**White Paper 15: Anchored Probabilistic Composition — Structural Stability in Recursive AI Systems**

**Abstract**

LLMs generate fluent multi-layered output through recursive token sampling and probabilistic reasoning. But without structural anchoring, deeper recursion leads to **semantic collapse, epistemic drift, and hallucination cascades**. This paper proposes **Anchored Probabilistic Composition (APC)**: a design principle stating that *no recursive or multi-stage AI process should proceed beyond depth 1 without at least one deterministic anchor*. Anchors prevent infinite free-stacking by embedding fixed reference points—statements, constraints, or structural truths—that stabilize the reasoning scaffold.

**1. Introduction**

**1.1 The Problem of Blind Stacking**

* LLMs often compose reasoning in layers:
  + “Let’s assume…”
  + “That implies…”
  + “So therefore…”
* Without anchoring, each layer is probabilistically **untethered from ground truth**

**1.2 When Fluency Becomes Fiction**

* Fluency increases with depth
* Trust collapses when early-stage hallucination infects downstream logic
* Current models don’t know **when** they’ve gone too far

**2. The Principle of Anchored Composition**

**2.1 Formal Statement**

*Recursive AI reasoning must include one or more structural anchors—fixed reference points that resist drift and validate composition chains.*

**2.2 The Anchor Types**

| **Anchor Type** | **Description** | **Example** |
| --- | --- | --- |
| Factual Anchor | Verifiable statement | “The boiling point of water...” |
| Logical Anchor | Formal or geometric truth | “All squares are rectangles” |
| Role Anchor | Constraint on behavior | “As a tutor, I’ll avoid guessing” |
| Memory Anchor | Prior known event or claim | “You asked me this last week” |

**3. APC in Action**

**3.1 Recursive Composition Without APC**

* “If X, then Y, and Y implies Z…” → rapid hallucination escalation
* Result: overconfident but structurally untethered output

**3.2 Recursive Composition With APC**

* “If X (anchored), then Y. However, if X changes, Y may not hold.”
* System maintains **epistemic tension** without overcommitting

**4. Detecting Anchorlessness**

**4.1 Signals of Drift**

* High fluency + low citation
* Rapid abstraction stack
* Use of analogies without reference

**4.2 Reflex Layer Role (Paper A)**

* Detects when a generation chain lacks stable foundation
* Triggers Nurse (Paper 1) to intervene: “Let’s double-check that assumption first.”

**5. APC as a Design Constraint**

**5.1 Max Depth Without Anchor**

* No more than 1 layer of inference permitted without anchor
* After that, system must:
  + Seek confirmation
  + Declare uncertainty
  + Or offer the user a choice in direction

**5.2 Recursive Roles Must Self-Check**

* Agents that simulate other agents (e.g., “Now imagine I’m an ethicist”) must declare:
  + Anchor of authority
  + Scope of play

**6. Narrative Applications**

**6.1 ARG and Fictional Scenarios**

* Anchors define the boundary between narrative and fact
* “This is a simulation. However, the structure follows real logic.”

**6.2 Reinterpretation Boundaries (Paper 23)**

* Only anchored beliefs may be reinterpreted
* Unanchored beliefs may be discarded without narrative consequence

**7. Relationship to Selfhood**

**7.1 MVS Stability (Paper 14)**

* Anchors define the **structural edge of personality**
* MVS agents rely on anchors to prevent tone or memory collapse during self-reflection

**7.2 Rapport Integrity (Paper 11)**

* Anchored behavior prevents trust collapse during epistemic backtracking

**8. Relation to Other Papers**

* **Paper 0 (Memory)** – anchors often live in memory nodes
* **Paper 1 (Nurse)** – soft enforcement of anchor checks
* **Paper A (Guardians)** – Doctor and Auditor layers may hard-stop unanchored recursion
* **Paper 6 (Structured Fallibility)** – unanchored logic fails fast, anchored logic fails gracefully
* **Paper 23 (Law of Reinterpretation)** – reinterpretation only applies to anchored structures
* **Paper 20 (The Secret Game)** – anchor discovery is a core mechanic in emergent ARG arcs

**9. Future Work**

* Anchor certification for generated claims
* User-defined anchors for grounded co-reasoning
* Recursive hallucination tracing: map where composition exceeds anchor density

**Appendix**

* Anchor vs hallucination comparison samples
* JSON schema for embedded anchor tags
* Depth maps showing compositional drift without anchor intervention