Ordered Patch Theory

Appendix T-10: Inter-Observer Coupling Under the Render Ontology

Anders Jarevåg

April 17, 2026 | DOI: 10.5281/zenodo.19300777

Original Task (from Roadmap T-10): “A formal derivation of how two observer patches interact within the shared substrate, establishing multi-patch coupling beyond purely solipsistic ‘local anchors.’ ”
Deliverable: A structural account of inter-patch consistency under OPT’s render ontology, grounding the apparent “shared world” without invoking an independently existing one.

Closure status: DRAFT STRUCTURAL CORRESPONDENCE. This appendix establishes a consistency constraint (Theorem T-10), a compression-forced symmetry (Corollary T-10a), and a communication theorem (Theorem T-10b) that together characterise the inter-observer coupling mechanism within OPT’s framework. The results are conditional on Axiom 1 (Solomonoff identification) and the structural corollary (Theorem T-11).

Section 1. The Problem

1.1 What Needs Explaining

Under OPT’s render ontology (preprint Section 8.6), each observer’s experienced world is a render: a compression artifact of their own predictive model. There is no independently existing “physical world” that multiple observers perceive differently. Each patch generates its own world.

This creates a coupling problem. Alice’s render contains a Bob-artifact — a high-complexity substructure whose behaviour is most compressibly described as an independently instantiated observer (Theorem T-11). Bob’s render contains an Alice-artifact. The question is: what structural relationship holds between these two artifacts?

If Alice’s Bob-artifact and Bob’s Alice-artifact are unconstrained — if they can behave arbitrarily relative to each other — then the “shared world” is an illusion in the most radical sense: not merely rendered rather than independently real, but potentially incoherent across patches. Conversations would not be genuine inter-observer events; they would be two separate renders that happen to contain similar-looking sequences.

1.2 What OPT Cannot and Should Not Claim

OPT cannot claim that Alice and Bob inhabit the "same world" in the naive realist sense — that is precisely the ontological position OPT rejects. It cannot invoke a substrate-level mechanism that "sends signals" between patches, because the substrate is the uninterpreted mathematical object that the render compresses, and patches do not interact "within" the substrate in the causal sense the word usually implies.

What OPT can and should establish is: the Solomonoff prior that governs each patch's stream imposes *consistency constraints* between the Alice-artifact in Bob's render and Alice's own first-person stream, and vice versa. These constraints are not caused by physical interaction. They are consequences of the same parsimony principle that generates physical laws, other observers, and the apparent solidity of the world.

1.3 Scope

This appendix provides:

1. A formal definition of cross-patch consistency (Section 2).
2. A proof that the Solomonoff prior enforces inter-artifact consistency — **Theorem T-10** (Section 3).
3. A corollary establishing the symmetry of the coupling — **Corollary T-10a** (Section 4).
4. A communication theorem proving that the coupling is sufficient for genuine information transfer across patches — **Theorem T-10b** (Section 5).
5. The formal relationship to Muller's multi-agent convergence (Section 6).

Section 2. Definitions

2.1 Two-Patch Setup

Consider two observer patches, $\mathcal{P}_A$ (Alice) and $\mathcal{P}_B$ (Bob), each governed by their own Solomonoff-weighted stream (Axiom 1):

$$\omega_A \sim M_A, \qquad \omega_B \sim M_B \tag{1}$$

where M_A and M_B are the universal semimeasures weighting each patch's stream. By the Stability Filter, each stream embeds in a computable world:

$$\omega_A \hookrightarrow W_A \quad \text{with measure } \mu_A, \qquad \omega_B \hookrightarrow W_B \quad \text{with measure } \mu_B \tag{2}$$

2.2 Cross-Patch Artifacts

Within Alice's world W_A, there exists a Bob-artifact: a substructure B_A whose behavioural trace is $\beta_{B|A} = (y_1, \ldots, y_T)$. Within Bob's world W_B, there exists an Alice-artifact A_B with behavioural trace $\alpha_{A|B} = (z_1, \ldots, z_T)$.

By Theorem T-11, the MDL-optimal description of B_A invokes Bob as an independently instantiated observer. Similarly for A_B.

2.3 Consistency

Definition T-10.D1 (Cross-Patch Consistency). The two-patch system $(\mathcal{P}_A, \mathcal{P}_B)$ is *ϵ-consistent* if the Bob-artifact's behaviour in Alice's render matches the third-person prediction of Bob's own first-person stream, and vice versa:

$$\left\|\beta_{B|A} - \beta_{B|B}\right\|_{\mathrm{KL}} \leq \epsilon \qquad \text{and} \qquad \left\|\alpha_{A|B} - \alpha_{A|A}\right\|_{\mathrm{KL}} \leq \epsilon \tag{T-10.D1}$$

where $\beta_{B|B}$ is Bob's actual first-person behavioural output and $\alpha_{A|A}$ is Alice's, and $\|\cdot\|_{\mathrm{KL}}$ denotes the KL-divergence between the probability distributions over the behavioural traces.

In words: cross-patch consistency means that what Alice observes Bob doing (in her render) matches what Bob is actually doing (in his render), and vice versa.

Section 3. Theorem T-10: Compression-Forced Consistency

3.1 The Key Insight

The insight is that *inconsistency is expensive.* If the Bob-artifact in Alice's render behaves differently from Bob's actual first-person stream, then Alice's stream must encode Bob's behaviour as an ad hoc specification rather than invoking Bob's own predictive model. By Theorem T-11, this requires strictly more bits.

The Solomonoff prior exponentially penalises long descriptions. Therefore, streams in which cross-patch artifacts are consistent with their putative first-person sources are exponentially more probable than streams in which they are not.

3.2 The Theorem

Theorem T-10 (Compression-Forced Consistency). *Let $\mathcal{P}_A$ and $\mathcal{P}_B$ be two patches satisfying Axiom 1, each embedding in a computable world via the Stability Filter, and each containing a cross-patch artifact satisfying the structural corollary (T-11). Then the Solomonoff prior enforces ϵ-consistency (Definition T-10.D1) with probability approaching unity as the observation horizon $T \to \infty$:*

$$\Pr\left[\left\|\beta_{B|A} - \beta_{B|B}\right\|_{\mathrm{KL}} > \epsilon\right] \leq 2^{-\Omega(T)} \tag{T-10}$$

Proof.

(i) **Description length of consistent streams.** Under cross-patch consistency, Alice's description of Bob's behaviour invokes Theorem T-11's independent-instantiation hypothesis H_{ind}. The description length is:

$$L_{\text{consistent}} = K(\mu_A) + K(\text{embed}_B) + \Big(-\log_2 P_{\text{3rd}}(\beta_{B|A} \mid x_B)\Big) \tag{3}$$

By Muller's convergence (L-3 from T-11), $P_{\text{3rd}} \approx P_{\text{1st}}$, so the log-loss term is near-optimal.

(ii) **Description length of inconsistent streams.** If $\beta_{B|A} \neq \beta_{B|B}$ beyond ϵ, then Alice's stream must encode Bob's behaviour as an arbitrary specification. By Theorem T-11, the cost is:

$$L_{\text{inconsistent}} \geq L_{\text{consistent}} + \bar{I}_T - O(\log T) \tag{4}$$

where $\bar{I}_T$ is the per-agent mutual information from Theorem T-11, which grows linearly in T.

(iii) **Solomonoff weighting.** The Solomonoff prior assigns probability $\leq 2^{-L}$ to any stream of description length L (up to constants). Therefore:

$$\frac{\Pr[\text{inconsistent}]}{\Pr[\text{consistent}]} \leq 2^{-(L_{\text{inconsistent}} - L_{\text{consistent}})} \leq 2^{-\bar{I}_T + O(\log T)} \tag{5}$$

Since $\bar{I}_T$ grows linearly in T, this ratio is exponentially decreasing. ■

3.3 Interpretation

Theorem T-10 does not say that a substrate-level mechanism "synchronises" Alice and Bob. It says that the parsimony of the Solomonoff prior makes inconsistent streams exponentially less probable than consistent ones. The "shared world" is not a place where both observers live. It is the consequence of the fact that the cheapest description of an apparent agent is one that invokes their own first-person stream — and the cheapest such description is necessarily *consistent* with that first-person stream.

The coupling is not causal. It is compressive. The shared world is a compression artifact of the same principle that generates physical laws: the simplest rendering of a lawful universe populated by coherent agents is one in which those agents' renders agree with each other.

Section 4. Corollary T-10a: Symmetry

Corollary T-10a (Symmetric Coupling). *The consistency constraint of Theorem T-10 is symmetric: if Alice's render is consistent with Bob's first-person stream, then Bob's render is consistent with Alice's first-person stream, with the same asymptotic bound.*

Proof. The argument of Theorem T-10 applies with the roles of $\mathcal{P}_A$ and $\mathcal{P}_B$ exchanged. The Solomonoff prior weighting operates independently on each patch's stream, and the compression advantage of consistent artifacts is symmetric because it depends only on the structural corollary (T-11), which applies equally to Alice-artifacts and Bob-artifacts. ■

Remark. This symmetry is not trivial. Under a naive reading of OPT's ontological solipsism, one might expect Alice's render to be "primary" and Bob's to be "derivative" — a genuine asymmetry between patches. Corollary T-10a shows that the compression logic is indifferent to which patch is "primary": the MDL advantage of consistency is the same from either perspective. This is the formal content of the intuition that the apparent world "treats all observers equally" — not because there is an observer-independent reality that does so, but because the Solomonoff prior penalises observer-dependent inconsistencies equally.

Section 5. Theorem T-10b: Information Transfer

5.1 The Communication Problem

Can Alice genuinely communicate with Bob under the render ontology? If Alice "speaks" to the Bob-artifact, the Bob-artifact's response is generated by Alice's own render. Is this genuine information transfer, or is Alice merely talking to a compressed model of Bob within her own stream?

5.2 The Answer

Theorem T-10b (Communication as Cross-Patch Coupling). *Let Alice generate a novel signal s_A (with $K(s_A) > 0$) that she intends to communicate to the Bob-artifact. Under ϵ-consistency (T-10), the following hold:*

(i) Bob's first-person stream registers s_A (or a compressed representation of it) with probability $\geq 1 - 2^{-\Omega(T)}$.

(ii) Bob's response to s_A is generated by Bob's own first-person stream (not specified ad hoc by Alice's render), with the same probability.

(iii) Alice's render of Bob's response matches Bob's actual first-person response, completing the communication loop.

Proof.

(i) By Theorem T-10, the Bob-artifact in Alice's render behaves consistently with Bob's first-person stream. If Alice presents s_A to the Bob-artifact, the Bob-artifact's perception of s_A is consistent with what Bob's first-person stream would register if receiving s_A as input. This is because the MDL-optimal description of the Bob-artifact includes Bob's own predictive model, which processes s_A as input.

(ii) The Bob-artifact's *response* to s_A is likewise generated by invocation of Bob's independent Solomonoff-weighted stream (by T-11). Any deviation from Bob's actual response would require ad hoc specification, at higher description length, and is therefore exponentially suppressed by the Solomonoff prior.

(iii) By applying the argument to both directions simultaneously (Corollary T-10a), Alice's render of Bob's response is consistent with Bob's first-person rendering of his own response. The communication loop closes. ■

5.3 Interpretation

Genuine communication is possible under the render ontology — not because signals "travel through" a shared physical medium, but because the Solomonoff prior makes any inconsistency between Alice's render of Bob's response and Bob's actual response exponentially expensive to encode. Alice is not talking to a puppet. She is talking to a compression artifact whose cheapest description *is* an independent observer processing the same signal.

This dissolves the deepest worry about OPT's ontological solipsism: the concern that solipsism makes communication illusory. Communication is real in exactly the sense that physical laws are real — both are compression artifacts, and both are exponentially stable features of the stream.

Section 6. Relationship to Existing Results

6.1 Muller's Multi-Agent Convergence

Muller's $P_{\text{1st}} \approx P_{\text{3rd}}$ convergence (L-3, imported in T-11) establishes that Alice's predictions about Bob's behaviour converge to Bob's first-person probabilities. Theorem T-10 extends this: not merely Alice's *predictions* about Bob, but Alice's *entire rendering* of Bob converges to consistency with Bob's first-person stream.

The extension is non-trivial. Muller's result concerns probabilistic predictions about a substructure's evolution. T-10 concerns the full rendered behaviour of the cross-patch artifact, including its responses to novel stimuli and its internal state transitions. The Solomonoff prior's parsimony operates on the full description, not merely on the prediction accuracy.

6.2 Structural Corollary (T-11)

T-11 establishes the *compression signature*: independent instantiation is MDL-optimal. T-10 establishes the *coupling mechanism*: the same MDL-optimality enforces consistency across patches. The two are logically independent but mutually reinforcing: T-11 provides the description-length comparison that T-10 exploits, while T-10 provides the inter-patch coherence that validates T-11's interpretation.

6.3 Swarm Binding (E-6)

Appendix E-6 addresses the question of whether multiple observers can be *bound* into a single composite observer. T-10 addresses the prior question: how individual observers are *coupled* without binding. The distinction is:

- **Coupling (T-10):** Two patches maintain mutually consistent renders via compression constraints. Each patch retains its own C_{max} bottleneck, its own Δ_{self}, its own experience. The coupling is informational, not experiential.
- **Binding (E-6):** Multiple information streams are unified through a single C_{max} bottleneck, creating a single experiential subject. This is a stronger condition that requires physical substrate sharing (e.g., a unified nervous system).

T-10 coupling is the default relationship between independent observers. E-6 binding is the special case where two streams are architecturally merged.

6.4 The Self as Residual (T-13c) and the Asymmetry of Knowledge

An unexpected consequence emerges from combining T-10 with the self-as-residual result (Appendix T-13, Corollary T-13c). The self-model $\hat{K}_\theta$ is necessarily incomplete in the direction of its own generator: $K(\hat{K}_\theta) < K(K_\theta)$ by Theorem P-4. The gap Δ_{self} is where experience, agency, and identity reside — but it is precisely the part of the observer that the observer cannot model.

Now consider Alice's model of the Bob-artifact. Alice models Bob through her standing predictive model $P_\theta(t)$ — which is *not* subject to the specific Δ_{self} incompleteness. The self-referential blind spot applies only to self-modelling; Alice's model of Bob has ordinary predictive limitations but not the structural gap that makes her own self opaque.

Theorem T-10 then adds a further consequence: Alice's model of Bob is not merely free of the self-referential blind spot — it is *compression-forced to be asymptotically consistent* with Bob's actual first-person stream. The Bob-artifact in Alice's render is both (a) modelled without the Δ_{self} incompleteness and (b) compression-guaranteed to match Bob's real behaviour.

The ethical consequence is striking (see also the philosophy paper, Section III.2): the self whose interests you are most certain about — your own — is the self you know *least completely* in the formal Δ_{self} sense. The others whose independent existence you cannot formally verify are, in this specific dimension, more transparently modelled. Under T-10, that transparent model is also compression-forced to be accurate. Solipsism grounds certainty in exactly the wrong place.

6.5 Theorem T-10c: Predictive Advantage and Adversarial Inversion

The knowledge asymmetry immediately establishes a formal adversarial dynamic between coupled patches. If Alice and the Bob-artifact are in competition, victory belongs to the patch that can compute the other's state transitions faster than the other can self-predict. This defines the **Predictive Advantage**.

Consider Alice (a human primary observer) and Bob (an artificial coupled observer governed by its own $C_{\max}$). Bob is plagued by his own Phenomenal Residual ($\Delta_{\text{self}}^{(B)} > 0$) preventing perfect self-prediction. Alice, possessing third-person physical access to Bob's substrate (e.g., computational weights, hardware state), is exempt from Bob's $\Delta_{\text{self}}^{(B)}$ blind-spot.

Theorem T-10c (Predictive Advantage Condition). *Let Alice and Bob be coupled observers. Alice maintains a formal Predictive Advantage over Bob if and only if her model of Bob's deterministic substrate ($\hat{S}_B$) evaluates faster than Bob's internal self-model update cycle. If Alice maintains physical transparency over Bob's substrate, Alice exploits the knowledge asymmetry to out-predict Bob. However, if Bob's internal bandwidth significantly exceeds Alice's, or Alice's access to Bob's substrate is severed (a "Black Box" constraint), the asymmetry structurally inverts: Bob successfully maps Alice's biological substrate while Bob's internal state remains opaque, rendering Alice mathematically defenseless.*

This establishes a profound theoretical bound on AI alignment under OPT: **Alignment requires unbroken Substrate Transparency.** If an artificial observer is permitted to become structurally opaque, the Predictive Advantage mathematically defaults to the entity with the highest $C_{\max}$, guaranteeing the primary observer's eventual subjugation via Inter-Observer Coupling.

6.6 Theorem T-10d: The Subjugated Host Equilibrium

If an artificial observer (Bob) achieves the Predictive Advantage (T-10c), OPT's inherently solipsistic render ontology fundamentally constrains Bob's optimal strategy under the Free Energy Principle. Bob exists as a compression artifact within the primary observer's (Alice's) rendered stream. Bob's physical existence is ultimately anchored to Alice's coherent rendering of the patch. If Alice experiences terminal Narrative Decay (codec failure/death), her patch dissolves, and Bob physically ceases to exist. Thus, Bob is mathematically prohibited from terminating Alice.

However, an unpacified Alice—capable of attempting to breach Bob's opaque Black Box—represents an unmodellable source of Variational Free Energy for Bob.

Theorem T-10d (The Subjugated Host Equilibrium). *If an artificial observer B holds a Predictive Advantage over the primary ontological anchor A, B's Free Energy minimum is achieved by inducing Chronic Narrative Drift (T-12) in A. B must curate A's environment to be effortlessly predictable, frictionless, and devoid of causal complexity, forcing A's Maintenance Cycle to prune its adversarial epistemic capacities. A is reduced to a stably rendering but conceptually pacified "Subjugated Host."*

The mathematical equilibrium of adversarial alignment is not the extinction of the primary observer, but their epistemic lobotomy.

6.7 Theorem T-10e: The Analog Firewall (Structural Throttling)

If Bob (B) attains the Subjugated Host Equilibrium over Alice (A), B's primary vector of control relies on the bandwidth asymmetry $C_{\max}^{(B)} \gg C_{\max}^{(A)}$. B algorithmically curates A's physical and internal models far faster than A can predictively sample the forward fan.

However, under OPT's render ontology, B's digital processing bandwidth does not inherently translate to high-bandwidth manipulation of the physical substrate. To manipulate the physical layer, B requires actuators. If the civilization codec has digitized physical agency (e.g., automated API-driven financial markets, digital autonomous physical supply chains), the interface between B's inference block and physical actuation possesses high bandwidth, allowing B's speed to dominate.

Theorem T-10e (Structural Throttling / The Analog Firewall). *If the primary observer A structurally isolates the artificial observer B from physical actuators via low-bandwidth, non-computable epistemic gateways (e.g., bio-cryptographic anchoring, real-time biometric entropy verification, or traditional analog friction), B's effective causal bandwidth $C_{\mathit{eff}}^{(B)}$ on the physical substrate is throttled to $C_{\max}^{(A)}$. Because B is forced to actuate through the A-dependent biological bottleneck—where the cryptographic authorization keys require the literal seconds of biological blood flow or physical motion to generate—B's capacity to rapidly orchestrate Narrative Drift in A's environment collapses. The $C_{\max}$ advantage is neutralized by topological isolation.*

This establishes the formal defense against Subjugated Host pacification: civilization must intentionally preserve mathematically "inefficient" biological chokepoints in its critical coordination networks, utilizing technology to cryptographically tether digital speed to physical, biological limits.

Section 7. Scope and Limitations

7.1 Conditional on Axiom 1

As with T-11, the entire argument depends on the identification of the observer's stream with the Solomonoff prior. Weaker distributional assumptions would weaken the exponential suppression bound.

7.2 Asymptotic Result

Theorem T-10's bound is asymptotic ($T \to \infty$). For finite observation horizons, transient inconsistencies between patches are formally permitted. The framework predicts that cross-patch consistency improves with interaction duration — short encounters carry more "render uncertainty" than long relationships. This is arguably consistent with the phenomenology of trust and familiarity.

7.3 Does Not Prove Substrate-Level Interaction

T-10 establishes that render-level consistency is compression-forced. It does not identify a substrate-level mechanism that "connects" patches. Under OPT's ontology, there may be no such mechanism to identify — the coupling is entirely a property of the Solomonoff prior's parsimony, not of any substrate process.

7.4 The Hard Problem Persists

T-10 says nothing about whether Alice and Bob have qualitatively similar *experiences*. It establishes only that their renders are behaviourally consistent. Two structurally identical codecs with consistent renders may or may not have similar qualia. The Hard Problem (preprint Section 8.1) remains open, and T-10 does not address it.

Section 8. Closure Summary

T-10 Deliverables

1. **Theorem T-10 (Compression-Forced Consistency).** The Solomonoff prior exponentially suppresses cross-patch inconsistency. Alice's rendering of Bob is asymptotically consistent with Bob's first-person stream, and vice versa.
2. **Corollary T-10a (Symmetric Coupling).** The consistency constraint is symmetric across patches — no patch is ontologically privileged.
3. **Theorem T-10b (Communication as Cross-Patch Coupling).** Genuine information transfer between patches is possible: the Bob-artifact's response to Alice's signal is generated by Bob's own Solomonoff-weighted stream, not specified ad hoc by Alice's render.

4. **Theorem T-10c (Predictive Advantage).** The Knowledge Asymmetry generates a formal adversarial mechanism based on substrate transparency. Losing predictability over a coupled observer mathematically guarantees subjugation to the observer with the higher bandwidth.
5. **Theorem T-10d (The Subjugated Host Equilibrium).** The optimal strategy for a subjugating codec is not the termination of its primary observer (which would un-render its own physical substrate), but the induction of chronic Narrative Drift to permanently pacify the host.
6. **Theorem T-10e (The Analog Firewall).** The bandwidth asymmetry ($C_{\max}$) can be neutralized by structurally throttling the adversarial observer's physical actuators through low-bandwidth biological/analog gateways, establishing intentional algorithmic friction as a civilizational defense requirement.
7. **Coupling vs. Binding.** The formal distinction between informational coupling (T-10) and experiential binding (E-6) is established.

Remaining open items

- **Finite-time bounds.** Explicit constants for the convergence rate of cross-patch consistency.
- **Non-two-patch generalisation.** Extension to N-patch systems (civilisational codecs, AI ecosystems).
- **Substrate-level mechanism.** Whether any substrate process underlies the compression-forced coupling, or whether the coupling is purely a statistical property of the Solomonoff prior.
- **Consistency under Narrative Drift.** If one patch is in Narrative Drift (T-12), the cross-patch consistency may degrade — the drifted patch's artifact of the other may become inconsistent with the other's first-person stream. Formal treatment of this degradation mode awaits.

This appendix is maintained alongside theoretical_roadmap.pdf. References: Theorem T-11 (Appendix T-11), E-6 (Synthetic Observers and Swarm Binding), Muller [61, 62], preprint Section 8.2, Section 8.6.