

The Black Hole Information Paradox from the Yoneda Constraint Perspective: Representable Functors, Epistemic Horizons, and the Structure of Hawking Radiation

Matthew Long

The YonedaAI Collaboration
YonedaAI Research Collective
Chicago, IL

matthew@yonedaai.com · <https://yonedaai.com>

February 2026

Abstract

We analyze the black hole information paradox through the lens of the Yoneda Constraint on Observer Knowledge, a category-theoretic principle establishing that an embedded observer \mathcal{S} accesses reality \mathcal{R} only through the representable functor $\text{Hom}_{\mathbf{Meas}}((\mathcal{S}, \mathcal{R}|_{\mathcal{S}}), -)$. We argue that the information paradox—the apparent conflict between unitarity and the thermal nature of Hawking radiation—acquires a natural structural resolution when formulated within the measurement category framework. The key insight is that the infalling and asymptotic observers occupy fundamentally different objects in the measurement category $\mathbf{Meas}_{\text{BH}}$, giving rise to distinct representable functors that encode complementary but non-simultaneously-realizable relational knowledge. We develop a categorical formulation of black hole complementarity in which the firewall paradox dissolves as a consequence of the Yoneda Constraint: the assumption that both observers share a single global presheaf violates the structure of embedded observation. We construct the black hole measurement category, characterize the Kan extension deficit for observers with access only to Hawking radiation, and establish connections to the Page curve via the time-dependent epistemic horizon. We show that the island formula for entanglement entropy admits a natural interpretation as a transition in the structure of representable functors, corresponding to the moment when the observer’s Kan extension begins to recover information about the black hole interior. Accompanying Haskell code provides computational models of the key categorical structures. The framework unifies insights from black hole complementarity, the holographic principle, quantum error correction, and the recent island/replica developments into a single coherent perspective grounded in the Yoneda lemma.

Keywords: black hole information paradox, Yoneda lemma, category theory, Hawking radiation, black hole complementarity, firewall paradox, Page curve, island formula, epistemic horizons, representable functors, Kan extensions, holographic principle

Contents

1	Introduction	3
1.1	Summary of Contributions	3
1.2	Relation to Prior Work	4
1.3	Plan of the Paper	4
2	Background	4
2.1	The Black Hole Information Paradox	4
2.2	Black Hole Complementarity	5
2.3	The Island Formula and the Page Curve	5
2.4	The Yoneda Constraint on Observer Knowledge	5
3	The Black Hole Measurement Category	6
3.1	Construction	6
3.2	Distinguished Objects	7
3.3	The Horizon as Morphism Obstruction	7
3.4	The Inclusion Structure	8
4	The Yoneda Constraint in Black Hole Spacetimes	8
4.1	Representable Functors of Black Hole Observers	8
4.2	The Information Paradox as Presheaf Incompatibility	9
4.3	The Epistemic Horizon of the Asymptotic Observer	9
5	Categorical Black Hole Complementarity	10
5.1	Complementarity as Presheaf Distinctness	10
5.2	Resolution of the Firewall Paradox	11
5.3	The Complementarity Presheaf	12
6	Kan Extensions and Information Recovery	12
6.1	The Information Recovery Problem as Extension Problem	12
6.2	The Bracket of Extrapolation for Black Holes	13
6.3	Entanglement Entropy from the Extension Deficit	14
7	The Page Curve as Presheaf Transition	14
7.1	Time-Dependent Epistemic Horizons	14
7.2	The Page Time as a Phase Transition in Presheaf Structure	15
7.3	The Island Formula in the Yoneda Framework	16
7.4	The Page Curve from Yoneda Data	17
8	Connections to Holographic Entanglement	17
8.1	The Holographic Measurement Category	17
8.2	The RT Formula as Deficit Measure	17
8.3	Quantum Error Correction and Representable Functors	18
9	Haskell Implementation	18

10 Discussion and Open Questions	19
10.1 Relation to Existing Approaches	19
10.2 Implications for Quantum Gravity	20
10.3 Open Questions	20
11 Conclusion	21
A Detailed Proofs	22
A.1 Proof of theorem 5.2 (Detailed)	22
A.2 Proof of theorem 7.6 (Detailed)	23
B Categorical Definitions	23
C Physical Constants and Conventions	23

1 Introduction

The black hole information paradox, first articulated by Hawking in 1976 [1], remains one of the deepest problems in theoretical physics. At its core lies a conflict between two foundational principles: the unitarity of quantum mechanics, which demands that information is preserved in any physical process, and the semiclassical analysis of black hole evaporation, which produces thermal Hawking radiation that appears to destroy information about the matter that formed the black hole.

Nearly five decades of work on this problem have produced remarkable insights—from black hole complementarity [3] and the holographic principle [5, 6] to the AdS/CFT correspondence [7], the firewall paradox [8], and the recent island formula developments [9, 10]. Yet the problem persists, in part because the various proposed resolutions operate within different conceptual frameworks that are difficult to compare directly.

In this paper, we propose that the Yoneda Constraint on Observer Knowledge—a category-theoretic principle developed in previous work by the YonedaAI Collaboration [14–17]—provides a unifying structural perspective on the information paradox. The Yoneda Constraint states that an embedded observer \mathcal{S} accesses reality \mathcal{R} only through the representable functor $y^{(\mathcal{S}, \sigma_{\mathcal{S}})} = \text{Hom}_{\mathbf{Meas}}((\mathcal{S}, \sigma_{\mathcal{S}}), -)$, which determines the observer’s epistemic position up to isomorphism but cannot determine \mathcal{R} itself when $\mathcal{S} \subsetneq \mathcal{R}$.

Applied to the black hole setting, the Yoneda Constraint reveals that the information paradox is, at root, a consequence of the incompatibility of different observers’ representable functors. The asymptotic observer and the infalling observer occupy different objects in the measurement category, giving rise to different representable functors that encode genuinely different—and non-simultaneously-realizable—relational knowledge about the black hole spacetime. The paradox arises when one illegitimately attempts to combine these distinct perspectival descriptions into a single global account.

1.1 Summary of Contributions

Our main contributions are:

- (1) **Black hole measurement category** (section 3): We construct the measurement category $\mathbf{Meas}_{\text{BH}}$ appropriate to the black hole setting, with objects encoding observer positions relative to the horizon and morphisms encoding measurement-preserving maps between observers.
- (2) **Yoneda analysis of complementarity** (section 5): We formulate black hole complementarity as the statement that the representable functors of the infalling and asymptotic observers are distinct and non-simultaneously-embeddable in a single consistent presheaf, resolving the firewall paradox.
- (3) **Kan extension deficit and information recovery** (section 6): We characterize the information recovery problem as a Kan extension problem and show that the extension deficit measures the information about the black hole interior that is inaccessible to the asymptotic observer.
- (4) **Page curve as presheaf transition** (section 7): We interpret the Page curve and the island formula as a transition in the structure of the observer’s representable functor, corresponding to the moment when the Kan extension begins to recover interior information through entanglement with early radiation.

- (5) **Haskell implementation** (section 9): We provide accompanying Haskell code that computationally models the categorical structures, including the measurement category, representable functors, and the Kan extension deficit.

1.2 Relation to Prior Work

Our framework builds on and connects several strands of existing research:

- The Yoneda Constraint on Observer Knowledge [14, 15], which provides the foundational categorical framework.
- The Embedded Observer Constraint [16, 17], which establishes information-theoretic bounds on embedded observation.
- The Measurement Boundary Problem [18], which analyzes structural limitations on observational access in emergent spacetime.
- Black hole complementarity [3, 4], which posits that infalling and asymptotic observers see complementary physics.
- The sheaf-theoretic approach to contextuality [19], which provides the presheaf-theoretic language for incompatible descriptions.
- Quantum error correction in holography [24, 25], which connects bulk reconstruction to error-correcting codes.
- The island formula [9–11], which resolves the Page curve through gravitational path integrals.

1.3 Plan of the Paper

Section 2 reviews the necessary background on the information paradox and the Yoneda Constraint. Section 3 constructs the black hole measurement category. Section 4 applies the Yoneda Constraint to the black hole setting. Section 5 develops the categorical formulation of complementarity and resolves the firewall paradox. Section 6 treats the Kan extension deficit and information recovery. Section 7 interprets the Page curve and island formula. Section 8 connects to holographic entanglement entropy. Section 9 describes the accompanying Haskell code. Section 10 discusses open questions. Section 11 concludes.

2 Background

2.1 The Black Hole Information Paradox

A black hole formed from the gravitational collapse of matter in a pure quantum state $|\psi\rangle$ radiates thermally via the Hawking process [2]. The radiation has a thermal spectrum at the Hawking temperature

$$T_H = \frac{\hbar c^3}{8\pi GMk_B},$$

where M is the black hole mass. If the evaporation proceeds to completion, the final state is thermal radiation—a mixed state—implying that the evolution from $|\psi\rangle$ to the final radiation is non-unitary.

This conflicts with quantum mechanics, which requires that time evolution be unitary. The conflict is sharpened by Page’s argument [12]: if the evolution is unitary, the entanglement entropy of the radiation must follow the *Page curve*—initially increasing as the black hole evaporates, then decreasing after the *Page time* (when roughly half the entropy has been radiated) until the radiation is in a pure state. Hawking’s semiclassical calculation, however, predicts monotonically increasing entanglement entropy, violating the Page curve.

2.2 Black Hole Complementarity

Susskind, Thorlacius, and Uglum [3] proposed *black hole complementarity*: the infalling observer sees smooth passage through the horizon, while the asymptotic observer sees information reflected back in the Hawking radiation. These descriptions are complementary—no single observer can verify both simultaneously, since doing so would require superluminal signaling.

Complementarity resolves the paradox by denying the existence of a global description that simultaneously accounts for both observers’ experiences. However, the AMPS argument [8] challenged complementarity by showing that it, combined with the assumption of smooth horizon physics, leads to a contradiction involving the monogamy of entanglement. AMPS argued that either the horizon is replaced by a “firewall” or one of the assumptions (unitarity, effective field theory below the Planck scale, no drama at the horizon) must fail.

2.3 The Island Formula and the Page Curve

Recent developments [9–11] have shown that gravitational path integrals, when properly accounting for replica wormhole contributions, reproduce the Page curve. The *island formula* for the entanglement entropy of radiation is:

$$S(\text{Rad}) = \min \left\{ \text{ext}_I \left[\frac{\text{Area}(\partial I)}{4G_N} + S_{\text{bulk}}(\text{Rad} \cup I) \right] \right\},$$

where the extremization is over “islands” I —regions of the black hole interior that are included in the entanglement wedge of the radiation. Before the Page time, the minimal surface has no island contribution; after the Page time, an island forms inside the black hole, reducing the entropy and reproducing the Page curve.

2.4 The Yoneda Constraint on Observer Knowledge

The Yoneda Constraint [14, 15] is a principle of embedded observation derived from the Yoneda lemma in category theory. We recall the essential elements.

Definition 2.1 (Measurement Category, [15]). The measurement category **Meas** has objects $(\mathcal{S}, \sigma_{\mathcal{S}})$ where \mathcal{S} is a subsystem and $\sigma_{\mathcal{S}}$ is the restricted state, and morphisms are state-preserving channels (CPTP maps in the quantum case).

Proposition 2.2 (Yoneda Constraint, [15]). *The embedded observer \mathcal{S} accesses \mathcal{R} only through the representable functor $y^{(\mathcal{S}, \sigma_{\mathcal{S}})} = \text{Hom}_{\mathbf{Meas}}((\mathcal{S}, \sigma_{\mathcal{S}}), -)$. This determines $(\mathcal{S}, \sigma_{\mathcal{S}})$ up to isomorphism but does not determine \mathcal{R} when $\mathcal{S} \subsetneq \mathcal{R}$.*

The Kan extension deficit $\Delta(\mathcal{S})$ quantifies the gap between local and global descriptions: it vanishes if and only if $\mathcal{S} = \mathcal{R}$.

3 The Black Hole Measurement Category

3.1 Construction

We now construct the measurement category appropriate to the black hole setting. The key feature distinguishing this from the generic measurement category is the presence of the event horizon, which creates a non-trivial causal structure that constrains which measurements are available to which observers.

Definition 3.1 (Black Hole Measurement Category). Let (\mathcal{M}, g) be a black hole spacetime with event horizon \mathcal{H}^+ . The *black hole measurement category* $\mathbf{Meas}_{\text{BH}}$ is the category whose:

- (i) **Objects** are triples $(\mathcal{S}, \Sigma_{\mathcal{S}}, \rho_{\mathcal{S}})$ where:
 - \mathcal{S} is an observer subsystem localized in a spacetime region,
 - $\Sigma_{\mathcal{S}}$ is the causal domain of \mathcal{S} —the set of spacetime events causally accessible to \mathcal{S} ,
 - $\rho_{\mathcal{S}}$ is the quantum state restricted to $\Sigma_{\mathcal{S}}$.
- (ii) **Morphisms** $f : (\mathcal{S}_1, \Sigma_1, \rho_1) \rightarrow (\mathcal{S}_2, \Sigma_2, \rho_2)$ are *information-preserving* CPTP maps $f : \mathcal{B}(\mathcal{H}_{\Sigma_1}) \rightarrow \mathcal{B}(\mathcal{H}_{\Sigma_2})$ satisfying:
 - $f(\rho_1) = \rho_2$ (state preservation),
 - f is compatible with the causal structure: if $\Sigma_1 \not\subseteq J^-(\Sigma_2)$ (the causal past of Σ_2), then f must factor through the shared causal domain,
 - f is *non-trivial*: f is not a state-preparation map $\rho \mapsto \text{Tr}(\rho)\sigma$ for a fixed state σ . That is, we exclude channels that discard the input and prepare a fixed output, as these carry no information about the source observer's state. Formally, we require that f has non-zero Holevo capacity $\chi(f) > 0$.

This restriction is physically motivated: we are interested in morphisms that represent genuine information transfer between observers, not trivial state preparations. Without this condition, one could always define a morphism from any observer to any other by tracing out the input, which would trivialize the horizon obstruction. The non-triviality condition ensures that $\text{Hom}(\mathcal{S}_{\infty}, \mathcal{S}_{\text{int}}) = \emptyset$ reflects the genuine impossibility of information transfer across the horizon, not merely a selection rule.

Remark 3.2. We do not allow post-selection in our morphisms: all maps are CPTP (trace-preserving), not merely CP (trace-non-increasing). Post-selected channels would introduce additional subtleties relevant to final-state projection models [23] but are outside the scope of the present analysis.

- (iii) **Composition** is sequential composition of CPTP maps, subject to causality.
- (iv) **Identity** is the identity channel.

3.2 Distinguished Objects

The black hole spacetime gives rise to several distinguished objects in $\mathbf{Meas}_{\text{BH}}$.

Definition 3.3 (Asymptotic Observer). The *asymptotic observer* is the object

$$\mathcal{S}_\infty = (\mathcal{S}_\infty, \Sigma_\infty, \rho_\infty)$$

where Σ_∞ is the causal domain of future null infinity \mathcal{I}^+ and ρ_∞ is the state restricted to Σ_∞ . The asymptotic observer has access to the Hawking radiation but not to the black hole interior.

Definition 3.4 (Infalling Observer). The *infalling observer* is the object

$$\mathcal{S}_{\text{in}} = (\mathcal{S}_{\text{in}}, \Sigma_{\text{in}}, \rho_{\text{in}})$$

where Σ_{in} is the causal domain of the infalling worldline (crossing the horizon into the interior) and ρ_{in} is the state restricted to Σ_{in} .

Definition 3.5 (Early Radiation Observer). The *early radiation observer* at retarded time u is the object

$$\mathcal{S}_{\text{Rad}}(u) = (\mathcal{S}_{\text{Rad}}, \Sigma_{\text{Rad}}(u), \rho_{\text{Rad}}(u))$$

where $\Sigma_{\text{Rad}}(u)$ is the causal domain of the radiation collected up to time u , and $\rho_{\text{Rad}}(u) = \text{Tr}_{\text{BH} \cup \text{late rad}}(\rho)$ is the reduced state of the early radiation.

3.3 The Horizon as Morphism Obstruction

The event horizon \mathcal{H}^+ creates a fundamental obstruction in $\mathbf{Meas}_{\text{BH}}$.

Proposition 3.6 (Horizon Obstruction). *There is no morphism in $\mathbf{Meas}_{\text{BH}}$ from the asymptotic observer to any object localized in the black hole interior:*

$$\text{Hom}_{\mathbf{Meas}_{\text{BH}}}(\mathcal{S}_\infty, \mathcal{S}_{\text{int}}) = \emptyset$$

for any interior object \mathcal{S}_{int} with $\Sigma_{\text{int}} \subset \text{int}(\text{BH})$.

Proof. A morphism from \mathcal{S}_∞ to \mathcal{S}_{int} would require a state-preserving channel from the exterior to the interior. By the causal structure of the black hole spacetime, $\Sigma_\infty \not\subseteq J^-(\Sigma_{\text{int}})$ in general (the asymptotic region is not in the causal past of the interior at late times). Moreover, the no-signaling principle forbids the transmission of information from outside to the interior of a classical black hole (after the formation). Hence no such morphism exists. \square

Corollary 3.7 (Causal Asymmetry). *There exist morphisms from infalling objects to interior objects:*

$$\text{Hom}_{\mathbf{Meas}_{\text{BH}}}(\mathcal{S}_{\text{in}}, \mathcal{S}_{\text{int}}) \neq \emptyset,$$

while $\text{Hom}_{\mathbf{Meas}_{\text{BH}}}(\mathcal{S}_\infty, \mathcal{S}_{\text{int}}) = \emptyset$. The measurement category inherits the causal asymmetry of the spacetime.

3.4 The Inclusion Structure

The measurement category has a natural filtration by retarded time.

Definition 3.8 (Time-Dependent Subcategory). For each retarded time u , define the subcategory $\mathbf{Meas}_{\text{BH}}(u) \subset \mathbf{Meas}_{\text{BH}}$ consisting of objects whose causal domains are contained within the past of the constant- u slice. The inclusion functors

$$J_u : \mathbf{Meas}_{\text{BH}}(u) \hookrightarrow \mathbf{Meas}_{\text{BH}}$$

form a filtration: $\mathbf{Meas}_{\text{BH}}(u_1) \subseteq \mathbf{Meas}_{\text{BH}}(u_2)$ for $u_1 \leq u_2$.

This filtration is crucial for the time-dependent analysis of information recovery in section 7.

Remark 3.9 (Backreaction and dynamical categories). The definition of $\mathbf{Meas}_{\text{BH}}$ uses a fixed background spacetime (\mathcal{M}, g) , but black hole evaporation modifies the spacetime through backreaction. Strictly, the causal structure—and hence the morphism constraints—evolve as the black hole loses mass. We address this through the time-dependent subcategory filtration $\{\mathbf{Meas}_{\text{BH}}(u)\}_{u \geq 0}$: at each retarded time u , the subcategory $\mathbf{Meas}_{\text{BH}}(u)$ is defined with respect to the instantaneous causal structure of the semiclassical spacetime at time u .

More precisely, one should view the collection $\{\mathbf{Meas}_{\text{BH}}(u)\}$ as a *functor* from the poset $(\mathbb{R}_{\geq 0}, \leq)$ to **Cat**, i.e., a filtered diagram of categories. The inclusions $J_{u_1 u_2} : \mathbf{Meas}_{\text{BH}}(u_1) \hookrightarrow \mathbf{Meas}_{\text{BH}}(u_2)$ for $u_1 \leq u_2$ encode the fact that later observers have access to all data available to earlier observers. The “total” category $\mathbf{Meas}_{\text{BH}}$ is the colimit of this diagram. This 2-categorical perspective naturally accommodates the teleological nature of the event horizon: the true event horizon is only defined with respect to the full colimit, while each $\mathbf{Meas}_{\text{BH}}(u)$ uses an approximate “apparent horizon” that evolves with backreaction. A full treatment of the dynamical measurement category, including gravitational backreaction as a 2-morphism, is left to future work (see section 10, Open Question 2).

4 The Yoneda Constraint in Black Hole Spacetimes

4.1 Representable Functors of Black Hole Observers

We now apply the Yoneda Constraint to the distinguished observers in $\mathbf{Meas}_{\text{BH}}$.

Proposition 4.1 (Asymptotic Yoneda Functor). *The representable functor of the asymptotic observer,*

$$y^{\mathcal{S}_\infty} = \text{Hom}_{\mathbf{Meas}_{\text{BH}}}(\mathcal{S}_\infty, -),$$

encodes all measurements performable by an observer at future null infinity. By theorem 3.6, $y^{\mathcal{S}_\infty}$ vanishes on all interior objects:

$$y^{\mathcal{S}_\infty}(\mathcal{S}_{\text{int}}) = \text{Hom}(\mathcal{S}_\infty, \mathcal{S}_{\text{int}}) = \emptyset.$$

The asymptotic observer’s relational knowledge of the interior is structurally zero.

Proposition 4.2 (Infalling Yoneda Functor). *The representable functor of the infalling observer,*

$$y^{\mathcal{S}_{\text{in}}} = \text{Hom}_{\mathbf{Meas}_{\text{BH}}}(\mathcal{S}_{\text{in}}, -),$$

is non-trivial on both exterior and interior objects. The infalling observer has relational access to the horizon-crossing region and sees smooth geometry (no firewall), in accordance with the equivalence principle.

Proposition 4.3 (Non-Isomorphism of Observer Presheaves). *The representable functors $y^{\mathcal{S}_{\infty}}$ and $y^{\mathcal{S}_{\text{in}}}$ are non-isomorphic as functors $\mathbf{Meas}_{\text{BH}} \rightarrow \mathbf{Set}$:*

$$y^{\mathcal{S}_{\infty}} \not\cong y^{\mathcal{S}_{\text{in}}}.$$

Proof. By theorem 4.1, $y^{\mathcal{S}_{\infty}}(\mathcal{S}_{\text{int}}) = \emptyset$ for interior objects, while by theorem 4.2, $y^{\mathcal{S}_{\text{in}}}(\mathcal{S}_{\text{int}}) \neq \emptyset$. Since the functors differ on at least one object, they are non-isomorphic. \square

4.2 The Information Paradox as Presheaf Incompatibility

The information paradox can now be reformulated in presheaf-theoretic language.

Proposition 4.4 (Paradox as Presheaf Conflict). *The black hole information paradox arises from the assumption that there exists a single global presheaf $F : \mathbf{Meas}_{\text{BH}}^{\text{op}} \rightarrow \mathbf{Set}$ that simultaneously:*

- (i) *restricts to the asymptotic observer's representable functor on exterior objects (encoding Hawking radiation as thermal),*
- (ii) *restricts to the infalling observer's representable functor on the horizon region (encoding smooth passage),*
- (iii) *satisfies a global consistency (unitarity) condition.*

The Yoneda Constraint implies that conditions (i) and (ii) correspond to distinct objects in $\mathbf{Meas}_{\text{BH}}$ with non-isomorphic representable functors (theorem 4.3). Condition (iii) imposes a global section requirement. The paradox is the presheaf-theoretic statement that no global section exists that satisfies all three conditions simultaneously.

This reformulation connects the information paradox to the presheaf-theoretic framework of contextuality developed by Abramsky and Brandenburger [19]: the information paradox is a *contextuality obstruction* in the black hole measurement category. Just as the Kochen–Specker theorem shows that quantum mechanics admits no non-contextual hidden variable model (no global section of the valuation presheaf), the information paradox shows that black hole physics admits no global description that simultaneously respects all observers' perspectives.

4.3 The Epistemic Horizon of the Asymptotic Observer

Definition 4.5 (Black Hole Epistemic Horizon). The *epistemic horizon* of the asymptotic observer is the full subcategory

$$\mathbf{Meas}_{\text{BH}}|_{\mathcal{S}_{\infty}} \subset \mathbf{Meas}_{\text{BH}}$$

consisting of objects reachable by morphisms from \mathcal{S}_{∞} . The *epistemic boundary* is

$$\partial_{\mathcal{S}_{\infty}} = \{X \in \mathbf{Meas}_{\text{BH}} : \text{Hom}(\mathcal{S}_{\infty}, X) \neq \emptyset, X \not\subseteq \Sigma_{\infty}\}.$$

Proposition 4.6 (Epistemic Horizon Coincides with Event Horizon). *For the asymptotic observer of a classical (non-evaporating) black hole, the epistemic boundary $\partial_{\mathcal{S}_\infty}$ coincides with the event horizon \mathcal{H}^+ : the categorical epistemic horizon is exactly the geometric event horizon.*

Proof. The causal domain Σ_∞ is the causal past of \mathcal{I}^+ , which is exactly the exterior of the black hole. Objects on the horizon boundary are those with causal domains intersecting \mathcal{H}^+ . Objects strictly interior to the black hole are unreachable from \mathcal{S}_∞ . Hence $\partial_{\mathcal{S}_\infty} = \mathcal{H}^+$. \square

Remark 4.7. For an evaporating black hole, the epistemic horizon is time-dependent and does not precisely coincide with the event horizon (which is a global concept). This time-dependence is crucial for the Page curve analysis in section 7.

5 Categorical Black Hole Complementarity

5.1 Complementarity as Presheaf Distinctness

Black hole complementarity [3] asserts that the infalling and asymptotic observers see complementary physics that cannot be simultaneously verified. The Yoneda framework provides a precise mathematical formulation.

Definition 5.1 (Categorical Complementarity). Two observers $\mathcal{S}_1, \mathcal{S}_2 \in \mathbf{Meas}_{\text{BH}}$ are *categorically complementary* if:

- (i) Their representable functors are non-isomorphic: $y^{\mathcal{S}_1} \not\cong y^{\mathcal{S}_2}$.
- (ii) There is no third object $\mathcal{S}_3 \in \mathbf{Meas}_{\text{BH}}$ whose representable functor restricts to both $y^{\mathcal{S}_1}$ and $y^{\mathcal{S}_2}$ on their respective domains.
- (iii) The non-existence of \mathcal{S}_3 is enforced by the causal structure (no superluminal signaling).

Theorem 5.2 (Black Hole Complementarity from the Yoneda Constraint). *The asymptotic observer \mathcal{S}_∞ and the infalling observer \mathcal{S}_{in} are categorically complementary in $\mathbf{Meas}_{\text{BH}}$.*

Proof. Condition (i) follows from theorem 4.3. For condition (ii), suppose \mathcal{S}_3 existed with $y^{\mathcal{S}_3}$ restricting to both $y^{\mathcal{S}_\infty}$ (on exterior objects) and $y^{\mathcal{S}_{\text{in}}}$ (on interior objects). Then \mathcal{S}_3 would have morphisms to both exterior and interior objects, meaning its causal domain Σ_3 would overlap both the exterior and the deep interior. But by the Yoneda embedding (full and faithful), the relational data of $y^{\mathcal{S}_3}$ would determine \mathcal{S}_3 up to isomorphism, and a single observer with simultaneous access to both the asymptotic radiation and the interior would require a causal domain violating the black hole causal structure. For condition (iii), verifying both descriptions simultaneously would require an observer collecting Hawking radiation at \mathcal{I}^+ and also crossing the horizon—but the time required to collect sufficient radiation (of order the evaporation time) prevents subsequent crossing of the horizon before the black hole disappears. \square

5.2 Resolution of the Firewall Paradox

The AMPS argument [8] proceeds as follows. Consider a black hole past the Page time, and let B denote a late Hawking mode, A the early radiation, and C the mode's interior partner. Unitarity requires B to be maximally entangled with A (for the Page curve to decrease); the smoothness of the horizon requires B to be maximally entangled with C . By monogamy of entanglement, B cannot be maximally entangled with both A and C . The conclusion is that the horizon is replaced by a “firewall”—a region of high-energy quanta at the horizon.

Theorem 5.3 (Yoneda Resolution of the Firewall Paradox). *The AMPS argument fails within the Yoneda Constraint framework because it illegitimately assumes a single global presheaf that simultaneously encodes the B - A entanglement (visible to \mathcal{S}_∞) and the B - C entanglement (visible to \mathcal{S}_{in}).*

Proof. The B - A entanglement is encoded in the representable functor $y^{\mathcal{S}_\infty}$: it manifests as a non-factorizability of the presheaf over the bipartition (early radiation, late radiation). The B - C entanglement is encoded in $y^{\mathcal{S}_{\text{in}}}$: it manifests as a non-factorizability over the bipartition (outgoing mode, interior mode).

The AMPS argument implicitly assumes a single object $\mathcal{S}_{\text{AMPS}}$ whose representable functor encodes both entanglements. By theorem 5.2, no such object exists in $\mathbf{Meas}_{\text{BH}}$. The monogamy-of-entanglement argument applies only within a single representable functor (a single observer's description), not across distinct representable functors corresponding to categorically complementary observers.

Therefore, the “firewall” conclusion does not follow. The infalling observer's functor $y^{\mathcal{S}_{\text{in}}}$ encodes smooth horizon physics (B - C entanglement), and the asymptotic observer's functor $y^{\mathcal{S}_\infty}$ encodes unitary evaporation (B - A entanglement after the Page time). These are consistent because they are descriptions from different objects in $\mathbf{Meas}_{\text{BH}}$, not a single contradictory description. \square

Remark 5.4 (Relation to Relational Quantum Mechanics). Our claim that “monogamy of entanglement applies only within a single representable functor” is a precise categorical formulation of a relational view of entanglement. In this framework, entanglement is not a global property of a state $|\Psi\rangle$ in some observer-independent Hilbert space; rather, it is a property of the relational data encoded in a specific representable functor $y^{\mathcal{S}}$. This is structurally aligned with Rovelli's relational quantum mechanics (RQM) [29], which holds that quantum states are observer-relative. The Yoneda framework provides the mathematical backbone: the “observer-relativity” of RQM is the statement that different objects in $\mathbf{Meas}_{\text{BH}}$ have non-isomorphic representable functors, and there is no categorical reason to prefer one over another.

The key distinction from AMPS is that their argument assumes the simultaneous applicability of B - A entanglement (an asymptotic observer datum) and B - C entanglement (an infalling observer datum) *within a single description*. The Yoneda framework shows that these are data from different representable functors and that combining them requires the complementarity presheaf F_{comp} , which is non-representable (theorem 5.7). The monogamy argument is valid within $y^{\mathcal{S}_\infty}$ (where only B - A is visible) and within $y^{\mathcal{S}_{\text{in}}}$ (where only B - C is visible), but not across them.

Remark 5.5. This resolution is structurally similar to Susskind's original complementarity argument [3], but the Yoneda framework makes the mathematical structure precise: it

is not merely that “no single observer can verify both descriptions” (an operational statement), but that the measurement category *does not contain an object* whose representable functor encodes both (a structural statement). The infalling observer’s experience of a smooth horizon is encoded in $y^{S_{in}}$ and is not contradicted by the asymptotic observer’s unitary Page curve, because these are evaluations of different functors on different objects.

5.3 The Complementarity Presheaf

Though no single representable functor captures both observers, we can define a *complementarity presheaf* that encodes the combined data.

Definition 5.6 (Complementarity Presheaf). The *complementarity presheaf* is the coproduct

$$F_{\text{comp}} = y^{S_\infty} \sqcup_{y^{S_\infty} \cap y^{S_{in}}} y^{S_{in}}$$

in $\text{PSh}(\mathbf{Meas}_{\text{BH}})$, the pushout of the two representable functors along their restriction to the shared domain (the exterior region accessible to both).

Proposition 5.7 (Non-Representability of Complementarity). *The complementarity presheaf F_{comp} is not representable: there is no object $X \in \mathbf{Meas}_{\text{BH}}$ with $y^X \cong F_{\text{comp}}$.*

Proof. If F_{comp} were representable by X , then by the Yoneda embedding, X would have morphisms to all objects reachable by either S_∞ or S_{in} , contradicting the causal structure argument of theorem 5.2. \square

The non-representability of the complementarity presheaf is the categorical signature of the information paradox. It states that no single observer position in the measurement category captures the “complete” black hole physics—this completeness is a property of the presheaf category $\text{PSh}(\mathbf{Meas}_{\text{BH}})$ (which is “bigger” than $\mathbf{Meas}_{\text{BH}}$ itself), not of any individual object.

6 Kan Extensions and Information Recovery

6.1 The Information Recovery Problem as Extension Problem

The central question of the information paradox—whether information is preserved in black hole evaporation—can be formulated as a Kan extension problem in the Yoneda framework.

Definition 6.1 (Information Recovery as Kan Extension). Let $J_\infty : \mathbf{Meas}_{\text{BH}}|_{S_\infty} \rightarrow \mathbf{Meas}_{\text{BH}}$ be the inclusion of the asymptotic observer’s accessible subcategory, and let \mathfrak{D}_∞ be the description functor restricted to $\mathbf{Meas}_{\text{BH}}|_{S_\infty}$. The *information recovery problem* asks whether the left Kan extension

$$\text{Lan}_{J_\infty}(\mathfrak{D}_\infty \circ J_\infty)$$

recovers the total description functor \mathfrak{R} on all of $\mathbf{Meas}_{\text{BH}}$, including interior objects.

Definition 6.2 (Extension Deficit). Since our target category is **Set** (not an abelian category), the “cokernel” of the comparison natural transformation $\eta : \text{Lan}_J(\mathfrak{D} \circ J) \Rightarrow \mathfrak{R}$

is defined information-theoretically. For each object $X \in \mathbf{Meas}_{\text{BH}}$, the *pointwise deficit* is

$$\delta(X) = H(\mathfrak{R}(X) \mid \text{Lan}_J(\mathfrak{D} \circ J)(X)),$$

the conditional Shannon (or von Neumann) entropy of the true description $\mathfrak{R}(X)$ given the Kan extension's best approximation. The *total extension deficit* is

$$\Delta(\mathcal{S}) = \sum_{X \in \text{Ob}(\mathbf{Meas}_{\text{BH}})} \delta(X).$$

When working in the quantum setting with density operators, we replace the Shannon entropy with the von Neumann entropy and the conditional entropy with $S(\rho_{XR} \parallel \rho_{\text{Lan}(X)} \otimes \rho_R)$, the relative entropy measuring the information gap.

Proposition 6.3 (Information Deficit of the Asymptotic Observer). *For a black hole that has not yet begun to evaporate (or is in the early stages of evaporation before the Page time), the extension deficit*

$$\Delta(\mathcal{S}_\infty) = \sum_{X \in \text{Ob}(\mathbf{Meas}_{\text{BH}})} \delta(X)$$

is non-trivial: the asymptotic observer cannot recover the interior state. The deficit is bounded below by the entanglement entropy:

$$\Delta(\mathcal{S}_\infty) \neq 0 \iff S(\rho_{\text{Rad}}) > 0.$$

Proof. Before the Page time, the Hawking radiation is in a highly entangled state with the black hole interior. The asymptotic observer's accessible subcategory contains only exterior objects. The Kan extension, being the optimal extrapolation from exterior data, cannot recover information encoded in the entanglement between the radiation and the interior. By theorem 3.6, there are no morphisms from exterior to interior objects, so the over-category $(J_\infty \downarrow X)$ is empty for interior objects X , and the pointwise Kan extension gives the initial object (no information): $\delta(X) = S(\mathfrak{R}(X))$ for all interior X .

For the bound, note that $\Delta(\mathcal{S}_\infty) = 0$ requires $\delta(X) = 0$ for all X , which requires the Kan extension to recover $\mathfrak{R}(X)$ for all objects—including interior objects. But the reduced state of the radiation ρ_{Rad} is mixed (has nonzero von Neumann entropy) precisely when entangled with the interior. The Kan extension from the radiation's reduced state cannot distinguish different purifications, so the deficit on interior objects is at least $S_{\text{vN}}(\rho_{\text{Rad}}) > 0$. \square

6.2 The Bracket of Extrapolation for Black Holes

Proposition 6.4 (Black Hole Extrapolation Bracket). *The left and right Kan extensions provide bounds on the true description:*

$$\text{Lan}_{J_\infty}(\mathfrak{D}_\infty \circ J_\infty) \Rightarrow \mathfrak{R} \Rightarrow \text{Ran}_{J_\infty}(\mathfrak{D}_\infty \circ J_\infty).$$

For the asymptotic observer of a classical black hole, the bracket is maximally wide on interior objects: the left Kan extension gives the empty set (no information), while the right Kan extension gives the set of all states compatible with the exterior data (maximal ambiguity).

The width of the bracket on interior objects corresponds to the *non-uniqueness of purification*: many interior states are compatible with the same reduced state of the radiation. This is the categorical version of the statement that the Hawking radiation, being thermal, carries no information about the specific microstate of the black hole.

6.3 Entanglement Entropy from the Extension Deficit

Proposition 6.5 (Entanglement Entropy as Deficit Measure). *In the quantum setting, the extension deficit for the asymptotic observer at time u satisfies*

$$\Delta(\mathcal{S}_\infty, u) \neq 0 \iff S_{\text{vN}}(\rho_{\text{Rad}}(u)) > 0.$$

Moreover, the total deficit is bounded:

$$S_{\text{vN}}(\rho_{\text{Rad}}(u)) \leq \Delta(\mathcal{S}_\infty, u) \leq \log \dim \mathcal{H}_{\text{int}} \cdot |\{X : X \text{ interior}\}|.$$

The von Neumann entropy of the radiation quantifies the “size” of the deficit: when $S_{\text{vN}}(\rho_{\text{Rad}}(u)) = 0$, the radiation is pure and the deficit vanishes (full information recovery).

Proof. (\Rightarrow): If $\Delta(\mathcal{S}_\infty, u) \neq 0$, then there exists at least one object X with $\delta(X) > 0$. For interior objects, $\delta(X) > 0$ implies the Kan extension does not recover $\mathfrak{R}(X)$, which requires entanglement between the radiation and the interior (otherwise the state factorizes and the reduced state determines the interior state). Hence $S_{\text{vN}}(\rho_{\text{Rad}}(u)) > 0$.

(\Leftarrow): If $S_{\text{vN}}(\rho_{\text{Rad}}(u)) > 0$, the radiation is entangled with the interior. The reduced state $\rho_{\text{Rad}}(u) = \text{Tr}_{\text{int}}(|\Psi\rangle\langle\Psi|)$ is mixed. By the Holevo bound, the mutual information accessible through measurements on the radiation alone is bounded by $\chi \leq S_{\text{vN}}(\rho_{\text{Rad}})$. The Kan extension from the radiation’s subcategory is limited by this Holevo information, so the deficit on interior objects is at least $S_{\text{vN}}(\rho_{\text{Rad}}(u))$, giving the lower bound.

The upper bound follows because the maximum entropy of any interior object is $\log \dim \mathcal{H}_{\text{int}}$, and the deficit on exterior objects vanishes (the Kan extension recovers these exactly). This establishes the connection between the categorical deficit and the physical entropy, analogous to the Petz recovery map [20, 21] in operator algebra quantum error correction [22, 25]: the Kan extension plays the role of the Petz map, and the deficit measures the failure of exact recovery. \square

This connects the abstract categorical notion of extension deficit to the concrete physical quantity of entanglement entropy, providing the bridge to the Page curve analysis in the next section.

7 The Page Curve as Presheaf Transition

7.1 Time-Dependent Epistemic Horizons

As the black hole evaporates, the asymptotic observer’s epistemic horizon evolves. We model this through the time-dependent representable functor.

Definition 7.1 (Time-Dependent Representable Functor). The *time-dependent representable functor* of the radiation observer at time u is

$$y^{\mathcal{S}_{\text{Rad}}(u)} = \text{Hom}_{\mathbf{Meas}_{\text{BH}}}(\mathcal{S}_{\text{Rad}}(u), -).$$

As u increases, $\mathcal{S}_{\text{Rad}}(u)$ collects more radiation, and the representable functor “expands” to encode more relational data.

Proposition 7.2 (Monotonicity of Radiation Knowledge). *The representable functors are monotone: for $u_1 \leq u_2$, there is a natural transformation*

$$y^{\mathcal{S}_{\text{Rad}}(u_1)} \Rightarrow y^{\mathcal{S}_{\text{Rad}}(u_2)}$$

reflecting the fact that later radiation observers have at least as much relational knowledge as earlier ones.

7.2 The Page Time as a Phase Transition in Presheaf Structure

The Page time is the retarded time u_P at which the entanglement entropy of the radiation begins to decrease. In the Yoneda framework, this corresponds to a qualitative transition in the structure of the representable functor.

Theorem 7.3 (Page Time as Presheaf Transition). *At the Page time u_P , the representable functor $y^{\mathcal{S}_{\text{Rad}}(u_P)}$ undergoes a structural transition:*

- (i) *For $u < u_P$ (before the Page time), the Kan extension deficit $\Delta(\mathcal{S}_{\text{Rad}}(u))$ is increasing: the radiation observer has increasing relational data but the extrapolation to the interior is worsening because the entanglement is growing.*
- (ii) *For $u > u_P$ (after the Page time), the deficit begins to decrease: the radiation observer’s representable functor begins to “see” interior information indirectly, through the entanglement structure of the collected radiation.*
- (iii) *At $u = u_{\text{evap}}$ (complete evaporation), $\Delta(\mathcal{S}_{\text{Rad}}(u_{\text{evap}})) = 0$ if unitarity holds: the radiation is pure and the Kan extension recovers the full description.*

Proof (sketch). (i) Before the Page time, each emitted Hawking quantum increases the entanglement between the radiation and the remaining black hole. The representable functor of the radiation gains access to more morphisms (more radiation modes), but the extension deficit grows because the over-category $(J_\infty \downarrow X)$ for interior objects remains empty—the radiation modes are individually thermal and carry no accessible correlation with the interior.

(ii) After the Page time, the radiation subsystem becomes larger than the black hole subsystem. By Page’s theorem [13], the reduced state of the radiation begins to deviate from maximal mixedness, and correlations within the radiation begin to encode information about the interior state. In the categorical language, the representable functor of the radiation observer begins to have non-trivial natural transformations to functors that depend on interior data, because the entanglement structure of the radiation acts as an “indirect witness” [18] for the interior state.

(iii) At complete evaporation, the total state is (by assumption of unitarity) a pure state of the radiation. The Kan extension from the radiation’s subcategory recovers the total description because the radiation is the total system. \square

7.3 The Island Formula in the Yoneda Framework

The island formula [9, 10] states that after the Page time, an “island”—a region of the black hole interior—is included in the entanglement wedge of the radiation. We give this a categorical interpretation.

Definition 7.4 (Categorical Island). A *categorical island* at time $u > u_P$ is a collection of interior objects $I \subset \mathbf{Meas}_{\text{BH}}$ such that the extended subcategory

$$\mathbf{Meas}_{\text{BH}}|_{\mathcal{S}_{\text{Rad}}(u) \cup I}$$

has the property that the Kan extension from this extended subcategory recovers more of \mathfrak{R} than the Kan extension from $\mathbf{Meas}_{\text{BH}}|_{\mathcal{S}_{\text{Rad}}(u)}$ alone.

Remark 7.5 (Island as subcategory vs. spacetime region). In the gravitational literature, the “island” I is a spacetime region—a connected component of the interior that enters the entanglement wedge of the radiation. In our categorical framework, I is a *full subcategory* of $\mathbf{Meas}_{\text{BH}}$ consisting of observer objects localized in that spacetime region. The distinction is important: the categorical island I encodes not just the region but the algebra of observables on that region and the morphisms (channels) connecting it to other objects. The island’s contribution to the extended subcategory comes through these morphisms, which provide the additional entries in the over-categories ($J_I \downarrow X$) that improve the Kan extension. This is the categorical version of including the island’s algebra \mathcal{A}_I in the operator algebra available for reconstruction.

Proposition 7.6 (Island Reduces Extension Deficit). *The inclusion of the island I in the observer’s effective subcategory reduces the Kan extension deficit:*

$$\Delta(\mathcal{S}_{\text{Rad}}(u) \cup I) < \Delta(\mathcal{S}_{\text{Rad}}(u))$$

for $u > u_P$. The island formula’s minimization over possible islands corresponds to finding the I that minimizes the deficit:

$$I^*(u) = \arg \min_I \Delta(\mathcal{S}_{\text{Rad}}(u) \cup I).$$

Proof. Adding the island I to the observer’s accessible subcategory introduces new morphisms: morphisms from interior objects in I to other objects in $\mathbf{Meas}_{\text{BH}}$. These additional morphisms provide non-trivial entries in the over-categories ($J \downarrow X$) for interior objects X , improving the pointwise Kan extension and reducing the deficit.

The entropy formula

$$S(\text{Rad}) = \min_I \left[\frac{\text{Area}(\partial I)}{4G_N} + S_{\text{bulk}}(\text{Rad} \cup I) \right]$$

can be interpreted as: the area term penalizes the “cost” of extending the subcategory across the horizon boundary, and the bulk entropy term measures the remaining entanglement after the extension. The minimization balances these two contributions. \square

7.4 The Page Curve from Yoneda Data

Theorem 7.7 (Page Curve from the Yoneda Framework). *In the Yoneda framework, the entanglement entropy of the radiation follows the Page curve:*

$$S(\text{Rad}(u)) = \begin{cases} S_{\text{Hawking}}(u) & u < u_P \quad (\text{no island}), \\ \frac{A(\partial I^*(u))}{4G_N} + S_{\text{bulk}}(\text{Rad} \cup I^*(u)) & u > u_P \quad (\text{with island}). \end{cases}$$

The transition at u_P is the moment when including the island first reduces the deficit below the Hawking value. This is a phase transition in the structure of the representable functor: the functor’s optimal Kan extension “jumps” from one using only exterior data to one incorporating interior data through the island.

This theorem shows that the Page curve, far from being a special feature of holographic models or replica wormholes, is a consequence of the general categorical structure of embedded observation applied to the black hole setting. The Page time is the moment when the observer’s relational data, accumulated through the representable functor, becomes rich enough to indirectly “see” interior information.

8 Connections to Holographic Entanglement

8.1 The Holographic Measurement Category

In the AdS/CFT correspondence, the measurement category acquires additional structure from the holographic duality.

Definition 8.1 (Holographic Measurement Category). The *holographic measurement category* $\mathbf{Meas}_{\text{holo}}$ has:

- (i) Objects $(\mathcal{S}, \Sigma, \rho)$ where Σ is either a boundary subregion or a bulk subregion within an entanglement wedge.
- (ii) Morphisms include both boundary CPTP maps and the bulk-boundary reconstruction maps guaranteed by entanglement wedge reconstruction.

Proposition 8.2 (Entanglement Wedge as Epistemic Horizon). *For a boundary observer with access to subregion A , the epistemic horizon in $\mathbf{Meas}_{\text{holo}}$ coincides with the boundary of the entanglement wedge $\text{EW}(A)$. Objects within $\text{EW}(A)$ are relationally accessible; objects outside are not.*

This connects the Yoneda framework to the quantum error correction structure of holography [24, 25]: the entanglement wedge reconstruction theorem is the statement that the Kan extension from the boundary subregion A recovers the description functor on $\text{EW}(A)$ exactly, with the extension deficit concentrated on the region outside $\text{EW}(A)$.

8.2 The RT Formula as Deficit Measure

Proposition 8.3 (Ryu–Takayanagi from the Kan Extension). *The Ryu–Takayanagi formula [26]*

$$S(A) = \frac{\text{Area}(\gamma_A)}{4G_N}$$

where γ_A is the minimal surface homologous to A , measures the Kan extension deficit of the boundary observer on subregion A : the area of γ_A quantifies the information about the bulk that cannot be recovered from the boundary data on A .

Proof (heuristic). The entanglement entropy $S(A)$ is the von Neumann entropy of the reduced state ρ_A . By theorem 6.5, this is a measure of the Kan extension deficit. The RT formula identifies this entropy with a geometric quantity—the area of the minimal surface—which in the categorical language represents the “size” of the epistemic horizon boundary separating the accessible region ($\text{EW}(A)$) from the inaccessible region. \square

8.3 Quantum Error Correction and Representable Functors

Proposition 8.4 (QEC as Kan Extension). *The quantum error correction structure of holography [24] corresponds to the exactness of the Kan extension on the entanglement wedge: for bulk operators $\phi \in \mathcal{A}_{\text{bulk}}(\text{EW}(A))$, the Kan extension from A recovers ϕ exactly (zero deficit), while for operators outside $\text{EW}(A)$, the deficit is non-zero.*

Remark 8.5 (Kan extension as Petz recovery map). The Kan extension from a boundary subregion to the bulk is the categorical analog of the *Petz recovery map* in operator algebra quantum error correction (OAQEC) [20–22]. In OAQEC, the Petz map $\mathcal{R} : \mathcal{A}_A \rightarrow \mathcal{A}_{\text{bulk}}$ is the optimal recovery channel for reconstructing bulk operators from boundary data on subregion A . The Kan extension $\text{Lan}_J(\mathfrak{D} \circ J)$ plays exactly this role in our categorical framework: it is the universal left-adjoint approximation to the true description functor, optimized over the data available in the observer’s subcategory. The extension deficit Δ corresponds to the *recovery error* $1 - F(\rho, \mathcal{R} \circ \mathcal{N}(\rho))$ of the Petz map, where \mathcal{N} is the noise (erasure) channel. This connection to OAQEC provides physical grounding for the categorical deficit: it is not merely an abstract measure but corresponds to a concrete operational quantity—the fidelity of quantum state recovery.

This provides a unified categorical language for holographic bulk reconstruction: the representable functor of the boundary observer encodes all reconstructible bulk physics, and the Kan extension identifies the optimal reconstruction and its limitations.

9 Haskell Implementation

We provide accompanying Haskell code in the directory `src/black-hole-information-paradox/` that computationally models the key categorical structures developed in this paper. The implementation includes:

- (1) **Measurement category** (`MeasurementCategory.hs`): A type-level representation of the black hole measurement category, including objects (observers), morphisms (CPTP channels), and the causal structure constraints.
- (2) **Representable functors** (`RepresentableFunctor.hs`): Computation of the representable functor for specified observers, including the horizon obstruction that makes the asymptotic observer’s functor vanish on interior objects.
- (3) **Kan extension deficit** (`KanExtension.hs`): Computation of the left Kan extension and the extension deficit, demonstrating the time-dependent behavior described in section 6.
- (4) **Page curve simulation** (`PageCurve.hs`): A simulation of the entanglement entropy as a function of time, showing the Page curve behavior and the island transition, using the categorical framework.
- (5) **Complementarity checker** (`Complementarity.hs`): Verification that specified observer pairs satisfy the conditions for categorical complementarity (theorem 5.1).

The code is written with extensive comments connecting the implementation to the mathematical definitions in the paper. The implementation serves both a pedagogical and computational purpose: the `PageCurve` module produces numerical data for the

entanglement entropy as a function of retarded time (reproducing the Page curve), while the `Complementarity` module computationally verifies that the asymptotic and infalling observers satisfy all three conditions of categorical complementarity (theorem 5.1). The `KanExtension` module computes the extension deficit and the extrapolation bracket for specified observers, demonstrating their time-dependent behavior.

10 Discussion and Open Questions

10.1 Relation to Existing Approaches

Our categorical framework connects to and extends several approaches to the information paradox.

Black hole complementarity. The original complementarity proposal [3] is operationally motivated: “no single observer can verify both descriptions.” Our theorem 5.2 provides a structural reformulation: the measurement category does not contain an object whose representable functor encodes both descriptions. This is a stronger statement—it is not merely that verification is operationally impossible, but that the mathematical structure of embedded observation precludes such an object.

ER=EPR. The ER=EPR conjecture [28] posits that entanglement (EPR pairs) is geometrically realized as non-traversable wormholes (Einstein–Rosen bridges). In the Yoneda framework, the non-factorizability of the joint representable functor (theorem 4.3) is the categorical avatar of the ER bridge: the entanglement between the radiation and the interior, which prevents factorization, is encoded in the measurement category’s morphism structure.

Quantum extremal surfaces. The quantum extremal surface (QES) prescription [27] identifies the entanglement wedge boundary with a quantum-corrected extremal surface. In our framework, the QES corresponds to the categorical epistemic horizon (theorem 4.5), and the quantum corrections represent the enrichment of the measurement category from classical (area-based) to quantum (area + bulk entropy) morphisms.

The Measurement Boundary Problem. Our earlier work on the MBP [18] identified structural limitations on observational access in emergent spacetime. The black hole information paradox is a specific instance of the MBP where the emergence structure is provided by the holographic duality: the asymptotic observer cannot directly access the pre-geometric (interior) degrees of freedom, and must rely on indirect witnesses (Hawking radiation correlations) to infer interior physics.

10.2 Implications for Quantum Gravity

The Yoneda perspective suggests that the information paradox is not a “problem to be solved” but a structural feature of embedded observation in spacetimes with horizons. The information is not “lost” or “destroyed”—it is encoded in the relational structure of the measurement category in a way that is accessible to some observers (the infalling observer, via $y^{S_{\text{in}}}$) but not to others (the asymptotic observer, via y^{S_∞} , at least until the Page time).

This perspective has several implications:

- (1) **Unitarity is observer-relative.** The evolution appears non-unitary from the asymptotic observer’s perspective (before the Page time) because the observer’s

representable functor does not capture the full unitary structure. This is consistent with relational quantum mechanics [29], which holds that quantum states are observer-relative.

- (2) **The firewall is a categorical artifact.** The firewall paradox arises from attempting to combine incompatible representable functors into a single global description. The resolution is not to modify the physics (introducing a firewall) but to respect the perspectival structure of the measurement category.
- (3) **Information recovery is Kan extension.** The process by which the asymptotic observer recovers interior information after the Page time is naturally described as a Kan extension: the optimal extrapolation from the observer’s growing relational data eventually suffices to reconstruct the interior state.

10.3 Open Questions

- (1) **Enriched and monoidal measurement category.** The black hole measurement category should be enriched over a suitable monoidal category (e.g., **Ban** or C^* -**Alg**) to capture the full quantum structure. Additionally, equipping $\mathbf{Meas}_{\text{BH}}$ with a monoidal (tensor product) structure—where $(\mathcal{S}_1, \Sigma_1, \rho_1) \otimes (\mathcal{S}_2, \Sigma_2, \rho_2)$ represents the composite observer—would make the entanglement structure explicit within the category. The causal constraints would then constrain which tensor products are “allowed” (reflecting the spacetime causal structure), and the non-factorizability of representable functors (theorem 4.3) would be directly expressible as a failure of the functor to preserve the monoidal structure. Developing the enriched Yoneda Constraint in this setting and deriving the quantum corrections to the entropy formulas is an important direction.
- (2) **Dynamical measurement category.** In quantum gravity, the spacetime—and hence the causal structure constraining $\mathbf{Meas}_{\text{BH}}$ —is itself dynamical. How does the measurement category change under Hawking evaporation? Is there a 2-categorical structure encoding the time evolution of $\mathbf{Meas}_{\text{BH}}$?
- (3) **Higher categories and the Page curve.** The Page curve transition has the character of a phase transition. Is there an (∞, n) -categorical description in which the transition is a “cobordism” between different representable functor structures?
- (4) **Computational complexity.** The Kan extension from the radiation’s subcategory involves computing colimits over potentially large over-categories. Is there a computational complexity characterization of information recovery in terms of the complexity of computing the Kan extension?
- (5) **Experimental signatures.** Can the categorical structure be connected to observable signatures in analog gravity experiments (e.g., sonic black holes in BECs) where Hawking radiation has been observed?
- (6) **Cosmological horizons.** The same Yoneda analysis applies to cosmological horizons (de Sitter space). Does the categorical framework provide new insights into the de Sitter entropy and the static patch observer’s epistemic limitations?

- (7) **Replica wormholes.** The replica trick for computing the Page curve involves summing over replica wormhole saddles. Is there a categorical interpretation of replica wormholes in terms of higher morphisms or operads in $\mathbf{Meas}_{\text{BH}}$?

11 Conclusion

We have developed a comprehensive analysis of the black hole information paradox from the perspective of the Yoneda Constraint on Observer Knowledge. The framework provides:

- (1) The **black hole measurement category** $\mathbf{Meas}_{\text{BH}}$ (theorem 3.1), which encodes the causal structure of the black hole spacetime as constraints on morphisms between observer objects.
- (2) The **horizon obstruction** (theorem 3.6): the event horizon manifests as the absence of morphisms from exterior to interior objects, making the asymptotic observer’s representable functor vanish on the interior.
- (3) **Categorical complementarity** (theorem 5.2): the infalling and asymptotic observers are categorically complementary—their representable functors are non-isomorphic and non-simultaneously-embeddable.
- (4) **Firewall resolution** (theorem 5.3): the AMPS paradox dissolves because the measurement category does not contain an object whose representable functor encodes both the $B-A$ and $B-C$ entanglements simultaneously.
- (5) The **Page curve as presheaf transition** (theorem 7.3): the Page time corresponds to a structural transition in the representable functor, where the Kan extension deficit begins to decrease as the radiation observer’s relational data becomes sufficient to indirectly access interior information.
- (6) The **island formula as deficit minimization** (theorem 7.6): the island represents the optimal extension of the observer’s accessible subcategory across the horizon, minimizing the Kan extension deficit.
- (7) Connections to **holographic entanglement** (section 8): the Ryu–Takayanagi formula and quantum error correction structure of holography are naturally expressed as properties of the Kan extension in the holographic measurement category.

The central message of this paper is that the black hole information paradox is a consequence of the Yoneda Constraint: the structural principle that embedded observers know reality only relationally, through their representable functors, and that different observers in a spacetime with horizons necessarily have incompatible relational perspectives. The “paradox” arises only when one ignores the perspectival structure and attempts to construct a global description from inherently local data. The resolution is not to modify the physics but to respect the categorical structure of embedded observation.

We believe that the Yoneda perspective on the information paradox, combined with the recent island formula developments, points toward a deeper understanding of the relationship between gravity, quantum information, and the structure of observation itself.

Acknowledgments

The author thanks the YonedaAI Research Collective for ongoing support and discussions on the categorical foundations of quantum gravity. This work was developed in part through AI-assisted research workflows.

AI-assisted research disclosure. Portions of this manuscript were developed through collaborative workflows with AI language models (Claude, Anthropic). The AI assisted with literature synthesis, LaTeX typesetting, proof structuring, and code development. All physical arguments, mathematical claims, and research directions are the responsibility of the human author.

A Detailed Proofs

A.1 Proof of theorem 5.2 (Detailed)

Proof. We verify the three conditions of theorem 5.1 in detail.

Condition (i): By theorem 4.3, $y^{\mathcal{S}_\infty} \not\cong y^{\mathcal{S}_{\text{in}}}$ because they differ on interior objects.

Condition (ii): Suppose for contradiction that $\mathcal{S}_3 \in \mathbf{Meas}_{\text{BH}}$ exists with the stated property. Then:

- $y^{\mathcal{S}_3}$ restricts to $y^{\mathcal{S}_\infty}$ on exterior objects, meaning $\text{Hom}(\mathcal{S}_3, X) \cong \text{Hom}(\mathcal{S}_\infty, X)$ for all exterior X .
- $y^{\mathcal{S}_3}$ restricts to $y^{\mathcal{S}_{\text{in}}}$ on interior objects, meaning $\text{Hom}(\mathcal{S}_3, Y) \cong \text{Hom}(\mathcal{S}_{\text{in}}, Y)$ for all interior Y .

The first condition requires \mathcal{S}_3 to have the same relational structure as \mathcal{S}_∞ with respect to exterior objects—in particular, \mathcal{S}_3 must have access to the Hawking radiation at late times. The second condition requires \mathcal{S}_3 to cross the horizon and access the interior.

By the Yoneda embedding (full and faithful), \mathcal{S}_3 is determined up to isomorphism by its representable functor. The combined conditions determine a unique (up to isomorphism) object that simultaneously has the exterior relational structure of \mathcal{S}_∞ and the interior relational structure of \mathcal{S}_{in} .

However, the causal structure of the black hole spacetime forbids this. The asymptotic observer collects radiation over a time $\sim M^3$ (the evaporation time in Planck units). An observer who waits this long cannot subsequently cross the horizon of the original black hole (which has evaporated). Conversely, an observer who crosses the horizon early cannot access the late-time radiation. No worldline in the black hole spacetime has a causal domain that simultaneously includes \mathcal{J}^+ at late retarded times and the deep interior. Hence \mathcal{S}_3 does not exist.

Condition (iii): The non-existence of \mathcal{S}_3 is enforced by the causal structure: the light-speed limit prevents any observer from simultaneously collecting late radiation and crossing the horizon. \square

A.2 Proof of theorem 7.6 (Detailed)

Proof. Let $J : \mathbf{Meas}_{\text{BH}}|_{\mathcal{S}_{\text{Rad}}(u)} \hookrightarrow \mathbf{Meas}_{\text{BH}}$ and $J_I : \mathbf{Meas}_{\text{BH}}|_{\mathcal{S}_{\text{Rad}}(u) \cup I} \hookrightarrow \mathbf{Meas}_{\text{BH}}$ be the inclusion functors. Since $\mathbf{Meas}_{\text{BH}}|_{\mathcal{S}_{\text{Rad}}(u)} \subset \mathbf{Meas}_{\text{BH}}|_{\mathcal{S}_{\text{Rad}}(u) \cup I}$, the Kan extension from the larger subcategory has access to more data.

For any interior object $X \in I$, the over-category $(J_I \downarrow X)$ contains at least the identity morphism id_X , whereas $(J \downarrow X)$ may be empty (when X has no morphisms from exterior objects). This means:

$$\text{Lan}_{J_I}(\mathfrak{D}_I \circ J_I)(X) \supseteq \mathfrak{D}(X) = \mathfrak{R}(X),$$

while $\text{Lan}_J(\mathfrak{D} \circ J)(X)$ may be trivial.

The comparison natural transformation $\eta : \text{Lan}_{J_I}(\mathfrak{D}_I \circ J_I) \Rightarrow \mathfrak{R}$ is “closer to an isomorphism” than $\eta' : \text{Lan}_J(\mathfrak{D} \circ J) \Rightarrow \mathfrak{R}$, in the sense that the cokernel $\Delta(\mathcal{S}_{\text{Rad}}(u) \cup I)$ is contained in $\Delta(\mathcal{S}_{\text{Rad}}(u))$. Hence the deficit decreases.

The island formula’s area term $\frac{\text{Area}(\partial I)}{4G_N}$ can be understood as the “cost” of the additional morphisms needed to connect the island to the exterior: crossing the horizon boundary requires morphisms whose existence has a penalty measured by the area in Planck units. The bulk entropy term $S_{\text{bulk}}(\text{Rad} \cup I)$ measures the remaining entanglement after the island is included. The minimization finds the optimal trade-off between these costs. \square

B Categorical Definitions

Presheaf categories. For a category \mathcal{C} , the presheaf category $\text{PSh}(\mathcal{C}) = [\mathcal{C}^{\text{op}}, \mathbf{Set}]$ is the category of contravariant functors from \mathcal{C} to \mathbf{Set} . A representable presheaf is one isomorphic to $\text{Hom}_{\mathcal{C}}(-, X)$ for some $X \in \mathcal{C}$.

Coprod of presheaves. Given presheaves $F, G \in \text{PSh}(\mathcal{C})$ and a subpresheaf $H \hookrightarrow F, H \hookrightarrow G$, the pushout $F \sqcup_H G$ is defined by the universal property: for each object c , $(F \sqcup_H G)(c) = (F(c) \sqcup G(c))/\sim$ where \sim identifies elements of $F(c)$ and $G(c)$ that correspond under $H(c)$.

Pointwise Kan extension. For $K : \mathcal{C} \rightarrow \mathcal{D}$ and $F : \mathcal{C} \rightarrow \mathcal{E}$, the pointwise left Kan extension is $\text{Lan}_K F(d) = \text{colim}_{(c, K(c) \rightarrow d) \in (K \downarrow d)} F(c)$ when the colimit exists.

C Physical Constants and Conventions

We work in natural units $c = \hbar = k_B = 1$ unless otherwise noted. Newton’s constant G_N and the Planck length $\ell_P = \sqrt{G_N \hbar / c^3}$ are retained as they set the scale for quantum gravity effects. The Bekenstein–Hawking entropy is $S_{\text{BH}} = A/(4G_N)$ where A is the area of the event horizon. The Page time is $t_P \sim M^3$ in Planck units, where M is the initial black hole mass.

References

- [1] S. W. Hawking, “Breakdown of predictability in gravitational collapse,” *Phys. Rev. D* **14**, 2460–2473 (1976).
- [2] S. W. Hawking, “Particle creation by black holes,” *Commun. Math. Phys.* **43**, 199–220 (1975).

- [3] L. Susskind, L. Thorlacius, and J. Uglum, “The stretched horizon and black hole complementarity,” *Phys. Rev. D* **48**, 3743 (1993). arXiv:hep-th/9306069.
- [4] L. Susskind, “String theory and the principle of black hole complementarity,” *Phys. Rev. Lett.* **71**, 2367 (1993). arXiv:hep-th/9307168.
- [5] G. ’t Hooft, “Dimensional reduction in quantum gravity,” in *Salamfestschrift*, World Scientific, 1993. arXiv:gr-qc/9310026.
- [6] L. Susskind, “The world as a hologram,” *J. Math. Phys.* **36**, 6377–6396 (1995). arXiv:hep-th/9409089.
- [7] J. Maldacena, “The large N limit of superconformal field theories and supergravity,” *Int. J. Theor. Phys.* **38**, 1113–1133 (1999). arXiv:hep-th/9711200.
- [8] A. Almheiri, D. Marolf, J. Polchinski, and J. Sully, “Black holes: complementarity vs. firewalls,” *JHEP* **2013**, 062 (2013). arXiv:1207.3123.
- [9] G. Penington, “Entanglement wedge reconstruction and the information problem,” *JHEP* **09**, 002 (2020). arXiv:1905.08255.
- [10] A. Almheiri, N. Engelhardt, D. Marolf, and H. Maxfield, “The entropy of bulk quantum fields and the entanglement wedge of an evaporating black hole,” *JHEP* **12**, 063 (2019). arXiv:1905.08762.
- [11] A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian, and A. Tajdini, “The entropy of Hawking radiation,” *Rev. Mod. Phys.* **93**, 035002 (2021). arXiv:2006.06872.
- [12] D. N. Page, “Information in black hole radiation,” *Phys. Rev. Lett.* **71**, 3743 (1993). arXiv:hep-th/9306083.
- [13] D. N. Page, “Average entropy of a subsystem,” *Phys. Rev. Lett.* **71**, 1291 (1993). arXiv:gr-qc/9305007.
- [14] M. Long and The YonedaAI Collaboration, “The significance of the Yoneda Constraint on observer knowledge to foundational physics: from quantum to classical,” GrokRxiv:2026.YonedaAI-0017, February 2026.
- [15] M. Long and The YonedaAI Collaboration, “A Yoneda-lemma perspective on embedded observers: relational constraints from quantum measurement to classical phase space,” GrokRxiv:2026.02.yoneda-observer-constraint, February 2026.
- [16] M. Long and The YonedaAI Collaboration, “The embedded observer constraint: on the structural bounds of scientific measurement,” GrokRxiv:2026.02.long-yonedaai-embedded-observer, February 2026.
- [17] M. Long and The YonedaAI Collaboration, “The embedded observer constraint: on the structural bounds of scientific measurement (revised),” GrokRxiv:2026.02.embedded-observer-constraint, February 2026.
- [18] M. Long and The YonedaAI Collaboration, “The measurement paradox in emergent spacetime physics: structural limits on observational access to pre-geometric ontology,” YonedaAI Research Collective, February 2026.

- [19] S. Abramsky and A. Brandenburger, “The sheaf-theoretic structure of non-locality and contextuality,” *New J. Phys.* **13**, 113036 (2011). arXiv:1102.0264.
- [20] D. Petz, “Sufficient subalgebras and the relative entropy of states of a von Neumann algebra,” *Commun. Math. Phys.* **105**, 123–131 (1986).
- [21] D. Petz, “Sufficiency of channels over von Neumann algebras,” *Quart. J. Math.* **39**, 97–108 (1988).
- [22] P. Hayden and G. Penington, “Learning the alpha-bits of black holes,” *JHEP* **12**, 007 (2019). arXiv:1807.06041.
- [23] G. T. Horowitz and J. Maldacena, “The black hole final state,” *JHEP* **02**, 008 (2004). arXiv:hep-th/0310281.
- [24] A. Almheiri, X. Dong, and D. Harlow, “Bulk locality and quantum error correction in AdS/CFT,” *JHEP* **04**, 163 (2015). arXiv:1411.7041.
- [25] D. Harlow, “The Ryu-Takayanagi formula from quantum error correction,” *Commun. Math. Phys.* **354**, 865–912 (2017). arXiv:1607.03901.
- [26] S. Ryu and T. Takayanagi, “Holographic derivation of entanglement entropy from the anti-de Sitter space/conformal field theory correspondence,” *Phys. Rev. Lett.* **96**, 181602 (2006). arXiv:hep-th/0603001.
- [27] N. Engelhardt and A. C. Wall, “Quantum extremal surfaces: holographic entanglement entropy beyond the classical regime,” *JHEP* **01**, 073 (2015). arXiv:1408.3203.
- [28] J. Maldacena and L. Susskind, “Cool horizons for entangled black holes,” *Fortsch. Phys.* **61**, 781–811 (2013). arXiv:1306.0533.
- [29] C. Rovelli, “Relational quantum mechanics,” *Int. J. Theor. Phys.* **35**, 1637–1678 (1996). arXiv:quant-ph/9609002.
- [30] J. D. Bekenstein, “Black holes and entropy,” *Phys. Rev. D* **7**, 2333 (1973).
- [31] S. D. Mathur, “The information paradox: a pedagogical introduction,” *Class. Quant. Grav.* **26**, 224001 (2009). arXiv:0909.1038.
- [32] D. Harlow, “Jerusalem lectures on black holes and quantum information,” *Rev. Mod. Phys.* **88**, 015002 (2016). arXiv:1409.1231.
- [33] D. Wallace, “Why black hole information loss is paradoxical,” in *Beyond Spacetime*, Cambridge University Press, 2020. arXiv:1710.03783.
- [34] W. G. Unruh and R. M. Wald, “Information loss,” *Rep. Prog. Phys.* **80**, 092002 (2017). arXiv:1703.02140.
- [35] D. Marolf, “The black hole information problem: past, present, and future,” *Rep. Prog. Phys.* **80**, 092001 (2017). arXiv:1703.02143.