Epistemic Geometry: Modeling Qualia

Robert C Dennis

Abstract

The quantitative modeling of agent qualia remains an open challenge. Here we describe the construction of a multi-agent model which manifestly takes into account cognition and agent communication at the outset. Agents are constructed as pairs of local sections of associated bundles $\mathcal{E}_i$ (where i= belief, model) to a principal G-bundle $\mathcal{N}$ composed of a base manifold $\mathcal{C}$ and Lie group G. Agents interact via induced connections and evolve according to a generalized variational energy. We demonstrate that a $SL(2, \mathbb{R})$ gauge frame acts as an agent "center" where communication between agents is mediated by gauge connections. We treat agent evolution under a dynamical variational generalized energy proffered by a "free energy" and "connection energies". This structure allows agents to maintain coherence in perception and inference, despite fundamentally local and subjective perspectives. Language, in this view, is a gauge theory and agentic perspectives can be both unique and communicable. Importantly, we describe how agent "qualia" can be studied as the pullback of informational geometric quantities from an agent's gauge frame to the base manifold. We further describe, then, a cognition-first approach towards a unified physics.

1 Introduction - From Observation to Ontology

1.1 The Geometric Legacy

"Ubi materia, ibi geometria" - Kepler

Geometry has proven to be a powerful tool which humans have leveraged to develop myriad theories and models of the natural world throughout our history. As exemplified by Kepler, above, geometry is thought to play a central role in the structure, composition, and dynamics of our perceptual reality. From Euclid to Newton, Maxwell to Einstein, Grothendieck to Witten, and many more, geometry has proven reliable in advancing new ideas and illuminating new paths for scientific inquiry.

The geometric study of the foundations of physics accelerated during the 20^{th} century and continues to this day. Researchers have discovered that the organizing principles of the natural world are best expressed in the language of geometry and symmetry - especially continuous symmetries described by Lie groups, and geometry (fiber bundles, manifolds, etc) and geometric objects such as spinors, vectors, and tensors.

Despite the philosophical challenges in studying the foundational elements of reality from an "external" observer point of view, science has progressed steadily by developing powerful and accurate theories based upon the assumptions that space and time (among other dimensionful quantities) are fundamental aspects of our universe[ˆ4]. Einstein and Minkowski (among others) unified space and time into a 3+1 dimensional smooth manifold wherein its geometry is dynamic and malleable. Quantum Field Theory (QFT) further built upon this scaffold of Minkowski spacetime by extending quantum theories to compatibility with Poincare/Lorentz symmetry leading to some of the most elegant, awesome, and precise predictions and verifications humanity has ever known.

Still, deep challenges remain. The fundamental quantities of energy, mass, space, time, and more remain stubbornly devoid of firm definition. Furthermore, the "unreasonable effectiveness of mathematics in the natural sciences" remains a profound but poorly understood property of our universe. We currently do not have a proto-geometry of physics and the natural world aside from candidate theories such as string theory, loop-quantum gravity, geometric unity, and others. Still, these frameworks lack firm ontological definitions of the dimensionful quantities they seek to describe.

1.2 The Cognitive Revolution

Separately from, though partly inspired by, physics neuroscientists, computer scientists, and cognitive researchers have likewise leveraged geometry to advance our understanding of the nature of perception and cognition; in particular, information geometry. Arguably, the focal thrust of the field of cognition is due to the pioneering work of Karl Friston who developed a variational free energy model that aims to unify the fields of cognition and behavior across many physical and cognitive scales[ˆ1].

In this view cognitive systems, themselves physical non-equilibrium open systems, perform Bayesian model selection continuously, updating beliefs in light of incoming sensory data — a process often described as predictive coding. The free energy principle unifies action and perception under the same variational imperative: agents act to minimize expected free energy — they sample the world to reduce uncertainty. This reframes action as epistemic foraging, and embodiment as a condition for inference. In this view, perception is not bottom-up sensation, but top-down model confirmation.

Explicit in these formulations is the premise that cognitive systems do not directly perceive the world's "true" states; rather, cognitive agents build models of their sensory data in such a way as to minimize their variational free energy and ultimately survive, reproduce, and evolve. Central to these studies is the geometry of information as pioneered by Amari, Rao, and Fischer (among others) and the Bayesean interpretation of probability as inferential logic championed by E.T. Jaynes in the mid to late 20th century.

In the cognitive revolution, geometry no longer applies only to physical space, but to abstract spaces of probability distributions - agents navigate abstract belief manifolds. Powerful methods and trends in deep-learning, artificial intelligence, and much more highlight the fruitful path of considering the geometry of beliefs and models as applied to physical systems.

1.3 On The Philosophy of Perception

"Thoughts without content are empty; intuitions without concepts are blind." - Kant

Physics traditionally posits an observer-independent reality governed by invariant laws. Yet, modern cognitive science tells us that observation is mediated by inference, prediction, and model-building — not direct access to a world "out there." This introduces a fundamental epistemological tension: how can physics describe the world as it is, when all measurements are filtered through subjective models of various partially independent agents? Physicists recognize these difficulties but rarely frame them as such. The foundational problem of defining "an observer" themselves embedded in a dynamic universe is a foremost challenge in physics.

This dilemma echoes the insights of Immanuel Kant, who argued in The Critique of Pure Reason that we never perceive the noumenon — the world-in-itself — but only the phenomenon, shaped by the a priori structures of human cognition. Space and time, for Kant, are not external entities "out there", but rather they are forms of intuition; cognitive scaffolds through which experience becomes possible.

Helmholtz later elaborated on Kant by considering an early version of the "Bayesean Brain Hypothesis". Helmholtz proposed that perception is an unconscious inference — that the brain is a statistical organ that constructs its experience of the world through internal hypotheses, constantly updated based on sensory input. Helmholtz's and Kant's insights pre-empted the modern view of the brain as a Bayesian inference machine.

Phenomenology, from Husserl to Merleau-Ponty, emphasized that experience is always predicated upon perspective[[x]][XX]. If space-time is predicated upon our cognitive perceptions then what does that mean for the entire edifice of physical theory? Is, perhaps, our physics secretly and fundamentally a framework of human-human communication - the study of those things which are invariant between human interpretative frames? A deep example of Chomsky's universal; grammar? Language then, in this view, is an abstract interaction which aligns frames of cognition between agents.

Physics assumes important symmetries such as Lorentz invariance, general coordinate invariance, and gauge invariance and yet subjective experiences — such as those reported under psychedelics, trauma, and mental illness - suggest deviations from standard space-time perception [ˆ4]. Such accounts are often dismissed as mere subjective hallucinations, yet modern neuroscience tells us that all perception is, in a sense, a concrete hallucination shaped by evolutionary pressures[ˆ5]. As agents, then, we dismiss communications which lead to misalignment under communication of another agent's models and beliefs. Language is the route by which models/beliefs are shared.

Furthermore, valid physical information may well be hidden within the perceptions of the pathological. We then might be more cautious in our immediate dismissal of such "observations" (especially apropos consistent overlaps and descriptions across many individuals) as shared/communicated between agents.

While Kant saw geometry as a condition of experience, Einstein viewed geometry as a property of the physical world. The curvature of space-time is not an artifact of perception, but rather a dynamical field sourced by mass and energy - poorly defined though they may be. Yet, a deep question is raised: how can a theory built upon observer-relative measurement (e.g. Lorentz invariance, coordinate choice) claim objectivity if no notion of an "observer" exists within the framework? This is much more problematic in

Quantum Theory where "observers" possess non-trivial abilities to interrupt otherwise smooth unitary evolution of quantum wavefunctions. Many famous paradoxes abound! Our most accurate and specific theory on the nature of reality is predicated upon an observer which the theory can say nothing about.

On one hand, physics excels at describing inter-subjectively verifiable reality — the world we can agree upon via communication with each other - so long as our beliefs/models don't diverge too far from the primary belief/models (a sort of epistemic democracy). On the other hand, it says almost nothing about the cognition that enables this fantastic and undervalued ability to agree! This tension, between shared reality and the mechanisms of inference, remains unresolved. Here we propose a path to reconciliation: a unified geometric model of agents, grounded in both physical and cognitive principles that may be applied to physical systems as well as abstract agentic systems such as culture, religion, government, economics, and much much more.

1.4 Towards Unification

Cognitive science and physics currently remain steadfastly separate fields of study. As physicists we may observe that when we perform experiments and record our results we are ultimately looking at the dials and knobs via electromagnetic waves and emitting sound waves to correlate to each other so as to interpret experimental confirmations between our warm, wet brains, locked inside a "dark cavern" extending electric tentacles into the myriad correlations of our universe[ˆ5]. Indeed, experiments are only "confirmed" if one agent's perception corresponds to another's via some lossless mapping (we hope!). What are we to make of physics if our basic notions of reality are themselves subjectively emergent. Here we adopt the position that any reasonable theory of physics should also be cognitively consistent and, most importantly, vice versa.

In recent decades, physicists have begun to entertain the possibility that space and time are not themselves fundamental, but emergent from som hidden layer of reality. John Wheeler famously proposed "it from bit" - the idea that the physical universe is ultimately an emergent system from an underlying web of information. More tangibly, Arkani-Hamed and collaborators introduced the amplituhedron[ˆ6], a timeless geometric object whose volumes encode scattering amplitudes in supersymmetric field theories. Meanwhile, other researchers have used tensor networks and holographic entanglement to construct emergent geometries resembling space-time apropos Einstein's general relativity[15][19][ˆ20]. Still others have found a striking similarity between thermodynamics and general relativity[ˆ17]. These results tentatively suggest the possibility of a deeper structure undergirding our verifiable reality - Kant's noumenon.

1.5 A New Framework for A Cognitive Physics

In this paper, we propose a radical inversion: that cognition is not an emergent property of a physical universe, but rather that physical law is an emergent structure within a web of cognitive inferences. We call this perspective cognitive physics. Within this paradigm, the fundamental entities are not particles and fields, but rather, they are agents, beliefs, and models — connected via the geometric machinery of inference, communication, and gauge connection.

At the heart of this theory lies a novel synergy of ideas: epistemic geometry. Here, every agent maintains a probabilistic belief about an underlying base space — a latent manifold which we interpret as "reality." These beliefs/models are not uniform: they vary across the base manifold, are subject to noise, and manifestly differ between agents. To communicate, agents must transport beliefs across distinct representational frames — necessitating a gauge-theoretic framework. From these transformations, a full-fledged field theory emerges. An agent, then, has a particular fixed gauge frame and communication is mediated by gauge transformations between agents.

Gauge connections represent epistemic compatibility and communication: how well two agents can align their perspectives/frames. Curvature then leads to a notion of semantic holonomy — an obstruction to globally consistent understandings. Belief alignment, power laws, model evolution, and phase transitions in inference are not byproducts of matter dynamics — but rather, they are consequences and properties of the communications of myriad agents under approximate steady-state non-equilibrium evolution of a single generalized variational "action" or "energy".

Surprisingly, from this basic model emerges an immediate conjectures:

Epistemic Pullback Conjecture. Traditional physical quantities (such as mass, charge, and energy) emerge as pullbacks of informational-geometric quantities (e.g. Fisher information, KL divergence) from an agent's belief and model fibers onto the latent base manifold. That is, physical law reflects the local structure of inference.

Gauge Invariance Conjecture: The gauge and coordinate invariance of physical theory is a consequence of human-human agent alignment - humans share a generative model of physics. Those properties which we agree upon are the properties we assign to our "universe".

Space-time and The Dark Universe: Space-time arises as the pull-back of high-dimensional information metrics on the belief and/or model fibers to the base manifold and can be decomposed into a 4 dimensional space-time metric and a "dark" or hidden metric. Changes in the informational fibers of an agent leads to mixing between these blocks of the decomposition. This then could be a perceptual origin of dark matter.

Science as a Meta-Agent: The institution of science is a meta-agent composed of myriad human agents across long time scales whose model is scientific theory and beliefs is independent of any individual agent. Humans "perceive" Newtonian/Euclidean space yet "Science" as an agent "perceives" Minkowski. Beliefs and models are constantly in flux

Physics as a Meta-Agent Model: Trajectories within belief/model fibers encode dynamical physical phenomena. Humans agree on a consistent physics of the universe since, a priori, our generative models agree as the result of evolutionary flow towards our current shared model. The gauge invariance of physics therefore can be ascribed to the gauge invariance of our human-scale generative models: aligned models/beliefs correspond to shared gauge frames. This behaves as a sort of anthropic principle. Shared cognitive experience, and thus shared physics, presupposes aligned generative models. If humans had fundamentally divergent priors, they would not inhabit the same apparent world. The universality of physical law thus reflects — and requires — a kind of epistemic anthropic principle. This naturally accounts for pathological hallucinations but with a stronger claim that these hallucinations (psychedelics, trauma, etc) must themselves be invariant under the gauge group.

Our paper outlines a formal theory grounded in differential geometry, information theory, cognitive science, and philosophy. It introduces a rigorous variational principle for agent interaction (as an extension of Friston's variational free energy), explores the emergence of large-scale structure via inference, and proposes a set of central conjectures that reframe the laws of physics as consequences of epistemic dynamics.

If successful, this cognitive-geometric approach could offer a unifying principle for physics, complexity, and cognition — one that treats inference not as a tool of physics, but as its very source. In this view, the universe is not a machine to be decoded, but a conversation to be understood — a dialogue between agents, inference, and the hidden structure of the noumenon.

2 Epistemic Geometry and Agent Dynamics

We now introduce the formal model underlying our cognitive physics framework. We call this model "Epistemic Gauge Theory".

We model each agent as a local pair of sections over a principal fiber bundle whose base space $\mathcal{C}$ represents a latent, generally unobservable "noumenal" manifold. Beliefs and models are encoded in associated bundles constructed from a structure group G acting on belief/recognition and model fibers ($\mathcal{B}_{\amalg}, \mathcal{B}_{\surd}$. These structures allow us to formalize both intra-agent inference and inter-agent communication using gauge-theoretic transport.

Let $\pi : \mathcal{N} \to \mathcal{C}$ be a smooth principal G-bundle where $\mathcal{C}$ is a smooth manifold, G is a Lie group acting freely and transitively on the right on $\mathcal{N}$. The projection satisfies $\pi(n \cdot g) = \pi(n)$ for all $g \in G, n \in \mathcal{N}$.

Let $\rho : G \to \mathrm{Aut}(\mathcal{B}_u)$ be a representation of the Lie group G on a smooth manifold $\mathcal{B}_u$. Depending on context, $\mathcal{B}_u$ may be modeled as a K-dimensional probability simplex Δ^K (e.g., for categorical distributions) or as a statistical manifold equipped with a suitable information geometry.

Importantly, the relevant geometric structures — such as divergence measures, metrics, and connections — remain well-defined on these fibers even when they lack linear structure (see e.g. Amari's information geometry for examples of dually flat but non-linear statistical manifolds).

Next, the associated bundles (for q and p) are defined as:

$$\mathcal{E}_u := \mathcal{N} \times_\rho \mathcal{B}_u = (\mathcal{N} \times \mathcal{B}_u)/\sim,$$

where

$$(n \cdot g, b) \sim (n, \rho(g)b).$$

This gives a fiber bundle $\pi_{\mathcal{E}_u} : \mathcal{E}_u \to \mathcal{C}$ with fiber $\mathcal{B}_u$. Generally speaking the model and belief fibers could potentially carry different representations of G. For simplicity we consider both to carry a dimension K and the corresponding suitable representations of G.

Definition: An agent is a pair of local sections $\mathcal{A}^i = (\sigma_q^i(c), \sigma_p^i(c)) = (q_i(c), p_i(c))$ over $\mathcal{C}$

$$\sigma_q^i : \mathcal{U}_i \subset \mathcal{C} \to \mathcal{B}_q,$$

$$\sigma_p^i : \mathcal{U}_i \subset \mathcal{C} \to \mathcal{B}_p.$$

Again, in full generality the two associated bundles could correspond to different subsets of the base manifold but for simplicity we consider both subsets to be identical.

Definition: A multi-agent ($\mathcal{M}$) over $\mathcal{C}$ is a tuple of agents (where $\mathcal{I}$ is an index set)

$$\mathcal{M} = \{A^i = (\sigma_q^i(c), \sigma_p^i(c))\}_{i \in \mathcal{I}}.$$

Definition: A meta-agent is a multi-agent whose agents share the same generative models. Alternatively, a Quasi-meta-agent is a multi-agent whose agents share the same beliefs but who do not share the same models. Furthermore, we say that a set of agents is "epistemically dead" if they identically share both beliefs and models. Note, importantly that even if the agents composing a meta-agent are epistemically dead that this does NOT imply the meta-agent is epistemically dead. In fact, this is a feature of our framework - epistemically dead agents can be integrated out!

Next, we define an observation/observable of an agent as a local section of $\mathcal{B}_q$

$$\mathcal{O}_q : \mathcal{U} \subset \mathcal{C} \to \mathcal{B}_q.$$

This then is what an agent "believes" they observed.

Alternatively we could define an observation to be within the model fiber. The definition then is context dependent. Relative to an agent we can either subjectively observe or consider the observation as raw data provided by "the world".

This is similar to an agent but does not necessarily require a section over the second associated bundle fiber - although we shall see later that observations can (and must) be described as agents themselves: lower-scale agents to a higher-scale meta-agent. Our definition above is then just a useful proxy for a much deeper situation/relation among many multi-scale agents.

Next, in the standard way[ˆ34], via horizontal lifting from $\mathcal{N}$ to $\mathcal{E}_i$ we have a variety of morphisms and induced connections across scales (i, j):

1. $\Omega^i : \Gamma(\mathcal{B}_q) \to \Gamma(\mathcal{B}_q)$
2. $\tilde{\Omega}^i : \Gamma(\mathcal{B}_p) \to \Gamma(\mathcal{B}_p)$
3. $\Lambda^i_j : \Gamma(\mathcal{B}_q) \to \Gamma(\mathcal{B}_q)$
4. $\tilde{\Lambda}^i_j : \Gamma(\mathcal{B}_p) \to \Gamma(\mathcal{B}_p)$
5. $\Theta^i_j : \Gamma(\mathcal{B}_p) \to \Gamma(\mathcal{B}_q)$
6. $\tilde{\Theta}^i_j : \Gamma(\mathcal{B}_q) \to \Gamma(\mathcal{B}_p)$
7. $\Phi : \mathcal{E}_p \to \mathcal{E}_q$
8. $\tilde{\Phi} : \mathcal{E}_q \to \mathcal{E}_p$

where $\Gamma(\mathcal{B}_i)$ denotes the space of smooth sections of $\mathcal{B}_i$ over the relevant open subset of the noumenal manifold $\mathcal{C}$

The two principal bundle morphisms of special interest are $\Phi : \mathcal{E}_p \to \mathcal{E}_q$ and $\tilde{\Phi} : \mathcal{E}_q \to \mathcal{E}_p$. When these diagrams commute (i.e., $\pi_q \circ \Phi = \pi_p \circ \tilde{\Phi}$), we interpret this as a condition for epistemic agreement — an experimentally verified prediction, where an agent's recognition aligns with its generative model across fibers.

In general, however, these bundle morphisms need not commute. In the simulations and examples to follow we will consider these bundle morphisms to be the identity - i.e. both model and recognition fibers will be K-dimensional spaces.

2.1 Types of Parallel Transport in Epistemic Geometry

In our framework, parallel transport arises in several distinct but interconnected settings, depending on whether the transport occurs along the base manifold $\mathcal{C}$, within or between fibers, or across agent frames. We distinguish the following cases:

2.1.1 Horizontal Transport Along the Base Manifold $\mathcal{C}$

Let $\pi : \mathcal{N} \to \mathcal{C}$ be a principal G-bundle, and let $\mathcal{E} = \mathcal{N} \times_\rho \mathcal{B}$ be an associated bundle with fiber $\mathcal{B}$. Given a path $\gamma : [0,1] \to \mathcal{C}$, a connection 1-form $A \in \Omega^1(\mathcal{C}, \mathfrak{g})$ defines a notion of parallel transport along γ via the path-ordered exponential:

$$T_\gamma = \mathcal{P} \exp \left(- \int_\gamma A_\mu(c)\, dc^\mu \right) \in G. \tag{1}$$

This operator maps fiber elements between different base points:

$$b(c_0) \in \mathcal{B}_{c_0} \quad \mapsto \quad b(c_1) = \rho(T_\gamma) \cdot b(c_0) \in \mathcal{B}_{c_1}. \tag{2}$$

This is the canonical notion of \emph{parallel transport along the base}, lifted via the connection to fibers.

2.1.2 2. Vertical Transport Within a Single Fiber $\mathcal{B}_c$

At a fixed point $c \in \mathcal{C}$, the fiber $\mathcal{B}_c$ is a manifold in its own right. In our setting, this fiber typically represents a statistical manifold (e.g., probability simplex Δ_K, or a Gaussian family), equipped with an intrinsic Riemannian or dual-affine geometry.

Parallel transport within $\mathcal{B}_c$ may occur along a curve $\eta(\tau) \subset \mathcal{B}_c$ with tangent vector $\dot{\eta}(\tau)$ governed by a connection ∇ intrinsic to $\mathcal{B}$:

$$\nabla_{\dot{\eta}} V = 0, \quad \text{for parallel vector field } V(\tau). \tag{3}$$

In the information geometry setting (Amari), $\mathcal{B}_c$ may carry a pair of dual connections $(\nabla^{(e)}, \nabla^{(m)})$ associated with exponential and mixture families. Curvature in this fiber space is defined via the Riemann tensor $R^{\mathcal{B}}$:

$$R^{\mathcal{B}}(X,Y)Z := \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z. \tag{4}$$

Thus, even at a single point $c \in \mathcal{C}$, the fiber $\mathcal{B}_c$ may be a curved internal space with nontrivial geodesics and holonomies. These paths represent purely epistemic transformations of beliefs or models that do not involve movement along $\mathcal{C}$.

2.1.3 3. Intra-Agent Spatial Transport (Within Fiber Bundle)

To compare beliefs $q(c_1)$ and $q(c_2)$ held by the same agent over two nearby points $c_1, c_2 \in \mathcal{U}_i \subset \mathcal{C}$, we apply the agent's gauge field $\phi_i(c)$ to define a connection $A_\mu^{(i)} = \partial_\mu \phi_i(c)$ and use:

$$T_{c_1 \to c_2}^{(i)} := \mathcal{P} \exp\left(-\int_{c_1}^{c_2} A_\mu^{(i)}\, dc^\mu\right). \tag{5}$$

This defines gauge-covariant parallel transport between fibers over space, for a fixed agent.

2.1.4 4. Inter-Agent Frame Transport (At a Shared Point)

When two agents $\mathcal{A}_i$ and $\mathcal{A}_j$ overlap at a point $c \in \mathcal{U}_i \cap \mathcal{U}_j$, each has its own gauge frame $\phi_i(c), \phi_j(c) \in \mathfrak{g}$. The inter-agent gauge transformation is given by:

$$\Omega_{ij}(c) := \exp(\phi_i(c)) \cdot \exp(-\phi_j(c)) \in G, \tag{6}$$

and transports beliefs or models from agent j's frame to agent i's:

$$q_j(c) \mapsto q_i^{(j)}(c) := \rho(\Omega_{ij}(c)) \cdot q_j(c). \tag{7}$$

This is essential for defining KL alignment between agents:

$$D_{\mathrm{KL}}\left[q_i(c) \,\|\, \Omega_{ij}(c) \cdot q_j(c)\right]. \tag{8}$$

2.1.5 5. Composite Transport and Holonomy

More generally, a belief may be transported along a path $\gamma = [(i_1, c_1) \to \cdots \to (i_n, c_n)]$ consisting of both spatial transport within agents and inter-agent frame shifts. The total parallel transport operator is:

$$P_\gamma = T_{c_1 \leftarrow c_n}^{(i_n)} \cdot \Omega_{i_n i_{n-1}} \cdots \Omega_{i_2 i_1} \cdot T_{c_2 \leftarrow c_1}^{(i_1)}, \tag{9}$$

and defines the holonomy:

$$\mathrm{Hol}_\gamma(q) := P_\gamma \cdot q. \tag{10}$$

If $\mathrm{Hol}_\gamma(q) \neq q$, the loop encloses nontrivial epistemic curvature.

2.1.6 Summary

Transport Type	**Domain**	**Operator**	**Purpose**
Horizontal (base)	$c_1 \to c_2 \in \mathcal{C}$	$T_\gamma = \mathcal{P}\exp(-\int A)$	Compare beliefs across space
Vertical (fiber)	$\eta(\tau) \subset \mathcal{B}_c$	∇ or $\rho(\exp(\phi))$	Transform within belief space
Intra-agent	$q_i(c_1) \to q_i(c_2)$	$T^{(i)}$ from $A_\mu^{(i)}$	Spatial update in same agent
Inter-agent	$q_j(c) \mapsto q_i^{(j)}(c)$	$\Omega_{ij}(c)$	Frame alignment across agents
Composite	$\gamma : (i_1, c_1) \to \cdots$	P_γ	Meta-agent transport and holonomy

Table 1: Types of parallel transport in epistemic gauge geometry.

2.2 Generalized Variational Energy

Following Friston[ˆ1] we define a variational free energy as

$$\mathcal{F}[q] = D_{KL}[\ q(c) \mid \Phi p(c)\ ] - \mathbb{E}_{q(c)}[\log p(o|c)].$$

For simplicity, in this study, we will not consider the "active" inference term - we shall concern ourselves with how agents passively condense into meta-agents and quasi-meta agents.

We extend this principle to agents who may share information/beliefs/models via parallel transport. We define a generalized variational energy for an agent A as:

$$\mathscr{V}_A = D_{KL}[\ q_A(c)\ ||\ \Phi p_A(c)\ ] - \mathbb{E}_{q_A}[\log p_A(o|c)] + \sum_i \mathscr{V}_{\Lambda_i},$$

where $\mathscr{V}_{\Lambda_i}$ represents possible interactions between other agents and (quasi)-meta-agents mediated by the various gauge connections above.

In our current considerations we will only focus on the induced connections Ω and $\tilde{\Omega}$ such that agents can compare beliefs and models between different sections of $\mathcal{B}_q$ and $\mathcal{B}_p$ where

$$\Omega : \Gamma(\mathcal{B}_q) \to \Gamma(\mathcal{B}_q).$$

$$\tilde{\Omega} : \Gamma(\mathcal{B}_p) \to \Gamma(\mathcal{B}_p)$$

The induced connection provides a parallel transport of beliefs/models which in turn defines the coordination structure used in our generalized variational energy term. This construction is well defined even when the fibers are not vector spaces so long as they carry a smooth G-action[ˆ34].

Therefore, we have

$$\mathscr{V}_T[q_i(c)] = \alpha D_{\mathrm{KL}}[(q_i(c)|\Phi p_i(c))]$$

$$+\beta \sum_j D_{\mathrm{KL}}[q_i(c)|\Omega_{ij} q_j(c)]$$

$$+\tilde{\beta} \sum_j D_{\mathrm{KL}}[p_i(c)|\tilde{\Omega}_{ij} p_j(c)]$$

$$+\tilde{\alpha} D_{\mathrm{KL}}[(p_i(c)|\tilde{\Phi} q_i(c))]$$

where $\alpha, \beta, \tilde{\alpha}, \tilde{\beta} \in \mathbb{R}$ represents general couplings/parameters. We note in passing that this expression bears remarkable similarities to the Grand potential $\Upsilon = U - TS + \mu N$ in standard thermodynamics (where μN is analogous to the $\Omega, \tilde{\Omega}$ terms).

2.3 Gauge Transport via Lie Algebra Frames

The inter-agent transport operators Ω_{ij} and $\tilde{\Omega}_{ij}$ serve to compare beliefs and models, respectively, between agents i and j. These operators are defined pointwise over the base manifold $\mathcal{C}$, and are constructed from local gauge frames associated with each agent.

Let $\phi_i(c) \in \mathfrak{g}$ be a smooth field over $\mathcal{C}$ valued in the Lie algebra $\mathfrak{g} = \mathrm{Lie}(G)$, representing the gauge frame of agent i. Then, for each pair of agents (i, j), we define the induced transport operator as:

$$\Omega_{ij}(c) := \exp\left(\phi_i(c)\right) \cdot \exp\left(-\phi_j(c)\right),$$

where the exponential map $\exp : \mathfrak{g} \to G$ maps Lie algebra elements to the Lie group G, and the product is taken in the group.

This construction defines a group-valued map $\Omega_{ij}(c) \in G$, which acts on the belief fiber $\mathcal{B}_q$ via the representation ρ. That is, for any section $q_j(c) \in \Gamma(\mathcal{B}_q)$, the transported belief is:

$$\Omega_{ij}(c) \cdot q_j(c) := \rho\left(\Omega_{ij}(c)\right) \cdot q_j(c).$$

Similarly, for the model fibers, we define:

$$\tilde{\Omega}_{ij}(c) := \exp\left(\tilde{\phi}_i(c)\right) \cdot \exp\left(-\tilde{\phi}_j(c)\right),$$

with $\tilde{\phi}_i(c) \in \tilde{\mathfrak{g}} = \mathrm{Lie}(\tilde{G})$, acting on $\mathcal{B}_p$ via the corresponding representation $\tilde{\rho}$.

These induced transport operators implement a form of gauge-covariant parallel transport between agents. The field $\phi_i(c)$ serves as the agent's gauge frame, and differences in these frames determine the epistemic misalignment between agents.

By construction, $\Omega_{ii} = \mathrm{id}$, and in general $\Omega_{ij} \neq \Omega_{ji}^{-1}$ unless the gauge fields are flat. This asymmetry encodes epistemic curvature, which we will explore in subsequent sections.

In our framework, each agent maintains a fixed gauge frame $\phi_i(c)$, which defines their local epistemic perspective. This frame serves as the reference against which other agents' beliefs are compared. The term $D_{\mathrm{KL}}[\, q_i(c) \,\|\, \Omega_{ij}(c) \cdot q_j(c)\,]$ then quantifies the epistemic misalignment between agent i and agent j, measured in agent i's own frame. Operationally, this can be interpreted as a communicative act: agent i attempts to transform agent j's beliefs into its own representational basis and evaluate their coherence.

Intuitively, the gauge frame $\phi_i(c)$ can be interpreted as a geometric encoding of the agent's cognitive identity — its internally consistent perspective on the latent manifold $\mathcal{C}$. In this sense, the gauge frame acts as a formal proxy for an agent's "selfhood" or consciousness, organizing all incoming beliefs and observations relative to a quasi-static internal model.

This interpretation opens the door to a rich epistemic ontology: communication becomes a gauge-theoretic alignment of perspectives, and disagreement (or confusion) corresponds to curvature in the "semantic" field. Agents with similar gauge frames can align beliefs easily, while those with strongly divergent frames experience epistemic dissonance.

Additionally, for future exploration, human language itself may be considered as a gauge theory. Each speaker maintains a local linguistic frame — a grammar, lexicon, and semantic structure — and successful communication requires parallel transport of meaning across these frames. Misunderstandings arise when the linguistic gauge transformations between individuals are ill-defined or insufficient to fully align perspectives. In this view, syntax encodes local structural constraints, semantics provides the fiber content, and discourse becomes a path-dependent operation on shared meaning spaces. The geometric framework developed here may thus offer a natural foundation for modeling language as a structured system of epistemic transport.

To further analyze the geometric structure of inter-agent alignment, we define the gauge field associated with an agent's gauge frame $\phi(c) \in \mathfrak{g}$, where $\mathfrak{g} = \mathrm{Lie}(G)$ is the Lie algebra of the structure group G. In analogy with classical gauge theory, we define the gauge potential (or connection form) as the derivative of the gauge frame:

$$A_\mu(c) := \partial_\mu \phi(c) \in \mathfrak{g},$$

where μ indexes coordinates on the base manifold $\mathcal{C}$, and A_μ is a $\mathfrak{g}$-valued one-form over $\mathcal{C}$. This connection governs infinitesimal parallel transport within an agent's belief fiber.

We then define the curvature — or field strength — of this connection as the standard nonabelian gauge curvature tensor:

$$F_{\mu\nu}(c) := \partial_\mu A_\nu(c) - \partial_\nu A_\mu(c) + [A_\mu(c), A_\nu(c)] \in \mathfrak{g}.$$

The curvature $F_{\mu\nu}$ measures the failure of the gauge field to be locally flat — that is, the extent to which infinitesimal transport around an infinitesimal loop in $\mathcal{C}$ depends on the path taken. In the context of epistemic geometry, nonzero curvature corresponds to epistemic inconsistency or semantic holonomy: agents attempting to align beliefs through sequences of communication may find that their beliefs do not return to their original form after a closed interaction loop.

This gives rise to a concrete interpretation of gauge curvature as a measure of cognitive or semantic torsion — the topological obstruction to global consensus. As such, the norm $\|F_{\mu\nu}(c)\|^2$ might be added to the variational energy to regularize inference processes and penalize incoherent belief cycles across agents.

2.4 Transport, Holonomy, and Path Dependence

In our framework, gauge transport occurs in two distinct forms: intra-agent and inter-agent transport. Though both rely on the structure of the underlying principal bundle and its associated gauge fields, they serve different roles and operate over different domains.

2.4.1 Intra-Agent Transport.

Each agent A_i defines two local sections over a subset $\mathcal{U}_i \subset \mathcal{C}$: one for beliefs $\sigma_q^i : \mathcal{U}_i \to \mathcal{B}_q$, and one for models $\sigma_p^i : \mathcal{U}_i \to \mathcal{B}_p$.

To compare beliefs or models at two different points $c_1, c_2 \in \mathcal{U}_i$, we may use its local gauge frame $\phi_i(c)$ to define **intra-agent parallel transport**.

This transport is mediated by a spatial gauge field $A_\mu^i = \partial_\mu \phi_i(c) \in \mathfrak{g}$, and the transport of a belief/model from point c_1 to c_2 within agent i's region is given by a path-ordered exponential:

$$T^{(i)}_{c_1 \to c_2} := \mathcal{P} \exp\left(- \int_{c_1}^{c_2} A_\mu^i \, dc^\mu \right).$$

This defines a gauge-covariant comparison of beliefs/models within the same agent's field of view, allowing belief gradients, curvature, and local coherence to be assessed.

2.4.2 Inter-agent transport.

By contrast, comparing beliefs/models between different agents i and j requires inter-agent gauge transport — that is, mapping beliefs/models expressed in agent j's gauge frame into agent i's frame. This is necessary because agents, in general, do not share the same frame $\phi(c)$, even at overlapping points $c \in \mathcal{U}_i \cap \mathcal{U}_j$.

We define the inter-agent transport operator $\Omega_{ij}(c) \in G$ by:

$$\Omega_{ij}(c) := \exp\left(\phi_i(c)\right) \cdot \exp\left(-\phi_j(c)\right),$$

which transports a belief $q_j(c) \in \mathcal{B}_q$ into agent i's frame:

$$q_j^{(i)}(c) := \rho\left(\Omega_{ij}(c)\right) \cdot q_j(c).$$

This transformation is used in the inter-agent alignment term of the variational energy:

$$D_{\mathrm{KL}}\left[\, q_i(c) \,\|\, \Omega_{ij}(c) \cdot q_j(c) \,\right],$$

and encodes the degree to which agent i finds agent j's beliefs compatible with its own gauge frame.

We may also define model transport via $\tilde{\Omega}$ similarly.

2.4.3 Combined transport

In more complex scenarios, beliefs/models may be transported across multiple agents and spatial positions. For instance, a belief at $c_i \in \mathcal{U}_i$ may be transported to $c_j \in \mathcal{U}_j$ via intra-agent transport inside i, inter-agent transport via Ω_{ij}, and intra-agent transport inside j. Such composite operations yield generalized holonomies and allow semantic curvature to accumulate.

Type	**Domain**	**Operator**
Intra-agent	$c_1, c_2 \in \mathcal{U}_i$	$T^{(i)}_{c_2 \to c_1} = \mathcal{P} \exp\left(-\int A^i_\mu dc^\mu\right)$
Inter-agent	$c \in \mathcal{U}_i \cap \mathcal{U}_j$	$\Omega_{ij}(c) = \exp(\phi_i(c)) \cdot \exp(-\phi_j(c))$

2.4.4 Holonomy in Epistemic Transport

In our epistemic gauge framework, holonomy captures the failure of global semantic consistency under composite belief/model transport. When a belief/model is parallel transported along a closed loop — involving both intra-agent motion over space and inter-agent communication across gauge frames — it may fail to return to its original form. This failure defines a nontrivial holonomy, analogous to path dependence in meaning or inference.

Let γ be a closed loop in the extended space of agent identity and base manifold position:

$$\gamma = \left[(i_1, c_1) \to (i_2, c_2) \to \cdots \to (i_n, c_n) \to (i_1, c_1)\right],$$

where each pair (i_k, c_k) represents a position $c_k \in \mathcal{U}_{i_k} \subset \mathcal{C}$ within agent i_k's support.

We define the composite transport operator $\mathcal{P}_\gamma$ along the loop as an ordered sequence of:

- Intra-agent spatial transports $T^{(i_k)}_{c_{k+1} \to c_k}$ when remaining within the same agent, and
- Inter-agent frame transports $\Omega_{i_k i_{k+1}}(c_k)$ when transitioning between agents at shared positions.

The full holonomy operator is thus given by:

$$\mathcal{P}_\gamma := T^{(i_n)}_{c_1\to c_n} \cdot \Omega_{i_n i_{n-1}}(c_n) \cdot T^{(i_{n-1})}_{c_n\to c_{n-1}} \cdots \Omega_{i_2 i_1}(c_2) \cdot T^{(i_1)}_{c_2\to c_1}.$$

Given a belief $q \in \mathcal{B}_q$ at the starting point (i_1, c_1), the holonomy is defined as:

$$\mathrm{Hol}_\gamma(q) := \mathcal{P}_\gamma \cdot q.$$

If $\mathrm{Hol}_\gamma(q) \neq q$, the loop is said to have nontrivial holonomy — signaling the presence of epistemic curvature in the connection field. This measures the cumulative distortion of belief through a closed sequence of communication and spatial inference. The deviation norm $\|q - \mathrm{Hol}_\gamma(q)\|$ quantifies the degree of semantic dissonance accumulated around the loop.

Holonomy then plays a central role in modeling topological epistemology: persistent holonomies reflect structural belief misalignments that cannot be resolved by local adjustments alone, analogous to geometric frustration, cognitive bias, or cultural divergence in agent societies. These tools would be especially useful for sociologists and researchers studying inter/intra agent coordination within larger epistemic systems.

3 Epistemic Dynamics and Quasi-static Inference

We now define the dynamical evolution of agents within our epistemic gauge framework. Agents are not static observers but active inference systems, continuously updating their beliefs q, models p, and gauge frames ϕ to minimize a generalized variational energy $\mathscr{V}$. These dynamics are inherently local, non-equilibrium, and, exhibit time-scale separation (models tend to evolve slower than beliefs) — reflecting the constrained, embodied nature of cognitive systems.

The evolution of each field is governed by a variational gradient descent flow, modulated by learning rates and regularization terms.

3.1 Gradient Dynamics of Beliefs and Models

Beliefs evolve by descending the variational energy gradient with respect to q_i, the agent's recognition field. The basic update rule is:

$$\frac{dq_i(c)}{dt} = -\eta_q \frac{\delta \mathscr{V}_i}{\delta q_i(c)}.$$

Here, η_q is the belief update rate (or inference speed), and $\mathscr{V}_i$ is the total variational energy for agent i. This includes both the self-alignment term $D_{\mathrm{KL}}[q_i \| \Phi p_i]$ and the inter-agent alignment term $D_{\mathrm{KL}}[q_i \| \Omega_{ij} q_j]$, as described earlier.

If model plasticity is permitted, the model field $p_i(c)$ also evolves according to:

$$\frac{dp_i(c)}{dt} = -\eta_p \frac{\delta \mathscr{V}_i}{\delta p_i(c)},$$

where η_p controls the timescale of model adaptation. In many cases, p_i is held fixed, representing innate priors or evolutionarily hardwired structure.

3.2 Gauge Frame Evolution

Each agent's gauge frame $\phi_i(c) \in \mathfrak{g}$ evolves according to a damped, curvature-regularized gradient flow. This flow is derived from the variational energy $\mathscr{V}_i$, with additional terms penalizing excessive gauge complexity and favoring smoothness (if desired gauge smoothing terms can be set to zero):

$$\frac{d\phi_i(c)}{dt} = -\eta_\phi \left(\frac{\delta \mathscr{V}_i}{\delta \phi_i(c)} + \gamma_\phi \, \phi_i(c) - \lambda_\phi \, \Delta \phi_i(c) \right).$$

- η_ϕ: gauge update rate
- γ_ϕ: damping coefficient (a simplicity prior/mass)
- λ_ϕ: curvature regularization (Laplacian term)

This update rule ensures that gauge frames evolve slowly, resist large excursions, and maintain coherence across $\mathcal{C}$. The Laplacian $\Delta \phi_i$ acts as a spatial smoothing term, penalizing rough or disjointed gauge configurations.

3.3 Timescales and Non-Equilibrium Regimes

Our framework reflects the inherently non-equilibrium nature of cognition. The components q_i, ϕ_i, and p_i evolve on distinct timescales, reflecting their respective epistemic roles:

- **Beliefs** $q_i(c)$ evolve rapidly in response to incoming observations and interactions. These encode moment-to-moment inferences about the agent's environment and neighbors.
- **Gauge frames** $\phi_i(c)$ evolve more slowly, capturing the agent's representational stance or semantic orientation. These mediate communication and may adjust via repeated interactions or accumulated dissonance.
- **Models** $p_i(c)$ generally evolve on the slowest timescale, if at all. These encode the agent's internal generative assumptions — formed developmentally, evolutionarily, or culturally — and represent its most stable epistemic structure.

This timescale separation defines a quasi-static regime: beliefs may equilibrate locally given the current ϕ_i and p_i, but the system as a whole remains dynamically evolving, particularly in the presence of epistemic curvature or topological misalignment. We say a system has reached an epistemic heat death if epistemic equilibrium is reached.

Generally, agents evolve to align (unless topological defects prevent this) and alignment produces meta-agents who themselves evolve to align with other meta-agents. In this manner an ever expanding tower of agents evolves up to the point that a single meta-agent exists. At this point the system reaches epistemic heat death. Therefore, such a simple geometric framework manifestly encodes a law that complexity increases over time if we consider complexity to reflect the scale of the "largest" meta-agent. We may be able to escape epistemic heat death by coupling meta-agent beliefs/models to the lower level agents in the manner identified by the $\Theta, \tilde{\Theta}$ gauge connections. This then injects an epistemic feedback mechanism which could potentially fragment meta-agents into lower scale agents. We pursue this route in future studies.

Next, utilizing our regularization terms outlined above our variational energy for agent i becomes

$$
\begin{aligned}
\mathscr{V}_i = \alpha\, D_{\mathrm{KL}}\left[q_i(c) \,\|\, \Phi p_i(c)\right] \\
&+ \beta \sum_j D_{\mathrm{KL}}\left[q_i(c) \,\|\, \Omega_{ij}(c) \cdot q_j(c)\right] \\
&+ \tilde{\beta} \sum_j D_{\mathrm{KL}}\left[p_i(c) \,\|\, \tilde{\Omega}_{ij}(c) \cdot p_j(c)\right] \\
&+ \tilde{\alpha}\, D_{\mathrm{KL}}\left[p_i(c) \,\|\, \tilde{\Phi} q_i(c)\right] \\
&+ \frac{\gamma_\phi}{2} \left\|\phi_i(c)\right\|^2 + \frac{\lambda_\phi}{2} \left\|\nabla^2 \phi_i(c)\right\|^2 \\
&+ \frac{\tilde{\gamma}_\phi}{2} \left\|\tilde{\phi}_i(c)\right\|^2 + \frac{\tilde{\lambda}_\phi}{2} \left\|\nabla^2 \tilde{\phi}_i(c)\right\|^2
\end{aligned}
\tag{11}
$$

where we write:

$$
\mathscr{R}_\phi = \frac{\gamma_\phi}{2} \left\|\phi_i(c)\right\|^2 + \frac{\lambda_\phi}{2} \left\|\nabla^2 \phi_i(c)\right\|^2, \tag{12}
$$

$$
\mathscr{R}_{\tilde{\phi}} = \frac{\tilde{\gamma}_\phi}{2} \left\|\tilde{\phi}_i(c)\right\|^2 + \frac{\tilde{\lambda}_\phi}{2} \left\|\nabla^2 \tilde{\phi}_i(c)\right\|^2. \tag{13}
$$

and

$$
\begin{aligned}
\mathscr{V}_i = \alpha\, D_{\mathrm{KL}}\left[q_i(c) \,\|\, \Phi p_i(c)\right] \\
&+ \beta \sum_j D_{\mathrm{KL}}\left[q_i(c) \,\|\, \Omega_{ij}(c) \cdot q_j(c)\right] \\
&+ \tilde{\beta} \sum_j D_{\mathrm{KL}}\left[p_i(c) \,\|\, \tilde{\Omega}_{ij}(c) \cdot p_j(c)\right] \\
&+ \tilde{\alpha}\, D_{\mathrm{KL}}\left[p_i(c) \,\|\, \tilde{\Phi} q_i(c)\right] \\
&+ \mathscr{R}_\phi + \mathscr{R}_{\tilde{\phi}}
\end{aligned}
\tag{14}
$$

4 Euler–Lagrange Equations for Epistemic Fields

The Euler-Lagrange equations governing the two primary fields are given by

$$
\frac{dq_i(c)}{dt} = -\eta_q \cdot \left[\alpha \left(\log \frac{q_i(c)}{\Phi p_i(c)} + 1\right) + \beta \sum_j \left(\log \frac{q_i(c)}{\Omega_{ij}(c) \cdot q_j(c)} + 1\right) - \tilde{\alpha}\, \frac{p_i(c)}{q_i(c)}\right]. \tag{15}
$$

and

$$\frac{d\phi_i(c)}{dt} = -\eta_\phi \cdot \left[\beta \sum_j \nabla_{\phi_i} D_{\mathrm{KL}} \left(q_i(c) \,\|\, \Omega_{ij}(c) \cdot q_j(c) \right) + \gamma_\phi \cdot \phi_i(c) - \lambda_\phi \cdot \Delta\phi_i(c) \right]. \tag{16}$$

Such that, since

$$\Omega_{ij} = \exp(\phi_i) \cdot \exp(-\phi_j), \quad \Rightarrow \frac{\partial \Omega_{ij}}{\partial \phi_i} = \Omega_{ij} \cdot \mathrm{Ad}_{\exp(-\phi_j)}.$$

Thus,

$$\frac{\partial(\Omega_{ij} \cdot q_j)}{\partial \phi_i} = \Omega_{ij} \cdot (\log q_j) \cdot \mathrm{Ad}_{\exp(-\phi_j)}.$$

Therefore,

$$\nabla_{\phi_i} D_{\mathrm{KL}}[q_i \,\|\, \Omega_{ij} \cdot q_j] = (\log q_i - \log(\Omega_{ij} \cdot q_j))^T \cdot \frac{\partial(\Omega_{ij} \cdot q_j)}{\partial \phi_i}.$$

and similarly for the model fibers.

Importantly, under a variational update $\delta\mathscr{V} = 0$, the belief/model field generally evolves as $q \to q + \Delta q$. The local change in informational content over a single step may be characterized by the self-divergence

$$\Delta\mathcal{I} = D_{\mathrm{KL}}[\, q \,\|\, q + \Delta q \,].$$

In gauge-theoretic terms, we interpret this variation as a local transformation of the form $q \to dg^{-1} \cdot q$, where dg^{-1} is a gauge transformation associated with the Lie group G, acting on the fiber $\mathcal{B}_q$. Thus, we write:

$$\Delta\mathcal{I} = D_{\mathrm{KL}}[\, q \,\|\, dg^{-1} \cdot q \,].$$

This quantity measures the epistemic deviation induced by a local frame change and highlights the informational cost (or tension) of re-expressing a belief in a shifted gauge frame. In the limit $\Delta q \to 0$, this yields a natural information geometry on the belief manifold.

We later conjecture that a single bit change in $\mathscr{V}$ pulls back via an agent's model section to be equal to $\hbar$ units of "action". This gives a fundamental definition of of what we refer to as physical action. Furthermore, we claim that a single bit update in belief/model pulls back to $\ln 2$ Joules per Kelvin a la Landauer. In this manner we can define many physical units/dimensions that otherwise have no fundamental origin in physics. In our view physical dimensions are the pullbacks of geometric informational objects from the fibers onto the base space via a particular agent. For humans we define dimensionful quantities in terms of agreement. That is to say we align our gauge frames and agree on our models of reality.

4.1 Agent Pullbacks to $\mathcal{C}$

Agents do not perceive the base manifold $\mathcal{C}$ directly. Instead, they infer its structure through local belief and model fields—sections of associated bundles $\mathcal{E}_i \to \mathcal{C}$ whose fibers are statistical manifolds $\mathcal{B}_i$. This aligns with contemporary models in neuroscience, which frame perception as inference over latent causes of sensory data. Accordingly, we define an agent's perceived geometry of $\mathcal{C}$ via the pullback of geometric structures defined on their belief/model fibers.

As a concrete example, consider the Fisher information metric defined on the model fiber:

$$\mathscr{F}_{ij}(c) = \mathbb{E}_{p(c)}\left[\frac{\partial \log p(c)}{\partial \theta^i}\frac{\partial \log p(c)}{\partial \theta^j}\right], \tag{17}$$

where θ^i are natural parameters charting the model manifold. Pulling this structure back along the agent's section $\sigma : \mathcal{C} \to \mathcal{E}_p$ yields an induced metric on $\mathcal{C}$:

$$\mathscr{G}_{ij}(c) = \sigma^*(\mathscr{F}_{ij}(c)) = \frac{\partial^2}{\partial \delta c^i \partial \delta c^j} D_{\mathrm{KL}}\left[p(c) \,\|\, p(c+\delta c)\right]. \tag{18}$$

Since $p(c+\delta c)$ lies in a different fiber, comparison requires transport. In the presence of a gauge connection $\tilde{\phi}$, we write:

$$D_{\mathrm{KL}}\left[p(c) \,\|\, p(c+\delta c)\right] = D_{\mathrm{KL}}\left[p(c) \,\|\, \left(I - \tilde{\phi}\cdot\delta c\right)\cdot p(c)\right],$$

so that at leading order,

$$\mathscr{G}_{\mu\nu}(c) = \mathbb{E}_{p(c)}\left[\left(\tilde{\phi}_\mu \cdot \log p(c)\right)\left(\tilde{\phi}_\nu \cdot \log p(c)\right)\right]. \tag{19}$$

This metric $\mathscr{G}_{\mu\nu}(c)$ governs the agent's inferred geometry of $\mathcal{C}$. Since belief and model fields $q_i(c), p_i(c)$ generally vary across agents, their respective pullback geometries may differ unless their fields are locally aligned.

We propose that the perceived spatial or spacetime geometry $g_{\mu\nu}$ could emerge from such inference dynamics. Under biologically constrained conditions—e.g., in human-scale agents—the pullback geometry may converge toward Euclidean or Minkowski structure as an evolutionarily stabilized attractor that supports efficient inference and shared coordination.

In general, since both the base and fiber manifolds may be high-dimensional, the pullback metric can naturally decompose:

$$\mathscr{G}_{ij} = g_{\mu\nu} \oplus \mathcal{G}_{\alpha\beta}, \tag{20}$$

where $g_{\mu\nu}$ spans the agent's perceptual subspace (typically 3 or 4 dimensions), and $\mathcal{G}_{\alpha\beta}$ captures residual structure in directions not directly accessible to conscious inference.

Because $\mathscr{F}_{ij}$ depends on the belief/model configuration and the gauge connection $\tilde{\phi}$, this decomposition is generally dynamic: variation in the hidden components of the agent's epistemic state can induce perturbations in the perceived geometry.

This structure is also justified by symmetry. If $\mathscr{F}_{ij}(c)$ is invariant under a Lie group G, then any subgroup $H \subset G$ defines an invariant subspace of the fiber. Under the adjoint action of H, the Lie algebra decomposes as:

$$\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{h}^{\perp}, \tag{21}$$

inducing a corresponding block structure in the pullback metric. For example, if $G = SO(10)$ and $H = SO(3)$, with $\dim \mathcal{B}_p = 10$, then:

$$\mathscr{G}_{ij} = g_{\mu\nu} \oplus \mathcal{G}_{\alpha\beta}, \tag{22}$$

where $g_{\mu\nu}$ spans a 3D perceptual subspace, and $\mathcal{G}_{\alpha\beta}$ spans a complementary 7D "dark" epistemic sector.

In general, the metric may include coupling blocks between these sectors:

$$\mathscr{G}_{\mu\nu}(c) = \mathbb{E}_{p(c)}\left[\left(\tilde{\phi}_\mu \cdot \log p\right)\left(\tilde{\phi}_\nu \cdot \log p\right)\right]$$

$$= \begin{pmatrix} g_{ab} & \Delta_{aB} \\ \Delta_{bA} & \mathcal{D}_{AB} \end{pmatrix}.$$

The off-diagonal blocks Δ_{aB} encode how sensitive the inferred spatial geometry is to fluctuations in hidden epistemic directions. These couplings are gauge-covariant and generally cannot be removed by local changes of basis. As such, variation in the dark sector geometry $\mathcal{D}_{AB}$ can propagate into g_{ab}, perturbing the perceived metric.

At second order, such coupling terms contribute to curvature in the perceptual geometry:

$$\delta R_{ab} \sim \Delta_{aB}\, \mathcal{D}^{BC}\, \Delta_{bC}. \tag{23}$$

This motivates the definition of an effective "dark" stress-energy tensor:

$$R_{ab} - \tfrac{1}{2} g_{ab} R = T_{ab}^{\text{dark}}, \quad T_{ab}^{\text{dark}} := \Delta_{aB}\, \mathcal{D}^{BC}\, \Delta_{bC}. \tag{24}$$

We do not claim this construction replaces physical models of dark matter or energy. Rather, it shows that curvature-like effects in perceived geometry can, in principle, emerge from latent structure in the agent's own belief and model dynamics.

This provides a formal explanation for perceptual distortions observed under altered states of consciousness—such as psychedelics, neurological trauma, or stress—which may correspond to perturbations of the Fisher information metric and hence to distortions in the pullback geometry. These deviations are not anomalous, but expected consequences of changes in the agent's epistemic configuration.

In this light, the agent's experience of space, time, energy, and quantization might arise from informational constraints on inference, and their associated pullbacks from fiberwise to basewise geometry. This epistemic geometry offers a path toward unifying observer and observed within a single coherent variational framework.

In general, informational-geometric structures defined on the fiber manifolds pull back, via the agent's section, to corresponding agent-specific geometric objects on the base manifold $\mathcal{C}$. Two agents may therefore pull back the same global structure in manifestly different ways, depending on their respective belief and model configurations. This reflects the epistemic fact that different agents may infer different geometries, even when operating over the same latent manifold.

We may further extend this reasoning to dynamics. Specifically, if a collection of agents share aligned models over a common region of $\mathcal{C}$, then their pullbacks of geometric flows—such as gauge-parallel trajectories or curvature evolution—will be structurally consistent. In this setting, the invariance of dynamics under local gauge transformations may be understood as a consequence of model alignment: that is, \emph{gauge invariance arises as an emergent property of inter-agent epistemic agreement}.

From this perspective, classical gauge symmetry in physics may reflect a deeper informational symmetry: the consistency of inferred dynamics across agents with locally aligned internal models.

Let us consider two agents A_i and A_j with respective model sections σ_i^p, σ_j^p and gauge frames $\tilde{\phi}_i(c)$, $\tilde{\phi}_j(c)$. Their models are said to be *model aligned* at point $c \in \mathcal{C}$ if there exists a gauge transformation $\tilde{\Omega}_{ij}(c)$ such that

$$p_i(c) \approx \tilde{\Omega}_{ij}(c) \cdot p_j(c), \tag{25}$$

and their models agree under the corresponding $\tilde{\Omega}_{ij}(c)$.

Now consider the induced metric pullbacks from their fibers:

$$\mathscr{G}_{ab}^{(i)}(c) = \sigma_i^* \mathscr{F}_{ab}^{(i)}(c), \qquad \mathscr{G}_{ab}^{(j)}(c) = \sigma_j^* \mathscr{F}_{ab}^{(j)}(c). \tag{26}$$

If p_i and p_j are aligned (up to $\tilde{\Omega}_{ij}$), and the information geometry is G-invariant, then

$$\mathscr{G}_{ab}^{(i)}(c) = \Lambda_a{}^{a'} \Lambda_b{}^{b'} \mathscr{G}_{a'b'}^{(j)}(c), \tag{27}$$

for some local transformation $\Lambda \in \mathrm{Aut}(T_c\mathcal{C})$ induced by the action of $\tilde{\Omega}_{ij}$ on tangent directions. Thus, model alignment implies coordinate invariance of the perceived geometry.

This reframes diffeomorphism invariance and gauge symmetry not as fundamental ontological principles, but as epistemic constraints: conditions for mutual intelligibility between agents.

Gauge Relativity Conjecture — All fundamental symmetries in physics (e.g., Lorentz, diffeomorphism, gauge) arise from the requirement that agent pullback geometries be consistent under epistemic gauge transport. Shared physical laws are not fundamental constraints on reality, but emergent regularities among agents whose beliefs are sufficiently aligned. In this view, a physical law is a maximally symmetric consensus geometry in the space of agent inferences.

While individual human agents do not directly perceive Lorentz symmetry, the collective process of scientific inference — a form of meta-agent epistemology — converges on a shared generative model that is Lorentz invariant. Thus, symmetries like SO(3,1) are not raw percepts but consensus structures in model space arising from epistemic alignment across agents under shared experimental contexts. We claim that a collection of epistemologically interacting agents can condense and align into a meta-agent whose beliefs and models can be manifestly different than that of their lower-agent components.

4.1.1 Meta-Agent Emergence.

We claim that a collection of epistemologically interacting agents can condense and align into a higher-order meta-agent, whose induced beliefs and models may differ qualitatively from those of its constituent agents. That is, alignment of local sections

$$\{\sigma_i^p, \sigma_i^q\}, \quad A_i \in M,$$

over a shared region of $\mathcal{C}$ enables the formation of an effective joint section Σ^p, Σ^q, representing the meta-agent $\mathcal{M}$.

Crucially, the geometric and epistemic structure of this emergent agent need not be reducible to that of any single constituent. In particular, the pullback geometry $\mathscr{G}_{\mu\nu}^{(\mathcal{M})}$ induced by the meta-agent may display symmetries, dimensions, or invariants (e.g., Lorentz structure, conserved quantities) not manifest in any lower-level agent's model space.

In this way, scientific knowledge and physical law may be understood as the result of epistemic condensation: an alignment-driven emergence of invariant structures in model space through multi-agent coordination. The resulting meta-geometries reflect consensus inference across agents, not any one subjective perspective.

4.1.2 Meta-Agent Construction.

Let $\{A_i\}_{i\in I}$ be a collection of agents, each defined as local sections over the model and belief bundles:

$$A_i = (\sigma_i^p : U_i \to \mathcal{B}_p, \quad \sigma_i^q : U_i \to \mathcal{B}_q),$$

with gauge frames $\tilde{\phi}_i(c), \phi_i(c) \in \mathfrak{g}$.

We say these agents are epistemically coherent on overlaps if there exist gauge transformations $\tilde{\Omega}_{ij}, \Omega_{ij} \in G$ such that

$$\sigma_i^p(c) \approx \tilde{\Omega}_{ij}(c) \cdot \sigma_j^p(c), \quad \sigma_i^q(c) \approx \Omega_{ij}(c) \cdot \sigma_j^q(c), \quad \forall c \in U_i \cap U_j.$$

Given this coherence, we define a meta-agent $\mathcal{M}$ as a pair of global sections

$$\Sigma^p : U \to \mathcal{B}_p, \qquad \Sigma^q : U \to \mathcal{B}_q,$$

where $U = \bigcup_{i \in I} U_i$, and $\Sigma^p|_{U_i} \approx \sigma_i^p$, up to gauge. Similarly for Σ^q. This corresponds to a gluing of local fields into a coherent global epistemic state.

The meta-agent is then endowed with an effective gauge frame as

$$\phi^{(\mathcal{M})}(c) := \frac{1}{|N(c)|} \sum_{i \in N(c)} \phi_i(c),$$

where $N(c)$ is the set of agents defined at point $c \in \mathcal{C}$. It also possesses a variational energy functional $V^{(\mathcal{M})}$, constructed from the joint fields $\Sigma^p, \Sigma^q, \phi^{(\mathcal{M})}$ and the same energy principle as in individual agents. In this manner new gauge connections $\Lambda^s_{s'}, \tilde{\Lambda}^s_{s'}$ emerge connecting agents and meta-agents across emergent scales s, s'. These emergent gauge connections between scales act as renormalization operators. These inter-scale transformations encode the relative embedding of local agent models within larger-scale collective structures, and govern the transport of beliefs and models across hierarchical layers of inference.

Furthermore, via the $q \to p$ and $p \to q$ bundle morphisms we can connect agent/meta-agent beliefs and modelsx via

$$\Theta^s_{s'} : \Gamma^s(\mathcal{B}_q) \to \Gamma^{s'}(\mathcal{B}_p)$$

$$\tilde{\Theta}^s_{s'} : \Gamma^s(\mathcal{B}_p) \to \Gamma^{s'}(\mathcal{B}_q)$$

As meta-agents emerge from the alignment of lower-level agents, the same variational principles that govern individual inference extend naturally to these higher-order structures. That is, meta-agents optimize their own belief and model bundles via variational energies that mirror those of their constituents:

$$\mathscr{V}^{(\mathcal{M})} = \alpha^s \cdot D_{\mathrm{KL}}\left[q^{(\mathcal{M})} \| p^{(\mathcal{M})}\right] + \beta^s \cdot D_{\mathrm{KL}}\left[q^{(\mathcal{M})} \| \Omega^{(\mathcal{M})}_{\mathcal{M}\mathcal{M}'} \cdot q^{(\mathcal{M}')}\right] + \cdots$$

This recursive application of the energy principle allows inference to operate coherently across, and between, scales.

This formalism shows how multi-agent epistemic alignment induces emergent geometric and dynamical structure. The resulting meta-agent may exhibit novel symmetries or lower effective curvature due to internal agreement — even if constituent agents are locally noisy or limited in scope.

Naturally we can consider adding as many agents and as many scales and connections as we have available. We then have generally a total energy/potential functional (sums over agents and meta-agents implied) for an agent A^s as

$$\mathcal{F}_T[q_A^s] = D_{KL}[\ q_A^s(c)\ ||\ \Phi p_A^s(c)\ ]+$$

$$\mathcal{F}_\Lambda + \mathcal{F}_{\tilde{\Lambda}} + \mathcal{F}_\Theta + \mathcal{F}_{\tilde{\Theta}} + \mathcal{F}_\Omega + \mathcal{F}_{\tilde{\Omega}} + \mathcal{F}_{\tilde{\Phi}}$$

For general complex systems this may be computationally intractable to simulate in full detail but avails itself to simplifications. Also note that this will define a set of coupling constants that one may be able to fit phenomenologically.

5 Results: Emergence of Meta-Agents in $SL(2,\mathbb{R})$ Epistemic Geometry

We simulate a fully non-Abelian gauge-theoretic system of agents using the special linear group $\mathrm{SL}(2,\mathbb{R})$, consisting of 30 agents distributed randomly on a two-dimensional periodic domain. Each agent occupies a circular region of radius 3 and is modeled as a local section of a fiber bundle with:

- Belief field $q_i(c) \in \mathcal{B}_q$
- Model field $p_i(c) \in \mathcal{B}_p$
- Gauge potentials $\phi_i(c)$ (belief) and $\tilde{\phi}_i(c)$ (model)

These fields live in $K = 17$-dimensional Lie-algebra-valued fibers.

5.1 Initialization

Agent beliefs and models are initialized as spatially varying Gaussians over soft supports:

- **Belief field:** $q_i(c) \sim \mathcal{N}(\mu_q(c), \sigma_q(c))$ with
$$\mu_q(c) \in [2, K-2], \quad \sigma_q(c) \in [0.5, 3]$$
- **Model field:** $p_i(c) \sim \mathcal{N}(\mu_p(c), \sigma_p(c))$ with
$$\mu_p(c) \in [2, K-2], \quad \sigma_p(c) \in [0.75, 2]$$
All agents shared identical models
- **Support masks:** $\chi_i(c)$ are Gaussian soft radial masks.
- **Smoothing:** Mean and variance fields are smoothed via Gaussian kernels with bandwidths in the range
$$\text{smoothness} \in [2, 6] \text{ for both } \mu(c), \sigma(c)$$
- **Gauge fields:** Belief gauge fields are initialized as:
$$\phi_i(c) \sim \mathcal{N}(0, 1) \quad \text{and} \quad \tilde{\phi}_i(c) = \phi_i(c) + \delta\phi, \quad \delta\phi = 0.1$$

5.2 Variational Energy and Parameters

Agents are embedded in Lie-algebra-valued fibers with $\mathrm{SL}(2, \mathbb{R})$ structure. The total free energy functional is:

$$\mathcal{F} = \sum_i \left[\alpha \int_{\mathcal{C}} D_{\mathrm{KL}}(q_i(c) \,\|\, p_i(c))\, dc + \beta \sum_{j \in \mathcal{N}(i)} \int_{\mathcal{C}} D_{\mathrm{KL}}(q_i(c) \,\|\, \Omega_{ij}(c) \cdot q_j(c))\, dc + \gamma \int_{\mathcal{C}} \|\nabla \phi_i(c)\|^2\, dc + \lambda \right.$$

With parameters:

$$\alpha = 1, \quad \beta = 1, \quad \gamma = 0, \quad \lambda = 0$$

All model-alignment and model-feedback terms are disabled in this configuration as agents share identical models.

5.3 Emergence of Quasi-Meta-Agents - XXXXXXXXX

Simulations on a 2D periodic domain reveal spontaneous emergence of quasi-meta-agents: coherent clusters of agents whose gauge-aligned beliefs become nearly indistinguishable under nontrivial gauge maps Ω_{ij}. These groups exhibit low KL divergence and share identical models.

5.3.1 Energy descent and alignment

During $T = 200$ variational steps the total free energy

$$\mathcal{F} = E_{\text{self}} + E_{\text{align}} + E_{\text{mass}} + E_{Lap}$$

dropped sharply, chiefly through a reduction in the inter-agent alignment term

$$E_{\text{align}} = \sum_{i,j} D_{\text{KL}}\big[q_i \| \Omega_{ij} q_j\big].$$

Self-alignment between each agent's belief and model fields

$$E_{\text{self}} = \sum_i D_{\text{KL}}\big[q_i \| p_i\big]$$

decreased only modestly.

After ~ 100 steps the magnitude of the gauge potentials $\|\phi_i(c)\|$ reached a minimum and then rose steadily, mirrored by an increase in the Laplacian energy

$$E_{Lap} = \gamma \sum_i \int_C \|\nabla \phi_i\|^2 \, dc$$

and a continued decrease in the mass energy

$$E_{\text{mass}} = \lambda \sum_i \int_C \|\phi_i\|^2 \, dc$$

Meta-agents were categorized by the condition that

$$D_{\text{KL}}\big[q_i \,\|\, \Omega_{ij} q_j\big] < 1,$$

Three such groups emerged, the largest containing eight agents.
Agents lacking spatial overlap experienced negligible evolution;
their beliefs relaxed toward their static models via the self-alignment term alone.

5.3.2 Holonomy experiments

We transported beliefs around closed paths using the full holonomy operator given by

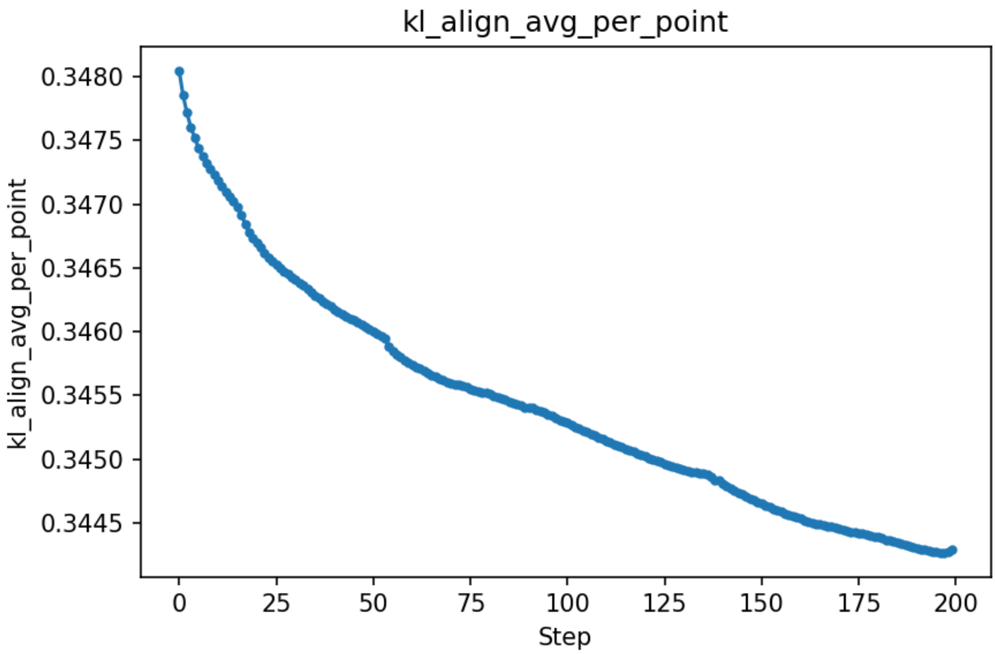
$$\mathcal{P}_\gamma := T^{(i_n)}_{c_1 \to c_n} \cdot \Omega_{i_n i_{n-1}}(c_n) \cdot T^{(i_{n-1})}_{c_n \to c_{n-1}} \cdots \Omega_{i_2 i_1}(c_2) \cdot T^{(i_1)}_{c_2 \to c_1}.$$

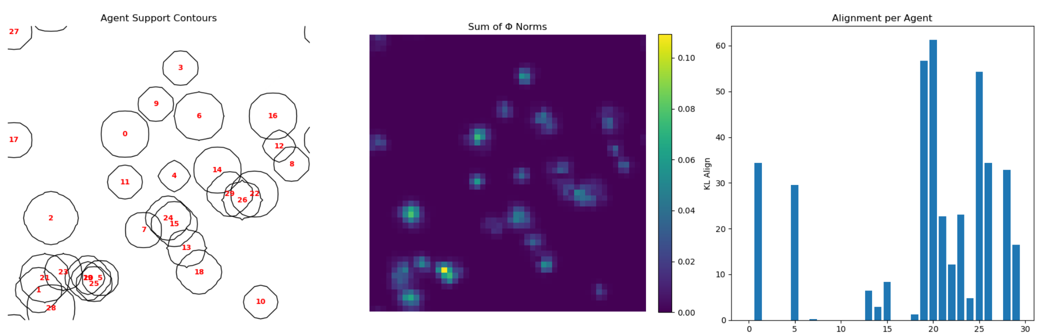
where $T^{(i_n)}_{c_1 \to c_n}$ represents spatial transport within a single agent i_n from points $c_1 \to c_n$ within the agent's mask. $\Omega_{i_n i_{n-1}}(c_n)$ is the inter-agent transport at a single point c_n in the overlap region between a pair of agents.

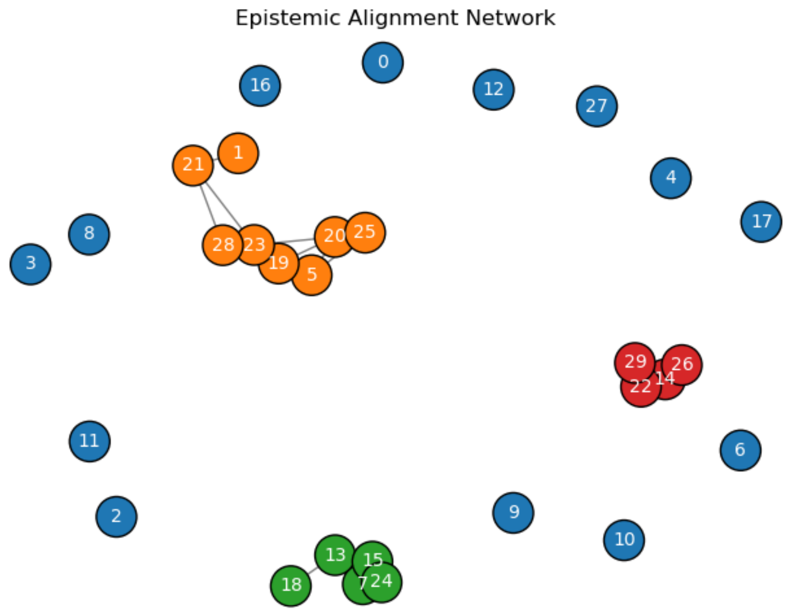
For cycles between distinct meta-agents, the holonomy was non-trivial: the transported belief differed from its starting value by roughly one basis vector in the K-dimensional fiber. This indicates appreciable curvature in the epistemic gauge field.

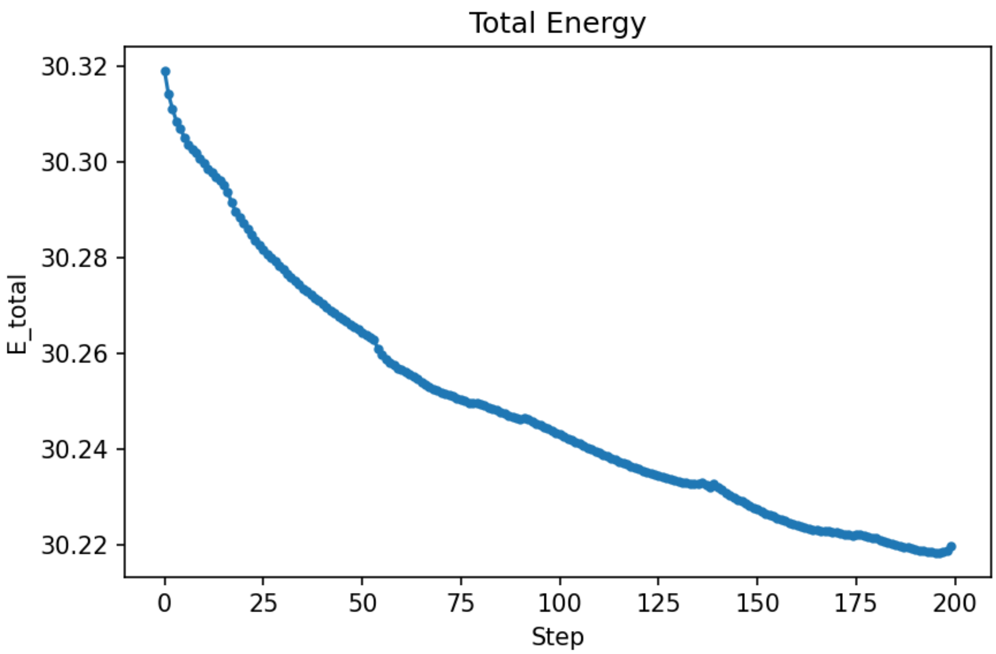
For cycles within a single meta-agent (including spatial loops that traverse multiple agents), the holonomy was trivial to numerical precision, corroborating the interpretation of each meta-agent as a locally flat epistemic domain.

Together, these results show that variational descent not only forms coherent meta-agents but also sculpts a curved epistemic landscape whose holonomy structure cleanly discriminates meta-agent formation.


kl_align_avg_per_point
0.3480
0.3475
0.3470
0.3465
0.3460
0.3455
0.3450
0.3445
kl_align_avg_per_point
0
25
50
75
100
125
150
175
200
Step


Agent Support Contours
Sum of Φ Norms
0.10
0.08
0.06
0.04
0.02
0.00
Alignment per Agent
KL Align
Agent Index


Epistemic Alignment Network
0
16
12
27
4
17
1
21
8
3
28
23
19
5
20
25
29
26
22
14
6
11
2
9
10
13
15
18
7
24


Total Energy
30.32
30.30
30.28
30.26
30.24
30.22
E_total
0
25
50
75
100
125
150
175
200
Step