www.ijarp.com ISSN 2456-9992 Page: 01-22

International Journal Advanced Research Publications

Research Article

Volume: 02
Issue: 08

INFORMATION-GEOMETRIC GRAMMAR PRINCIPLE

***Angelo Plastino**

Universidad Nacional de La Plata 1900 La Plata, Argentina August 19, 2026.

Article Received: 11 July 2026, Article Revised: 31 July 2026, Published on: 20 August 2026

***Corresponding Author: Angelo Plastino**

Universidad Nacional de La Plata 1900 La Plata, Argentina August 19, 2026.

Doi: https://doi-doi.org/101555/ijarp.2121

ABSTRACT

We introduce the **Information-Geometric Grammar Principle (IGGP)**, a mathematical framework that unifies concepts from for-mal language theory, information geometry, Bayesian inference, and complex adaptive systems. In conventional grammar-based descrip-tions of adaptive systems, individual agents constitute the alphabet, interaction rules define the grammar, and collective behaviors emerge as the language generated by repeated applications of these rules. We generalize this paradigm by replacing symbolic representations with probabilistic ones. Each agent is identified with a local statistical man-ifold endowed with a Fisher information metric, while communication acts as a geometric production operator that transforms probability distributions through Bayesian updating and information exchange. Repeated communication generates an evolving collective statistical manifold whose geometry encodes the emergent organization of the system.

Within this formulation, communication assumes the role of a probabilistic grammar: the elementary statistical states of individual agents constitute the alphabet, communication operators define the production rules, and the family of admissible collective probability distributions forms the language generated by the adaptive dynam-ics. The resulting Fisher geometry provides a quantitative description of syntactic organization, with off-diagonal metric components mea-suring communication-induced statistical dependencies among agents. Emergent collective structures are therefore interpreted as geometric properties of a statistical language rather than as purely symbolic constructions.

The Information-Geometric Grammar Principle naturally extends traditional grammar-based models by introducing intrinsic geometric observables, including curvature, geometric entropy, and communication-induced coupling measures, thereby linking

symbolic organization with differential geometry. The framework is sufficiently general to encom-pass biological populations, social and communication networks, dis-tributed robotic systems, scientific communities, and human–artificial intelligence ecosystems. More broadly, it suggests that the emergence of collective organization can be understood as the generation of an increasingly structured statistical language whose syntax is encoded by an evolving Fisher geometry. This perspective establishes a unified mathematical language for studying adaptive organization across dis-ciplines and opens new directions for the geometric analysis of learn-ing, communication, and emergence in complex systems.

KEYWORDS: Complex Adaptive Systems; Computational Sci-ences; Financial markets; Emergent Collective behavior.

1 INTRODUCTION

Complex adaptive systems (CAS) arise throughout the natural, social, and computational sciences, encompassing ecosystems, neural networks, scien-tific communities, financial markets, robotic swarms, and human–artificial intelligence ecosystems. Despite their diversity, these systems share several characteristic features: they consist of multiple interacting agents, contin-uously exchange information, adapt to changing environments, and exhibit emergent collective behavior that cannot be understood solely from the prop-erties of their individual components [2, 3, 4].

One influential approach to describing CAS has been through grammar-based models. Within this perspective, individual agents are viewed as the elementary symbols of an alphabet, while interaction rules play the role of grammatical production rules that generate increasingly complex collective structures. The resulting ”language” represents the family of admissible sys-tem configurations and provides a useful qualitative description of emergence and self-organization [5, 1].

Although grammar-based approaches successfully capture the combina-torial organization of adaptive systems, they remain fundamentally symbolic. The symbols themselves possess no intrinsic quantitative structure, and the grammatical rules describe admissible transformations without providing a natural measure of statistical similarity, uncertainty, or adaptation. Con-sequently, traditional grammars cannot directly quantify the geometry of learning or the evolution of collective information.

A complementary viewpoint has emerged from information geometry, where families of probability distributions are regarded as differentiable man-ifolds endowed with the Fisher information metric [6, 7, 8]. Within this framework, probability distributions become geometric objects, and statisti-cal inference may be interpreted as motion on curved manifolds. Information geometry has found important applications in statistics, machine learning, thermodynamics, quantum information, and biological systems, providing a natural language for studying learning and adaptation [9, 10].

The present work seeks to unify these two perspectives. Rather than regarding agents as abstract symbols, we represent each agent by a local probability distribution or, equivalently, by a point on a statistical manifold. Communication among agents then acts not merely as symbolic rewriting but as a transformation of probability distributions. In this way, communication becomes a probabilistic production operator that continuously modifies the collective statistical manifold.

This observation motivates what we call the *Information-Geometric Gram-mar Principle* (IGGP). The central idea is that grammar itself possesses an intrinsic geometric structure once its elementary symbols are replaced by statistical states. Individual agents constitute the statistical alphabet, com-munication operators define the production rules, and the evolving family of collective probability distributions forms the language generated by repeated adaptive interactions. The Fisher information metric endows this language with an intrinsic geometry, allowing grammatical organization to be char-acterized quantitatively through geometric observables such as curvature, entropy, and communication-induced statistical couplings [12].

The Information-Geometric Grammar Principle naturally generalizes clas-sical grammar-based descriptions of complex adaptive systems. Instead of generating symbolic expressions, the grammar generates an evolving statisti-cal manifold whose geometry reflects the organization of the communicating population. Emergent collective behavior is thus interpreted as the progres-sive geometric structuring of a statistical language through communication and Bayesian adaptation.

The framework developed here establishes a bridge between formal lan-guage theory, information geometry, Bayesian inference, and the theory of complex adaptive systems. Besides providing a unified mathematical descrip-tion of adaptive organization, it suggests that many apparently unrelated systems—including biological populations, social networks, distributed arti-ficial intelligence, scientific communities, and multi-agent learning systems—can be understood within a common geometric grammar generated by prob-abilistic communication.

2 The Statistical Alphabet and Information-Geometric Grammar

The Information-Geometric Grammar Principle (IGGP) begins by replacing the symbolic alphabet of conventional grammars with a statistical alphabet. In this formulation, the elementary constituents of a complex adaptive sys-tem are no longer abstract symbols but adaptive probability distributions. Communication acts upon these statistical objects, generating an evolving probabilistic language whose structure is described by information geome-try.

2.1 The Statistical Alphabet

Consider a population of N adaptive agents,

$$\mathcal{A} = \{A_1, A_2, \ldots, A_N\}.$$

Each agent is characterized by an internal probabilistic description of its environment,

$$p_i(x; \vartheta_i), \qquad i = 1, \ldots, N,$$

where x denotes the observable variables and

$$\vartheta_i = (\vartheta_i^1, \ldots, \vartheta_i^{m})$$

is a vector of statistical parameters.

Each probability distribution defines a statistical manifold

$$(\mathcal{M}_i, g_i),$$

whose intrinsic geometry is determined by the Fisher information metric.

The collection

$$\mathrm{A} = \{(\mathcal{M}_1, g_1), \ldots, (\mathcal{M}_N, g_N)\}$$

constitutes the *statistical alphabet* of the adaptive system.

Unlike the symbols of a conventional grammar, the elements of the statis-tical alphabet possess an intrinsic differential-geometric structure. Distances between letters are measured by the Fisher metric, reflecting their statistical distinguishability.

2.2 Communication as Grammar

In formal language theory, grammatical production rules specify how sym-bols may be combined to generate admissible words and sentences. The Information-Geometric Grammar Principle extends this concept by replac-ing symbolic rewriting with probabilistic transformations.

Communication between two agents,

$$A_i \longrightarrow A_j,$$

is represented by a communication operator

$$L_{ij},$$

which acts on the corresponding probability distributions,

$$L_{ij} : (p_i, p_j) \longrightarrow (P_i, P_j).$$

The transformed distributions incorporate the information exchanged during communication and generally differ from their initial states.

Communication therefore functions as a probabilistic production rule. Rather than generating new symbolic expressions, it generates new statistical states on the collective manifold.

2.3 The Statistical Language

Initially, independent agents satisfy

$$P_0(x_1, \dots, x_N) = \prod_{i=1}^{N} p_i(x_i). \qquad (1)$$

Successive communication events progressively generate statistical corre-lations,

$$P(x_1, \dots, x_N) \neq \prod_{i=1}^{N} p_i(x_i). \qquad (2)$$

The family $L = \{P(x_1, \dots, x_N)\}$ of all probability distributions reachable through repeated communication defines the *statistical language* generated by the communication grammar.

Each element of this language corresponds to a possible collective orga-nization of the adaptive population.

2.4 Communication Grammar

The communication grammar consists of the collection

$$G = (\mathbb{A}, L_C, L),$$

where

- A is the statistical alphabet,
- L_C denotes the set of communication operators,
- L is the statistical language generated by repeated communication.
- This triplet plays a role analogous to the classical grammar

$$(\Sigma, P, L),$$

where Σ denotes the alphabet, P the production rules, and L the gener-ated language. The essential difference is that the present grammar is defined on a space of probability distributions rather than symbolic strings.

2.5 Emergence of Collective Structure

Repeated communication modifies not only individual probability distribu-tions but also the global organization of the statistical language.

Initially independent statistical letters become progressively coupled, pro-ducing increasingly structured collective probability distributions.

The grammar therefore generates organization at two complementary lev-els.

First, it determines which collective statistical states are admissible. Second, it continuously reshapes the geometric relationships among those states.

The latter aspect distinguishes the Information-Geometric Grammar Prin-ciple from conventional symbolic grammars and provides the basis for the geometric theory developed in the following sections.

2.6 Discussion

The central innovation introduced in this section is the replacement of sym-bolic alphabets by statistical alphabets. Communication is thereby reinter-preted as a probabilistic production mechanism acting on statistical mani-folds rather than on discrete symbols. The resulting statistical language pos-sesses an intrinsic information geometry that quantifies the distinguishabil-ity, organization, and evolution of collective adaptive states. This viewpoint naturally prepares the introduction of the Fisher information metric, which provides the mathematical structure underlying the Information-Geometric Grammar Principle. Thus, we have

* **agents** → statistical alphabet,
* **communication** → grammar,
* **collective probability distributions** → language,
* **information geometry** → syntax of the language.

3 Information-Geometric Grammar Principle

The preceding section established the correspondence between a grammar and a statistical manifold. Each admissible sentence generated by the gram-mar determines a probability distribution $p(x|\theta)$, while the Fisher informa-tion metric endows the resulting manifold M with a Riemannian structure.

The remaining question is therefore dynamical: among all admissible gram-matical transformations, which ones are naturally selected?

The central postulate of the present work is that grammatical evolution follows the intrinsic geometry of the statistical manifold. Grammar produc-tion is therefore not merely combinatorial but geometric.

3.1 The Principle

Consider two successive grammatical states represented by probability dis-tributions

$$p(x|\vartheta) \quad -\!\!\longrightarrow \quad p(x|\vartheta + d\vartheta). \tag{3}$$

Their infinitesimal statistical distinguishability is measured by the Fisher line element

$$ds^2 = g_{ij}(\vartheta)\, d\vartheta^i d\vartheta^j, \tag{4}$$

where

$$g_{ij} = \sum_x p(x|\vartheta)\, \partial_i \ln p\, \partial_j \ln p. \tag{5}$$

We propose the following principle.

Information-Geometric Grammar Principle (IGGP).

Among all grammatically admissible transformations, the realized evo-lution is the one that extremizes an information-geometric action built from the Fisher metric.

Equivalently, grammatical evolution follows geodesics of the statistical manifold whenever no external constraints are imposed.

The corresponding action is

$$\mathcal{A} = \int g_{ij}\frac{d\vartheta^i}{dt}\frac{d\vartheta^j}{dt}\,dt, \qquad (6)$$

Application of the Euler–Lagrange equations yields the geodesic equa-tions

$$\delta\mathcal{A} = 0. \qquad (7)$$

$$\frac{d^2\vartheta^k}{dt^2} + \Gamma^k_{ij}\frac{d\vartheta^i}{dt}\frac{d\vartheta^j}{dt} = 0, \qquad (8)$$

$$\Gamma^k_{ij} = \frac{1}{2}g^{kl}\left(\partial_i g_{jl} + \partial_j g_{il} - \partial_l g_{ij}\right) \qquad (9)$$

Where

Are the Christoffel symbols associated with the Fisher metric. Thus grammatical production becomes a geometric flow.

3.2 Interpretation

Traditional formal grammars specify which productions are allowed but re-main agnostic regarding which admissible production is actually selected. The Information-Geometric Grammar Principle supplies precisely this miss-ing selection rule.

Suppose several production rules can be applied at a given stage,

$$R_1, R_2, \ldots, R_n.$$

Each rule induces a displacement on the statistical manifold,

$$\Delta\vartheta_1, \Delta\vartheta_2, \ldots, \Delta\vartheta_n.$$

Their corresponding Fisher lengths are

$$L_i^2 = g_{jk}\Delta\vartheta_i^j\Delta\vartheta_i^k. \qquad (10)$$

The IGGP predicts that the preferred grammatical transformation mini-mizes the information-geometric action rather than an arbitrary combinato-rial criterion. In this sense, grammar acquires a variational structure analo-gous to Hamilton's principle in mechanics.

3.3 Complex Adaptive Systems

Within the framework of Complex Adaptive Systems (CAS), the grammar alphabet is identified with the collection of interacting agents, while pro-duction rules represent possible interactions among them. The probability distribution over grammatical strings therefore becomes a probability distri-bution over collective system configurations.

Learning, adaptation, and evolution correspond to trajectories on the statistical manifold. The Fisher metric measures the distinguishability of collective system states, whereas geodesic evolution identifies the statistically most economical adaptive pathway.
Consequently, adaptation appears as an emergent geometric phenomenon rather than as an externally imposed optimization procedure. The geometry itself determines which transitions are statistically natural.

3.4 Relation to Existing Variational Principles

The proposed principle is closely related to several well-known geometric optimization principles.
First, in information geometry, geodesics represent paths of minimal sta-tistical distance between probability distributions. Second, in optimal trans-port, evolution minimizes a suitable transportation cost. Third, in statisti-cal mechanics, equilibrium distributions arise through entropy-maximization principles.
The Information-Geometric Grammar Principle differs from these ap-proaches in that the objects undergoing optimization are grammatical pro-ductions. The optimization variable is therefore not a physical trajectory but a sequence of admissible symbolic transformations. Consequently, the grammar itself becomes a dynamical object whose evolution is governed by the geometry induced by probability distributions.
This establishes a unified bridge between formal language theory, infor-mation geometry, statistical inference, and complex adaptive systems.
This section:

- Defines the **Information-Geometric Grammar Principle (IGGP)** as a variational principle;
- Derives the **geodesic equation** governing grammar evolution;

- Explains how **grammar acquires dynamics**, not just syntax;
- Explicitly connects the framework to **Complex Adaptive Systems (CAS)**, where agents act as the alphabet and interactions as production rules.

4 Worked Examples

To illustrate the Information-Geometric Grammar Principle (IGGP), we now consider two simple examples. The first admits an exact analytical treatment, while the second demonstrates how the same geometric ideas naturally extend to adaptive multi-agent systems.

4.1 Example I: Bernoulli Grammar

Consider the simplest possible grammar generated by a binary alphabet

$$\Sigma = \{A, B\}.$$

Suppose that every generated sentence consists of a single symbol. The statistical grammar is therefore completely characterized by one parameter

$$\vartheta = P(A), \qquad P(B) = 1 - \vartheta,$$

with

$$0 < \vartheta < 1.$$

The associated probability distribution is

$$p(x\,\vartheta) = \begin{cases} \vartheta, & x = A, \\ 1 - \vartheta, & x = B. \end{cases}$$

4.1.1 Information Geometry

The Fisher metric is

$$g(\vartheta) = \frac{1}{\vartheta(1-\vartheta)}. \tag{11}$$

Hence the statistical line element becomes

$$ds^2 = \frac{d\vartheta^2}{\vartheta(1-\vartheta)}. \tag{12}$$

The manifold is one-dimensional and positively curved only through its coordinate representation.

4.1.2 Competing Productions

Assume that the grammar admits two possible production rules

$$R_1 : \vartheta \to \vartheta + \delta,$$

and

$$R_2 : \vartheta \to \vartheta + 2\delta,$$

where $\delta > 0$ is small.

The corresponding Fisher lengths are

$$L_1 = \sqrt{\frac{\delta}{\vartheta(1-\vartheta)}}, \quad (13)$$

$$L_2 = \sqrt{\frac{2\delta}{\vartheta(1-\vartheta)}}. \quad (14)$$

Since

$$L_1 < L_2,$$

The Information-Geometric Grammar Principle predicts that R_1 is the preferred grammatical transformation because it minimizes the information-geometric action.

Notice that this conclusion does not depend on any externally assigned utility or fitness function. Selection follows solely from the intrinsic geometry of the probability manifold.

4.1.3 Geodesic Evolution

The action is

$$\mathcal{A} = \int \sqrt{g(\vartheta)}\,|\dot{\vartheta}|\,dt.$$

Introducing the coordinate transformation

$$u(\vartheta) = 2 \arcsin \sqrt{\vartheta},$$

one finds

$$ds^2 = du^2.$$

Thus geodesics are simply straight lines,

$$u(t) = u_0 + vt,$$

which correspond to

$$\vartheta(t) = \sin^2\left(\frac{u_0 + vt}{2}\right). \qquad (15)$$

The evolution of the grammar is therefore completely determined by the geometry of the Fisher manifold.

4.2 Example II: Adaptive-Agent Grammar

We now consider a minimal Complex Adaptive System.

Suppose two interacting agents

$A_1, \quad A_2$

may each choose one of two actions,

C (cooperate), or

D (defect).

The alphabet of the grammar is therefore

$\Sigma = \{C, D\},$

while the admissible strings are

$CC, \; CD, \; DC, \; DD.$

These four strings represent the collective system states.

4.2.1 Statistical Grammar

Assume probabilities

$$p_i(\vartheta), \qquad i = 1, \ldots, 4,$$

parameterized by

$$\vartheta = (\vartheta^1, \vartheta^2).$$

The statistical manifold is now two-dimensional. Its Fisher metric is

$$g_{ij} = \sum_{k=1}^{4} p_k \, \partial_i \ln p_k \, \partial_j \ln p_k.$$

4.2.2 Competing Adaptive Rules

Suppose two admissible interaction rules exist.

Rule R_A favors cooperation,

Rule R_B favors defection,

$CD \to CC.$

$CC \to DD.$

Each rule induces a displacement

$\Delta\theta_A, \qquad \Delta\theta_B.$

Their Fisher lengths are

$L^2 = g_{ij}\Delta\theta^i \ \Delta\theta^j ,$

And

$A \ A \ A$

$L^2 = g_{ij}\Delta\theta^i \ \Delta\theta^j .$

Suppose that

$B \ B \ B$

$L_A < L_B.$

Then IGGP predicts that cooperative adaptation is selected because it corresponds to the shorter geodesic displacement on the statistical manifold.

The decision is therefore geometric rather than algorithmic.

4.2.3 Adaptive Interpretation

Each grammatical production changes the probability distribution of the collective system.

Consequently,

$$P_t \to P_{t+1}$$

becomes a trajectory on the statistical manifold.

The Fisher metric measures the distinguishability between successive col-lective configurations, while the geodesic equations determine the most sta-tistically economical adaptive path,

$$\frac{d^2\vartheta^k}{dt^2} + \Gamma^k_{ij}\frac{d\vartheta^i}{dt}\frac{d\vartheta^j}{dt} = 0.$$

Learning, adaptation, and self-organization therefore become geometric flows.

4.3 Comparison with Classical Grammar

Traditional grammars specify only whether a production is legal.

The Information-Geometric Grammar Principle additionally specifies *which* legal production is dynamically preferred.

Thus the grammar acquires three distinct layers:

1. syntactic admissibility;
2. probabilistic weighting;
3. geometric selection.

The first determines the space of possible productions. The second as-signs probabilities to these productions. The third selects the evolution by extremizing the information-geometric action.

4.4 Discussion

The two examples illustrate complementary aspects of the Information- Ge-ometric Grammar Principle.

The Bernoulli grammar demonstrates that geodesics can be computed analytically and that grammatical evolution follows paths of minimum sta-tistical distance.

The adaptive-agent grammar shows that exactly the same geometric prin-ciple extends naturally to Complex Adaptive Systems. In this setting the grammar alphabet consists of interacting agents, grammar productions rep-resent possible interactions, and adaptive evolution corresponds to geodesic motion on the manifold of collective probability distributions.

These examples suggest that IGGP provides a universal geometric selec-tion principle applicable across symbolic systems, adaptive networks, evolu-tionary dynamics, and statistical inference.

5 General Properties of the Information-Geometric Grammar Principle

Having illustrated the Information-Geometric Grammar Principle through simple examples, we now establish several of its general mathematical prop-erties. These results demonstrate that the principle is independent of the particular choice of coordinates or grammar representation and therefore constitutes an intrinsic geometric law.

5.1 Coordinate Invariance

A statistical manifold admits many possible parameterizations. Suppose

$$\vartheta \longrightarrow \phi(\vartheta)$$

is a smooth invertible coordinate transformation. The Fisher metric transforms according to

$$g'_{ab} = \frac{\partial \vartheta^i}{\partial \phi^a} \frac{\partial \vartheta^j}{\partial \phi^b} g_{ij} \tag{16}$$

so that the line element satisfies

$$ds^2 = g_{ij} d\vartheta^i d\vartheta^j = g'_{ab} d\phi^a d\phi^b. \tag{17}$$

Consequently, the information-geometric action

$$\mathcal{A} = \int ds$$

is invariant under every smooth reparameterization.

The Information-Geometric Grammar Principle therefore depends only on the underlying statistical geometry and not on the coordinates used to describe the grammar.

5.2 Selection by Minimum Statistical Distance

Consider several admissible productions

$$R_1, \ldots, R_n,$$

originating from the same grammatical state.

Each production induces an infinitesimal displacement

$$\Delta\vartheta_i.$$

Their Fisher lengths are

$$L_i^2 = g_{jk} \Delta\vartheta_i^j \Delta\vartheta_i^k.$$

The Information-Geometric Grammar Principle predicts

$$R_{\text{selected}} = \arg\min_i L_i.$$

Thus grammatical evolution is determined locally by the geometry of the statistical manifold.

5.3 Theorem (Local Optimality)

Theorem. Assume that the Fisher metric is positive definite. Then every sufficiently short geodesic minimizes the information-geometric action among all neighboring admissible grammatical paths having the same endpoints.

Proof.

The Fisher metric endows the statistical manifold with a Riemannian structure. Classical results from Riemannian geometry imply that geodesics are locally length minimizing.

Since the IGGP identifies grammatical evolution with stationary curves of every sufficiently short geodesic possesses minimum Fisher length. Therefore the selected grammatical production is locally optimal in the information-geometric sense.

$$A = \int ds,$$

5.4 Uniqueness

If the statistical manifold is geodesically convex, every pair of sufficiently close grammatical states is connected by a unique geodesic.

Hence the Information-Geometric Grammar Principle assigns a unique preferred production whenever the manifold contains no conjugate points.

Non-uniqueness may occur only when multiple geodesics possess identical Fisher length, analogous to degeneracies in classical variational problems.

5.5 Relation with Information Theory

The Fisher metric is the second-order approximation of several statistical divergences.

For nearby probability distributions,

$$D_{\mathrm{KL}}(p(\vartheta)\|p(\vartheta + d\vartheta)) = \frac{1}{2} g_{ij} d\vartheta^i d\vartheta^j + O(d\vartheta^3), \qquad (18)$$

where D_{KL} denotes the Kullback–Leibler divergence.

Consequently, minimizing Fisher length is equivalent, to second order, to minimizing statistical distinguishability.

The Information-Geometric Grammar Principle therefore selects the ad-missible grammatical production producing the smallest information gain between successive grammatical states.

5.6 Emergent Complexity

The Information-Geometric Grammar Principle introduces a hierarchy of descriptions.

At the syntactic level, production rules determine the admissible symbolic transformations.

At the probabilistic level, these productions generate probability distri-butions over grammatical strings.

At the geometric level, the Fisher metric induces a statistical manifold whose geodesics determine preferred grammatical evolution.

Complex adaptive behavior emerges naturally from the interaction of these three layers.

The geometry is therefore not an auxiliary mathematical structure but an emergent consequence of probabilistic grammar itself.

5.7 Toward Universal Information-Geometric Gram-mars

The preceding results suggest that grammars may be classified according to the geometry of their associated statistical manifolds rather than solely by their syntactic production rules.

Two grammars possessing different symbolic representations but inducing isometric Fisher manifolds exhibit identical information-geometric dynamics. Conversely, grammars with distinct Fisher geometries generate fundamen-tally different evolutionary behaviors, even when their syntax is similar.

This observation motivates the concept of an *information-geometric equiv-alence class* of grammars, in which the essential dynamical properties are determined by intrinsic geometry rather than symbolic realization.

We have demonstrated that IGGP is:

- coordinate invariant;
- intrinsically geometric;
- locally optimal;
- related to KL divergence;
- capable of classifying grammars via their induced Fisher geometry.

6 Applications and Connections to Complex Adaptive Systems

The Information-Geometric Grammar Principle (IGGP) was introduced as a geometric selection principle for probabilistic grammars. Although moti-vated within formal language theory, its scope extends well beyond symbolic systems. The central idea—that admissible state transitions are selected by the geometry of the underlying statistical manifold—appears naturally in a wide variety of adaptive, physical, and computational systems.

6.1 Complex Adaptive Systems

A Complex Adaptive System (CAS) consists of many interacting agents whose local interactions give rise to emergent collective behavior.

Within the present framework, the correspondence is immediate:

Grammar alphabet	Agents
Production rules	Agent interactions
Sentences	Collective configurations
Probability distribution	Distribution over configurations
Fisher metric	Geometry of adaptation
Geodesics	Preferred adaptive trajectories

Each interaction among agents changes the probability distribution over collective states, thereby generating a trajectory on the statistical manifold. The Fisher metric measures the statistical distinguishability between suc-cessive configurations, while the IGGP selects the evolution requiring the smallest information-geometric action.

Adaptation therefore emerges as a geometric phenomenon rather than as an externally imposed optimization process.

6.2 Learning Systems

Machine learning algorithms iteratively modify model parameters in response to incoming information. In many modern optimization methods, parameter updates already possess a geometric interpretation through natural-gradient methods.

Within the IGGP framework, learning becomes a grammatical process. Each parameter update corresponds to a production rule acting on the gram-mar of possible hypotheses, while the Fisher information defines the natural geometry of the hypothesis space.

Consequently, learning trajectories may be interpreted as geodesics on the statistical manifold, providing a geometric criterion for selecting among multiple admissible updates.

6.3 Bayesian Inference

Bayesian inference transforms prior distributions into posterior distributions according to Bayes' theorem.

From the viewpoint of IGGP,

$$P_{\text{prior}} \to P_{\text{posterior}}$$

is simply another grammatical production acting on probability distribu-tions.

Successive Bayesian updates therefore define a trajectory on the statistical manifold, whose geometry is again determined by the Fisher metric. The information-geometric viewpoint suggests interpreting Bayesian inference as an evolutionary grammar in which evidence acts as a production operator.

6.4 Evolutionary Dynamics

Evolutionary systems continually modify the frequencies of competing species, strategies, or genotypes.

If $p_i(t)$ denotes the frequency of the *i*-th type, then the evolving population itself forms a probability distribution.

The Information-Geometric Grammar Principle interprets evolutionary change as a sequence of grammatical productions acting on this distribu-tion. Competition among admissible evolutionary pathways is resolved by the geometry of the Fisher manifold rather than by symbolic rules alone.

This viewpoint naturally complements information-geometric formula-tions of replicator dynamics and evolutionary game theory.

6.5 Statistical Physics

Statistical mechanics associates every equilibrium state with a probability distribution over microscopic configurations.

$$p_i = \frac{e^{-\beta E_i}}{Z},$$

The Gibbs distribution, defines a statistical manifold parameterized by thermodynamic variables such as temperature, chemical potential, or external fields.

Within the present framework, thermodynamic transformations corre-spond to grammatical productions acting on probability distributions. Near equilibrium, the Fisher metric is directly related to equilibrium fluctuations, implying that thermodynamic evolution may also be interpreted as geodesic motion on an information manifold.

Thus IGGP naturally complements information-geometric approaches to thermodynamics without assuming any specific microscopic dynamics.

6.6 Networks and Collective Intelligence

Many adaptive systems are naturally represented as networks whose topology changes in response to interactions among their nodes.

Examples include social networks, communication systems, neural net-works, and biological regulatory networks.

Each rewiring event changes the probability distribution over network states. The resulting evolution generates a trajectory on the associated sta-tistical manifold.

The Information-Geometric Grammar Principle predicts that the pre-ferred network evolution minimizes information-geometric action, providing a geometric criterion for adaptive rewiring and self-organization.

6.7 Toward a Universal Grammar of Adaptation

The preceding examples suggest a common conceptual structure.

Regardless of the specific discipline, one identifies:

1. A set of admissible states;
2. Rules generating transitions among those states;
3. Probability distributions over the admissible states;
4. The Fisher geometry induced by those distributions;
5. Evolution governed by geodesics of the statistical manifold.

This hierarchy is remarkably independent of the underlying physical, bi-ological, or computational substrate.

Consequently, the Information-Geometric Grammar Principle may be viewed as a candidate universal principle of adaptive organization, replac-ing purely syntactic descriptions with an intrinsic geometric dynamics.

6.8 Limitations and Future Directions

The present work establishes the geometric foundations of the Information-Geometric Grammar Principle but leaves several important questions open. First, the examples

considered here involve finite-dimensional statisti-
cal manifolds. Extensions to infinite-dimensional manifolds associated with stochastic processes, quantum field theories, or continuous grammars remain to be developed.

Second, the present framework assumes that the Fisher metric adequately captures statistical distinguishability. More general information metrics, in-cluding quantum monotone metrics and Wasserstein-type geometries, may lead to alternative grammatical dynamics.

Finally, practical implementations in adaptive networks, reinforcement learning, evolutionary biology, and distributed artificial intelligence will re-quire efficient numerical algorithms for computing geodesics on large statis-tical manifolds.

These directions suggest that the Information-Geometric Grammar Prin-ciple is best viewed not as a completed theory but as the foundation of a broader research program linking formal grammars, information geometry, and complex adaptive systems.

We remark that this section significantly broadened our scope by demon-strating that the proposed framework applies to:

- Complex Adaptive Systems,
- machine learning and natural-gradient optimization,
- Bayesian inference,
- evolutionary dynamics,
- statistical physics, adaptive and evolving networks.

7 CONCLUSIONS AND OUTLOOK

This work has introduced the *Information-Geometric Grammar Principle* (IGGP), a new variational principle that unifies probabilistic grammars, in-formation geometry, and adaptive dynamics within a single mathematical framework.

The starting point was the observation that every probabilistic gram-mar naturally generates a family of probability distributions. Endowing this family with the Fisher information metric transforms the grammar into a statistical manifold, thereby providing an intrinsic geometric representation of symbolic evolution. Within this setting, grammatical productions cease to be purely combinatorial operations and become geometric displacements on an information manifold.

The central postulate of the present work was that admissible gram-matical transformations are selected according to an information-geometric variational principle. Among all syntactically permissible productions, the realized evolution is the one that extremizes the Fisher information action. In the absence of external constraints, this principle reduces grammatical evolution to geodesic motion on the statistical manifold. Our treatment sug-gests that geometry, rather than syntax alone, provides the fundamental organizing principle governing probabilistic grammars.

REFERENCES

1. J. H. Holland, *Adaptation in Natural and Artificial Systems*, University of Michigan Press (1975).
2. J. H. Holland, Complex Adaptive Systems, *Daedalus* **121**, 17–30 (1992).
3. J. H. Holland, *Signals and Boundaries: Building Blocks for Complex Adaptive Systems*, MIT Press (2012).
4. M. Mitchell, *Complexity: A Guided Tour*, Oxford University Press (2009).
5. N. Chomsky, *Syntactic Structures*, Mouton, The Hague (1957).
6. C. R. Rao, Information and the accuracy attainable in the estimation of statistical parameters, *Bull. Calcutta Math. Soc.* **37**, 81–91 (1945).
7. S. Amari and H. Nagaoka, *Methods of Information Geometry*, American Mathematical Society and Oxford University Press (2000).
8. N. Ay, J. Jost, H. Vàn Lê, and L. Schwachhöfer, *Information Geometry*, Springer (2017).
9. T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed., Wiley (2006).
10. Caticha, *Entropic Inference and the Foundations of Physics*, Brazilian Center for Physics Research (2012).
11. J. H. Holland, *Hidden Order: How Adaptation Builds Complexity* , Basic Books, 1995 M08 21.
12. Angelo Plastino, Research Gate Preprints DOI 10.0163/1. 3456789 and 10.0163/1. 4567891.