergodic systems couple: their confinement to distinct phase-space regions may be partially released, activating interaction coordinates that were structurally present but dynamically inaccessible under the uncoupled dynamics.*

3.4 Superadditive Accessible Complexity

The preceding theorem establishes conditions for rank increase under coupling. We now formalize “the whole exceeds the sum of its parts” as a precise inequality: the coupled accessible family can have higher generic rank than the independent product family.

Definition 10 (Component accessible families). *For each subsystem $i = 1, \dots, n$, let \mathcal{B}_i be its parameter manifold and let $F_i : \mathcal{B}_i \rightarrow \mathcal{P}_i$ be its dynamics-to-distribution map. Define the component accessible family $\mathcal{M}_i := \text{Image}(F_i)$ and its generic rank*

$$r_i := \max_{\beta_i \in \mathcal{B}_i} \text{rank } dF_i(\beta_i).$$

Definition 11 (Independent composite and additive baseline). *Let $\mathcal{B} := \prod_{i=1}^n \mathcal{B}_i$ and $\mathcal{P} := \prod_{i=1}^n \mathcal{P}_i$. Define the independent composite map*

$$F_{\oplus} : \mathcal{B} \rightarrow \mathcal{P}, \quad F_{\oplus}(\beta_1, \dots, \beta_n) := (F_1(\beta_1), \dots, F_n(\beta_n)).$$

Let $\mathcal{M}_{\oplus} := \text{Image}(F_{\oplus}) = \prod_i \mathcal{M}_i$ be the independent composite accessible family, and define the additive baseline rank

$$r_{\oplus} := \max_{\beta \in \mathcal{B}} \text{rank } dF_{\oplus}(\beta).$$

Lemma 12 (Additivity under independence). *For the independent composite map F_{\oplus} ,*

$$r_{\oplus} = \sum_{i=1}^n r_i.$$

Proof. The differential dF_{\oplus} is block-diagonal with blocks dF_i on $T_{\beta_i} \mathcal{B}_i$. Hence $\text{rank } dF_{\oplus}(\beta) = \sum_i \text{rank } dF_i(\beta_i)$. Maximizing over β yields $r_{\oplus} = \sum_i r_i$. \square

This lemma provides the “sum of parts” baseline against which we measure superadditivity.

Definition 13 (Superadditive accessible complexity). *Let $F_{\kappa} : \mathcal{B} \rightarrow \mathcal{P}$ be the coupled dynamics-to-distribution map with generic rank $r(\kappa) := \max_{\beta \in \mathcal{B}} \text{rank } dF_{\kappa}(\beta)$. The coupled system exhibits superadditive accessible complexity at coupling κ if*

$$r(\kappa) > r_{\oplus} = \sum_i r_i.$$

Proposition 14 (Sufficient condition via non-factorizable coordinates). *Suppose there exists a smooth map $\pi : \mathcal{P} \rightarrow \mathbb{R}^k$ (a “collective observable coordinate”) such that:*

- (A1) $\pi \circ F_{\oplus}$ is locally constant on some open $U \subset \mathcal{B}$ (the independent product family cannot vary these coordinates);
- (A2) $\pi \circ F_{\kappa}$ has Jacobian rank k on U (the coupled family varies freely in these coordinates).

Then $r(\kappa) \geq r_{\oplus} + k$, so the coupled system is superadditive whenever $k \geq 1$.

Proof. By (A1), the image of dF_{\oplus} lies in $\ker(d\pi)$ on U , so dF_{\oplus} cannot span the k directions detected by π . By (A2), $d(\pi \circ F_{\kappa}) = d\pi \circ dF_{\kappa}$ has rank k on U , so dF_{κ} spans at least k independent directions beyond those available under F_{\oplus} . \square

We now connect superadditivity to the mechanisms identified in Theorem 5.

Corollary 15 (Superadditivity from constraint release). *Let $\mathcal{M}_0 := \mathcal{M}_{\oplus} = \prod_i \mathcal{M}_i$ be the independent composite accessible family. If the coupled family satisfies the constraint-release hypotheses of Theorem 5 Part II with k transverse directions released relative to \mathcal{M}_0 , then for $\kappa > \kappa_c$:*

$$r(\kappa) \geq r_{\oplus} + k.$$

Corollary 16 (Superadditivity from symmetry breaking). *Let $\mathcal{M}_0 := \mathcal{M}_{\oplus}$ and suppose at $\kappa = 0$ the accessible family is G -invariant ($\mathcal{M}_{\text{acc}}(0) \subseteq \text{Fix}(G)$), while for $\kappa > \kappa_c$ symmetry is broken with k -dimensional orbits. If the orbit accessibility condition (S4) holds, then for $\kappa > \kappa_c$:*

$$r(\kappa) \geq r_{\oplus} + k.$$

Remark 17 (Interpretation). *These corollaries formalize “the whole exceeds the sum of its parts” as a rank inequality: coupling activates collective coordinates not attributable to any subsystem alone. This is not a violation of information-theoretic bounds—it is a statement about model geometry. The accessible statistical family of the coupled system has more identifiable directions than the product of component families.*

3.5 Example 1: Coupled Ornstein–Uhlenbeck (Transversality)

This example demonstrates Theorem 5 Part II (transversality criterion) with $\kappa_c = 0$.

Dynamics. Consider two coupled Ornstein–Uhlenbeck processes:

$$dX_1 = -\gamma_1 X_1 dt + \kappa(X_2 - X_1) dt + \sigma_1 dW_1 \quad (17)$$

$$dX_2 = -\gamma_2 X_2 dt + \kappa(X_1 - X_2) dt + \sigma_2 dW_2 \quad (18)$$

where $\gamma_i > 0$ are mean-reversion rates, $\sigma_i > 0$ are noise intensities, and $\kappa \geq 0$ is the coupling strength.

Stationary distribution. The stationary distribution is bivariate Gaussian with zero mean and covariance matrix Σ satisfying the Lyapunov equation $A\Sigma + \Sigma A^T + D = 0$, where the drift matrix is

$$A = \begin{pmatrix} -(\gamma_1 + \kappa) & \kappa \\ \kappa & -(\gamma_2 + \kappa) \end{pmatrix}$$

and $D = \text{diag}(\sigma_1^2, \sigma_2^2)$. Solving the Lyapunov equation (a 3×3 linear system in $\Sigma_{11}, \Sigma_{22}, \Sigma_{12}$) yields:

$$\begin{aligned} \Sigma_{11} &= \frac{\sigma_1^2(\gamma_2 + \kappa) + \kappa^2\sigma_2^2/(\gamma_2 + \kappa)}{2(\gamma_1 + \kappa)(\gamma_2 + \kappa) - 2\kappa^2} \cdot (\gamma_2 + \kappa) \\ &\approx \frac{\sigma_1^2}{2\gamma_1} + O(\kappa) \quad \text{as } \kappa \rightarrow 0 \end{aligned} \quad (19)$$

For the symmetric case $\gamma_1 = \gamma_2 = \gamma$ and $\sigma_1 = \sigma_2 = \sigma$, the closed-form solution simplifies to:

$$\Sigma_{11} = \Sigma_{22} = \frac{\sigma^2}{2\gamma}, \quad \Sigma_{12} = \frac{\kappa\sigma^2}{2\gamma(\gamma + 2\kappa)} \quad (20)$$

giving correlation

$$\rho = \frac{\Sigma_{12}}{\Sigma_{11}} = \frac{\kappa}{\gamma + 2\kappa} \quad (21)$$

This confirms $\rho(\kappa = 0) = 0$ and shows the rank transition explicitly. Computing $\partial\rho/\partial\gamma$:

$$\frac{\partial\rho}{\partial\gamma} = -\frac{\kappa}{(\gamma + 2\kappa)^2} \neq 0 \quad \text{for } \kappa > 0 \quad (22)$$

This nonzero derivative demonstrates that changing γ (a physical parameter in \mathcal{B}) changes the correlation ρ (the transverse coordinate η), confirming kernel activation: a direction that left ρ unchanged at $\kappa = 0$ now varies ρ at $\kappa > 0$.

The dynamics-to-distribution map. Following our convention that κ is an *index* (not a parameter in \mathcal{B}), we define the physical parameter space as $\mathcal{B} = \{(\gamma_1, \sigma_1, \gamma_2, \sigma_2) : \gamma_i, \sigma_i > 0\}$. For each fixed $\kappa \geq 0$:

$$F_\kappa : \mathcal{B} \rightarrow \mathcal{P}, \quad (\gamma_1, \sigma_1, \gamma_2, \sigma_2) \mapsto \mathcal{N}(0, \Sigma(\kappa, \beta)) \quad (23)$$

The ambient manifold \mathcal{P} of centered bivariate Gaussians is 3-dimensional, parameterized by (v_1, v_2, ρ) where $v_i > 0$ are marginal variances and $\rho \in (-1, 1)$ is the correlation. The independence submanifold is $\mathcal{M}_0 = \{(v_1, v_2, 0) : v_1, v_2 > 0\}$, which has codimension $c = 1$.

Jacobian structure. At $\kappa = 0$: The map $\beta \mapsto (v_1, v_2, \rho)$ satisfies $\rho = 0$ identically, so $F_0(\mathcal{B}) \subseteq \mathcal{M}_0$. The Jacobian of F_0 is:

$$dF_0 = \begin{pmatrix} \partial v_1 / \partial \gamma_1 & \partial v_1 / \partial \sigma_1 & 0 & 0 \\ 0 & 0 & \partial v_2 / \partial \gamma_2 & \partial v_2 / \partial \sigma_2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (24)$$

The first two rows have rank 2 (generically), while the third row is zero. Hence $r(0) = 2$.

At $\kappa > 0$: The correlation ρ depends on the parameters, so the third row becomes nonzero. Specifically, $\partial\rho/\partial\gamma_1 \neq 0$ generically, so the Jacobian has rank 3.

Verification of Theorem 5 Part II.

Using the notation of the theorem: the submanifold coordinates are $f = (v_1, v_2)$ and the transverse coordinate is $\eta = \rho$.

- (C1): At $\kappa = 0$, $\rho = 0$ for all β , so $F_0(\mathcal{B}) \subseteq \mathcal{M}_0$. ✓
- (C2): For $\kappa > 0$, $\rho > 0$, so $F_\kappa(\mathcal{B})$ escapes \mathcal{M}_0 . ✓
- (C3) **Kernel activation:** The Jacobian df_0 has rank 2, so $\ker(df_0)$ is 2-dimensional. In particular, there exist directions v in parameter space that change neither v_1 nor v_2 at $\kappa = 0$. For $\kappa > 0$, some of these directions activate the transverse coordinate: $d\eta_\kappa(v) \neq 0$. ✓

Conclusion. The rank at $\kappa = 0$ is $r(0) = 2$. The theorem predicts $r(\kappa) \geq r(0) + 1 = 3$ for $\kappa > 0$. Direct computation confirms $\text{rank}(dF_\kappa) = 3$ for $\kappa > 0$, verifying the bound.

Superadditivity analysis. Each OU process in isolation has a 1-dimensional accessible family (variance only), so $r_1 = r_2 = 1$ and $r_\oplus = 2$. The coupled system achieves $r(\kappa) = 3$ for $\kappa > 0$. Thus:

$$r(\kappa) = 3 > r_\oplus = 2$$

This is *strict superadditivity*: the coupled system has one more identifiable direction (the correlation ρ) than the product of independent systems. The correlation is a *collective coordinate* that is not attributable to either subsystem alone.

Observation map. The analysis assumes observation of (X_1, X_2) jointly. If only X_1 were observed, ρ would remain unidentifiable even for $\kappa > 0$ —the correlation affects the joint distribution but not the X_1 marginal. Identifiability requires sufficient observation richness.

3.6 Example 2: Kuramoto Synchronization (Symmetry Breaking)

This example demonstrates Theorem 5 Part III (symmetry-breaking criterion) with $\kappa_c > 0$. We work in the **mean-field limit** ($N \rightarrow \infty$) with added noise, where the stationary density is unique up to symmetry [12, 14].

Mean-field dynamics. Consider the noisy Kuramoto model in the thermodynamic limit. Let $\rho(\theta, \omega, t)$ be the density of oscillators with natural frequency ω at phase θ . With phase diffusion of strength $D > 0$, the density evolves according to the nonlinear Fokker–Planck equation:

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial \theta} [\rho (\omega + Kr \sin(\Psi - \theta))] = D \frac{\partial^2 \rho}{\partial \theta^2} \quad (25)$$

where the order parameter $r e^{i\Psi} = \int e^{i\theta} \rho(\theta, \omega, t) \nu(\omega) d\omega d\theta$ is determined self-consistently, with $\nu(\omega)$ the frequency distribution. The system has S^1 symmetry: if $\rho(\theta)$ is a stationary solution, so is $\rho(\theta - \phi)$ for any ϕ .

Stationary solutions. For $K < K_c$ (where K_c depends on D and ν), the stable stationary solution is the incoherent S^1 -invariant density with $r = 0$. For $K > K_c$, there exists an S^1 -orbit of synchronized stationary solutions with $r > 0$, related by rotation: $\rho_\Psi(\theta) = \rho_0(\theta - \Psi)$ for $\Psi \in S^1$. The diffusion ensures uniqueness of the stationary density within each symmetry class.

The accessible family. Following our convention that coupling K is an *index*, define the parameter space $\mathcal{B} = S^1$, where $\Psi \in S^1$ selects a point on the orbit of stationary solutions. The dynamics-to-distribution map is:

$$F_K : \mathcal{B} \rightarrow \mathcal{P}, \quad \Psi \mapsto \rho_{K,\Psi} \quad (26)$$

where $\rho_{K,\Psi}$ is the stationary solution with mean phase Ψ (or the unique incoherent distribution, independent of Ψ , if $K < K_c$).

Why $\mathcal{B} = S^1$ here. In this example, \mathcal{B} does not represent microscopic physical parameters (noise strength, frequency distribution, etc.) but rather the *emergent orbit coordinate* that becomes identifiable once symmetry breaks. This is precisely what Theorem 5 Part III addresses: a coordinate that parameterizes the symmetry orbit, which is unidentifiable below K_c (maps to a single point) and identifiable above K_c (maps onto the orbit). Readers expecting \mathcal{B} to contain physical knobs should note that one could alternatively let \mathcal{B} include such parameters; the symmetry-breaking criterion would then show that breaking S^1 adds one dimension (the orbit direction) to the accessible family beyond whatever dimension the physical parameters contribute.

The group action. Let $G = S^1$ act on \mathcal{P} by rotation: $(e^{i\phi} \cdot \rho)(\theta) = \rho(\theta - \phi)$. The fixed-point set $\text{Fix}(S^1)$ consists of rotationally invariant densities.

Verification of Theorem 5 Part III.

- (S1): For $K < K_c$, the image $F_K(\mathcal{B})$ is the single incoherent distribution, which is S^1 -invariant. ✓
- (S2): For $K > K_c$, the synchronized stationary solutions break S^1 symmetry ($r > 0$). ✓
- (S3): The S^1 -orbit of synchronized solutions is 1-dimensional. ✓
- (S4): For $K < K_c$, the map $F_K : \Psi \mapsto \rho_{K,\Psi}$ is constant (all Ψ give the same incoherent distribution), so $dF_K = 0$. For $K > K_c$, $\partial F_K / \partial \Psi \neq 0$ and maps onto the orbit direction (kernel-to-orbit activation). ✓
- $\dim(S^1) = 1$, so the theorem predicts: $r(K) \geq r(0) + 1$ for $K > K_c$. Since $r(0) = 0$ (constant map), we get $r(K) \geq 1$. ✓

The new identifiable coordinate. Below K_c , the parameter Ψ is unidentifiable—all values give the same uniform distribution. Above K_c , Ψ becomes identifiable: different mean phases produce distinguishable synchronized distributions. The Fisher information in the Ψ direction transitions from zero to positive at $K = K_c$.

Observation map. Identifiability of Ψ requires an observation that breaks the S^1 symmetry. If the observation is S^1 -invariant (e.g., reporting only the coherence r), then Ψ remains unidentifiable even above K_c . We assume observation relative to an external phase reference.

Remark 18 (Pinned-phase alternative formulation). *Readers concerned that “ $\Psi \in \mathcal{B}$ ” amounts to “adding a parameter” may prefer the following equivalent formulation: introduce a small symmetry-breaking field $\varepsilon \sin(\theta - \theta_{\text{ref}})$ in the dynamics, so that the stationary distribution is unique for each K . Let \mathcal{B} contain only physical parameters (noise, frequency distribution). For $K < K_c$, the unique stationary distribution is nearly uniform and θ_{ref} has negligible effect; for $K > K_c$, the distribution concentrates near θ_{ref} . Taking $\varepsilon \rightarrow 0$, the rank transition persists: the accessible family dimension increases from 0 (below K_c) to 1 (above K_c) due to the emergent collective coordinate. This confirms that the rank increase is a property of the coupling transition, not an artifact of parameterization choice.*

Key difference from OU. In the OU example (Part II), any $\kappa > 0$ immediately releases the constraint: $\kappa_c = 0$. In Kuramoto (Part III), the symmetry persists for $0 < K < K_c$; manifold expansion occurs at a *genuine* critical threshold $K_c > 0$, driven by symmetry breaking in the mean-field limit.

Superadditivity analysis. This example demonstrates *strict* superadditivity. Consider the independent oscillators: the phase of each is defined only modulo S^1 symmetry, so no individual phase is identifiable. The independent composite r_{\oplus} includes no mean-phase direction.

For $K > K_c$, the coupled system gains an identifiable global phase Ψ . This coordinate is *not attributable to any subsystem*—it emerges only at the collective level. By Corollary 16:

$$r(K) \geq r_{\oplus} + 1 \quad \text{for } K > K_c$$

The whole exceeds the sum of its parts: the synchronized oscillator ensemble has an identifiable degree of freedom that no individual oscillator possesses.

3.7 When Expansion Dominates

Manifold expansion is most pronounced when:

1. **High internal dimensionality:** Each system has many coupled degrees of freedom
2. **Coherent dynamics:** Internal correlations are strong (non-ergodic, structured)
3. **Weak-to-intermediate coupling:** Strong enough to create collective modes, weak enough to avoid synchronization collapse
4. **Heterogeneity:** Systems are different enough that coupling is non-trivial

In the opposite regime (low-D, incoherent, very weak or very strong coupling), the standard product assumption holds.

4 Consequences for Complexity Growth

4.1 Faster Than Fixed-Class Bounds Predict

Standard bounds on complexity growth (learning rate, adaptation speed, structure formation rate) assume:

- Fixed parameterization
- Fixed sufficient statistics
- Fixed dimensionality

When coupling creates new coordinates:

- The Fisher geometry changes because the model class changes—new significant Fisher directions appear
- KL-based learning rates underpredict adaptation
- Channel capacity arguments don't apply (the “channel” itself is being restructured)

Definition 19 (Model complexity). *We define complexity C of a statistical family \mathcal{M} as the Fisher rank: the dimension of the identifiable tangent space (equivalently, the number of nonzero Fisher eigenvalues). This is coordinate-invariant and intrinsic to the manifold geometry. For “soft” complexity measures, one can use participation ratio in a fixed parameterization; this is related to stochastic complexity in the MDL framework [8].*

Conjecture 20 (Eigenvalue emergence under coupling). *In regimes where the accessible family $\mathcal{M}_{\text{acc}}(\kappa)$ can be embedded in a fixed parameterization across κ (e.g., via a common ambient family like “all bivariate Gaussians”), eigenvalue emergence provides a concrete signature. As coupling strength κ crosses a critical threshold κ_c , one or more Fisher eigenvalues emerge*

from zero (or cross a detectability threshold $\epsilon > 0$). Specifically, if $\lambda_1(\kappa) \leq \lambda_2(\kappa) \leq \dots$ are the ordered eigenvalues of $I(\kappa)$ in this fixed parameterization, then for some k :

$$\lambda_k(\kappa_c^-) < \epsilon \quad \text{and} \quad \lambda_k(\kappa_c^+) > \epsilon \quad (27)$$

This implies rank increase (or effective dimensionality increase) as coupling activates previously degenerate interaction coordinates. In contrast, for a fixed statistical family, the eigenvalue spectrum is fixed.

Remark 21 (Explicit mismatch: PAC bounds). Standard PAC generalization bounds take the form

$$L(h) \leq \hat{L}(h) + O\left(\sqrt{\frac{d}{n}}\right) \quad (28)$$

where d is the VC dimension (or effective dimension) of the hypothesis class \mathcal{H} , and n is sample size. The key premise: d is fixed by the chosen hypothesis class.

Disclaimer: We do not claim that Fisher rank equals VC dimension; we use PAC bounds as a representative instance of bounds that assume a fixed-dimension model class. The analogy is structural: both Fisher rank and VC dimension count “effective degrees of freedom,” and both can change if the model class changes.

Under coupling-induced manifold expansion, the identifiable family $\mathcal{M}_{\text{acc}}(\kappa)$ grows with κ . If one treats the effective dimension as constant ($d = d_0$), the bound underestimates true complexity. The correct bound requires $d_{\text{eff}}(\kappa)$:

$$L(h) \leq \hat{L}(h) + O\left(\sqrt{\frac{d_{\text{eff}}(\kappa)}{n}}\right) \quad (29)$$

If $d_{\text{eff}}(\kappa) > d_0$ due to coupling, systems can appear to “learn faster than the bound allows”—they are exploiting newly accessible coordinates that weren’t counted in the fixed- d analysis.

Remark 22 (Mutual information vs. model dimension). A potential confusion: mutual information $I(X_1; X_2)$ measures dependence between variables, and one might expect it to bound “complexity growth.” But MI and model dimension measure different things:

- **Mutual information:** Given a fixed variable space (X_1, X_2) , how much does observing X_1 tell you about X_2 ? This quantifies shared information on a fixed manifold.
- **Model dimension:** How many independent parameters are needed to specify a distribution in the accessible family? This quantifies the size of the space of distinguishable models.

Manifold expansion is about the second quantity: coupling can increase model dimension even when mutual information is small. Consider the OU example at small κ : the correlation ρ is small, so $I(X_1; X_2) \approx -\frac{1}{2} \log(1-\rho^2) \approx \rho^2/2$ is negligible. Yet the generic rank increased from 2 to 3—a new coordinate direction (ρ) became identifiable. The manifold expanded discretely (rank jumped by 1) even though the “amount of dependence” is continuous and small. MI bounds information within a fixed model; it does not bound which models are accessible.

4.2 Heuristic Scaling: Colony Dynamics

The following is a *heuristic* extension of the preceding results to many-body systems; it provides intuition but is not derived from Theorem 5.

Consider N high-D systems in a “colony” (weakly coupled network). Here “colony” denotes a coupled system whose joint statistical family admits additional Fisher-identifiable collective observables beyond those of individual components.

The *ambient space* of possible interaction parameters grows rapidly: with N subsystems of internal dimension n , pairwise cross-covariances alone contribute $O(N^2 \cdot n^2)$ potential coordinates. However, the *accessible* submanifold dimension is constrained by:

- Number of independent coupling parameters in the dynamics
- Symmetry constraints (e.g., permutation-invariant couplings)
- Attractor structure (many parameter combinations may lead to same stationary distribution)
- Observation map h (unobserved modes don’t contribute to Fisher rank)
- Identifiability degeneracies

A rough upper bound, if each pairwise coupling can create k new identifiable coordinates:

$$\text{rank } I_{\text{colony}} \leq N \cdot \text{rank } I_{\text{single}} + |E| \cdot k \quad (30)$$

where $|E|$ is the number of coupling edges. For dense networks $|E| \sim N^2$; for sparse networks $|E| \sim N$. The actual rank depends on coupling architecture and may be much smaller than this bound. Synchronization collapse (strong coupling driving all subsystems to identical states) and redundancy can saturate growth well below the upper bound.

Remark 23 (Thermodynamic cost). *Superadditive rank growth has a cost: maintaining collective modes at a critical coupling regime requires continuous dissipation [19, 20]. The net accessible complexity is the interaction dividend minus maintenance cost. This trade-off constrains which coupling architectures are sustainable.*

5 Discussion and Outlook

5.1 Summary of Contributions

This paper provides:

1. A formal definition of **manifold expansion** as Fisher-rank increase under coupling (Definition 1)
2. **Theorem 5:** Coupling-induced rank transitions, with three parts:
 - Part I: Pullback identity (Fisher rank = Jacobian rank)

- Part II: Transversality criterion for rank increase
 - Part III: Symmetry-breaking criterion for rank increase
3. Worked examples verifying the theorem: coupled OU (Part II, $\kappa_c = 0$) and Kuramoto oscillators (Part III, $\kappa_c > 0$)
 4. A conjecture on **eigenvalue emergence** under coupling (Conjecture 20)

The core technical contribution is identifying **coupling-induced identifiability activation**: interaction coordinates that are structurally present but dynamically inaccessible can become identifiable under coupling, increasing the Fisher rank of the observable family.

5.2 Testable Predictions

The framework makes specific predictions amenable to experimental or computational test:

1. **Fisher rank increase**: For coupled dynamical systems, estimate Fisher information from time-series data at varying coupling strengths κ . Predict: rank $I(\kappa)$ increases as κ crosses critical thresholds.
2. **Eigenvalue emergence**: New Fisher eigenvalues should emerge continuously from zero (or from below numerical precision) as coupling increases.

Connection to Theorem 5. The eigenvalue emergence test operationalizes the theorem's criteria:

- *Constraint release (Part II)*: An eigenvalue emerging from zero corresponds to a parameter direction v that was in $\ker(dF_0)$ acquiring nonzero transverse variation $d\eta_\kappa(v) \neq 0$.
- *Symmetry breaking (Part III)*: An eigenvalue emerging from zero corresponds to an orbit direction becoming statistically distinguishable once $F_\kappa(\mathcal{B})$ leaves the fixed-point set.

In practice, one observes eigenvalue emergence without needing to verify the geometric criteria directly; the criteria explain *why* the emergence occurs.

Practical Fisher estimation procedure.

- (a) **Choose a parametric approximation** $q(y; \kappa, \beta)$ for the observable distribution (e.g., Gaussian, mixture model, or kernel density estimate with bandwidth as a nuisance parameter).
- (b) **Collect trajectories** at multiple coupling strengths $\kappa_1 < \kappa_2 < \dots < \kappa_m$ spanning the putative critical threshold κ_c .
- (c) **Estimate the FIM** at each κ via the expected outer product of score vectors:

$$\hat{I}_{ij}(\kappa) = \frac{1}{T} \sum_{t=1}^T \frac{\partial \log q(y_t; \kappa, \hat{\beta})}{\partial \beta^i} \frac{\partial \log q(y_t; \kappa, \hat{\beta})}{\partial \beta^j}$$

where $\hat{\beta}$ is the MLE or MAP estimate from the trajectory.

- (d) **Compute eigenvalues** $\lambda_1(\kappa) \leq \lambda_2(\kappa) \leq \dots$ of $\hat{I}(\kappa)$.
- (e) **Track smallest eigenvalues** vs. κ : a rank transition appears as one or more eigenvalues crossing from $\lambda_k \approx 0$ to $\lambda_k > 0$ near κ_c .

For high-dimensional systems, dimensionality reduction (PCA on the score vectors) may be needed before eigenvalue computation.

3. **Superlinear scaling**: For colonies of N coupled subsystems, the identifiable parameter count should grow faster than N in the manifold-expansion regime.
4. **Compression transition**: If expansion eventually saturates, the accessible family should transition from growing to stable (or shrinking), corresponding to collective modes becoming fixed conventions.

5.3 Scope and Limitations

Our analysis assumes that the accessible family forms a smooth parametric model. In practice, the set of stationary distributions induced by dynamics may have irregular structure; our results apply where smooth-family regularity holds at least locally.

The mechanism is geometric, not substrate-specific. It applies wherever high-dimensional coherent systems couple and satisfy the regularity conditions of Theorem 5: smooth dependence on parameters, accessible constraint submanifolds, and observation maps with sufficient richness.

6 Conclusion

Information geometry has been remarkably successful because the fixed-manifold assumption usually holds. We have identified a regime where it does not: when high-dimensional coherent systems couple, the identifiable statistical family can expand.

This is not a rejection of information-theoretic bounds. It is recognition that such bounds apply to fixed model classes, and coupling can change which parameters are identifiable.

The key contribution is formalizing this as **manifold expansion**—defined via Fisher-rank increase—and establishing conditions under which superadditive dimensionality occurs.

The formalization we provide—Definition 1 for manifold expansion and Theorem 5 for rank-change criteria—gives the intuition “coupling creates new degrees of freedom” precise geometric content. The theorem provides checkable conditions: transversality to a constraint submanifold (Part II) or symmetry breaking (Part III). The worked examples verify these criteria for coupled OU processes ($\kappa_c = 0$) and Kuramoto oscillators ($\kappa_c > 0$). The predictions are testable: estimate Fisher information at varying coupling strengths and observe rank changes.

In the regime described here, a useful description is constraint exchange rather than token exchange: high-dimensional systems reshape each other’s accessible state spaces rather than transmitting discrete symbols through a fixed channel.

Code availability: Supplementary simulation code is available at <https://github.com/todd866/manifold-expansion>

Statements and Declarations

Funding. The author did not receive support from any organization for the submitted work.

Competing interests. The author has no relevant financial or non-financial interests to disclose.

Author contributions. Single author.

Data availability. No datasets were generated or analyzed. Simulation code is available at the repository above.

References

- [1] S.-I. Amari. *Information Geometry and Its Applications*. Springer, 2016.
- [2] S.-I. Amari. Information geometry on hierarchy of probability distributions. *IEEE Transactions on Information Theory*, 47(5):1701–1711, 2001.
- [3] N. Ay, J. Jost, H. V. Lê, and L. Schwachhöfer. Information geometry and sufficient statistics. *Probability Theory and Related Fields*, 162(1):327–364, 2015.
- [4] N. Ay, J. Jost, H. V. Lê, and L. Schwachhöfer. *Information Geometry*. Springer, 2017.
- [5] F. Nielsen. An elementary introduction to information geometry. *Entropy*, 22(10):1100, 2020.
- [6] S. Watanabe. *Algebraic Geometry and Statistical Learning Theory*. Cambridge University Press, 2009.

- [7] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley, 2nd edition, 2006.
- [8] J. Rissanen. Fisher information and stochastic complexity. *IEEE Transactions on Information Theory*, 42(1):40–47, 1996.
- [9] P. Gao, E. Trautmann, B. Yu, G. Santhanam, S. Ryu, K. Shenoy, and S. Ganguli. A theory of multineuronal dimensionality, dynamics and measurement. *bioRxiv*, 214262, 2017.
- [10] E. Litwin-Kumar, K. D. Harris, R. Axel, H. Sompolinsky, and L. F. Abbott. Optimal degrees of synaptic connectivity. *Neuron*, 93(5):1153–1164, 2017.
- [11] C. Stringer, M. Pachitariu, N. Steinmetz, C. B. Reddy, M. Carandini, and K. D. Harris. High-dimensional geometry of population responses in visual cortex. *Nature*, 571(7765):361–365, 2019.
- [12] S. H. Strogatz. From Kuramoto to Crawford: exploring the onset of synchronization in populations of coupled oscillators. *Physica D*, 143(1–4):1–20, 2000.
- [13] A. Pikovsky, M. Rosenblum, and J. Kurths. *Synchronization: A Universal Concept in Nonlinear Sciences*. Cambridge University Press, 2001.
- [14] J. A. Acebrón, L. L. Bonilla, C. J. Pérez Vicente, F. Ritort, and R. Spigler. The Kuramoto model: A simple paradigm for synchronization phenomena. *Reviews of Modern Physics*, 77(1):137–185, 2005.
- [15] M. J. Panaggio and D. M. Abrams. Chimera states: coexistence of coherence and incoherence in networks of coupled oscillators. *Nonlinearity*, 28(3):R67–R87, 2015.
- [16] A. H. Knoll and M. A. Nowak. The timetable of evolution. *Science Advances*, 3(5):e1603076, 2017.
- [17] D. H. Erwin, M. Laflamme, S. M. Tweedt, E. A. Sperling, D. Pisani, and K. J. Peterson. The Cambrian conundrum: early divergence and later ecological success in the early history of animals. *Science*, 334(6059):1091–1097, 2011.
- [18] E. A. Sperling, C. A. Frieder, A. V. Raman, P. R. Girguis, L. A. Levin, and A. H. Knoll. Oxygen, ecology, and the Cambrian radiation of animals. *Proceedings of the National Academy of Sciences*, 110(33):13446–13451, 2013.
- [19] J. L. England. Statistical physics of self-replication. *Journal of Chemical Physics*, 139(12):121923, 2013.
- [20] J. M. Horowitz and T. R. Gingrich. Thermodynamic uncertainty relations constrain non-equilibrium fluctuations. *Nature Physics*, 16(1):15–20, 2020.
- [21] J. Hoffmann et al. Training compute-optimal large language models. *arXiv preprint arXiv:2203.15556*, 2022.

- [22] J. Kaplan et al. Scaling laws for neural language models. *arXiv preprint arXiv:2001.08361*, 2020.
- [23] I. A. Chen and P. Walde. From self-assembled vesicles to protocells. *Cold Spring Harbor Perspectives in Biology*, 2(7):a002170, 2010.
- [24] P. Adamski, M. Eleveld, A. Sood, A. Kun, A. Szilágyi, T. Czárán, E. Szathmáry, and S. Otto. From self-replication to replicator systems *en route* to de novo life. *Nature Reviews Chemistry*, 4(8):386–403, 2020.
- [25] J. Maynard Smith and E. Szathmáry. *The Major Transitions in Evolution*. Oxford University Press, 1995.
- [26] P. Turchin. *Ultrasociety: How 10,000 Years of War Made Humans the Greatest Cooperators on Earth*. Beresta Books, 2016.