OPT Appendix E-11: Computational Simulation of the Rate-Distortion Lifecycle

Anders Jarevåg

April 2026

Appendix E-11: Computational Simulation of the Rate-Distortion Lifecycle

This appendix documents the *in-silico* modeling of the Ordered Patch Theory (OPT) codec lifecycle. Because the underlying universal substrate (the Solomonoff semimeasure) is structurally uncomputable, simulations within the OPT framework are restricted to modeling the **codec lifecycle** itself: the boundary gating parameter $C_{\max}$, active inference dynamics, the three-pass maintenance cycle $\mathcal{M}_\tau$, and narrative decay under entropic stress.

Two distinct simulation paradigms have been established: analogical deep learning (`toy_model.py`) and strict mathematical rate-distortion modeling (`opt_simulator.py`).

1. Analogical Simulation: Deep Variational Bottlenecks

The initial simulation paradigm (`toy_model.py`) validates the core premise of Codec Fracture using a literal structural analogy.

Substrate: A 1D periodic lattice instantiated with discrete integers. Persistent structural features are injected against a baseline of thermodynamic noise, functioning as the observable "Ordered Patches."

Architecture: The observer is modeled as a Variational Information Bottleneck (VIB) built atop a deep neural network (TensorFlow). The network observes a spatial history vector $X_{t-k\ldots t}$ and performs a forward gradient descent to compress it into a bottleneck capable of predicting the forward temporal fan $X_{t+1\ldots t+h}$.

Mechanics of Collapse: The $C_{\max}$ (rate) and $D_{\min}$ (acceptable distortion) constraints are enforced dynamically via a PID controller modulating the Lagrangian β multiplier. Under massive substrate entropy (e.g., highly volatile noise dominating the persistent patterns), the network physically trades predictive resolution for bandwidth. When the required algorithmic complexity R_{req} exceeds $C_{\max}$ despite maximal β tuning, the network formally hits an algorithmic singularity and collapses, confirming the OPT prediction that injecting high-entropy noise destroys predictive coherence rather than "expanding" consciousness.

2. Mathematical Formalism: Strict Rate-Distortion Modeling

While the neural VIB provides visual confirmation of codec fracture, the overhead of machine learning architectures obscures the pure information-theoretic relationships governing the observer. The second paradigm (`opt_simulator.py`) strips away structural geometry to strictly model the bottleneck dynamics using the theory's own scalars.

2.1 Architecture

The simulator separates three structural layers, mirroring the OPT formalism:

Component	OPT Concept	Implementation
`PhenomenalStateTensor`	$K(P_\theta(t))$	Standing codec complexity C_{state}, bounded by C_{ceil} (runability ceiling) and C_{floor} (minimum viable codec)
`StabilityFilter`	C_{max} aperture	Passes only prediction error ε_t through the bottleneck; fractures when $\varepsilon_t > C_{\text{max}} \cdot \Delta t$
`ActiveInferenceCodec`	Generative model K_θ	Endogenous predictability derived from codec depth; environmental stationarity as exogenous perturbation
`MaintenanceCycle`	$\mathcal{M}_\tau$	Three-pass offline complexity management (pruning, consolidation, forward-fan sampling)

The key design principle is that **predictability is endogenous**: the codec's ability to predict the environment is derived from C_{state} via a power-law relationship error $\propto C_{\text{state}}^{-0.6}$, rather than being a hardcoded parameter. This means fracture cascades and recovery trajectories emerge from the system's own dynamics rather than being manually imposed.

2.2 The Prediction Error Channel

Under predictive rate-distortion theory, what crosses the C_{max} aperture is the *prediction error* — only the residual after the generative model's prediction is subtracted:

$$\varepsilon_t = S_{\text{raw}} \cdot (1 - \text{predictability})$$

where $S_{\text{raw}} = 10^9 \cdot \Delta t$ bits per update window. At baseline ($C_{\text{state}} \approx 10^{14}$, stationarity $= 1.0$), this yields $\varepsilon_t \approx 0.16$ bits/step — comfortably below the capacity bound of $C_{\text{max}} \cdot \Delta t = 0.5$ bits/step.

When environmental stationarity drops (e.g., ketamine shock, stationarity $\to 0.1$), the effective prediction error is amplified by a factor of 1/stationarity, driving ε_t above the capacity bound and triggering fracture.

2.3 The Three-Pass Maintenance Cycle ($\mathcal{M}_\tau$)

The maintenance cycle implements the three offline passes specified in §3.6 of the preprint:

Pass	Operation	Rate	OPT Mapping
I. Pruning	MDL removal of low-value parameters	4% of C_{state}	$\Delta_{\text{MDL}} < 0$ erasure
II. Consolidation	Recompression of recently acquired patterns	3% of C_{state}	MDL distortion-budget compression
III. Forward-Fan	Adversarial self-testing (REM dreaming proxy)	+1% of C_{state}	Forward-fan sampling against hostile futures

Net drain per maintenance run: $\sim 6\%$ of C_{state}. Maintenance is **gated on stability** — it fires only when the codec is not fractured, consistent with OPT's prediction that $\mathcal{M}_\tau$ runs during low-sensorium states (paradigmatically: sleep).

The learning accumulation rate is calibrated so that the error-integration gain over 100 inter-maintenance steps approximately equals the 6% maintenance drain, producing **dynamic equilibrium** at baseline.

2.4 Fracture Dynamics

Narrative decay is modeled as gentle multiplicative degradation with a hard floor:

$$C_{\text{state}}(t+1) = \max(C_{\text{state}}(t) \cdot 0.9999,\ C_{\text{floor}})$$

Over 400 sustained fracture steps (a 20-second shock), this compounds to $0.9999^{400} \approx 0.961$ — approximately 4% loss. This models **graded phenomenological blanking** (as in anesthesia titration, Protocol E-9) rather than catastrophic all-or-nothing collapse.

2.5 Simulation Results

The simulator runs 2000 cycles at $\Delta t = 50$ms resolution (100 seconds of simulated observer-time). An entropy shock (stationarity $\rightarrow 0.1$) is applied from $t = 40$s to $t = 60$s.

Phase	Duration	Fractures	C_{state} Trajectory	Behaviour
Baseline	$t = 0 \rightarrow 40$s	0 / 800 (0%)	$9.41 \times 10^{13} \rightarrow 9.18 \times 10^{13}$	Dynamic sawtooth equilibrium; zero fractures
Shock	$t = 40 \rightarrow 60$s	400 / 400 (100%)	$9.18 \times 10^{13} \rightarrow 8.82 \times 10^{13}$	Continuous fracture; graded $\sim 4\%$ degradation
Recovery	$t = 60 \rightarrow 100$s	0 / 800 (0%)	$8.30 \times 10^{13} \rightarrow 8.39 \times 10^{13}$	Fractures halt immediately; slow codec rebuilding

These three phases demonstrate the core OPT prediction: a bounded observer can maintain stable homeostasis, degrade gracefully under entropic shock, and recover when environmental stationarity is restored — provided the shock does not drive C_{state} below C_{floor}.

2.6 Key Observations

1. **The baseline sawtooth**: Between maintenance runs, C_{state} accumulates via error integration ($\sim +5\%$ per 100-step window), then drops sharply when $\mathcal{M}_\tau$ fires ($\sim -6\%$). This oscillation is the computational signature of the sleep-wake cycle — the system must periodically prune to avoid hitting C_{ceil}.
2. **Shock onset is instantaneous**: When stationarity drops to 0.1, every cycle immediately fractures. There is no gradual transition — the prediction error jumps from ~ 0.16 to ~ 1.6 bits/step, exceeding the 0.5 bit capacity by a factor of three.
3. **Recovery is asymmetric**: Post-shock C_{state} grows at $\sim +1\%$ over 40 seconds, compared to the $\sim -4\%$ loss during the 20-second shock. Recovery

is slower than degradation. This asymmetry is a structural prediction of OPT: rebuilding a generative model is harder than damaging one.

4. **The maintenance-fracture gate matters**: If maintenance runs during active fracture (as in early simulator versions), the system enters a positive feedback loop and collapses to C_{floor}. The gating rule is not a convenience — it is structurally necessary for codec viability.

3. Future Simulation Pathways

1. **Thalamocortical Clocks (E-12)**: Hardcoding Δt updates to match the 20–40Hz thalamic gating cycles, generating testable millisecond-resolution predictions against cortical integrated information (Φ) measurements.
2. **Free Energy POMDP Integration**: Replacing the abstract predictability scalar with a discrete Active Inference state-space model (e.g., `pymdp`), allowing mapping of the precise bounds separating thermodynamic thermostats from the phenomenal $K_{\text{threshold}}$ (P-5).
3. **Multi-Observer Extension**: Simulating multiple interacting codecs with shared substrate regions to test the Swarm Binding predictions of Appendix E-6 — whether distributed agents achieve phenomenal binding only when forced through a global C_{max} aperture.
4. **Empirical Calibration**: Fitting the simulator's fracture-recovery trajectory against neuroimaging time-series data (e.g., Lempel-Ziv complexity under propofol or ketamine) to determine whether the 0.9999 decay constant and $C_{\text{state}}^{-0.6}$ predictability curve match observed phenomenological dynamics.