

Defining the Observer

A General Physical Criterion, Its Relation to Existing Models, Impossibility of the final ascent, and the Impossibility of the Ideal Limit

Sergej Materov* · 2026

Abstract

No treatment of quantum gravity that takes observers seriously can avoid the question of what an observer physically is, and existing treatments - including the author's own (Materov 2026a,b,c) - each state a modelling assumption for the observer without deriving it. This paper's contribution is narrow and specific, not a claim to be the first physical account of an observer: a particular combination of requirements (D1–D4) on the trajectory of an embedded tensor factor, a separately-argued self-capacity criterion for individuation (D3') tied to the no-cloning theorem, and a short, self-contained no-go argument (Proposition 4.1) for complete reconstruction of an observer's complement. We show that the observer-notions already in use in Materov (2026a,b,c) relate to this combination in three structurally different ways - as a more general object, as a concrete realization, and as a structurally parallel system not yet unified with it - and work out one concrete case, the photonic Friend of Local-Friendliness experiments, in detail. Since no observer can be ideal in this omniscient sense, we instead investigate uniqueness for a weaker, separately-defined notion (a self-dominant observer, whose own representational capacity exceeds its observing capacity); no dimension-counting argument we have found forces this either, and the question remains, honestly, open.

1. Motivation

No physical theory of quantum gravity can be built on an unexamined notion of “observer”: every serious treatment of the subject - including our own (Materov 2026a,b,c) - states a modelling assumption for what an observer is without deriving it. This is not a flaw specific to those papers; it is the default in the field. The present paper proposes a candidate for that missing prior step: a general, minimal physical criterion for what makes a subsystem an observer at all, from which the observer-notions already in use - a bare tensor factor with access to a local algebra, a system holding a persistent redundant record, a bounded worldline with a causal patch - can be related to a single definition, as a more general object, a concrete realization, or a structurally parallel system requiring a bridging assumption not currently adopted.

Postulate (shared with Materov 2026a,b). No physical observer is external to the physical universe; every observer is itself a physical subsystem, composed of the same kind of stuff as whatever it observes. This is not derived here: no physical theory has ever postulated an observer outside physics. It is the ground this paper stands on, not a result it proves.

1.1 Related work

Physical models of measurement, records, and embedded observers already exist along several lines, though none formulates observerhood itself as a minimal dynamical criterion on the continued evolution of a tensor factor. Fields (2012) requires an observer to satisfy the information-identification requirements of classical automata theory - closest in spirit to (D2)–(D3) below - and, in later work (Fields 2018), analyzes the thermodynamic cost of system identification for a finite embedded observer, a physical constraint this paper does not impose. Quantum Darwinism (Zurek 2009) explains how information about a system becomes objectively accessible through redundant proliferation into its environment - the same physical mechanism underlying (D3)–(D4) below and Materov (2026b)'s persistent record - but asks how classical objectivity emerges for many observers, not what makes a subsystem an observer in the first place. Anderson (2017) treats information as an observer-relative physical quantity, jointly defined by a system and the reference frame that reads it, close in spirit to the relational character of (D1)–(D3) but not formulated as a criterion on the observer subsystem's own dynamics. Closest in aim is Parker (2024), which models an observer inside a fully quantum system using spacetime-localized projection operators and quantum Darwinism, explicitly avoiding an artificial system/observer/environment split - the same target as the Postulate above, reached by a different route.

*Independed Researcher. E-mail: sergejmaterov2@gmail.com

Closest of all, and contemporaneous, is Uchida (2026), which individuates a record-bearing observer relative to a subsystem boundary, domain, timescale, tolerance and resource-bounded operation profile, and characterizes observerhood through four capacities - record acquisition and stabilization, integrated retention across acquisitions, endogenous self-legibility, and counterfactually record-sensitive response - a structure that maps closely onto (D3)–(D4) and (D3') below. Distinctively, Uchida (2026) grounds subsystem structure in a net of local observable algebras rather than an assumed Hilbert-space tensor factorization, precisely to avoid the type-III von Neumann algebra obstruction this paper leaves as an open bridge in §3.3 and §6; on that specific point, its approach is more general than ours. What we take Definition 1 to add, against this closest relative, is not the four-capacities architecture itself but a physically transparent, finite-dimensional trajectory formalism in which a no-go for complete reconstruction (Proposition 4.1) and an explicit uniqueness investigation (§4, Appendix A) can both be carried out concretely - at the cost of the tensor-factorization restriction Uchida (2026) avoids. None of the works surveyed here gives a minimal set of conditions on a subsystem's own finite-dimensional trajectory, together with a proved reconstruction no-go, in the specific combination proposed below.

Status of what follows. For orientation: the Postulate above is assumed, not derived. Definition 1 (D1–D4, §2) is a proposed criterion, not a theorem. The further condition (D3', §2.1) is explicitly additional and less well established than D1–D4. Proposition 4.1 (§4) is a proved result, resting only on Definition 1 and a proposition already proved in Materov (2026a). Proposition 4.1 leaves no ideal observers for a uniqueness question to be about; §4 instead investigates uniqueness for a weaker, separately-defined self-dominant notion, which is not established and is reported as open (§4–§5), not as a theorem. §6 states the scope this paper does and does not claim.

2. Definition 1: the embedded observer

Let $H = H_O \otimes H_E$ be a finite-dimensional closed quantum system with unitary global dynamics U . Throughout, “ O_i ” denotes the pair of a chosen tensor factor and its reduced state ρ_{O_i} - not a vector or operator in isolation. A sequence of proper tensor factors $O_0, O_1, O_2, \dots$ (not necessarily of fixed dimension) is a trajectory of an embedded observer relative to a class of admissible interactions $\mathcal{C}$ if, at each step:

(D1) Embedded. O_i is a proper tensor factor of H : $H = H_{O_i} \otimes H_{E_i}$.

(D2) Unitarily self-evolving under $\mathcal{C}$. Writing $O_{i+1} = U(c) \cdot O_i$ as shorthand: there is a global unitary $U(c)$ (depending on the incoming interaction $c \in \mathcal{C}$) such that $\rho_{O_{i+1}} = \text{Tr}_{E_{i+1}}[U(c)(\rho_{O_i} \otimes \rho_c)U(c)^\dagger]$, with the coupling not a product unitary in the sense of Theorem 4.1 of Materov (2026a) - O_i 's evolution is genuinely informed by c , not autonomous.

(D3) Distinguishable. The map $c \mapsto \rho_{O_{i+1}}$ is non-degenerate on $\mathcal{C}$: distinct $c, c' \in \mathcal{C}$ give $\rho_{O_{i+1}}(c) \neq \rho_{O_{i+1}}(c')$ for a non-trivial subset of $\mathcal{C}$. (Entropy equality $S(\rho(c)) = S(\rho(c'))$ does not imply $\rho(c) = \rho(c')$ - distinct states with the same spectrum satisfy it trivially - so (D3) is stated directly on ρ , not via $S(\rho)$.)

(D4) Branching. The distinction in (D3) propagates: O_{i+2} 's further evolution depends qualitatively - not merely quantitatively, e.g., as a rescaled parameter - on which c occurred at step i .

Persistence of a distinguishing trace along the trajectory (“memory”) is not a fifth condition; it is what it means for (D3)–(D4) to hold over more than one step, and is a matter of degree, not a threshold. (D1)–(D4) are the definition of an embedded observer used throughout the rest of this paper.

Example 2.2 (a minimal D1–D4 instance). Let $H_O = H_{\text{input}} = H_F = \mathbb{C}^2$, with $O_0 = |0\rangle_O$ fixed and $\mathcal{C} = \{c_0, c_1\}$ realized by preparing an input qubit in $|c\rangle$. (D1): O is the tensor factor H_O of $H = H_O \otimes H_{\text{input}} \otimes H_F$. (D2): $U(c)$ is CNOT with the input qubit as control and O as target, giving $\rho_{O_1}(c) = |c\rangle\langle c|$ - genuinely conditioned on c ; CNOT is not a product unitary on the relevant subspace. (D3): $\rho_{O_1}(0) = |0\rangle\langle 0| \neq |1\rangle\langle 1| = \rho_{O_1}(1)$, an exactly distinguishing pair, $r = 2$. (D4): a further, c -independent interaction - CNOT with O_1 as control and a fresh qubit F as target - sends the $c = 0$ history to $O_2 \otimes F = |0\rangle \otimes |0\rangle$ and the $c = 1$ history to $O_2 \otimes F = |1\rangle \otimes |1\rangle$: the two histories remain exactly distinguishable, and whether O_2 becomes correlated with F in this particular way is a qualitative fact about which branch occurred, not a continuously-varying parameter of one fixed response. This is deliberately the simplest possible instance; §2.1's individuation condition and Example 3.2.1 (§3.2) show, symmetrically, how a system can fail (D3') while still satisfying (D1)–(D4).

Remark (trajectories need not continue). Definition 1 does not require the trajectory $O_0, O_1, O_2, \dots$ to be infinite, or even to extend past a single step; (D2)–(D4) apply at whichever steps exist. Termination should be distinguished from gradual degeneration in branching capacity (D4 holding ever more weakly as (D3)'s distinguishing power shrinks): termination is a structural fact - no tensor factor O_{i+1} satisfying (D1) exists that continues O_i at all - not a statement about how well O_i distinguishes its inputs. A concrete example: a radiation sensor whose entire material substrate is vaporized by the same high-energy event it registers, inside a nuclear reactor. The event is an interaction $c \in \mathcal{C}$ satisfying (D2)–(D3) - the sensor's

final state is genuinely, non-autonomously informed by, and distinguishes, what it measured - but the global unitary evolution disperses the sensor's degrees of freedom into the reactor's plasma so thoroughly that no nontrivial factorization $H = O_1 \otimes E_1$ exists with O_1 retaining enough of O_0 's history to count as its continuation. O_0 observes and, in the same interaction, ceases to be identifiable as a subsystem at all. Definition 1 accommodates this without modification: a length-one trajectory satisfies (D1)–(D3) and vacuously satisfies (D4), there being no O_2 for its absence to fail against.

2.1 A further condition: individuation

(D1)–(D4) say when a trajectory of physical states counts as observing at all. They do not say when successive states O_i , O_{i+1} , ... count as the same observer, as opposed to a sequence of freshly prepared instruments that each happen to satisfy (D1)–(D4) once. The condition below addresses that separate question. It is additional to Definition 1, not a fifth member of it: nothing in §3–§4 depends on it, and its physical necessity is less well established than (D1)–(D4) - a detector that fully spends its degrees of freedom on one measurement and only later acquires fresh capacity from its environment is a real physical possibility this condition would, at that moment, exclude.

(D3') Self-capacity. H_{O_i} admits a factorization $H_{record} \otimes H_{self}$ with $\dim(H_{self}) > 1$, where H_{record} is defined as the smallest tensor factor of H_{O_i} such that the map $c \mapsto \rho_{O_{i+1}}$ of (D3) factors through it: there exists a map f from H_{record} alone with $\rho_{O_{i+1}}(c) = f(\pi_{record}(c)) \otimes \sigma$ for some fixed σ on the complementary factor, π_{record} the record obtained from c . Concretely, H_{record} is the span of the (at most r) mutually orthogonal projectors distinguishing the r values (D3) requires - the minimal subsystem that can carry the record without residue - and H_{self} is its tensor complement in H_{O_i} . (D3') requires this *minimal* record factor to leave a non-trivial complement; $\dim(H_{O_i}) > r$ is necessary for this (it fails if $\dim(H_{O_i}) \leq r$) but not sufficient on its own - the minimal record factor must actually be exhibited, not merely shown to be smaller than H_{O_i} on a dimension count.

Remark (motivation for D3', and its limit). The distinction pursued informally in the title of Materov (2026b): observer or instrument? - has a precise answer under (D3'). If no such factorization exists (in particular if $\dim(H_{O_i}) = r$), O 's state after recording c is identical, not merely similar, to that of a bare r -dimensional register holding the same value: nothing survives that is not already determined by c alone, so a third party has no basis, even in principle, for re-identifying O across observations, as opposed to treating each encounter as a fresh instrument. (D3') removes this collapse by requiring residual capacity H_{self} not determined by c - but capacity alone is not sufficient for identity across time: if H_{self} carries a known or fixed state, a third party could simply prepare a fresh instrument in that same state; and even an unknown H_{self} state must itself evolve under (D2)/(D4), so unclonability of a single instant's state does not by itself establish that H_{self} functions as a stable marker of the same O across the trajectory. A candidate mechanism for the first part is the no-cloning theorem (Wootters–Zurek 1982): an unknown state cannot be reproduced by a third party without access to O itself, the same fact that makes quantum money unforgeable (Wiesner 1983). This blocks impersonation by a third party at a single instant; it does not by itself establish diachronic identity, which is left open (§5).

2.1 A further condition: individuation

(D1)–(D4) say when a trajectory of physical states counts as observing at all. They do not say when successive states O_i , O_{i+1} , ... count as the same observer, as opposed to a sequence of freshly prepared instruments that each happen to satisfy (D1)–(D4) once. The condition below addresses that separate question. It is additional to Definition 1, not a fifth member of it: nothing in §3–§4 depends on it, and its physical necessity is less well established than (D1)–(D4) - a detector that fully spends its degrees of freedom on one measurement and only later acquires fresh capacity from its environment is a real physical possibility this condition would, at that moment, exclude.

(D3') Self-capacity. H_{O_i} admits a factorization $H_{record} \otimes H_{self}$ with $\dim(H_{self}) > 1$, where H_{record} is defined as the smallest tensor factor of H_{O_i} such that the map $c \mapsto \rho_{O_{i+1}}$ of (D3) factors through it: there exists a map f from H_{record} alone with $\rho_{O_{i+1}}(c) = f(\pi_{record}(c)) \otimes \sigma$ for some fixed σ on the complementary factor, π_{record} the record obtained from c . Concretely, H_{record} is the span of the (at most r) mutually orthogonal projectors distinguishing the r values (D3) requires - the minimal subsystem that can carry the record without residue - and H_{self} is its tensor complement in H_{O_i} . (D3') requires this *minimal* record factor to leave a non-trivial complement; $\dim(H_{O_i}) > r$ is necessary for this (it fails if $\dim(H_{O_i}) \leq r$) but not sufficient on its own - the minimal record factor must actually be exhibited, not merely shown to be smaller than H_{O_i} on a dimension count.

Remark (motivation for D3', and its limit). The distinction pursued informally in the title of Materov (2026b): observer or instrument? - has a precise answer under (D3'). If no such factorization exists (in particular if $\dim(H_{O_i}) = r$), O 's state

after recording c is identical, not merely similar, to that of a bare r -dimensional register holding the same value: nothing survives that is not already determined by c alone, so a third party has no basis, even in principle, for re-identifying O across observations, as opposed to treating each encounter as a fresh instrument. (D3') removes this collapse by requiring residual capacity H_{self} not determined by c - but capacity alone is not sufficient for identity across time: if H_{self} carries a known or fixed state, a third party could simply prepare a fresh instrument in that same state; and even an unknown H_{self} state must itself evolve under (D2)/(D4), so unclonability of a single instant's state does not by itself establish that H_{self} functions as a stable marker of the same O across the trajectory. A candidate mechanism for the first part is the no-cloning theorem (Wootters–Zurek 1982): an unknown state cannot be reproduced by a third party without access to O itself, the same fact that makes quantum money unforgeable (Wiesner 1983). This blocks impersonation by a third party at a single instant; it does not by itself establish diachronic identity, which is left open (§5).

3. Relation to existing observer-notions

3.1 The No-Go Theorem's tensor-factor observer

Materov (2026a) studies O as a proper tensor factor with access to the local algebra A_O , imposing only (D1) of Definition 1. Conditions (D2)–(D4) are neither required nor excluded: Theorem 4.1 of that paper explicitly allows the degenerate case where the coupling is a product unitary (autonomous O , failing D2), since autonomy is one of the two poles the theorem characterizes, not a case ruled out by the modelling assumption. The relation is therefore not a restriction of Definition 1 but the reverse: No-Go's O is the more general object obtained by working prior to (D2)–(D4) being imposed at all. Proposition 4.1 below (§4) is the specialization of that more general framework to the case O also satisfies (D1) with a non-empty complement - which is automatic.

3.2 Observer-or-Instrument's persistent record

Materov (2026b) gives an explicit operational construction realizing (D3)–(D4): the orthogonality condition $\langle M_h | M_v \rangle = 0$ on distinct memory outcomes is a concrete instance of (D3), and the redundant environmental encoding $|M_h\rangle \rightarrow |M_h\rangle|E_h\rangle$ with $\langle E_h | E_v \rangle \approx 0$ is a concrete instance of the persistence that (D4) requires along a trajectory. This is not a restriction of Definition 1 (dropping conditions) but a realization of it (adding operational structure Definition 1 leaves abstract): Definition 1 nowhere requires that the interaction take the form of a discrete measurement with orthogonal outcomes, only that some non-degenerate, branching map exist. Observer-or-Instrument is one physically motivated way to build such a map; it is not the only one Definition 1 admits.

Example 3.2.1 (the photonic Friend has no self-capacity). The photon used as “Friend” in Local-Friendliness experiments (Proietti et al.; Bong et al., discussed in Materov 2026b) satisfies (D1)–(D3) - it is embedded, its evolution is non-autonomous, and its two polarization outcomes are exactly distinguishable - but fails (D3'): its entire two-dimensional Hilbert space is exhausted by encoding the measurement outcome itself, leaving no residual capacity - no H_{self} that could constitute a persisting “self” distinct from the record. This is a sharper diagnosis than a failure of persistence (D4) alone would give: no factorization $H_{record} \otimes H_{self}$ with $\dim(H_{self}) > 1$ exists inside a two-dimensional H_O , so the failure is at the level of (D3'), before the question of persistence under a given class $\mathcal{C}$ of subsequent interactions can even be posed. Reading this as “the photon is a pure information channel, not an observer” is an interpretation of that formal result, not a further consequence of Definition 1 itself - the definition only pronounces on (D3'), not on what a system that fails it “is”. With that qualification, “instrument”, in that paper's title, is a natural name for exactly this case: an embedded, non-autonomously-evolving, distinguishing physical system that fails (D3').

3.3 Principle of Significance's bounded observer

Materov (2026c) does not use a tensor-factor decomposition at all: the observer is modelled as a bounded worldline with a causal patch $D(O)$, and §4.6.0 of that paper explicitly avoids assuming a global $H = H_O \otimes H_E$ factorization, precisely because continuum quantum field theory across a causal horizon does not admit one (Witten 2018). Condition (D1) as stated therefore does not directly apply to this framework. Connecting the two requires an additional postulate — the ontological Banks–Fischler claim, referred to as S3a in Materov (2026c) - that Materov (2026c) deliberately declines to adopt, deriving only the weaker operational bound $N(O) \leq \exp\left(\frac{A_{max}(O)}{4l_p^2}\right)$ instead (S3b). This paper does not adopt S3a either, and so does not claim Principle of Significance's observer as a case of Definition 1: it is a structurally parallel, geometrically motivated notion of observer capacity, compatible in spirit but not currently unified with (D1)–(D4). Whether such a unification is possible, and at what cost, is left as an open problem (§5).

4. The ideal observer

We define the ideal observer as one that could fully know - reconstruct - the state of its complement. Its impossibility requires no construction beyond Definition 1 and a result already proved in Materov (2026a).

Proposition 4.1 (impossibility of the ideal observer). No embedded observer O (Definition 1, condition D1) can fully know or reconstruct the state of its complement E .

Proof. By (D1), O is a proper tensor factor, hence $d_E \geq 2$. Proposition 3.1 of Materov (2026a) gives $\dim_{\mathbb{R}} F(\Psi) \geq 2(d_E - 1) > 0$. Therefore, O 's data are compatible with infinitely many distinct global states differing on E . Hence O cannot uniquely reconstruct the state of E . ■

This holds regardless of the relative sizes of O and E , and requires no further assumption. The obstruction is therefore informational, not dimensional: increasing $\dim(H_O)$ relative to $\dim(H_E)$ does not remove the nontrivial fibre - Proposition 3.1 holds for every d_O , not only for $d_O \leq d_E$.

Corollary 4.2 (no terminal ascent). Let $O^{(0)} \subsetneq O^{(1)} \subsetneq O^{(2)} \subsetneq \dots$ be any sequence of proper embedded observers (D1), each a tensor factor of some closed system $H^{(n)}$, with nonempty complement $E^{(n)}$ at every stage. Then no $O^{(n)}$ at any finite n is ideal in the sense of Proposition 4.1: each retains a reconstruction defect of dimension $\geq 2(d_{E^{(n)}} - 1) > 0$, regardless of the size, composition, or algebraic type of $O^{(n)}$.

Proof. Immediate: each $O^{(n)}$ satisfies (D1) by hypothesis, so Proposition 4.1 applies to it directly. ■

Remark. This forecloses a specific, recurring class of attempted escapes: enlarging O , changing its internal composition, or passing to a different algebraic description (type I, type II, or a general non-commutative L^1 space) does not terminate the obstruction at any finite stage of such an ascent - the same dichotomy recurs: either the complement is trivial (nothing left to observe), or it is not (the defect persists, essentially unchanged in kind). Whether it can be evaded only in an infinite limit (e.g., an inductive limit of an ascending chain of observer-algebras) is a genuinely different technical question, not addressed here.

On what is new here. The mathematics of the proof is entirely Materov (2026a)'s: Proposition 3.1 is proved, and the informational-not-dimensional reading of it is already implicit there. What Proposition 4.1 adds is the observation that (D1) supplies exactly the hypothesis ($d_E \geq 2$) that proposition needs, and nothing more - so impossibility of the ideal observer follows for *every* system satisfying Definition 1, automatically, without a separate argument. This is a corollary in the precise sense, not a new class of no-go theorem: its content is Materov (2026a)'s; what is new is establishing that Definition 1's observers are, without exception, instances to which it applies.

Proposition 4.1 leaves nothing further to ask about the ideal observer itself: since none exist, asking whether the empty set of ideal observers is unique is not a live question - an empty set is trivially unique. What follows is therefore not an attempt to salvage or refine the ideal observer, but an investigation of a different, explicitly weaker notion, defined for its own sake and named accordingly so as not to reuse a term whose corresponding class Proposition 4.1 has just emptied.

Call O self-dominant if $\dim(H_O)$ exceeds that of its complement (or, in the self-referential form used below, if an observer's own $\dim(H_{self})$ exceeds $\dim(H_{obs})$). A natural question, distinct from anything Proposition 4.1 addresses, is whether at most one embedded observer within a given closed system can be self-dominant - so that the postulate of §1 rules out not just an external observer but a plurality of internal near-omniscient ones. Proposition 4.1 is silent on this: it is a statement about each observer individually - O cannot reconstruct its own complement E - and says nothing about a relation between two different observers O_1 and O_2 , self-dominant or not. Uniqueness of self-dominant observers is therefore an independent question, not a refinement of Proposition 4.1.

We investigated one candidate route to such a relational claim - replacing informational reconstruction with a comparison of Hilbert-space dimensions, global (O against the whole of its complement) and self-referential (an observer's own self-representing capacity against its observing capacity), tested for both disjoint and nested pairs of observers - and found that dimension-counting alone does not deliver it: for nested observers $D(O_1) \subset D(O_2)$ in particular, no linking postulate we could formulate between the two levels' capacity allocations forces exclusion. Appendix A works this out explicitly: under either of the two natural ways of assigning the inner observer's Hilbert space to the outer observer's self or its complement, an explicit finite-dimensional instance exists in which both levels simultaneously satisfy the proposed self-dominance condition - enough to show this route does not force exclusion, though not enough to establish that simultaneous satisfaction is typical across all models of this kind. We report uniqueness as an open problem (§5), not a theorem. The same discipline applies to Corollary 4.2: its proof is immediate from Proposition 4.1, applied once per stage of a chain

- no new mathematics there either. What is new is only the framing as a statement about *chains* of observers rather than a single one, and the Remark's role in foreclosing, by name, four specific evasions - enlarging O 's dimension, removing entanglement, adding degrees of freedom foreign to the constituent subsystems, and passing to a non-Hilbert/von-Neumann-algebra description - none of which is addressed by Proposition 4.1 as originally stated, but all of which reduce to it directly once the chain is set up.

In summary: proper embedded observer $\Rightarrow$ impossible perfect (ideal) reconstruction is proven (Proposition 4.1) - there is nothing left to be unique there; at-most-one-self-dominant-observer is a separate claim, and it is not established - dimension dominance does not by itself imply it. We regard the first as this paper's main result and the second as a genuinely open question (§5), not an unfinished proof.

5. Open problems

The bridge to Principle of Significance. §3.3 leaves open whether, and at what cost, Materov (2026c)'s causal-patch observer can be unified with Definition 1. Doing so by adopting S3a would license $\dim(H_O)$ as a literal, finite Hilbert-space dimension bounded by horizon area - but S3a is an ontological claim Materov (2026c) explicitly leaves open, not a theorem. Doing so without S3a would require reformulating (D3)–(D3') in terms of the operational count $N(O)$ rather than a literal Hilbert-space dimension, which has not been attempted here.

Growth and forgetting. Nothing in Definition 1 bounds $\dim(H_{O_i})$ as i grows: (D2)'s injectivity on $\mathcal{C}$ forces $\dim(H_{O_{i+1}})$ to grow at least in proportion to the number of distinguishable c encountered, so an observer accumulating distinctions indefinitely requires either unbounded capacity or a mechanism for forgetting (Landauer-cost erasure) that Definition 1 does not currently supply. A holographic cap of the kind discussed in §3.3 would give one such mechanism, at the cost noted there.

The scope of $\mathcal{C}$. Definition 1 is stated relative to a class of admissible interactions $\mathcal{C}$ whose own scope - fixed globally for a system, or allowed to vary contextually, as in the photonic-Friend case of §3.2, where persistence relative to a generic decohering $\mathcal{C}$ differs from persistence relative to the specifically constructed $\mathcal{C}$ of a Local-Friendliness protocol - is not fixed by Definition 1 itself. Whether $\mathcal{C}$ should be taken as a further parameter of the definition, or fixed by additional physical input, is left open.

Uniqueness of self-dominant observers. As §4 reports, no dimension-counting argument we have found forces uniqueness for nested or disjoint self-dominant observers, and this appears to be a genuine limit of that method rather than a failure of proof: Proposition 4.1 is a per-observer statement, and dimension comparisons inherit that limitation. A more promising route, not attempted here, would replace static reconstruction with a $\mathcal{C}$ -relative and dynamical notion - e.g. requiring a self-dominant O to also reconstruct every degree of freedom of its complement accessible through $\mathcal{C}$, and to keep doing so at every step under (D2)–(D4), rather than at one instant. Two observers each attempting this simultaneously would need mutually updated models of each other, a self-referential structure that Proposition 3.1's mutual-information identity $I(O:E) = 2S(\rho_O)$ may bound but that we have not worked out. We record this as the candidate direction rather than a claim.

Diachronic identity. (D3') and the no-cloning mechanism discussed in §2 block a third party from impersonating O at a single instant, given knowledge of H_{self} 's state at that instant. They do not establish that H_{self} remains a marker of the same O across a trajectory in which H_{self} itself evolves under (D2)/(D4) - nor do they specify what, if anything, would let O or a third party recognize that continuity. This is left open.

6. Limitations and scope

This paper's claims are bounded as follows. (i) Definition 1 and Proposition 4.1 are stated for finite-dimensional tensor-factor decompositions $H = H_O \otimes H_E$; the generalization to continuum quantum field theory, where such a global factorization is not generally available (Witten 2018, §3.3), is not attempted. (ii) (D3') is proposed as a further, separately-motivated condition (§2.1), not established as physically necessary for observerhood in general - §2.1 gives an explicit physical counterexample to its universal necessity. (iii) Proposition 4.1 settles the ideal observer (none exist); uniqueness of the separately-defined, weaker self-dominant observer is investigated but not proved or disproved - §4 reports a negative result about one method (dimension-counting), not a resolution. (iv) The criterion is stated relative to a class of admissible interactions $\mathcal{C}$, and Definition 1 does not fix how $\mathcal{C}$ itself is chosen or scoped (§5); different physically reasonable choices of $\mathcal{C}$ can change which systems satisfy (D3)–(D4), as §3.2's photonic-Friend example shows. (v) Whether Definition 1 is genuinely novel relative to the full observer-modelling literature has not been established by a systematic review; §1.1 surveys the closest lines of prior work we are aware of, not an exhaustive one.

7. Conclusion

We proposed a general, minimal physical criterion - embeddedness, non-autonomous evolution under a class of admissible interactions, distinguishability, and branching (D1–D4) for what makes a physical subsystem an observer at all, prior to and independent of any particular gravitational or measurement-theoretic model. Relative to this criterion, the observer-notions already in use in Materov (2026a,b,c) turned out to relate to it in three structurally different ways: as a more general object not yet specialized by (D2)–(D4), as a concrete operational realization of it, and as a geometrically motivated but not yet unified parallel construction. A further, separately-motivated condition (D3') let us give a precise, non-speculative answer to the question posed by the title of Materov (2026b) for its simplest case, the photonic Friend of Local-Friendliness experiments. On the ideal observer, impossibility of complete reconstruction is proved outright, in a short, self-contained argument resting on nothing beyond Definition 1 and a proposition already established in Materov (2026a) - leaving no ideal observers for any further question, uniqueness included, to be about. What we investigated in some depth instead is uniqueness for a separately-defined, weaker notion, the self-dominant observer - far enough to rule out dimension-counting, in both its static and inheritance-based forms, as a viable route to it - and reported honestly as open rather than forced to a conclusion the argument does not support. What remains is concrete rather than vague: a bridge to Materov (2026c) contingent on an ontological assumption we have chosen not to adopt, a growth-and-forgetting mechanism Definition 1 does not yet supply, the scope of the interaction class $\mathcal{C}$ itself, diachronic identity of the residual self across a trajectory, and uniqueness under a dynamical rather than static notion of ideality. We take the criterion itself - not any single result built on it - to be this paper's main contribution: a first answer, not the last word, to a question quantum gravity cannot keep postponing.

Appendix A: the nested-uniqueness counterexamples in full

Setup. Let $D(O_1) \subset D(O_2)$ with $H_{D(O_1)} = H_{self_1} \otimes H_{obs_1}$, $\dim(H_{self_1}) = a$, $\dim(H_{obs_1}) = b$. Let $H_{D(O_2)} = H_{D(O_1)} \otimes H_{rest}$, $\dim(H_{rest}) = c$, so $\dim(H_{D(O_2)}) = abc$. O_1 is (self-dominant) ideal iff $a > b$. The outer observer's own self/obs split depends on how the inner lab's Hilbert space $H_{D(O_1)}$ is assigned:

Reading I (O_2 observes the inner lab from outside, as in the Wigner's-friend framing of Materov 2026b, §1): $obs_2 = H_{D(O_1)}$ (dimension ab), $self_2 = H_{rest}$ (dimension c). O_2 is ideal iff $c > ab$.

Reading II (O_1 's material substrate is physically part of O_2 , e.g. neurons inside a body): $self_2 = H_{D(O_1)}$ (dimension ab), $obs_2 = H_{rest}$ (dimension c). O_2 is ideal iff $ab > c$.

Example A.1 (Reading I). $a = 3$, $b = 2$ ($a > b$: O_1 is self-dominant), $c = 7$. $ab = 6 < c = 7$, so O_2 is self-dominant under Reading I. Both O_1 and O_2 are simultaneously self-dominant.

Example A.2 (Reading II). $a = 3$, $b = 2$ (O_1 is self-dominant, as before), $c = 5$. $ab = 6 > c = 5$, so O_2 is self-dominant under Reading II. Both O_1 and O_2 are simultaneously self-dominant. (Note $c = 7$ from Example A.1 does

A.1 Why no linking postulate between the levels rescues uniqueness.

The natural candidate is inheritance: require $self_2$ to contain $self_1$ rather than being chosen freely, i.e., $self_2 = self_1 \otimes self_{extra}$, $obs_2 = obs_1 \otimes obs_{extra}$, with $self_{extra} \cdot obs_{extra} = c$ (the newly available capacity at the outer level splits between an extension of the self and an extension of the observing part). O_1 ideal: $a > b$. O_2 ideal: $a \cdot c_s > b \cdot c_o$, where $c_s \cdot c_o = c$.

Nothing in Definition 1 fixes the split (c_s, c_o) . Taking $c_s = c$, $c_o = 1$ (all new capacity assigned to $self_{extra}$) gives the O_2 -ideal condition $a \cdot c > b$, which holds whenever $a > b$ and $c \geq 1$ - i.e. whenever O_1 is already ideal, some split of any nontrivial extension ($c \geq 1$) makes O_2 ideal too. Simultaneous ideality cannot be excluded by this inheritance postulate for any c .

The universally-quantified (worst-case) version of the same postulate - require O_2 to fail ideality for every possible split - fares no better, in the opposite direction: it demands $a \cdot c_s \leq b \cdot c_o$ for all $c_s \cdot c_o = c$, which already fails at the trivial split $c_s = c_o = 1$ whenever $a > b$ (since $a \cdot c \leq b$ would require $a \leq \frac{b}{c} \leq b$, contradicting $a > b$ for any $c \geq 1$). So, the worst-case version excludes nothing either - it is violated before any freedom in the split is even used.

Both directions of the inheritance postulate therefore fail to force exclusion; the same failure, examined in the main text (§4), holds for every other partition-based linking rule we attempted. We take this as evidence that uniqueness, if it holds at all, is not a fact about how Hilbert-space dimension is partitioned across nesting levels, and requires a postulate of a different (dynamical, not static) kind - the direction sketched, but not developed, in §5.

References

- Anderson, N. G. (2017). Information as a physical quantity. *Information Sciences* 415–416, 397–413.
- Fields, C. (2012). If physics is an information science, what is an observer? *Information* 3(1), 92–123.
- Fields, C. (2018). Some consequences of the thermodynamic cost of system identification. *Entropy* 20(10), 797.
- Parker, B. (2024). Natural probability. *arXiv:2412.04689*.
- Bong, K.-W. et al. (2020). A strong no-go theorem on the Wigner's friend paradox. *Nat. Phys.* 16, 1199–1205.
- Materov, S. (2026a). The Embedded-Observer No-Go Theorem. September 2026. DOI 10.5281/zenodo.23066381.
- Materov, S. (2026b). Observer or Instrument? The Physical Status of the Friend in Local Friendliness Experiments. September 2026. DOI: 10.2139/ssrn.7537882
- Materov, S. (2026c). Toward a Principle of Significance. DOI 10.5281/zenodo.23066624
- Proietti, M. et al. (2019). Experimental test of local observer independence. *Sci. Adv.* 5, eaaw9832.
- Uchida, T. (2026). The Observer as a Record-Bearing Subsystem: Local Access, Integrated Memory, and Record-Conditioned Response Without a Global Reader. *PhilArchive*.
- Wiesner, S. (1983). Conjugate coding. *SIGACT News* 15, 78–88.
- Witten, E. (2018). Notes on some entanglement properties of quantum field theory. *Rev. Mod. Phys.* 90, 045003.
- Wootters, W. K., Zurek, W. H. (1982). A single quantum cannot be cloned. *Nature* 299, 802–803.
- Zurek, W. H. (2003). Decoherence, einselection, and the quantum origins of the classical. *Rev. Mod. Phys.* 75, 715–775.
- Zurek, W. H. (2009). Quantum Darwinism. *Nature Physics* 5, 181–188.