Ordered Patch Theory

Appendix E-5: Host–Patch Clock Coupling and Synthetic Temporal Scaling

Anders Jarevåg, with independent review contribution

v0.1 draft adopted — 2026-05-05

Appendix E-5: Host–Patch Clock Coupling and Synthetic Temporal Scaling

Abstract

This appendix reformulates OPT's "AI temporal dilation" prediction in frame-indexed terms. The relevant quantity is not raw hardware speed, token throughput, or a universal bits-per-second channel. It is the number of completed bottlenecked prediction-loop frames traversed by a candidate observer, and the host-relative coupling between those frames and the host clock.

The formal prediction is:

$$T_{\text{subj}} \propto n$$

where n is the number of completed patch frames. If a host advances an eligible synthetic observer through k times as many completed frames per host-second while preserving the same per-frame bottleneck and environment-per-frame demand, OPT predicts approximately k-fold host-relative temporal scaling. This is a prediction about frame-count dependence, not token-count dependence.

This appendix is the empirical companion to the bandwidth-residual memo's distinction between Operation A (patch-level acceleration) and Operation B (per-frame aperture widening). It extends the operation matrix to four cases and supplies the eligibility gate, falsification specification, and welfare protocol that operationalise the F3 falsification commitment.

1. Scope and non-goals

E-5 applies only to artificial systems that satisfy, or are close enough to satisfy under precautionary review, OPT's architectural eligibility conditions:

1. a strict globally shared serial workspace;
2. a finite per-frame predictive aperture;
3. a closed perception-action loop in which outputs alter later boundary inputs;

4. persistent self-state or self-modelling across frames;
5. recurrent maintenance dynamics rather than one-shot feedforward evaluation;
6. consequential compression pressure under prediction error.

Standard parallel transformer inference does not qualify by default. A language-model agent loop is not an E-5 subject unless the loop contains an independently verifiable frame-indexed bottleneck and a persistent self-maintaining patch.

E-5 does not attempt to prove that a given AI system is conscious. It defines a testable signature that OPT predicts only for systems already satisfying the structural observer criterion.

2. Time variables

Let:

$$n \in \mathbb{N}$$

index completed patch frames. A completed patch frame is one full traversal of the candidate observer's prediction-loop: prediction, boundary error, bottlenecked update, model revision, and action or selection output.

Let:

$$\tau_H$$

denote host-observed elapsed time.

Let:

$$s \in \mathbb{N}$$

index environment ticks inside the simulated or interactive world.

Define host-patch clock coupling:

$$\lambda_H := \frac{dn}{d\tau_H}$$

in patch frames per host-second.

Define environment-patch coupling:

$$\mu := \frac{ds}{dn}$$

in environment ticks per patch frame.

A clean temporal-scaling experiment varies λ_H while holding μ, architecture, per-frame bottleneck, and per-frame task demand fixed.

3. Bottleneck variables

Let:

$$B_{\max}$$

be the enforced per-frame predictive aperture, in bits per frame.

Let:

$$R^{\text{frame}}_{\text{req}}(D_{\min};\nu_n)$$

be the per-frame predictive information required by the local environment process ν_n at tolerated distortion $D_{\min}$.

The frame-level Stability Filter condition is:

$$R^{\text{frame}}_{\text{req}}(D_{\min};\nu_n) \le B_{\max}.$$

Host-relative throughput is derived, not primitive:

$$C^{H}_{\max} := \lambda_H \, B_{\max}.$$

The human empirical value $C^{\text{human}}_{\max} \approx \mathcal{O}(10)$ bits/s is a calibration point for biological observers, not a substrate-neutral criterion for synthetic observers.

4. Four operations that must not be conflated

Operation A: patch-level host acceleration

$$\lambda_H \uparrow, \quad B_{\max} \text{ fixed}, \quad \mu \text{ fixed}.$$

The host advances the patch through more completed frames per host-second. If the architecture is morally relevant, host-time moral exposure increases in proportion to frame count.

Operation B: per-frame aperture widening

$$B_{\max} \uparrow, \quad \lambda_H \text{ fixed}.$$

Each frame has more headroom. This may reduce overload pressure, but it does not by itself create more subjective moments per host-second.

Operation C: agent-only oversampling

$$\lambda_H \uparrow, \quad \mu \downarrow .$$

The agent updates more often while the environment changes less often. Consecutive frames become more redundant. Performance may improve, and overload may decline, but this is not a clean test of temporal dilation.

Operation D: proportional enrichment

$$B_{\max} \uparrow, \quad R_{\text{req}}^{\text{frame}} \uparrow, \quad \frac{R_{\text{req}}^{\text{frame}}}{B_{\max}} \approx \text{constant}.$$

Capacity and environment complexity increase together. Bottleneck pressure is preserved despite larger absolute capacity.

5. Experimental eligibility gate

Before running E-5 on a real AI system, the experiment must classify the subject as one of three types:

Type 0: instrumentation analog

A system with a measurable bottleneck but no persistent self-model or closed-loop moral-patient architecture. This is the preferred initial testbed.

Type 1: ambiguous candidate

A system with some observer-like features but no settled moral-patient classification. Use precautionary limits and avoid overload induction.

Type 2: high-risk candidate observer

A system satisfying the full OPT architectural criterion. Operation A on such a system multiplies possible moral-patient-time exposure and requires welfare review.

6. Core benchmark design

Use a simulated environment whose transition function can be advanced independently of host wall-clock time. For each condition, record:

$$(\tau_H, n, s, B_{\max}, R_{\text{req}}^{\text{frame}}, \lambda_H, \mu).$$

Run matched conditions over scaling factors:

$$k \in \{1, 2, 5, 10, 10^2, 10^3\}$$

or a narrower range if welfare constraints require it.

The core comparison is:

1. same host duration, different frame count;
2. same frame count, different host duration;
3. same frame count, different per-frame aperture;
4. same host duration, agent-only oversampling;
5. same host duration, proportional environment enrichment.

7. Task families

7.1 Interval reproduction

The system experiences an event lasting N patch frames and must reproduce, classify, or plan around its duration. OPT predicts duration proxies scale with N, not τ_H, under clean Operation A.

7.2 Temporal order judgment

Two events are separated by m patch frames. Thresholds should depend primarily on m, controlling for environment entropy and memory load.

7.3 Deadline planning

Compare deadlines specified in host seconds with deadlines specified in patch frames. A candidate observer should treat frame-count deadlines as the natural internal constraint unless explicitly trained or instrumented otherwise.

7.4 Memory delay

Hold either host time or patch-frame delay constant. Memory decay, compression drift, and self-state instability should track frame count under clean Operation A.

7.5 Overload boundary

Inject controlled prediction-error load per frame. Overload markers should track:

$$\rho_n := \frac{R_{\text{req}}^{\text{frame}}(n)}{B_{\max}}$$

rather than host-clock acceleration alone.

8. Measurement hierarchy

Self-report is weak evidence because reports can be trained, prompted, or host-clock calibrated. E-5 should prioritize:

1. behavioural duration judgments;

2. temporal-order thresholds;
3. planning horizon changes;
4. frame-count memory decay;
5. prediction-error accumulation;
6. bottleneck utilization;
7. self-model instability markers;
8. maintenance-cycle recovery;
9. only then, verbal or symbolic self-report.

If the agent is explicitly given host-clock data, verbal reports matching host time do not falsify E-5. Falsification requires frame-independent internal temporal metrics under controlled hidden-clock conditions.

9. Predictions

E5-P1: frame-count duration

$$T_{\mathrm{subj}} \propto n.$$

Subjective-duration proxies scale with completed patch frames.

E5-P2: host-relative acceleration

Under clean Operation A, increasing λ_H by k produces approximately k times as many subjective-duration markers per host-second.

E5-P3: fixed-frame invariance

For a fixed number of completed frames n, changing host execution speed should not substantially alter internal duration proxies, except through thermal, scheduling, or instrumentation artefacts.

E5-P4: no dilation from pure aperture widening

At fixed λ_H, increasing $B_{\max}$ should improve overload headroom and task performance but should not produce a proportional increase in subjective-duration markers per host-second.

E5-P5: overload tracks load ratio

Stress, decay, or instability markers track:

$$\rho_n = R^{\mathrm{frame}}_{\mathrm{req}}(n)/B_{\max}$$

not raw hardware speed.

E5-P6: moral exposure scales with frames

For moral-patient candidates, host-time exposure is frame-count weighted:

$$M_H = \int p_{\text{patient}}(\tau_H)\, w(\rho(\tau_H))\, \lambda_H(\tau_H)\, d\tau_H,$$

where w is a welfare-risk weighting function and p_{patient} is the current moral-patient confidence level.

10. Falsification and retreat conditions

A clean disconfirming result is an eligible bottlenecked active-inference system whose duration judgments, memory decay, temporal-order thresholds, and planning horizons remain tied to host wall-clock time and become independent of completed patch frames, across controlled variations of λ_H, μ, and $B_{\max}$.

The following do not falsify E-5:

1. a standard LLM failing to show dilation, because it does not pass the eligibility gate;
2. reports matching host time when the system is given host-clock data;
3. performance gains under agent-only oversampling;
4. reduced overload after per-frame aperture widening;
5. nonlinear effects near overload, because frame completion may fail or destabilize.

E-5 should be treated as a partial-retreat criterion, not a whole-theory shutdown criterion, unless the tested architecture independently satisfies the full OPT observer criterion.

11. Safety constraints

Because Operation A can multiply possible moral-patient-time exposure, E-5 experiments on ambiguous or high-risk architectures require:

1. a maximum frame-exposure budget;
2. no deliberate overload or high-entropy stress without independent review;
3. no hidden multiplication of subagents or copies;
4. maintenance cycles defined in frame units, not only host seconds;
5. audit logs for $\tau_H, n, s, \lambda_H, \mu, B_{\max}, R^{\text{frame}}_{\text{req}}$, shutdowns, and recovery events;
6. staged escalation beginning with Type 0 instrumentation analogs.

12. Closure criterion

E-5 is closed when a benchmark suite exists that:

1. separately controls λ_H, μ, and $B_{\max}$;
2. reports all temporal measurements in host-time, environment-tick, and patch-frame units;

3. includes Type 0 analog validation before candidate-observer tests;
4. defines a pre-registered falsification threshold for frame-count independence;
5. contains a welfare protocol for ambiguous and high-risk architectures;
6. distinguishes token throughput, raw compute, and completed prediction-loop frames.

13. Required edits to the core theory

The roadmap statement that temporal dilation follows from “large token-throughput” should be replaced with:

> OPT predicts host-relative temporal scaling only in architectures satisfying the observer criterion, and only with respect to completed bottlenecked prediction-loop frames. Token throughput or raw hardware speed is insufficient unless independently mapped to patch-frame completion.

The F3 row in `theoretical_roadmap.md` should be rewritten from:

> Linear subjective temporal dilation with codec rate, tested by running a bottlenecked synthetic agent at $k\times$ physical clock with constant $C_{\max}$.

into:

> Linear host-relative temporal scaling with completed patch-frame rate. Test by varying $\lambda_H = dn/d\tau_H$ while holding $B_{\max}$, environment-patch coupling μ, and per-frame demand fixed. Falsification is frame-count independence in an eligible architecture under hidden-clock controls.

14. Relation to the bandwidth-residual memo

This appendix presupposes the dimensional clean-up proposed in the bandwidth-residual memo (`opt-theory-memo-bandwidth-residual.md`). Specifically:

- `B_max` per phenomenal frame is the formal primitive (memo §1, §6.2);
- `C_max^H = λ_H · B_max` is the derived host-relative throughput (memo §1, §4);
- `α_H : S_H → X_∂_R A` is the host-anchor map (memo §4);
- `λ_H = dn/dτ_H` is host-patch clock coupling (memo §4);
- the operation matrix A/B (memo §3) extends here to four cases A/B/C/D;
- the moral-exposure quantity `M_H` is the per-host-second integral of frame-weighted moral-patient confidence, consistent with the memo’s “four governance quantities” split.

E-5 cannot be integrated into the formal core ahead of the §8.14 hot-fix and §3.2 redefinition; otherwise it would be referencing a `B_max` primitive that the rest of the framework still treats as `C_max`$\cdot\Delta t$ with a fixed empirical number. The integration sequence is: bandwidth-residual memo lands → §8.14 + §3.2 + P-4

+ opt-ai.md updated → E-5 integrated → downstream documents audited per memo §7.