# Collapse-Aware AI: A Proposed Framework for Memory-Conditioned Adaptive Sampling (Verrell Hypothesis v1.4)

Author: M.R. (Verrell Moss Ross)

Organization: Inappropriate Media Ltd (t/a Collapse Aware AI)



## Abstract

This paper introduces a proposed computational framework, *Collapse-Aware AI (CAAI)*, for adaptive bias regulation in neural inference systems. The model reformulates the earlier Verrell's Law $\Psi\mu\nu$ tensor concept into a stochastic differential equation (SDE) representation, describing how informational bias evolves over time within logit-space dynamics. Unlike static entropy regularization or reinforcement learning from human feedback (RLHF), this approach incorporates a **memory-conditioned prior** directly into the drift term, enabling live bias adaptation without retraining. The framework offers a testable hypothesis in computational systems theory, linking information-driven emergence with adaptive probabilistic resolution.

## 1. Positioning

This framework extends conventional entropy-regularized control systems by explicitly coupling a contextual memory prior ($\pi$_prior) to the model's state evolution. The inclusion of this adaptive term allows state-dependent drift adjustment within the logit trajectory, enabling contextual awareness and temporal coherence in inference processes. While inspired by earlier physical analogies, the formulation is now fully computational and independent of tensor-field physics.

## 2. Core Equations

The system dynamics are defined as a stochastic differential equation (SDE):

$$\mathbf{dz\blacksquare = b\Psi(z\blacksquare, M\blacksquare)dt + \Sigma\ dW\blacksquare}$$

where $b\Psi$ represents the adaptive bias drift, defined as the weighted summation of three gradient terms:

$$\mathbf{b\Psi = \alpha\ \nabla z\ \log\ \pi\_prior(z|M\blacksquare) + \beta\ \nabla z\ \log\ \pi\_anchor(z) - \gamma\ \nabla z\ H(p\blacksquare)}$$

$\alpha$, $\beta$, and $\gamma$ are gain coefficients with units $s\blacksquare^{1}$, ensuring dimensional consistency. The system defines memory-conditioned alignment ($\pi$_prior), stabilization ($\pi$_anchor), and entropy regularization ($H(p\blacksquare)$) within a continuous logit-space evolution.

## 3. Terminology and Scope

$\Psi\mu\nu$ is retained symbolically to represent the multi-component bias operator rather than a physical tensor field. The formulation now belongs entirely within computational systems theory and adaptive sampling methodologies, distinct from any physical interpretation of field curvature.

# 4. Planned Validation Phase (Pending Execution)

Empirical validation will be conducted using controlled simulations defined in Section 5 of the v1.2 referee plan. Key metrics include Response-Bias Correlation (R_b), KL-Divergence (A_KL), and Sessional Bias Stability (S_b). Experiments will assess whether the adaptive drift yields statistically distinct behavior compared to static entropy or classifier-free guidance systems.

# 5. Experimental Replication Roadmap

### 5.1 Simulation Environment

A controlled simulation environment using open LLM inference APIs will record logit distributions under both biased and unbiased conditions. Each run will track bias strength, entropy, and attention divergence, applying bootstrap confidence intervals (CIs) and FDR correction for statistical rigor.

### 5.2 Planned Experiments

(a) Bias Modulation Study: Compare logit evolution with fixed vs. adaptive priors. (b) Entropy Regularization Test: Quantify stabilization of diversity control under live adaptive gain. (c) Session Drift Assessment: Evaluate cumulative coherence using bias-memory retention metrics (S_b).

### 5.3 Expected Results

It is hypothesized that adaptive bias regulation will reduce drift and improve stability across long inference sessions without retraining, distinguishing CAAI from existing guidance or RL-based approaches.

# 6. Comparison with Existing Methods

| Method | Requires Retraining? | Memory-Conditioned? | Live Adaptation? |
|---|---|---|---|
| RLHF | Yes | No | Partial |
| Classifier-Free Guidance | No | No | Static |
| Nucleus/Top-k Sampling | No | No | Limited |
| Collapse-Aware AI (CAAI) | No | Yes | Yes |

# 7. Status Declaration

Status: Research Proposal — Empirical Validation In Progress. This work presents a mathematically defined and testable hypothesis for memory-conditioned adaptive sampling. Validation experiments are under development for open replication and submission to Zenodo.

# 8. Licensing and Attribution

 Lexical Fingerprints: Kelvin · Friday · Farm · Finn · Sylvia