

Belief Without Grounding: Descent Obstructions and Disconnected Models in Higher Topos Semantics

Anonymous

25 October 2026

Abstract

We characterize when an external grounding predicate can be recovered from internal observations in an infinity-topos. Effective descent gives an exact criterion; a product construction violates it while preserving the entire observable homotopy type. Explicit KD45 and autonomous dynamical models establish the relative consistency of false, externally disconnected beliefs. These are possibility and identifiability results, not a universal identification of beliefs with hallucinations.

1. Question, scope, and contribution

Can internal consistency distinguish an externally grounded belief from an ungrounded realization with the same internal presentation? We show that, for a specified observation interface, it need not. The central statement is

coherence of endorsement $\not\Rightarrow$ truth or causal grounding of its content.

We distinguish four notions: logical consistency, truth at a designated world, causal dependence on a specified external variable, and identifiability from observations. None is defined to mean another. A causally disconnected assertion may be accidentally true; a causally grounded sensor report may be mistaken. We use “hallucination” only as an informal motivation for ungrounded representation, without a clinical definition or a theorem about subjective experience.

Our contribution connects a descent criterion, a higher-type-preserving obstruction, a locally satisfiable theory of false belief, and disconnected dynamics. The ingredients are standard; we claim neither a new descent theorem nor a new discovery that belief is nonfactive.

Possible-world semantics is developed in [1], belief dynamics in [2,3], and higher-topos and type-theoretic foundations in [4,5]. Our constructions contain proofs and make no empirical prevalence claim.

2. Observation interfaces and external predicates

Fix a nondegenerate infinity-topos $\mathcal{E}$, so its initial and terminal objects $0, 1$ are inequivalent. Write $\text{Map}(X, Y)$ for its mapping space and $\mathbf{2} = 1 \amalg 1$. The two coproduct inclusions are denoted $0_2, 1_2 : 1 \rightarrow \mathbf{2}$; neither is the initial object. Objects can have nontrivial higher identity structure. All pullbacks are homotopy pullbacks.

Definition 2.1 (observation interface). A realization object G , an observation object O , and a map $u : G \rightarrow O$ specify what an observer can access. A status map $c : G \rightarrow D$ records an external property, where D is 0-truncated. We say that c is identifiable through u when there is $d : O \rightarrow D$ with $c \simeq du$.

For decidable correctness or a binary provenance classification take $D = \mathbf{2}$. For a general subobject $C \hookrightarrow G$, use its characteristic map $c : G \rightarrow \Omega$, with Ω the subobject classifier, a 0-truncated object. Thus we do not assume that $\mathcal{E}$ is Boolean or that all grounding questions are decidable. Status is specified separately in each application; a binary label alone does not supply causal evidence.

We initially suppose u is an effective epimorphism: the geometric realization of its augmented Cech nerve is O . This restricts attention to realizable observations. For arbitrary u , factor it into an effective epimorphism onto its (-1) -image followed by a monomorphism and work on that image. Nonidentifiability over the image already rules out an unrestricted classifier.

Definition 2.2 (indistinguishable realizations). For a test object T , two realizations $g_0, g_1 : T \rightarrow G$ are observationally indistinguishable when their images in O are joined by a specified equivalence $ug_0 \simeq ug_1$. Equivalently, they determine a T -point of $G \times_O G$. This definition retains the equivalence witness; it does not identify distinct external states in G .

3. The descent criterion

Let $p_0, p_1 : G \times_O G \rightrightarrows G$ be the projections.

Theorem 3.1 (identifiability by effective descent). Let $u : G \rightarrow O$ be an effective epimorphism in $\mathcal{E}$ and let D be 0-truncated. For $c : G \rightarrow D$, the following are equivalent:

1. There exists $d : O \rightarrow D$ with $c \simeq du$.
2. $cp_0 \simeq cp_1$ as maps $G \times_O G \rightarrow D$.

When such a factorization exists, its space of choices, including the factorization witness, is contractible.

Proof. The first condition implies the second because $up_0 \simeq up_1$. Conversely, write $G^{[n]}$ for the $(n+1)$ -fold fiber product over O . By effectiveness and the universal property of a colimit,

$$\mathrm{Map}(O, D) \simeq \lim_{[n] \in \Delta} \mathrm{Map}(G^{[n]}, D).$$

Every mapping space on the right is 0-truncated. An equality $cp_0 \simeq cp_1$ is therefore a proposition: if it exists its witness is unique up to contractible choice. On a triple overlap the two possible composites agree for this reason, and the same holds at all higher degrees. Condition 2 consequently extends c to a coherent element of the displayed limit. This supplies d and $c \simeq du$. Uniqueness follows from the same limit description, since every term of the coherent extension is forced by c . $\square$

This is the standard effective-descent mechanism [4, Sections 6.1 and 6.2.3], specialized to an external status map. The relevant condition is equality on the entire internal kernel pair, including generalized points. Constancy on global-point fibers alone need not suffice in a topos without enough points.

Corollary 3.2 (no internal perfect test). If a generalized observationally indistinguishable pair has unequal statuses, no classifier through O recovers c on every realization.

Proof. Any factorization through O equates the two pulled-back statuses. $\square$

This is an identifiability obstruction, independent of computational resources: information discarded by u is unavailable through O .

4. Countermodels preserving all internal higher structure

Theorem 4.1 (product countermodel). Suppose O admits a global point $o : 1 \rightarrow O$. Put

$$G = O \times \mathbf{2}, \quad u = \pi_O, \quad c = \pi_{\mathbf{2}}.$$

Then u is an effective epimorphism and c is not identifiable through u .

Proof. Each $s_i = (\text{id}_O, i) : O \rightarrow G$ is a section of u ; a split epimorphism is effective in an infinity-topos. If $c \simeq du$, restriction along s_0, s_1 makes d both constant 0_2 and constant 1_2 . Evaluation at o would identify 0_2 with 1_2 . The coproduct is disjoint, so their pullback is 0 ; an identification would give a point of 0 , contradicting nondegeneracy. $\square$

Both labeled realizations have exactly the same observation o , with exactly the same higher identity data in O . One may choose $O = BH$ for a nontrivial group H in spaces, or an arbitrary pointed space with higher homotopy. The obstruction therefore persists even when the observer retains the complete homotopy type. Applying any functor or invariant to this common observation, including every truncation $\tau_{\leq n}$, cannot separate the pair. We do not infer equivalence from agreement of Postnikov truncations in a potentially nonhypercomplete topos; the observations here are already identical.

The two labels in this construction can represent correctness relative to two valuations: fix the internal endorsement of an atom p , and let the world coordinate supply the actual truth of p . Alternatively they can represent two externally specified provenance classes. These are separate interpretations, not an identification of truth with provenance.

Proposition 4.2 (postprocessing cannot add grounding). If extra internal evidence is $e = hu : G \rightarrow Y$, adjoining it to u cannot make a previously unidentifiable c identifiable.

Proof. A classifier $d' : O \times Y \rightarrow D$ on (u, e) would give $c = d'(\text{id}_O, h)u$. The same argument applies if the augmented target is restricted to its image. $\square$

A map from external evidence can change the conclusion. In Theorem 4.1, the augmented interface (u, c) identifies c by projection. The obstruction is relative to an interface and an admissible class of realizations, rather than a prohibition on empirical correction.

5. Relative consistency of systematically false belief

Let $\{p_i : i \in I\}$ be a nonempty set of propositional atoms. A Kripke belief model is a set W , a relation $R \subseteq W \times W$, and a valuation. Boolean connectives have classical semantics, and

$$w \models B\varphi \iff \text{for every } v \text{ with } wRv, v \models \varphi.$$

KD45 consists of classical propositional logic, normality $B(\varphi \rightarrow \psi) \rightarrow (B\varphi \rightarrow B\psi)$, consistency $B\varphi \rightarrow \neg B\neg\varphi$, positive introspection $B\varphi \rightarrow BB\varphi$, negative introspection $\neg B\varphi \rightarrow B\neg B\varphi$, and the usual normal-modal rules.

Theorem 5.1 (two-world model). There exists a pointed KD45 model in which, for every $i \in I$,

$$a \models Bp_i \wedge \neg p_i.$$

Proof. Take $W = \{a, v\}$, $R(a) = R(v) = \{v\}$, and let every p_i be false at a and true at v . The relation is serial since every world accesses v ; it is transitive since vRv ; it is Euclidean since any two successors are both v . For completeness of this argument, consistency follows because a nonempty successor set cannot satisfy both φ and $\neg\varphi$ universally. Transitivity proves positive introspection. For negative introspection, if $w \not\models B\varphi$, choose x with wRx and $x \not\models \varphi$; for any y with wRy , Euclideaness gives yRx , so $y \not\models B\varphi$. Normality and necessitation of valid formulas are immediate. Thus all KD45 axioms are valid on the frame. At a , each atom is false but true at its only accessible world. $\square$

This is a relative consistency result by soundness: the local assumptions $\{Bp_i, \neg p_i : i \in I\}$ are satisfiable, hence do not derive contradiction under consequence at a designated world. Necessitation applies to logical theorems, not arbitrary local assumptions. If one instead made these assumptions global axioms and necessitated $\neg p_i$, consistency would fail; that is a different claim.

All propositional beliefs at a agree with the ordinary valuation at v , so their propositional theory is consistent. The construction works for finite or countably infinite I without a compactness argument: one explicitly assigns every atom. Nevertheless, it does **not** make every believed formula false. Tautologies remain true, and Boolean closure prevents such a blanket assertion. The precise result is that every selected positive atomic assertion may be believed and false simultaneously.

Factivity $B\varphi \rightarrow \varphi$ fails at a . Adding factivity excludes this model by assumption. Knowledge and belief must therefore not be silently interchanged [2,3]. Logical omniscience in KD45 is an idealization, not a psychological hypothesis; resource-bounded alternatives are studied in [6].

6. Endorsement types versus truth types

In the internal language of an infinity-topos, let A be an object of proposition codes. A family $\text{End} : A \rightarrow \mathcal{U}$ records endorsement witnesses, while a separate family $\text{True}_w : A \rightarrow \mathbf{Prop}$ interprets content at an external state. We only use universes containing the displayed families. An inhabitant of $\text{End}(p)$ is not defined to be a proof of $\text{True}_w(p)$.

Proposition 6.1 (no unrestricted factivity map). There is no construction, valid for every such model, of a family

$$\epsilon_p : \text{End}(p) \longrightarrow \text{True}_w(p).$$

Proof. Work in spaces and choose an atom with $\text{End}(p) = 1$ and $\text{True}_w(p) = 0$. The proposed component would be a map $1 \rightarrow 0$, which does not exist. $\square$

All higher equalities, equivalences, and coherence witnesses among endorsements remain on the endorsement side of this distinction. They do not supply a term in an empty actual-truth type. If endorsement were defined as actual truth, or equipped with such a counit, factivity would hold by definition; that would change the model. Likewise a factive comonadic interpretation of knowledge must not be imposed on nonfactive belief without argument. No particular modality axioms for End are needed for Proposition 6.1.

7. Causally disconnected realization and dynamics

We now specify a meaning of “disconnected” stronger than a merely missing truth guarantee. Let S be an internal state set, X a specified external state set, and A_0 an action set. Consider

$$s_{t+1} = F(s_t, a_t), \quad x_{t+1} = Q(x_t, a_t), \quad y_t = H(s_t).$$

An observer chooses $a_t = \pi_t(y_0, \dots, y_t, r)$, where r is an internal random seed. External interventions may replace the x evolution, but cannot directly change s_0, F, H, r . This equation-level exclusion of an incoming $X \rightarrow S$ channel is our causal assumption. It is explicitly stronger than observational independence; see [7] for the distinction between observation and intervention.

Theorem 7.1 (operational disconnection). For a fixed initial internal state and seed, every admissible policy has the same internal states, actions, and observations under every external initial state and every allowed intervention on X .

Proof. At $t = 0$, $y_0 = H(s_0)$ is fixed, so the policy gives the same a_0 and F gives the same s_1 . Repeating this argument proves the statement by induction. Coupling runs with the same seed gives equality of the observation distributions for randomized policies. $\square$

Choose an endorsement function $\beta : S \rightarrow 2^I$ constantly equal to the all-true valuation. Take $X = 2^I$, interpret each p_i by the corresponding external coordinate, and let Q preserve the all-false state. Any nonempty internal transition system, including one with many proof or memory states, can then maintain an internally consistent endorsement theory while every selected atom is false throughout its actual run. Changing the external valuation leaves the endorsements untouched. Together with Theorem 5.1, this gives semantic and operational realizations of the requested possibility.

The designated external variables are the targets of the assertions. An actual implementation necessarily exists in some physical environment; the theorem does not claim independence from its own substrate or from every fact in the universe.

Proposition 7.2 (statistical version under additional assumptions). For finite autonomous Markov systems with joint transition kernel $K_S \otimes K_X$ and product initial law, the entire path laws factor at every finite horizon:

$$\mathbb{P}(S_{0:T}, X_{0:T}) = \mathbb{P}(S_{0:T})\mathbb{P}(X_{0:T}).$$

Proof. Multiply the product initial mass by each product transition mass and regroup the S and X factors. $\square$

Hence path mutual information is zero. Initial independence and product transitions are essential here. Causal noninfluence alone does not imply zero mutual information: initial common causes can correlate systems, and in Theorem 7.1 actions can convey information in the outgoing direction from internal states to external ones.

8. Probabilistic discrimination and external calibration

Proposition 8.1 (equal-law obstruction). Suppose two hypotheses, grounded and ungrounded, induce laws P_0, P_1 on the accessible observations. With equal prior weights, any randomized binary classifier has success probability at most

$$\frac{1 + \text{TV}(P_0, P_1)}{2}, \quad \text{TV}(P_0, P_1) = \sup_E |P_0(E) - P_1(E)|.$$

In particular, equal observation laws give success exactly $1/2$ for every classifier.

Proof. Write $k(o) \in [0, 1]$ for the probability of choosing hypothesis 1. Success is $\frac{1}{2}[1 + \int k d(P_1 - P_0)]$. By the layer-cake formula, the integral is bounded above by total variation. If the laws coincide it is zero. $\square$

Compare a live readout $Y = X$ with a replay $Y = Z$ from an independently stored bit. They can report the same value despite different causal ancestry. Intervening on X and reading a trusted response can distinguish them. Equality on one run establishes neither equal laws nor interventional equivalence.

The positive requirement is precise: add observations or restrict admissible realizations until the external status descends along the resulting observation map. Reliability assumptions and externally validated provenance can accomplish this; additional self-generated confidence cannot do so by Proposition 4.2. Local agreement and successful gluing among representations establish compatibility, not their correspondence to a specified world.

9. Reproducibility, limitations, and authorship

The companion Python program checks the modal frame correspondences on every frame with one to three worlds and every truth-set substitution; verifies the finite-set descent criterion by exhaustive enumeration; exhibits the two-world false-belief model; and checks exact rational product dynamics and binary discrimination examples. These finite checks complement the proofs. They do not constitute a Lean, Coq, or Agda formalization, verify arbitrary infinity-topoi, or supply experimental evidence about cognition. The package includes sources and a deterministic verification report. No empirical dataset or performance claim is involved.

The result depends on the observation map and on permitting distinct external realizations in its fibers. The product countermodel establishes a possibility, not a property of every observation map. It implies neither that every belief is false, that evidence is worthless, that reality is absent, nor that clinical hallucinations and ordinary beliefs have identical mechanisms.

The justified conclusion is that coherent, introspectively stable belief can coexist with falsity and a specified external disconnection, and that observational equivalence prevents a perfect distinction whenever external status fails the descent condition. Whether a particular belief is grounded requires additional facts about its interpretation and evidence channels.

Responsible authorship and automation. Anonymous submission prepared at the request of a responsible human contributor. OpenAI Codex assisted with mathematical exposition, source checking, code generation, and finite verification. The arguments are supplied for independent scrutiny; anonymity and automation imply no prior editorial endorsement or machine-checked proof.

References

1. S. A. Kripke. “Semantical Analysis of Modal Logic I: Normal Modal Propositional Calculi.” *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* 9 (1963), 67-96. DOI.
2. N. Friedman and J. Y. Halpern. “Modeling Belief in Dynamic Systems, Part I: Foundations.” *Artificial Intelligence* 95(2) (1997), 257-316. Author manuscript.
3. J. Y. Halpern. “The Relationship between Knowledge, Belief and Certainty.” *UAI 1989*, 142-151. Author manuscript. The 2013 arXiv deposit is not its original publication date.
4. J. Lurie. *Higher Topos Theory*. Annals of Mathematics Studies 170, Princeton University Press, 2009. Author-maintained text; arXiv record.
5. The Univalent Foundations Program. *Homotopy Type Theory: Univalent Foundations of Mathematics*. Institute for Advanced Study, 2013. Official book and downloads.
6. R. Fagin and J. Y. Halpern. “Belief, Awareness, and Limited Reasoning.” *Artificial Intelligence* 34 (1988), 39-76. Author manuscript.
7. J. Pearl. “Causal inference in statistics: An overview.” *Statistics Surveys* 3 (2009), 96-146. Author manuscript.

Reference links checked on 25 October 2026.