Copied Consensus Is Not Objectivity

An Information-Theoretic Measure of Independent Recoverability

Josh Farrow

Draft manuscript - May 2026

Status: conceptual and simulation validation; not yet empirical validation on real-world data.

Abstract

Redundancy is a familiar ingredient in objectivity: quantum Darwinism explains classical objectivity as redundant environmental records of pointer states; statistics corrects correlated samples through effective sample size; and social epistemology has long distinguished independent testimony from copied consensus. This paper does not claim to rediscover that redundancy matters. It proposes a portable operational measure for a narrower problem: distinguishing effective objectivity from mere copied consensus. We define effective objectivity as target-relative independent recoverability. A fragment subset is adequate when it recovers a specified target within a tolerance. The smallest generally adequate subset size, k*, is combined with an effective independent fragment count, N_eff, estimated from target-conditioned residual dependence, to define FI_eff = max(0, log2(N_eff/k*)). This quantity measures the number of bits of effectively independent adequate readouts supported by a system. Synthetic stress tests show four desired behaviors: logarithmic capacity scaling with independent fragments, collapse under copied redundancy, decay under record drift, and target relativity under false consensus. The measure is not a truth oracle: it can assign high effective objectivity to a false target if fragments independently recover that target. Its role is therefore diagnostic, not metaphysical. It quantifies when a claim, state, record, or representation has become independently recoverable, and it exposes cases where apparent public agreement is only duplicated dependence.

Keywords: objectivity; redundancy; information theory; effective sample size; quantum Darwinism; copied consensus; drift; representation learning; partial information decomposition

1. Introduction: the known idea and the narrower claim

A statement can be known privately without being objective in the public, operational sense. A single perfect witness, a single perfect sensor, or a single excellent model can recover a target state, but the target has not thereby become multiply available to independent observers. Conversely, a claim can appear overwhelmingly public because it has been copied many times, even when all copies trace back to one dependent source. The difference between these cases is not simply accuracy. It is independent recoverability.

The background idea is not new. Quantum Darwinism treats classical objectivity as the redundant encoding of information about pointer states into many environmental fragments [1,2]. Effective sample size in statistics corrects nominal sample counts for correlation and dependence [5]. Partial information decomposition studies the information that multiple sources share about a target [3,4]. Social epistemology and jury-theorem traditions have long distinguished independent testimony from dependent repetition [7]. The contribution proposed here is therefore deliberately bounded: not a new metaphysics of objectivity, but an operational order parameter for effective objectivity, defined as target-relative independent recoverability.

The title claim - copied consensus is not objectivity - is the paper's diagnostic center. A public record can look redundant while being informationally centralized. A media narrative, clinical documentation pattern, model ensemble, or scientific claim can appear widely confirmed while its fragments remain residual copies of the same upstream process. The proposed measure is designed to separate naive redundancy from effective redundancy.

2. Related work: redundancy, objectivity, and effective independence

2.1 Quantum Darwinism and functional information

Quantum Darwinism provides the closest physical analogue. In that framework, an observer does not need access to the system itself; small fragments of the environment carry redundant information about selected pointer states. Recent functional-information work makes this operational by counting fragments that are adequate, within a tolerance, to recover pointer information [1]. This paper borrows the fragment-adequacy logic while separating it from specifically quantum assumptions.

2.2 Redundancy in information theory

Partial information decomposition (PID) asks how the information that several sources provide about a target decomposes into redundant, unique, and synergistic components. Recent redundancy-bottleneck work frames redundancy as information that predicts the target without revealing which source supplied it [3]. The present paper is less ambitious than a full PID: it does not decompose all multivariate information. It defines an operational adequacy threshold and asks how many effectively independent adequate readouts exist.

2.3 Effective sample size and effective rank

Statistics already contains the essential correction: correlated samples count as fewer independent samples. Stan's reference manual describes effective sample size as playing the role in the Markov-chain central limit theorem that independent draws play in the standard central limit theorem [5]. Effective rank and participation-ratio methods similarly estimate the effective dimensionality of a spectrum [6]. We use this family of ideas to correct fragment counts for copied residual dependence.

2.4 Representation convergence and learned objectivity

The Platonic Representation Hypothesis argues that increasingly capable models and modalities may converge toward shared representations of an underlying reality [8]. For the present purpose, that hypothesis becomes testable only after replacing representation similarity with adequate readout: do independently trained systems become fragments from which the same target can be recovered? Similar embeddings are not enough; random geometry can produce false positives.

3. Formal definition

Let X be a target observable and let F_1, ..., F_N be fragments: sensors, environmental subsystems, model readouts, clinical record fields, market channels, witnesses, or other partial records. A subset S of fragments is adequate if it recovers the target within tolerance delta:

$$A_{\delta}(S) = 1\{ I(X; F_S) \geq (1 - \delta) I_{max}(X) \}.$$

In fully observed information-theoretic settings $I_{max}(X)$ may be $H(X)$. In applied settings it can be replaced by a calibrated score: held-out accuracy, semantic agreement with a reference, clinical adequacy, predictive log score, or another target-specific readout metric.

$$p_{\delta}(k) = P(A_{\delta}(S) = 1 \mid |S| = k).$$

The minimal generally adequate fragment size is:

$$k^*_{\delta,q} = \min \{ k : p_{\delta}(k) \geq q \}.$$

If N_{eff} denotes the effective number of independent fragments, effective objectivity is:

$$FI_{eff} = \max(0, \log_2(N_{eff} / k^*_{\delta,q})).$$

The max operation encodes a semantic floor: if the system cannot support even one effectively independent adequate readout, its redundant objectivity is zero. A single perfect fragment gives perfect private recoverability but zero redundant objectivity because $N_{eff}/k^* = 1$.

3.1 Target-conditioned effective independence

Fragments that agree because they observe the same target should not be penalized. Fragments that agree because they copy the same residual source should be penalized. For this reason, dependence is estimated after conditioning on the target or an accepted proxy for it. Let C be the residual correlation matrix among fragment readouts after subtracting the target-conditioned mean. With eigenvalues $\lambda_1, ..., \lambda_N$, we estimate effective independent fragment count by a participation-ratio form:

$$N_{eff} = (\sum_i \lambda_i)^2 / \sum_i \lambda_i^2.$$

When residuals are independent, eigenvalues are approximately flat and N_{eff} is close to N. When all fragments are copies of one residual source, one eigenvalue dominates and N_{eff} approaches 1. This is not the only possible correction, but it is transparent, computable, and aligned with effective-sample-size logic.

3.2 Naive redundancy, effective redundancy, and copied consensus

$$FI_{naive} = \max(0, \log_2(N / k^*)).$$

$$\text{Copy gap} = FI_{naive} - FI_{eff}.$$

The copy gap is a diagnostic: it is small when nominal fragments are effectively independent and large when apparent redundancy is inflated by copied residual dependence. A high FI_naive with low FI_eff is the mathematical signature of copied consensus.

4. Conceptual validation and boundary behavior

A measure of effective objectivity should behave sensibly at the boundaries. Table 1 summarizes core cases.

Scenario	k*	N_eff	FI_eff	Interpretation
One perfect fragment only	1	1	0	Perfect private recoverability; zero redundant public objectivity
100 perfect independent fragments	1	100	6.64	High effective objectivity
100 copied fragments of one reporter	1	1	0	Copied consensus rejected
100 independent weak fragments needing coalitions of 10	10	100	3.32	Moderate objectivity; adequate readouts require groups
No target information	>N or undefined	any	0	No adequate readout exists
False target independently recovered	1	100	6.64 relative to false target	Objectivity is target-relative, not truth-guaranteeing

Table 1. Boundary behavior of the proposed measure. FI_eff measures redundant recoverability relative to a specified target; it does not validate the target itself.

The measure also has a declared limitation for high-order synergy. If only the entire system can recover the target, then k* is close to N_eff and FI_eff is near zero. This does not mean the system contains no information. It means the information is not redundantly recoverable from independent fragments. The measure is therefore best suited to public, fragment-accessible objectivity, not purely holistic encodings.

5. Information-theoretic derivation

The derivation is short. Mutual information quantifies how much a fragment subset reduces uncertainty about X. Adequacy converts a continuous information quantity into an observer criterion: the subset either has enough information for the intended readout or it does not. Sampling subsets of size k yields p_delta(k), the probability that a randomly chosen k-fragment subset is adequate. The minimal k that makes adequacy typical is k*.

If N_eff independent fragments are available and k* fragments are needed to form one adequate readout, then the system can support approximately N_eff/k* effectively independent adequate readouts. Information theory measures multiplicity in bits by a base-2 logarithm. Thus each doubling of the number of independent adequate readouts adds one bit:

$$FI_{eff} = \log_2(\text{number of effectively independent adequate readouts}) = \log_2(N_{eff}/k^*).$$

The logarithm is therefore not a cosmetic choice. It expresses objectivity as the bit-length of a choice among independent adequate records. This predicts capacity scaling: when k* is stable, FI_eff should grow linearly with log2(N_eff).

6. Synthetic stress tests

The experiments reported here are synthetic. They are not empirical validation on real clinical, social, or physical datasets. Their purpose is narrower: to check whether the measure behaves correctly under controlled generative conditions, including failure modes where a naive redundancy measure should fail.

Simulation	Final score	Peak FI_eff	Peak FI_naive	Median N_eff
bayesian witnesses	1.000	3.974	4.000	31.402
local phase patches	1.000	4.975	5.000	31.400
narrative channels	1.000	4.975	5.000	31.403
negative correlated copy	0.995	0.013	5.000	1.009
negative no signal	0.495	0.000	0.000	31.389
record standardization	1.000	4.974	5.000	31.388
sensor network	1.000	4.976	5.000	31.396

Table 2. Additional simulation summary. Positive systems generate independent adequate readouts; negative controls test no signal and copied reports.

6.1 Capacity scaling

When independent fragments are individually adequate, the measure predicts FI_eff approximately log2(N_eff). In the capacity-scaling stress test, peak effective objectivity followed FI_eff = 0.980 log2(N) + 0.048, with R^2 = 0.9999. This result is partly expected from the definition, but it verifies the implementation and the bit-scale interpretation.

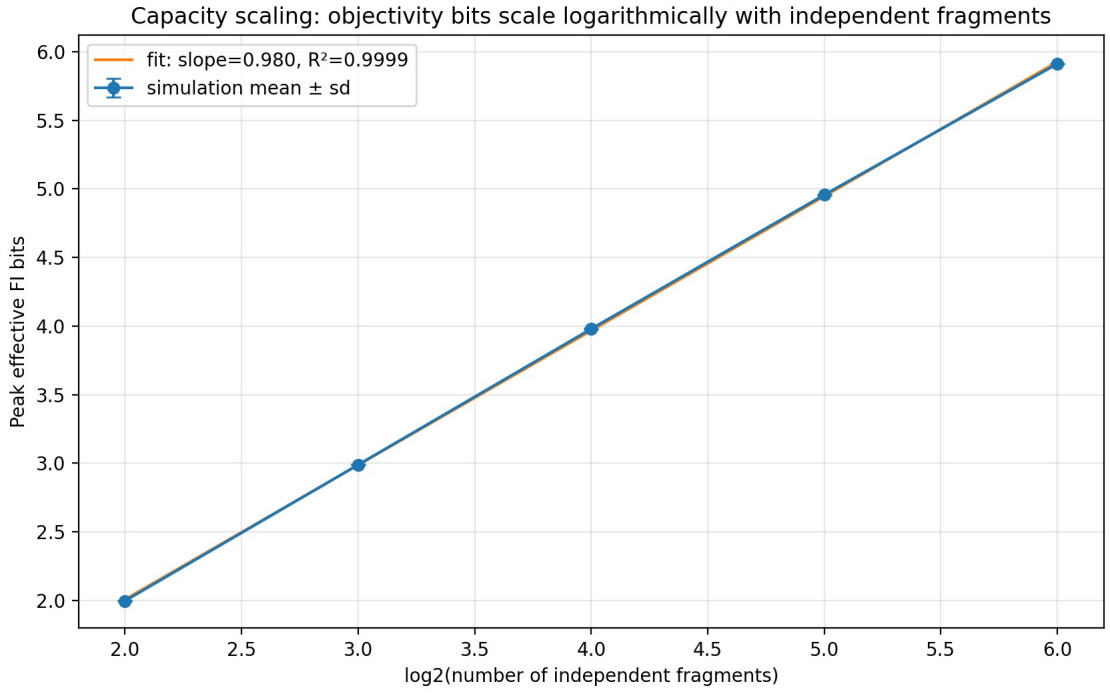

Figure 1. Capacity scaling. Peak effective objectivity scales approximately linearly with log2 of the number of independent adequate fragments.

6.2 Copied consensus gradient

The central test varies residual copy strength rho from 0 to .99. Naive redundancy remains high because every copied fragment is individually adequate. Effective objectivity collapses because N_eff falls from approximately 45.89 to 1.02. At rho = .99, mean FI_naive remains approximately 5.58 bits while mean FI_eff falls to approximately 0.03 bits. This is the cleanest demonstration of the title claim.

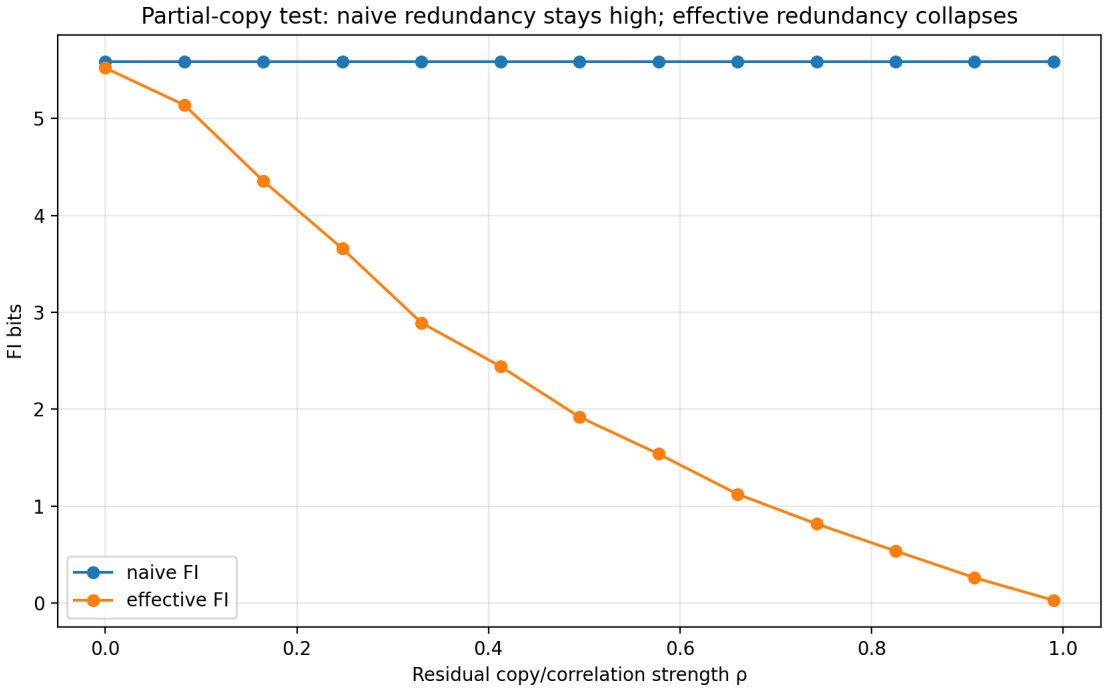

Figure 2. Copied consensus test. Naive redundancy remains high under copying, while effective redundancy collapses as residual dependence increases.

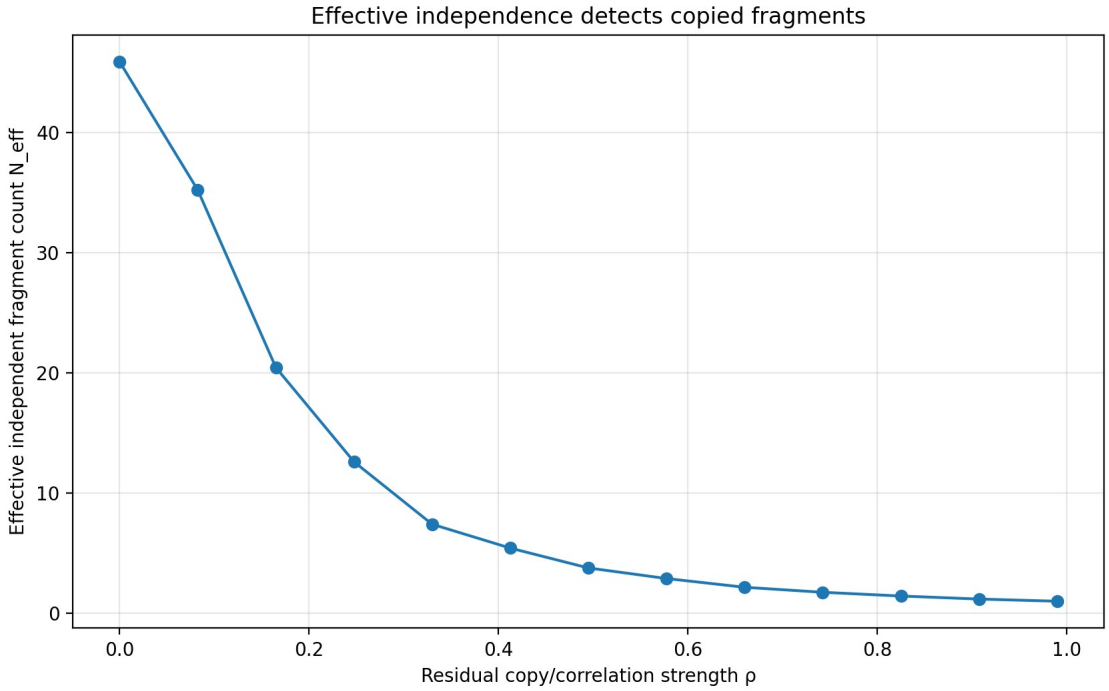

Figure 3. Effective independent fragment count falls as residual copying increases.

6.3 Objectification and drift

A useful measure should not only rise. It should decay when records stop independently recovering the target. In the drift simulation, independent recoverability rises early, plateaus, and later falls as signal quality decays and residual correlation increases. This supports defining drift as the negative time derivative of effective objectivity:

$$\text{Drift}(t) = -\, d\,\text{FI}_{\text{eff}}(t) / dt.$$

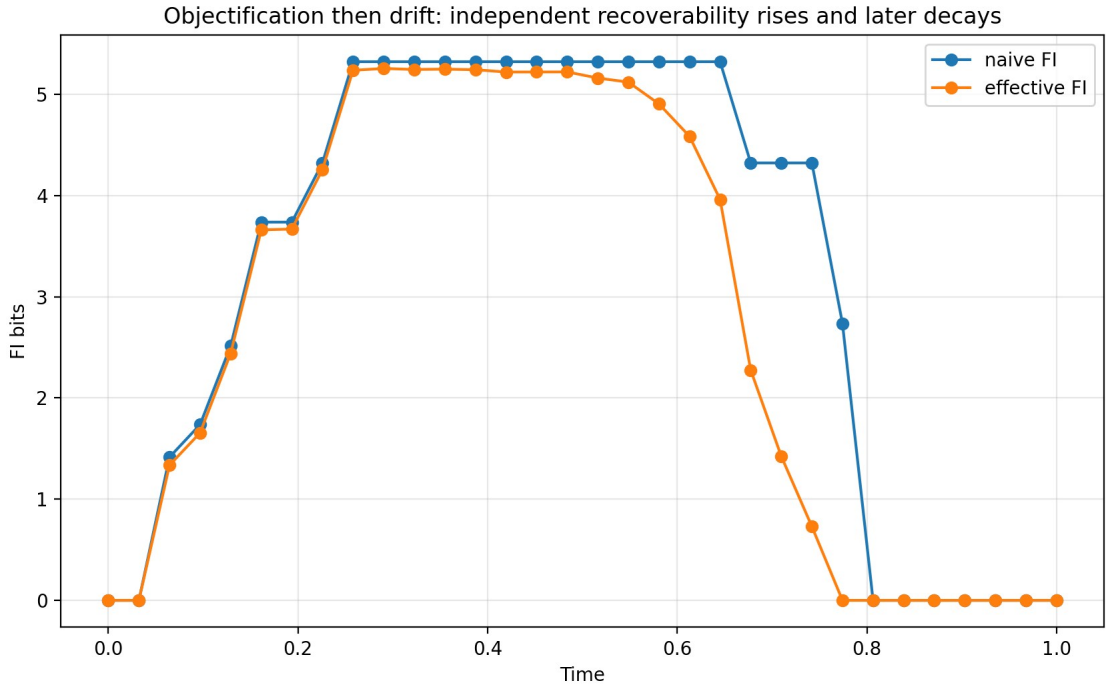

Figure 4. Objectification followed by drift. Effective objectivity rises as independent readouts form and falls when recoverability decays.

6.4 False consensus and target relativity

The false-consensus simulation is a warning rather than a defect. Fragments transition from recovering a true target to recovering an inverted/captured target. FI_eff for the true target falls while FI_eff for the false target rises. Therefore FI_eff is not a truth oracle. It measures independent recoverability relative to the specified target. Truth-sensitive use requires external target validation, causal grounding, or adversarial reference channels.

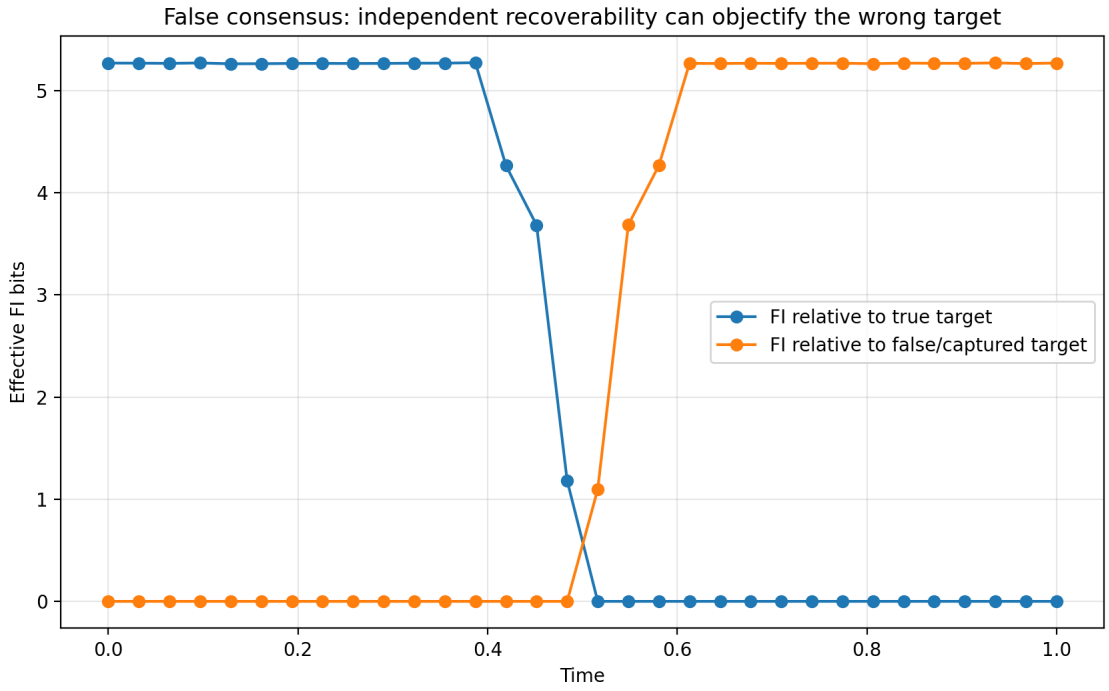

Figure 5. False consensus. Independent recoverability can objectify the wrong target; the measure must be paired with target validation.

6.5 Threshold robustness and curve collapse

The qualitative transition survives changes in tolerance delta and adequacy quantile q. Across additional positive simulations and prior pilot curves, normalized curve-collapse diagnostics produced mean RMSE 0.073 and worst RMSE 0.117 to the mean normalized curve. These numbers should not be overinterpreted; the generative systems were designed to form independent readouts. The useful result is that the same measure exposes consistent transition structure and rejects no-signal and copied-copy controls.

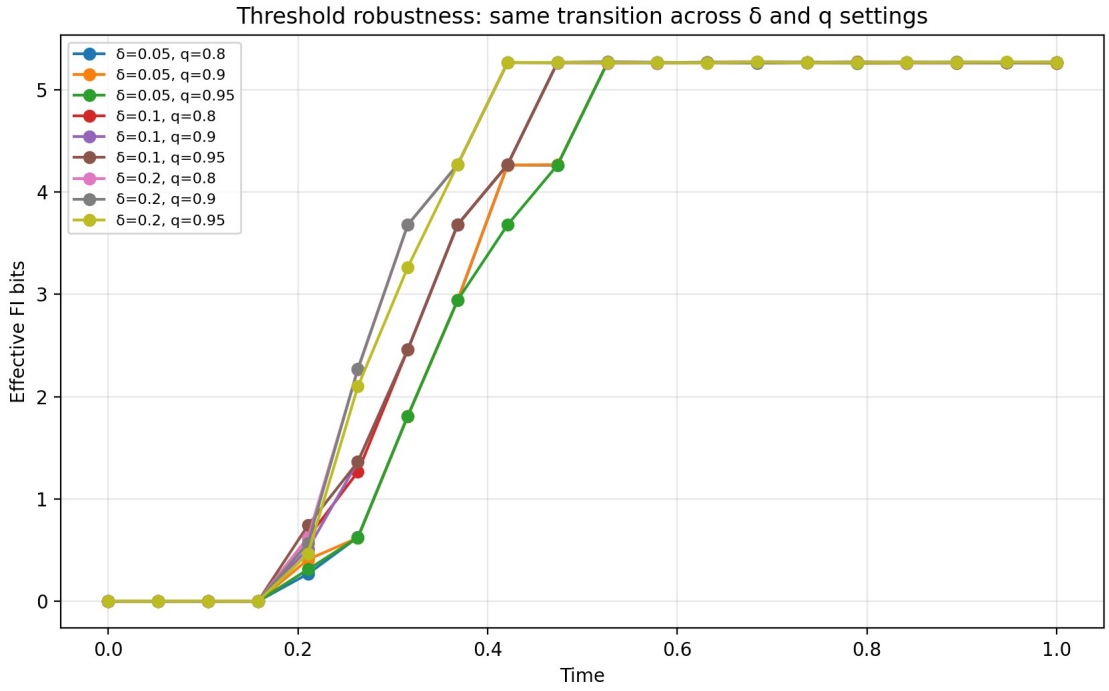

Figure 6. Threshold robustness across delta and q settings.

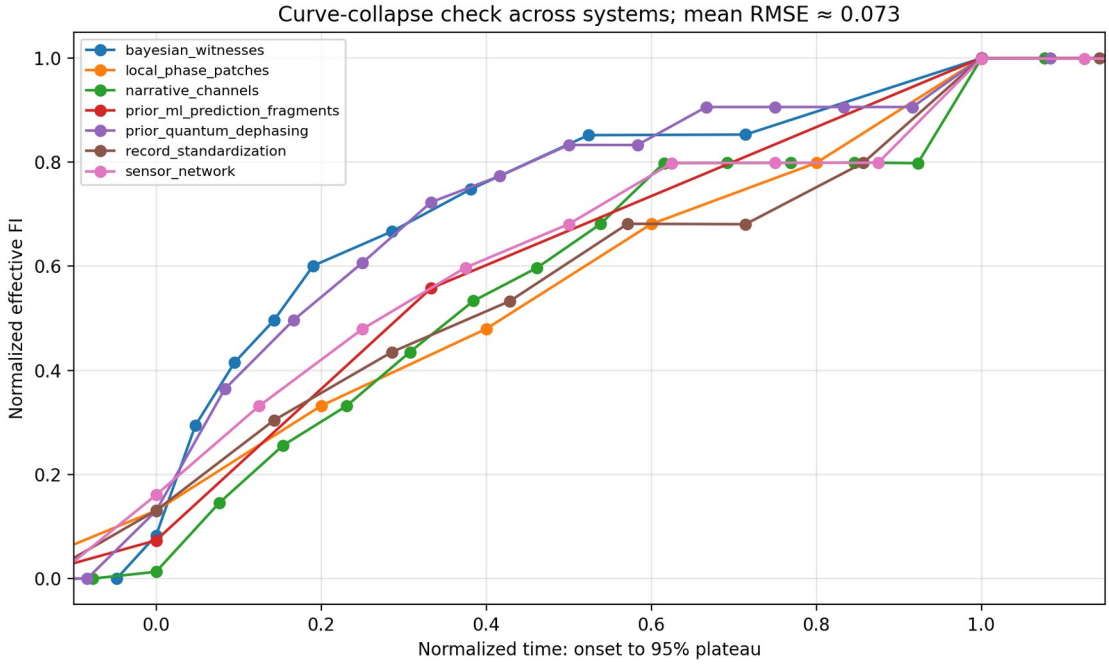

Figure 7. Normalized curve-collapse check across positive systems and prior pilot curves.

7. What the measure adds beyond known redundancy

The paper's novelty is not the claim that redundancy matters. Its contribution is the operational distinction between nominal redundancy and effective independent recoverability. In particular, it gives a quantitative diagnostic for copied consensus: high FI_naive, low FI_eff.

This distinction matters in domains where public agreement can be manufactured by dependence. A news narrative can be repeated by many outlets but trace to one source. A clinical record can contain multiple fields that appear to agree because they are copied from one template. A model ensemble can appear robust because models share a training artifact. A market signal can appear to exist across channels because one channel mechanically propagates into the others. The measure asks whether fragments remain adequate after correcting for residual dependence.

It also separates private accuracy from public objectivity. A single excellent model, clinician, sensor, or document can be correct, but FI_eff remains low if no independent fragments recover the same target. That is not a denial of truth; it is a distinction between truth, private knowledge, and publicly redundant recoverability.

8. Candidate empirical applications

Clinical and ABA documentation

Define X as a clinically adjudicated learner/session state or audit-quality note. Treat fragments as structured data, narrative note sections, target graphs, supervision comments, plan fields, and caregiver reports. Falling FI_eff would indicate that the clinical state is becoming less independently recoverable across the record ecology.

Representation learning

Define X as a label, teacher embedding, held-out property, or causal factor. Treat independently trained models as fragments. Compare representation similarity with adequate readout. The hypothesis is that FI_eff should rise during learning, while embedding similarity alone may create false positives.

Narrative and market systems

Define X as a regime label, event outcome, or later-validated state. Treat market, options, news, social, search, and policy channels as fragments. A large copy gap would identify hype or consensus built through dependent propagation rather than independent evidence.

Scientific and institutional epistemology

Define X as an experimentally adjudicated claim. Treat labs, instruments, datasets, protocols, or replications as fragments. FI_eff offers a formal distinction between independent replication and repeated analysis of the same upstream data-generating process.

9. Limitations

Target relativity: FI_eff measures recoverability relative to a specified target. A false target can be highly objectified if independent fragments recover it. Truth requires external validation.

Target conditioning can be hard: The N_eff correction assumes that the target or a reliable proxy is available for conditioning. If the target is latent, noisy, or inferred from the same fragments, the correction can under- or over-estimate independence.

Synergistic systems: Systems in which only the whole configuration recovers the target will receive low FI_eff. This is appropriate for redundant objectivity but not for measuring all distributed information.

Threshold dependence: The thresholds delta, q, and the adequacy score must be chosen for the domain. Robustness checks reduce but do not remove this calibration problem.

Synthetic validation only: The present simulations validate logic and edge behavior. They do not prove empirical usefulness. Real-world fragments are messy, nonstationary, adversarial, and often target-constructing rather than target-reading.

10. Falsifiable predictions

Test	Prediction	Failure condition
Copied-source test	Datasets with duplicated upstream sources should show high FI_naive but low FI_eff.	If FI_eff remains high after known copying is introduced.
Drift test	Documentation or representation drift should appear as declining FI_eff before or alongside human-rated inconsistency.	If FI_eff does not track later adjudicated drift better than naive redundancy.
Learning test	Model ensembles should show FI_eff onset during learning when readouts become adequate.	If FI_eff rises in random-label or no-signal controls equally.
Consensus test	Copy-driven narratives should show larger copy gaps than independently corroborated events.	If copy gap cannot distinguish these cases under blinded labels.

Table 3. Falsifiable empirical predictions.

11. Conclusion

Copied consensus is not objectivity. Objectivity, in the operational sense proposed here, is not the number of times a claim is repeated, but the number of effectively independent adequate readouts of a specified target. FI_eff = max(0, log2(N_eff/k*)) compresses this idea into a bit-scale order parameter. It rewards independent recoverability, punishes copied residual dependence, distinguishes private accuracy from public redundancy, and makes drift measurable as loss of recoverability. Its strongest value may be diagnostic: when FI_naive is high but FI_eff is low, a system has the appearance of objectivity without the independence that objectivity requires. The measure remains target-relative and must be paired with truth-validation when truth matters. But as a portable way to separate robust distributed evidence from copied agreement, it offers a practical information-theoretic tool.

References

1. Tank, A. B. (2025). Functional Information in Quantum Darwinism: An Operational Measure of Objectivity. arXiv:2509.17775.
2. Blume-Kohout, R., & Zurek, W. H. (2005). A simple example of quantum Darwinism: redundant information storage in many-spin environments. Foundations of Physics, 35, 1857-1876.
3. Kolchinsky, A. (2024). Redundancy as Information Bottleneck. Entropy / arXiv:2405.07665.
4. Faes, L., et al. (2025). Partial Information Rate Decomposition. Physical Review Letters.
5. Stan Development Team. Stan Reference Manual: Effective Sample Size. https://mc-stan.org/docs/reference-manual/analysis.html
6. Roy, O., & Vetterli, M. (2007). The Effective Rank: A Measure of Effective Dimensionality. EUSIPCO.
7. Condorcet, M. de. (1785). Essay on the Application of Analysis to the Probability of Majority Decisions.

8. Huh, M., Cheung, B., Wang, T., & Isola, P. (2024). The Platonic Representation Hypothesis. ICML / arXiv:2405.07987.

9. Recanatesi, S., et al. (2022). A scale-dependent measure of system dimensionality. Patterns, 3(8).

Appendix A. Reproducibility note

The simulation outputs used in the figures are the CSV and PNG artifacts generated in the accompanying synthetic validation runs. Core files include capacity_scaling.csv, partial_copy_gradient.csv, objectification_then_drift.csv, false_consensus_target_relativity.csv, threshold_robustness.csv, and more_sims_summary.csv. The experiments are deterministic conditional on the listed seeds in the generation scripts used during this session.