# Unpacking Black Hole Complementarity

Siddharth Muthukrishnan

Department of History and Philosophy of Science, University of Pittsburgh,
Pittsburgh PA 15260, USA
siddharth@pitt.edu

Oct-22-2022

**Abstract**

"Black hole complementarity" is a label attached to an influential set of ideas that have emerged in response to the black hole information paradoxes. However, many different claims come under that label in the literature, and it can often be confusing what an appeal to "black hole complementarity" is meant to do. Unpacking this literature, I argue that black hole complementarity is about the consistency of quantum characterizations of an evaporating black hole. To bring this into focus, I delineate two consistency claims—i.e., two principles of black hole complementarity: operational complementarity and descriptive complementarity. These two principles are more or less attractive depending on one's background philosophy of science. Instrumentalists will lean towards operational complementarity while scientific realists will lean towards descriptive complementarity. If one resists instrumentalism (as many do, for good reason), one has a *prima facie* reason to adopt the descriptive principle and reject the operational principle. However, a series of thought experiments in the physics literature on black hole complementarity gives us strong reasons to adopt the operational principle and reject the descriptive principle. This shows that if we can stomach instrumentalism, then operational complementarity may suffice to resolve the black hole information paradox.

## 1 Introduction

The black hole information paradox identifies a central, productive tension within quantum mechanical treatments of black holes. It suggests that something goes wrong if one treats a black hole as a quantum statistical object and also treats the physics near the horizon using quantum field theory. Both the fact that the paradox involves a tension in the application of quantum mechanics to general relativistic systems and the fact that the paradox can be articulated in regimes with good theoretical control has made it a central problem in modern theoretical physics.[1]

"Black hole complementarity" is a label attached to an influential set of ideas (Susskind et al., 1993; Lowe et al., 1995; Almheiri et al., 2013; Hayden and Preskill, 2007; Harlow, 2016) that have emerged in response to the paradox. Many different claims come under that label in the literature, and it can often be confusing what an appeal to "black hole complementarity" is meant to do. Simplyfing somewhat, black hole complementarity aims to be a way to sand off the bite of the black hole information paradox by arguing that the inconsistency identified by the information paradox does not ramify up into a physically or an empirically relevant sort of inconsistency. To make this more precise, I will tease apart the different threads in this literature and argue that black hole complementarity is best understood as a principle about the *consistency* of characterizations of the physics of an evaporating black hole.

To bring this out clearly, I will first distinguish two separate principles of black hole complementarity embedded in these discussions: an operational principle and a descriptive principle. The operational principle says that experiments conducted by any *single* observer in or near an evaporating black hole will always be consistent with quantum mechanics, as long as these experiments cannot probe physics beyond the Planck scale. The princi-

[1]See, e.g., Belot et al. (1999); Wallace (2017); Marolf (2017) for systematic overviews.

ple is *operational* because it makes ineliminable appeal to what is empirically accessible to *observers*.[2] The descriptive principle says that the *infalling* and *exterior* descriptions of the physics of an evaporating black hole are consistent descriptions of the same physics. Holography, a popular approach to the black hole information paradox, may, on some readings, be seen as subscribing to the descriptive principle of complementarity (Raju, 2020, pp. 37-72). The operational principle has been stated more-or-less explicitly in the physics literature (Hayden and Preskill, 2007; Bousso, 2013; Nomura et al., 2013), while the descriptive principle is only implicit in parts of the physics and philosophy literature (Lowe et al., 1995; Belot et al., 1999; van Dongen and de Haro, 2004; Wallace, 2017), and part of my contribution is to identify it and set it out clearly.

Our background philosophy of science guides which of these two principles we have *prima facie* reasons to adopt. If you are of a realist inclination, then you will lean towards a descriptive principle. If you are of an instrumental persuasion, you might find the operational principle attractive. However, these are just *prima facie* reasons, and hence defeasible by evidence. I will argue that the recent physics literature indicates that the operational principle succeeds and the descriptive principle fails. More precisely, the physics of quantum black holes seems to *describe* scenarios that violate some tenets of quantum mechanics, leading to the failure of descriptive complementarity. However, we are unable to *operationalize* these violations—no single observer is able to see a violation of quantum mechanics. Consequently, we are forced to at least tentatively adopt operational complementarity.

Where does one go from there? The realisitically inclined should take the operational principle as data to be explained by future descriptively-interpretable physics. However, if

[2] I use the term "operational" as it is used in the physics literature on black hole complementarity. I'm not using this term in the way Percy Bridgman does. (See Chang (2021) for an overview of the Bridgmanian sense of the term.)

one is comfortable with operationalism in physics, then the black hole information paradox plausibly dissolves away.

*Plan.* First, I will briefly review the black hole information paradox, distinguishing between the total evaporation paradox and the Page-time paradox (Sec. 2). Our focus will be on the latter. Next, I will introduce black hole complementarity, and distinguish and define the operational and descriptive principles of complementarity (Sec. 3). This will then allow us see why black hole complementarity is best understood as a kind of consistency. Following this, I will discuss three families of thought experiments, each attempting to identify contradictions in black hole physics. The first will be about what happens to a horizon crosser (Sec. 4). Here, I will argue that both operational and descriptive complementarity are successful. The second will involve a potential violation of the quantum no-cloning theorem (Sec. 5). Here, descriptive complementarity fails while operational complementarity is vindicated. The final thought experiment I study involves a potential violation of entanglement monogamy (Sec. 6). Here too, descriptive complementarity fails whereas operational complementarity succeeds. I conclude by discussing some the consequences of these arguments for the black hole information paradox (Sec. 7).

## 2 The black hole information paradox

Consider a black hole that forms from the collapse of matter and then evaporates via Hawking radiation (Hawking, 1975). (See Fig. 1.) At least two different, but related, paradoxes can be identified in this scenario: the *total evaporation paradox* and the *Page-time* paradox (Wallace, 2017). (Thus, "'the' black hole information paradox" is somewhat of a misnomer.) In this work, I will focus on the Page-time paradox. This is because the case for the total evaporation 'paradox' being *really* paradoxical is rather weak (Hawking (1976); Mathur (2009); Unruh and Wald (2017); Maudlin (2017); Wallace (2017)). Moreover, black hole complementarity only makes sense when seen as a response to the Page-time paradox.

We will now examine these two paradoxes in turn in the context of an evaporating black hole. Note, however, that the Page-time paradox arises even in black holes that do not evaporate *away*—e.g., by being in thermal equilibrium with their own radiation. It is not clear how to state the total-evaporation paradox in such a case. So, for uniformity of exposition, I will stick to the case of a black hole that is evaporating away.

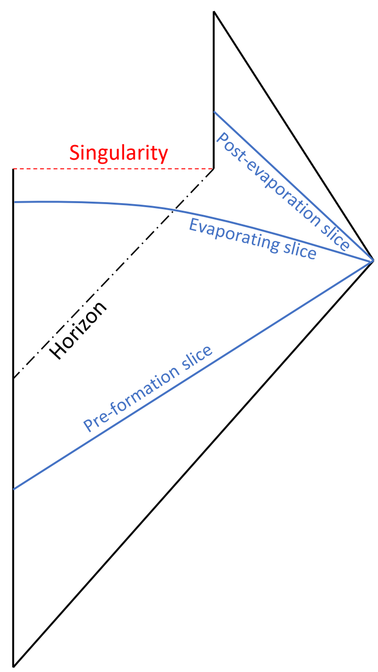


**Figure 1:** A Penrose diagram for a black hole that forms and then evaporates away. (See, e.g., Strominger (1995, pp. 4-14) for how to draw and interpret Penrose diagrams.)

### *2.1 Total evaporation paradox*

The *total evaporation paradox* is the conflict between the following two statements: (1) the dynamical evolution of the evaporating black hole is *unitary*—and hence, reversible and information-preserving; and (2) the Hawking radiation that is emitted by the black hole during its lifetime is thermal.

Statement (1) is plausible because it is an axiom of quantum mechanics that all closed systems evolve unitarily. And it is plausible that the system under study is closed because we have not restricted the system to be in some sub-region of space, and thus there is no "outside" system to speak of.

Statement (2) is plausible because it is a consequence of Hawking's argument (Hawking (1975)). Moreover, Hawking's argument is supported, though not replicated, by various other arguments that make slightly different assumptions.[3] To be sure, some of these other arguments are strictly weaker than Hawking's argument, in the following way. Hawking's argument concludes that the radiation from a black hole is *perfectly* thermal, i.e., every subsystem of the radiation is also thermal; consequently, there is no entanglement between subsystems of the radiation (cf. Sec. 6). In contrast, consider arguments relying on the self-consistency of the semiclassical Einstein field equations in the black hole exterior (Candelas (1980); Sciama et al. (1981)). Consistent with Hawking's argument, they conclude that individual quanta of radiation will be thermal. However, they do not rule out non-thermal subsystems of radiation. Thus, such arguments provide evidence for Hawking's conclusion by supporting an implication of Hawking's conclusion.

The conflict between (1) and (2) can be brought out as follows. Assume that the quantum state of the pre-formation degrees of freedom is pure. If Hawking radiation is thermal (i.e., statement (2)), then the quantum state of the degrees of freedom outside the horizon must be mixed throughout evaporation. This, along with the assumption of unitarity (i.e., statement (1)), implies that the post-evaporation state must be mixed as well because the degrees of freedom behind the horizon are forever inaccessible, and hence we must inevitably discard

[3] See (Wallace, 2018a, p. 61) for a summary of these arguments.

the behind-the-horizon degrees of freedom to obtain the post-evaporation state. Thus, (1) and (2) together imply that the transition between pre-formation degrees of freedom and post-evaporation degrees of freedom will be a pure-to-mixed transition. That is, the physics connecting epochs preceding the formation of and following the extinction of a black hole will be essentially non-unitary. This would be a paradox—say, in Quine's sense (Quine (1962))—to the extent one finds the above argument compelling and the conclusion absurd.

Many, however, do not find the conclusion absurd (Hawking (1976); Mathur (2009); Unruh and Wald (2017); Maudlin (2017); Wallace (2017)). Why not? Because the conclusion is absurd only to the extent that we expect unitarity to hold in this context, but we ought not to expect it here. Unitarity in quantum field theory is expected to hold only between different Cauchy slices. Here, however, between the pre-formation and post-evaporation regions of the spacetime, is a singularity. This prevents a foliation of the spacetime by Cauchy slices.[4] Thus, some (Hawking (1976); Maudlin (2017); Unruh and Wald (2017)) take the total evaporation "paradox" to just be a successful argument for the failure of global unitarity.

Others do not find the argument compelling. They argue that because we have an incomplete understanding of the physics of the singularity, we have no strong reason to believe in the thermality of Hawking radiation *throughout* the evaporation of the black hole. Arguments for the thermality of Hawking radiation rely on semiclassical physics—quantum field theory on a curved background—but this obviously breaks down towards the end of the evaporation process as we near the singularity. This has lead to proposals like remnants[5] and thunderbolts (Hawking and Stewart (1993)) to explain how unitarity can be preserved in the physics of late-stage evaporation.

[4]A Cauchy slice is a spacelike hypersurface such that every timelike curve can be extended to a timelike curve which intersects with the hypersurface exactly once.

[5]See Chen et al. (2015) for a recent review.

### 2.2 *Page-time paradox*

In contrast, the *Page-time paradox* (Page (1993); Wallace (2017)) does present a compelling argument to a seemingly absurd conclusion. The argument is compelling because its premises rely on physics in regimes believed to be under good theoretical control. Consequently, the Page-time paradox has dominated recent discussion by physicists (Mathur (2009); Polchinski (2015); Harlow (2016); Marolf (2017)). As we will see below, black hole complementarity can then be seen as a way of attenuating the absurdity of the conclusion.

According to the Page-time paradox, we get a contradiction between three statements: (A) the evaporation process is unitary; (B) the black hole is a quantum statistical mechanical system with its von Neumann entropy (i.e., its fine-grained entropy) bounded above by its Bekenstein-Hawking entropy (i.e., its microcanonical entropy); [6] and (C) Hawking radiation is perfectly thermal throughout evaporation.

Let us bring out the contradiction. (A) entails that the von Neumann entropy of the radiation has to be equal to the von Neumann entropy of the black hole degrees of freedom. Thus, from (C), we can conclude that the entropy of the black hole keeps increasing throughout the evaporation as more and more thermal photons are emitted as radiation. However, the Bekenstein-Hawking entropy—which is proportional to the area of the horizon—of the black hole will keep decreasing as the evaporation proceeds. Consequently, (B) implies that the von Neumann entropy of the black hole will have to start decreasing as well at some point in time so as to avoid becoming larger than the Bekenstein-Hawking entropy. Thus, (A), (B), and (C) together imply that the von Neumann entropy of the black hole both keeps increasing and has to start decreasing at some time during evaporation. We have a

[6] See, Wallace (2018b) for an extended defense of the aptness of a statistical-mechanical description of black holes.

contradiction.[7]

To be a little more quantitative, Page (1993) argued that if a black hole evaporates unitarily and is a statistical mechanical system with the Bekenstein-Hawking entropy as its microcanonical entropy, then the total state of the radiation it emits has to start deviating from thermality starting roughly halfway into the lifetime of the black hole. This latter time is called the *Page time*. (See Fig. 2.) On the other hand, Hawking's argument tells us that the radiation from a black hole has to be thermal throughout evaporation, even past the Page time. Thus, the paradox.

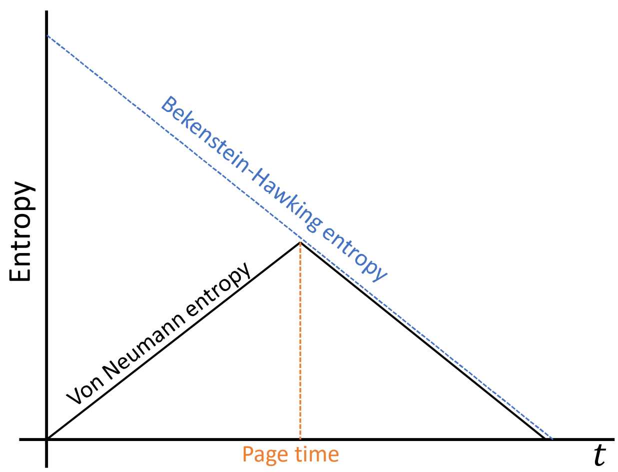


**Figure 2:** The Page curve for the entropy of an evaporating black hole. The Page time is the time at which the von Neumann entropy of the black hole has to start decreasing so as to not exceed the Bekenstein-Hawking bound.

It is easy to see why we take the conclusion of the argument to be absurd: It is a contradiction! To see why the argument is compelling, note that the argument was made without any reference to the singularity or to post-evaporation quantum states. We are invoking a

[7]Note that we cannot bring out a paradox by only employing (A) and (C), as the total evaporation paradox attempted to do. Because without (B), there is nothing to prevent unitary dynamics from increasing the entropy of the black hole without bound.

quantum statistical description of a black hole in a regime where time evolution between slices is perfectly well-defined and hence there is no problem with defining unitarity between slices.[8] Moreover, it is also in a regime where we have no reason to expect Hawking's argument for the thermality of black hole radiation to fail. These reasons make the Page-time paradox much more compelling than the total evaporation paradox.

### *2.3 The stretched horizon*

The *stretched horizon* is a striking consequence of a unitary quantum statistical description (i.e., assumptions (A) and (B)) of a black hole. It will feature in the discussion to follow, so let me briefly describe it here. To hold on to a quantum statistical description of the black hole (i.e., (B)) we need an effective field theory for the exterior which has an entropy at the horizon bounded above by the Bekenstein-Hawking entropy. Moreover, to hold on to a unitary description of the black hole (i.e., (A)), the exterior field theory requires a boundary that is strictly above the horizon (since things that cross the horizon cannot re-emerge) which can absorb and re-emit the information that crosses it. Both of these can be satisfied by choosing a boundary surface located one Planck length above the horizon. (This location is set by the demand that the entropy of the exterior field theory not exceed the Bekenstein-Hawking entropy at the horizon.) This is the *stretched horizon*. Thus, the stretched horizon is timelike unlike the true horizon, which is null. Note that the encoding of the infalling degrees of freedom onto the stretched horizon will necessarily require Planck-scale physics;[9] nevertheless, the stretched horizon can maintain unitarity for exterior physics by the way it interacts with the low-energy physics of the exterior. The stretched horizon

[8]These slices need not be Cauchy, since a spacetime admits of a Cauchy slice if and only if it is globally hyperbolic (Geroch, 1970). And in the context of the Page-time paradox, we are not requiring global hyperbolicity.

[9]See, e.g., (Banks, 1995, pp. 9-11).

is a real entity in the reference frame of observers hovering outside the black hole.[10] In the classical limit, it will have its own distinctive viscosity and electrical resistance; it will respond in a local way to external perturbations—it will radiate, carry electrical currents, and oscillate.

In the classical limit, the stretched horizon becomes the membrane in the membrane paradigm of Thorne et al. (1986).[11] The membrane paradigm was introduced by Macdonald, Price, Thorne, Blandford, and others in order to understand black holes as real astrophysical objects that can interact with their surrounding electrical and magnetic fields. They found they could interpret many of the results of classical general relativity combined with electromagnetism in a way that allowed exterior hovering observers to view a black hole as a three-dimensional spheroidal object that had a surface which could carry currents and develop charges in response to external fields and support mechanical oscillations, much like stars do. This paradigm was justified by the fact that they could translate all their results derived in the membrane paradigm into the standard spacetime picture of black holes.[12] This helped them better understand how black holes could be quasars. One can view the quantum stretched horizon as a quantization of the classical membrane. However, this is not the way I introduced the stretched horizon above, for it can be justified in entirely quantum terms.

With the stage set, let's turn now to main topic of this paper: black hole complementarity.

[10]See (Susskind and Lindesay, 2005, Ch. 7) for a pedagogical introduction to the stretched horizon. Susskind et al. (1993) provided the first systematic treatment of the idea of a stretched horizon. However, 't Hooft (1985) is an important precursor (see also (Harlow, 2016, pp. 18-19) for a summary of 't Hooft's argument).

[11]See (Price and Thorne, 1988)for an accessible overview.

[12]So not really a Kuhnian paradigm, even though they took the word from Kuhn.

## 3 Black hole complementarity

The Page-time paradox points to a contradiction between a statistical mechanical application of quantum mechanics and a field-theoretic application of quantum mechanics, with both applications happening in regimes where we believe quantum mechanics works well. It is to assuage this contradiction that we now turn to black hole complementarity.

You might think: A *contradiction* is not the sort of thing that one "assuages"; if your theory yields a contradiction, then so much the worse for your theory. But this response is unsatisfying because the history of science provides many examples of seemingly inconsistent theories that are extremely successful.[13] However, an inconsistency becomes much more troubling if it can be directly subjected to empirical test. Thus, in our context, the natural question to ask is: Can we point to some way in which the Page-time paradox leads to a clear observable violation of the predictions of quantum mechanics? If not, then that is the sense in which we would have "assuaged" the contradiction.

How might we extract an observable violation of quantum mechanics from the Page-time paradox? Consider an observer who hovers outside a black hole, patiently collecting Hawking radiation. Such an observer is often termed a *fiducial observer* (see, e.g., (Susskind and Lindesay, 2005, p. 21)), and in this paper we shall call such an observer "Bob". (Bob will frequently be contrasted with *Alice*, who will be an observer who freely falls into the black hole.) Bob waits past the Page time and he observes the total state of the radiation deviate from thermality, as predicted by the Page curve. He infers from this, following the logic of the Page-time paradox, that something has gone wrong with semiclassical quantum field theory near the horizon. Bob decides to go to the horizon and empirically test what has gone wrong with Hawking's argument. If he can do this, he will

[13] See, e.g., Vickers (2013) for several examples of inconsistent but successful scientific theories.

have an empirical test of the inconsistency.

To this end, he will want to hover above the horizon to examine near-horizon modes. However, to do so he would need to sustain Planck-scale accelerations.[14] And even if he could sustain such accelerations, his experiments would require Planck-length sensitivity because the physics of the true horizon would be behind the stretched horizon, and the gap between the two is just one Planck length. Thus, Bob would need control over Planck-scale physics to conduct his test (Susskind and Thorlacius (1994)). So it seems impossible that Bob will experimentally detect the failure of some prediction of quantum mechanics.

But perhaps cleverer attempts can succeed? In much of the rest of this paper, we will discuss *attempts* to show that the discrepancies arising from the Page-time paradox result in a single-observer violation of quantum mechanics—and how those attempts fail. That such attempts will continue to fail is the import of black hole complementarity. In other words, that a consistent way of characterizing the physics of a quantum black hole is available despite the inconsistency delineated by the Page-time paradox is the core tenet of black hole complementarity.

### *3.1 Operational and descriptive complementarity*

As I have described it so far, black hole complementarity is an operational principle, i.e., a principle that makes essential, ineliminable reference to the capabilities of observers. Let me set it out explicitly:

> **Operational complementarity**: No experiment attempting to create the observation of a direct contradiction of the rules of quantum mechanics by a single observer near, or in, black holes will succeed, as long as that observer cannot

[14] Even though the surface gravity can be small as you like at the horizon of a black hole, the surface gravity is defined as the local proper acceleration times the gravitational redshift. So, because at the horizon the redshift goes to zero, the acceleration needed to hover has to diverge.

empirically access Planck-scale physics.

A word on what we mean by "observer" here. Obviously we don't mean actual an human, even though that might be suggested by names such as "Alice" and "Bob". After all, it is implausible for a human to survive and conduct measurements in extreme regions like that between the stretched horizon and the true horizon. Rather, an "observer" here is some kind of observational system that can collect and process quantum mechanical data. I will leave the notion of "observer" underspecified beyond this admittedly vague characterization. However, once we start considering examples, it should be clear what sorts of things might be observers in the relevant sense. And despite these systems not usually being humans, we will stick with the human-name convention common in this literature. But if that bothers you, it might help you to imagine some exquisitely sensitive automated instrument rather than a human.

To highlight the prevalence of something like operational complementarity in the recent physics literature, let me present some quotes:

- (Hayden and Preskill, 2007, p. 2): "..."black hole complementarity," according to which no violations of the accepted principles of quantum physics can be detected by any observer, whether outside or inside the black hole."
- (Bousso, 2013, p. 1): "Complementarity distinguishes the viewpoint of an observer who remains far from the black hole, Bob, from that of an infalling observer, Alice. These viewpoints have to be consistent as long as they can be operationally compared."
- (Nomura et al., 2013, p. 1): "Black hole complementarity asserts that there is no contradiction between the two pictures [i.e., the infalling and the exterior], since the statements by the two observers cannot be operationally compared..."

Clearly an operational principle is being employed in the literature. One may welcome this, especially if one is already sympathetic to some kind of instrumentalism about scientific theories. For, if you think that our scientific theories are essentially predictive tools, then it shouldn't be surprising that some of the principles involved in our theories make essential reference to the capabilities of observers—after all, guiding observers is what theories are for. And such a view isn't at all alien in the context of quantum mechanics. If one prefers an instrumentalist response to the quantum measurement problem, then an instrumentalist principle in the context of the quantum physics of black holes is just so much more grist for one's mill.

Typically however, in philosophy of physics, and in philosophy of science more broadly, there is a wariness surrounding operational principles. This is of a piece with the popularity of scientific realism in these communities.[15] This is because it's unclear how principles that refer to observers could be interpreted as describing a mind-independent reality.

The relative antipathy towards instrumentalist views in broader philosophy of science is related to the difficulties of making work a verificationist semantics for scientific theories, as was attempted by the logical empiricists.[16] Within philosophy of physics specifically, the following argument has had significant purchase: Observers are physical systems, constituted by atoms and the like, and thus what they can or cannot do should ultimately be explicable in terms of observer-free physics. Hence, any principle that takes observers as primitive incurs the burden of explaining why we cannot state the capabilities of observers in observer-free terms given their physical constitution.[17]

[15]The 2020 Philpapers Survey found that about 60% of the survey's target faculty in philosophy of science accept or lean towards scientific realism, with that fraction rising to about 70% of the target faculty in philosophy of physical science (Bourget and Chalmers, 2020).

[16]A particularly influential critique of verificationism was advanced by Quine (1951, Sec. V).

[17]This is a version of an argument advanced by Bell (1990) in the context of the quantum measurement problem.

Given all this, philosophers may find the operationalism in operational complementarity unpalatable. And so one might ask whether the reference to observers in operational complementarity is really ineliminable. One might seek to eliminate the reference to observers in one's statement of a principle of complementarity and to state it in a way that does not appeal to observers and their capabilities. I will call such a way of stating complementarity *descriptive*. On this approach, instead of constraining what is possible for observers, we want to constrain our theories, interpreted as *descriptions* of the world. I will now supply one natural way of articulating a descriptive principle of complementarity.

If we are to constrain descriptions, we need to first decide which descriptions we'll be constraining. Well, our goal is to eliminate the appeal to observers in our operational principle, so a natural way to generate the descriptions to be constrained would be the descriptions that are natural to the two kinds of observers that often come up in this context—fiducial and infalling. So let us then clearly have on the table the two different descriptions that attach to these two kinds of observers.

- *Exterior description*—This is a consistent low-energy quantum mechanical description of the black hole and its exterior that includes degrees of freedom outside the horizon with the stretched horizon as the boundary. On this description, the black hole evolves unitarily and is a statistical-mechanical object with entropy bounded above by the Bekenstein-Hawking entropy. This description will be attributed to the black hole by observers hovering outside the horizon, i.e., fiducial observers.

- *Infalling description*—This is another consistent low-energy quantum mechanical description of the black hole that includes the exterior but also includes degrees of freedom in the interior of the black hole (though not all the way down to the singularity). This description does *not* include the stretched horizon. This is the description seen by observers falling into the black hole.

Note that while we have used the fiducial and infalling observers as ostensive devices to generate these two descriptions, these two descriptions can be specified and understood without any appeal to observers. (You can just delete the last sentence in each of the definitions above if you don't like talk of observers.)

With these two descriptions in place, we can now state the descriptive principle of complementarity.

> **Descriptive complementarity**: The exterior and infalling descriptions are *descriptively* consistent (as opposed to just operationally consistent) with each other and with quantum mechanics.

The principle states that both descriptions can simultaneously be accurate representations of the way the world is while being consistent with quantum mechanics. While operational complementarity only required that there be no possible experimental way for the inconsistency to become salient to a single observer, descriptive complementarity takes the descriptions as representations of the world and asks if both could simultaneously be true.

Descriptive complementarity can also be found in the literature phrased in terms of *holography*[18]. On one common reading of holography, the two descriptions are just two different descriptions of the same underlying physics; consequently, the two descriptions are consistent. Thus, holography, read this way, is a stronger claim than descriptive complementarity. For it is at least in principle possible for two descriptions to be consistent without them being different descriptions of the same state of affairs.

Note that descriptive complementarity is *not* saying that, in manner similar to Bohrian complementarity about quantum mechanics, we are only allowed to use one of the descriptions depending on one's context, i.e., depending on whether one is hovering or one is infalling. This is because the restriction on the use of a description from a given context is

[18] See, e.g., (Raju, 2020, pp. 59-67) and references therein.

still an *operational* restriction since it relies on saying what kind of observer one is. Further, this kind of Bohrian move does not take the descriptions representationally seriously, much like in the case of quantum mechanics.

### *3.2 Complementarity as consistency*

Both my complementarity principles are *consistency* claims. These are what *I* am referring to as *complementarity*. Several writers in the literature include under the label "black hole complementarity" the claim that there *are* the two descriptions mentioned above, or the claim that there is a stretched horizon. In this they follow the seminal paper of Susskind et al. (1993), in which the existence of an exterior description with the stretched horizon as the boundary is the import of their three "postulates of black hole complementarity".

I prefer to reserve the term "black hole complementarity" for the consistency conditions that I have delineated. I do so because one does not need to appeal to an extra principle to argue for the existence of these descriptions or the stretched horizon. That much follows from combining low-energy field theory and general relativity. The claim that the two descriptions are consistent—operationally or descriptively—*does* amount to a further claim that is worth elevating to a principle and exploring its consequences, because what is precisely at stake here is whether quantum mechanics in the vicinity of a black hole is consistent.

In the literature, the claim about the consistency of the descriptions is often left implicit; moreover, it is not always made clear whether the consistency of these two descriptions is being judged on operational grounds or on descriptive grounds. Thus, by reserving "complementarity" for the consistency claim, and by distinguishing between operational and descriptive consistency, I hope to have at least added some clarity to the discussion.

To the best of my knowledge, there has not yet been an explicit delineation between operational and descriptive principles of complementarity. The handful of philosophy-of-physics treatments of this topic ((Belot et al., 1999, pp. 211-216), van Dongen and de Haro (2004), and Bokulich (2005)) have been sensitive to the operationalism implicit in black hole complementarity. However, they view operational complementarity as being useful only insofar as it is a starting point for efforts to de-operationalize it. Even if one holds this view, it is worth appreciating the breadth of the principle's applicability and the ways in which it has been employed in the recent literature.[19]

As we shall see with the examples to follow, distinguishing the two principles will make clear the limits of descriptive complementarity, and the power of operational complementarity. I will show that the descriptive principle fails when it encounters some recent examples, while the operational principle succeeds. But first we will consider an example where both principles succeed.

## 4 What is the fate of the horizon crosser?

Alice is an infalling observer who crosses the horizon and Bob is an observer hovering outside the black hole. Let us look at the experience of Alice as she crosses the horizon. From her perspective, the experience is very smooth. She cannot distinguish her observations from that of traversing empty space as long as she considers physics at scales that are small compared to the local curvature. However, from the perspective of Bob, Alice will be thermalized as soon as she encounters the stretched horizon.[20]

While this might be counterintuitive, there is nothing mathematically inconsistent here. There are just these two different descriptions that are valid here: we can map one description to the other via a standard coordinate change from Schwarzschild to infalling coordi-

[19]Note that (Wallace, 2017, pp. 21-22) also has a brief discussion of complementarity. He formulates complementarity in the descriptive register and does not discuss operational versions of complementarity.

[20]This case from Susskind et al. (1993). See also (Wallace, 2017, pp. 21-22).

nates. This can be seen classically as well: in Schwarzschild coordinates Alice will appear to be getting closer and closer to horizon forever, while in infalling coordinates (such as Gullstrand-Painleve), she will smoothly cross the horizon. And so the infalling description is just a *redescription* of the exterior description, and consequently the two descriptions are descriptively consistent. Thus, descriptive complementarity holds in this case.[21]

What about operational complementarity? Can we point to some *observable* problem here? You might think Alice can provide evidence to Bob that she was not thermalized at the stretched horizon. If she sends a signal to Bob saying that she is fine after she crosses the stretched horizon, then Bob will know that something is wrong about his model of the black hole. For he would then both see Alice thermalized at the stretched horzion but also have confirmation that Alice safely crossed the stretched horizon.

The trouble with this suggestion, of course, is that once Alice crosses the true horizon, she will be unable to send any signals out. Thus, she has to send a signal to Bob when she is between the stretched horizon and the true horizon. Because the stretched horizon is only one Planck length above the true horizon, she has to encode and send her message in field modes of Planck-length frequencies (if it even makes sense to talk about "field modes" at those frequencies). If she does not send it at those frequencies, it will be too late: she would have crossed the horizon before the message gets out. Conversely, if Bob wants to probe what is happening to Alice as she crosses the stretched horizon, he would have to send in modes of Planck-scale frequencies to be able to resolve what is happening.

As far as this case goes, then, both descriptive complementarity and operational complementarity succeed because neither are we able to locate a descriptive contradiction nor an operational contradiction. However, as we turn to more involved situations that have

[21]It might be questioned whether the existence of a mathematical mapping between the two descriptions suffices for physical equivalence. However, this is a common assumption made in this literature, and I shall take it in on board here.

appeared in the recent literature, the value of operational complementarity becomes clear.

## 5 Quantum cloning?

Consider some quantum information that is encoded in infalling matter or radiation. Now, from the perspective of infalling observers, this information uneventfully crosses the horizon and continues on towards the singularity. However, this information will be accessible to exterior observers in the radiation after the Page time but before the black hole finishes evaporating (see Sec. 2). This seems to imply that the quantum information that was present in the infalling matter has been *cloned* at two different locations: in the interior of the black hole and in the radiation coming out of the black hole. This looks like a violation of the no-cloning theorem of quantum mechanics, which states that there is no unitary transformation (indeed, no linear transformation) that can create a copy of an arbitrary quantum state.[22] In this particular case, the no-cloning theorem says that there cannot be a linear transformation connecting the slice containing the quantum information encoded in the infalling matter *before* it fell in (this slice could be defined even before the formation of the black hole) and the slice that contains (unitary transformations of) both the quantum state of the fallen-in matter and the same quantum state coming out in the Hawking radiation (see Fig. 3). However, it seems as if this must indeed be the case if we believe that the black hole evaporation is unitary.

On this scenario, straightforwardly and immediately, descriptive complementarity fails. That is, if we take both the exterior and infalling descriptions representationally seriously, we have an inconsistency with quantum mechanics. After all, the exterior description will say that Alice's clone exists in the exterior and the infalling description will say that Alice

[22] This possibility was first considered in Susskind and Thorlacius (1994) and significantly sharpened by Hayden and Preskill (2007). (See, e.g., (Nielsen and Chuang, 2010, pp. 24-25) for review of the no-cloning theorem).

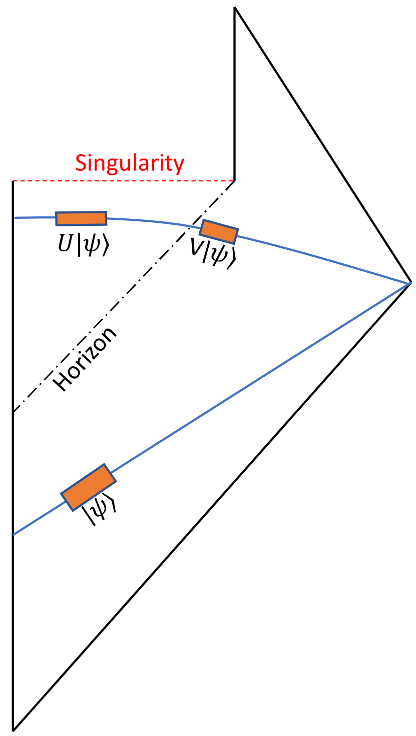


**Figure 3:** Penrose diagram depicting that the quantum information in the infalling matter $|\psi\rangle$ is cloned. That is, unitary transformations of it, $U|\psi\rangle$ and $V|\psi\rangle$, are generated in the interior *and* in the Hawking radiation in the exterior.

exists in the interior. And these two clones can be located on a slice that's smooth across the horizon. Consequently, if both descriptions are simultaneously valid, we have, upon radiation of the relevant degrees of freedom, a quantum cloning process.

The situation, as yet, seems consistent with operational complementarity because no single observer can compare the two systems that are supposed to contain the two copies and see the violation of the no-cloning theorem because one system is behind the horizon while the other is not. However, as Susskind and Thorlacius (1994) have pointed out, a single observer *might* be able to observe a violation of quantum mechanics if the exterior

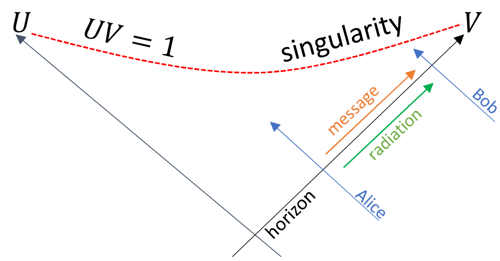


**Figure 4:** $U$ and $V$ are Kruskal coordinates. Alice has to send her information to Bob before Bob hits the singularity. (Adapted from (Hayden and Preskill, 2007, Fig. 2).)

observer waits until the infalling information comes out as Hawking radiation, and then jumps into the black hole and compares their record with the interior record. To see more clearly how this might work, let us return to Alice and Bob: Alice is the infalling observer and Bob is the exterior observer. Let us say that Alice carries with her a quantum bit (or a qubit) as she falls into the black hole. Bob waits, patiently collecting Hawking radiation, until, past the Page time, the information that Alice carried in reappears in the Hawking radiation. He then jumps into the black hole. Alice then sends her qubit to Bob (whom she knows will jump in) via a photon. *If* Bob can intercept Alice's photon and compare it with the information that he recovered from the Hawking radiation, then it looks as if he should be able to directly see that a quantum state has been cloned, and thus directly detect a violation of the no-cloning theorem.[23]

However, if operational complementarity is right, then there will be an operational barrier to conducting this experiment. And indeed, there is. For this proposal to work, it has to be the case Bob can receive the message from Alice before he crashes into the singularity. The longer Bob waits to jump in, the shorter the time Alice has after she crosses the horizon to send her message to Bob if he is to be able to recieve it before he hits the singularity

[23]This can be done, for instance, if they pre-decide that Alice will carry the quantum state $|\psi\rangle$, and then all Bob has to do is make a projective measurement with projectors $\{|\psi\rangle\langle\psi|, \mathbb{I}-|\psi\rangle\langle\psi|\}$.

(see Fig. 4). This is the source of the operational barrier. More precisely, Susskind and Thorlacius (1994) argued that Alice only has an extremely short time to send her message: she must send the information within a time-scale that is exponentially small in the square of the black hole mass if the message is to reach Bob before he encounters the singularity. This would then require her to encode her information in a signal of frequency that is exponentially large in the square of the black hole mass (in Planck units). Thus, unless Alice has access to Planck-scale frequencies, she cannot send a message to Bob that will allow him to verify a violation of no-cloning. Thus, operational complementarity is vindicated.

Or so it seems, for the plot thickens. One might think that operational complementarity only *seems* to be vindicated because we were not clever enough. For there is a significant weakness in Susskind and Thorlacius (1994)'s argument: They assumed that Bob would have to wait on the order of the Page time before the information that Alice possessed would come back out as Hawking radiation. They needed to assume this because Page's theorem says that the Hawking radiation before the Page time is almost indistinguishable from thermal radiation, and hence will not contain sufficient detail to allow Bob to extract Alice's qubit. And the Page time is very long: it is of the order of the black hole evaporation time. However, Hayden and Preskill (2007) showed that, if information falls in *after* the Page time, then black holes re-emit the information that falls in very quickly. Indeed, they showed that such information comes back out on a time-scale of the order of $M \log M$ (in Planck units). For solar mass black holes, this is on the order of $10^{-4}$ seconds![24] This is extremely short compared to the evaporation time or the Page time of such black holes, which is of the order of $M^3$, which, for solar mass black holes, is of the order of $10^{63}$

[24]To obtain this number, we can put the constants back in. So $M \log M$ becomes $\frac{GM}{c^3} \log\left(\sqrt{\frac{G}{\hbar c}}\right)$, which we can then compute.

years.[25] Thus, Hayden and Preskill call old black holes—i.e., black holes past their Page time—*information mirrors*. Because after the Page time, any information you throw in comes right back to you very quickly.

This seems to restore hope that there might indeed be a way for Bob to directly observe cloning. If Alice jumps into a black hole *after* the Page time, then, as argued above, Bob receives the information in Hawking radiation soon after she jumps in; following which, he jumps in as well. What is striking is that Hayden and Preskill show that even in this most forgiving scenario, Alice still does not have enough time to send the signal after her horizon-crossing so that it reaches Bob before he hits the singularity. They argued that if Alice sends the signal after a time that is longer than $O(M \log M)$, then it would be too late for Bob to receive the signal. But that is precisely the time-scale for which Bob has to wait before he has to jump in if he wants to recover Alice's information from the radiation! More carefully, the time difference between when Alice has to send the message and when Bob has to jump in is of the order of the Planck time, meaning that Alice has to encode her qubit in Planck-scale modes, preventing Bob from seeing the clone. Thus we see that operational complementarity really is vindicated.

Thus, the case quantum cloning provides a clear example wherein descriptive complementarity fails while operational complementarity succeeds. So we should lean more towards accepting operational complementarity, even if we are uncomfortable with operational principles in our physics.

Before we move on, it's worth pausing here a moment to consider why the problem of cloning discussed in this section could not have been raised in the scenario we considered in the previous section: the scenario of the horizon crosser. In that case too, doesn't the information in the degrees of freedom constituting Alice get encoded in the stretched horizon

[25] As above, $M^3$ becomes $\frac{G^2 M^3}{\hbar c^4}$.

after she crosses the horizon, resulting in clones located in the interior and on the stretched horizon? But the descriptive complementarist can respond that the description of Alice as smeared on the horizon is not part of the exterior description for it is not a low-energy description. More precisely, we don't really know if we can describe the way in which Alice is encoded on the stretched horizon as a quantum field state localized on a certain spacelike slice. At the Planck scale, we may not even be able to distinguish between spacetime and field-states on spacetime. However, once Alice's information appears in the radiation, then a low-energy description becomes available, allowing us to identify a problem for descriptive complementarity. For similar reasons, the operational complementarist will be untroubled by the appearance of cloning in this scenario: the clone on the stretched horizon is only empirically accessible to observers with the ability to manipulate Planck scale degrees of freedom.

## 6 A violation entanglement monogamy?

Let's consider another potential experiment that suggests that one might be able to set up a violation of quantum mechanics observable by a single observer. This is the famous AMPS paradox (named after its discoverers Almhieri, Marolf, Polchinski, and Sully (Almheiri et al. (2013)); this is sometimes also called "the firewall paradox"). Suppose, again, that we have unitary evaporation. This then means that the early Hawking radiation is going to be near-maximally entangled with the late (i.e., post-Page-time) Hawking radiation. This has to be the case because the inclusion of the late radiation starts purifying the total state of the radiation, as seen in the Page curve. Now do the following. Collect all the early radiation until after the Page-time. Then go close to the horizon and collect the radiation that ought to be near-maximally entangled with all the early radiation, as predicted by the Page curve. Given the large amount of entanglement between the early and the late radiation, we should be able to distill, from all the radiation that we have collected, a quantum state that is close to a pure state. However, we also expect that the modes near the horizon—i.e., the late-time radiation that we just collected—will be highly entangled with modes just behind the horizon, since Hawking radiation arises from tracing over the highly entangled vacuum of a relativistic quantum field theory. However, we know from the principle of monogamy of entanglement that the same quantum system cannot be highly entangled with two different systems.[26] So it seems that an observer could observe a violation of the monogamy of entanglement by distilling a large amount of the entanglement between the early and late radiation into a pure state, and then crossing the horizon and checking if the resultant state is still entangled with modes behind the horizon. Thus, it seems as if we can, in principle, have a single observer observe a violation of quantum mechanics.

This thought experiment results in a failure of descriptive complementarity. We have large amounts of entanglement between the late radiation and the early radiation, while simultaneously also having large amounts of entanglement between the late radiation with interior modes. All these three quantum systems can be located on a single spacelike slice that smoothly traverses the horizon. Thus, we have a violation of monogamy on one spacelike slice. This means the exterior and infalling descriptions taken together contradict quantum mechanics. Thus, descriptive complementarity is false in this scenario.

Operational complementarity continues to succeed. Harlow and Hayden (2013) have argued that if an observer attempts to perform the AMPS experiment, they will fail because the task of distilling the entanglement between the early and the late radiation will almost certainly take much longer than the evaporation time of the black hole, thus destroying any modes behind the horizon that would allow us to observe a violation of monogamy. The argument for this is based on computational complexity theory. A strengthened version of

[26] See, e.g., (Horodecki et al., 2009, p. 917).

Harlow and Hayden's argument was given by Aaronson,[27] who has shown that if the task of distilling the Hawking radiation—the so-called Hawking distillation problem—could be performed efficiently—i.e., in a time that is polynomial in the entropy of black hole—then a complexity-theoretic conjecture that is widely believed to be true, and widely employed in the security proofs for cryptographic protocols, would be false.[28] Therefore, our observer cannot see a direct violation of entanglement monogamy, for the black hole would have finished evaporating before they are ready to jump in and compare their distilled state with the modes behind the horizon. (They could try to collect and manage all the physical processes that occur during the final phase of evaporation, but that would require the observer to access Planck-scale physics, for in the final stages of evaporation, the horizon intersects with the singularity.)

(It's worth remarking here that the fact that the above result relies on a conjecture is no mark against it. Computational complexity[29] is a very successful field, with important practical applications, but much of it relies on conjectures—the most of famous of which is the conjecture that $P \neq NP$—that while unproven, are widely believed to be true, and hence frequently relied upon in proofs. See Aaronson (2016b) for a review of progress on and barriers to proving the $P \neq NP$ conjecture, and a defense of why one ought to think it true. See Williams (2019) for a discussion of many of the conjectures in the field and their plausibility.)

So we see that operational complementarity is secure. We are unable to identify an operational contradiction in this scenario. Attempting to directly observe the violation of monogamy is foiled by computational complexity. The fact that the barrier is *computational*

[27] Aaronson has not published this argument; see (Harlow, 2016, pp. 48-49) for a version in print.

[28] See Kim et al. (2020) for an important recent strengthening of the result.

[29] See, e.g., Arora and Barak (2009) for an introduction.

highlights the value of a truly operational principle here.

### *6.1 The dialectic following Harlow and Hayden*

Harlow and Hayden's result has been influential in the field. It is instructive examine some of the discourse that has followed their result. Oppenheim and Unruh (2014) argued that one can avoid the computational limitation identified by Harlow and Hayden by *precomputing* the distillation. On their proposal, one takes a large number of Bell pairs; performs a certain precomputation on one half of these pairs; collapses the other half of these pairs into a black hole; and then the collects the Hawking radiation from the resultant black hole. Their precomputation allows one to perform the Hawking distillation task before the black hole finishes evaporating. This suggests that one might be able to directly verify the violation of entanglement monogamy after all, and hence falsify operational complementarity.

However, Oppenheim and Unruh's arguments have received pushback: on the grounds that the kinds of states required by them will be difficult to construct and maintain in the vicinity of a black hole or lead to otherwise unphysical results (Harlow (2016); Kim et al. (2020)) and on the grounds that their proposal cannot apply to astrophysical black holes (Aaronson (2016a); Kim et al. (2020)).

Others have suggested that even if one is able to avoid the Harlow and Hayden's computational barrier to verifying the violation of monogamy, other barriers remain. For instance, Bao et al. (2016) argue, using some weak added assumptions, that even if Bob is able to distill the entanglement in time and jump in, he will not have enough time to verify the entanglement in the interior before he crashes into the singularity. A different barrier is suggested by Yoshida (2019), who argues that the gravitational backreaction due to Bob crossing the horizon will disentangle the late radiation and the early radiation.

So operational complementarity has not yet been refuted. Of course, it has by no means achieved universal acceptance, and it could still turn out that someone comes up with a way of obtaining a single-observer violation of quantum mechanics in the context of quantum black holes which doesn't face any barriers. Nevertheless, given the various proposals and counter-proposals and counter-counter-proposals that are looking for, or attempting to refute, *operational* barriers to verifying violations of quantum mechanics, it is clear that operational complementarity is crucial in guiding this discussion. Thus, give the literature so far, operational complementarity remains a plausible and powerful principle.

## 7 Lessons

An operational formulation of black hole complementarity has been essential in the recent literature surrounding the black hole information paradox. This literature has showed that attempts to extract observable violations of quantum mechanics out of the paradox fail, as long as we restrict to above-Planck-scale physics. Very promising proposals to generate observable violations of quantum mechanics have been thwarted for subtle reasons. Meanwhile a descriptive version of complementarity is unsuccessful: the exterior and infalling descriptions taken simultaneously results in violations of quantum mechanics.

So where do we go from here? For scientific realists, the failure of descriptive complementarity is perhaps none too surprising. After all, the black hole information paradox identified an *inconsistency* in the application of quantum mechanics to a black hole. Given an inconsistency, it is no suprise that the inconsistency reappears in different guises in different thought experiments. You can run but you can't hide.

Nevertheless, realists ought to be surprised by the success of operational complementarity. The fact that the inconsistency cannot be ramified up to an experimental problem likely signals something about the deeper descriptive theory that would resolve the information paradox. This is analogous to how, in the case of special relativity, the inability of observers to agree on which events are simultaneous signals the geometry of Minkowski spacetime. Or how, in quantum mechanics, the inability of observers to simultaneously measure precise values of position and momentum signals the nature of the wavefunction. Thus, even realists must take seriously the success of operational complementarity for it provides both clues towards, and new explananda for, future physics.

On the other hand, for those with no objection to operationalism in physics, the success of operational complementarity indicates the black hole information paradox has been resolved. For what is a paradox? It's a compelling argument to an absurd conclusion. If one is operationalist, then the absurdity of the conclusion has to be cashed out in operational terms. The success of operational complementarity suggests that there is no operational absurdity arising from the black hole information paradox—no experiment can be done to bring out the contradiction. Consequently, for operationalists, as long as future work doesn't invalidate operational black hole complementarity, there's no paradox left.

*Acknowledgments*—I am indebted to David Wallace for several enlightening discussions and comprehensive feedback on drafts. I am grateful to Ignacio J. Araya for many discussions on the physics of black holes. Thanks to Sam Fletcher, Bixin Guo, John Norton, and especially Nick Huggett, for helpful comments on drafts and for valuable discussions.

## References

Aaronson, S. (2016a). The complexity of quantum states and transformations: from quantum money to black holes. *arXiv preprint arXiv:1607.05256*.

Aaronson, S. (2016b). *P ?= NP*, pp. 1–122. Cham: Springer International Publishing.

Almheiri, A., D. Marolf, J. Polchinski, and J. Sully (2013). Black holes: complementarity or firewalls? *Journal of High Energy Physics 2013*(2), 62.

Arora, S. and B. Barak (2009). *Computational complexity: a modern approach*. Cambridge University Press.

Banks, T. (1995). Lectures on black holes and information loss. *Nuclear Physics B - Proceedings Supplements 41*(1), 21–65.

Bao, N., A. Bouland, A. Chatwin-Davies, J. Pollack, and H. Yuen (2016). Rescuing complementarity with little drama. *Journal of High Energy Physics 2016*(12), 1–21.

Bell, J. (1990). Against 'measurement'. *Physics world 3*(8), 33.

Belot, G., J. Earman, and L. Ruetsche (1999). The Hawking information loss paradox: The anatomy of a controversy. *British Journal for the Philosophy of Science 50*(2), 189–229.

Bokulich, P. (2005). Does black hole complementarity answer Hawking's information loss paradox? *Philosophy of Science 72*(5), 1336–1349.

Bourget, D. and D. Chalmers (2020). Philosophers on philosophy: The 2020 philpapers survey.

Bousso, R. (2013, Jun). Complementarity is not enough. *Phys. Rev. D 87*, 124023.

Candelas, P. (1980, Apr). Vacuum polarization in Schwarzschild spacetime. *Phys. Rev. D 21*, 2185–2202.

Chang, H. (2021). Operationalism. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2021 ed.). Metaphysics Research Lab, Stanford University.

Chen, P., Y. Ong, and D. h. Yeom (2015). Black hole remnants and the information loss paradox. *Physics Reports 603*, 1–45.

Geroch, R. (1970). Domain of dependence. *Journal of Mathematical Physics 11*(2), 437–449.

Harlow, D. (2016, Feb). Jerusalem lectures on black holes and quantum information. *Rev. Mod. Phys. 88*, 015002.

Harlow, D. and P. Hayden (2013). Quantum computation vs. firewalls. *Journal of High Energy Physics 2013*(6), 85.

Hawking, S. and J. Stewart (1993). Naked and thunderbolt singularities in black hole evaporation. *Nuclear Physics B 400*(1), 393–415.

Hawking, S. W. (1975). Particle creation by black holes. *Communications in mathematical physics 43*(3), 199–220.

Hawking, S. W. (1976, Nov). Breakdown of predictability in gravitational collapse. *Phys. Rev. D 14*, 2460–2473.

Hayden, P. and J. Preskill (2007, Sep). Black holes as mirrors: quantum information in random subsystems. *Journal of High Energy Physics 2007*(09), 120–120.

Horodecki, R., P. Horodecki, M. Horodecki, and K. Horodecki (2009, Jun). Quantum entanglement. *Rev. Mod. Phys. 81*, 865–942.

Kim, I., E. Tang, and J. Preskill (2020). The ghost in the radiation: Robust encodings of the black hole interior. *Journal of High Energy Physics 2020*(2003.05451), 1–65.

Lowe, D. A., J. Polchinski, L. Susskind, L. Thorlacius, and J. Uglum (1995, Dec). Black hole complementarity versus locality. *Phys. Rev. D 52*, 6997–7010.

Marolf, D. (2017). The black hole information problem: past, present, and future. *Reports on Progress in Physics 80*(9), 092001.

Mathur, S. D. (2009, Oct). The information paradox: a pedagogical introduction. *Classical and Quantum Gravity 26*(22), 224001.

Maudlin, T. (2017). (Information) paradox lost. *arXiv preprint arXiv:1705.03541*.

Nielsen, M. A. and I. Chuang (2010). *Quantum computation and quantum information*. Cambridge University Press.

Nomura, Y., J. Varela, and S. J. Weinberg (2013). Complementarity endures: no firewall for an infalling observer. *Journal of High Energy Physics 2013*(3), 59.

Oppenheim, J. and B. Unruh (2014). Firewalls and flat mirrors: An alternative to the AMPS experiment which evades the Harlow-Hayden obstacle. *Journal of High Energy Physics 2014*(3), 120.

Page, D. N. (1993, Dec). Information in black hole radiation. *Phys. Rev. Lett. 71*, 3743–3746.

Polchinski, J. (2015). *The Black Hole Information Problem*, Chapter 6, pp. 353–397.

Price, R. H. and K. S. Thorne (1988). The membrane paradigm for black holes. *Scientific American 258*(4), 69–77.

Quine, W. V. (1962). Paradox. *Scientific American 206*(4), 84–99.

Quine, W. V. O. (1951). Two dogmas of empiricism. *Philosophical Review 60*(1), 20–43.

Raju, S. (2020). Lessons from the information paradox. *arXiv preprint arXiv:2012.05770*.

Sciama, D. W., P. Candelas, and D. Deutsch (1981). Quantum field theory, horizons and thermodynamics. *Advances in Physics 30*(3), 327–366.

Strominger, A. (1995). Les Houches lectures on black holes. *arXiv preprint hep-th/9501071*.

Susskind, L. and J. Lindesay (2005). *An introduction to black holes, information and the string theory revolution: The holographic universe*. World Scientific.

Susskind, L. and L. Thorlacius (1994, Jan). Gedanken experiments involving black holes. *Phys. Rev. D 49*, 966–974.

Susskind, L., L. Thorlacius, and J. Uglum (1993, Oct). The stretched horizon and black hole complementarity. *Phys. Rev. D 48*, 3743–3761.

't Hooft, G. (1985). On the quantum structure of a black hole. *Nuclear Physics B 256*, 727–745.

Thorne, K. S., K. S. Thorne, R. H. Price, and D. A. MacDonald (1986). *Black holes: the membrane paradigm*. Yale University Press.

Unruh, W. G. and R. M. Wald (2017, Jul). Information loss. *Reports on Progress in Physics 80*(9), 092002.

van Dongen, J. and S. de Haro (2004). On black hole complementarity. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 35*(3), 509–525.

Vickers, P. (2013). *Understanding inconsistent science*. Oxford University Press.

Wallace, D. (2017). Why black hole information loss is paradoxical. *arXiv preprint arXiv:1710.03783*.

Wallace, D. (2018a). The case for black hole thermodynamics part I: Phenomenological thermodynamics. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 64*, 52–67.

Wallace, D. (2018b). The case for black hole thermodynamics, part II: Statistical mechanics. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 66*, 103–117.

Williams, R. R. (2019). *Some Estimated Likelihoods for Computational Complexity*, pp. 9–26. Cham: Springer International Publishing.

Yoshida, B. (2019). Firewalls vs. scrambling. *Journal of High Energy Physics 2019*(10), 1–50.