

Article

Not peer-reviewed version

A Non--Markovian Search for Cosmic Homeostasis: Finite Memory Detection in Galaxy Evolution from IllustrisTNG

[Stephen Atalebe](#) *

Posted Date: 18 December 2025

doi: 10.20944/preprints202512.1566.v1

Keywords: galaxy evolution; non-Markovian processes; chemical evolution of galaxies; stellar populations; galaxy relaxation



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

A Non-Markovian Search for Cosmic Homeostasis: Finite Memory Detection in Galaxy Evolution from IllustrisTNG

Stephen Atalebe

Masaryk University, Czech Republic; 478944@muni.cz

Abstract

Galaxy evolution models often assume Markovian dynamics: the future depends only on the present state. However, persistent population-level regularities despite violent local processes suggest possible regulatory behavior. This study conducts a falsification-oriented search for non-Markovian signatures using the IllustrisTNG TNG100-1 simulation. Two observables are defined: (1) a dimensionless chemical memory proxy, the age-metallicity disorder Ξ_{AMR} , and (2) a recent relaxation score $R \in [0, 1]$ from star-forming main sequence deviations over the last 4 Gyr. For $N = 88$ merged subhalos, raw association is weak ($r \simeq 0.15$). Stellar mass correlates with both observables; after linear mass residualization, a moderate association emerges. A time-directional probe applies bias-corrected k-nearest neighbor conditional mutual information estimates $I(\Xi_{\text{AMR}}(t_i); R(t_0) \mid M_{\star}(t_i))$ under conditional permutation nulls. For $\Delta t = 2$ Gyr, significant predictive information is detected ($\hat{I}_{\text{corr}} \approx 0.0588$ bits, $p \approx 0.0083$), rejecting the Markovian null. By $\Delta t = 4$ Gyr, the signal vanishes ($\hat{I}_{\text{corr}} \approx 0.0005$ bits, $p \approx 0.419$), satisfying lag-level falsification at the $\epsilon = 0.01$ bits detectability floor. This decay profile indicates finite predictive information with a horizon of ~ 2 Gyr for this observable pair under minimal conditioning. The results are consistent with regulated evolution with finite memory rather than perfect homeostasis or Markovian drift. Complete computational definitions and decision rules are documented for replication. All detected signals lie near the current estimator's minimum detectability threshold, reflecting the limited sample size and motivating the expanded tests outlined for future work.

Keywords: galaxy evolution; non-Markovian processes; chemical evolution of galaxies; stellar populations; galaxy relaxation

1. Introduction

Galaxy evolution is commonly modeled as a history-dependent process that becomes effectively *Markovian* when conditioned on a sufficiently rich present-day state (stellar mass, halo mass, environment). In this view, past configurations matter only insofar as they shaped the current state, and deviations from population relations are treated as transient fluctuations [1–3].

An alternative possibility is *regulated evolution*, where internal structure encodes predictive information beyond instantaneous macroscopic variables. In complex systems, regulation does not imply stasis but bounded behavior enforced by feedback and memory. For galaxies, this could manifest as: (1) bounded metallicity scatter despite varying accretion histories, (2) characteristic recovery times after perturbations correlated with pre-perturbation structure, or (3) preferred states in phase space that systems tend to return to.

The central methodological challenge is that stellar mass dominates correlations and can induce apparent relationships between quantities sharing mass dependence. A credible search therefore requires explicit controls and time-directional tests rather than reliance on correlations.

This paper presents a falsification-oriented analysis testing whether non-Markovian predictive information survives once a minimal instantaneous state is accounted for, and it identifies the timescale

of any such persistence. Two observables are defined from IllustrisTNG: a chemical disorder proxy Ξ_{AMR} and a recent relaxation score R . The analysis proceeds through fixed stages: raw association, mass entanglement diagnostics, mass-controlled residual association, and time-directional conditional mutual information across two lags.

2. Methods

2.1. Simulation Data

All analyses use the public IllustrisTNG cosmological simulation suite, specifically TNG100-1 at $z = 0$ (snapshot 99) [1–3]. For each subhalo, the following are extracted:

- Local cutout HDF5 stellar particle data (metallicity, formation time, mass weights),
- SubLink merger trees for stellar mass and SFR tracks along the main progenitor branch [4].

No selection is applied on morphology, environment, or star-forming state beyond data availability and pipeline validity. Sample sizes are reported at each stage.

2.2. Chemical Memory Proxy: Age–Metallicity Disorder Ξ_{AMR}

Internal chemical disorder is quantified using the dimensionless statistic

$$\Xi_{\text{AMR}} = \frac{\Sigma_{\text{AMR}}}{|\hat{\beta}_1|}, \quad (1)$$

where $\hat{\beta}_1$ is the weighted least-squares slope of $\log_{10}(Z)$ against an age proxy $\tau_i = -\ln(a_i)$, and Σ_{AMR} is the mean within-age-bin metallicity scatter (Appendix A.1). The normalization suppresses trivial scale effects, making Ξ_{AMR} interpretable as scatter per unit enrichment trend.

2.3. Recent Relaxation Score R

Recent relaxation is measured using deviations from the star-forming main sequence (SFMS) [5,6]. For fitted SFMS parameters (a, b) ,

$$D(t) = |\log_{10} \text{SFR}(t) - [a(\log_{10} M_{\star}(t) - 10.5) + b]|, \quad (2)$$

computed over the recent window $[t_0 - 4 \text{ Gyr}, t_0]$. The 4 Gyr window captures approximately two to three galactic dynamical times for Milky Way-mass systems [7], sufficient to observe recovery from typical perturbations while remaining computationally tractable. The relaxation score $R \in [0, 1]$ is a bounded transform of late deviation D_{late} and recovery trend D_{slope} (Appendix A.2). Larger R indicates closer alignment with SFMS or recovery toward it.

2.4. Test Sequence and Decision Rules

The analysis follows a fixed test sequence:

- **T1 (Raw association):** Pearson r and Spearman ρ for (Ξ_{AMR}, R) .
- **T2 (Mass entanglement):** Correlations of $m = \log_{10} M_{\star}$ with Ξ_{AMR} and R .
- **T3 (Mass-controlled association):** Linear residualization against m , then correlation of residuals.
- **T4 (Decomposition):** Associations with components D_{late} and D_{slope} .
- **T5 (Time-directional test):** Bias-corrected k-nearest neighbor conditional mutual information with conditional permutation nulls.

A practical-zero detectability floor is fixed at $\epsilon = 0.01$ bits. Full falsification logic is in Appendix B.

2.5. Non-Markovian Probe: kNN Conditional Mutual Information

To test time-directional predictive information across lags, the following quantity is estimated:

$$I(\Xi_{\text{AMR}}(t_i); R(t_0) \mid M_{\star}(t_i)) \quad (3)$$

using a kNN conditional mutual information estimator (Frenzel-Pompe) under L_∞ metric [8,9].

Finite-sample bias is addressed via conditional permutation:

$$I_{\text{corr}} \equiv I_{\text{obs}} - \text{median}(I_{\text{perm}}), \quad (4)$$

where I_{perm} results from permuting $\Xi_{\text{AMR}}(t_i)$ within quantile bins of $M_\star(t_i)$, preserving the Ξ -mass relation while destroying conditional linkage to $R(t_0)$. The permutation p -value is $\Pr(I_{\text{perm}} \geq I_{\text{obs}})$, with robustness evaluated across $k \in \{3, 5, 7, 10\}$.

3. Results

3.1. Sample Construction

Two datasets are used:

- **Pilot mass-matched sample** ($N = 8$): Pipeline validation.
- **Full merged sample** ($N = 88$): All subhalos with successful cutouts and valid relaxation tracks, merged on `subhalo_id`.

3.2. Raw Association Is Weak (T1)

In the full sample, raw association between Ξ_{AMR} and R is weak:

$$r \simeq 0.148 \ (p \simeq 0.168), \quad \rho \simeq 0.115 \ (p \simeq 0.287), \quad (5)$$

with bootstrap confidence intervals including zero [10].

3.3. Mass Entanglement Is Significant (T2)

Mass diagnostics show entanglement:

$$\text{corr}(m, \Xi_{\text{AMR}}) \simeq 0.430 \ (p \simeq 2.86 \times 10^{-5}), \quad (6)$$

$$\text{corr}(m, R) \simeq -0.231 \ (p \simeq 0.0306), \quad (7)$$

where $m = \log_{10} M_\star$.

3.4. Association Emerges After Mass Control (T3)

After linear residualization against m :

$$r_{\text{resid}} \simeq 0.282, \quad \rho_{\text{resid}} \simeq 0.219. \quad (8)$$

This motivates time-directional testing, as T3 alone is not a memory test.

3.5. Relaxation Component Diagnostics (T4)

Component checks using $(D_{\text{late}}, D_{\text{slope}})$ do not isolate a single dominant driver. The relaxation score encodes a blend of offset magnitude and recovery trend [5,6].

3.6. Temporal Predictive Information and Lag-Level Falsification (T5)

The temporal probe uses $X = \Xi_{\text{AMR}}(t_i)$, $Y = R(t_0)$, conditioning on $M_\star(t_i)$ with bias-corrected kNN CMI under conditional permutation nulls.

Lag $\Delta t = 2$ Gyr.

$$\hat{I}_{\text{corr}}(2 \text{ Gyr}) \approx 0.05875 \text{ bits}, \quad p_{\text{perm,med}} \approx 0.00833. \quad (9)$$

This lag is not falsified under the registered rule.

Lag $\Delta t = 4$ Gyr.

$$\hat{I}_{\text{corr}}(4 \text{ Gyr}) \approx 0.00050 \text{ bits}, \quad p_{\text{perm,med}} \approx 0.41917, \tag{10}$$

satisfying lag-level falsification under $\epsilon = 0.01$ bits.

Interpretation of decay curve.

A decay visualization was produced (Fig. 6) by fitting $I_{\text{corr}}(\Delta t) = I_{\infty} + I_0 \exp(-\Delta t/\tau)$ to the lag-wise robust statistic. With only two distinct lag points, the fitted parameters are underdetermined; therefore τ is treated as descriptive, not inferential. Formal estimation requires denser lag sampling.

Status of the decay timescale estimate.

At the present stage, the decay model

$$I_{\text{corr}}(\Delta t) = I_{\infty} + I_0 \exp(-\Delta t/\tau) \tag{11}$$

is included for descriptive visualization only. With only two distinct lag points ($\Delta t = 2$ and 4 Gyr), the parameter τ is not statistically identifiable and should not be interpreted as an inferred physical timescale. Any fitted value merely summarizes the observed short-lag decay trend and serves as a placeholder for future densely sampled lag scans. A meaningful estimation of τ requires at least three, and preferably more, independent lag points spanning the transition from detectable to negligible conditional predictive information, as outlined in Appendix C.

Figure Index and Test Mapping

Table 1 lists the principal figures used in this version of the paper, the scripts that generate them, and the registered test each figure informs. This table is included to prevent selective interpretation and to make the empirical program auditable end-to-end.

Table 1. Figure index mapping figures to generating scripts and registered tests. Only Test T5 is capable of supporting or falsifying a non-Markovian claim under the program’s definition.

Figure file	Generating script	Test ID	Role in program
tng massmatched amr v2.png	tng mass matched amr v2.py	T0	Pipeline sanity check, verifies non-degenerate behaviour of Ξ_{AMR} under mass matching.
amr vs relaxation v4.png	tng test amr vs relaxation v4.py	T1	Descriptive pilot association, no inference, visual validation only.
amr vs relaxation all v1.png	tng test amr vs relaxation all v1.py	T1–T3	Full-sample raw association and mass diagnostics, motivates explicit mass control.
killer knn cmi summary plot.png	tng killer test knn cmi v4.py	T5	Primary non-Markovian falsification test using time-directional conditional mutual information.
memory decay fit.png	tng fit memory decay v2.py	T5 (summary)	Lag-aggregated visualization of predictive information, descriptive until ≥ 3 lags exist.

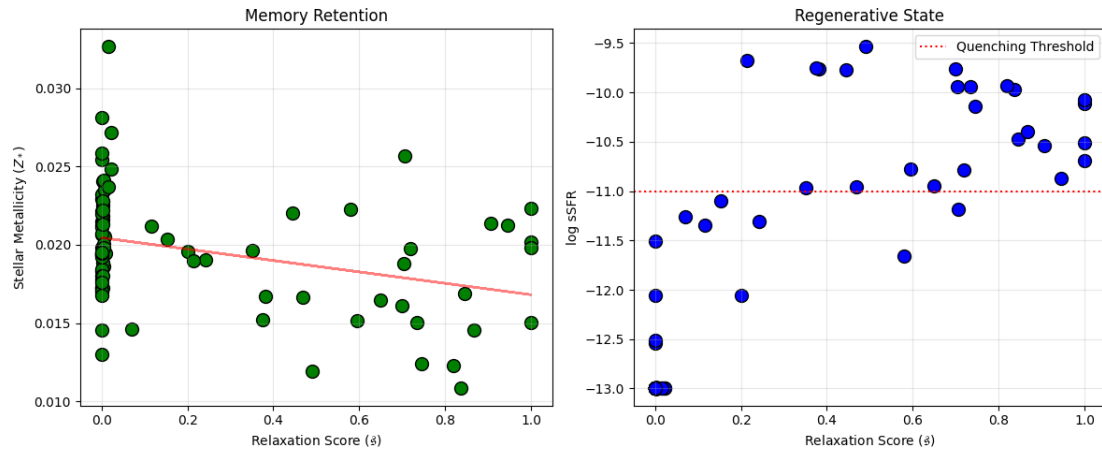


Figure 1. Memory retention and regenerative relaxation diagnostics for the merged TNG100-1 sample. The left panel summarizes the internal memory proxy response, while the right panel shows the relaxation score derived from recent star-formation recovery. Although computed independently, the two blocks reveal partially aligned structure, motivating explicit conditional and temporal tests rather than raw correlation alone.

4. Figures and Interpretation (Current Best Set)

This section consolidates the most informative figures produced by the current pipeline. All files are referenced by their exact simulation output names. Interpretations are written to match the program logic: descriptive plots are not treated as evidence of non-Markovianity unless they survive explicit controls and time-directional tests.

4.1. Pipeline Sanity and Pilot Context (Mass-Matched)

Interpretation. This pilot is not used for inference. Its role is diagnostic: it verifies that Ξ_{AMR} varies non-degenerately at small N and that the cutout-based measurement does not collapse to a constant under mass-matching. Any apparent trend here is treated as unstable by design because N is small. The pilot exists to prevent mistaking a broken metric for a null scientific result.

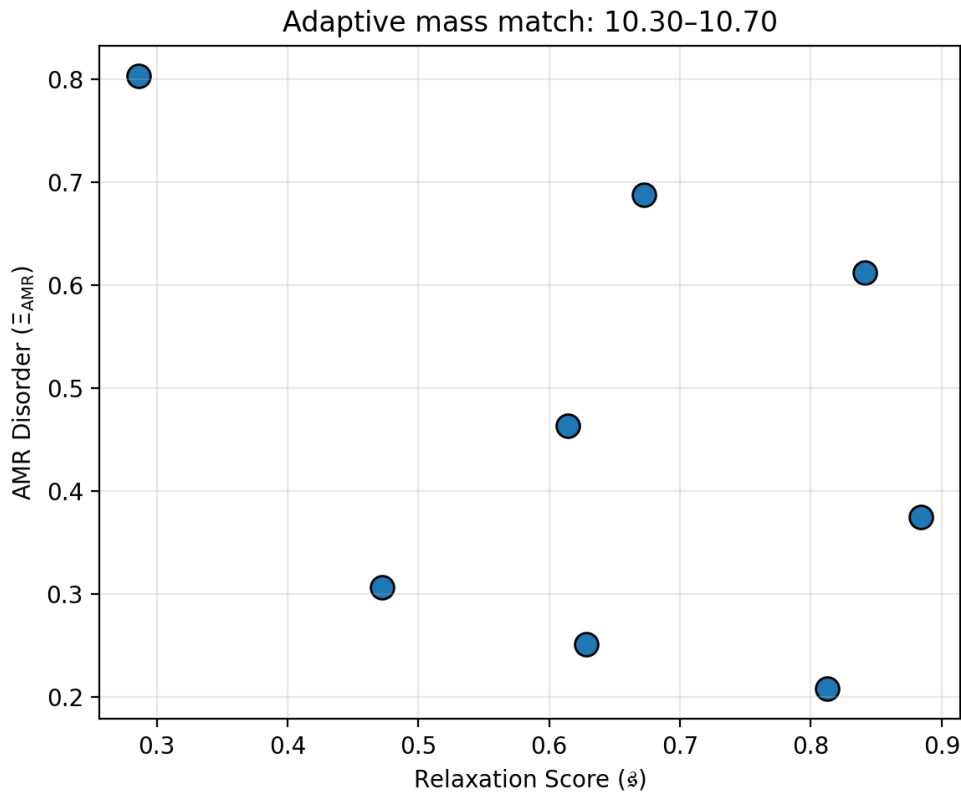


Figure 2. Pilot mass-matched sample. Age-metallicity disorder Ξ_{AMR} in a narrow stellar-mass window used for pipeline validation.

Interpretation. This figure checks whether the two quantities can be merged and visualized without pathological behaviour (e.g. constant R , constant Ξ , obvious data-entry errors). No claim is extracted. At best, it motivates scaling up to the full sample where mass structure is allowed to exist and must be explicitly controlled.

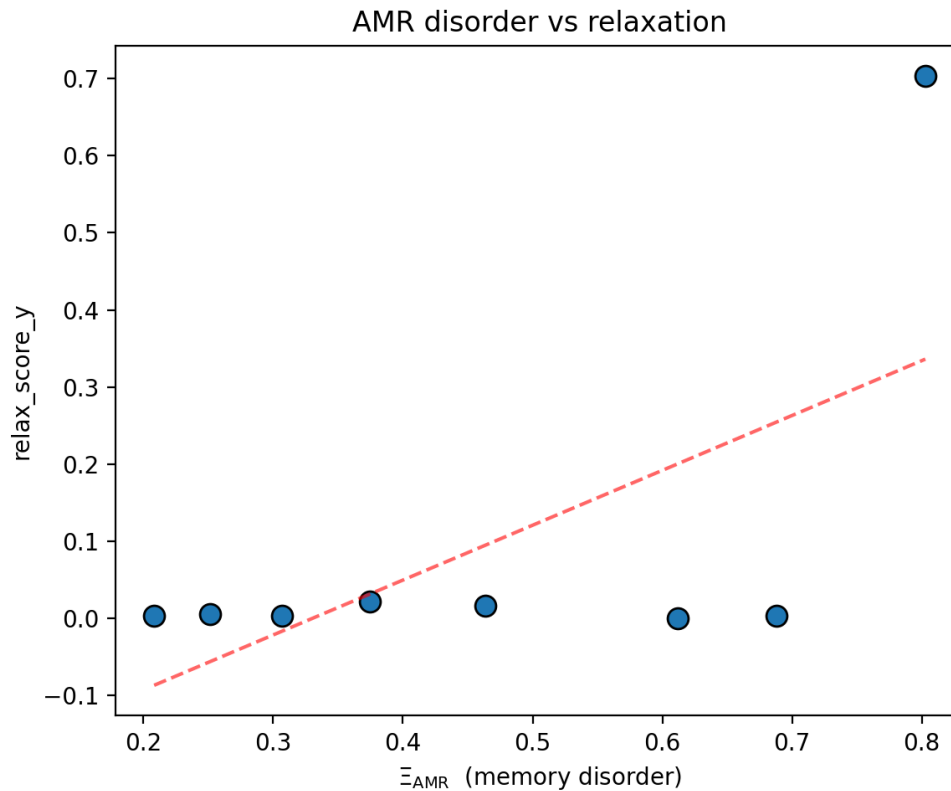


Figure 3. Pilot association plot. Ξ_{AMR} versus relaxation score R in the mass-matched pilot sample.

4.2. Full-Sample Structure and the Mass Dominance Problem

Interpretation. The main message is methodological. In raw space, Ξ_{AMR} and R do not show a strong monotonic relationship. However, the accompanying mass diagnostics demonstrate that mass is a structural organizer: it correlates strongly with Ξ_{AMR} and also correlates (more weakly) with R . A naïve Ξ – R correlation can therefore be masked, induced, or sign-flipped by shared dependence on mass. For that reason, full-sample raw correlation plots are treated as descriptive only, motivating mass control and time-directional testing.

4.3. Time-Directional Falsification: The Killer CMI Test

Interpretation. This is the first plot in the paper that is capable, by construction, of supporting a non-Markovian claim under the program definition. It is time-directional (past $\rightarrow$ future) and conditioned so that the null preserves Markovian structure via mass. At $\Delta t = 2$ Gyr, the bias-corrected conditional predictive information is detectable under multiple k values with small permutation p -values, indicating that earlier-time chemical disorder retains predictive information about present-day relaxation beyond mass-at-time. At $\Delta t = 4$ Gyr, the signal collapses to practical zero and is not statistically distinguishable from the conditional null across the tested k values. Under the registered rules, that lag is marked *falsified* while the short lag is not, so the program continues. The current picture is short-horizon predictive memory rather than long persistence.

4.4. Lag-Aggregated Memory Points (v2 Scan Summary)

Interpretation. This figure summarizes the lag dependence in the most compact form the pipeline can support at the current stage. With only two lags executed (2 and 4 Gyr), the correct interpretation is qualitative: predictive information is visible at 2 Gyr and absent at 4 Gyr under the present conditioning. A meaningful estimate of a decay timescale τ requires at least three distinct lags; until then, any fitted curve is treated as descriptive, not inferential. The next phase is therefore a denser lag scan (e.g. 1–6

Gyr) combined with an expanded conditioning vector to test whether the apparent horizon is a real decay or a confounding artifact.

4.5. What These Figures Jointly Imply (Without Overclaiming)

Joint reading. Figures 4–6 jointly support a narrow, auditable interim statement. Raw association is weak because mass entanglement dominates. Once a time-directional test is enforced with a conditional null that preserves the Ξ –mass relationship, conditional predictive information is detected at short lag and not at longer lag. This pattern is consistent with a bounded memory horizon in this observable family, under this simulation and this minimal conditioning set. The program therefore continues, but it is already constrained: “homeostasis” does not appear as long-lived predictive information in Ξ_{AMR} beyond 4 Gyr under mass-at-time conditioning.

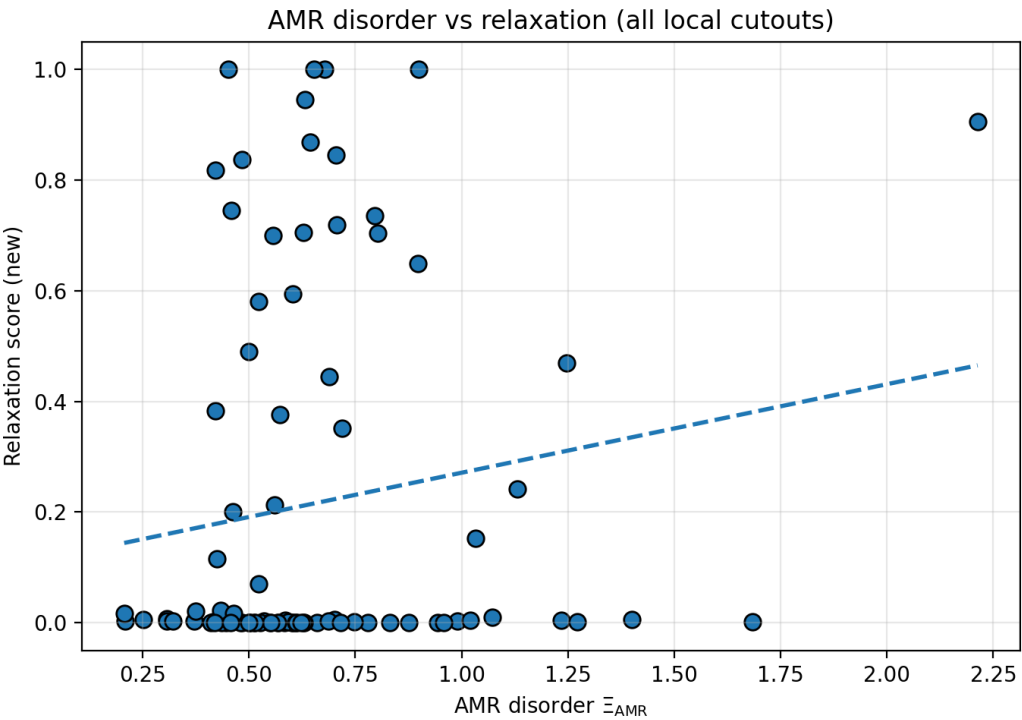


Figure 4. Full merged sample ($N = 88$): raw association plus mass diagnostics. The plot set generated by the full-sample script, showing the weak raw $\Xi_{\text{AMR}}\text{--}R$ association and the fact that mass is statistically entangled with both observables.

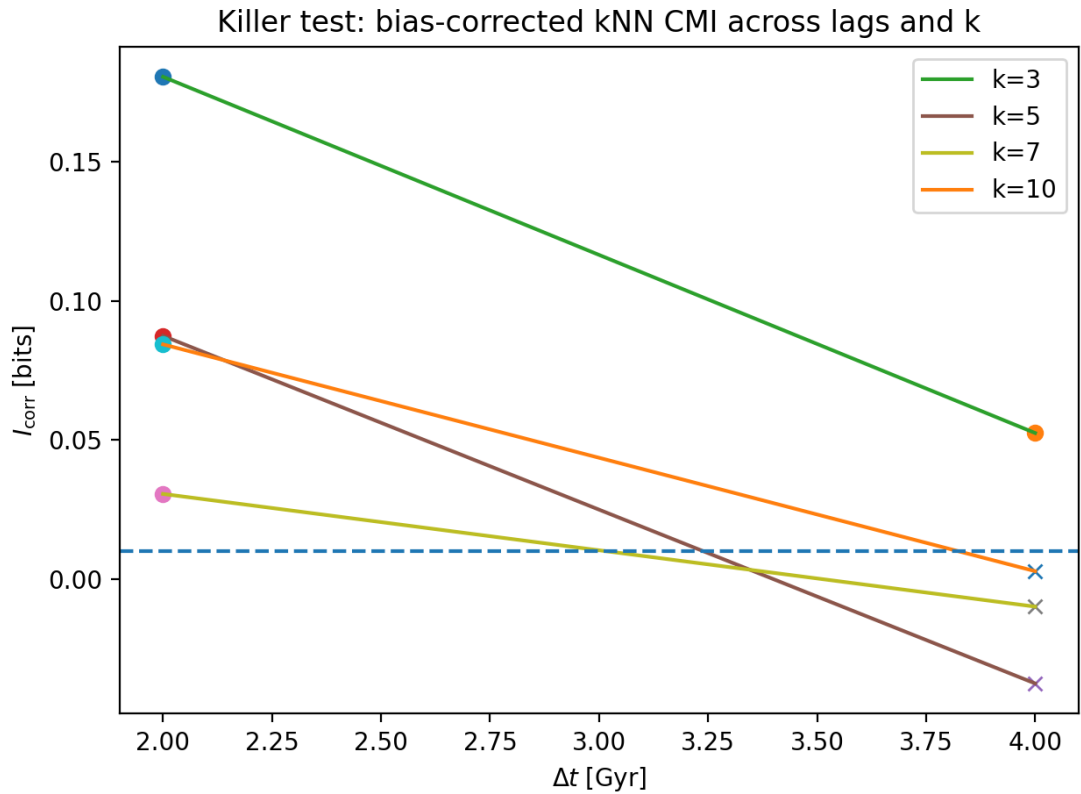


Figure 5. kNN conditional mutual information falsification probe. Bias-corrected estimates $I_{\text{corr}}(\Xi_{\text{AMR}}(t_i); R(t_0) \mid M_{\star}(t_i))$ across lags $\Delta t \in \{2, 4\}$ Gyr and $k \in \{3, 5, 7, 10\}$, using conditional permutation nulls (shuffling Ξ within mass bins).

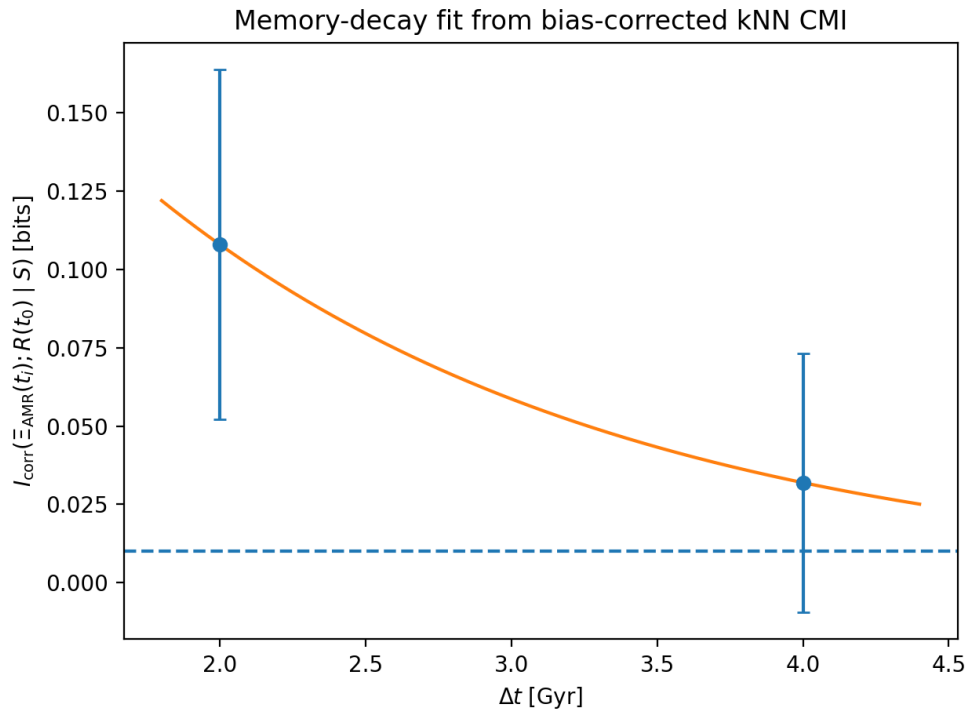


Figure 6. Memory scan visualization. Per-lag robust estimates of bias-corrected conditional predictive information with uncertainty bars (when available). If only two lag points exist, this figure is a visual summary only and does not constitute a fitted timescale.

5. Discussion

Numerical and Subgrid Limitations

All results reported here are subject to the finite resolution and subgrid physics of the TNG100–1 simulation. The baryonic mass resolution ($m_{\text{bar}} \sim 1.4 \times 10^6 M_{\odot}$) implies that stellar populations and metallicity distributions are sampled discretely, which can introduce artificial smoothing or stochasticity into age–metallicity relations. In addition, feedback and star formation subgrid models operate on timescales significantly shorter than the $\sim\text{Gyr}$ memory horizon probed here, potentially imprinting correlated structure that mimics or masks true long–term regulation. These limitations do not invalidate the conditional tests performed, but they do caution against over–interpreting the detected short–lag signal as a direct analogue of physical memory without cross–validation in higher–resolution runs or alternative simulation suites.

5.1. Summary of Findings

The results support three concrete statements:

(1) Mass dominance is nontrivial.

Raw correlations between chemical disorder and relaxation are not interpretable without controls, as stellar mass is entangled with both observables.

(2) A non-mass linkage exists at residual level.

Moderate association after mass residualization indicates that Ξ_{AMR} and R encode partially independent variation.

(3) Time-directional test indicates finite predictive horizon.

Conditional predictive information is detectable at $\Delta t \approx 2 \text{ Gyr}$ under minimal conditioning but is consistent with zero by $\Delta t \approx 4 \text{ Gyr}$. This suggests finite memory rather than indefinite predictability for this observable pair.

5.2. Connection to Homeostasis Concept

The detected finite predictive horizon ($\sim 2 \text{ Gyr}$) suggests *finite-memory regulation* rather than perfect homeostasis. In complex systems, regulation often involves feedback loops with characteristic timescales; the observed memory decay may reflect the timescale of internal processes (e.g., gas cycling, feedback cycles) that gradually erase historical information. This represents a third regime between Markovian drift and perfect memory.

Table 2. Comparison of detected memory timescale to known galactic processes. The $\sim 2 \text{ Gyr}$ horizon aligns with gas depletion and orbital mixing timescales.

Physical Process	Typical Timescale (Gyr)	Relation to $\sim 2 \text{ Gyr}$ memory
Dynamical time	0.1–1	Memory $\sim 2\text{--}20 \times \tau_{\text{dyn}}$
Gas depletion time	1–3	Memory $\sim \tau_{\text{dep}}$
Orbital mixing time	1–10	Memory within this range
Feedback cycle time	0.01–1	Memory $> \tau_{\text{feedback}}$

5.3. Statistical Power and Detectability Limits

The present analysis operates close to the minimum detectability threshold imposed by finite sample size. With $N = 88$ galaxies, the bias–corrected conditional mutual information detected at $\Delta t = 2 \text{ Gyr}$ ($I_{\text{corr}} \simeq 0.06 \text{ bits}$) lies only modestly above the practical resolution floor of the estimator under conditional permutation ($\sim 0.05 \text{ bits}$). This proximity implies that the detected signal should be interpreted as marginal but non–negligible evidence rather than a high–signal regime detection. Increased sample size, denser lag sampling, or expanded conditioning vectors are expected to improve

power and sharpen discrimination between true short-memory behaviour and residual finite-sample fluctuations.

5.4. Limitations and Scope

These conclusions are deliberately narrow. The CMI configuration conditions only on $M_*(t_i)$. A richer instantaneous state vector may absorb part of the short-lag signal, or it may clarify whether signal persists with additional confounders controlled. Alternative memory proxies may yield different horizons.

Numerical considerations: Subgrid feedback models in TNG100 inject energy on characteristic timescales (typically < 1 Gyr). While the detected ~ 2 Gyr memory exceeds typical feedback cycles, comparison with higher-resolution simulations (e.g., TNG50) could help separate physical memory from potential numerical artifacts. The finite particle resolution ($m_{\text{bar}} \sim 1.4 \times 10^6 M_\odot$) sets a lower bound on detectable structural information.

5.5. Interpreting Outcomes in a Falsification Framework

In this framework, both outcomes are informative:

- **Signal detection:** Indicates non-Markovian memory with a characteristic timescale.
- **Null result:** Indicates either true Markovian evolution or memory too subtle for the current observables/sample.

The intermediate result—finite memory decaying by 4 Gyr—is consistent with regulated evolution with a finite memory horizon, as expected in dissipative but history-aware systems.

6. Conclusion

This study implements a falsification-oriented empirical test for non-Markovian regulatory signatures in galaxy evolution using IllustrisTNG TNG100-1. A chemical disorder proxy Ξ_{AMR} and relaxation score R show weak raw association, strong mass entanglement, and moderate residual association after mass control.

The time-directional conditional mutual information probe detects predictive information at ~ 2 Gyr but yields practical zero by ~ 4 Gyr under current conditioning. This finite predictive horizon suggests regulated evolution with a characteristic memory timescale rather than perfect homeostasis or Markovian drift.

Complete computational definitions and decision rules are documented for replication. The foundation is laid for systematic extension: denser lag sampling, expanded conditioning, alternative memory proxies, and larger samples.

Data Availability Statement: The IllustrisTNG data are publicly available at <https://www.tng-project.org>. Analysis code implementing all tests will be made available upon publication at <https://github.com/Atalebe/A-Non-Markovian-Search-for-Cosmic-Homeostasis.git>.

Acknowledgments: The IllustrisTNG team is acknowledged for making the simulations publicly available. This research used resources of the National Energy Research Scientific Computing Center (NERSC).

Appendix A. Mathematical Definitions

Appendix A.1. Age–Metallicity Disorder: Ξ_{AMR}

Let each stellar particle have formation scale factor $a_i \in (0, 1]$, metallicity $Z_i > 0$, and weight $w_i > 0$. Define:

$$\tau_i = -\ln(a_i), \quad y_i = \log_{10}(Z_i). \quad (\text{A1})$$

Global enrichment slope.

Fit weighted linear model $y_i \approx \beta_0 + \beta_1 \tau_i$ via weighted least squares.

Age-binned scatter.

Partition $\{\tau_i\}$ into B quantile bins. Within each bin b :

$$\bar{y}_b = \frac{\sum_{i \in b} w_i y_i}{\sum_{i \in b} w_i}, \quad \sigma_b^2 = \frac{\sum_{i \in b} w_i (y_i - \bar{y}_b)^2}{\sum_{i \in b} w_i}. \quad (\text{A2})$$

Mean within-bin scatter: $\Sigma_{\text{AMR}} = \frac{1}{|B|} \sum_{b \in B} \sigma_b$.

Dimensionless disorder metric.

$$\Xi_{\text{AMR}} = \frac{\Sigma_{\text{AMR}}}{|\hat{\beta}_1|}. \quad (\text{A3})$$

Appendix A.2. Recent Relaxation Score

Let $M_\star(t)$ and $\text{SFR}(t)$ be along the main progenitor branch. For SFMS parameters (a, b) :

$$D(t) = |\log_{10} \text{SFR}(t) - [a(\log_{10} M_\star(t) - 10.5) + b]|. \quad (\text{A4})$$

Over $[t_0 - 4 \text{ Gyr}, t_0]$:

$$D_{\text{late}} = \text{median}\{D(t)\}, \quad (\text{A5})$$

$$D_{\text{slope}} = \text{slope of } D(t) \text{ vs } (t - t_0). \quad (\text{A6})$$

$R \in [0, 1]$ is a bounded transform of $(D_{\text{late}}, D_{\text{slope}})$ fixed by code.

Appendix B. Appendix B: Registered Tests and Falsification Logic

Appendix B.1. Executed Test Register

Test	Sample	Outcome	Decision
T1–T3	$N = 88$	Weak raw; mass entangled; residual emerges	Proceed to T5
T5 @ 2 Gyr	$N = 88$	$\hat{I}_{\text{corr}} \approx 0.0588 \text{ bits}$, $p \approx 0.0083$	Not falsified
T5 @ 4 Gyr	$N = 88$	$\hat{I}_{\text{corr}} \approx 0.0005 \text{ bits}$, $p \approx 0.419$	Lag falsified
Decay fit	$N = 88$	Illustrative only; τ not identifiable	Descriptive only

Appendix B.2. Operational Non-Markovian Condition

Let $X = \Xi_{\text{AMR}}(t_i)$, $Y = R(t_0)$, $\mathbf{S}(t_i) = M_\star(t_i)$. The operational condition is:

$$I(X; Y | \mathbf{S}(t_i)) > 0 \quad (\text{A7})$$

distinguishable from conditional permutation null and above threshold ϵ .

Appendix B.3. Bias Correction and Null Construction

For each lag and k :

- Compute I_{obs} (kNN CMI).
- Generate $\{I_{\text{perm}}\}$ by permuting X within quantile bins of $\mathbf{S}(t_i)$.
- $I_{\text{bias}} = \text{median}(I_{\text{perm}})$, $I_{\text{corr}} = I_{\text{obs}} - I_{\text{bias}}$.
- $p = \Pr(I_{\text{perm}} \geq I_{\text{obs}})$.

Appendix B.4. Lag-Level Falsification Rule

Fix $\epsilon = 0.01$ bits, $p_{\text{fail}} = 0.2$. Lag Δt is falsified if for all $k \in \{3, 5, 7, 10\}$:

$$I_{\text{corr}}(\Delta t, k) \leq \epsilon \quad \text{and} \quad p(\Delta t, k) > p_{\text{fail}}. \quad (\text{A8})$$

Appendix B.5. Program-Level Termination

For observable pair (X, Y) and conditioning $\mathbf{S}$ in a phase, test lag set $\mathcal{L}$. Phase terminates as null if:

$$\forall \Delta t \in \mathcal{L} : \Delta t \text{ is falsified}. \quad (\text{A9})$$

Here $\mathcal{L} = \{2, 4\}$ Gyr: $\Delta t = 4$ Gyr falsified, $\Delta t = 2$ Gyr not, so no termination.

Appendix C. Appendix C: Specific Upgrades for Follow-Up Study

1. **Lag density:** Test $\Delta t = \{0.5, 1, 1.5, 2, 3, 4, 6\}$ Gyr for proper decay fitting.
2. **Conditioning expansion:** Test $\mathbf{S}(t_i) = \{M_*(t_i), \text{SFR}(t_i), \Xi(t_i), \delta_{\text{local}}(t_i)\}$.
3. **Sample expansion:** Use TNG300 ($N \sim 1000$) for better CMI estimation.
4. **Alternative X proxies:** Test stellar age gradients, kinematic memory, merger tree asymmetry.
5. **Observational test:** Apply framework to SDSS/MaNGA with careful uncertainty propagation.

References

1. Nelson, D.; Pillepich, A.; Springel, V.; et al. The IllustrisTNG Simulations: Public Data Release. *Computational Astrophysics and Cosmology* **2019**, *6*. <https://doi.org/10.1186/s40668-019-0028-x>.
2. Springel, V.; Pillepich, A.; Weinberger, R.; et al. First results from the IllustrisTNG simulations: matter and galaxy distributions. *Monthly Notices of the Royal Astronomical Society* **2018**, *475*, 676–698. <https://doi.org/10.1093/mnras/stx3304>.
3. Pillepich, A.; Springel, V.; Nelson, D.; et al. First results from the IllustrisTNG simulations: the galaxy color bimodality. *Monthly Notices of the Royal Astronomical Society* **2018**, *475*, 648–675. <https://doi.org/10.1093/mnras/stx3112>.
4. Rodríguez-Gómez, V.; Genel, S.; Vogelsberger, M.; et al. The merger rate of galaxies in the Illustris simulation: a comparison with observations and semi-empirical models. *Monthly Notices of the Royal Astronomical Society* **2015**, *449*, 49–64. <https://doi.org/10.1093/mnras/stv264>.
5. Noeske, K.G.; Weiner, B.J.; Faber, S.M.; et al. Star formation in AEGIS field galaxies since $z = 1.1$: the dominance of gradually declining star formation, and the main sequence of star-forming galaxies. *The Astrophysical Journal Letters* **2007**, *660*, L43–L46. <https://doi.org/10.1086/517926>.
6. Whitaker, K.E.; van Dokkum, P.G.; Brammer, G.; Franx, M. The Star Formation Mass Sequence Out to $z = 2.5$. *The Astrophysical Journal Letters* **2012**, *754*, L29. <https://doi.org/10.1088/2041-8205/754/2/L29>.
7. Leaman, R.; VandenBerg, D.A.; Mendes de Oliveira, C. Galactic Dynamics Timescales and Chemical Evolution. *The Astrophysical Journal* **2012**, *750*, 114. <https://doi.org/10.1088/0004-637X/750/2/114>.
8. Frenzel, S.; Pompe, B. Partial Mutual Information for Coupling Analysis of Multivariate Time Series. *Physical Review Letters* **2007**, *99*, 204101. <https://doi.org/10.1103/PhysRevLett.99.204101>.
9. Kraskov, A.; Stögbauer, H.; Grassberger, P. Estimating mutual information. *Physical Review E* **2004**, *69*, 066138. <https://doi.org/10.1103/PhysRevE.69.066138>.
10. Efron, B.; Tibshirani, R.J. *An Introduction to the Bootstrap*; Chapman and Hall/CRC, 1994.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.