The Black Hole Information Paradox

Stefano Antonini, John Martyn, Gautam Nambiar

December 14, 2018

1 Introduction and motivation

The 1970's saw the emergence of many revolutionary ideas in black hole physics. In 1973, Jacob Bekenstein predicted that a black hole has a finite entropy that scales as the area of its event horizon. Moreover, in 1974, Stephen Hawking confirmed this result by proving, within the framework of quantum field theory in curved spacetime, that black holes emit thermal radiation, ultimately leading to their "evaporation". It is likely, at least at a first glance, that this phenomenon violates a fundamental property of quantum mechanics: unitarity. In short, unitarity means that the (quantum) information stored in a closed system is preserved over time, akin to the conservation of energy in classical mechanics. The violation of unitarity poses many problems to our understanding of quantum mechanics and quantum information theory.

Indeed, Hawking's argument suggests that part of the information associated with the matter falling into a black hole (or that present in it since its formation) is destroyed. The thermal radiation carries information only about the temperature of the body emitting it (the black hole), i.e. about a mixed state. Hence, the thermal radiation emitted by a black hole as it evaporates does not contain any information about the microscopic state of the system. For instance, if a particle in a pure state falls into a black hole, the ensuing radiation emitted will be associated with a mixed state, which does not provide direct information about the original pure state. This is the essence of the black hole information paradox (BHIP): unlike any other classical or quantum system, black holes may not conserve information, thus violating unitarity.

We expect that the resolution to this paradox, or its confirmation, lies in a full quantum theory of gravity, which has yet to be developed and remains one of the most challenging issues of contemporary physics. This theory will supersede both general relativity and quantum mechanics on scales on the order of the Planck length. Therefore, as a black hole evaporates and decreases in size, this theory should govern the nature of the black hole when its size is comparable to the Planck length and describe what happens to the information that entered the black hole.

In spite of the lack of a consistent theory of quantum gravity, many hypotheses about the information paradox have been formulated. Some physicists speculate that quantum gravity may actually be non-unitary, while others uphold unitarity and develop solutions to the BHIP through the framework of string theory. Neither of these proposals are conclusive so far. A powerful tool to investigate the black hole thermodynamics relevant to the BHIP is provided by the AdS/CFT correspondence, which allows one to compute the entropy of a black hole in a (d+1)-dimensional spacetime by computing the entanglement entropy of a state of a conformal field theory living on the d-dimensional boundary of this spacetime. AdS/CFT correspondence is also an extremely useful tool for computing entanglement entropies in quantum systems by performing easier calculations in the bulk dual.

Relying on one of the proposed solutions to the paradox, one paper that we will study [1] assumes unitarity is preserved and demonstrates how the information that enters a black hole in the latter half of its evaporation process is almost immediately re-emitted through the Hawking radiation. In this case, a black hole can be viewed as a "mirror" for information.

In this work, we will present a survey of the BHIP and its connections to quantum information. Specifically, we will first review black hole thermodynamics and present the black hole information paradox, starting from Hawking's original arguments outlined in [2], and thereafter delving into other points of view of the paradox discussed in [3, 4]. These sources will be useful in describing how the paradox emerges and explaining its various implications. We will briefly discuss the relations between the BHIP and the AdS/CFT correspondence, which can be used to help to solve the BHIP, as described in [3, 5, 4]. Finally, we will analyze in detail how black holes can behave as information mirrors in one theorized resolution to the BHIP, as shown in [1].

2 Black hole thermodynamics and black hole evaporation

2.1 The thermodynamics of black holes

When we have a system composed of many degrees of freedom, there is room for disorder, and one can talk about its entropy. Furthermore, if we assume that the system is in equilibrium with a heat bath, (alternatively if the system acts as a heat bath for its subsystems), then we can define the temperature of the system. Can we then talk about the entropy or temperature of a black hole too? Classically, we wouldn't expect a black hole to be described by smaller degrees of freedom composing it. Roughly, we suspect this because anything that falls into the black hole is causally disconnected from the outside, and an observer on the outside cannot know that the black hole is composed of smaller parts. This would lead us to believe that the entropy of a black hole vanishes and that it doesn't make sense to talk about its temperature. However, even classically, there exist black hole phenomena which look surprisingly analogous to thermodynamics, in which temperature and entropy are well-defined. Here, we will give a qualitative description of one of them, the Penrose process. [2]

This is a classical process that enables one to extract energy out of a black hole. In this scenario, the symmetries in spacetime are described by Killing vectors, which are the directions along which the metric tensor does not change. One important property of a black hole is that, what is a time-like Killing vector outside the horizon becomes space-like inside. Now if we make a further assumption that the black hole is rotating, then this flipping of time-like and space-like vectors happens in a region outside the horizon called an ergoregion. If a particle splits into two and one of them enters this region, its energy (a time-like vector) becomes a component of spatial momentum, which can be negative. This energy deficit is carried away by the particle outside the ergoregion, which means we have extracted energy from the rotating black hole. This energy extracted comes at the cost of a decrease in the black hole's rotational energy, as the infalling particle has an angular momentum opposite to the black hole's. It turns out that the efficiency of this process (the ratio of energy extracted to the angular momentum loss) is maximized when *the area of the horizon remains constant.*[2]

This has a resemblance to the statement that the highest efficiency of a thermodynamic process is achieved when it is reversible, i.e. *when its entropy remains constant.* Further, there exists an Area Theorem proved by Hawking, which states that the area of an event horizon can never decrease. These analogies are made more concrete in the four laws of black hole mechanics, in which the area is the analogue of entropy, and surface gravity (which is the acceleration of a particle just outside the horizon with respect to the Killing time) is the analogue of temperature.

However, the analogy to thermodynamics still has some serious shortcomings, one of which being that it doesn't tell us how to make sense of the temperature of a classical black hole. Another problem is that area doesn't have the same dimensions as entropy. A solution to this would be to introduce a quantum mechanical length scale $l_p = \sqrt{\frac{\hbar G}{c^3}}$ while doing a counting of microstates. However, to actually perform this counting, one needs a quantum theory of gravity which we don't have yet. So, instead we will give an argument (from Chapter VII.3 of [6]) for what the temperature of the black hole should be.

A quantum system governed by a Hamiltonian H and in equilibrium with a heat bath of inverse temperature β is described by the density operator $\rho_{\text{canonical}} = \frac{e^{-\beta H}}{Z}$. Here Z is the partition function which equals $Z = \operatorname{Tr} e^{-\beta H} = \sum_n \langle n|e^{-\beta H}|n\rangle$. This is analogous to the amplitude of making a transition from state $|n\rangle$ back to itself in time t, which is $\langle n|e^{-iHt/\hbar}|n\rangle$. We notice that if we sum this amplitude over n and replace t by $-i\beta$, we get the partition function. Thus, the inverse temperature β is the recurrence period in the imaginary time (divided by $\hbar$). Now, a photon near the horizon experiences a spacetime described by the Schwarzschild metric

$$ds^2 = \left(1 - \frac{r_S}{r}\right) c^2 dt^2 - \left(1 - \frac{r_S}{r}\right)^{-1} dr^2 - r^2 d\theta^2 - r^2 \sin^2\theta d\phi^2$$

where r_S is the Schwarzschild radius $2GM/c^2$.

To avoid the coordinate singularity at $r = r_S$, we change variables to $\rho^2 = 4r_S(r - r_S)$ (this singularity is not a physical singularity; the only physical singularity is located at $r = 0$). Then, in the approximation that the photon is close to the horizon, the metric becomes $ds^2 = \frac{\rho^2}{4r_S^2}c^2dt^2 - d\rho^2 - r_S^2 d\Omega^2$, where Ω is the solid angle. We now use the trick of rotating to imaginary time by defining $t = -it_E$. Performing this manipulation, the metric appears Euclidean: $ds^2 = -\left(d\rho^2 + \rho^2\frac{c^2dt_E^2}{4r_S^2} + r_S^2d\Omega^2\right)$. Here, imaginary time takes the role of an angular coordinate, in analogy with Euclidean polar coordinates in 4 dimensions. We now impose a period of 2π on the coordinate $\frac{ct_E}{2r_S}$ to avoid a conical singularity. From this, we conclude that the temperature is $\frac{\hbar c}{4\pi r_S}$. We are thus led to guess that the black hole has a temperature T_H, called the Hawking temperature,

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

We can use this fact to give a rough argument of what the entropy of a black hole should be. From the thermodynamic law $dU = TdS$, applied to a black hole, we have that $d(Mc^2) = T_H dS$ (here we use $U = Mc^2$ because all of the energy of a Schwarzschild black hole comes from its mass). This gives the Bekenstein-Hawking entropy:

$$S = \frac{A}{4l_P^2} \tag{1}$$

where $l_p = \sqrt{\frac{\hbar G}{c^3}}$ is the Planck length and A is the surface area of the black hole horizon. We can argue that such an entropy may be somehow related to the microscopic description of the black hole. Nonetheless, we are unable to correctly identify these microscopic degrees of freedom due to the lack of a full theory of quantum gravity.

2.2 Hawking radiation and black hole evaporation

We have argued that black holes have an effective temperature, T_H. This brings one idea to mind: objects with finite temperature radiate energy. For instance, a hot stove glows red and emits blackbody radiation. Analogously, could a black hole with nonzero temperature radiate away energy, and "glow red"?

In the case of a blackbody at temperature T, it is of interest to analyze the average number of photons per mode (per frequency) emitted as radiation. This quantity is given by the Bose-Einstein distribution, with energy $E = \hbar\omega$ [7]:

$$\langle N(\omega)\rangle = \frac{1}{e^{\frac{\hbar\omega}{k_BT}} - 1}. \tag{2}$$

This relation quantifies the spectrum of the thermal radiation emitted by a blackbody.

If a black hole were to radiate energy, we would expect a similar distribution to describe the radiation it emits. This idea crossed Stephen Hawking's mind in the 1970's; in 1974, he showed that a black hole indeed radiates, and that the spectrum of this radiation is described by a blackbody distribution with temperature $T = T_H$. Explicitly, the spectrum of particles emitted by a black hole due to Hawking radiation is [3]

$$\langle N(\omega)\rangle = \frac{\Gamma}{e^{\frac{\hbar\omega}{k_BT_H}} - 1}. \tag{3}$$

In this equation Γ is the greybody factor, which quantifies the lower intensity of radiation emitted by the black hole [4] with respect to the radiation emitted by an ideal black body at temperature T_H. We present a derivation of this relation in Appendix A. This equation is nearly the same as equation (2); its interpretation is that the number of particles per mode seen by the asymptotic observer is given by a blackbody spectrum, with temperature T_H. We therefore conclude that an observer far away from the black hole perceives the black hole to emit thermal radiation akin to an object at temperature T_H.

Furthermore, one can demonstrate (see Appendix A) that black hole radiation is similar to pair production. In particular, for every outgoing particle of radiation, there is another infalling particle, traveling toward the singularity, which has negative energy. Such a pair of radiation particles is known as a Hawking pair. This process is not unphysical, because, as we have already stated, energy inside the horizon becomes a momentum. In fact, an analysis of this situation indicates that each photon belonging to the Hawking radiation outside the horizon is entangled with a corresponding particle inside the horizon. This property will be fundamental to understanding the black hole information problem.

A distant observer will thus view particles emitted from a black hole with nonzero energy. Since the black hole is the only source of mass-energy in this scenario, it must be that these particles carry energy away from the black hole, and its mass decreases as radiation is emitted. As demonstrated in [4], this is indeed the case: Hawking radiation causes black holes to lose energy and mass, ultimately resulting in their "evaporation", at which they cease to exist. For typical astrophysical black holes, this process is incredibly slow. It is often the case that $T_H \ll 1$, and so these black holes are quite cold. Over humanly time scales, the energy carried away from these black holes in their Hawking radiation is negligible compared to their total mass-energy. For instance, even the smallest astrophysical black holes will take at least a few billion years to evaporate, and larger black holes take even longer. Nevertheless, the fact that black holes do evaporate has profound implications on our interpretation of information and quantum physics.

3 The black hole information paradox

We are now ready to face the black hole information problem. We can gain intuition of the paradox by means of a thought experiment. Let us suppose that we have a pure state of n EPR pairs, and that the black hole is initially in a pure state. Since the state of the system is pure, its (Von Neumann) entanglement entropy, which quantifies information, vanishes. Then, suppose we throw one of the qubits of each pair into a black hole. Then the new entanglement entropy of the state outside the black hole (performed by tracing out the interior of black hole) will be $n \ln 2$ ($\ln 2$ for each pair), and the entanglement entropy of the black hole will be the same. After a period of time, the black hole evaporates and disappears, but the entanglement entropy of the external state is necessarily unchanged. Indeed, the qubits that had fallen behind the horizon of the black hole cannot influence the matter outside the black hole, which would result in a violation of causality. Therefore, in this thought experiment, we started with an initial entanglement entropy of 0, and at the end, we obtained a total entanglement entropy $S^f_{ent} = n \ln 2$. In other words, we started with a pure state and we ended up with a mixed state.

Such a behavior strongly contrasts the principles of ordinary quantum mechanics. In fact, the time evolution of an ordinary quantum system is represented by a unitary matrix U, and can be written in an arbitrary basis as:

$$\psi^f_m = U_{mn}\psi^i_n. \tag{4}$$

Since U is unitary, it maps pure states to pure states. On the other hand, in the thought experiment explained above, the time evolution of the black hole system can be described using a matrix \$ that can map the density matrix of a pure state to the density matrix of a mixed state:

$$\rho^f_{mm'} = \$_{mm',nn'}\rho^i_{nn'}. \tag{5}$$

In the particular case $\$_{mm',nn'} = U_{mn}U^*_{m'n'}$, we recover the unitary evolution of ordinary quantum mechanics.

The thought experiment previously outlined is useful for understanding the paradox, but it is not really necessary. In fact, the same result can be achieved by taking into account only Hawking radiation. Let us consider again a black hole that is initially in a pure state. As we have pointed out in Section 2.2 (see also Appendix A), when a Hawking pair is generated, it exists in a pure, entangled state. One of the modes, $\tilde{a}$, exists inside the horizon, while the other one, a, exists outside and contributes to Hawking radiation. In effect, this means that Hawking radiation is entangled with the black hole. However, since this "new part" of Hawking radiation is produced in a pure state $|\Psi_{a\tilde{a}}\rangle$, the state outside the horizon must have the form $|\Psi_a\rangle \otimes |\Psi_E\rangle$ where $|\Psi_E\rangle$ is the state describing all the previously emitted Hawking radiation. In other words, there cannot be entanglement between the newly emitted Hawking radiation and earlier emitted radiation. This implies that the entanglement entropy of the Hawking radiation always increases. This fact is quite intuitive: the earlier Hawking radiation is entangled with modes behind the horizon of the black hole. But those modes cannot exit the horizon and influence the state of the early Hawking radiation, since this process would violate causality. Therefore, we cannot "go back" to a pure state outside the horizon. This can be formalized by using the strong subadditivity of entanglement entropy, that we will not demonstrate, but which guarantees:

$$S_{a\tilde{a}} + S_{aE} \geq S_a + S_{a\tilde{a}E}. \tag{6}$$

It can be shown that ordinary subadditivity provides the opposite inequality in this scenario, and thus equality must hold in this relation. Since $|\Psi_{a\tilde{a}}\rangle$ is a pure state, then $S_{a\tilde{a}} = 0$, $S_{a\tilde{a}E} = S_E$, and we obtain the expected result:

$$S_{aE} = S_a + S_E. \tag{7}$$

Note that this behavior is very different from an ordinary system. For instance, let us consider a piece of burning coal and suppose that it is initially in a pure state. The photons first emitted can be thought as entangled with the particles still in the coal. Then we can imagine that, initially, the entanglement entropy increases linearly, just as in the black hole case. But when almost half of the coal is already burnt, we can think of the newly generated photons as encoding the information contained in the coal particles that are entangled with the radiation emitted earlier. Therefore, pure states will be "recomposed" in the emitted radiation, and entanglement entropy will start decreasing. When the coal has burnt up, we end up with a pure state and the entanglement entropy will be 0. This behavior of the entanglement entropy is described by the Page curve (see Figure 1). The difference between this system and a black hole is that the presence of the event horizon prevents the information stored in the modes behind the horizon from influencing the state of the Hawking radiation. Therefore, as we proved, the entanglement entropy of the Hawking radiation, i.e. of the outside of the black hole, follows the Hawking curve, increasing monotonically until the black hole disappears. If we want to recover a pure state in the end, an exotic, quantum gravity-driven process should allow for an extremely steep drop of the entanglement entropy in the last instants of the black hole evaporation, when the Hawking's argument is not valid anymore.

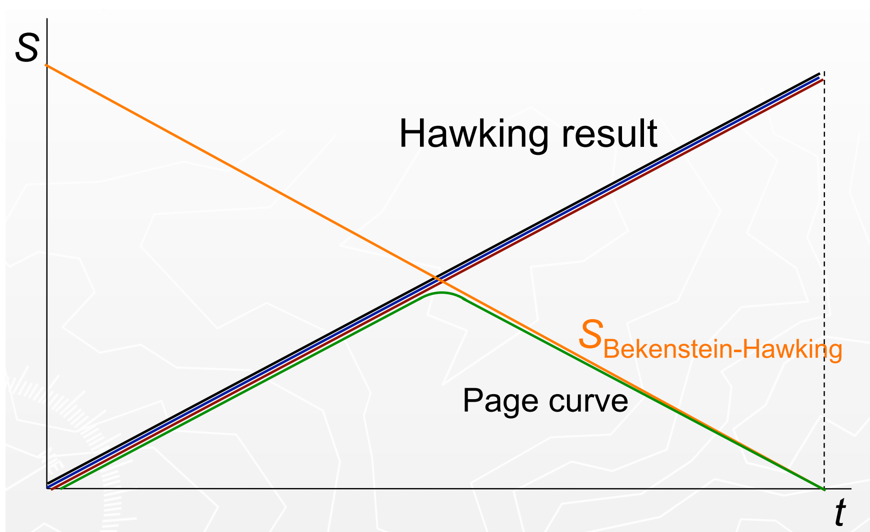

Figure 1: Hawking curve (Von Neumann entropy of the black hole and of the radiation, i.e. entanglement entropy of the two), Bekenstein-Hawking entropy and Page curve. Figure from [8].

From Figure 1, we note also that when the evaporation of the black hole is at the halfway point, the Bekenstein-Hawking entropy becomes smaller than the entanglement entropy of the black hole. This is another warning of a possible breakdown of ordinary physics. Indeed, the Bekenstein-Hawking entropy is a thermodynamic entropy, and can therefore be regarded as a "coarse-grained" version of the entanglement entropy [3]. Since entropy can be viewed as a count of the microstates associated to the black hole macrostate, logically the Bekenstein-Hawking entropy should be always larger than the entanglement entropy. Therefore, the situation after the middle point of the evaporation is clearly of difficult interpretation. We remark that, for the reasons we explained earlier, the Page curve follows the Hawking curve for the first half of the evaporation process, and then it drops following the Bekenstein-Hawking curve. At the end of the process, the difference between an ordinary system and a black hole is visually highlighted by the discrepancy between the Hawking curve and the Page curve.

The fact that entanglement entropy of the Hawking radiation always increases leads again to the paradox that we pointed out in our first thought experiment: when the evaporation process ends and the black hole disappears, we are left only with Hawking radiation, a mixed state which now represents the whole system. Therefore, we again started with a pure state and ended up with a mixed state: information has been lost during the evaporation process.

3.1 "Classical" possible explanation of the paradox

The dollar matrix introduced in the previous section, first proposed by Stephen Hawking, gives the first of three "classical" possible explanations of the black hole information paradox, probably the most shocking one: the time evolution of a black hole may violate unitarity. In essence, this means that quantum gravity may violate the principles of quantum mechanics and destroy information. This is a large issue since unitarity is fundamental to the construction of quantum circuits and the theory of quantum information. Additionally, it can be shown that time evolution governed by the dollar matrix would violate the conservation of energy. Even if quantum gravity effects may be visible only at the Planck scale, it seems unlikely that such violations are not detected at lower energies, and we have very strong experimental limits on these effects.

The second possible explanation of the black hole information paradox is that information is not lost, but it is somehow carried by Hawking radiation. Nonetheless, from the analysis we carried out earlier, it is clear that such a possibility requires superluminal transport of information.

Finally, the "remnant" hypothesis has also been proposed: when an evaporating black hole decreases to a size of the order of the Planck length, all of the missing information may still be stored into such a Planckian black hole. Indeed, when the horizon size is comparable to the Planck length, Hawking's analysis, which is based on the adiabatic hypothesis, is not reliable anymore. Thus, the evaporation could stop when the black hole reaches this length scale. Nonetheless, unless quantum gravity induces extremely exotic phenomena at the Planckian scale, it seems unlikely that all of the information contained in the matter fallen into an arbitrarily large black hole could be contained in an object of Planckian size.

Therefore, by approaching the problem from a "classical" perspective, as far as we know, there is no resolution to the information paradox that does not require a breakdown of our fundamental theories. Nevertheless, different approaches have enabled physicists to make progress on possible solutions to the black hole information problem. In particular, string theory and the AdS/CFT correspondence may give hints about the resolution of the paradox.

3.2 A solution without an explanation: AdS/CFT correspondence

In this section we will give a basic and qualitative understanding of how the information paradox seems to be avoided by means of the AdS/CFT correspondence. A detailed dissertation of this topic is extremely demanding both in terms of space and of difficulty, and it is beyond the purposes of this report. Nonetheless, we find it useful to roughly review the idea leading to the solution of the paradox, given that in the rest of the report we will assume it to be solved in order to explain the behavior of black holes as mirrors for information [1].

The dependence of the Bekenstein-Hawking entropy on the surface of the black hole instead of its volume gives a hint that gravity could be a holographic theory, i.e. that all the information needed to describe it can be localized on a hypersurface of one fewer dimension. Based on this idea and on similar reasoning arising from string theory, the AdS/CFT correspondence was formulated. This theory correlates a "bulk" $d+1$-dimensional gravitational theory describing an asymptotically AdS spacetime with a d-dimensional conformal field theory living on the boundary of such a universe, given by a spatial hypersphere S^{d-1} times a time interval. The two theories are dual, which means that every state of the CFT corresponds to a geometry of the bulk spacetime, and every observable of the CFT corresponds to an observable of the gravitational theory. A "dictionary" allows to find the desired quantity in one of the two theories by computing the corresponding quantity in the other theory. It is not clear what are the necessary conditions for a CFT to have a holographic dual gravitational theory, nor if the duality is complete, i.e. if the gravitational theory is completely defined by the dual CFT, and vice versa. Nonetheless, even if a rigorous proof of the latter question has not been found, it is generally considered true.

If we assume that the gravitational theory is completely defined by the dual CFT, then the solution of the information paradox is immediate. Indeed, the CFT is an ordinary quantum system, for which unitarity is preserved. If we consider a high energy state of the CFT, which is dual to a black hole in the bulk, the evolution of such a state will always be unitary. Therefore, the evolution of the black hole must be unitary as well, since it is completely determined by the dual evolution of the CFT state. It is remarkable how the AdS/CFT correspondence suggests that the black hole information paradox is solved and unitarity is safe, but it does not explain how this can happen in the gravitational side of the story. All we know is that the full quantum gravitational theory, whatever it is, must preserve unitarity because the dual CFT does it. It is a solution without an explanation.

This argument is certainly not satisfying, and its exposition is not detailed, but it is enough to understand that there are significant hints of the non-violation of unitarity in the black hole evaporation process.

3.3 Page Curve, if unitarity is preserved

If unitarity is somehow preserved, then the Page Curve does describe the entanglement entropy of the Hawking radiation *throughout* the process of evaporation (this time, just like burning coal). Before the halfway point of evaporation, the radiation R is maximally entangled with *a part* of the black hole B. The entanglement entropy $S(R) = S(\frac{1}{d_R}I)$, where d_R is the dimension of the Hilbert Space describing R. The information that can be obtained from the Hawking radiation, i.e. $\mathrm{Inf}(R) = S(\frac{1}{d_R}I) - S(R)$ is thus zero. However, after the halfway point, the black hole is maximally entangled with a part of the radiation. So, $S(R) = S(B) = S(\frac{1}{d_B}I)$. This time the information in the radiation

$$\begin{aligned}
\mathrm{Inf}(R) &= S(\frac{1}{d_R}I) - S(R) \\
&= S(\frac{1}{d_R}I) - S(\frac{1}{d_B}I) \\
&= \log d_R - \log d_B \\
&= 2\log d_R - \log(N)
\end{aligned}$$

where $N = d_B d_R =$ Dimension of the total Hilbert Space containing the black hole and its radiation.

This preliminary analysis suggests that for any information to be retrieved from the Hawking radiation, one must wait till the half-way point, and from this point onwards, the information increases to its maximum value $\log N$ once the evaporation is complete (no information loss). We will return to this result in the next section where, relying on some sort of solution of the black hole information paradox, we will assume unitarity to be preserved.

4 Black holes as mirrors – the issue of quantum cloning

Even if we assume that unitarity is preserved, we only trade one paradox for another – information loss with the violation of the no cloning theorem. To illustrate this, we consider the following scenario (Figure 2). Alice falls into a black hole with a quantum state $|\phi\rangle$. Meanwhile, Bob waits outside collecting the Hawking radiation. Since we assume that unitarity is preserved, if he waits long enough, he can extract $|\phi\rangle$ from the Hawking radiation by performing a unitary transformation. But from the perspective of Alice who is inside the black hole, the quantum state $|\phi\rangle$ never escapes the event horizon, because if it did, then causality would be violated. In fact, it is possible to show that there exists a space-like surface intersecting both Alice's $|\phi\rangle$ inside the black hole and the $|\phi\rangle$ obtained from the Hawking radiation. This means that we have created a copy of a quantum state, thus violating the no-cloning theorem.

A way out of this paradox is the Black Hole Complementarity hypothesis (BHC)[9]. According to this, cloning is unverifiable because one cannot have a simultaneous quantum mechanical description of both the interior and exterior of black holes. However, it is not obvious that this implies that cloning is unverifiable, because one can imagine an experiment (Figure 3) in which Bob waits for time Δt collecting Hawking radiation to recover Alice's state $|\phi\rangle$. He then jumps into the black hole, and Alice shoots a light beam at Bob which has the message required to verify cloning. If Bob can receive this message before hitting the singularity, then cloning is verifiable. If this experiment can succeed, then BHC is inconsistent. In their paper [1], John Preskill and Patrick Hayden tried to find if this experiment can succeed. We will briefly review their results here.

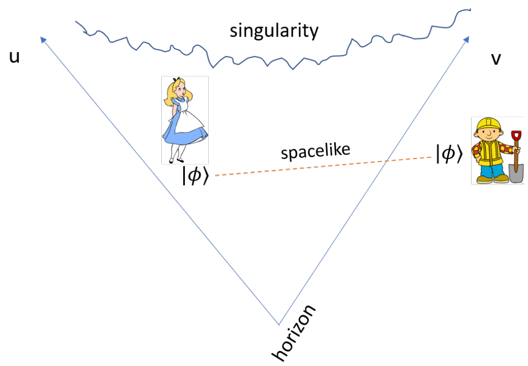

Figure 2: The cloning problem.

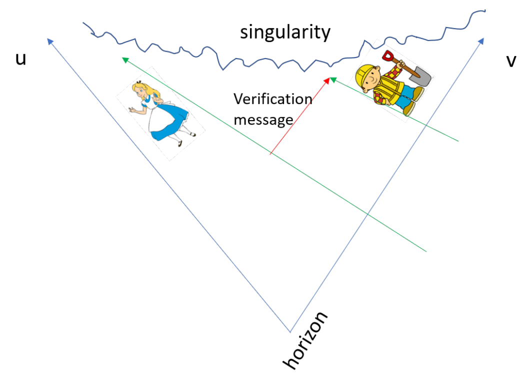

Figure 3: Experiment to verify cloning.

First of all, we see that if Bob waits for too long outside before jumping in, Alice's message will not reach Bob. It is a calculation in general relativity to find a bound on Δt so that he is just in time to receive Alice's message. This bound is $\Delta t \leq O(r_S \log r_S)$. Next, we ask how long Δt should be so that Bob can actually recover $|\phi\rangle$ from the Hawking Radiation. If BHC is consistent, then we want the experiment to fail, i.e. $\Delta t \geq r_S \log r_S$.

Looking at the Page curve discussed in Section 3, it seems that for any information to be available at all from the Hawking radiation, we need to wait for half the black hole to evaporate, i.e. Δt is of the order of the lifetime of the black hole, which goes as $O(r_S^3)$ and is enough to protect BHC. However, by posing the question more sharply, Preskill and Hayden found that in some scenarios, the relevant time-scale can shift from the black-hole lifetime to the black-hole thermalization time.

4.1 Black Hole as a quantum randomizer

Say Alice drops a k-qubit message into a *newly* formed $n-k$-qubit black hole which is in a pure state. In our model, we view the black-hole as an object that performs a random unitary transformation V on the n-qubit system, and then releases qubits one by one as Hawking radiation (see Figure 4). Waiting outside, Bob collects s qubits. Bob knows the initial state of the black-hole (somehow) and the unitary V. By performing a unitary operation on the s-qubit radiation system, Bob should be able to recover $|\phi\rangle$ as a k-qubit subsystem.

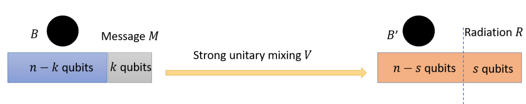

Figure 4: The quantum randomizer model.

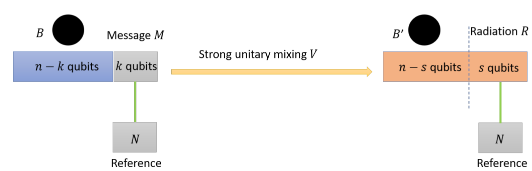

Figure 5: Restating it in terms of a reference N.

This can be rephrased in terms of an equivalent condition. For *one* message M that is maximally entangled with a reference N (which remains outside the black hole), if Alice drops M into the black hole, Bob should be able to recover a k-qubit subsystem that is maximally entangled with the reference N.

Now, we will argue that this is in turn equivalent to the condition that the post-evaporation black hole state B' is completely uncorrelated with the reference N (Figure 6). To see this, first we note that the joint state of N and B', $\sigma_{NB'}$, is mixed because of the Hawking Radiation. So, the s qubit system R (the Hawking Radiation) can be used to purify $\sigma_{NB'}$. We break up R into a k-qubit system M' and an $s-k$-qubit system Q. So, let $\mathrm{tr}_{M',Q}\left(|\xi_{NB'M'Q}\rangle\langle\xi_{NB'M'Q}|\right) = \sigma_{NB'}$. Additionally if we succeed in decoupling N and B', then

$\sigma_{NB'} = \sigma_N \otimes \sigma_{B'}$. Now, we can find a purifier for σ_N in M' and for $\sigma_{B'}$ in Q.

$$\sigma_{NB'} = \mathrm{tr}_{M',Q} |\alpha\rangle\langle\alpha| \text{ where } |\alpha\rangle = |\psi\rangle_{NM'} |\lambda\rangle_{B'Q}$$
$$\implies |\alpha\rangle = I_{NB'} \otimes U_{M'Q} |\xi_{NB'M'Q}\rangle \text{ by unitary equivalence of purifications.}$$

Therefore if N and B' are decoupled, then Bob can use the unitary operation given by the equivalence of purifications to recover a system maximally entangled with N.

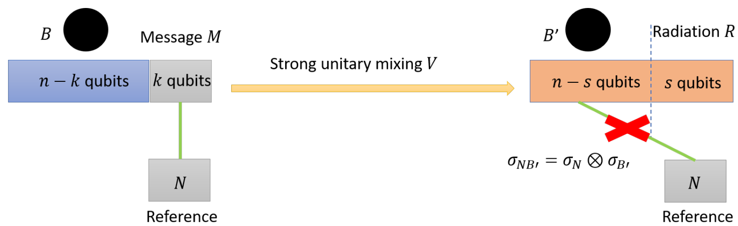

Figure 6: Equivalent condition

We have thus reduced the problem to finding how close $\sigma_{NB'}$ is to $\sigma_N \otimes \sigma_{B'}$, i.e. finding $\|\sigma_{NB'} - \sigma_N \otimes \sigma_{B'}\|_1$. Preskill and Hayden calculate this quantity averaged over unitary matrices V chosen uniformly from a Haar measure on 2^n-dimensional unitaries. They make an additional simplification by assuming that the post-evaporation black hole state $\sigma_{B'}$ is maximally mixed. The result that we need to bound this average distance is

$$\int dV \|\sigma_{NB'} - \sigma_N \otimes \sigma_{B'}\|_1 \leq \frac{d_N d_{BM}}{d_R^2} \mathrm{tr}\left((\omega_{BMN})^2\right) \tag{8}$$

where $d_N = 2^k$ is the dimension of N, and so on, and ω_{BMN} is the joint state of B, M and N. Since B was in a pure state to start with, and MN is pure too, ω_{BMN} is pure. Hence, $\mathrm{tr}\left((\omega_{BMN})^2\right) = 1$. Substituting the numbers into the RHS of the inequality in Equation (8), we get $\frac{2^k 2^n}{2^{2s}} = \frac{1}{2^{2s-n-k}}$. This means that if Bob collects a constant number of qubits more than $(n+k)/2$, he can recover the message. This is in agreement with the Page analysis which says that Bob needs to wait for half of the black hole to evaporate.

While we haven't proved the inequality in Equation (8) here, we would like to present an argument by Patrick Hayden [10] that motivates it. Going through this argument will lead us to a scenario where the results are drastically different from the Page analysis!

4.2 Using entropies to motivate Inequality (8)

We recall the definition of Renyi entropy

$$S_\alpha(\rho) = \frac{1}{1-\alpha} \log\left(\mathrm{tr}\, \rho^\alpha\right)$$

We look at S_α for some special α

1. $S_0(\rho) = \log(\mathrm{rank}(\rho))$

2. $S_1(\rho) = S(\rho)$, the von-Neumann entropy.

3. $S_2(\rho) = -\log\left(\mathrm{tr}\, \rho^2\right)$

Now, we (naively) think of dimension as a rank and use the above definitions to rewrite the RHS of Inequality (8) as

$$2^{S_0(N)+S_0(BM)-S_2(BMN)}\ 2^{-2s} \tag{9}$$

Now, we (very roughly) think of $S_0(N) + S_0(BM) - S_2(BMN)$ as $S(N) + S(BM) - S(BMN)$, which is the mutual information $I(N : BM)$ between N and BM. Then Inequality 8 seems to say that the average distance of $\sigma_{NB'}$ from being a product state is $\leq 2^{I(N:BM)-2s}$. It is easier to make sense of this – the average distance is a measure of correlation between N and B'. All the initial correlation between N and the black hole is given by $I(N : BM)$ and this correlation slowly decreases as s increases, i.e. as we throw away more and more qubits into the Hawking Radiation.

However, we see that something is not right with the above numbers – the mutual information between N and BM shouldn't depend on the size of the black hole, but the RHS of Inequality (8) clearly does! This begs the question – can we modify the scenario so that the RHS is indeed independent of n?

4.3 The Mirror Effect

If we revisit the above calculation, we see that we have equated the rank of the BM state with its dimension. But in our scenario, since the black hole started in a pure state B, the only "impurity" of BM is coming from the fact that M is entangled with N. So the rank can at most be 2^k and not 2^n, as we incorrectly assumed. What should we do to make the rank 2^n? We should make the initial black hole B mixed. This can be achieved if we modify our scenario so that the black hole is *old* (see Figure 7).

1. Alice waits for half the black hole to evaporate before throwing in her message into the black hole.
2. Meanwhile Bob collected the early Hawking radiation E that was emitted before Alice threw her qubits in.
3. Now, Bob tries to recover Alice's state from E and the new Hawking radiation R.

Figure 7: Modified scenario: Bob collected the early radiation E apart from the later Hawking Radiation R.

Since Alice waited for half the black-hole to evaporate, we can assume that B is maximally entangled with the early radiation, and hence ρ_B maximally mixed. Thus, $\operatorname{tr}\left((\omega_{BMN})^2\right) = 2^{-(n-k)}$. Next, we observe that the rank calculations done above are now correct, i.e. $\operatorname{rank}\rho_{BM} = 2^n$. Now, we can plug the numbers back into the RHS of Inequality 8 to get $2^k 2^n 2^{-(n-k)} 2^{-2s} = 2^{-2(s-k)}$.

$$\int dV \|\sigma_{NB'} - \sigma_N \otimes \sigma_{B'}\|_1 \leq \frac{1}{2^{2(s-k)}} \tag{10}$$

This is drastically different from the Page analysis! This means that just by collecting $O(1)$ qubits, i.e. a constant number of qubits more than k, Bob can recover Alice's state. So in this model for a black hole, the black hole acts as a mirror for any quantum information that is thrown into it after half of it got evaporated!

4.4 Is Black Hole Complementarity safe any more?

It takes $O(r_S)$ time to emit $O(1)$ photons, and $O(r_S) \leq O(r_S \log r_S)$! So, BHC is no longer protected by the Page time of $O(r_S^3)$, hinting at an inconsistency in BHC. However there is one more time-scale which we haven't considered yet, which will serve as BHC's saviour. We have assumed that the black hole performs a random unitary transformation instantaneously. However, the black-hole takes a non-zero time to thermalise - i.e. to scramble the information that Alice threw in. Since we do not know the internal dynamics of a black hole (for want of a quantum theory of gravity), estimating this time-scale is difficult. Hayden and Preskill, in Section 4 of [1] do this by assuming that the black hole's qubits are distributed uniformly just outside (one Planck length away) the horizon. They estimate the thermalization time to be $\log(r_S)$.

If this calculation is correct, then the estimate for Δt is $r_S \log r_S$, which means that BHC could be safe, but only precariously so!

5 Conclusions

In this work, we have developed the Black Hole Information Paradox (BHIP) and discussed how black holes behave as information mirrors in a possible resolution to the BHIP. We saw that there is a thermodynamics analogy to classical black hole phenomena. Using quantum field theory in curved spacetime of a black hole, we found that black holes are indeed characterized by the temperature $T_H = \frac{\hbar c^3}{8\pi GM}$. This indicates that black holes emit Hawking radiation, leading to their ultimate evaporation. When this phenomenon is analyzed closer, we discover that it takes pure states to mixed states, a violation of unitarity, a fundamental property of quantum physics.

Resolutions to this paradox have been developed, and numerous are still being explored. One resolution is provided by the AdS/CFT correspondence, which dictates that unitarity is preserved because the black hole can be described by a dual conformal field theory. However, if we assume that unitarity is preserved, we are still in trouble because we find that black holes can effectively clone quantum states. The Black Hole Complementarity hypothesis tries to resolve this by arguing that cloning is unverifiable. Hayden and Preskill investigated this issue to check if cloning is actually unverifiable. Using a model in which black holes act as instantaneous quantum randomizers, they found that black holes act as a mirrors for quantum information. However, using estimates of black hole thermalization time, they argue how this result is not inconsistent with black hole complementarity.

Another important problem in this context is the Firewall Paradox. Here, there is an apparent violation of the Principle of Monogamy of Entanglement, provided that a certain task, called the HH decoding task, can be done. This task amounts to separating a Bell Pair out of a system, by applying a unitary transformation just on the part of the system that doesn't contain the first qubit of the Bell Pair. Harlow and Hayden showed [11] that the computational hardness of this task is related to a problem in complexity theory on whether $SZK \subseteq BQP$ (SZK = Statistical Zero Knowledge class). This has opened up research directions at the interface of BHIP and computer science [12]. In case the HH-decoding task can be performed in polynomial time, then to prevent the paradox of Monogamy violation, one has to theorize the existence of firewalls at the horizon – where a smooth spacetime at the horizon is replaced by Planck energy particles! This would be a marked departure from our current understanding of the event horizon.

In spite of much progress in this field, many questions have yet to be answered. For example, what is the process which leads to preservation of unitarity in the evolution of black holes? What precisely is the thermalization time of a black hole? We expect that the physics of quantum information scrambling, often used to study condensed matter systems, could provide insights into calculating this. To say the least, BHIP has opened doors to new concepts in general relativity, quantum information and computer science!

References

[1] Patrick Hayden and John Preskill. Black holes as mirrors: quantum information in random subsystems. `http://iopscience.iop.org/article/10.1088/1126-6708/2007/09/120`, 2007.

[2] Ted Jacobson. Introductory lectures on black hole thermodynamics. `"https://www.physics.umd.edu/grt/taj/776b/lectures.pdf`, 1996.

[3] Joseph Polchinski. The black hole information problem. In *New Frontiers in Fields and Strings: TASI 2015 Proceedings of the 2015 Theoretical Advanced Study Institute in Elementary Particle Physics*, pages 353–397. 2017.

[4] Ted Jacobson. Introduction to quantum fields in curved spacetime and the hawking effect. `https://arxiv.org/abs/gr-qc/0308048`, 2003.

[5] Mark Van Raamsdonk. Lectures on gravity and entanglement. In *New Frontiers in Fields and Strings: TASI 2015 Proceedings of the 2015 Theoretical Advanced Study Institute in Elementary Particle Physics*, pages 297–351. World Scientific, 2017.

[6] Anthony Zee. Einstein gravity in a nutshell, 2013.

[7] Sean Carroll. Spacetime and geometry: An introduction to general relativity, 2004.

[8] Joseph Polchinski. https://physics.princeton.edu/strings2014/slides/polchinski.pdf, 2014.

[9] Leonard Susskind, Larus Thorlacius, and John Uglum. The stretched horizon and black hole complementarity. *Physical Review D*, 48(8):3743, 1993.

[10] Patrick Hayden. Lecture at us-india advanced studies institute on thermalization: From glasses to black holes. `https://www.youtube.com/watch?v=qom1jAPlycQ&t=4292s`.

[11] Daniel Harlow and Patrick Hayden. Quantum computation vs. firewalls. *Journal of High Energy Physics*, 2013(6):85, 2013.

[12] Scott Aaronson. The complexity of quantum states and transformations: from quantum money to black holes. *arXiv preprint arXiv:1607.05256*, 2016.

[13] Michael Peskin and Daniel Schroeder. An introduction to quantum field theory, 1995.

A A Brief Derivation of Hawking Radiation

A.1 Motivation and Planck's law

Above, we mentioned that Stephen Hawking showed that a black hole radiates, with a spectrum described by a blackbody distribution with temperature $T = T_H$. Here, we present a brief derivation of Hawking radiation and black hole evaporation.

Our first clue to unraveling black hole radiation is that the Hawking temperature is proportional to $\hbar$: $T_H = \frac{\hbar c^3}{8\pi G M k_B}$. In the classical regime, wherein one neglects quantum effects, we take the limit $\hbar \to 0$. In this limit, the Hawking temperature goes to 0, at which temperature we would not expect an object to radiate. Therefore, it is conceivable that black hole radiation is a quantum effect that only arises when we restore the nonzero value of $\hbar$. In order to examine the possibility of black hole radiation, we must therefore analyze the quantum physics of black holes.

A.2 Quantum field theory in flat spacetime

Unfortunately, if we try to examine the quantum physics of black holes, we encounter at a problem: we do not know the correct theory of quantum gravity, and so we cannot analyze the quantum physics of black holes exactly. In view of this shortcoming, we take our second best option: apply quantum field theory (QFT) to the curved spacetime of general relativity. In other words, we impose quantum physics onto a classical curved spacetime background. Since the effects of quantum gravity are irrelevant at length scales greater than the Planck length, we expect that QFT in curved spacetime will provide an accurate description of our situation on scales attainable by present-day experiments.

Before studying QFT in curved spacetime, let us first review QFT in flat spacetime. For simplicity, we will analyze the free (Klein-Gordon) scalar field $\phi(\vec{x}, t)$, and follow the analysis presented in [13]. In the following, we will set $\hbar = c = 1$.

Classically, our field is governed by a Lagrangian density $\mathcal{L}$, which obeys the Euler-Lagrange equation:

$$\mathcal{L} = \frac{1}{2}\eta^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - \frac{1}{2}m^2\phi^2, \qquad \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} - \frac{\partial \mathcal{L}}{\partial \phi} = 0. \tag{11}$$

where $\eta^{\mu\nu}$ is the Minkowski metric, and the repeated indices are summed over 0 to 3. In the canonical formalism of classical field theory, we define the field conjugate to $\phi(\vec{x})$ as the momentum density: $\pi(\vec{x}) = \frac{\partial \mathcal{L}}{\partial(\partial_0 \phi(\vec{x}, t))}$. This is essentially an additional coordinate used to describe our system.

Next, in order to build a QFT, we quantize our classical field theory by promoting ϕ and π to operators acting on a Hilbert space that obey the commutation relation $[\phi(\vec{x},t),\pi(\vec{x}',t)] = i\hbar\delta(\vec{x}-\vec{x}')$. $\phi(\vec{x},t)$ and $\pi(\vec{x},t)$ are now operators acting on the quantum state of the system, which we denote by $|\Psi\rangle$.

In order to study the action of these operators on a state, one expands $\phi(\vec{x})$ into its Fourier modes (wave-like solutions to the equations of motion) with coefficients $a_{\vec{p}}$ and $a^\dagger_{\vec{p}}$:

$$\phi(\vec{x},t) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\vec{p}}}}\big(a_{\vec{p}}\ e^{-i(\omega_{\vec{p}}t-\vec{p}\cdot\vec{x})} + a^\dagger_{\vec{p}}\ e^{i(\omega_{\vec{p}}t-\vec{p}\cdot\vec{x})}\big), \quad \omega_{\vec{p}} = \sqrt{|\vec{p}|^2+m^2}, \tag{12}$$

$$\pi(\vec{x},t) = \int \frac{d^3p}{(2\pi)^3}\sqrt{\frac{\omega_{\vec{p}}}{2}}\big(a_{\vec{p}}\ e^{-i(\omega_{\vec{p}}t-\vec{p}\cdot\vec{x})} - a^\dagger_{\vec{p}}\ e^{i(\omega_{\vec{p}}t-\vec{p}\cdot\vec{x})}\big). \tag{13}$$

Once this procedure is performed, one can use the canonical commutation relations to find that $a^\dagger_{\vec{p}}$ and $a_{\vec{p}}$, act as creation and annihilation operators, respectively. In particular, these operators act on the state of a system by creating or annihilating one particle with momentum $\vec{p}$. In addition, the operator that measures the number of particles with momentum $\vec{p}$ is the number operator $N_{\vec{p}} = a^\dagger_{\vec{p}}a_{\vec{p}}$. The expected number of particles in the mode $\vec{p}$ is then

$$\langle N_{\vec{p}}\rangle = \langle\Psi|\, N_{\vec{p}}\,|\Psi\rangle\,. \tag{14}$$

If we use this procedure to count the number of photons of the quantized electromagnetic field surrounding a body at temperature T, we discover that $\langle N_{\vec{p}}\rangle$ is given by equation 2, which describes the radiation emitted by the body. Therefore, in order to analyze the possibility of black hole radiation, we aim to apply quantum field theory to the curved spacetime of a black hole, and determine $\langle N_{\vec{p}}\rangle$ in the quantum state of the black hole.

A.3 QFT in curved spacetime

For concreteness, we will analyze the quantized Klein-Gordon field in curved spacetime, following the approach given in [4] and [7]. Nevertheless, our procedure applies for arbitrary quantum fields in general curved spacetimes.

In curved spacetime, QFT is slightly modified. In particular, in curved spacetime, we must take into account the metric tensor $g_{\mu\nu}$, which replaces the Minkowski metric $\eta_{\mu\nu}$, and describes the geometry of curved spacetime. In the Lagrangian density, we replace partial derivatives by covariant derivatives ($\partial_\mu \mapsto \nabla_\mu$), which account for the curvature of spacetime, and multiply by $\sqrt{-\det(g)}$ to incorporate the proper volume element:

$$\mathcal{L} = \sqrt{-\det(g)}\Big(\frac{1}{2}g^{\mu\nu}\nabla_\mu\phi\nabla_\nu\phi - \frac{1}{2}m^2\phi^2\Big). \tag{15}$$

We then quantize our fields by imposing commutation relations, and again expand $\phi(\vec{x},t)$ in an orthonormal basis of plane waves, like the Fourier decomposition above. However, we must be careful in doing so. In particular, in order for the particles of the quantum field to have well-defined, positive definite energies, we require the plane waves to have positive frequencies with respect to a future directed timelike Killing vector (i.e. positive frequency with respect to the direction of time). In curved spacetime, the timelike and spacelike directions may be dependent on one's location. For instance, in the spacetime of a black hole, the spacelike and timelike direction differ on opposite sides of the event horizon.This means that a quantity which is identified with an energy outside the horizon becomes a momentum inside, and vice versa. Therefore, if we are to have well defined positive frequency modes, they must be dependent on the region of spacetime one is analyzing. In particular, the creation and annihilation operators will be different for observers in different regions of the black hole spacetime.

In view of this observation, let $\{a^\dagger_{\vec{p}}, a_{\vec{p}}\}$ and $\{b^\dagger_{\vec{p}}, b_{\vec{p}}\}$ be two sets of creation and annihilation operators for observers in two different regions of spacetime. Since these sets of operators satisfy similar algebraic relations, induced by the canonical commutation relations, these sets of operators are related to each other by Bogoliubov transformations:

$$b_{\vec{p}} = \int \frac{d^3p'}{(2\pi)^3}\big(\alpha_{\vec{p}\vec{p}'}\ a_{\vec{p}'} - \beta_{\vec{p}\vec{p}'}\ b^\dagger_{\vec{p}'}\big) \tag{16}$$

$$a_{\vec{p}} = \int \frac{d^3p'}{(2\pi)^3} \left(\alpha^*_{\vec{p}'\vec{p}}\, a_{\vec{p}'} + \beta^*_{\vec{p}'\vec{p}}\, b^\dagger_{\vec{p}'}\right) \tag{17}$$

Here, $\alpha_{\vec{p}\vec{p}'}$ and $\beta_{\vec{p}\vec{p}'}$ are functions that relate the sets of operators. Their specific form is specified by the system under analysis.

Since different observers have different sets of creation/annihilation operators that are related to each other by such a transformation, observers in different spacetime regions perceive differing quantum environments. This is the essence of the Hawking effect: an infalling observer near the black hole horizon will perceive the quantum state to be a vacuum, whereas observers far from the black hole perceive the quantum state to contain particles, emitted as radiation from the black hole.

A.4 Hawking radiation

Let us demonstrate the last claim explicitly. Consider the quantized Klein-Gordon field in the spacetime of a Schwarzschild black hole of mass M. Let there be an asymptotic observer far away from the black hole with creation and annihilation operators $\{a^\dagger_{\vec{p}}, a_{\vec{p}}\}$, and an infalling observer who crosses the event horizon and has a set of operators $\{b^\dagger_{\vec{p}}, b_{\vec{p}}\}$. For the asymptotic observer, the Fourier modes may have time dependence $e^{i\omega t}$, since this wave solution has positive frequency with respect to a future directed timelike Killing vector outside the horizon. Hence, $a_{\vec{p}}$ and $a^\dagger_{\vec{p}}$ are well-defined only outside the black hole horizon. On the other hand, for the infalling observer, the time dependence of the Fourier modes must go as $e^{i\omega u}$, where $u = t - r^*$ is the advanced time, and $r^* = r + 2GM \ln |\frac{r}{2GM} - 1|$ is the tortoise coordinate of Schwarzschild spacetime [3].

Furthermore, since the geometry we analyze is not strongly time dependent, the adiabatic hypothesis states that the quantum state of the black hole will be a vacuum with respect to the infalling modes, which vary on a much smaller time scale. Letting $|\Psi\rangle$ denote this quantum state, this claim amounts to the relation $b_{\vec{p}}\,|\Psi\rangle = 0$. The number of particles seen by the asymptotic observer is $\langle\Psi|\, a^\dagger_{\vec{p}} a_{\vec{p}}\, |\Psi\rangle$. Use of the Bogoliubov transformation for this scenario then indicate that the spectrum of particles in this spacetime is [3]

$$\langle N(\omega)\rangle = \langle\Psi|\, a^\dagger_{\vec{p}}\, a_{\vec{p}}\, |\Psi\rangle = \frac{\Gamma}{e^{\frac{\hbar\omega}{k_B T_H}} - 1}. \tag{18}$$

In this equation Γ is the greybody factor, which quantifies the lower intensity of radiation emitted by the black hole [4] with respect to the radiation emitted by an ideal black body at temperature T_H. This is nearly the same as equation (2): the number of particles per mode seen by the asymptotic observer is given by a blackbody spectrum, with temperature T_H. We therefore conclude that an observer far away from the black hole perceives the black hole to emit thermal radiation akin to an object at temperature T_H.

Furthermore, one can demonstrate that black hole radiation is similar to pair production, as we noted previously. For every outgoing particle of radiation, there is another infalling particle, traveling toward the singularity, which has negative energy and is produced by a creation operator $\tilde{a}^\dagger_{\vec{p}}$. Such a pair of radiation particles is known as a Hawking pair. This process is not unphysical, because, as we have already stated, energy inside the horizon becomes a momentum. $\tilde{a}^\dagger_{\vec{p}}$ is defined as the equivalent of $a^\dagger_{\vec{p}}$, but acting inside the horizon. It can be shown that the vacuum state of the infalling observer can be written as:

$$|\Psi\rangle = \mathcal{N} \exp\left(\int_0^\infty \frac{d\omega}{2\pi} \mathrm{e}^{-\frac{\omega}{2T_H}}\, a^\dagger_{\vec{p}} \tilde{a}^\dagger_{\vec{p}}\right) |0\rangle_a \equiv |\Psi_{a\tilde{a}}\rangle \tag{19}$$

where $|0\rangle_a$ is the vacuum for the asymptotic observer. $|\Psi_{a\tilde{a}}\rangle$ is an entangled state. Therefore, each photon belonging to the Hawking radiation outside the horizon is entangled with a corresponding particle inside the horizon. This property was fundamental to development of the black hole information problem above.