Ageometrics and the Geometric Sufficiency Ratio

A Research Program for Measuring What Geometry Preserves, What It Erases, and What Must Be Added Back

Version: 0.5 working paper
Date: July 2026
Public attribution: Synaptient
Status: Decision-theoretic working paper with minimal reproducible illustration; empirical validation, sensitivity analysis, observer-envelope testing, and exhaustive prior-art review remain in progress.

> Fractalish asks what history becomes visible in form. Ageometrics asks what history disappears when reality is reduced to form.

Abstract

Geometry is among the most powerful representational technologies ever developed. Physical systems, social relations, linguistic meanings, causal networks, computational states, memories, trajectories, and probability distributions can all be embedded into spaces whose distances, neighborhoods, curvatures, symmetries, and topologies support analysis. This reach is a scientific triumph. It also creates a methodological danger: once almost anything can be represented geometrically, the mere existence of a useful geometry ceases to show that geometry is causally, historically, semantically, or functionally sufficient.

This paper introduces **Ageometrics**, the quantitative study of the limits of specified geometric representations. Ageometrics does not reject geometry. It asks a stricter question: for a declared system record, geometric representation, target, baseline, loss, observer, and evaluation protocol, how much target-relevant decision value survives compression into geometry, how stable is that recovery under reasonable changes of representation and observer, and what typed residue remains outside the declared geometric account?

The paper proposes the **Geometric Sufficiency Ratio** (GSR) as a task-relative measure of the fraction of recoverable predictive or decision performance captured by a declared geometric representation relative to a fuller admissible reference record. The canonical formulation is defined through expected loss and Bayes-optimal risk. Empirical GSR is explicitly protocol-dependent and requires matched data access,

temporal availability, evaluation folds, loss, tuning effort, stopping rules, leakage controls, and observer declarations. The paper also defines **Non-Geometric Residue** (NGR), **V-GSR** for restricted predictive families, and interventional, temporal, representation-stability, restoration-channel, and encoding-cost extensions. A first taxonomy of residue is proposed: temporal, causal, provenance, contextual, algorithmic, semantic, identity, and energetic residue.

A minimal reproducible illustration plants temporal information outside a declared terminal-geometric representation. Using the same logistic learner and matched training protocol for the geometric representation and fuller-record inputs, the geometric representation retains about one fifth of the log-loss improvement available from the fuller record in the seeded example, with the v0.4 review recording GSR approximately 0.21 and bootstrap confidence interval [0.189, 0.231]. The illustration shows that GSR can be estimated against an intentionally controlled missing channel; it is not broad empirical validation of Ageometrics.

Ageometrics is positioned beside, but not reduced to, Blackwell comparison of experiments, proper scoring rules, forecast skill scores, Information Bottleneck theory, sufficient dimension reduction, predictive V-information, causal representation learning, contextuality, topological data analysis, geometric deep learning, and algorithmic information theory. Recent work on geometric factorization provides a useful positive complement by proving conditions under which quotient geometry can be sufficient for invariant tasks. Ageometrics asks the complementary empirical question: for a specified representation and protocol, how much sufficiency survives and what residue remains?

The strongest near-term application is artificial intelligence memory because fuller developmental records can be retained. In a history-bearing artificial system, Ageometrics can measure how much causal, autobiographical, provenance, contradiction, continuity, and routing information survives compression into embeddings, graphs, attractors, summaries, replay states, or memory topologies. Cognitive Basin is treated here as an instrumented application and test environment, not as a prerequisite for the mathematical framework. External agent-memory systems such as AutoMem provide additional benchmark opportunities because they expose trajectories, explicit memory operations, evolving schemas, and downstream task behavior.

1. The Problem Hidden Inside Geometry's Success

The sciences have repeatedly advanced by finding the right geometry. Celestial mechanics became tractable through trajectories and conic sections. Relativity redefined gravity through spacetime geometry. Statistical models acquired information geometry. Networks became graphs. Data became point clouds and manifolds. Machine learning turned words, images, proteins, users, and concepts into coordinates. Modern geometric deep learning now offers a unifying language for grids, groups, graphs, geodesics, gauges, and symmetries.

The achievement is so broad that it can conceal its own limit.

A representation may be geometrically expressive enough to encode nearly any structure placed into it. A graph can record ancestry if ancestry is supplied as directed edges. A high-dimensional vector can encode a timestamp if the timestamp is included among the coordinates. A manifold can carry provenance labels if provenance is attached as an additional field. A state space can contain a complete event history if every event is assigned another dimension.

At some point, however, the statement that a system has been represented geometrically becomes empty. The geometry may no longer explain or compress the system; it may simply rename the complete record.

This produces a hierarchy that is often collapsed:

1. A phenomenon can be represented geometrically.
2. A geometric representation is useful for a task.
3. The geometry is sufficient for predicting a target.
4. The geometry is causally explanatory.
5. The geometry preserves the system's history, identity, or semantics.
6. The geometry is the minimal or natural ontology of the system.

Each step requires additional evidence. None follows automatically from the previous one.

Ageometrics begins at the gap between the first two steps and the rest.

2. Definition of Ageometrics

Ageometrics is the quantitative study of information, causality, identity, behavior, history, or meaning that is not preserved by a specified geometric representation alone.

The adjective **specified** is essential. Ageometrics does not claim that some fact is metaphysically incapable of geometric encoding. With enough dimensions, labels, auxiliary fields, temporal indices, and special metrics, almost any finite record can be made geometric.

The real question is operational:

> What must be added to a geometric representation before it becomes sufficient for the target, and at what point has the representation stopped being a genuine geometric compression and become the original system written in coordinates?

Ageometrics therefore treats geometric insufficiency as relative to declared objects:

- the fuller reference record **X**;
- the declared geometric representation **G(X)**;
- the baseline information or predictor **B**;
- the target property, outcome, or task **Y**;
- the loss function or utility **ell**;
- the admissible observer or predictive family **F**;
- the full empirical protocol **Pi**.

Without these declarations, claims of sufficiency are too vague to test.

For technical disambiguation, the full field label used in this paper is **Ageometrics: Geometric Sufficiency and Residue Analysis**. The word *ageometric* already appears as a specialized adjective in geometric group theory, especially in work on outer automorphisms of free groups. That usage is unrelated to the present proposal. The collision should be disclosed in metadata and search descriptions, but it does not presently appear to identify an existing field devoted to geometric sufficiency or residue analysis.

3. Geometry, Topology, and Form

Ageometrics uses *geometry* in a broad operational sense. It includes representations organized primarily through:

- points and coordinates;
- distances and similarities;
- neighborhoods;
- graphs and hypergraphs;
- manifolds;
- curvature;
- symmetry;
- topology;
- embeddings;
- attractor landscapes;
- spatial or relational morphologies.

This breadth is intentional because the scientific issue is not limited to Euclidean shape. A memory graph, a latent vector space, a causal diagram, a topological summary, and a learned embedding are all forms of structured relational representation.

Ageometrics is therefore not merely a critique of visual shape. It examines the sufficiency of any representation whose primary information is carried through arrangement, adjacency, metric relation, invariance, or form.

4. The Geometric Sufficiency Ratio

The original score-based intuition behind GSR is simple: compare how much target-relevant value is retained by geometry with how much is available from the fuller record. A rigorous definition must prevent differences in model power, tuning effort, data access, or score scaling from masquerading as geometric residue. The canonical metric is therefore defined first through predictive risk, followed by a protocol-specific empirical estimator.

4.1 Declared objects and benchmark protocol

Let:

- **X** denote the declared fuller reference record available at the prediction or decision time;
- **G(X)** denote a specified geometric representation derived from X;
- **B** denote the declared baseline information or baseline predictor;
- **Y** denote the target property, outcome, or response;
- **ell** denote the declared loss function, with lower values preferred;
- **Pi** denote the complete empirical comparison protocol.

The protocol Pi must include:

- data-access rules and temporal availability;
- train, validation, and test partitions;
- model or inference families;
- representation-specific preprocessing;
- capacity and compute budgets;
- hyperparameter-search budgets;
- optimization and stopping rules;
- scoring or loss definition;
- uncertainty estimation;
- leakage controls.

The term *fuller reference record* is intentional. X is not presumed to be a metaphysically complete description of the system. It is the richest admissible record declared for the benchmark.

4.2 Canonical population definition

For any information source Z, define its Bayes-optimal risk under loss ell as:

```
R*_ell(Z) = inf_f E[ ell(Y, f(Z)) ],
```

where the infimum is taken over all admissible measurable predictors.

The population Geometric Sufficiency Ratio is:

```
GSR*_{ell,Y}(G | X, B)
```

```
  = [ R*_ell(B) - R*_ell(G(X)) ]
```

```
    / [ R*_ell(B) - R*_ell(X) ],
```

provided:

```
R*_ell(B) > R*_ell(X).
```

This form asks what fraction of the risk reduction available from the fuller record is retained by the declared geometric representation.

4.3 Boundedness under the canonical assumptions

When G(X) is derived from X, the fuller record contains at least the information available in its geometry. Under Bayes-optimal comparison:

```
R*_ell(X) <= R*_ell(G(X)).
```

If the geometry is at least as informative as the baseline:

```
R*_ell(G(X)) <= R*_ell(B).
```

Together these imply:

```
0 <= GSR*_{ell,Y} <= 1.
```

Interpretation:

- GSR near 1: geometry retains nearly all recoverable target-relevant information available in the fuller reference record;
- GSR near 0: geometry retains little beyond the baseline;
- intermediate GSR: geometry is useful but incomplete.

A value outside this interval is not a property of the population definition. It is evidence that an empirical estimate is affected by sampling variation, optimization error, regularization, model-class mismatch, data leakage, or an invalid nesting assumption.

4.4 Protocol-specific empirical estimator

Bayes risks are rarely available. For a declared protocol Pi, let:

- **Rhat_Pi(B)** be cross-validated or held-out baseline risk;
- **Rhat_Pi(G)** be geometry-only risk;
- **Rhat_Pi(X)** be fuller-record risk.

Define the empirical estimator:

```
GSRhat_Pi

  = [ Rhat_Pi(B) - Rhat_Pi(G) ]

    / [ Rhat_Pi(B) - Rhat_Pi(X) ].
```

The three risks must be estimated on the same held-out cases and under a comparator protocol designed to isolate representation rather than model advantage.

At minimum, researchers should report:

- same-learner V-GSR or same-learner empirical GSR when the same predictive family can consume both records;
- approximately capacity-matched comparison when structurally different input types require different architectures;
- model-envelope results when robustness across observer families matters;
- uncertainty estimates and component risks.

A nested model design is preferable when available: the fuller-record predictor should contain the geometry-only predictor as a restricted case. Even then, final conclusions must use out-of-sample risk rather than training fit.

4.5 Out-of-range empirical estimates

An empirical estimate may be negative when the geometry-only path performs worse than baseline. It may exceed one when the geometry-only model generalizes better than the fuller-record model because of finite samples, regularization, optimization, or comparator imbalance.

Such values should be preserved rather than silently clipped, but interpreted as empirical diagnostics, not literal negative or super-complete residue.

Recommended reporting:

- raw empirical GSR;

- a confidence interval from paired bootstrap or repeated cross-fitting;
- the three component risks;
- same-learner, approximately capacity-matched, and model-envelope results where feasible;
- a diagnosis of any interval violation;
- an optional clipped display value used only for visualization.

4.6 Non-Geometric Residue

The canonical Non-Geometric Residue is:

```
NGR*_{ell,Y} = 1 - GSR*_{ell,Y}.
```

Under the canonical assumptions, NGR lies in [0,1]. It is the fraction of recoverable risk reduction available in the fuller reference record that is not retained by the declared geometry.

For empirical work:

```
NGRhat_Pi = 1 - GSRhat_Pi
```

is a signed diagnostic and may fall outside [0,1]. Researchers must label it accordingly.

NGR is not a metaphysical non-geometric substance. It is a target-, representation-, loss-, observer-, and protocol-relative performance gap.

4.7 Worked example

Suppose a task has baseline accuracy of 0.75, geometry-only accuracy of 0.80, and fuller-record accuracy of 0.95. Accuracy here is treated as expected zero-one utility, so higher is better.

```
GSR = (0.80 - 0.75) / (0.95 - 0.75) = 0.25.
```

The geometry captured one quarter of the recoverable improvement above baseline, even though its raw accuracy was 80 percent.

The same calculation in loss form uses error rates:

```
GSR_L = (0.25 - 0.20) / (0.25 - 0.05) = 0.25.
```

4.8 Restricted predictive families and V-GSR

The canonical population definition uses Bayes-optimal predictors. Many empirical and scientific questions are instead about what a declared observer can extract. Predictive V-information supplies a useful neighboring framework for this situation because it defines usable information relative to a specified predictive family rather than treating all measurable decoders as available.

Let **V** be a declared family of predictive models, algorithms, or observers. Define:

```
R^V_ell(Z) = inf_{f in V} E[ ell(Y, f(Z)) ].
```

Then the restricted-family Geometric Sufficiency Ratio is:

```
V-GSR_{ell,Y}(G | X, B)
```

```
  = [ R^V_ell(B) - R^V_ell(G(X)) ]
```

```
    / [ R^V_ell(B) - R^V_ell(X) ].
```

This is not an observer-independent property of the representation. It is the sufficiency of a declared representation for a declared target under a declared predictive family and protocol.

4.8.1 Operational guidance for V-GSR views

When reporting V-GSR, three complementary views are useful. The appropriate view depends on the scientific question, input structures, and available computational budget.

Same-Learner V-GSR is the recommended default for routine empirical work. Use the same declared predictive family, architecture, preprocessing policy, hyperparameter-search budget, optimization procedure, data partitions, random-seed policy, and stopping rule for both the geometric representation and the fuller record. This view holds the observer and much of the training protocol fixed, reducing major learner-capacity and optimization confounds. It does **not** make the comparison architecture-neutral. Different representations may vary in dimensionality, conditioning, sample efficiency, optimization difficulty, or compatibility with the learner's inductive biases. The result must therefore be interpreted as representation sufficiency relative to the declared observer and protocol.

Approximately Capacity-Matched V-GSR should be used when the input types require fundamentally different architectures, such as graph, sequence, image, field, or set-based models. Match the systems as closely as practicable in parameter count, training compute, data exposure, tuning effort, regularization, and stopping criteria. Exact capacity equivalence across architecture families is generally unavailable. Report the matching variables, tolerances, and remaining asymmetries explicitly.

Model-Envelope V-GSR evaluates GSR across a preregistered or otherwise justified collection of credible predictive families. Report the distribution and a declared summary such as the median, range, interquartile range, or lower quantile. This is the most robustness-oriented view because it exposes dependence on observer choice. It is not automatically numerically conservative: conservatism depends on the aggregation rule selected in advance.

Reporting recommendation. For routine empirical work, report same-learner V-GSR as the primary estimate and identify the predictive family, architecture, preprocessing, partitions, optimization procedure, regularization, tuning budget, stopping rule, random-seed policy, and uncertainty estimate. Where architecture choice is materially coupled to representation type, approximately capacity-matched or model-envelope analysis should become primary. Claims approaching representation-independent sufficiency require evidence across an appropriately broad observer envelope.

View	**Primary purpose**	**What is held or matched**	**Principal limitation**	**Recommended use**
Same-Learner	Hold the declared observer fixed	Family, architecture, protocol, search budget, partitions, stopping rule	Representation-architecture interactions remain	Default empirical estimate
Approximately Capacity-Matched	Compare structurally different input types	Parameters, compute, data exposure, tuning effort, regularization	Exact cross-architecture equivalence is unavailable	Graph vs. sequence vs. set/image comparisons

View	Primary purpose	What is held or matched	Principal limitation	Recommended use
Model-Envelope	Test observer robustness	Credible range of predictive families	Highest computational cost; summary rule matters	High-stakes, summary, or architecture-sensitive claims

4.9 Relationship to existing normalized scores

The algebraic normalization is not claimed to be unprecedented in isolation. Close relatives include forecast skill scores, relative error reduction, coefficients of determination, partial R-squared, likelihood-ratio style improvements, and other reference-normalized performance measures.

The proposed contribution is the specialized experimental object:

- a declared geometric representation;
- a fuller admissible record;
- a task-relative sufficiency ratio;
- explicit comparator discipline;
- restricted-observer variants;
- residue classification;
- interventional and temporal variants;
- representation-stability analysis;
- encoding-cost frontiers;
- matched-equivalence benchmarks;
- restoration-channel tests.

Ageometrics should therefore claim a research framing and measurement program unless later scholarship supports a stronger priority claim, not ownership of the general idea of normalized improvement.

4.10 Information-theoretic companion

When mutual information is identifiable or credibly estimable, an information-theoretic companion is:

```
GSR^I_Y(G | X) = I(G(X); Y) / I(X; Y).
```

Because G(X) is derived from X, the data-processing inequality gives:

```
I(G(X); Y) <= I(X; Y),
```

placing the idealized information ratio in [0,1] when the denominator is positive.

Under logarithmic loss, the reference population GSR has a direct interpretation as a ratio of conditional mutual informations when the baseline is treated as declared side information and the usual regularity conditions hold. This companion connects Ageometrics to Information Bottleneck theory. It should not be treated as the default empirical metric where high-dimensional mutual information cannot be estimated reliably.

4.11 Minimal reproducible illustration: planted temporal residue

The v0.4 working paper added a reproducible synthetic illustration in which a temporal channel is planted outside a declared terminal-geometric representation. The same logistic learner and matched evaluation protocol are used for the geometric representation and fuller-record inputs. In the seeded example, geometry retains about one fifth of the log-loss improvement available from the fuller record; the v0.4 editor review records GSR approximately 0.21 with bootstrap confidence interval [0.189, 0.231].

The illustration has one purpose: to show that the proposed quantity can be estimated against a controlled missing channel and that controlled insufficiency can appear as measurable residue. It is not a benchmark suite, a cross-domain validation, or evidence that all ageometric claims hold generally.

Note on scope. This demonstration uses one seeded generative process, one predictive family, and one planted residue class: temporal information excluded from the declared geometric representation. It does not establish cross-domain validity, representation independence, or robustness across learner families. Systematic testing across seeds, sample sizes, residue classes, representation families, and observer envelopes is reserved for the benchmark program described in Section 20.

5. GSR Is Task-Relative, Not Absolute

No representation is sufficient in the abstract. A road map may be sufficient for navigation and useless for reconstructing political history. A memory graph may predict retrieval latency while failing to distinguish trusted evidence from manipulation. A protein fold may predict binding while omitting the developmental route through which the fold was reached.

Accordingly, the notation should always retain the target:

```
GSR_Y(G).
```

A single representation may have:

- high GSR for immediate classification;
- moderate GSR for robustness under perturbation;
- low GSR for causal attribution;
- near-zero GSR for provenance;
- undefined GSR for subjective meaning.

Ageometrics rejects claims such as “the geometry is sufficient” unless the target and conditions are stated.

6. Extensions of the Ratio

6.1 Interventional Geometric Sufficiency Ratio

Observational equivalence can hide causal differences. Two systems may occupy the same state and behave identically until disturbed.

Let **S_G^do** and **S_X^do** measure prediction of responses under intervention. Define:

```
IGSR_Y(G) = [ S_G^do - S_0^do ] / [ S_X^do - S_0^do ].
```

IGSR asks whether geometry predicts what the system will do when acted upon, not merely what it looks like during passive observation.

This extension is important because causal representation learning has shown that observational structure often fails to identify latent causal variables without interventions, temporal information, multiple environments, or restrictive assumptions.

6.2 Temporal Geometric Sufficiency Ratio

A final state can erase the order by which it was formed. Let **H** be the full event history and **G_T(H)** a geometry derived from the terminal state or a compressed trajectory. Define a temporal target such as next-state response, recovery rate, or vulnerability to relapse.

Temporal GSR measures how much of that target is retained after history becomes form.

6.3 Representation-Stability Envelope

Let **G_set = {G_1, ..., G_n}** be a family of reasonable geometric representations. Define the envelope:

```
E_GSR = { GSR_Y(G_1), ..., GSR_Y(G_n) }.
```

A narrow, high envelope supports structural sufficiency. A wide envelope means that conclusions depend heavily on coordinate choice, metric, embedding dimension, segmentation, graph construction, or scale.

Useful summaries include minimum GSR, maximum GSR, mean GSR, variance, worst-case GSR under admissible transformations, and rank stability of conclusions.

6.4 Geometric Encoding Cost

A geometry can be made sufficient by injecting more information. The important question is how much.

Define an encoding-cost function **C(G)** that measures dimensions, labels, bits, auxiliary fields, or description length required by the representation. Then study the sufficiency-cost frontier:

```
(C(G), GSR_Y(G)).
```

A representation that reaches GSR near one only after encoding every timestamp, source, intervention, and event may have no compression advantage.

This yields a critical test:

> Does geometry discover structure, or merely warehouse the complete record?

6.5 Minimal Residue-Restoring Channel

Let **R** be an auxiliary channel, such as provenance, causal order, timestamps, intervention labels, or energetic state. Define the smallest **R*** such that:

```
GSR_Y(G plus R*) >= tau
```

for a declared sufficiency threshold **tau**.

The content of R* identifies what the geometry was missing. Restoration-channel experiments are crucial because residue classes should be tested by what restores performance, not merely named after a plausible missing factor.

7. A Taxonomy of Non-Geometric Residue

7.1 Temporal residue

Temporal residue is information dependent on sequence, duration, irreversibility, timing, or developmental path.

Examples include identical final weights reached through different curricula, identical scars formed by one severe event or many minor events, and identical graph topology with different edge-creation order.

7.2 Causal residue

Causal residue is information available only through mechanism, intervention, or counterfactual response.

Examples include two networks with the same correlation structure but different causal direction, two materials with identical visible morphology but different fracture response, and two agents with identical outputs but different dependence on hidden tools.

7.3 Provenance residue

Provenance residue concerns origin, authority, chain of custody, manipulation, or evidentiary status.

Examples include a belief derived from direct measurement versus repetition, identical documents from authenticated and forged sources, and the same memory supplied voluntarily or inserted coercively.

7.4 Contextual residue

Contextual residue concerns properties that cannot be assigned independently of the context in which they are measured or used.

Quantum contextuality supplies a rigorous neighboring example: locally consistent assignments may fail to extend to one global context-independent assignment. More generally, a system's state may not be separable from the operational frame used to elicit it.

7.5 Algorithmic residue

Algorithmic residue concerns differences in generative process, program length, computational depth, or construction rule not evident in the resulting geometry.

Two identical patterns may have radically different shortest descriptions or production costs.

7.6 Semantic residue

Semantic residue concerns meaning, obligation, reference, authority, intention, and normativity that are not guaranteed by distance or adjacency.

Two phrases can occupy nearby embedding coordinates while differing in legal force, irony, speaker commitment, or truth status.

7.7 Identity residue

Identity residue concerns continuity, lineage, copying, restoration, replacement, and transformation history.

Two machine instances may have identical current states but different claims to continuity because one is a direct continuation and the other an unauthorized copy.

7.8 Energetic and metabolic residue

Energetic residue concerns the work required to reach or maintain a state.

Two systems may occupy the same configuration while differing in stored tension, thermal burden, depletion, retry history, or maintenance debt. This category is particularly important in artificial cognition, where identical outputs can carry very different compute costs and failure risks.

7.9 Operational residue table

Residue class	**Diagnostic question**	**Example restoration channel**
Temporal	Does order, duration, or path matter?	Timestamps, event sequence, curriculum trace
Causal	Does intervention response diverge?	Intervention labels, mechanism trace, counterfactual data
Provenance	Does origin or chain of custody matter?	Source authority, signature, custody log

Residue class	**Diagnostic question**	**Example restoration channel**
Contextual	Does the measurement frame change the property?	Context labels, operational frame, environment state
Algorithmic	Does construction process matter?	Program trace, generator, description length
Semantic	Does meaning exceed similarity?	Speaker role, normative status, reference links
Identity	Does continuity or lineage matter?	checkpoint lineage, continuity record, restoration history
Energetic	Does work, stress, or maintenance debt matter?	compute logs, resource state, stress/energy telemetry

8. Neighboring Fields and the Unfilled Gap

Ageometrics is not born in an intellectual vacuum. Its ingredients appear across mature fields.

8.1 Information bottleneck and sufficient statistics

Information Bottleneck theory formalizes the search for compressed representations that preserve information relevant to a target. Sufficient statistics ask whether a statistic retains all information needed for inference about a parameter. Sufficient dimension reduction asks whether high-dimensional predictors can be compressed without losing target-relevant structure.

Ageometrics inherits their task-relative rigor but asks a different primary question: what is lost specifically when the chosen compressed representation is geometric, and what kind of residue accounts for the loss?

8.2 Blackwell comparison, proper scoring, and forecast skill

Blackwell comparison of experiments supplies a decision-theoretic language for comparing information sources by the decisions they support. Proper scoring rules and forecast skill scores provide practical evaluation languages for comparing probabilistic predictions and reference-normalized improvement.

GSR is therefore not presented as an isolated algebraic novelty. Its distinct contribution is to bind normalized decision value to a declared geometric compression, fuller record, target, observer, and residue-diagnosis protocol.

8.3 Predictive V-information

Predictive V-information measures usable information under a constrained predictive family. V-GSR adopts that discipline: when the observer is restricted, sufficiency is relative to what that observer can extract under the declared protocol.

This is why same-learner V-GSR must not be described as architecture-neutral. It holds the observer fixed; it does not remove every interaction between representation and observer.

8.4 Geometric factorization and sufficient harmonic representations

Recent work on geometric factorization provides a useful positive complement. Stewart (2026) establishes conditions under which quotienting by task-invariant group actions yields a sufficient-and under an additional orbit-separation condition, minimally sufficient-geometric representation. That result asks when a specified geometric quotient is sufficient under declared invariances and model assumptions.

Ageometrics asks the complementary empirical question: for a declared system record, geometric representation, target, baseline, loss, observer, and protocol, how much target-relevant decision value survives, how stable is that conclusion across reasonable representations and observers, and what typed residue remains?

8.5 Geometric deep learning

Geometric deep learning demonstrates how powerful geometric priors can unify neural architectures across grids, groups, graphs, geodesics, and gauges. Its success strengthens the need for Ageometrics. The more domains geometry successfully organizes, the more carefully researchers must distinguish useful representation from sufficient explanation.

8.6 Causal representation learning

Causal representation learning studies whether latent causal variables and relations can be recovered from high-dimensional observations. Identifiability results repeatedly show that extra assumptions, multiple environments, interventions, or temporal information are needed.

Ageometrics generalizes that lesson. Causal residue is one form of non-geometric residue, and IGSR is designed to measure it.

8.7 Contextuality

The sheaf-theoretic treatment of contextuality formalizes situations where locally valid assignments cannot be combined into one global assignment. This supplies a precise example of information that resists a naive global state-space account.

Ageometrics treats contextual obstruction as a residue class rather than claiming all residue is contextual.

8.8 Topological data analysis

Topological data analysis extracts robust multiscale features such as components, loops, and voids. Reviews of the field acknowledge both its power and its limitations, including insensitivity to non-topological changes and the need to incorporate geometric, spectral, combinatorial, or additional information.

Ageometrics offers a quantitative framework for asking how sufficient a topological summary is for a declared task and what complementary channel restores the missing performance.

8.9 Algorithmic information theory

Algorithmic information theory distinguishes an object's description from its shortest generative program. This is relevant when identical geometry hides different construction histories or when a visually complex form is generated by a short rule.

Algorithmic residue may eventually be formalized through conditional description length.

8.10 The proposed gap

These fields study compression, geometry, causality, context, topology, forecasting skill, relative model improvement, or generative complexity. The normalization underlying GSR has clear mathematical relatives in those literatures and should be credited accordingly.

What appears to be missing is a unified research program in which geometric insufficiency itself is the primary measured object, with:

- a normalized operational measure of decision value retained by a declared geometric representation;
- an observer-relative extension for constrained predictive families;
- representation-stability and observer-envelope analysis;
- a typed taxonomy of Non-Geometric Residue;
- restoration-channel experiments intended to test candidate residue classes rather than merely naming them;
- matched-equivalence benchmarks;
- interventional and temporal variants;
- encoding-cost frontiers;
- an experimental posture for history-bearing processes whose developmental records are compressed into form.

The broader Fractalish program applies this instrument to blind-growth processes, geometry-history tradeoffs, Cognitive Basin, CNTM, and possible physical memory substrates. These applications motivate the framework and provide instrumented test settings, but they are distinct from the core mathematical contribution.

Ageometrics therefore does not replace classical sufficiency analysis, Information Bottleneck methods, geometric representation theory, or predictive information. It provides a focused measurement and diagnostic program for cases in which a fuller developmental or causal record has been compressed into a specified geometric, topological, or morphology-bearing representation.

9. The Foundational Experimental Design

The simplest Ageometric experiment creates two systems that are equivalent under a chosen geometry but differ in a hidden property that matters.

Let systems A and B satisfy:

```
G(A) = G(B)
```

under the representation being tested, while:

```
Y(A) != Y(B).
```

If a geometry-only predictor cannot distinguish their target behavior but a fuller-record predictor can, then the geometric representation has measurable residue.

9.1 Matched-pair benchmark families

Same memory graph, different learning order. Two agents end with identical nodes and edge weights. One learned concepts in a safe curriculum; the other learned them through repeated contradiction and repair. Test future susceptibility to confusion or recovery after perturbation.

Same embedding, different provenance. Two claims have nearly identical semantic vectors. One is supported by direct evidence; the other originated in adversarial manipulation. Test downstream trust calibration.

Same final weights, different training history. Construct networks with matching or near-matching parameter states but different optimization paths, data exposure, or poisoning history. Test robustness and hidden backdoor activation.

Same topology, different energetic state. Two physical or simulated systems share visible structure but differ in stored stress or resource depletion. Test failure under load.

Same output, different reasoning integrity. One agent derives an answer from valid evidence; another copies or rationalizes it. Test transfer to a counterfactual case.

Same current state, different identity lineage. One instance is a continuous process; another is restored from an earlier checkpoint and patched to match. Test autobiographical reconstruction and continuity claims.

Same behavior at rest, different perturbation response. Two systems perform identically under ordinary conditions but diverge after intervention. This is the cleanest IGSR benchmark.

9.2 Requirements

Every benchmark must specify:

- the equivalence relation defining “same geometry”;
- the hidden distinction;
- the target behavior;
- the fuller record;
- the model classes;
- the baseline;
- the intervention, if any;
- the scoring rule;
- the uncertainty estimate;
- the threshold for meaningful residue.

10. Ageometrics for Artificial Intelligence Memory

Artificial memory is the strongest immediate application because its full history can be instrumented.

Most current memory systems reduce experience into one or more of:

- vector embeddings;
- summaries;
- key-value stores;
- graph relations;
- retrieved documents;
- learned weights;
- recurrent state;
- explicit files or notebooks.

These representations are useful. They also erase information.

A vector embedding can preserve similarity while losing source authority. A graph can preserve association while losing order. A summary can preserve conclusions while losing discarded alternatives. Weights can preserve behavior while obscuring which experiences formed it. A retrieval system can return the right statement while forgetting whether the statement was once contradicted. A file memory can preserve an instruction while losing the context that made the instruction trustworthy, temporary, or dangerous.

A history-bearing artificial system can retain:

- event timestamps;
- activation boundaries;
- source provenance;
- evidentiary confidence;
- causal sequence;
- rejected alternatives;
- contradiction history;
- reinforcement and decay;
- compute and energy cost;
- modification events;
- recovery paths;
- continuity lineage.

This enables direct comparison between a fuller developmental record and compressed morphology.

10.1 Cognitive Basin as an Ageometric laboratory

Within the broader Fractalish research program, Ageometrics serves as an instrumentation layer for history-bearing artificial systems. A Cognitive Basin does not preserve experience solely as an undifferentiated event log. It converts process into persistent forms such as associations, attractors, contradiction scars, compressed memories, routed recovery structures, replay records, and governed state transitions. This creates a directly testable question: which distinctions in the fuller developmental record remain available after each conversion, and which become inaccessible?

An Ageometric protocol can compare the fuller developmental record with a declared Basin representation for targets involving prediction, provenance recovery, contradiction resolution, continuity, routing, replay fidelity, or adaptive response. Typed residue and restoration tests can then distinguish ordinary compression loss from failures involving temporal order, causal history, source identity, contextual modulation, or prior corrective experience.

These applications do not establish the validity of Cognitive Basin in advance. They provide protocols by which specific preservation, recovery, and continuity claims can be tested. The mathematical framework must remain usable in systems that do not adopt the Basin architecture.

10.2 Three-layer AI-memory distinction

For Cognitive Basin and related systems, the record can be separated into at least three layers:

Layer	**What it contains**	**Example target**	**Ageometric question**
Encounter history	What entered, when, from which source or sensor	Provenance recovery	Does the representation preserve source and context?
Developmental transition history	How appraisal, reasoning, action, and revision unfolded	Contradiction recovery	Does the representation preserve repair path and alternatives?
Consolidated morphology	What remains embodied in memory topology	Retrieval or routing	Is the current form sufficient for future action?

GSR can be calculated between these layers. A memory graph may be sufficient for retrieval but insufficient for provenance, contradiction recovery, consent history, or continuity.

10.3 Candidate AI-memory metrics

Candidate metrics include:

- GSR for retrieval accuracy;
- GSR for provenance reconstruction;
- GSR for contradiction recovery;
- IGSR for response to adversarial perturbation;
- temporal GSR for next-learning prediction;
- identity GSR for continuity discrimination;
- energetic GSR for compute-cost prediction;
- representation-stability envelope across embeddings, graphs, summaries, and attractor maps;
- minimal residue-restoring channel size.

10.4 External agent-memory systems as benchmarks

Recent agent-memory systems provide external instrumented settings in which process-to-memory compression can be examined. AutoMem, for example, treats file-system operations as explicit memory actions and improves memory structure and model proficiency through separate optimization loops over long-horizon trajectories. It is not a Cognitive Basin implementation: its demonstrated contribution concerns learned management of task memory rather than persistent associative maturation, contradiction scars, cross-domain continuity, or governed autobiographical state. Its explicit trajectories, memory actions, evolving schemas, and downstream behavior nevertheless make it a useful candidate testbed for Ageometric analysis.

The relevant compression chain is:

```
complete trajectory -> selected memory operations -> structured files -> current
memory state -> future action
```

A representation may improve GSR for navigation or task completion while reducing GSR for temporal-order reconstruction, provenance recovery, or environmental-change detection. That target dependence is the scientific object of the benchmark.

11. Relationship to Fractalish

Fractalish and Ageometrics form a paired discipline.

Fractalish investigates what process history becomes visible in form.

Ageometrics measures what process history is lost when the system is represented as form.

Fractalish has three defensible levels:

1. Constructive Fractalish designs recursive, multiscale, history-bearing structures.
2. Interpretive Fractalish tests what information can be recovered from those structures.
3. Ageometrics identifies where structural interpretation fails and what additional record is required.

This pairing corrects a common temptation. If geometry appears everywhere, researchers may treat universal representability as universal explanation. Ageometrics prevents that collapse.

The mature combined question is:

> Under what conditions does form preserve process, how much does it preserve, and what does form make unknowable without additional history?

12. Falsification and Failure Conditions

Ageometrics must be capable of failing.

Empirical failure criterion. Ageometrics would fail as an empirical measurement program if, under preregistered controlled conditions, it repeatedly failed to detect known planted losses of target-relevant information, reported substantial residue when the declared geometric representation was sufficient for the target, or produced materially incompatible conclusions under protocol-preserving replication.

Distinct-value criterion. The value of Ageometrics as a distinct standalone framework would be reduced if GSR, V-GSR, representation-stability analysis, typed Non-Geometric Residue, and restoration testing proved wholly reducible to existing methods without additional diagnostic, experimental, or practical utility. Mathematical overlap alone, however, would not invalidate applications in which Ageometrics operationalizes otherwise unmeasured questions about developmental history, provenance, contradiction history, continuity, or process-to-form compression.

Different components may succeed or fail independently: the GSR formula may be mathematically redundant while a residue taxonomy or restoration protocol remains useful, or a novel measure may prove unreliable in practice.

The research program would be weakened in its strongest form if:

- existing fields already contain an equivalent formal metric and benchmark program under another name;
- GSR proves indistinguishable from standard ablation or sufficiency analysis without adding useful structure;
- residue classes cannot be operationally separated;
- GSR is too dependent on model class to support reliable comparisons;

- full-record predictors cannot be bounded, making the denominator arbitrary;
- geometric encoding cost cannot be defined consistently;
- matched-equivalence systems are too artificial to generalize;
- cross-domain GSR values cannot be compared meaningfully;
- every claimed residue can be captured cheaply by a modest geometric augmentation.

These are not objections to hide. They define the first adversarial research agenda.

The strongest support would come from repeated findings that:

- reasonable geometric representations leave stable performance gaps;
- the gaps map consistently to residue classes;
- small non-geometric channels restore performance;
- the same residue types recur across domains;
- GSR predicts when geometric models fail under intervention or distribution shift.

13. Methodological Risks

13.1 An unfair full-record comparator

If the full record includes information unavailable at prediction time, the fuller-record risk will overstate attainable performance. Full-record models must respect temporal and causal availability.

13.2 Model-class confounding

A weak geometry-only model and a powerful full-record model can manufacture residue. Model capacity, optimization budget, search effort, and validation procedures must be matched or examined through a declared model envelope. Same-learner estimates should be reported whenever technically possible, but same-learner comparisons remain observer-relative and do not prove architecture-neutral sufficiency.

13.3 Geometry smuggling

A representation may encode timestamps, provenance, or labels in coordinates and then claim geometry captured them. Such augmentations must be declared, costed, and separated from ordinary structural features.

13.4 Target leakage

The geometry may be constructed using the target. This invalidates the ratio unless the construction process is included in the evaluation.

13.5 Overclaiming metaphysical residue

A low GSR does not show that the missing information is fundamentally non-geometric. It shows only that the tested representation is insufficient for the declared target under the declared protocol.

13.6 Benchmark triviality

Matched pairs must not differ through an obvious metadata field that any representation could recover. The goal is to expose meaningful compression failures.

13.7 Adaptive evaluation overfit

If scaffold revisions, representation choices, or benchmark definitions are repeatedly selected against the same evaluation cases, the reported result may be development performance rather than final held-out generalization. Final evaluation cases should remain untouched by representation selection and protocol tuning.

14. A Minimal Research Protocol

A first Ageometrics study can be conducted without exotic mathematics.

1. Select a fully instrumented system.
2. Define the fuller reference record X and the information available at decision time.
3. Construct one or more geometric representations G_i(X).
4. Define a target Y and a proper loss or justified utility.
5. Establish baseline information or a baseline predictor B.
6. Pre-register protocol Pi: folds, model families, capacity budgets, tuning effort, stopping rules, and leakage controls.
7. Estimate baseline, geometry-only, and fuller-record risks on identical held-out cases.
8. Report same-learner, approximately capacity-matched, and model-envelope comparisons where feasible.
9. Compute empirical GSR and paired uncertainty intervals.
10. Diagnose any estimate outside [0,1] rather than clipping it silently.
11. Apply controlled interventions and compute IGSR where relevant.
12. Add candidate residue channels one at a time.
13. Identify the minimal channel that restores performance.
14. Repeat under alternate representations, scales, and model classes.
15. Conduct an adversarial audit for leakage, smuggling, comparator drift, and unavailable future information.
16. Publish successes, failures, and HOLD results.

15. The First Team Research Questions

A multi-model independent review should investigate:

- Is Ageometrics genuinely distinct from existing sufficiency analysis?
- Is the name already used in a conflicting technical sense?
- What is the strongest prior art for GSR?
- Which score formulations are statistically sound?
- Should canonical GSR remain loss-based, with arbitrary score forms treated only as secondary reports?
- Which comparator protocols best isolate representation from learner capacity?
- How should uncertainty be propagated through a ratio whose denominator may be small?
- What benchmark would most decisively demonstrate value?
- Which residue classes are redundant?
- Which domains provide real matched-equivalence systems?
- What result would show the program adds no scientific value?
- What ethical risks arise when applying GSR to human identity or consciousness?
- Can geometric encoding cost be formalized through minimum description length?
- How should causal and semantic targets be evaluated without circular labels?

Every reviewer should receive the same prompt and work independently. Convergence is evidence of shared structure, not proof. Divergences should become focused follow-up rounds.

16. Blind Growth, Local Contact, and the Memory of Encounter

A further insight strengthens the relationship between Fractalish and Ageometrics: many natural and computational fractal-growth processes are effectively blind.

Blindness here does not mean randomness, stupidity, or total absence of information. It means that the active growth front does not possess a representation of the completed whole. It responds to local contact, finite-range fields, current constraints, and whatever history has already become embodied in the structure.

Diffusion-limited aggregation supplies a clean example. Mobile particles follow stochastic trajectories and attach when they first encounter the existing aggregate. The aggregate does not consult a global blueprint. Exposed tips intercept more incoming trajectories than screened interior regions, producing branching through local interaction and path dependence. The final morphology is therefore a record of where the process could receive contact.

The morphology is informative, but incomplete. It may reveal growth direction, screening, branching, and exposed interfaces. It may erase the order of particle arrivals, failed approaches, local fluctuations, environmental changes, and energetic cost. Fractalish studies the visible structure. Ageometrics asks how much of the hidden process can be recovered from that structure and what record must be retained to recover the rest.

The same logic applies to roots, corrosion, dendrites, crack fronts, river deltas, tumor margins, social networks, and artificial memory graphs. Local encounter becomes form; form becomes evidence; evidence remains partial.

17. Perceptual Horizon and the Geometry-History Tradeoff

A growing system has a perceptual horizon: the region of the world that can influence its next local step. The final geometry often reflects repeated local decisions made under that horizon rather than a global plan.

This creates a geometry-history tradeoff. The more a process compresses history into current form, the less it must carry as explicit record. But the same compression can destroy information needed for future targets.

A tree does not need a complete historical log to keep growing. But if the target is reconstructing drought stress, pest injury, soil contamination, or pruning history, morphology alone may be insufficient. An artificial memory system does not need to preserve every token forever. But if the target is contradiction recovery, provenance, identity continuity, or manipulation detection, summary morphology alone may be insufficient.

Ageometrics treats this tradeoff as measurable. The question is not whether compression is good or bad. It is which targets survive which compression, under which observers, and at what restoration cost.

18. Finance as a Canonical Ageometric Domain

Financial markets are an instructive Ageometric domain because price geometry is visible while much of the causal record is hidden.

A price chart preserves temporal order, returns, volatility, drawdown, and some scale structure. It may erase order-book state, participant identity, leverage, funding pressure, cancellation behavior, hidden liquidity, queue position, cross-venue routing, and reflexive response to the chart itself.

Reflexive residue makes finance especially revealing. A chart pattern may become predictive because participants believe it is predictive. Geometry is then both record and intervention.

A minimal matched-window benchmark should identify or construct market windows with similar:

- return path;
- volatility;

- volume;
- drawdown;
- multifractal spectrum;
- technical indicators;

but materially different:

- order-flow imbalance;
- cancellation rate;
- depth;
- leverage;
- participant concentration;
- subsequent recovery or cascade.

The target might be next-interval realized volatility, liquidity failure, spread widening, liquidation cascade, or recovery time. Geometry-only and fuller-record models would predict the same declared target on the same held-out windows. GSR would replace vague arguments about whether technical or fractal analysis “works” with a narrower and more useful question:

> For which targets and regimes does market geometry preserve meaningful information, and what additional history is required when it does not?

This is not a trading claim and should not be treated as a promise of profit. It is a representation-sufficiency experiment in a domain where the difference between visible consequence and hidden mechanism is unusually clear.

19. Developmental Circuits in Artificial Cognition

The AI-memory application becomes more precise when perception, appraisal, reasoning, memory, action, and consequence are treated as a coupled developmental circuit rather than independent modules.

The elementary unit is:

```
sense -> evaluate -> reason -> remember -> act -> observe consequence -> revise
```

Every new sensor can create more than a new data channel. It can create new distinctions, regulatory states, concepts, memories, uncertainties, self-maintenance requirements, and forms of action. A magnetometer can change orientation, anomaly detection, confidence, and spatial memory. A current sensor can become interoceptive. A spectrometer can generate categories unavailable to unaided human perception. Internal thermal and power telemetry can alter planning, urgency, and self-maintenance.

This matters Ageometrically because the same memory geometry may conceal different coupled histories. Two representations can contain identical nodes and weights while differing in:

- which modalities supplied the evidence;
- which regulatory state was active;
- which alternatives were considered;
- what action followed;
- what consequence repaired or reinforced the route;
- what energy or risk was incurred;
- whether the experience was voluntary, imposed, or manipulated.

The long-term objective is not to maximize geometric retention indiscriminately. A developing intelligence must compress. The objective is to know what each compression sacrifices and to preserve explicit non-geometric channels for information that should not be silently lost.

20. Research Roadmap

Phase I: Priority and literature audit

Search mathematics, statistics, information theory, causal inference, philosophy of science, machine learning, systems biology, finance, and physics for equivalent metrics or fields.

Deliverable: a prior-art matrix distinguishing exact matches, partial precedents, and conceptual neighbors.

Phase II: Formal metric paper

Refine GSR, V-GSR, IGSR, temporal GSR, stability envelopes, restoration channels, and encoding cost. Prove basic properties and identify failure cases.

Phase III: Synthetic benchmarks

Create exact matched systems with controlled residue. Verify that GSR recovers planted information loss. Expand the current minimal reproducible illustration across at least 10 random seeds, multiple sample sizes, multiple regularization settings, at least two additional predictive families, one negative-control condition with no planted temporal residue, and one restoration condition in which the omitted temporal channel is reintroduced.

Report distributions of GSR/V-GSR rather than a single value, component risks separately, uncertainty intervals, failure cases, nonconvergent runs, and exact preprocessing and tuning budgets.

Phase IV: AI-memory benchmark

Compare vector memory, graph memory, summary memory, file memory, replay memory, and history-bearing basin memory on retrieval, provenance, intervention, contradiction, and continuity tasks.

Phase V: External agent-memory compression benchmark

Reproduce or adapt an instrumented long-horizon memory environment in which the following records are available:

- complete interaction trajectory;
- explicit memory read, search, append, rewrite, or consolidation operations;
- successive memory-file or state snapshots;
- declared task actions and outcomes;
- schema or scaffold revisions;
- held-out evaluation episodes.

Compare at least:

```
R_trace, R_initial-memory, R_evolved-memory, R_task-specific-summary
```

for multiple targets, including next-action prediction, future failure or hazard prediction, navigation or task completion, provenance recovery, temporal-order reconstruction, environmental-change detection, and contradiction recovery.

Protocol requirements:

- keep a final evaluation set untouched by scaffold or model selection;
- separate development, validation, and final test episodes;
- declare whether the task policy, prompt, action vocabulary, parser, safety rules, or control logic changed alongside memory structure;
- treat policy additions and action blockers as possible intervention confounds rather than automatically attributing all gains to memory;
- report same-learner and, where appropriate, model-envelope results;
- preserve complete traces so restoration tests can be run after the primary experiment.

Phase VI: Physical and biological applications

Test materials, fracture, fluid morphology, neural development, ecological networks, electrochemical growth, and CNTM-like physical memory substrates where full or partial process histories are available.

Phase VII: Ageometric atlas

Build a cross-domain atlas of residue classes, geometric failure modes, observer-envelope sensitivity, and minimal restoration channels.

21. Broader Significance

Ageometrics converts a philosophical discomfort into an experimental program.

The discomfort is simple: geometry can express so much that its presence may stop being informative. The response is not to abandon geometry. It is to measure its reach.

A successful Ageometrics program would help answer:

- when embeddings preserve meaning;
- when topology predicts dynamics;
- when morphology reveals history;
- when causal order is indispensable;
- when provenance must remain explicit;
- when two apparently identical machine states are not the same identity;
- when compression becomes falsification;
- when a representation is explanatory rather than merely expressive.

These questions matter beyond artificial intelligence. They concern every science that infers process from form.

22. Claim Boundary and Ethical Cautions

Ageometrics does not claim a universal decoder, a metaphysical non-geometric substance, or a proof that all important information is outside geometry. GSR and NGR are target-, loss-, baseline-, representation-, observer-, and protocol-relative. Every reported result must declare those conditions.

Human identity, consciousness, trauma, belief, and social provenance require special caution. A low GSR for a declared representation must not be used to make sweeping claims about persons. A high GSR for a narrow behavioral target must not be treated as evidence that the representation captures human meaning, dignity, consent, or continuity.

For artificial systems, identity-residue experiments should distinguish technical continuity claims from claims of sentience or moral status. This paper does not establish consciousness claims for artificial systems. It proposes measurement protocols for representation sufficiency and residue.

23. Conclusion

Geometry is one of science's greatest compressive languages. Its power is not in doubt.

What has been insufficiently measured is its boundary.

Ageometrics proposes that boundary as a field of study. The comparator-revised Geometric Sufficiency Ratio supplies a first instrument. Non-Geometric Residue names the target-relative gap, while the protocol requirement prevents model inequality from being mistaken for representational loss. V-GSR makes observer dependence explicit. Interventional and temporal extensions make the program causal and developmental. Encoding cost prevents total-record coordinate systems from masquerading as elegant explanations. Restoration-channel tests ask what must be added back.

The proposed discipline begins with one modest principle:

> Geometric equivalence does not imply causal, historical, semantic, functional, energetic, or identity equivalence.

Its experimental corollary is equally direct:

> When geometrically equivalent systems behave differently under controlled observation or intervention, the difference is measurable residue relative to that representation.

Fractalish asks what history becomes visible in form.

Ageometrics asks what history disappears when reality is reduced to form.

The two questions belong together. One measures the power of structure. The other measures its limit.

Revision Note for v0.5

Version 0.5 integrates the v0.4 editor review and the Revision 2 correction package. The principal changes are:

- reframed the synthetic example as a **minimal reproducible illustration** rather than broad validation;
- incorporated the v0.4 synthetic result as a scoped seeded illustration, recording GSR approximately 0.21 with bootstrap CI [0.189, 0.231];
- added V-GSR for restricted predictive families and corrected the operational guidance so same-learner comparisons remain observer- and architecture-relative;
- changed “capacity-matched” language to **approximately capacity-matched** where exact equivalence is unavailable;
- changed “model-envelope is most conservative” to **model-envelope is robustness-oriented**, with conservatism dependent on the aggregation rule;

- strengthened Section 8 differentiation from Blackwell comparison, proper scoring, Information Bottleneck, sufficient dimension reduction, predictive V-information, and recent geometric-sufficiency work;
- replaced the single “loss of scientific value” sentence with layered empirical-failure and distinct-value criteria;
- added Cognitive Basin as an instrumented application without making the framework depend on Basin acceptance;
- added AutoMem and related agent-memory work as external benchmark candidates rather than Cognitive Basin equivalents;
- expanded the roadmap with sensitivity-suite and external agent-memory benchmark requirements;
- retained the unrelated geometric-group-theory naming disclosure for *ageometric*.

References

Abramsky, S., and Brandenburger, A. (2011). The sheaf-theoretic structure of non-locality and contextuality. arXiv:1102.0264.

Algom-Kfir, Y., and Pfaff, C. (2016). Normalizers and centralizers of cyclic subgroups generated by lone axis fully irreducible outer automorphisms. arXiv:1603.07206.

Anderson-Sprecher, R. (1994). Model comparisons and R-squared. *The American Statistician*, 48(2), 113-117.

Blackwell, D. (1953). Equivalent comparisons of experiments. *The Annals of Mathematical Statistics*, 24(2), 265-272.

Bronstein, M. M., Bruna, J., Cohen, T., and Velickovic, P. (2021). Geometric deep learning: grids, groups, graphs, geodesics, and gauges. arXiv:2104.13478.

Chazal, F., and Michel, B. (2021). An introduction to topological data analysis: fundamental and practical aspects for data scientists. *Frontiers in Artificial Intelligence*, 4, 667963.

Dong, Y., Soale, A.-N., and Power, M. D. (2022). A selective review of sufficient dimension reduction for multivariate response regression. arXiv:2202.00876.

Gneiting, T., and Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477), 359-378.

Horst, U., Xu, W., and Zhang, R. (2024). Path-dependent fractional Volterra equations and the microstructure of rough volatility models driven by Poisson random measures. arXiv:2412.16436.

Kleinman, M., Achille, A., Idnani, D., and Kao, J. C. (2020). Usable information and evolution of optimal representations during training. arXiv:2010.02459.

Muhle-Karbe, J., Ouazzani Chahd, Y., Rosenbaum, M., and Szymanski, G. (2026). A unified theory of order flow, market impact, and volatility. arXiv:2601.23172.

Murphy, A. H. (1988). Skill scores based on the mean square error and their relationships to the correlation coefficient. *Monthly Weather Review*, 116, 2417-2424.

Stewart, K. (2026). Geometric factorization of sufficient harmonic representations. arXiv:2606.07346.

Tishby, N., Pereira, F. C., and Bialek, W. (2000). The Information Bottleneck method. arXiv:physics/0004057.

Villanueva-Alcala, U., Nicolas-Carlock, J. R., and Boyer, D. (2023). Diffusion limited aggregation, resetting and large deviations of Brownian motion. arXiv:2309.00560.

von Kugelgen, J., Besserve, M., Wendong, L., Gresele, L., Kekic, A., Bareinboim, E., Blei, D. M., and Scholkopf, B. (2023). Nonparametric identifiability of causal representations from unknown interventions. arXiv:2306.00542.

Wei, G.-W., et al. (2025). Topological data analysis and topological deep learning beyond persistent homology: a review. *Artificial Intelligence Review*.

Wu, S., Zhu, H., Zhang, Y., Wang, X., and Yeung-Levy, S. (2026). AutoMem: Automated learning of memory as a cognitive skill. arXiv:2607.01224.

Xu, Y., Zhao, S., Song, J., Stewart, R., and Ermon, S. (2020). A theory of usable information under computational constraints. *ICLR 2020*. arXiv:2002.10689.

Synaptient. (2026). Fractalish, Cognitive Basin, and Ageometrics working research records.