Communication Beyond Information: Manifold Expansion via High-Dimensional Coupling

Ian Todd

Abstract Many information-theoretic bounds assume a fixed statistical model class. We identify a regime where this assumption fails: when high-dimensional systems couple, the identifiable parameter set changes because the image rank of the dynamics-to-distribution map changes. We formalize this as *manifold expansion*: coupling-induced increase in Fisher rank. Operationally, this means new score directions become nonzero—statistics that were previously insensitive to parameter variation become informative under coupling. Our main theorem provides 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). We prove Fisher rank equals Jacobian rank, so transitions are coordinate-invariant and detectable via eigenvalue emergence in the Fisher information matrix. Two examples demonstrate superadditive complexity: coupled Ornstein–Uhlenbeck processes (transversality, $\kappa_c = 0$) and Kuramoto oscillators (symmetry breaking, $\kappa_c > 0$).

Keywords Fisher information metric · Parameter identifiability · Superadditivity · Coupled dynamical systems · Statistical manifolds

1 Introduction

Information theory provides powerful bounds on communication: channel capacity limits transmission rates, mutual information bounds learning, and data processing inequalities constrain inference (**?**). Many commonly used bounds—PAC generalization, channel coding theorems, Cramér–Rao bounds—assume a fixed model class, channel law, or hypothesis space.

I. Todd
Sydney Medical School, University of Sydney
Sydney, NSW, Australia
E-mail: itod2305@uni.sydney.edu.au

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 $\dim(\mathcal{H})$ (**?**).
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.

Scope clarification. This is an *identifiability/expressivity* claim, not a channel coding claim. We do not argue that mutual information through a coupling channel can exceed capacity bounds. Rather, we show that the *dimension of the identifiable parameter family* can increase under coupling. A concrete illustration: two weakly coupled OU processes may exchange negligible mutual information through the coupling, yet the joint system's identifiable covariance structure has dimension 3 versus dimension 2 for the uncoupled system. The coupling activates a new statistically distinguishable coordinate (the cross-covariance) without requiring high-bandwidth information transfer.

1.3 Paper outline

This paper provides:

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

2 Background: Information geometry

Information geometry studies the differential geometry of statistical manifolds (**???**). A statistical manifold $\mathcal{M}$ is a space of probability distributions, equipped with the Fisher information metric:

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

The Fisher metric induces a Riemannian structure on $\mathcal{M}$. A key feature of information geometry 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 (**?**).

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

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

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

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

? 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 information geometry, 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 (**?**). 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 (**?**). 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** $\mathcal{M}_{\mathrm{acc}}(\kappa)$: 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 $\mathcal{M}_{\mathrm{ind}} = \{p_1 \otimes p_2 : p_i \in \mathcal{M}_i\}$—the product distributions with zero interaction. In information geometry terminology, this is an *e-flat* submanifold corresponding to zero interaction coordinates in the hierarchical log-linear decomposition (**?**).

   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 $\phi_\kappa : X_1 \times X_2 \to X_1 \times X_2$ denote coupled dynamics with coupling strength $\kappa \geq 0$, and let $h : X_1 \times X_2 \to \mathbb{R}^m$ 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$\to$distribution map, not from expanding the parameter space.

For each fixed κ, we have a map:

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

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:

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

Role of the observation map. The observation map $h : X_1 \times X_2 \to \mathbb{R}^m$ 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.

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

Technical note on dimensionality. The ambient space $\mathcal{P}(\mathbb{R}^m)$ of all distributions is infinite-dimensional. For our purposes, we work locally inside a finite-dimensional parametric submanifold $\mathcal{P} \subset \mathcal{P}(\mathbb{R}^m)$ that contains the image of F_κ. This is the setting of classical information geometry: a *regular* statistical manifold where the Fisher information metric is a nondegenerate inner product on each tangent space (**?**). Our results apply to the finite-dimensional accessible family $\mathcal{M}_{\mathrm{acc}}(\kappa) = \mathrm{Image}(F_\kappa) \subset \mathcal{P}$; we do not require the full machinery of infinite-dimensional information geometry.

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

$$r(\kappa) := \sup_{\beta \in \mathcal{B}} \mathrm{rank}\, dF_\kappa(\beta) \tag{4}$$

By upper semicontinuity of rank (the sets $\{\beta : \mathrm{rank}\, dF_\kappa(\beta) \geq r\}$ are open for smooth maps), the set where rank equals the supremum is open. For analytic maps or generic smooth maps this set is also dense; in general applications we restrict to parameter regions where the supremum is attained. When $\mathcal{B}$ is compact, the supremum is a maximum.

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

$$r(\kappa_c^+) > r(\kappa_c^-) \tag{5}$$

Equivalently, the dimension of $\mathcal{M}_{\mathrm{acc}}(\kappa)$ 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.

The novelty is not that interaction parameters exist in the full parametric family—**?**'s hierarchical decomposition already accounts for this. 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**.

Manifold Expansion: Constraint Release
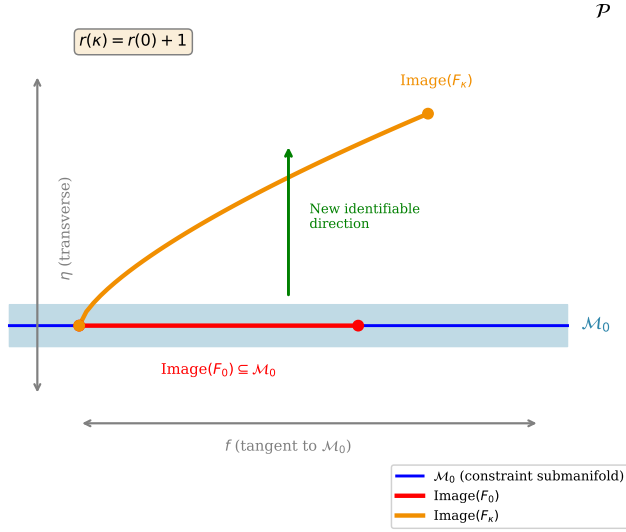
Fig. 1 Geometric schematic of manifold expansion. At $\kappa = 0$, the image of F_0 lies on the constraint submanifold $\mathcal{M}_0$. Under coupling ($\kappa > 0$), the image $F_\kappa(\mathcal{B})$ escapes into the transverse direction η, increasing Fisher rank by at least one.

3.2 Mechanism: collective coordinates

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

Definition 2 (Collective coordinate) A collective coordinate is a function $\psi : X_1 \times X_2 \to \mathbb{R}$ 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

Examples:

- Phase difference between coupled oscillators (**??**)
- Synchronization manifold coordinates (**?**)
- Order parameters of collective states (e.g., chimera states) (**?**)
- 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; 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.

Theorem 1 (Coupling-induced rank transitions) *Let $\mathcal{B}$ be a smooth d-dimensional parameter manifold, let $\mathcal{P}$ be a finite-dimensional **regular** statistical manifold on observation space Y equipped with the Fisher information metric g (assumed positive-definite on each tangent space), and let $F_\kappa : \mathcal{B} \to \mathcal{P}$ 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). *The Fisher information matrix on $\mathcal{B}$ induced by the accessible family $\mathcal{M}_{\mathrm{acc}}(\kappa) = \mathrm{Image}(F_\kappa)$ is the pullback of the Fisher metric:*

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

Consequently:

$$\mathrm{rank}\, I_{\mathcal{B}}(\kappa) = \mathrm{rank}\, dF_\kappa \tag{7}$$

Part II (Transversality 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 $r(0) = \mathrm{rank}\, dF_0(\beta_0)$. 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) ***Kernel activation at β_0:** There exists $v \in T_{\beta_0}\mathcal{B}$ with $df_0(v) = 0$ (hence $v \in \ker(dF_0(\beta_0))$) such that $d\eta_\kappa|_{\beta_0}(v) \neq 0$ for $\kappa > \kappa_c$*

Geometrically, (C2) says F_κ becomes transverse to $\mathcal{M}_0$ at $F_\kappa(\beta_0)$: the image acquires a component in the normal bundle $N\mathcal{M}_0$. Then for $\kappa > \kappa_c$:

$$r(\kappa) \geq r(0) + 1 \tag{8}$$

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

*   **Interpretation note.** When analyzing symmetry breaking, we work on the quotient $\mathcal{P}/G$ where the G-action has been quotiented out (equivalently, we gauge-fix by choosing a representative from each orbit). The rank increase arises because parameter directions in $\ker(dF_0)$ acquire nonzero images in the*

quotient geometry—typically via G-equivariant observables like order parameters that become parameter-sensitive when symmetry breaks.
    Suppose:

(S1) *$F_0(\mathcal{B}) \subseteq \mathrm{Fix}(G)$ (distributions are G-invariant at $\kappa = 0$)*

(S2) *For $\kappa > \kappa_c$, there exists β such that $F_\kappa(\beta) \notin \mathrm{Fix}(G)$ (symmetry broken)*

(S3) ***Kernel activation via symmetry breaking**: There exists $v \in \ker(dF_0(\beta_0))$ such that a G-invariant observable $\phi : \mathcal{P} \to \mathbb{R}$ (e.g., order parameter magnitude) satisfies $d(\phi \circ F_\kappa)|_{\beta_0}(v) \neq 0$ for $\kappa > \kappa_c$*

Then for $\kappa > \kappa_c$:

$$r(\kappa) \geq r(0) + 1 \tag{9}$$

More generally, if there are k independent invariant observables activated, the rank increase is at least k.

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. By the chain rule:

$$\frac{\partial \log q(y; \kappa, \beta)}{\partial \beta^i} = \sum_\alpha \frac{\partial \log q}{\partial \theta^\alpha} \frac{\partial \theta^\alpha}{\partial \beta^i} \tag{10}$$

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

$$(F_\kappa^* g)_{ij} = g\left(\frac{\partial F_\kappa}{\partial \beta^i}, \frac{\partial F_\kappa}{\partial \beta^j}\right) = (I_\mathcal{B}(\kappa))_{ij} \tag{11}$$

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 $\ker(I_\mathcal{B}(\kappa)) = \ker(dF_\kappa)$ and $\mathrm{rank}(I_\mathcal{B}(\kappa)) = \mathrm{rank}\, dF_\kappa$.

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

    Let $v \in \ker(df_0(\beta_0))$ be as in (C2). Since $\eta_0 \equiv 0$, we have $d\eta_0(v) = 0$, hence $dF_0(v) = (0, 0)$ and $v \in \ker(dF_0(\beta_0))$. By (C2), for $\kappa > \kappa_c$, $d\eta_\kappa|_{\beta_0}(v) \neq 0$, so $dF_\kappa(v)$ has a nonzero transverse component.

    The tangent space decomposes as $T_p\mathcal{P} \cong T_p\mathcal{M}_0 \oplus N_p\mathcal{M}_0$. At $\kappa = 0$, the image of dF_0 lies entirely in $T\mathcal{M}_0$. The vector $dF_\kappa(v)$ has a nonzero component in $N\mathcal{M}_0$, so it is linearly independent of the $r(0)$ tangential directions. By continuity, the original $r(0)$ directions persist, so $r(\kappa) \geq r(0) + 1$.

    Part III. At $\kappa = 0$, (S1) implies $F_0(\mathcal{B}) \subseteq \mathrm{Fix}(G)$, so all G-invariant observables ϕ are constant on $F_0(\mathcal{B})$: varying β does not change $\phi(F_0(\beta))$.

   Let $v \in \ker(dF_0(\beta_0))$ and ϕ be as in (S3). At $\kappa = 0$, $dF_0(v) = 0$, so certainly $d(\phi \circ F_0)(v) = 0$. For $\kappa > \kappa_c$, condition (S3) states that $d(\phi \circ F_\kappa)|_{\beta_0}(v) \neq 0$.

Since ϕ is G-invariant, it descends to the quotient $\mathcal{P}/G$. The nonzero derivative $d(\phi \circ F_\kappa)(v)$ implies that $dF_\kappa(v)$ has a nonzero component in directions that affect G-invariant observables. This direction was in $\ker(dF_0)$ but leaves the kernel under coupling, giving $r(\kappa) \geq r(0) + 1$.

Corollary 1 (Operational detection of rank transitions) *Given time-series data at coupling strengths $\kappa_1 < \kappa_2 < \cdots < \kappa_m$, rank transitions can be detected as follows:*

1. **Estimate Fisher information**: *For each κ_j, estimate the Fisher information matrix $\hat{I}(\kappa_j)$ from data—either directly via score function estimation, or via the Hessian of the log-likelihood at the MLE.*
2. **Track eigenvalues**: *Compute eigenvalues $\lambda_1(\kappa_j) \geq \lambda_2(\kappa_j) \geq \cdots$ of $\hat{I}(\kappa_j)$.*
3. **Detect emergence**: *A rank transition at κ_c manifests as one or more eigenvalues crossing from $\lambda_k \approx 0$ (within estimation noise) to $\lambda_k > 0$ significantly.*

For the OU example, the minimal sufficient statistics are the sample covariance entries $(\hat{\Sigma}_{11}, \hat{\Sigma}_{22}, \hat{\Sigma}_{12})$; rank increase corresponds to $\hat{\Sigma}_{12}$ becoming β-sensitive. For Kuramoto, the order parameter magnitude r and phase Ψ suffice; rank increase corresponds to Ψ becoming parameter-sensitive above K_c.
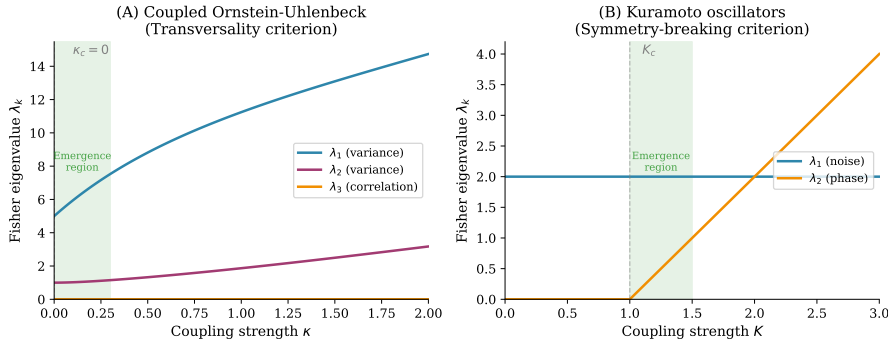
Fig. 2 Eigenvalue emergence under coupling. (A) Coupled Ornstein–Uhlenbeck: the third Fisher eigenvalue (correlation direction) emerges immediately at $\kappa > 0$. (B) Kuramoto oscillators: the phase-sensitivity eigenvalue emerges at the critical coupling K_c. Both demonstrate coupling-induced identifiability activation predicted by Theorem **??**.

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.

Definition 3 (Component accessible families) For each subsystem $i = 1, \ldots, n$, let $\mathcal{B}_i$ be its parameter manifold and let $F_i : \mathcal{B}_i \to \mathcal{P}_i$ be its dynamics-to-distribution map. Define the component accessible family $\mathcal{M}_i := \mathrm{Image}(F_i)$ and its generic rank $r_i := \max_{\beta_i \in \mathcal{B}_i} \mathrm{rank}\, dF_i(\beta_i)$.

Definition 4 (Superadditive accessible complexity) Let $F_\kappa : \mathcal{B} \to \mathcal{P}$ be the coupled dynamics-to-distribution map with generic rank $r(\kappa) := \max_{\beta \in \mathcal{B}} \mathrm{rank}\, dF_\kappa(\beta)$. The coupled system exhibits *superadditive accessible complexity* at coupling κ if

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

Lemma 1 (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 $\mathrm{rank}\, dF_\oplus(\beta) = \sum_i \mathrm{rank}\, dF_i(\beta_i)$. Maximizing over β yields $r_\oplus = \sum_i r_i$.

Proposition 1 (Sufficient condition via non-factorizable coordinates) *Suppose there exists a smooth map $\pi : \mathcal{P} \to \mathbb{R}^k$ such that:*

(A1) $\pi \circ F_\oplus$ is locally constant on some open $U \subset \mathcal{B}$;
(A2) $\pi \circ F_\kappa$ has Jacobian rank k on U.

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

3.5 Example 1: coupled Ornstein–Uhlenbeck (transversality)

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

Dynamics. Consider two coupled Ornstein–Uhlenbeck processes:

$$dX_1 = -\gamma_1 X_1\, dt + \kappa(X_2 - X_1)\, dt + \sigma_1\, dW_1 \tag{12}$$

$$dX_2 = -\gamma_2 X_2\, dt + \kappa(X_1 - X_2)\, dt + \sigma_2\, dW_2 \tag{13}$$

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

Parameter manifold and observation. The physical parameter manifold is $\mathcal{B} = \{(\gamma_1, \sigma_1, \gamma_2, \sigma_2) : \gamma_i > 0, \sigma_i > 0\} \cong (\mathbb{R}^+)^4$, which is 4-dimensional. The observation map is full state observation: $h(X_1, X_2) = (X_1, X_2)$, so the observable is the bivariate Gaussian with covariance Σ. The map $F_\kappa : \mathcal{B} \to \mathcal{P}$ sends $\beta = (\gamma_1, \sigma_1, \gamma_2, \sigma_2)$ to the centered bivariate Gaussian with covariance $\Sigma(\kappa, \beta)$. We compute the rank of the Jacobian $\partial(\Sigma_{11}, \Sigma_{22}, \Sigma_{12})/\partial\beta$.

Stationary distribution. The stationary distribution is bivariate Gaussian with zero mean and covariance matrix Σ satisfying the Lyapunov equation $A\Sigma + \Sigma A^T + D = 0$. For the symmetric case $\gamma_1 = \gamma_2 = \gamma$ and $\sigma_1 = \sigma_2 = \sigma$:

$$\Sigma_{11} = \Sigma_{22} = \frac{\sigma^2(\gamma + \kappa)}{2\gamma(\gamma + 2\kappa)}, \qquad \Sigma_{12} = \frac{\kappa\sigma^2}{2\gamma(\gamma + 2\kappa)} \tag{14}$$

giving correlation $\rho = \Sigma_{12}/\Sigma_{11} = \kappa/(\gamma + \kappa)$.

This confirms $\rho(\kappa = 0) = 0$ and $\rho \to 1$ as $\kappa \to \infty$ (strong coupling induces perfect correlation, as expected physically). Computing $\partial\rho/\partial\gamma$:

$$\frac{\partial\rho}{\partial\gamma} = -\frac{\kappa}{(\gamma + \kappa)^2} \neq 0 \quad \text{for } \kappa > 0 \tag{15}$$

This nonzero derivative demonstrates kernel activation: a direction that left ρ unchanged at $\kappa = 0$ now varies ρ at $\kappa > 0$.

Verification of Theorem ?? Part II. The ambient manifold $\mathcal{P}$ of centered bivariate Gaussians is 3-dimensional, parameterized by (v_1, v_2, ρ). The independence submanifold is $\mathcal{M}_0 = \{(v_1, v_2, 0)\}$, which has codimension $c = 1$.

At $\kappa = 0$: $\rho = 0$ for all β, so $F_0(\mathcal{B}) \subseteq \mathcal{M}_0$ and $r(0) = 2$.

At $\kappa > 0$: $\rho > 0$, so $F_\kappa(\mathcal{B})$ escapes $\mathcal{M}_0$ and $r(\kappa) = 3$.

Role of observation. The rank increase from 2 to 3 assumes the observation map h is rich enough to identify the cross-covariance Σ_{12}. If h only observes X_1 (marginal observation), or only observes variances $(\Sigma_{11}, \Sigma_{22})$ without cross-terms, the correlation ρ remains unidentifiable and the rank increase is not observable. This illustrates that both dynamics and observation contribute to the accessible family.

Superadditivity. 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 $r(\kappa) = 3 > r_\oplus = 2$: *strict superadditivity*.
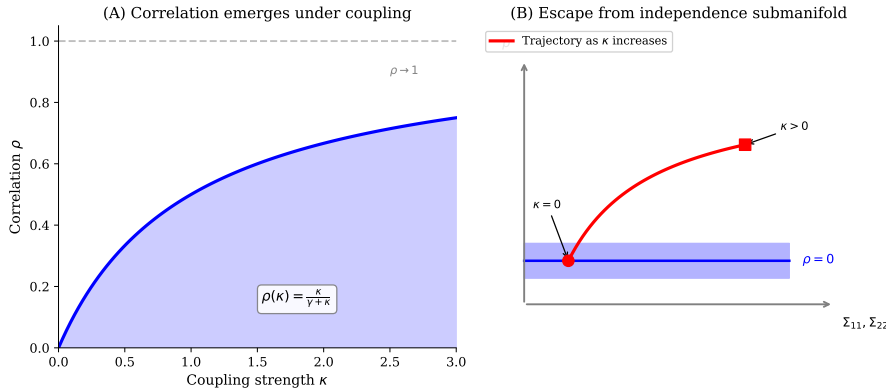
Fig. 3 Coupled OU: correlation emerges under coupling. (A) The correlation $\rho = \kappa/(\gamma + \kappa)$ increases monotonically from zero, asymptoting to 1 as $\kappa \to \infty$. (B) Geometric interpretation: at $\kappa = 0$, the accessible family lies on the independence submanifold ($\rho = 0$); coupling moves the trajectory into the transverse direction.

3.6 Example 2: Kuramoto synchronization (symmetry breaking)

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

Notation. Following the Kuramoto literature, we denote coupling strength by K (rather than κ) in this section. The role is identical: K indexes the dynamical system, not a coordinate in the parameter space $\mathcal{B}$.

Parameter space. The physical parameter manifold is $\mathcal{B} = \{(D, \nu) : D > 0, \nu \in \mathcal{P}(\mathbb{R})\}$ where D is noise intensity and ν is the frequency distribution. For concreteness, take $\nu = \mathcal{N}(\omega_0, \sigma_\omega^2)$ so that $\mathcal{B} \cong \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^+$ is 3-dimensional. The mean phase Ψ is *not* a coordinate in $\mathcal{B}$—it parameterizes the S^1-orbit of solutions.

Mean-field dynamics. 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:

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial \theta} \left[\rho \left(\omega + Kr \sin(\Psi - \theta) \right) \right] = D \frac{\partial^2 \rho}{\partial \theta^2} \tag{16}$$

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

Stationary solutions and gauge fixing. For $K < K_c$, the unique 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 $r > 0$. To satisfy assumption (A1), we define F_K to return the **gauge-fixed representative** with $\Psi = 0$. Equivalently, one can define F_K to return the equivalence class $[\rho]$ under S^1; the rank analysis is unchanged since we quotient by the symmetry.

Verification of Theorem ?? Part III.

- (S1): For $K < K_c$, $F_K(\mathcal{B})$ is contained in the S^1-fixed point set (the incoherent state with $r = 0$).
- (S2): For $K > K_c$, the synchronized stationary solutions have $r > 0$, breaking S^1 symmetry.
- (S3) **Kernel activation via order parameter**: Consider varying the noise intensity D or the frequency spread σ_ω. For $K < K_c$, these parameters do not affect the (unique) incoherent distribution, so they lie in $\ker(dF_K)$. For $K > K_c$, these parameters affect the order parameter magnitude: $\partial r/\partial D \neq 0$ and $\partial r/\partial \sigma_\omega \neq 0$. The S^1-invariant observable $\phi = r$ becomes parameter-sensitive.

The theorem predicts: $r(K) \geq r(0) + 1$ for $K > K_c$. The "+1" comes from the order parameter magnitude becoming identifiable—a G-invariant observable that was constant on the incoherent manifold but varies with parameters on the synchronized manifold.

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 $K_c > 0$.

4 Consequences for complexity growth

4.1 Faster than fixed-class bounds predict

Standard bounds on complexity growth assume fixed parameterization, fixed sufficient statistics, and 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 5 (Model complexity) We define *complexity C* of a statistical family $\mathcal{M}$ as the Fisher rank: the dimension of the identifiable tangent space. This is coordinate-invariant and intrinsic to the manifold geometry.

Clarification. Fisher rank measures *identifiability dimension*—the number of statistically distinguishable directions—not computational or algorithmic complexity. We use "complexity" in the sense of "accessible degrees of freedom under observation," following the information-geometric tradition where model dimension is a natural complexity measure (**?**).

Rank vs. effective rank. Fisher rank is a binary notion: a direction is either in the kernel (rank 0 contribution) or not (rank 1 contribution). It does not distinguish "sloppy" directions (small but nonzero eigenvalues) from "stiff" directions (large eigenvalues). In finite-sample settings, practical identifiability depends on whether eigenvalues exceed a noise floor. Our theoretical results concern the transition from $\lambda = 0$ to $\lambda > 0$; Conjecture **??** addresses the empirical signature via an ϵ-threshold. For applications, one would use an *effective rank* (e.g., number of eigenvalues above noise level) rather than exact rank.

Conjecture 1 (Eigenvalue emergence under coupling) In regimes where the accessible family $\mathcal{M}_{\mathrm{acc}}(\kappa)$ can be embedded in a fixed parameterization across κ, eigenvalue emergence provides a concrete signature. As coupling strength κ crosses a critical threshold κ_c, one or more Fisher eigenvalues emerge from zero:

$$\lambda_k(\kappa_c^-) < \epsilon \quad \text{and} \quad \lambda_k(\kappa_c^+) > \epsilon \tag{17}$$

This implies rank increase as coupling activates previously degenerate interaction coordinates.

4.2 Heuristic scaling: colony dynamics

Consider N high-D systems in a "colony" (weakly coupled network). 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 coupling parameters, symmetry constraints, attractor structure, and observation map.

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

$$\operatorname{rank} I_{\text{colony}} \leq N \cdot \operatorname{rank} I_{\text{single}} + |E| \cdot k \tag{18}$$

where $|E|$ is the number of coupling edges. For dense networks $|E| \sim N^2$; for sparse networks $|E| \sim N$.

5 Discussion

5.1 Summary of contributions

This paper provides:

1. A formal definition of **manifold expansion** as Fisher-rank increase under coupling (Definition **??**)
2. **Theorem ??**: 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, $\kappa_c > 0$)

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:

1. **Fisher rank increase**: For coupled dynamical systems, estimate Fisher information from time-series data at varying coupling strengths κ. Predict: $\operatorname{rank} I(\kappa)$ increases as κ crosses critical thresholds.
2. **Eigenvalue emergence**: New Fisher eigenvalues should emerge continuously from zero as coupling increases.
3. **Superlinear scaling**: For colonies of N coupled subsystems, the identifiable parameter count should grow faster than N in the manifold-expansion regime.

5.3 Practical estimation protocol

We outline a concrete protocol for detecting manifold expansion from data.

Setup. Given time-series observations $\{y_t^{(\kappa)}\}_{t=1}^T$ at coupling strengths $\kappa \in \{\kappa_1, \ldots, \kappa_m\}$, the goal is to detect whether Fisher rank increases across the κ sweep.

Step 1: Choose sufficient statistics. For each example:

- *Coupled OU*: Use sample covariance $\hat{\Sigma} = (\hat{\Sigma}_{11}, \hat{\Sigma}_{22}, \hat{\Sigma}_{12})$
- *Kuramoto*: Use order parameter (r, Ψ) estimated from phase time-series
- *General*: Use moments or cumulants up to order matching the expected model dimension

Step 2: Estimate Fisher information. Two approaches:

(a) *Score-based*: Estimate partial derivatives $\partial_\beta \log p(y|\beta)$ at the MLE $\hat{\beta}$; compute $\hat{I}_{ij} = \frac{1}{T} \sum_t s_i(y_t) s_j(y_t)$ where $s_i = \partial_{\beta_i} \log p$.

(b) *Hessian-based*: Compute $\hat{I} = -\nabla_\beta^2 \ell(\hat{\beta})$ where ℓ is the log-likelihood.

Step 3: Track eigenvalue spectrum. For each κ_j, compute eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots$ of $\hat{I}(\kappa_j)$. Plot $\lambda_k(\kappa)$ versus κ.

Step 4: Detect rank transition. A rank increase at κ_c appears as:

- One or more eigenvalues crossing from $\lambda_k < \epsilon$ to $\lambda_k \gg \epsilon$
- The threshold ϵ should account for finite-sample noise (bootstrap confidence intervals recommended)

Expected signatures by example:

- *OU (transversality, $\kappa_c = 0$)*: The third eigenvalue (corresponding to Σ_{12}) emerges immediately at any $\kappa > 0$
- *Kuramoto (symmetry breaking, $\kappa_c > 0$)*: An eigenvalue corresponding to phase sensitivity emerges at $K = K_c$, with the transition sharpening as system size increases

5.4 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 **??**.

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.

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.

Statements and declarations

References

Acebrón, J. A., Bonilla, L. L., Pérez Vicente, C. J., Ritort, F., Spigler, R. (2005). The Kuramoto model: A simple paradigm for synchronization phenomena. *Reviews of Modern Physics*, 77(1), 137–185

Amari, S.-I. (2001). Information geometry on hierarchy of probability distributions. *IEEE Transactions on Information Theory*, 47(5), 1701–1711

Amari, S.-I. (2016). *Information Geometry and Its Applications*. Springer

Ay, N., Jost, J., Lê, H. V., Schwachhöfer, L. (2015). Information geometry and sufficient statistics. *Probability Theory and Related Fields*, 162(1), 327–364

Ay, N., Jost, J., Lê, H. V., Schwachhöfer, L. (2017). *Information Geometry*. Springer

Cover, T. M., Thomas, J. A. (2006). *Elements of Information Theory* (2nd ed.). Wiley

Nielsen, F. (2020). An elementary introduction to information geometry. *Entropy*, 22(10), 1100

Panaggio, M. J., Abrams, D. M. (2015). Chimera states: coexistence of coherence and incoherence in networks of coupled oscillators. *Nonlinearity*, 28(3), R67–R87

Pikovsky, A., Rosenblum, M., Kurths, J. (2001). *Synchronization: A Universal Concept in Nonlinear Sciences*. Cambridge University Press

Strogatz, S. H. (2000). From Kuramoto to Crawford: exploring the onset of synchronization in populations of coupled oscillators. *Physica D*, 143(1–4), 1–20

Watanabe, S. (2009). *Algebraic Geometry and Statistical Learning Theory*. Cambridge University Press