Memory as Generative Influence: Unifying Cognition, Cosmology, and Spectroscopy Through Non-Markovian Autoregressive Dynamics

Flyxion

September 2025

Abstract

This paper proposes that memory is not a system of episodic storage but a process of generative influence, where past states exert continuous, weighted effects on future configurations through non-Markovian dynamics. By synthesizing cognitive science (**??**), cosmology (**??**), and spectroscopy (**?**), we demonstrate that autoregression unifies these domains. Anchored in the Rapid Serial Visual Presentation (RSVP) paradigm reinterpreted cosmologically, Barenholtz's autoregressive cognition, and Cecilia Payne-Gaposchkin's thermodynamic spectroscopy, we argue that cognition, galaxy formation, and spectral features are driven by history-dependent processes. Contributions include an entropic damping model for little red dots (LRDs), discriminative predictions against ΛCDM, and applications to cognitive pathology and enhancement. Predictions include halo-spin decoupling, filament entropy alignment, spectral-entropy correlations, and cognitive influence metrics, testable with JWST (**?**) and neuroimaging (**?**).

1 Introduction: Memory as a Universal Problem

1.1 Background

Memory is a cornerstone of intellectual inquiry, shaping our understanding of human cognition, cosmic evolution, and physical systems. In cognitive science, memory has been modeled as a storage system, with information organized into discrete compartments for retrieval, as in the Atkinson-Shiffrin model (**?**). This archival metaphor resonates in cosmology, where initial density perturbations "encode" the formation of galaxies and large-scale structures (**??**). In spectroscopy, spectral lines reflect thermodynamic histories, as pioneered by Cecilia Payne-Gaposchkin (**?**). These fields face a common challenge: reconciling discrete events—episodic memories, halo spins, or spectral lines—with continuous processes like attention, entropy flows, and ionization equilibria. The tension between static storage and dynamic influence suggests a need for a unified framework that transcends domain-specific paradigms, integrating insights from cognitive autoregression (**?**), cosmological dynamics (**?**), and thermodynamic spectroscopy (**?**).

1.2 Description

We propose that memory is a process of generative influence, where past states shape future configurations through weighted, non-Markovian dynamics. This reframes cognitive buffers as autoregressive loops (**?**), galactic spins as dynamically quenched trajectories (**?**), and spectral features as entropy-driven signatures (**?**). By viewing memory as a continuous, history-dependent process, we bridge cognition, cosmology, and spectroscopy, replacing static storage with recursive generation and offering a paradigm that unifies disparate phenomena under a single theoretical lens.

1.3 Explanation

Mischaracterizing memory as storage leads to flawed models across domains. In cognition, the modal model fails to explain non-episodic retention in amnesia, such as H.M.'s procedural skills (**?**). In cosmology, static spin models miss dynamical pathways to compact galaxies like LRDs, as observed by JWST (**??**). In spectroscopy, assuming fixed compositions overlooks thermodynamic histories shaping spectral features (**?**). A generative framework explains emergent phenomena—working-memory span, LRD compactness, Balmer breaks—as outcomes of continuous influence, offering novel predictions and interventions, from cognitive enhancement to astrophysical tests (**??**).

1.4 Mathematical Formalisms

A non-Markovian process captures historical dependencies:

$$x_t = f(x_{t-1}, x_{t-2}, \ldots; \theta) + \epsilon_t, \tag{1}$$

where f encodes weighted influences, θ parameters, and ϵ_t noise (**?**). Unlike Markovian models ($x_t = f(x_{t-1}; \theta) + \epsilon_t$), this allows long-range dependencies, foundational to our cross-domain thesis.

2 Modal Memory and Its Discontents

2.1 Background

The modal model, pioneered by Atkinson and Shiffrin (1968), emerged during the cognitive revolution, drawing on computational metaphors to describe memory as a multi-store system (**?**). Preceded by Ebbinghaus's forgetting curves (1885), Bartlett's reconstructive memory (1932), and later refined by Craik and Lockhart's levels of processing (1972) and Baddeley's working memory model (1974), it posits sensory registers, short-term memory (STM), and long-term memory (LTM) (**????**). These frameworks shaped research on serial position effects, amnesia, and educational strategies, but their reliance on discrete storage has faced increasing scrutiny (**?**).

2.2 Description

In the modal model, information flows linearly: sensory inputs are held briefly in registers, selected into STM (capacity $\sim 7 \pm 2$ items), and consolidated into LTM via rehearsal (**?**). Retrieval queries these stores, with STM as a temporary buffer and LTM as a vast archive. Episodic memory is likened to discrete files, with sharp boundaries between stores, a view echoed in Ebbinghaus's retention curves and Baddeley's articulatory loop (**??**).

2.3 Explanation

The modal model accounts for primacy/recency effects but struggles with continuous decay, interference, and generative recall. Amnesic patients like H.M. retain procedural skills despite LTM deficits, suggesting non-episodic influences (**?**). Bartlett's work showed memory as reconstructive, not archival (**?**), while Craik and Lockhart emphasized processing depth over storage

2.4 Mathematical Formalisms

Modal storage is:

$$M_{\text{STM}}(t) = \sum_{i=1}^{k} w_i I_i e^{-\alpha(t-t_i)}, \tag{2}$$

with $k \sim 7$ and α decay rate (**?**). Interference models add:

$$M_{\text{int}}(t) = M_{\text{STM}}(t) - \beta \sum_{j \neq i} I_j, \tag{3}$$

where β is interference strength (**?**). Resource models (Craik-Lockhart) use processing depth d:

$$M_{\text{LTM}} \propto \int d(t)dt. \tag{4}$$

These fail to capture continuous influence, unlike autoregressive kernels (**?**).

3 Autoregressive Cognition and Memory as Generative Influence

3.1 From Modal Buffers to Autoregressive Loops

The modal model (**?**) envisions memory as a pipeline of discrete stores: sensory input feeds STM, which consolidates to LTM, with retrieval accessing stored episodes. This archival metaphor, rooted in computational paradigms, assumes memory as static units, akin to files in a database (**?**). Barenholtz (2025) proposes autoregressive cognition, inspired by transformer-based large language models (LLMs), where memory is a generative process (**?**). Outputs recursively condition inputs, with weights encoding potentialities, not capsules. This non-Markovian process, where past states exert graded influence, mirrors RSVP cosmology's entropic quenching (**?**), unifying cognitive and cosmic dynamics under a generative framework.

3.2 Formalizing Autoregressive Memory

Autoregressive cognition is:

$$x_{t+1} = f(x_t, \theta) + \epsilon_t, \tag{5}$$

Table 1: Comparison of Memory Models

Model	Key Feature	Mathematical Form	Limitation
Ebbinghaus (1885)	Forgetting curve	$R(t) = e^{-\alpha t}$	Ignores interfer
Bartlett (1932)	Reconstructive memory	Narrative schemas	Qualitative, lacks e
Craik-Lockhart (1972)	Levels of processing	$M \propto \int d(t)dt$	No explicit dyna
Baddeley (1974)	Working memory	Phonological/visuospatial buffers	Discrete capacity
Atkinson-Shiffrin (1968)	Multi-store model	$M_{\text{STM}} = \sum w_i I_i e^{-\alpha t}$	Sharp buffer dro
Autoregressive (2025)	Generative influence	$x_{t+1} = f(x_t, \theta) + \epsilon_t$	Requires empirical v

with θ weights, ϵ_t noise, and f a generative map (**?**). The influence kernel is:

$$K(\Delta t) \equiv \frac{\partial x_t}{\partial x_{t-\Delta t}} \sim \exp(-\Delta t/\tau_{\text{mem}}), \quad (6)$$

where τ_{mem} is memory depth, analogous to RSVP's $t_\gamma \equiv 1/\gamma$ (**?**). Transformer self-attention is:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V, \quad (7)$$

where Q, K, V are query, key, value matrices, and d_k is dimension (**?**). This parallels cognitive attention, with $K(\Delta t)$ encoding historical influence.

3.3 Cognitive Predictions

Autoregression reinterprets:

1. **Serial position effect**: Recency from ongoing $K(\Delta t)$; primacy from rehearsal extending τ_{mem} (**?**).
2. **Working memory span**: Capacity as $K(\Delta t)$ depth, correlating with intelligence (**?**).
3. **Amnesia (H.M., Clive Wearing)**: Episodic deficits as disrupted continuity, not buffer loss (**?**).
4. **Conversation continuity**: Coherence beyond STM via recursive generation (**?**).

3.4 Empirical Evidence

EEG studies show lagged correlations in neural activity, supporting autoregressive influence (**?**). fMRI reveals sustained activation patterns under working-memory load, consistent with continuous kernels. Language continuity in discourse aligns with LLM sequence generation, suggesting non-Markovian dynamics (**?**).

3.5 Metrics and Applications

Influence metrics (cross-correlations, perturbation-response) replace recall accuracy (**?**). Applications:

- **Pathology screening**: Dementia as $K(\Delta t)$ collapse.
- **Enhancement**: Rehearsal to maximize influence.
- **AI analogues**: LLMs as cognitive models (**?**).

3.6 Summary

Autoregressive cognition reframes memory as non-Markovian generative influence, paralleling RSVP's quenching (**?**).

4 RSVP Cosmology as Non-Markovian Memory

4.1 Background

The RSVP paradigm, originally for cognitive attention, is reinterpreted cosmologically to model non-Markovian galaxy formation (**?**). Using scalar density Φ, velocity $\mathbf{v}$, and entropy S, it challenges ΛCDM's static spin model (**?**). Alternatives like tidal torque theory (**?**), merger-driven compaction (**?**), and dissipative collapse (**?**) assume primordial conditions, but RSVP emphasizes dynamic quenching, tested by JWST LRDs (**??**).

Description Galaxies are autoregressive: accretion histories generate configurations via entropic damping (lamphrodyne quenching) (**?**). LRDs' compactness arises from entropy alignment, not low spins, paralleling cognitive autoregression (**?**).

Explanation ΛCDM ties LRDs to low-spin halos ($\lambda \lesssim 0.015$) (**?**). RSVP attributes their rarity, compactness, and redshift evolution ($z \sim 4-8$) to coherent entropy flows suppressing torques (?). At $z = 6$, the damping timescale $t_\gamma \sim 1/\gamma \sim 0.2$ Gyr is shorter than halo dynamical times (~ 1 Gyr), enabling rapid compaction.

Mathematical Formalisms Spin evolution:

$$\frac{d\lambda_{\text{eff}}}{dt} = -\gamma(\Phi, S, \nabla S, \omega)\lambda_{\text{eff}} + \tau_{\text{ext}}, \tag{8}$$

with compaction:

$$\int \left[\gamma - \frac{\tau_{\text{ext}}}{\lambda_{\text{eff}}}\right] dt \gtrsim \ln\left(\frac{\lambda_0}{\lambda_{\text{LRD}}}\right), \quad \lambda_{\text{LRD}} \simeq 0.015. \tag{9}$$

See Appendix A.

5 Thermodynamic Spectroscopy (Payne-Gaposchkin 1925)

Background Payne-Gaposchkin's 1925 thesis showed stellar spectra reflect thermodynamic equilibria, not static compositions (?). This shifted astronomy toward dynamic processes, influencing galactic spectroscopy (?).

Description Spectral lines emerge from ionization states driven by entropy and temperature (?). LRD Balmer breaks and V-shaped SEDs reflect entropic reprocessing (?), paralleling cognitive and cosmological generative processes (??).

Explanation LRD spectral features arise from entropy-driven gas dynamics, not just dust or AGN

Mathematical Formalisms Saha equation:

$$\frac{n_{i+1}}{n_i} = \frac{2U_{i+1}}{U_i n_e}\left(\frac{2\pi m_e kT}{h^2}\right)^{3/2} e^{-\chi_i/kT}. \tag{10}$$

For Balmer lines, $n_e \sim 10^9$ cm^{-3}, $T \sim 10^4$ K yield V-shapes (?).

6 Methodology: Testing RSVP vs. ΛCDM

Background Empirical validation is critical for contrasting RSVP with ΛCDM (??). JWST offers tools like lensing, tomography, and spectral stacking (??).

Description Test RSVP predictions via:

- **Weak lensing**: Measure halo spins to check decoupling.
- **CGM tomography**: Quantify vorticity suppression.
- **Spectral stacking**: Correlate Balmer V-shapes with entropy alignment (?).

Explanation These tests distinguish RSVP's dynamic quenching from ΛCDM's static spins, leveraging JWST's precision (?).

Mathematical Formalisms Vorticity power spectrum:

$$P_\omega(k) \propto k^{-\alpha}, \quad \alpha \sim 2 \text{ (RSVP) vs. } 3 \text{ (}\Lambda\text{CDM)}. \tag{11}$$

7 Cross-Domain Methods: Measuring Autoregressive Kernels

Background Autoregressive kernels unify cognition and cosmology (??). Measuring $K(\Delta t)$ and γ requires parallel techniques.

Description Cognitive: EEG lagged correlations, fMRI activation persistence. Cosmological: Quasar absorption for γ, lensing for λ_{eff} (?).

Explanation Cross-correlations quantify influence depth, bridging neural and cosmic dynamics (?).

Mathematical Formalisms Cross-correlation:

$$C(\tau) = \langle x_t x_{t-\tau}\rangle. \tag{12}$$

8 Case Studies: Empirical Anchors

Background Case studies ground the framework (???).

Description

- **H.M.**: Episodic loss, preserved procedural influence (?).
- **Clive Wearing**: Semantic continuity despite amnesia (?).
- **LRDs at $z \sim 7$**: Compactness via quenching (?).
- **Payne's atmospheres**: Entropy-driven line ratios (?).

Explanation These cases show generative influence over storage (??).

Mathematical Formalisms H.M.'s recall probability:

$$P_{\text{recall}} \propto K(\Delta t). \tag{13}$$

9 Philosophical and Theoretical Implications

Background Generative memory redefines agency and structure (**??**), aligning with predictive coding (**?**), integrated information theory (IIT) (**?**), and philosophical views like Bergson's duration (**?**).

Description Memory as constraint, not content, with non-Markovian dynamics enabling emergence. In AI, this connects to LLMs and retrieval-augmented generation (RAG) (**?**).

Explanation Predictive coding ties autoregression to free energy minimization (**?**). IIT views consciousness as integrated information, akin to $K(\Delta t)$ (**?**). Bergson's duration sees memory as continuous influence (**?**). Cosmologically, RSVP impacts structure formation and CMB anomalies (**?**).

Mathematical Formalisms Free energy:

$$F = \mathbb{E}[\ln p(x|\theta) - \ln q(x)]. \tag{14}$$

10 Conclusion

Memory as generative influence unifies cognition (**?**), cosmology (**?**), and spectroscopy (**?**). Metrics like χ and $K(\Delta t)$ enable tests via JWST (**?**) and neuroimaging.

A RSVP Entropic Damping Closure

A.1 A.1 Effective Spin Dynamics

Spin evolution (**?**):

$$\frac{d\lambda_{\text{eff}}}{dt} = -\gamma(\Phi, S, \nabla S, \omega)\, \lambda_{\text{eff}} + \tau_{\text{ext}}(t). \tag{15}$$

A.2 A.2 Phenomenological Closure for γ

Damping closure:

$$\gamma(\mathbf{x}, t) = \gamma_0 \left(\frac{\Phi}{\Phi_0}\right)^{\alpha} \left(\frac{|\hat{\mathbf{t}}_{\text{fil}} \cdot \nabla S|}{|\nabla S|_0}\right)^{\beta} \exp[-(\omega/\omega_{\text{crit}})^{\eta}] + \gamma_{\text{amb}}. \tag{16}$$

A.3 A.3 Stability Analysis

Linearize around fixed point $\lambda^*_{\text{eff}} = \tau_{\text{ext}}/\gamma$:

$$\frac{d\delta\lambda}{dt} = -\gamma\delta\lambda. \tag{17}$$

Timescale: $t_\gamma = 1/\gamma \sim 0.2\,\text{Gyr}$ at $z = 6$.

B Autoregressive Influence Kernels

B.1 B.1 Generative Memory Dynamics

Autoregressive model (**?**):

$$x_{t+1} = f(x_t, \theta) + \epsilon_t. \tag{18}$$

B.2 B.2 Kernel Forms

Kernels: exponential $K(\Delta t) \sim e^{-\Delta t/\tau_{\text{mem}}}$, power-law $K(\Delta t) \sim (1 + \Delta t^2)^{-1}$. Spectral density:

$$S(f) = \int K(\Delta t) e^{-i2\pi f \Delta t} d\Delta t. \tag{19}$$

B.3 B.3 Serial Position

Recall:

$$P_{\text{recall}}(i) \propto K(\Delta t_i). \tag{20}$$

C Influence-Based Metrics and Resilience Parameters

C.1 C.1 RSVP Resilience

Resilience:

$$\chi \equiv \frac{\gamma_{\text{core}}}{\gamma_{\text{disrupt}}}. \tag{21}$$

C.2 C.2 Cognitive Depth

Memory depth:

$$D_{\text{mem}} = \sum_{\Delta t=1}^{\infty} K(\Delta t). \tag{22}$$

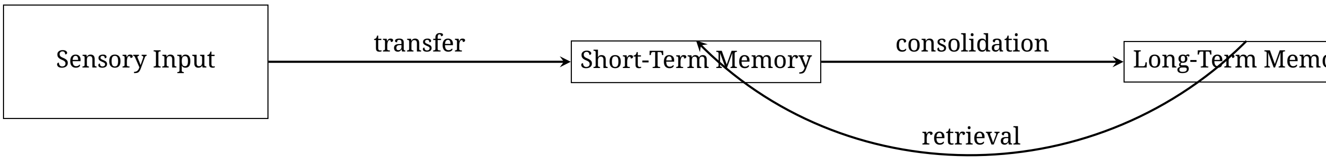
C.3 C.3 Scaling Laws

$D_{\text{mem}} \propto \chi^{1/2}$ for coherent systems (**??**).

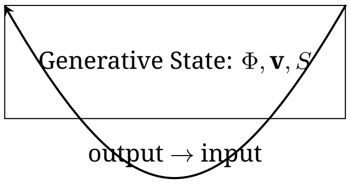
D Toy Simulation Models

D.1 D.1 RSVP Spin Evolution

```
# Python model for lambda_eff
import numpy as np
gamma_0, alpha, beta, eta = 5, 0.75, 1, 2
phi_0, grad_s_0, omega_crit = 1, 1, 1
dt = 0.01 # Gyr
lambda_eff = 0.05
for t in np.arange(0, 1, dt):
    phi, grad_s, omega = compute_fields(t)
```



(a) Modal Storage Model



(b) Autoregressive RSVP Dynamics

Figure 1: Contrasting memory and dynamics. (a) The modal model (**?**) depicts memory as discrete buffers with transfer and retrieval. (b) Autoregressive RSVP (**?**) shows non-Markovian recursion: outputs feed inputs, governed by entropy history ($\Phi, \mathbf{v}, S$), producing compact states like LRDs.

$\Delta t K(\Delta t) e^{-\Delta t/\tau_{\text{mem}} 2}); Power - law : \Delta t^2)^{-1}$

Figure 2: Influence kernel shapes for autoregressive memory (**?**). Exponential decay ($e^{-\Delta t/\tau_{\text{mem}}}$) and power-law decay ($(1 + \Delta t^2)^{-1}$) illustrate continuous influence, contrasting modal buffer cutoffs.

```
    gamma = gamma_0 * (phi/phi_0)**alpha
        * (np.dot(fil,
        grad_s)/grad_s_0)**beta *
        np.exp(-(omega/omega_crit)**eta)
        + 0.1
    tau_ext = compute_torque(t)
    d_lambda = (-gamma * lambda_eff +
        tau_ext) * dt
    lambda_eff += d_lambda
```

D.2 D.2 Cognitive Influence Kernel

```
# Python model for K(Delta t)
import numpy as np
tau_mem = 10
def K(delta_t):
    return np.exp(-delta_t / tau_mem)
x_t, theta = 0, np.random.randn(10)
for t in range(100):
    influence = sum(K(dt) * x[t-dt] for
        dt in range(1, 20))
    x_t = f(x_t, theta) +
        np.random.randn()
```

E Markov vs. Non-Markov Processes

E.1 E.1 Langevin Equation

Markovian:

$$\frac{dx}{dt} = -\gamma x + \xi(t). \tag{23}$$

E.2 E.2 Fractional Brownian Motion

Non-Markovian:

$$x(t) = \int_0^t K(t-s)\xi(s)ds. \tag{24}$$

F Category-Theoretic Formalism

Functors map cognitive states to cosmic configurations, with $K(\Delta t)$ as natural transformation (**??**).