Imperial College London

Imperial College London

Department of Physics

Towards a Resolution of the Black Hole Information Loss Problem

Author:
Yvette Chow

Supervisor:
Prof. Arkady Tseytlin

Submitted in partial fulfillment of the requirements for the degree of Master of Science of Imperial College London

September 2020

Abstract

Four and a half decades after the introduction of the information loss problem by Hawking in 1976, it is a current thought that now, in 2020, an aspect of it has been solved. This aspect relates to the recovery of the initial infalling matter state from the interior of the black hole through operations performed on the final radiation state. Arriving at the solution involved integrating key historical and recent works such as Page's 1993 study of entropies in black hole evaporation, Ryu-Takayanagi's 2006 holographic area relation, Faulkner, Lewkowycz and Maldacena's and Engelhardt and Wall's extensions to the area relations in 2013 and 2015 respectively, Penington's work on entanglement wedges in 2019 and Almheiri, Mahajan, Maldacena and Zhao's work on the island conjecture in 2019. This dissertation reviews these selected works.

Contents

1 Introduction **1**

1.1 Hawking Radiation . . . 3

1.2 The Laws of Thermodynamics and Black Hole Mechanics . . . 3

2 Entropies **6**

2.1 A Statistical Interpretation . . . 6

2.2 Entanglement Entropy . . . 7

2.3 Holographic Entanglement Entropy . . . 9

2.4 Thermodynamic Entropy . . . 12

2.5 Semiclassical Generalised Entropy . . . 13

2.6 Quantum Corrections to Semiclassical Generalised Entropy in Black Hole Thermodynamics and in AdS/CFT . . . 13

2.7 Quantum Extremal Surfaces and Gravitational Fine-Grained Entropy . . . 16

3 The Information Loss Problem **18**

3.1 Paradoxical Spacetime Geometries for the Black Hole . . . 20

3.2 The Central Dogma . . . 20

4 Page Curve for the Black Hole **22**

4.1 Quantifying the Page Curve . . . 24

4.2 Gravitational Fine-Grained Entropy Formula and the Page Curve for the Black Hole . . . 25

5 Islands and the Page Curve for Hawking Radiation **29**

5.1 The Euclidean Black Hole . . . 32

6 Entanglement Wedge Reconstruction **35**

7 Conclusion and Discussion **39**

7.1 A Few Notes on the Page Curve . . . 41

7.2 Looking Towards the Future . . . 42

1 Introduction

The subject area of the black hole information loss problem is a melting pot of various areas of physics – from general relativity, statistical mechanics, quantum mechanics to string theory, quantum information and chaos theory. These areas are integrated with the several layers of studying black holes – from a far away external perspective to an internal perspective going all the way to the singularity. These separate individual perspectives produces different descriptions and thus properties of the black hole. General relativity provides a surface-level description of the spacetime outside the black hole and does not reveal any fine-grained details. From the point of view of the infalling observer, such fine-grained details are also not entirely accessible as there is only a finite amount of information available to them in a finite amount of time before reaching the singularity. We will see that the tools of quantum gravity and string theory are required to describe the black hole interior and study its microscopic degrees of freedom.

This dissertation will take a historical route and progress towards providing a resolution to the information loss problem. Crucially, if black hole evaporation starts in a pure total state, then the thermodynamic entropy of Hawking radiation has to eventually decrease back to zero at the end of evaporation. This was missed in Hawking's original calculation in 1975. We will see this, mostly qualitatively, using Page's 1993 and 2013 analysis, and more recently, with significant work by Almheiri *et al.* [1, 2] and Penington *et al.* [3]. The latter research considered semiclassical descriptions of specific gravity theories. In reviewing this, the entropies involved in black hole evaporation and the role of entanglement wedges in information recovery from the black hole interior will also be discussed. We will begin by providing a historical overview of black hole mechancis, thermodynamics and the information loss problem.

Classical black holes were first predicted by Einstein's equations as its classical solutions, but physicists at the time did not place much value to its physical realisation. In 1916, Schwarzschild discovered the first exact spherically symmetric solution to Einstein's equations in vacuum $R_{\mu\nu} = 0$, called the Schwarzschild metric

$$ds^2 = -\left(1 - \frac{2M}{r}\right) dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2 d\Omega^2, \tag{1.1}$$

where M is the mass of the star that collapses to form a black hole and $d\Omega^2 = d\theta^2 + sin^2\theta d\phi^2$ is the metric on the unit two-sphere. Schwarzschild found two singularities associated to his solution. One was a coordinate singularity at the Schwarzschild radius $r_s = 2M$, called the event horizon, and the other was a curvature singularity at $r = 0$. Subsequent work by Eddington, Finkelstein, Kruskal and Szekeres from the 1920s to 1960s shed light on the former singularity. It was found that by changing the coordinate system from Minkowski to Kruskal, the geometry is smoothly extendable past $r = r_s$, where the metric is well defined and in fact for all $r > 0$. In other words, there is nothing special about the event horizon for a freely falling observer, which is the equivalence principle. In Kruskal coordinates it is also seen that $r = 0$ is not a position in space but rather an unavoidable moment of time if one passes the event horizon. A signal that has passed r_s will not be able to escape back out.

In 1971 Hawking's area theorem was published. He proved that the horizon area of a classical black hole can never decrease with time [4]. In 1972–1973, Bardeen, Carter and Hawking raised the analogy between a black hole's area and thermodynamic entropy. They suggested a mathematical relation between the laws of black hole mechanics and the laws of thermodynamics but did not draw any physical conclusions to the relations [5]. Also in 1973, papers by Bekenstein sparked the study into the meaning of black hole entropy and the mechanism behind the laws of black hole mechanics [6]. Bardeen, Carter and Hawking took the view that black holes have zero temperature, do not radiate and thus do not have a physical entropy. This was implied by the fact that classical black holes obey the no-hair theorem, which states that a classical black hole is completely characterised by its mass M, angular momentum J and charge Q. By this theorem, it would seem that solutions for black holes have no connection to thermodynamics. But Bekenstein conducted the following thought experiment: consider throwing a highly entropic system into the black hole. If black holes indeed do not radiate, the object would cool to absolute zero and this would result in a decrease in the total entropy of the external universe not including the black hole. This violates the second and third law of thermodynamics. In order for the second law of thermodynamics to hold, the increase in entropy of the black hole S_{BH} has to outweigh the decrease in entropy of the exterior S_{external}. He called this the *generalised second law* (GSL); in any process involving black holes, the generalised entropy S_{total} does not increase [7]:

$$dS_{\mathrm{total}} = dS_{\mathrm{BH}} + dS_{\mathrm{external}} \geq 0. \tag{1.2}$$

Bekenstein's GSL of thermodynamics holds true as if S_{BH} decreases via black hole evaporation, S_{ext} will increase due to the thermodynamic nature of Hawking radiation in a way such that $S_{\mathrm{total}} = S_{\mathrm{BH}} + S_{\mathrm{ext}}$ does not decrease with time. In this way, the second law of thermodynamics is not violated. More subtly, we see that to an external observer, black holes must have an entropy in order to avoid the violation of the second law of thermodynamics, which led to the GSL. Crucially, the GSL implies the dependence of the increase in the horizon area of the black hole horizon on the infalling matter's microscopic degrees of freedom. But this dependence is not implied by the Einstein's equations that governs black holes. It was shown in [7] and [8] that there is a bound which characterises the maximum value of entropy S that a system with finite energy E and size, characterised by a radius R, can have. This is the Bekenstein bound

$$S \leq \frac{2\pi}{\hbar} ER. \tag{1.3}$$

In other words, the quantum information of the system is related to its energy and geometry. Importantly, as pointed out by Bousso in [8], the absence of the constant G_N and black hole parameters A, M, Q, J in the bound relation implies its universality; the bound is associated with only the system.

1.1 Hawking Radiation

In 1975 Hawking discovered that black holes radiate with a temperature $T_{\mathrm{H}} = \hbar c^3/8\pi k_B G_N M$ [9], implying that the temperature of the black hole increases as it evaporates. The geometry of a black hole comprises an exterior and an interior region. A full quantum state near the black hole event horizon consists of entangled pairs of particles, which are created and annihilated due to quantum fluctuations in the vacuum. One particle of the pair falls into the singularity in the interior while the other particle escapes out to infinity. The latter is termed outgoing Hawking quanta or modes and constitute the outgoing Hawking radiation. In Hawking's 1976 paper the information loss problem was introduced [10]. In quantum mechanics, an isolated system evolves unitarily and information is conserved. Hawking's calculation revealed that radiation emitted from a black hole does not depend on the initial state of photons; different initial states can lead to the same final state. A black hole formed from the collapse of a pure state of matter ends up emitting radiation of mixed states that are entangled both with previously infalling matter states and with earlier radiation states, which gives the final mixed radiation state a high entanglement entropy (see section 2.2). As the black hole evaporates, its area decreases until it reaches zero, i.e. it completely evaporates into outgoing quanta such as photons and gravitons, which constitutes the outgoing Hawking radiation. This evaporation of pure infalling matter into mixed outgoing radiation implies the violation of unitarity in quantum mechanics and loss of information. We have that in the process of formation and evaporation of a black hole, ingoing quanta are deemed lost and unretrievable.

Bekenstein proved that the black hole entropy associated to T_{H} is proportional to the area of the horizon and Hawking later proved the proportionality constant by using QFT in curved spacetime. The Bekenstein-Hawking entropy was understood to measure the information inside the black hole according to information theory. We will now see the analogies between the laws of black hole mechanics and the laws of thermodynamics by re-expressing Einstein's equations in an indicative way.

1.2 The Laws of Thermodynamics and Black Hole Mechanics

The first law of black hole mechanics states that if an object is slowly added to a stationary black hole with event (Killing) horizon area A, mass M, charge Q, angular momentum J, surface gravity of the horizon κ, electric surface potential Φ_H and angular velocity Ω_H, its mass, charge and angular momentum will change in response to the change in its area as

$$dM = \frac{\kappa}{8\pi} dA + \Omega_H dJ + \Phi_H dQ. \tag{1.4}$$

In parallel, the first law of thermodynamics states

$$dM = T dS_{\mathrm{BH}} + \Omega_H dJ + \Phi_H dQ. \tag{1.5}$$

If one defines for a black hole a temperature $T \propto \kappa$ and an entropy $S_{\mathrm{BH}} \propto A$, one can rewrite equation (1.4) as (1.5) and infer the Hawking temperature $T = \hbar\kappa/2\pi k_B$ and the thermodynamic Bekenstein-Hawking entropy

$$S_{\mathrm{BH}} \sim \frac{A}{4{l_p}^2}, \tag{1.6}$$

where $A = 4\pi {r_s}^2$ is the horizon area and $l_p = \sqrt{G_N \hbar}$ is the Planck length. This entropy is of statistical origin and counts the logarithm of the number of microstates. The resemblance between (1.4) and (1.5) was viewed by Bardeen Carter and Hawking as only a mathematical one. The entropy (1.6) is massive; the black hole in the centre of the Milky Way has an entropy larger than that of all observable matter in the universe excluding black holes [11]. The most straightforward way to derive (1.6) was proposed by Gibbons and Hawking in 1977 in [12]. This Gibbons-Hawking method provides the classical answer $A/4\hbar$, which is the tree level (lowest order $\hbar$) contribution in the functional integral. In addition, (1.6) also appears in holographic entanglement entropy (see section 2.3). By attributing an area of l_p^2 to one pair of entangled particles, the entanglement entropy is proportional to the amount of entanglement among particle pairs across the event horizon. Entropy (1.6) also appears in the covariant entropy bound and in general codimension 2 surfaces [13] with the area being the leading order classical piece.

The second law of black hole mechanics, also referred to as Hawking's area theorem, states that the area A of the future event horizon of a black hole in asymptotically flat spacetime is a non-decreasing function of time:

$$dA \geq 0. \tag{1.7}$$

In the event of two black holes coalescing, the area of the final state is greater than the sum of the areas of the two initial states. Analogously, the second law of thermodynamics demands that the total entropy of a system S should not decrease: $dS \geq 0$. Whereas entropy can be transferred from one system to another, the same is not true for the area of black holes.

The zeroth law of black hole mechanics states that the surface gravity κ is constant over the future event horizon of a stationary black hole. This parallels the zeroth law of thermodynamics, which states that a system in thermal equilibrium has a constant temperature. Lastly, the third law of black hole mechanics states that it is impossible to reduce the surface gravity of a black hole to zero in a finite number of operations. In parallel, the third law of thermodynamics states that the temperature of any system cannot be reduced to absolute zero in a finite number of operations.

The study of black holes boils down to understanding the following areas: 1. The microscopic meaning of black hole entropy, in particular, from a gravity perspective; 2. The mechanism of the GSL and 3. The information loss problem. A resolution in these areas is important, on a basic level, to ensure that the fundamental laws of quantum mechanics and information preservation still holds in the case of black hole evaporation, but also for its potential to implicate future lessons on the black hole interior and singularity. The study of the information loss problem draws from and integrates different areas of physics, as well as produces major developments and extensions in those areas. Of notable example is the discovery of AdS/CFT in string theory, which has injected new ideas and driven discoveries in the area of black holes since the nineties. Initially introduced as a new way to count microscopic states of black holes, it also suggested that information does indeed escape and is not lost. In 2020 we are four decades into understanding the information loss problem, an aspect of which has been solved, as claimed in a recent paper by Almheiri *et al.* [1].

The plan of this dissertation is as follows. In this section an overview of the history of black hole mechanics, thermodynamics and the information loss problem were given. A brief statistical interpretation of the thermodynamic Hawking entropy is given in section 2.1. In section 2, two distinct notions of entropy are introduced. The first is the entanglement entropy or fine-grained entropy, which is covered in section 2.2. The Ryu-Takayanagi prescription [14–16] in AdS/CFT for computing the holographic entanglement entropy, which was inspired by Hawking's area term, is then introduced in section 2.3. The second notion of entropy introduced in section 2.4 is the thermodynamic or coarse-grained entropy. The semiclassical generalised entropy and its quantum corrections are then introduced in sections 2.5 and 2.6 respectively. The section ends with a review of quantum extremal surfaces and its use in computing the gravitational generalised semiclassical entropy in section 2.7. The information loss problem and some of its early suggested solutions and oppositions, along with the concept of the central dogma, are introduced in section 3. This leads to the discussion of the Page curve in section 4, where the evolution of the entanglement and thermodynamic entropies involved in black hole evaporation are explained. A brief quantification of the curve is provided in section 4.1, followed by using the gravitational fine-grained entropy formula on an evaporating black hole to obtain its Page curve in section 4.2. In section 5, the island conjecture for computing the entropy of Hawking radiation is discussed. The Euclidean black hole is then introduced in section 5.1 and sets the historical context for the concept of entanglement wedge reconstruction, which we will finally see in section 6 is involved in resolving an aspect of the information loss problem. The dissertation is concluded with a summary and discussion in section 7.

2 Entropies

In this section a brief statistical interpretation of black hole entropy is given. The notions of entanglement entropy, holographic entanglement entropy and thermodynamic entropy are introduced. The semiclassical generalised entropy and its quantum corrections are then covered. Lastly, the notion of a quantum extremal surface and gravitational fine-grained entropy are reviewed.

2.1 A Statistical Interpretation

Recall that in statistical mechanics, we obtain the macroscopic properties of a system, such as temperature and statistical entropy, by counting the microscopic degrees of freedom or microstates of the system. For example, the entropy is given by the logarithm of the number of microstates. However in black hole thermodynamics, it is a struggle to understand what microscopic degrees of freedom correspond to the temperature and entropy of the black hole. The presence of the physical constants $\hbar$ and G_N in the Bekenstein-Hawking temperature formula suggests that the microscopic degrees of freedom describing the statistical mechanics of black holes could be described by a full theory of quantum gravity [17]. However, it is a non-trivial task to find a statistical mechanical description of black hole degrees of freedom; one has to find candidates for a full theory of quantum gravity that produces entropy expressions that are in agreement with the Bekenstein-Hawking entropy. Investigations attempting to address the nature, description and interpretation of microscopic black hole degrees of freedom in the past were found to naturally lead one to use the tools of quantum gravity and holography. The different areas related to black hole statistical mechanics studied included entanglement entropy, the AdS/CFT correspondence, weakly coupled strings, "fuzzballs", and loop quantum gravity among others [17]. We will focus in particular on the first two areas in this dissertation. It has to be noted that the derivations in these different areas, which counts different microstates, gave the same macrostate properties. This was called "the problem of universality" by Carlip in [17].

In the area of weakly coupled strings, Stominger and Vafa in 1996 were the first to count black hole microstates and to reproduce the Bekenstein-Hawking entropy for a class of five-dimensional extremal black holes [18]. It was shown that their results are extendable to some other class of extremal and nonextremal black holes [19]. In 2007, Horowitz and Roberts reproduced the entropy for a four-dimensional Kerr black hole by counting microstates in string theory [20]. A characteristic of this method is that the computation of the entropy is dependent on the black hole geometry; a different calculation would have to be done for, say, a three-charge black hole in five dimensions and a four-charge black hole in six dimensions.

The AdS/CFT correspondence discovered in 1998 by Maldacena [21, 22] is a relation in string theory that conjectures a gauge/gravity duality or map between a quantum theory of gravity in d-dimensional asymptotically anti-de-sitter (AdS) bulk spacetimes $\mathcal{M}$ to a certain conformal field theory (CFT) on the conformal boundary, $\partial\mathcal{M}$ of $\mathcal{M}$, that is in a flat $(d-1)$-dimensional space. This correspondence naturally obeys the holographic principle [23–25], which relates the number of degrees of freedom in a theory of quantum gravity to the area of the system, i.e. AdS/CFT correspondence is a relation between entropy and area. This allows one to study entropy dependence on area instead of volume, which is usually the case for extensive entropies in QFTs. A black hole in AdS space can be described in terms of a QFT in one lower dimension; AdS/CFT allows one to count the microstates and compute the entanglement entropy in a one dimensional lower non-gravitational dual CFT associated with a subsystem of the boundary CFT. This implies that information in the bulk is stored in the one dimensional lower event horizon of the black hole, which can be regarded as a hologram.

One can also statistically interpret the black hole degrees of freedom by computing the quantum mechanical (holographic) entanglement entropy. This was realised in 1994 in [26, 27] and in subsequent years by [28–30]. The entanglement entropy characterises correlations between the degrees of freedom near and across the horizon that are associated to information loss.

2.2 Entanglement Entropy

We aim to answer the following question: how is the initial infalling data encoded in the outgoing Hawking radiation? First, we have to distinguish between two concepts of entropy. The first we will discuss is entanglement entropy, which is quantum mechanical in nature. This entropy is also called the von Neumann entropy, or fine-grained entropy. It arises due to the existence of a fundamental limit to the level of detail we can know about the microscopic behaviour of a state of the system; this limit exists regardless of the amount and complexity of the measurements performed on the state. This entropy is conserved by the principle of unitarity in quantum mechanics, which translates to the preservation of information. Hence von Neumann entropy plays an important role in the study of quantum information.

Consider a total quantum system composed of two subsystems A and B with $A \cup B$ being a pure state. The states of the total system lives in the bipartite Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. A general state in this space can be written as $|\psi\rangle = \sum_{a,b} \psi(a,b)\, |a\rangle \otimes |b\rangle$, where $\{|a(b)\rangle\}$ is an orthonormal basis for $\mathcal{H}_{A(B)}$. The state is invariant under unitary time evolution. The reduced density matrix of the subsystem $A(B)$ in the basis $\{|a(b)\rangle\}$, $\rho_{A(B)}$, is the partial trace over subsystem $B(A)$ of the total density matrix $\rho_{\text{tot}} = \langle\psi|\psi\rangle$ [31]:

$$\rho_A = \text{tr}_B\,\rho_{\text{tot}}, \tag{2.1}$$

$$\rho_B = \text{tr}_A\,\rho_{\text{tot}}. \tag{2.2}$$

The entanglement entropies of subsystems A and B are defined by

$$S_E(A) = -\,\text{tr}_A\,\rho_A \log \rho_A, \tag{2.3}$$

$$S_E(B) = -\,\text{tr}_B\,\rho_B \log \rho_B. \tag{2.4}$$

Entanglement entropy can increase or decrease with time and it is not the entropy that appears in the second law of thermodynamics, which would be the thermodynamic entropy (see section 2.4). Also note that the entanglement entropy can be defined at a finite temperature $T = \beta^{-1}$, by using a thermal density matrix $\rho_{th} = e^{\beta H}$, where H is the Hamiltonian. Then for a total system A, $S_A(\beta)$ is just the thermodynamic entropy. A general property in the high temperature limit is that $S_{A_1}(\beta) - S_{A_2}(\beta)$ would approach $S_{th,A_1} - S_{th,A_2}$, where S_{th} is the thermodynamic entropy [16].

Entanglement entropy is generally applicable to quantum mechanical lattice models and QFTs. Such an example would be a quantum mechanical spin chain that has a large number of degrees of freedom. At zero temperature the von Neumann entropy of the total state, given by the density matrix of the pure ground state, would be zero. Imaginarily divide the total system into two sub-chains, corresponding to subsystems A and B. The entanglement entropy of the subsystem A measured by an observer that has no access to subsystem B would measure the extent to which the state $|\psi\rangle$ is entangled [16].

For a pure state of $A \cup B$, $S_E(A) = S_E(B)$ and a subsystem has $S_E = 0$ only if there is zero entanglement between the two subsystems, i.e. the total state is not quantum. Assuming that the dimension of $\mathcal{H}_B$, D_B, is less than or equal to that of $\mathcal{H}_A$, then the maximum value of $S_E(B)$ (and thus of $S_E(A)$) is $S_E(B)_{\max} = -\ln(D_B)$ [31]. In non-static backgrounds, the density matrices are time dependent and the time at which the entropy is measured will have to be specified [16]. An important property of entanglement entropy is that it obeys the strong subadditivity condition [32], which are the inequality relations [16]:

$$S_{A+B+C} + S_B \leq S_{A+B} + S_{B+C}, \tag{2.5}$$

$$S_A + S_C \leq S_{A+B} + S_{B+C}, \tag{2.6}$$

where A, B and C are subsystems with no overlaps. From these relations, a quantity called the mutual information $I(A,B)$ can be defined; letting B in (2.6) be an empty set, we obtain [16]

$$I(A,B) = S_A + S_B - S_{A+B} \geq 0. \tag{2.7}$$

The mutual information between two subsystems for a specific state in AdS/CFT was shown to be proportional to the Bekenstein-Hawking entropy in [33]. In the context of black holes, the outgoing radiation located outside the black hole is represented by the subsystem A and the interior of the black hole is represented by subsystem B.

2.3 Holographic Entanglement Entropy

Entanglement entropy is useful in allowing one to study spacetime geometry emerging from entanglement. The initial studies of entanglement entropy later led to the discovery of holographic entanglement entropy (HEE), which was inspired by AdS/CFT correspondence. In particular, analogies between the holographic principle and the Bekenstein-Hawking entropy given by the area law motivated a gravitational interpretation of the entanglement entropy in QFTs. Ryu and Takayanagi (R-T) [14] and Hubeny, Rangamani, and Takayanagi (HRT) [15] provided the first tools to compute HEE using the ideas of bulk reconstruction and gravity quantisation to study the holographic properties of a d-dimensional black hole embedded on the $(d+1)$-dimensional boundary of asymptotic AdS space. By considering the geometric properties of an appropriate bulk spacetime, the HEE connects the dynamics of the bulk and the entanglement entropy. A gravitational perspective will be taken to explore the HEE. We have seen that the entanglement entropy S_A measures the extent of entanglement with subsystem B and thus the information encoded in B. We now reformulate this question as: where in AdS space contributes to the evaluation of S_A in the CFT, or which part of AdS encodes the information which is given in the CFT ?

In the previous subsection we saw how to compute the entanglement entropy among two subsystems in QFTs. A modification to this prescription is required when considering theories with a gravity dual [14, 16]. In such cases, we have to first look for a static (spatial) surface in the spatial boundary region A that extends up to its boundary which has the minimal area. This would be called the minimal surface and it splits the bulk into two regions – region A_b that resides in the bulk and its complement – corresponding to two subsystems. These regions are illustrated in figure 1. The area of the minimal surface in this prescription is associated to the classical bulk contribution to the HEE, which is given by [34]

$$S(A) = \frac{(\text{Area})_{\text{min}}}{4G_N}. \tag{2.8}$$

We now introduce the general form of the formula (2.8). In 2006, R-T motivated the first

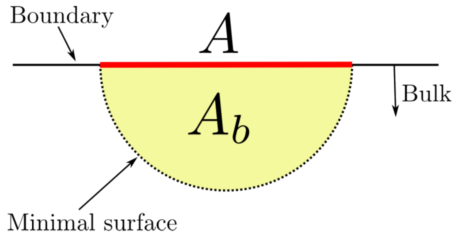

Figure 1: The minimal area associated to the classical bulk term in the holographic entanglement entropy is depicted as the dotted surface that splits the bulk region into two, which are represented by A_b and its complement. The minimal surface is bounded by the spatial region A, which is the boundary associated to the bulk region and thus has one dimension less. Figure taken from [34].

holographic formula for entanglement entropy in quantum CFTs using the $\text{AdS}_{d+2}/\text{CFT}_{d+1}$ correspondence and called it the (holographic) area law relation. R-T considered a bulk spacetime with a static asymptotic boundary that has a timelike Killing field, which enabled them to work in Euclidean spacetime. For a d-dimensional CFT in $\mathbb{R}\times S^d$ space with a $(d-1)$-dimensional boundary $\partial A \in \mathbb{R}^d$ space, R-T defined the HEE of a subsystem A, S_A, as [14]

$$S_A = \frac{\text{Area}(\gamma_A)}{4{G_N}^{(d+2)}}, \tag{2.9}$$

where γ_A is the d-dimensional static minimal surface in the bulk geometry (the bulk gravitational dual) AdS_{d+2}, whose boundary is given by the $(d-1)$-dimensional manifold ∂A, and ${G_N}^{d+2}$ is the $(d+2)$-dimensional Newton constant. Region A lives in the boundary CFT_{d+1}. This formula was actually a speculation from the Bekenstein-Hawking formula (1.6). Figure 2 illustrates the regions associated with the R-T formula. In figure 1 we had $d = 1$, in which case the extremal surface γ_A is the geodesic line of minimal area that lies in the bulk Cauchy slice in AdS_3. A local observer of subsystem A would identify γ_A as a hologram. This identification can also be seen by comparing (2.9) to the Bekenstein-Hawking entropy (1.6). As in the case for entanglement entropy, equation (2.9) also obeys the strong subadditivity condition and it is also defined for finite temperatures T, in which case the density matrix would represent a mixed state. This formula can also be used for asymptotically AdS static spacetimes. In the presence of an event horizon, the minimal surface wraps around it; in this sense, we see that (2.9) is a generalisation of (1.6). This was shown in [35]. The metric of the bulk AdS produces finite contributions to the expansions of G_N and the entanglement entropy. The R-T formula is the leading order result of such an expansion and arises from classical physics in the bulk.

The R-T formula has been proven to work in various applications [36], providing an understanding of the physical realisation of the microstates involved. R-T compared results in the AdS

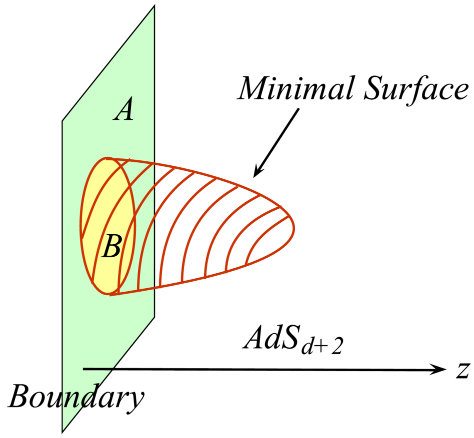

Figure 2: The minimal surface γ_A entering the computation of the holographic entanglement entropy for a CFT_{d+1} ends at the boundary of region B. Figure taken from [16].

space obtained using their formula for a general d with known CFT results in two dimensions, which showed agreement. It was also proven for AdS_3 space in [37, 38]. As soon as the R-T proposal was released, several derivations were provided. Fursaev's derivation in 2006, despite having many evidences [39–41], contained a mistake relating to a misuse of the codimensional 2 surface and its dual. Later in 2013 Lewkowycz and Maldacena provided a proof for a general version of the R-T formula in [42].

In 2007, HRT extended R-T's HEE formula covariantly. Whereas in the R-T proposal, a codimension 2 minimal area γ_A is the surface lying in a spatial slice in the bulk, in HRT's prescription, the codimension 2 minimal surface would be the extremal surface Σ_A that has the same boundary as γ_A but found using a different method. In the R-T proposal, we vary the surface in space since there is no time component in a static spacetime; in this case we would just have that Σ_A is constant on a time slice and is thus the same as γ_A. In the HRT proposal, one needs to vary the surface in both space and time directions in finding Σ_A. The maximin method [43] for finding Σ_A is useful. The procedure is as follows. For each spatial slice in the bulk geometry, find the minimal surface on ∂A. Then among all slices, choose the minimal surface that has the maximum surface area. This would be the extremal surface. Another method would be to use the action principle to find the extremal surface for which any first-order variation in its area produces zero. It was proposed that in the classical limit, the entanglement entropy of an area region A in the bulk CFT is related to the spacelike, codimension 2 extremal surface Σ in the bulk such that $\partial A = \partial \Sigma$ and Σ is homologous to A. Then for a static spacetime, one has $S(\Sigma) = \text{Area}(A)/4G\hbar$, valid to order $O(\hbar^{-1})$ in the CFT [44].

2.4 Thermodynamic Entropy

The second kind of entropy we will now discuss is the coarse-grained entropy. It characterises the ignorance we have about the system. An example of this type of entropy would be the thermodynamic entropy, which exists because we can never know the fine-grained or microscopic details of a system, which are blurred or smoothed over. In this case the density matrix that is associated to the system characterises this ignorance and is manifested as attributing a probability to each state. This is in contrast to the density matrix in the case of entanglement entropy, which characterises the extent of the subsystem's quantum entanglement with the rest of the system. A system in contact with a reservoir at a constant temperature $T = \beta^{-1}$ is described by a Maxwell-Boltzmann density matrix

$$\rho_{\mathrm{MB}} = Z^{-1} e^{-\beta H}, \tag{2.10}$$

where Z is the partition function and H is the Hamiltonian of the system. The thermodynamic entropy is then defined as

$$S_T = -\operatorname{Tr}\left[\rho_{\mathrm{MB}} \ln(\rho_{\mathrm{MB}})\right]. \tag{2.11}$$

It obeys the second law of thermodynamics and increases under unitary time evolution. Another way to obtain the coarse-grained entropy of a system described by the density matrix ρ is to measure a subset of coarse-grained observables A_i and evaluate all possible density matrices $\tilde{\rho}$ such that $\operatorname{Tr}[\tilde{\rho} A_i] = \operatorname{Tr}[\rho A_i]$. Then for each $\tilde{\rho}$, evaluate the fine-grained entropy associated to it, $S_{\mathrm{vN}}(\tilde{\rho})$, and the maximum value among all obtained would be the coarse-grained entropy [1].

By a direct consequence of the definitions, we have that $S_{\mathrm{vN}} \leq S_T$. This conveys the fact that the coarse-grained description of the state encompasses the macroscopic properties of the fine-grained details. In other words, coarse graining the microscopic degrees of freedom of the black hole would produce a single thermal state, where coarse graining means grouping e^{S_T} microscopic states with similar macroscopic behaviour into a single thermodynamic state that has thermodynamic entropy S_T. Therefore the coarse-grained entropy establishes the limit on the degree of entanglement of the interior and exterior of the black hole (the two subsystems). For a pure state, the fine-grained entropy is zero but the coarse-grained entropy is $\log N$ by ergodic theory[1]. An analogy for the two types of entropies can be illustrated with a ball made with many threads – the fine-grained volume of the ball would be represented by the volume of the threads, which is constant, while the coarse-grained volume of the ball would be the volume that is associated to the profile of the ball [45].

[1] This theory states that the probability for the state to be in all N states is the same.

2.5 Semiclassical Generalised Entropy

The Bekenstein-Hawking entropy does not capture the full entropy of the black hole system. For a full expression one has to include the entropy from all quantum fields outside the black hole, which accounts for the matter in the exterior region, such as gravitons, stars, Hawking radiation and vacuum contributions from the quantum fields. Let the state of such matter fields be described by the density matrix ρ. Then the fine-grained entropy of the external matter fields is $S_{\text{outside}} = -\operatorname{tr}[\rho \ln \rho]$. Adding this to our previous entropy (1.6), we obtain the generalised entropy of the event horizon S_{gen}, which represents the total entropy comprising the geometric black hole entropy, given by the area of the horizon Area_{hor}, plus the entanglement entropy of the quantum fields outside the event horizon S_{outside}:

$$S_{\text{gen}} = \frac{\text{Area}_{\text{hor}}}{4\hbar G_N} + S_{\text{outside}}. \tag{2.12}$$

This is also sometimes referred to as the thermodynamic entropy and it is finite. The second term in (2.12) was obvious to Bekenstein, who just added the area term. Including S_{outside} ensures that S_{gen} obeys the second law of thermodynamics when there is emission of Hawking radiation; if the first term in (2.12) decreases, the second term will increase in a way that ensures the increase of total entropy. The generalised entropy has a massive number of degrees of freedom – the black hole at the center of the Milky Way has $S_{\text{gen}} \approx 10^{85}$ and a black hole the size of a proton has $S_{\text{gen}} \approx 10^{40}$ [1].

In the semiclassical description we treat a gravity theory in the semiclassical approximation and we have a background geometry that is classical with quantum fields and perturbative gravitons defined on it. In this approximation we have effective coupling of the strength ${g_{\text{eff}}}^2 \propto \frac{G_N}{{r_s}^2} \propto \frac{1}{S}$, which implies a non-perturbatively small energy spacing between the energy states of the black hole that are of the order $e^{-S} \sim e^{-\frac{1}{g^2{}_{\text{eff}}}}$ [46]. The semiclassical entropy $S_{\text{semi-cl}}(\Sigma)$ of a spatial subregion Σ is defined as the fine-grained entropy of the quantum fields and gravitons on the semiclassical geometry. Referring back to equation (2.12), we identify the semiclassical entropy as the second term in the formula, S_{outside}.

2.6 Quantum Corrections to Semiclassical Generalised Entropy in Black Hole Thermodynamics and in AdS/CFT

We have established that in both the contexts of black hole thermodynamics and AdS/CFT, the entropy of a surface is proportional to its area. At first order in $\hbar$, we have $S_{\text{gen}} = S_{BH}$, i.e. the area term (1.6) is just the leading classical contribution to the black hole entropy. Recall that the underlying mechanism behind the proportionality between entropy and horizon area is the correlations across the event horizon. The fact that black holes radiate imply that they are quantum in nature and hence their entropy cannot be fully described using classical geometry. We thus expect the area theorem to have quantum corrections. In the R-T and HRT prescriptions, a calculation of the semiclassical generalised entropy involved evaluating it on an event horizon slice. In fact, a UV divergence in the entropy is encountered at the horizon that is due to a UV divergent contribution to the partition function. This divergent component is proportional to the area of the boundary ∂A of the subsystem A, which is a standard result in QFT. It arises due to the entanglement among particles that are produced from vacuum excitations across the horizon. The divergence is dealt with using a regulator and is to be offset by the renormalisation of G_N [26], which one can compute using Feynman diagrams for gravitational scattering. The renormalisation of G_N ensures that the entropy is independent of how one divides black hole in the computation. This means that unlike the thermodynamic entropy, the entanglement entropy is not an extensive quantity. The horizon also contains other subleading divergences, such as quantum corrections to the black hole entropy, which are to be absorbed into counterterms. Then for the semicalssical generalised entropy, one should really write: equation (2.12) + counterterms.

Faulkner, Lewkowycz and and Maldacena (FLM) [34] and Engelhardt and Wall (E-W) [44] extended the R-T and HRT prescriptions to include quantum effects. The former computed the first quantum corrections to HRT's formula. By considering order G_N^0 quantum effects in the bulk, FLM computed the quantum corrections to the bulk entanglement entropy between the bulk region A_b and other subregions in the bulk in figure 1, which we denote $S_{\text{bulk-ent}}(A_b)$. This gives the leading classical correction, which we denote $S_q(A)$, to the boundary entanglement entropy (2.8). In FLM's bulk field dependent computation, the bulk is essentially treated as an effective field theory that lives on a fixed background geometry, which parallels the framework in which entropy is computed in a normal QFT [34]. FLM wrote the quantum correction as [34]

$$S(A) = S_{cl}(A) + S_q(A) + O(G_N), \tag{2.13}$$

$$S_q(A) = S_{\text{bulk–ent}}(A_b) + ..., \tag{2.14}$$

where the dots in (2.14) represents one loop correction integrals, such as terms that offset the UV divergence of the bulk entanglement entropy, which ensures the finiteness of S_q [34].

In QFT, the density matrices describing quantum entanglement between subsystems are mostly studied using the tool of path integrals. There are two types of path integral prescriptions corresponding to a time-dependent or time-independent Hamiltonian and states. The former case uses a Euclidean path integral in Euclidean time while the latter case uses a Lorentzian one that involves integrating over Lorentzian time using the Schwinger-Keldysh contour. Higher powers of the density matrix can be calculated using the replica trick, which is a mathematical method to evaluate $-\operatorname{Tr}[\rho \log \rho]$ in a case where the density matrix ρ is not known. In particular, for time-independent states, the replica method can be used to compute the entropy to any order in G_N, thus allowing for the computation of the quantum corrections. By an analytic continuation, the entanglement entropy associated to a $\rho = \rho_A$, where A is a region in the boundary theory, can be formulated from the Renyi entropies [42]. Renyi entropies are a set of entropies defined as [47]

$$S_A^{(n)} = \frac{1}{1-n} \log \operatorname{Tr}_A(\rho_A^{(n)}) \tag{2.15}$$

The n^{th} Renyi entropies are found by computing the complete partition functions about the analytic bulk solutions for each integer n. For order G_N^0, this involves computing the one loop bulk quantum corrections to the classical solutions [34]. In [37], the Renyi entropies and related partition functions of (1+1)-dimensional CFTs were computed via the replica trick. An analytic continuation in n in this case reproduced the R-T formula. One-loop bulk corrections to this was calculated by [48]. In general, the entropy of n copies of the system, which could be the black hole or the radiation, obtained by analytically continuing the computation in n, is given by [42]

$$S = -\partial_n \left(\log Z_n - n \log Z_1\right)|_{n=1} = -\operatorname{Tr}[\rho_A \log \rho_A]. \tag{2.16}$$

Copies of the system can be connected via various topologies that fulfill the boundary conditions. All valid topologies are to be summed over in the calculation of $\operatorname{Tr}(\rho^n)$. The completely connected and disconnected topologies are the two extreme geometry choices. For $n = 1$ the system is disconnected and one just gets the Hawking saddle which gives the dominant contribution to $\operatorname{Tr}\rho$. For $n \neq 1$, one gets the subdominant contributions to the radiation outside the cutoff surface. Consider the case $n = 2$ where we want to compute $\operatorname{Tr}\big(\rho^2\big)$. The disconnected geometry would give the Hawking saddle where we have two disconnected copies of the black hole geometry, while the connected geometry gives the replica wormhole saddle in which the black hole interiors are connected. The latter geometry can be written as a product of two separate copies of $\operatorname{Tr}(\rho)$, indicating that this is a pure state. The replica wormhole saddle thus gives the dominant contribution to $\operatorname{Tr}\big(\rho^2\big)$ in this case [1]. Equation (2.16) is the standard method of computing the von Neumann entropy using the replica trick.

With knowledge of this classical replica trick, one can see that the Bekenstein-Hawking equation (1.6) is the "shortcut" answer obtained from using this method. After performing the replica trick at the quantum level, FLM found the following full expression for formula (2.14), the quantum correction to the entropy [34]

$$S_q = S_{\text{bulk–ent}} + \frac{\delta(\text{Area})}{4G_N} + \langle \Delta S_{\text{W–like}} \rangle + S_{\text{counterterms}}, \tag{2.17}$$

where the first entropy term is of the bulk entanglement, which obeys the strong subadditivity condition individually, the second term corresponds to a shift in the area due to quantum corrections and the third term represents the expectation value of a Wald-like entropy[2]. The final term is included to absorb other subleading divergences so that S_q is finite. At order G_N^0, these corrections are negligible. The last three terms in (2.17) represent the dots in (2.14) and are integrals of local quantities on the minimal surface [34]. Note that if these last three terms are added to the R-T formula (2.9), one still obtains the desired minimal value result. Lastly, S_q also obeys the subadditivity condition. At the lowest boundary order, or bulk order $O(\hbar^0)$, we therefore have [34]:

$$S_{\text{gen}}(A) = \frac{\text{Area}_{\text{hor}}}{4\hbar G_N} + S_{\text{bulk-ent}} + \text{counterterms}, \tag{2.18}$$

where $S_{\text{bulk-ent}}$ is the bulk entanglement entropy between region A_b and the remaining bulk. Since FLM only considered the case where the total state is pure, we have that $S_{\text{outside}} = S_{\text{bulk-ent}}$. We see a similarity between equation (2.12) counterterms and equation (2.18). Whereas in the former the generalised entropy was associated to the minimal area of an event horizon slice, the latter is associated to a surface extremised with respect to space and time.

2.7 Quantum Extremal Surfaces and Gravitational Fine-Grained Entropy

E-W in [44] pointed out that the black hole entropy is not always represented by the semiclassical generalised entropy of the horizon because in general, S_{gen} can be defined for an extensive class of arbitrary surfaces; other surfaces with a statistical interpretation can also be chosen. We now review E-W's argument. Consider a static codimension 2 surface E that divides a spatial surface Σ into two spatial regions denoted by $\text{Ext}(E)$ and $\text{Int}(E)$, as shown in figure 3. The former is contained outside E and the latter inside E. Define $S_{\text{out}}(E)$ and $S_{\text{in}}(E)$ to be the entanglement entropy in $\text{Ext}(E)$ and $\text{Int}(E)$ respectively. If S_{gen} is defined by (2.12) but evaluated on E rather than on a horizon spatial slice, then by unitarity, any Σ that passes through E would define the same entropy, rendering the outcome that $S_{\text{gen}}(E)$ is independent of the choice of E [44]. For a pure state, we have $S_{\text{out}}(E) = S_{\text{in}}$. But for a mixed state, one has to choose the side with region corresponding to the boundary entropy being computed; one is not allowed to choose the other side for the reason of locality violation [44]. Note that entanglement entropy is defined for both pure and mixed states.

Recall that in FLM's prescription, the area was first extremised before S_{outside} was added. Above we saw a setback in this formulation of choosing arbitrary surfaces. Noticing this, E-W proposed the following modification to computing the entropy of a holographic boundary region:

[2]See [34] and [49] for an explanation of the Wald-like entropy.

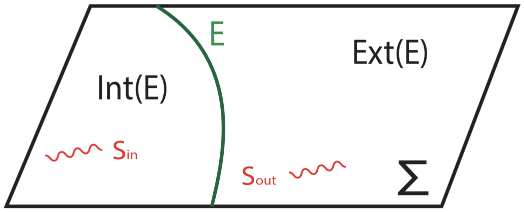

Figure 3: The setup used to define the generalised entropy on an arbitrary surface and to argue that the generalised entropy does not always represent the black hole entropy. The spatial surface Σ in AdS space is split into two by a codimension 2 static surface E into the regions denoted by $\text{Ext}(E)$ and $\text{Int}(E)$. Entropies S_{out} and S_{in} represent the entanglement entropy in the former and latter regions respectively. For a pure state, the generalised entropy computed on either region is the same while for a mixed state it depends on the side chosen to compute the generalised entropy on. Figure taken from [44].

extremise the total generalised entropy, which is defined on a *quantum extremal surface* (QES), a quantum corrected R-T or HRT surface. Despite its name, the QES is a classical surface on the spacetime geometry. E-W also extended FLM's formulation for non-static extremal surfaces. E-W showed that at order $O(\hbar^0)$ of the first quantum corrections, their prescription is equivalent to FLM's. Since the latter was only proved at this order, the equivalence does not tell us which prescription is correct at higher orders; E-W argued that at higher, even infinite, order in $\hbar$, FLM's prescription does not compute the entanglement entropy as it changes under boundary unitary transformations. In addition, E-W produced a few theorems for their prescription which do not hold for FLM's. These conclusions were viewed by E-W as evidence that their prescription is accurate and that QESs are physical realisations of quantum spacetimes [44].

The QES is defined as the surface on which the fine-grained entropy of the black hole has minimal extremal value; on the QES, the gravitational generalised entropy is minimised in the spatial direction but maximised in the time direction. In the case where more than one extremal surfaces are found, one should choose the global minimum. The maximin construction [43] is one way to carry this out. In finding the QES, one starts from a surface at a time on the cutoff surface outside the black hole and move inwards, even crossing the horizon to consider surfaces in the interior. If the QES lies totally in the interior of the black hole, then there will no area contribution to the total entropy. This prescription implies that the fine-grained entropies of black holes is dependent on the black hole interior geometry and thus that black holes that have the same exterior characteristics but different interiors would have different entropies.

For quantum systems coupled to gravity, the extremal value of the semiclassical generalised entropy obtained by choosing the surface which minimises its value gives the gravitational fine-grained entropy of the black hole. According to the QES proposal, this entropy is [44]

$$S(R) = \min_{Q} \left\{ \text{ext}_Q \left[\frac{\text{Area}(Q)}{4G_N} + S_{\text{semi-cl}}(\Sigma_Q) \right] \right\}, \tag{2.19}$$

where surface Q is the QES, a codimension 2 bulk surface homologous to the boundary region R on which the entropy is calculated on, $S_{\text{semi-cl}}$ is the semiclassical fine-grained entropy on a partial Cauchy surface extending from Q to R on the asymptotic boundary and represents the entropy of the quantum matter fields and gravitons in the fixed background geometry outside the black hole, and Σ_Q is the region bounded by the QES Q and the AdS boundary. The extremisation is with respect to the surface Q and the minimisation is with respect to the surface Q with the minimal generalised entropy. For states described using path integrals, formula (2.19) can be derived using a similar method to that of Gibbons-Hawking (see section 5.1). In later sections we will see that this gravitational entropy formula was successfully applied to compute the fine-grained entropy of an evaporating black hole, which is a gravitational system. This implies that Hawking was just not using the right formula to compute the entropy, which ultimately led him to the conclusion that black hole evaporation violates unitarity.

3 The Information Loss Problem

As a black hole radiates away energy, its mass decreases and eventually it will evaporate away completely. Consider a classical static black hole formed from the stellar collapse of some pure quantum state of matter, which in the process produces thermal Hawking radiation via evaporation. The exterior geometry of the black hole is almost stationary whereas the interior geometry is extended along a direction whilst shrinking to zero size in the angular direction. The angular direction and event horizon becomes the singularity at the end of the evaporation and outside remains a smooth spacetime containing thermodynamic radiation. It appears that in this process an initially pure state has been converted into a final thermal mixed state. Similarly, one can consider adding infalling matter to an evaporating black hole to balance the outgoing radiation such that it stays in equilibrium. In both scenarios, the evolution appears to violate unitarity.

The concept that black holes are described by microscopic degrees of freedom have been applied to the process of black hole evaporation in past suggestions and attempts to solve the information loss problem [50–53]. Many of the proposed solutions fell into the following main areas: 1. The black hole does not completely evaporate and information ends up encoded in a remnant [54, 55]; 2. Information is lost; 3. The end state of the radiation is pure, not mixed, and the information escapes the black hole through correlations among the outgoing radiation quanta; 4. The information escapes together with the bulk of the radiation and this gives an S-matrix that maps the in states to the out states [31, 56–58].

The first possibility is realised when the black hole stops evaporating when its mass and size reaches Planck scale and a remnant is leftover. But in this case the remnants would have to have unbounded degeneracy or be highly entropic if one started with an arbitrarily large black hole and the entanglement entropy of the remnant would exceed the Bekenstein-Hawking entropy. This is in contrast to expecting only a finite number of states; having an infinite number of states would revoke the interpretation of black holes as being described by microscopic degrees of freedom. Possibility two, that in the process of black hole formation and evaporation, quantum information is permanently lost and the von Neumann entropy of the universe increases, was advocated by Hawking [10, 59]. Computations performed by Hartle and Hawking suggested that information is lost via Hawking radiation. In this case the mixed state of radiation has entropy of the order of the initial black hole horizon area in Planck units, and there is no S-matrix to describe the transition. Possibility three suggests that the final state of the evaporation is a pure state of the radiation field, with a complete basis of dimensionality of the order of the exponential of the Bekenstein-Hawking entropy. Any small subsystem considered would appear thermal. In this possibility, the outgoing radiation and thus information that escapes are independent of the details of the initial state of infalling matter. However, quantum gravity as an effective field theory breaks down at small curvature.

The fourth and last possibility was investigated by Page in [56], which was eliminated as a potential solution by Giddings and Nelson in [52, 60]. Hawking's original argument for information loss claimed that the semiclassical approximation breaks down when the black hole reaches Planck mass. Giddings and Nelson studied two dimensional dilatonic black holes with N minimally coupled fields and showed that all correlation functions and the density matrix of the Hawking radiation can be exactly computed. In other words, the formation and evaporation processes for this setup can be explained until the breakdown of the semiclassical approximation; in [52], Giddings showed that working perturbatively in $1/N$ reveals that information escapes only after the Planck scale for a four dimensional black hole. It was concluded that quantum dilaton gravity would have to be futher quantised in order to obtain deeper understandings. However, considering the works of Giddings and Nelson, Page in [50] found that even a perturbative analysis would not recover the information. He argued that information does escape, but at a very slow rate or with long intervals, which would require an unphysical number of measurements to detect – even an analysis limited to perturbations in m_{planck}/M cannot recover the information [50]. We will explore the ideas of Page on the entropies involved in black hole evaporation in section 4 through the introduction of the Page curve.

3.1 Paradoxical Spacetime Geometries for the Black Hole

One could argue the following perspective of the information loss problem [1]: imagine dividing the spacetime of an evaporating black hole into two aymptotic regions – a parent universe and a baby universe. The latter is formed from the infalling matter and could be some future region that is near the singularity. During evaporation the two regions are still connected and the infalling matter creates the singularity. After the black hole has completely evaporated, the two regions will disconnect but still be highly entangled with each other. One will then be left with a baby universe that is the singularity, entangled with a parent universe that is filled with the previously escaped radiation. It now appears, to an external observer of the parent universe, that information is lost. In a semiclassical gravity theory, the argument is that an understanding of the interior has no bearing on the way that the information is lost. In fact, an observer of the whole universe, including both baby and parent universes, can conclude that information is not lost by using the wavefunction of the baby universe, which was shown in [61].

Consider another spacetime geometry – the vacuum solution corresponding to the two-sided maximally extended Schwarzschild geometry, which represents two black holes by two exteriors that are sharing an interior. The two black holes asymptotically looks like two disconnected $\mathbb{R}^3$ connected by a wormhole, which is called the Einstein-Rosen bridge. Now we consider a geometry called "bags of gold" [62] that is closely related to the geometry above. This object has the second $\mathbb{R}^3$ replaced by, say, an $\mathbb{S}^3$, which represents a highly entropic closed universe since it can contain a lot of matter [63]. This classical geometry is shown in figure 4. We have the Schwarzschild solution in the throat (neck) region, which is asymptotically flat and narrow. The argument is that this geometry looks like a black hole with an entropy given by the horizon area of the throat when viewed from the outside, which is in contrast to the possibility of having an arbitrarily larger entropy in the closed universe. This object has been used to argue that the area entropy does not measure the entropy on some spatial slice inside a black hole [1]. We will revisit both geometries at the end of section 6.

3.2 The Central Dogma

Considering the mechanics and thermodynamics of black holes, Almheiri *et al.* in [1] advocated for regarding the black hole as seen by an external observer as an ordinary quantum system with $S = \text{Area}/4G_N$ degrees of freedom, or qubits, and which obeys the laws of thermodynamics and evolves unitarily under time evolution. For this advocation they coined the name the "*central dogma*". One can visualise this by imagining a reflecting circular surface (which acts as a reflecting boundary condition) surrounding the black hole and the environment around it up to some cutoff surface a few r_s away from the black hole, and regard the system inside

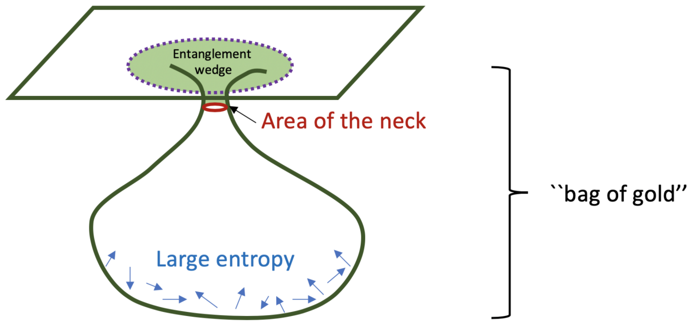

Figure 4: The "bags of gold" spatial geometry consists of an asymptotically flat space connected to a highly entropic closed universe by a narrow neck or throat. As viewed by an observer on the outside, it is argued that this geometry describes a black hole with an entropy computed by the area formula, which does not account for the entropy of the matter contained in the closed universe. In section 6, this is resolved by considering entanglement wedges; we will see that the entropy area formula computes the fine-grained entropy of only the exterior and not the whole interior. Figure taken from [1].

the cutoff surface as a quantum system. For example, this could be a set of qubits that has interactions with the exterior described by a strongly interacting unitary Hamiltonian that can generate a chaotic gravity evolution in time [1]. In AdS/CFT the cutoff surface is represented by a boundary coupled to degrees of freedoms living outside the cutoff surface. Another analogy would be to replace the black hole as viewed from the outside with a burning piece of coal, in the sense that it is just an ordinary quantum system [63], with the crucial difference of the presence of a horizon in the former. A piece of coal in an initially pure state, through burning, emits radiation quanta that are entangled with other quantum degrees of freedom in the coal. Whilst quantum degrees of freedom inside the coal can affect the radiation modes at late times, the same is not true for the interior modes inside black holes due to causality.

Almheiri *et al.* pointed out that the central dogma is not implied in the gravity description since we do not know how to derive the S degrees of freedom and the unitary Hamiltonian in this description. The central dogma thus has to be treated as an assumption of a full theory of quantum gravity and not a theorem; the assumption is that the degrees of freedom can describe arbitrarily precise measurements performed outside the black hole. Black hole thermodynamics and fine-grained entropies are characteristics of a full theory of gravity and do not rest on the validity of the central dogma [1]. The statement of the central dogma made above only describes the exterior of the black hole and not its interior. The information loss problem, if true, would imply that the central dogma is false.

Counting the microstates of supersymmetric black holes using strings and D-branes have provided evidence for the central dogma. These have been seen to reproduce the area formula and its corrections [18,64]. The BFSS matrix model and BFSS conjecture related to the 11-dimensional S-matrix [65] provides another piece of evidence. Lastly, the AdS/CFT correspondence is useful in studying the microstates of an AdS black hole through studying its thermal state in a unitary CFT [42,66].

4 Page Curve for the Black Hole

A quantitative form of the information loss problem is encapsulated in the Page curve, which is a plot of the entanglement entropy of the radiation as a function of time. We will now consider the evolution of the fine-grained and coarse-grained entropies of the radiation and the black hole horizon during the formation and evaporation process for a black hole formed from the collapse of a pure state.

At the beginning, the coarse-grained or thermodynamic entropy of the horizon is given by the area of the black hole, which starts at zero. Before matter starts falling into the black hole, its area increases exponentially during black hole formation. The area reaches a maximum just before infalling matter reaches the black hole and the evaporation process starts [31]. During evaporation, the area decreases linearly with time. Thus, the thermodynamic entropy decreases steadily until the black hole has completely evaporated at which point the thermodynamic entropy is zero. This evolution is depicted in figure 5 as the decreasing orange curve. In contrast, at the start of the evaporation process, no Hawking radiation has been emitted yet so its thermodynamic entropy starts at zero. As the black hole evaporates, more Hawking radiation is emitted so the thermodynamic entropy of the outgoing radiation increases steadily. This evolution is represented by the increasing green line in figure 5. The gradient of this line is constant as the radiation is characterised by the Hawking temperature. The entropy eventually reaches a maximum, at which point the black hole has completely evaporated and the entropy stays constant from then on. This is Hawking's calculation for the entropy of outgoing radiation. Note that we have already encountered a problem – at the point where the thermodynamic entropy of the outgoing radiation exceeds that of the black hole horizon, there is insufficient black hole microscopic degrees of freedom to entangle with the entropy of outgoing radiation.

Now we consider the fine-grained entropies, or the entanglement entropies. Since the total state at the start is pure, by unitarity, we must have that the black hole quantum system and radiation are described by pure states as well and also be entangled with each other. At the start of the evaporation process when the black hole horizon still has a small area, the black hole quantum system is still described by a pure state. Hence, the outgoing radiation, which is entangled with

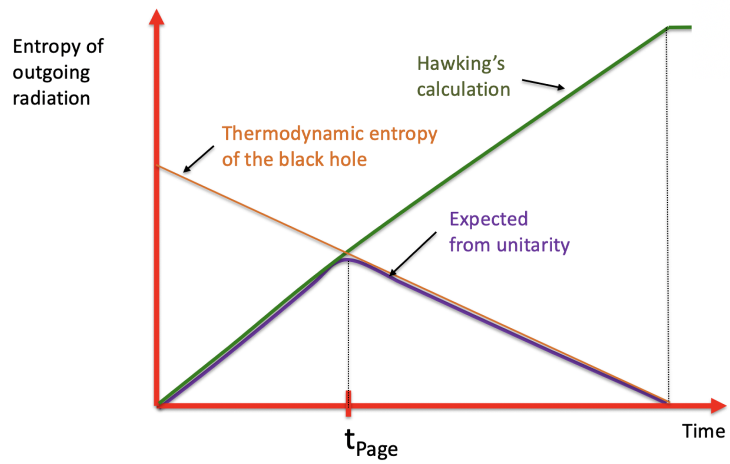

Figure 5: The themodynamic and fine-grained entropies of the outgoing radiation and black hole involved in the evaporation process of a black hole are schematically illustrated. The decreasing contribution in yellow represents the thermodynamic entropy of the black hole given by the area of its horizon. It starts to decrease as the black hole evaporates and its area shrinks. The increasing contribution in green represents the thermodynamic entropy of the outgoing radiation, which increases as more radiation is emitted and reaches a maximum value at the point of complete evaporation and remains constant after that. By unitarity, the fine-grained entropy of radiation cannot exceed the thermodynamic entropy of the black hole. Thus the entanglement entropy follows the Page curve, represented by the purple curve. The point of inflection occurs at the Page time, when the fine-grained entropy of radiation is equal to the thermodynamic entropy of the black hole and this quantity characterises the Page curve. Figure taken from [1].

the black hole, will also be described by a pure state and thus be approximately thermal during this time. Note that the fine-grained entropy of the radiation is that of the radiation that is outside the cutoff surface and neglects vacuum contributions to the entropy. The fine-grained entropies of both the black hole and outgoing radiation are equal. This entropy starts at zero and increases as the black hole evaporates and outgoing radiation enters the region outside the cutoff surface. However, this entropy cannot exceed the thermodynamic entropy of the black hole, which is given by the horizon area, or bounded by the dimension of the Hilbert space of the black hole horizon. Therefore we expect the fine-grained entropy of radiation to decrease back to zero as the horizon area shrinks back to zero at the end of the evaporation process. The point where the fine-grained entropy of radiation is equal to the thermodynamic entropy of the black hole before it starts decreasing is called the Page time. At this point the semiclassical gravity description is still applicable. Page argued that the fine-grained entropy of the black hole would follow the thermodynamic entropy of the outgoing radiation at the early stages and overall, would follow the Page curve, which is represented by the purple line in figure 5. Note that the exact shape of the curves would depend on the details of the black hole and radiation systems. Figure 5 just shows the approximate general feature.

4.1 Quantifying the Page Curve

In this section a brief quantification of the Page curve using Page's theorem is given. Page considered the black hole (or equivalently, the radiation in the black hole that has not yet been emitted) as one quantum subsystem BH and the radiation outside the imaginary sphere as another quantum subsystem R. The bipartite Hilbert space for the outgoing radiation can be written as $\mathcal{H}_{out} = \mathcal{H}_R \otimes \mathcal{H}_{BH}$. The total system is always in some pure state throughtout the evaporation process so we always have that $S_R^{\text{vN}} = S_{BH}^{\text{vN}}$. Each subsystem has an associated density matrix, which when traced over gives a generically mixed state of the other subsystem. Page studied the information encoded in the radiation subsystem throughout the evaporation process by studying the reduced density matrices at different times. Page confirmed that the fine-grained entropy of outgoing radiation, S_R^{vN}, follows the Page curve. It increases as the black hole evaporates, reaching a maximum value of $S_R^{\text{vN}} \approx S_0/2$ at a time $t_{\text{Page}} \approx t_{evap}/2$, where $S_0 = 4\pi M^2$ is the thermodynamic entropy of the black hole at the start of evaporation. It then decreases back to zero upon complete evaporation. Note that t_{Page} and S_R^{vN} are not exactly half of t_{evap} and S_0 respectively as they depend on the evaporation process, which is irreversible. Page performed numerical calculations for a large Schwarzschild black hole that started in a pure state and which evaporated into photons and gravitons in [67], where he found a value of $S_R^{\text{vN}} \approx 0.6 S_0$ at $t_{\text{Page}} \approx 0.54 t_{\text{evap}}$. Page derived the numerical equations governing the time dependence of the fine-grained entropy of outgoing radiation for both initially pure and mixed states of the same black hole.

We now follow Harlow's [11] approximate derivation for Page's numerical equation using Page's theorem. Page's theorem [68] states that a pure state randomly chosen is almost maximally entangled[3] if $D_R/D_{BH} \ll 1$, where $D_{R(BH)}$ is the dimension of $\mathcal{H}_{R(BH)}$. A random pure state is given by $|\psi(U)\rangle \equiv U\,|\psi_0\rangle$, where $|\psi_0\rangle$ is some particular state and U is a random unitary matrix. At early times we have $S_R^{\text{coarse}} = \log D_R < S_{BH}^{\text{coarse}} = \log D_{BH}$. By approximating the outgoing radiation as a (1+1)-dimensional photon gas, we can write for early times: $S_R^{\text{coarse}} \propto tT$, where t is the time that has elapsed since the start of evaporation and T is the temperature of the black hole [11]. While subsystem R has an order of M^{-1} characterised by the Hawking temperature, subsystem BH has order M^2 characterised by the thermodynamic entropy of radiation (1.6). Hence for $t \ll M^3$, we have: $S_R^{\text{coarse}} \ll S_{BH}^{\text{coarse}}$. Applying Page's theorem, we see that the fine-grained and thermodynamic entropies of the radiation are approximately equal at early

[3] This is characterised by a very small deviation of ρ_R from the maximally mixed state. A maximally mixed state is characterised by having a pure total density matrix, and a reduced density matrix that is proportional to the identity operator on the respective Hilbert space.

times [11]:

$$S_R^{\rm vN} \approx S_R^{\rm coarse} \quad t \ll M^3. \tag{4.1}$$

This explains the qualitative behaviour at the start of the evaporation process: the fine-grained entropy of radiation follows the thermodynamic entropy of radiation, reflecting the fact that the outgoing radiation is entangled with the black hole quantum system. At early times, there are many black hole degrees of freedom available for the radiation to form a pure state with. This initial behaviour is not in contradiction with the central dogma. Then when the fine-grained/thermodynamic entropy of the radiation is equal to the thermodynamic entropy of the black hole at the Page time, we have $S_R^{\rm coarse} \approx fS_0$, where f is a fraction of order one and S_0 is the initial thermodynamic entropy of radiation. Similarly applying Page's theorem (in reverse) after the Page time, one obtains $S_{BH}^{\rm vN} \approx S_{BH}^{coarse}$. Now we have that the outgoing radiation is entangled with radiation that was emitted at earlier time. After the Page time, the outgoing radiation can no longer form a pure state together with the black hole since the former region is greater than the latter. Past the Page time, one can say that the information of the initial state starts to escape. If the central dogma is true, we expect the Page curve to be followed. Overall, the approximate qualitative form of the Page curve is represented by [11, 67]:

$$S_R \approx S_{BH}^{coarse} \propto S_0 \left(1 - \frac{t}{t_{evap}}\right)^{2/3} \quad t_{Page} < t < t_{evap}. \tag{4.2}$$

In order to obtain a more accurate form and the exact values of $t_{\rm Page}$ and the entropy at that time, one has to take into account details such as greybody factors and particle helicities, which were done by Page in [67].

4.2 Gravitational Fine-Grained Entropy Formula and the Page Curve for the Black Hole

We will now apply the gravitational fine-grained formula (2.19) to the case of an evaporating black hole. We introduce the term *entanglement wedge* (EW) [69–71] which is the bulk region of the semiclassical spacetime that is described by or encoded in the given boundary system such as the radiation or the black hole. It is defined at a moment in time and depends on time. For a boundary region B, an EW is the bulk region, or domain of dependence, bounded by the QES Q (that is homologous to B) and the boundary region B, that has the minimal generalised entropy (2.19) by the maximin prescription, and has $Q \cup B = \partial Q$ [72]. From here we will call the QES that gives the minimal generalised entropy the quantum R-T surface.

Penington [69] and Almheiri, Engelhardt, Marolf and Maxfield (AEMM) [71] embedded the information loss problem in AdS/CFT by coupling a AdS black hole with absorbing boundary conditions to an auxiliary system $\mathcal{H}_{\rm rad}$ that absorbs the Hawking radiation. An example was done in [2] where an auxiliary system that is half of Minkowski space and containing no gravity is attached to the boundary of AdS, through which radiation into matter fields is allowed to pass through. To avoid backreaction, the auxiliary system is assumed to be a large and holographic. In this prescription one has two holographic boundary systems: $\mathcal{H}_{\text{CFT}}$ and $\mathcal{H}_{\text{rad}}$. The HEE of the black hole is then measuring the entanglement between these two systems. By the R-T formula, this HEE would be the generalised entropy of the quantum R-T surface associated to the boundary CFT.

The exact location of the QES depends on the amount of radiation that has escaped, and thus on the time along the cutoff surface at which we decide to calculate the entropy. To locate the QES, we travel a scrambling time $\beta/2\pi \log S_{\text{BH}}$[4] into the past along the cutoff surface and then send a light ray into the black hole. The point where this light ray crosses the horizon is where the surface is located close to [1]. The QES would be constant with respect to infinitesimal directional deformations. For example, consider a surface Q on the horizon and moving it inwards along an ingoing null direction, which would decrease its area. If one moves Q too inwards to the point where it includes more black hole interior degrees of freedom (that are entangled with the outgoing radiation) than those that are purifying the outgoing radiation, then this would cause the bulk entropy $S_{\text{semi-cl}}$ to increase. The location where Q accounts for just enough black hole degrees of freedom to purify all the outgoing radiation is where the QES lies. Any infinitesimal variations to its direction would change both its area and entropy in such a way that they offset each other [1].

An obvious QES homologous to the entire boundary in any Cauchy slice through the interior at any time is the empty surface. Before the Page time, this QES gives the minimal generalised entropy and is thus the quantum R-T surface during this time. This vanishing surface is pictured in figure 6. In such cases, the area term in (2.19) is zero and the generalised fine-grained entropy, or the bulk entanglement entropy S_{rad}, measures the entanglement between the outgoing radiation in $\mathcal{H}_{\text{rad}}$ and the interior black hole degrees of freedom in $\mathcal{H}_{\text{CFT}}$. In other words, the interior of the black hole is entirely encoded in the boundary $\mathcal{H}_{\text{CFT}}$, or in the EW of the CFT. During this time, information has not yet escaped the black hole interior [69]. The reduced density matrix of $\mathcal{H}_{\text{rad}}$ is thermal. Note that since the entropy depends on the interior geometry of the black hole and the area term is zero, if we started in a pure state then we have that the initial entropy is zero as well. As the black hole evaporates, the generalised entropy increases as the outgoing radiation modes are entangled with the previously infalling matter states. This contribution of

[4]This is the timescale at which information gets encoded in the black hole interior and is short compared to the evaporation time $r_s S_{\text{BH}}$. The scrambling time can be derived in Eddington-Finkelstein coordinates as the infalling time of the QES [69]. In later sections as will see that after a scrambling time, the information will be contained in the EW of the radiation and is thus recoverable by acting on the radiation with complex operators. From the point of view of the boundary, the information lives in the past.

entropy is called $S_{\text{gen}} = S_{\text{rad}}$ of the vanishing surface and is depicted as the increasing green line in figure 7. The growth of this entropy parallels the thermodynamic entropy of outgoing radiation.

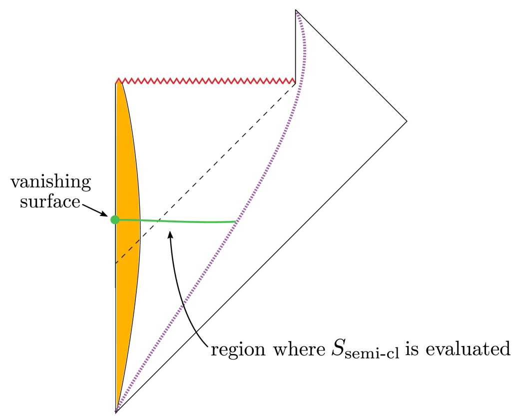

Figure 6: An empty surface homologous to the boundary that gives the minimal generalised entropy before the Page time is labelled as the vanishing surface. The generalised entropy is characterised by the bulk entanglement entropy of the radiation in the region $\mathcal{H}_{\text{rad}}$. Figure taken from [1].

Now we consider any Cauchy slice after the Page time and apply the maximin method. One would be able to find a non-empty surface homologous to the entire boundary that lies outside the event horizon with an area larger than that of the horizon [69]. This surface in $\mathcal{H}_{\text{rad}}$ contains outgoing radiation that encodes the interior degrees of freedom. The generalised entropy computed with this surface measures the degree of entanglement between $\mathcal{H}_{\text{rad}}$ and $\mathcal{H}_{\text{CFT}}$ and is given by the Bekenstein-Hawking entropy $S_{\text{BH}} = A_{\text{hor}}/4G_N$, which is now smaller than S_{rad}, the entropy computed with the vanishing surface. Therefore, this non-empty QES is the quantum R-T surface after the Page time. During this time, for increasing times on the cutoff surface, the QES moves along a spatial slice outwards towards the horizon. As the black hole evaporates and the area of the horizon decreases, the QES produces a decreasing $S_{\text{gen}} = S_{\text{BH}}$. This is depicted as the decreasing yellow line in figure 7.

In fact, a non-empty QES already exists a short time after the emission of outgoing radiation starts and before the escape of radiation. This surface lies a scrambling time inside the black hole near the horizon. However, it is not the quantum R-T surface since its generalised entropy exceeds that given by the vanishing surface. But at the Page time, the black hole undergoes a phase transition and this is when this QES becomes the quantum R-T surface [69].

To summarise, we have three types of quantum R-T surfaces that gives the minimum of the

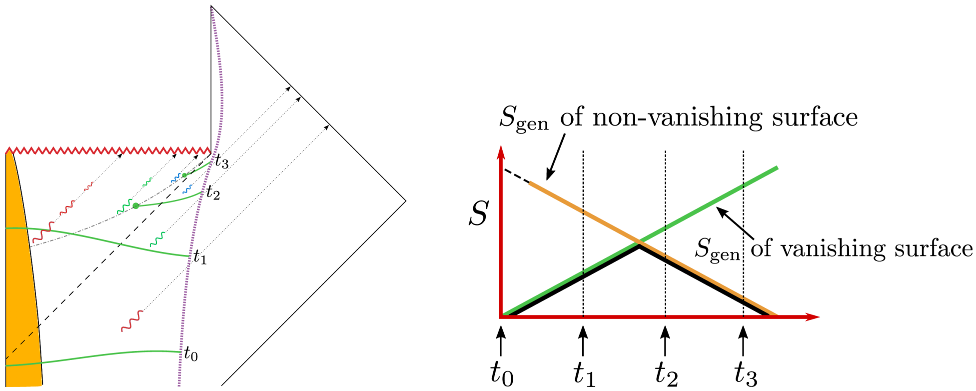

Figure 7: The green line represents the increasing generalised entropy of the black hole region with time as more outgoing radiation gets emitted. This rising entropy contribution is given by the vanishing surface. There is entanglement among the escaping modes that are represented by the same colour. The yellow line represents the decreasing generalised entropy of the non-vanishing extremal surface that appears near the event horizon in the interior a scrambling time after the black hole forms. For increasing times on the cutoff surface, we get a decreasing surface which moves along a spatial direction outwards towards the horizon. This gives a decreasing generalised entropy (since the black hole area is shrinking). The decreasing thermodynamic entropy of the black hole is represented by the dotted line which is overlapped by the yellow line. This is in contrast to the growing entropy of radiation. The fine-grained entropy of the black hole follows the Page curve in black. This curve is a result of transitioning between the contributions from the vanishing surface and the non-vanishing extremal surface. Figure taken from [1].

generalised entropy S_{gen} across all QESs at different times: the vanishing surface before the Page time, the non-empty surface inside the horizon at the Page time and the non-empty surface outside the horizon after the Page time. Using the R-T formula, the generalised entropy is initially the fine-grained entropy of the black hole S_{rad} until the time when the entropy contribution from the non-vanishing surface, S_{BH}, becomes lesser than the former. Overall, the generalised entropy ends up following the Page curve, represented by the black line in figure 7. The transition in the Page curve from increasing to decreasing entropy at the Page time corresponds to when the outgoing radiation starts to encode most of the black hole interior. Note that if we were still using the original classical R-T surface instead of a QES in the computation of S_{rad}, then we will always find the whole interior encoded in $\mathcal{H}_{\text{CFT}}$ and not $\mathcal{H}_{\text{rad}}$, which suggests that information never escapes the black hole interior. We have shown how the black hole gravitational fine-grained entropy follows the Page curve, which indicates that evaporation as a unitary process.

5 Islands and the Page Curve for Hawking Radiation

The EWs and Page curve considered in the previous section concerns the black hole entropy and not the entropy of Hawking radiation, the latter of which is involved with the information loss problem. We would like to account for the increasing entropy of radiation $S_{\text{semi-cl}}(\Sigma_{\text{Rad}})$, where Σ_{Rad} is the exterior region outside the cutoff surface. After complete evaporation, all the radiation resides outside the AdS space, where (quantum) gravity effects are negigible. (It could therefore be imagined that it is collected into a quantum computer far away from the black hole [1].) We are unable to describe this final radiation state, which we have obtained by gravity, by a density matrix since we do not have the information required to compute its individual matrix elements. In fact if we were to use the standard $S = -\operatorname{Tr}[\rho \log \rho]$ as for an ordinary QFT without gravity, we would obtain Hawking's answer.

We have seen with holographic entropy formula (2.19) how to compute the entropy of the outgoing radiation that is entangled with the interior of a black hole in gravity. Despite residing in a region with no gravity, it has been argued that the final radiation state could and should also be computed using formula (2.19) since it was obtained using gravity. In other words, the argument is that the same formula can be applied in the following two different regions: the interior of the cutoff surface in AdS space and some region without gravity outside the cutoff surface containing a quantum system [1]. Now we will apply formula (2.19) to a region containing no black holes – the region behind the horizon, which is not included in either of Σ_{Rad} or Σ_{Bulk}. This motivated the addition of "*islands*", Σ_{Island}, to the existing formula (2.19) and was first suggested by Penington and AEMM. Considering this addition alongside the QES prescription allows one to find the island regions. We will see that computing the Page curve using the revised formula (with islands) turns out to compute the Page curve for the radiation and not the black hole.

The idea to add contributions from islands to the computation of the fine-grained entropy for a system coupled to gravity, called the island conjecture, was realised by Almheiri, Mahajan, Maldacena and Zhao (AMMZ) [73] in 2019. The prescription computes the von Neumann entropy of a region A in a QFT by extremising the generalised entropy with respect to islands and then minimising with respect to all islands with extrema entropy. This approach was termed 'doubly holographic'. In formula (2.19), we considered connected QESs Q. In the island prescription, disconnected Qs are allowed. These are manifested as regions outside the black hole far away that contain entangled matter. An island is defined as any disconnected codimension 1 region found by extremising the entropy; the boundary of an island is the QES. The causal diamond of an island region is the causal domain of dependence of the region that the entropy is calculated in and is a part of the EW of the radiation [1]. Recall that outgoing radiation is entangled with the interior degrees of freedom. The island prescription thus allows one to increase the area of the boundary by including the black hole interior, which corresponds to including the disconnected regions, and simultaneously decrease the semiclassical entropy contribution. This is achieved by adding an area term, $S_{\text{semi-cl}}(\Sigma_{\text{Island}})$, corresponding to the boundary area of the island [1]. An island region with centre at the origin in shown in figure 8.

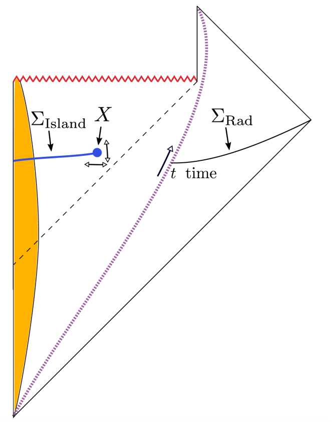

Figure 8: At late times, there is an island EW that is disconnected from the region Σ_{Rad} that contains the outgoing radiation. When calculating the semiclassical fine-grained entropy in the latter region, one finds that entropy in Σ_{Island} contributes. In two dimensions, the EW of the black hole describes only a part of the interior and the rest is described by the EW of the radiation, which is disconnected from the interior. In three dimensions, this disconnection is not manifested in the same way; a connection living in another dimension connects the island with the CFT interval in Σ_{Rad} on which we are computing the fine-grained entropy. This was shown in [73] by considering a two-dimensional black hole in a gravity theory that is coupled to a matter CFT_2. Figure taken from [1].

The island prescription states that the fine-grained entropy of the full exact quantum state of the radiation, $S_{\text{rad}}(R)$, is given by the "*island formula*" [1]

$$S_{\text{Rad}}(R) = \min_Q \left\{ \text{ext}_Q \left[\frac{A(Q)}{4G_N} + S_{\text{semi-cl}} \left[\Sigma_{\text{Rad}} \cup \Sigma_{\text{Island}} \right] \right] \right\}, \tag{5.1}$$

up to subleading corrections. We vary the area A of the boundary of the island to find an extremal value for the RHS of (5.1), then find the minimal value across all extremal positions as well as islands. Here Σ_{Rad} represents the entire region outside the cutoff surface to infinity, which is where all the outgoing radiation that has escaped from the black hole interior resides, and Σ_{Island} is some island region of the bulk inside the black hole. There can be any number of such island regions. Region Σ_{Rad} is described by the density matrix $\rho_{\Sigma_{\text{Rad}}}$ in the full theory of quantum gravity and a state in the region $\Sigma_{\text{Rad}} \cup \Sigma_{\text{Island}}$ is described by the density matrix $\tilde{\rho}_{\Sigma_{\text{Rad}} \cup \Sigma_{\text{Island}}}$. Such states can be evaluated by performing a semiclassical path integral on the Euclidean black hole saddle [2]. $S_{\text{semi-cl}} \left[\Sigma_{\text{Rad}} \cup \Sigma_{\text{Island}} \right]$ is the fine-grained entropy of the bulk, which is the entropy of the island plus the outgoing radiation in the semiclassical description. It was emphasised in [1] that the LHS is the full entropy for the full exact quantum state of radiation while the RHS is the quantum state of radiation in the semiclassical description; both states are different. The former is the radiation state that one would wish to obtain non-perturbatively from a complete theory of gravity. We see that formula (5.1) is a generalisation of (2.19); both formulas are based off the same principles of minimising and extremising. Whilst the latter was motivated as the generalised entropy of a black hole, the former has no relation to a black hole and concerns just the radiation [1]. In a case where the quantum system is entangled with fields inside a closed universe, or in a case where the interior of a black hole has evaporated completely, the area contribution from the boundary of the island is zero.

Now we will compute the Page curve for the radiation using the island formula. At early times of the evaporation there is no island. Thus extremising the island formula would yield an empty set. Then $S_{\text{semi-cl}}(\Sigma_{\text{Island}}) = 0$ and the contribution is just $S_{\text{semi-cl}}(\Sigma_{\text{Rad}})$. Note that, similar to the vanishing surface considered in section 4.2, the no island contribution also always extremises the generalised entropy, but does not always give the global minimal entropy. As the black hole evaporates, more outgoing radiation escapes and $S_{\text{semi-cl}}(\Sigma_{\text{Rad}})$ increases. This is represented by the increasing green line in figure 9. A scrambling time after the black hole forms, a non-vanishing island that extremises the generalised entropy surfaces. This island has a centre at the origin and its boundary is close to the black hole event horizon [1]. For increasing times on the cutoff surface, the island moves outwards towards the horizon. The $S_{\text{semi-cl}}$ term is small relative to the area term because now most of the interior modes that purify the outgoing radiation are included in the island [1]. In this case, the generalised entropy is dominated by the area term in (5.1), and decreases to zero with time. This behaviour is represented by the decreasing yellow line in figure 9.

To summarise, we found an increasing entropy contribution from the empty island and a decreasing entropy contribution from the non empty island which appears shortly after the black hole starts evaporating. As in the case of applying the QES prescription in section 4.2, here we also choose the island which gives the minimal entropy contribution at each time. Doing this, we obtain the form of the Page curve again. The overall fine-grained entropy of the outgoing radiation is depicted as the black curve in figure 9. Thus we see that the gravitational fine-grained entropy formula for the black hole, formula (2.19), and the formula for the entropy of exact state of radiation, (5.1), indeed point to the same result.

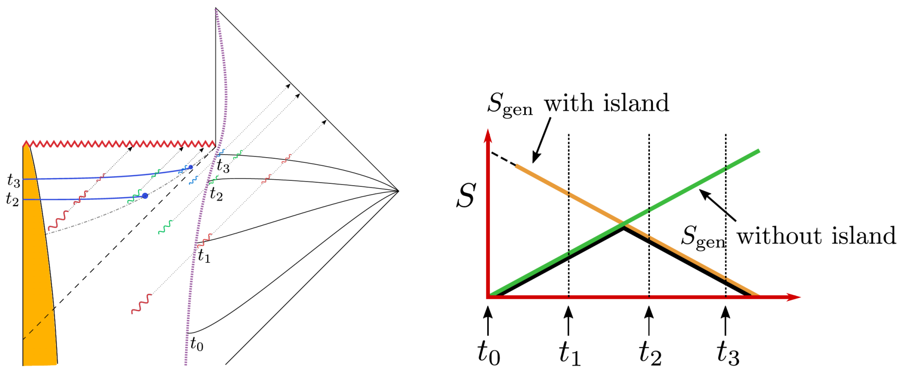

Figure 9: The green line represents the increasing generalised entropy with time from the no-island contribution, $S_{\text{semi-cl}}(\Sigma_{\text{Rad}})$, as more outgoing radiation escapes as the black hole evaporates . The yellow line represents the decreasing entropy from the the island contribution which appears some scrambling time after the black hole forms. It tracks the thermodynamic entropy of the black hole. When we consider both contributions and at each time and pick the minimal one, we get the final answer for the full entropy of radiation, which gives the Page curve. This is shown in black. Figure taken from [1].

5.1 The Euclidean Black Hole

Before moving on to the next section, we would like to first set some historical context in this section. We begin by introducing the Euclidean black hole and its thermal properties and lead up to a historical overview of the topic of reconstructions of the black hole interior. The geometry of the four-dimensional Euclidean black hole is obtained by analytically continuing the Schwarzschild solution to a Euclidean geometry by performing a Wick rotation $t = i\tau$, where τ is some imaginary time. The Euclidean metric is

$$ds_E^2 = \left(1 - \frac{r_s}{r}\right) d\tau^2 + \frac{1}{1 - \frac{r_s}{r}} dr^2 + r^2 d\Omega_2^2. \tag{5.2}$$

By choosing the periodicity of $\tau = \tau + \beta$ with period $2\pi/\kappa$, where $\kappa = 1/2r_s$, and defining a new radial polar coordinate, this metric is made regular everywhere; the conical singularity at the horizon $r_s = 2M$ becomes a coordinate singularity at the origin. Near the horizon, this metric has a $\mathbb{R}^2 \times S^2$ topology; the former is given by the first two terms in the metric. The periodicity condition in imaginary time required to make the metric regular at r_s defines a unique temperature $T_H = \kappa/2\pi$, the Hawking temperature of the black hole. At any other temperature, or with any other periodicity in τ, a conical singularity would exist. In other words, the black hole is in equilibrium with the quantum fields only at T_H [74].

Gibbons and Hawking in [12] used an Euclidean path integral for the imaginary-time evolution of the Euclidean black hole. They treated this path integral as a partition function $Z(\beta) = \text{Tr}\, e^{-\beta H}$. In QFT this is a standard argument; with knowledge of the partition function, one can compute the energy and entropy using

$$E = -\frac{\partial_\beta Z(\beta)}{Z(\beta)}, \tag{5.3}$$

$$S = (1 - \beta\partial_\beta) \log Z(\beta). \tag{5.4}$$

For the Euclidean black hole, we have the path integral [1, 74]

$$Z(\beta) = \int [\mathcal{D}\Phi] e^{-I_E[\Phi]} \sim e^{-I_{\text{classical}}} Z_{\text{quantum}}, \tag{5.5}$$

where the fields Φ have the same periodicity as τ and I_E is the Euclidean action that gives the Euclidean Einstein equations whose solutions give the metric (5.2). This thermal partition function has a contribution $I_{\text{classical}}$ from gravity and Z_{quantum} from the quantum fields on the geometry. More precisely, one can implement a cutoff at some finite radius such that the geometry within this cutoff is Euclidean and that outside the cutoff is flat, with each metric having its own Euclidean action [11]. It turns out that the latter, flat, geometry contributes to (5.5) more than the former Euclidean geometry. This is not alarming if one considers the leading contribution to (5.5) as coming from the Hawking radiation residing in the flat space region. The Euclidean metric then produces the subdominant contributions. By considering the solutions near r_s and applying equation (5.4) to (5.5), one can derive the generalised entropy (2.12) [12].

The metric (5.2) corresponds to a "cigar" shape with a regular tip at the horizon r_s. At a distance a few r_s away from the black hole, the metric has the geometrical shape of a cylinder, which is equivalent to having a strip with both its ends glued together with circumference β, where β is the inverse temperature measured by an observer far away [1]. The path integral will be calculated on this geometry with the angular coordinate $\theta = \theta + \beta$. Likewise, the fields in the path integral would have the same periodicity, as well as any observables computed. The cigar geometry (5.2) can be split into two and by computing the Euclidean path integral on the lower half, one can obtain the quantum state of the fields on the upper half, corresponding to a two-sided Schwarzschild black geometry with a smooth horizon. Tracing over one side would give a thermal distribution around T_H [11, 59]. In general, this is called the *thermofield double state* [75].

Similarly, in AdS/CFT, an eternal AdS black hole can be represented by two disconnected flat CFT exteriors (two asymptotic AdS boundaries), that are connected by a wormhole called the Einstein-Rosen bridge. The degrees of freedom on the two CFTs do not interact, but by gravity, bulk interactions are allowed via the wormhole. This manifests in terms of an entanglement entropy between the black holes characterised by the Bekenstein-Hawking entropy of the black hole [76]. It was shown in [77] that the bulk ground state of the wormhole, obtained by performing the CFT path integral on the boundary, can be written in terms of a product of the two CFTs. Two maximally entangled CFTs are represented by the state [76]

$$|\Psi\rangle \propto \sum_i e^{-\beta E_i/2} |\bar{i}\rangle |i\rangle \,, \tag{5.6}$$

where $|i\rangle$ is an eigenstate of the Hamiltonian of one CFT and $|\bar{i}\rangle$ is the state obtained by switching the two CFTs and reversing time in each CFT. This is the CFT version of the general thermofield double state.

This construction was used in introducing the concept of "ER=EPR" [76]. Unitarity required that the outgoing radiation modes be entangled with both the interior modes inside the horizon and the early outgoing radiation (which later becomes the late time outgoing radiation). But this is forbidden by the strong subadditivity condition. This implied that unitarity of the bulk dynamics is inconsistent with having a smooth horizon for an infalling observer after the Page time [78, 79]. According to ER=EPR, the Einstein-Rosen bridge in the two-sided AdS-Schwarzschild geometry is a manifestation of entanglement and connects the black hole and outgoing radiation regions. These two regions are entangled in the sense that if an observer were to collect the radiation and input it into a quantum computer, the output would be the second black hole that is maximally entangled with the first [77]. Even though an evaporating black hole is one-sided, at the Page time it becomes two-sided. At this time, there is maximal thermal entanglement between the two regions [76].

The constructions we have reviewed above are part of the historical development of studying the information loss problem and black hole interior reconstructions. In 1976, Hawking argued that correlations between outgoing radiation with the initial matter state decay exponentially in the entropy for computations performed on the Schwarzschild geometry. Late time correlations were largely studied as a means to probe the microscopic degrees of freedom of a black holes quantum system. However, they do not agree with the semiclassical description of the black hole horizon [78]. In AdS/CFT, large black holes do not spontaneously evaporate; radiation is reflected off the AdS boundary and back into the black hole. Thus correlations with the initial states in this context in the semiclassical description take arbitrarily long to decay. This is in fact the information loss problem in the context of AdS/CFT. Maldacena in [77] considered a thermal state of an eternal AdS black hole and showed that the two-point correlation function of quantum fields do fluctuate and eventually decay exponentially in the entropy. Noting that the entropy is inversely proportional to G_N, this suggests that the correlations are due to non-perturbative effects and thus are undetectable by a perturbative analysis. Such late time correlation functions were also studied in JT gravity [80]. Generally, in the past, investigating the information loss problem in the context of AdS/CFT was limited to studying effects that were exponentially small, which excludes the Page curve since it has a $O(1/G_N)$ effect [2].

A resolution to this was to construct a dictionary for interior operators behind the horizon and CFT observables [81]. This was studied by Almheiri in [79]. Papadodimas and Raju argued that such boundary operators must be state-dependent and explicitly constructed the interior operators for a given microstate [82, 83]. Different boundary operators of a bulk AdS space are related to different interior bulk operators; the latter can be reconstructed from the former. This is an example of state dependence, which was argued in [70] to be only required for certain types of reconstructions. In general, interior reconstructions of the black hole in the past involved the state-dependence of interior operators [82–85] and is an area of study in quantum information and correction. Past ideas for interior reconstruction served as motivation for reconstruction proposals today like the entanglement wedge reconstruction (EWR) [86–88]. A bulk operator can only be reconstructed if it is strictly in the EW of the boundary region and we use EWR to find the part of the boundary to use in the reconstruction. In [79], Almheiri applied EWR alongside ER=EPR for a two-sided black hole with absorbing boundary conditions. He noted that degrees of freedom in the EW of one side can be later found in the EW of the other side. Penington in [69] applied the works of [79] and [76] and used EWR to show the consistency between having a smooth horizon and ensuring unitarity for one-sided evaporating black holes.

6 Entanglement Wedge Reconstruction

Thus far we know that information is not lost and does escape. In the previous sections we found that both the black hole and radiation entropies follow the Page curve. However, understanding the Page curve is only an aspect of the black hole information loss problem. We still do not know: 1. How the infalling matter escapes as outgoing radiation from a bulk perspective; 2. If the black hole degrees of freedom in the central dogma (which characterise the geometry contained within the minimal surface) provide any information about the interior and 3. How the final state encodes the initial infalling information.

A peek into the answer to the above questions were given in section 4.2, where the EWs of the black hole and of radiation were introduced. Note that it was assumed that the outgoing radiation remained completely thermal till the end of the black hole evaporation and that it is purified by interior modes. The EWR proposal by Penington and AEMM states that the degrees of freedom of the black hole or radiation subsystem is described by its respective EW, or vice versa [69]. A part of the interior is encoded in the black hole degrees of freedom and a part to the radiation (represented by islands). When we say that the EW describes the subsystem, by describe we mean that one can perform a quantum operation on the radiation degrees of freedom in order to recover the state of a qubit (or a small number of qubits) inside the EW (the interior), with the recovery being state or subspace dependent, i.e. dependent on the state of the rest of the qubits [69]. Such operations have complexity of the order of an exponential of the black hole entropy. One complex measurement involves creating a wormhole (such as the Gao Jafferis Wall traversable wormhole [89]) that connects the black hole interior to the quantum computer that is conducting the measurement. Other operations include the Petz map [90–93] and a modular Hamiltonian that can act non-locally in the bulk. Petz maps were used in [3] to define a reconstruction operator using a gravitational path integral and the replica method.

In information recovery, the Hayden-Preskill decoding criterion [94] states that if a small diary was thrown into the black hole before the Page time, then its information can be reconstructed at the Page time. But if the diary was thrown in after the Page time, one would have to wait for a scrambling time before its information can be reconstructed from the outgoing radiation. Penington in [69] showed that the EWR precisely reproduces these results and further generalised the criterion to show that reconstruction of the diary depends on its energy and entropy. A diary that was thrown into the black hole before and after the Page time would reside in the EW of the radiation and the CFT respectively. The reconstruction of the diary that was thrown in before the Page time rests on knowing the state of the black hole. Note that although the entanglement between radiation and the interior is independent of the initial state that fell into the black hole, the encoding of the interior in the radiation does depend on the initial state. Hence, an observer with access to the radiation is able to recover information about the initial state.

Penington showed that at late times after the Page time, interior degrees of freedom are encoded in the early-time outgoing radiation. In other words, a large part of the interior is described by radiation degrees of freedom and not by the black hole degrees of freedom that is associated to the central dogma. We have in fact already used EWR in section 2.7. In the context of the island conjecture, EWR implies that information in islands Σ_{Island} is contained in the radiation region Σ_{Rad}. For late times after the Page time, we have outgoing radiation modes b that are entangled with both the early time outgoing radiation modes b' and their Hawking mode partner a that is in the interior. The clearer picture here is that b' actually resides in an island and is a part of the radiation (it is in the EW of radiation) and a is the future evolution of b', i.e. $a = b'$ [63]. We see that the interior operators do not act on the Hilbert space of the black hole associated with the central dogma, but on the Hilbert space of the radiation.

Actually if one noticed the use of the same surface Q in both the QES and island prescriptions, one would not be surprised by both producing the Page curve. Indeed, for a pure matter state on the whole Cauchy slice, we have $S_{\text{semi-cl}}(\Sigma_Q) = S_{\text{semi-cl}}(\Sigma_{\text{Rad}} \cup \text{Island})$ [1]. In particular, only if one assumes EWR, can the claim that the QES computes the radiation entropy be valid.

AMMZ considered black hole evaporation in AdS_2 with JT gravity embedded in a holographic theory in AdS_3 [73]. They considered a black hole coupled to a matter CFT on a half line and computed the entropy associated to regions in the CFT. It was found that after the Page time, regions lying deep in the black hole interior (and thus the degrees of freedom of the interior quantum state) are geometrically connected to the escaped radiation via an extra dimension. In other words, the EW for the radiation can reside deep inside the black hole interior in the form of an island. In fact, the appearance of an island was previously noted by Penington and AEMM when computing the EW of a black hole. The extra dimension in the holographic theory connects the island to the exterior where the radiation resides in [73]. A similar embedding was done for an AdS_4 black hole in [95].

We now illustrate EWR using a simple example. In figure 10, the green spacetime regions represent the EW of the black hole, which are associated to the quantum degrees of freedom that describe the black hole from the outside. The blue regions represent the EW of radiation. Consider the minimal vanishing surface of a spatial slice going all the way to $r = 0$ at a time before the Page time. The EW of the black hole in this case is represented by the green region in figure 10(a). Any point in this region can be reached with initial knowledge of that slice. After the Page time, we have a new QES and the spatial slice only goes up to the point labelled QES in figure 10(b). Now the blue regions includes part of the black hole interior. In other words, the black hole degrees of freedom only describes a part of the interior, given by the green region in figure 10(b). In computing the entropy of radiation, we were including an interval on the inner blue region in figure 10(b). Thus the EW of radiation includes this interval region of the black hole interior. The fine-grained entropy formula of the radiation after the Page time includes this portion of the interior (the interval) as part of the island. Thus formula (5.1) depends on the quantum state of that region. Lastly ,when the black hole has evaporated completely, the EW of radiation is the whole black hole interior (which is flat space), represented by the blue region in figure 10(c).

For some intuition on the QES in figure 10(b), consider decoupling the total quantum system at some time on the cutoff surface using some boundary condition. This stops the black hole evaporation at the point of decoupling, i.e. there is no longer interaction between the radiation outside and the black hole. Now, one can only recover information in the green region (the black hole) behind the horizon by altering the Hamiltonian, but not from the blue region in the interior (the island). Regardless of the amount and complexity of operations done on the black hole system, information in the island can never be retrieved. In this sense, the QES that connects the island and black hole systems acts as the boundary of the region that is accessible

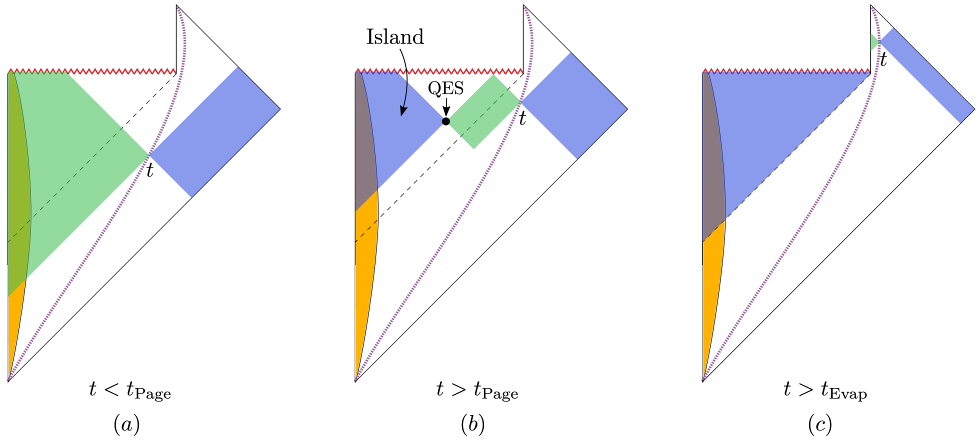

Figure 10: The green and blue regions represent the EWs of the black hole and escaped radiation respectively. The white region contains no gravity and is flat. (a) The black hole at early times before the Page time. A diary previously thrown into the black hole would reside in the EW of radiation represented by the inner green region. (b) The black hole after the Page time. A diary previously thrown into the black hole would reside in the EW of the CFT represented by the inner and outer blue regions that are disconnected. The diaries in (a) and (b) can be reconstructed by operating on the radiation if the initial state of the black hole is known. (c) The black hole after complete evaporation. Semiclassical description is no longer valid. The radiation and CFT systems are disconnected. All the outgoing radiation modes lie in the EW of the radiation represented by the inner blue region. Figure taken from [1].

through operations on the black hole degrees of freedom [63].

Almheiri *et al.* in [1] emphasised that there are two concepts associated to the black hole degrees of freedom. The first is that it refers to the quantum degrees of freedom associated to the central dogma. In the analogy with a piece of coal, this corresponds to all the degrees of freedom of the coal's surface and interior. The second concept is that it refers to the quantum interior degrees of freedom of the black hole in the semiclassical description. The degrees of freedom inside the coal are not included in this case. These degrees of freedom are either encoded in the Hilbert space associated to the central dogma or to that of the radiation region, depending on the QES containing the degrees of freedom [1].

We now revisit the two geometries that were introduced in section 3.1. In figure 10(b), consider some qubit that is inside the island, i.e. it is associated to the black hole degrees of freedom. In principle, this qubit is accessible by operating on the outside radiation (outer blue region in the figure 10(b)). Note that this operation ultimately does not affect the state of the qubit; initial perturbations can be quantum error corrected to ensure this [63]. Relating to the second concept of black hole degrees of freedom mentioned above, this qubit has no relation with degrees of freedom inside the coal. In the geometry of a disconnected parent and baby universe, the former only has access to a part of the latter, which is the middle green region in figure 10(b). The other part of the baby universe, the island region, can only be accessed from the outside [63].

Recall that in the bags of gold geometry, the fine-grained entropy of the object from the outside is just the area term given by the area of the throat, which is the QES. But this would not account for the large "extra" amount of entropy because the EW of the black hole degrees of freedom ends near the surface of the throat and does not include the whole interior, but only a part of it. The rest of the interior region is entangled with other matter that lives inside this closed universe. In fact, after the Page time late into black hole evaporation, the geometry of the black hole resembles that of a bag of gold [63]. In figure 10(b), consider a slice starting in the interior, passing through the QES and ending in the outer blue region. The geometry of this slice and its associated entropy parallels that of the bag of gold, where the portion of the slice that is in the green region and crossing the black hole horizon (QES) corresponds to a narrow throat. The entropy in the island in the case for an old black hole would be the radiation of the bag of gold, and the rest of the entropy in the bag would be entangled to some other system that depends on how the closed universe was created [63].

7 Conclusion and Discussion

We saw the analogies between the laws of black hole mechanics and laws of thermodynamics, which suggested the thermodynamic Bekenstein-Hawking entropy. Evidence that such an entropy is thermodynamic has to come from a precise statistical mechanical interpretation, like the counting of black hole microstates. To this day we have only achieved this for specific black holes in string theory and AdS/CFT. Since the first piece of such evidence was provided by Strominger and Vafa in 1996, many similar calculations have been done, even for the Schwarzschild black hole. This was done in various dimensions by Xiao in [96] by considering the black hole as composed of microscopic particles. In particular, he noted a resemblance of a Schwarzschild black hole to a long non-relativistic quantum mechanical string in one dimension. Futhermore, Xiao advocated for describing a black hole with an equation of state with cosmological constant $w = 1$ and suggested that a corresponding holographic fluid permeated the early universe. Another microscopic derivation of the Bekenstein-Hawking entropy for non-supersymmetric Schwarzschild black holes in four and five dimensions was performed in [97] and was also related to D-branes.

The Bekenstein-Hawking entropy together with AdS/CFT motivated its generalisation to higher dimensions for CFTs with a dual gravitational theory, the R-T formula, which computes the fundamental entanglement between the boundary CFT and a region in AdS space in the CFT's gravitational dual. The area of interest is the minimal bulk surface. For a black hole treated as two subsystems, the HEE is given by the entangling surface that separates the subsystems and is the event horizon. Solodukhin in [28] provided a holographic description of the horizon for two- and four-dimensional black holes; an attempt to interpret the Bekenstein-Hawking entropy as a HEE was also done. The minimal R-T surface in the HEE is generally a challenge to find and compute. Whilst relatively simpler to find for specific geometries in three-dimensional gravity AdS_3/CFT_2, Bao and Davies in [98] produced an algorithm with a polynomial complexity to achieve this task for n boundary subregions.

The beauty of the R-T formula is that its nature converts the computation of reduced density matrices to a geometric one; pure and mixed states of a black hole can be distinguished by the interior geometry. The approaches used to compute entropy in this dissertation do not involve knowing the precise quantum state of the subsystems; we only provided qualitative descriptions of the fine-grained entropies in semiclassical geometry using gravity as an effective theory, which showed us how the spacetime geometry is constructed from the degrees of freedom. AdS/CFT and holography have been used to test these results; the tools of string theory serves as a geometrical bridge between the entropies of gravitational systems and of quantum systems. AdS/CFT allows us to answer the question of why black holes with different interiors result in the same final state of radiation.

We do not know if the fine-grained formula (2.19) would work for cutoff surfaces lying near the black hole or in non static spacetimes outside the cutoff surface. Current derivations of formula (2.19) involves summing over degrees of freedom or microstates using the Euclidean gravitational path integral using a method similar to that of Gibbons and Hawking in [12]. It is an assumption, a constituent of the central dogma and also does not reveal an explicit picture of the Hilbert space associated to the computation. Nonetheless, current understanding using string theory is successful in explaining the emergence of black hole thermodynamics. In gravity, we still need to understand how to define or derive this formula, such as the saddle points one should use [1]. In [3], it was claimed that the Page curve for the black hole was derived from using only a gravitational path integral without requiring AdS/CFT.

We saw that the quantum extended R-T formula, which uses QESs, explains the Page curve. In the QES prescription, the quantum term in (2.19) can compete with the classical area term to produce new saddle points for the QES, which could describe the late time behaviour of the entanglement entropy. At the Page time the EW undergoes a phase transition in the quantum R-T surface, the QES that gives the minimal generalised entropy, of an evaporating black hole. The decrease in entanglement entropy is strictly less than that of the Hawking radiation because the quantum R-T surface lies strictly just inside the event horizon a scrambling time in the past. This is equivalent to the statement that the GSL is a strict inequality. It is remarkable that we could reproduce the Page curve from the extremising and minimising procedures in the semi-classical prescription without knowledge of the details of black hole degrees of freedom and how information is encoded in the radiation. EWR explains the Hayden Preskill decoding criterion as well as ensures consistency between the Page curve and the bulk entanglement structure. Crucial is the state dependence of the EWR, which explains how information is able to escape the black hole.

In the island prescription, we saw that the black hole interior is included in the final state of radiation, and that the Page curve was also reproduced. One could argue that the addition of islands is just equivalent to including the interior region, which included for more purifying modes and produced a pure state. But, as pointed out in [1], this was necessitated by gravity. In the island formula there is the presence of two types of radiation quantum states – the one that lives in the semiclassical geometry and the exact one that a full theory of quantum gravity is supposed to give. In the future, one would like to obtain the latter directly from a theory of gravity instead of through string theory in the semiclassical description.

In the last section, we qualitatively saw that information recovery is possible through performing operations on the outgoing radiation, which allowed for the extraction of information from the interior. More precisely, Penington *et al.* in [3] showed that this involves creating wormholes in the semiclassical geometry, which ensures unitarity. In such cases the exact state of radiation (the left hand side of (5.1)) and computations using the replica method are required. Crucial to the historical development of an understanding of the information loss problem was the AMPS paradox, of which an explicit discussion was omitted in this dissertation. However, this paradox was shown to be resolved in [1] and [69].

7.1 A Few Notes on the Page Curve

We saw that by the requirement of having a pure final state radiation, the entanglement entropy must decrease back to zero. Page noted that the rate of this decrease is bounded and the inflection point of the Page curve at the Page time is consistent with this. The transition at the inflection point is sharper and faster for more chaotic black hole systems; else, the evolution of the entropy takes on a flatter shape [99].

In the discussion of the Page curve, a detailed description of the segment of time within $O(t_{\text{Page}}/\sqrt{S_{\text{BH}}})$ of the turning point at the Page time was omitted. This was tackled by Penington in [69], which we will now follow. The curve has an $O(1)$ correction [68] at the turning point of the curve due to fluctuations in the area of the horizon and total energy of radiation that are of order $O(\sqrt{G_N})$ and $O(\sqrt{S})$ respectively. In other words, there is no well-defined R-T surface in this period of time and so the entanglement entropy remains stagnant. But it is argued to be possible to treat a state in the total system, which is a superposition of the two subsystems (the black hole and radiation), as having a well-defined R-T surface. Then upon performing EWR, one will find that a tiny, non-perturbative error exists for some of the states of the total system; these states would have the interior encoded in the radiation outside the cutoff surface, while the remaining states will have the interior encoded in the CFT. The number of former states increases as the black hole evaporates until the whole interior can be reconstructed from the radiation, albeit with a tiny error [69]. Penington pointed out that a plot of the fraction of states with interior encoded in the radiation region against time would allow one to discern the form of the peak of the Page curve. In addition, an "error" in the Page time of order $O(\beta\sqrt{S_{\rm BH}H})$ due to the uncertainty in the rate of evaporation implies that in principle it is possible to compute the entropy to order $O(1/\sqrt{S_{\rm BH}})$ accuracy at all times [69]. Hence, we see that the fluctuations allows for a more precise calculation of the Page curve.

7.2 Looking Towards the Future

The AdS/CFT framework, which is presently assumed to be accurate non-perturbatively, allows us to explore the low energy effective gravitational theory and its properties. It also explains observations such as the fate of the information, but not how the fate occurs. One would also desire to understand current formulations of HEE and the EWR beyond the context of AdS/CFT and in the context of an independent theory of quantum gravity in a general background. Such a full theory of quantum gravity will be reproducible from perturbative analyses in string theory, which will also extend the types of quantum mechanical systems we can study. Currently, AdS/CFT is limited to specific or special geometries and spacetimes with boundary conditions where we can define quantum mechanical systems or quantum fields on [99]. Of particular attention is the problem of microscopically distinguishing between the degrees of freedom encoding the black hole and the degrees of freedom encoding the radiation, which requires using the holographic tools beyond asymptotic boundaries. Achieving this would provide an explicit picture of how the black hole interior is encoded in the radiation even before one is able to extract the radiation [69].

Generalising current constructions in string theory to be applicable in cosmology is important for obtaining an accurate interpretation of the black hole singularity and the physics near that region. What does an infalling observer see? Past proposals for ensuring a smooth horizon is seen by such an observer were attempted without accurate knowledge of the black hole interior [100, 101]. Understanding quantum gravity in a cosmological background would allow us to study the information problem in our universe. In this context, we still do not know if information escape black holes. There are no timelike or lightlike asymptotic regions in de Sitter spacetime on which to base the Cauchy slice; the state-dependent encoding of the interior in radiation after the Page time may not be applicable [69]. Lastly, quantum cosmology would allow for the study of the big crunch of our universe. In similar analogy, Hawking radiation is like the early fluctuations that permeated the universe during inflation which later became the CMB. Our expanding universe also has a cosmological event horizon which we would like to understand fundamentally and quantum mechanically.

The topic of this dissertation connects many different areas of physics and in particular connects quantum information theory and the geometry of spacetime. Models of quantum systems resembling black holes have been used to study this area. Such models must have, among other features, a qubit number comparable to the black hole entropy, be chaotic, and be strongly interacting [102]. Quantum circuits are such an example; operating on technologies such as trapped ions, superconductors and silicon quantum dots, one would like to study the emergence of gravity and the spacetime in relation to quantum corrections in such systems. It will be interesting to discover if such systems are a universe of their own [102].

References

[1] A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian, A. Tajdini, "The entropy of Hawking radiation", 2020, [arXiv:2006.06872v1 [hep-th]].

[2] A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian and A. Tajdini, "Replica Wormholes and the Entropy of Hawking Radiation", *JHEP* **05**, 013, 2020, [arXiv:1911.12333 [hep-th]].

[3] G. Penington, S. H. Shenker, D. Stanford and Z. Yang, "Replica wormholes and the black hole interior", 2019, [arXiv:1911.11977 [hep-th]].

[4] S. W. Hawking, "Gravitational radiation from colliding black holes", *Phys. Rev. Lett.* **26**, 1344–1346, 1971.

[5] J. M. Bardeen, B. Carter and S. Hawking, "The Four laws of black hole mechanics", *Commun. Math. Phys.* **31**, 161–170, 1973.

[6] J. D. Bekenstein, "Black holes and entropy", *Phys. Rev.* **D7**, 2333–2346, 1973.

[7] J. D. Bekenstein, "Generalised second law of thermodynamics in black hole physics" *Phys. Rev.* **D9**, 3292, 1974.

[8] R. Bousso, "Black hole entropy and the Bekenstein bound", 2018, [arXiv:1810.01880v1 [hep-th]].

[9] S. Hawking, "Particle Creation by Black Holes", *Commun. Math. Phys.* **43**, 199–220, 1975.

[10] S. W. Hawking, "Breakdown of Predictability in Gravitational Collapse", *Phys. Rev.* **D14**, 2460–2473, 1976.

[11] D. Harlow, "Jerusalem Lectures on Black Holes and Quantum Information", *Rev. Mod. Phys.* **88**, 015002, 2016, [arXiv:1409.1231 [hep-th]].

[12] G. Gibbons and S. Hawking, "Cosmological Event Horizons, Thermodynamics, and Particle Creation", *Phys. Rev.* **D15**, 2738–2751, 1977.

[13] E. Bianchi and R. C. Myers, "On the Architecture of Spacetime Geometry", 2012, [arXiv:1212.5183 [hep-th]].

[14] S. Ryu and T. Takayanagi, "Holographic derivation of entanglement entropy from AdS/CFT", *Phys. Rev. Lett.* **96**, 181602, 2006, [arXiv:hep-th/0603001].

[15] V. E. Hubeny, M. Rangamani, and T. Takayanagi, "A Covariant holographic entanglement entropy proposal", *JHEP* **07**, 062, 2007, [arXiv:0705.0016].

[16] T. Nishioka, S. Ryu and T. Takayanagi, "Holographic Entanglement Entropy: An Overview", *J. Phys.* **A42**, 504008, 2009, [arXiv:0905.0932 [hep-th]].

[17] S. Carlip, "Black Hole Thermodynamics and Statistical Mechanics", *Lect. Notes Phys.* **769**, 89–123, 2009, [arXiv:0807.4520v1 [gr-qc]].

[18] A. Strominger and C. Vafa, "Microscopic origin of the Bekenstein-Hawking entropy", *Phys. Lett.* **B379**, 99–104, 1996, [arXiv:hep-th/9601029 [hep-th]].

[19] G. T. Horowitz and J. Polchinski, "A Correspondence Principle for Black Holes and Strings", *Phys. Rev.* **D55**, 6189, 1997, [arXiv:hep-th/9612146v2].

[20] G. T. Horowitz and M. M. Roberts, "Counting the Microstates of a Kerr Black Hole", *Phys. Rev. Lett.* **99**, 221601, 2007, [arXiv:0708.1346v1 [hep-th]].

[21] J. M. Maldacena, "The large N limit of superconformal field theories and supergravity", *Adv. Theor. Math. Phys.* **2**, 231–252, 1998, [arXiv:hep-th/9711200v3].

[22] O. Aharony, S. S. Gubser, J. M. Maldacena, H. Ooguri, and Y. Oz, "Large N field theories, string theory and gravity", *Phys. Rept.* **323**, 183–386, 2000, [arXiv:hep-th/9905111v3].

[23] G. 't Hooft, "Dimensional reduction in quantum gravity", 2009, [arXiv:gr-qc/9310026v2].

[24] L. Susskind, "The world as a hologram", *J. Math. Phys.* **36**, 6377–6396, 1995, [arXiv:hep-th/9409089v2].

[25] D. Bigatti and L. Susskind, "Tasi lectures on the holographic principle", 2000 [arXiv:hep-th/0002044v1].

[26] L. Susskind and J. Uglum, "Black hole entropy in canonical quantum gravity and superstring theory", *Phys. Rev.* **D50**, 270, 1994, [arXiv:hep-th/9401070v3].

[27] C. G. Callan, Jr. and F. Wilczek, "On geometric entropy", *Phys. Lett.* **B333**, 55, 1994, [arXiv:hep-th/9401072v3].

[28] S. N. Solodukhin, "Entanglement entropy of black holes", *Living Rev. Rel.* **14**, 8, 2011, [arXiv:1104.3712v1 [hep-th]].

[29] S. N. Solodukhin, "The Conical singularity and quantum corrections to entropy of black hole", *Phys. Rev.* **D51**, 609, 1995, [hep-th/9407001].

[30] D. V. Fursaev and S. N. Solodukhin, "On one loop renormalisation of black hole entropy", *Phys. Lett. B*, 365, 51, 1996, [arXiv:hep-th/9407001v1].

[31] L. Susskind, L. Thorlacius, and J. Uglum, "The stretched horizon and black hole complementarity", *Phys. Rev.* **D48**, 8, 3743–3761, 1993, [arXiv:hep-th/9306069v2].

[32] E. H. Lieb and M. B. Ruskai, "A fundamental property of quantum-mechanical entropy", *Phys. Rev. Lett.* **30**, 434–436, 1973; "Proof of the strong subadditivity of quantum mechanical entropy", *J. Math. Phys.* **14**, 1938–1941, 1973.

[33] T. Hartman and J. Maldacena, "Time Evolution of Entanglement Entropy from Black Hole Interiors", *JHEP* **05**, 014, 2013, [arXiv:1303.1080v2 [hep-th]].

[34] T. Faulkner, A. Lewkowycz and J. Maldacena, "Quantum corrections to holographic entanglement entropy", *JHEP* **11**, 074, 2013, [arXiv:1307.2892v2 [hep-th]].

[35] R. Emparan, "Black hole entropy as entanglement entropy: A holographic derivation", *JHEP* **0606**, 012, 2006, [arXiv:hep-th/0603081v3].

[36] D. Fursaev "Proof of the holographic formula for entanglement entropy", *JHEP* **0609**, 018, 2006, [arXiv:hep-th/0606184v2].

[37] T. Faulkner, "The Entanglement Renyi Entropies of Disjoint Intervals in AdS/CFT", 2013, [arXiv:1303.7221v1 [hep-th]].

[38] T. Hartman, "Entanglement Entropy at Large Central Charge", 2013 [arXiv:1303.6955v1 [hep-th]].

[39] S. N. Solodukhin, "Entanglement entropy of black holes and AdS/CFT correspondence", *Phys. Rev. Lett.* **97**, 201601, 2006, [arXiv:hep-th/0606205v3].

[40] M. Headrick and T. Takayanagi, "A holographic proof of the strong subadditivity of entanglement entropy", *Phys. Rev.* **D76**, 106013, 2007, [arXiv:0704.3719v1 [hep-th]].

[41] T. Hirata and T. Takayanagi, "AdS/CFT and strong subadditivity of entanglement entropy", *JHEP*, 0702, 042, 2007 [arXiv:hep-th/0608213].

[42] A. Lewkowycz and J. Maldacena, "Generalised gravitational entropy", *JHEP* **08**, 090, 2013, [arXiv:1304.4926v2 [hep-th]].

[43] A. C. Wall, "Maximin Surfaces, and the Strong Subadditivity of the Covariant Holographic Entanglement Entropy", *Class. Quant. Grav.* **31**, 225007, 2014, [arXiv:1211.3494v4 [hep-th]].

[44] N. Engelhardt and A. C. Wall, "Quantum Extremal Surfaces: Holographic Entanglement Entropy beyond the Classical Regime", JHEP **01**, 073, 2015, [arXiv:1408.3203v3 [hep-th]].

[45] L. Susskind, Lecture on "Statistical Mechanics Lecture 7", `https://www.youtube.com/watch?v=sg15UClUY48&t=83m0s`, May 13 2013. Originally presented in the Stanford Continuing Studies Program.

[46] J. Maldacena, Talk on "Review: Recent progress on the black hole information paradox", `https://indico.cern.ch/event/929434/contributions/3913390/attachments/2069777/3474397/Maldacena.pdf`. Presented at the Strings 2020 conference.

[47] A. Renyi. "On measures of entropy and information", *Hungarian Academy Of Sciences*, 1961.

[48] T. Barrella, X. Dong, S. A. Hartnoll and V. L. Martin, "Holographic entanglement beyond classical gravity", *JHEP* **1309**, 109, 2013, [arXiv:1306.4682v2 [hep-th]].

[49] V. Iyer and R. M. Wald, "Some properties of Noether charge and a proposal for dynamical black hole entropy", *Phys. Rev.* **D50**, 846, 1994, [arXiv:gr-qc/9403028v1].

[50] D. N. Page, "Information in black hole radiation", *Phys. Rev. Lett.* **71**, 3743–3746, 1993, [arXiv:hep-th/9306083v2 [hep-th]].

[51] J. A. Harvey and A. Strominger, "Quantum Aspects of Black Holes", 1992, [arXiv:hep-th/9209055v1].

[52] S. B. Giddings, "Toy Models For Black Hole Evaporation", 1992 [arXiv:hep-th/9209113v2].

[53] S. B. Giddings, "Black Holes and Quantum Predictability", 1993, [arXiv:hep-th/9306041v2].

[54] J. Preskill, "Do Black Holes Destroy Information?", 1992, [arXiv:hep-th/9209058v1].

[55] T. Banks, M. O'Loughlin, and A. Strominger, "Black Hole Remnants and the Information Puzzle", *Phys. Rev.* **D47**, 4476–4482, 1993, [arXiv:hep-th/9211030v2].

[56] D. N. Page, "Is Black-Hole Evaporation Predictable?", *Phys. Rev. Lett.*, **44**, 301, 1980.

[57] G. 't Hooft, "On the quantum structure of a black hole", *Nucl. Phys.* **B256**, 727, 1985.

[58] K. Schoutens, E. Verlinde and H. Verlinde, "Quantum Black Hole Evaporation", 1993, [arXiv:hep-th/9401081v1].

[59] J. B. Hartle and S. W. Hawking, "Path Integral Derivation Of Black Hole Radiance", *Phys. Rev.* **D13**, 2188, 1976.

[60] C. G. Callan, S. B. Giddings, J. A. Harvey, and A. Strominger, 'Evanescent Black Holes", *Phys. Rev.* **D45**, R1005, 1992, [arXiv:hep-th/9111056v1].

[61] A. Strominger, "Unitary rules for black hole evaporation", 1994, [arXiv:hep-th/9410187v2].

[62] J. Wheeler, *Relativity, groups and topology*, edited by B. S. DeWitt and C. M. DeWitt, Gordon and Breach, New York, 1964.

[63] J. Maldacena , Talk on "Review: Recent progress on the black hole information paradox", `https://www.youtube.com/watch?v=6UpV_C4RB4M`. Presented at Strings 2020 conference, 29 June 2020.

[64] A. Dabholkar, J. Gomes and S. Murthy, "Nonperturbative black hole entropy and Kloosterman sums", *JHEP* **03**, 074, 2015, [arXiv:1404.0033v3 [hep-th]].

[65] T. Banks, W. Fischler, S. Shenker and L. Susskind, "M theory as a matrix model: A Conjecture", *Phys. Rev.* **D55**, 5112–5128, 1997, [arXiv:hep-th/9610043v3].

[66] E. Witten, "Anti-de Sitter space and holography", *Adv. Theor. Math. Phys.* **2**, 253–291, 1998, [arXiv:hep-th/9802150v2 [hep-th]].

[67] D. N. Page, "Time Dependence of Hawking Radiation Entropy", *JCAP* **1309**, 028, 2013, [arXiv:1301.4995v3].

[68] D. N Page, "Average entropy of a subsystem", *Phys. Rev. Lett.* **7**, 9, 1291, 1993.

[69] G. Penington, "Entanglement Wedge Reconstruction and the Information Paradox", 2019, [arXiv:1905.08255v3 [hep-th]].

[70] P. Hayden and G. Penington, "Learning the Alpha-bits of Black Holes", 2018, [arXiv:1807.06041 [hep-th]].

[71] A. Almheiri, N. Engelhardt, D. Marolf and H. Maxfield, "The entropy of bulk quantum fields and the entanglement wedge of an evaporating black hole", 2019, [arXiv:1905.08762 [hep-th]].

[72] G. Penington, Talk on "Entanglement Wedge Reconstruction and the Information Paradox", `https://www2.yukawa.kyoto-u.ac.jp/~qist2019/slides/5th/Penington.pdf`. Presented at IFQ school/workshop, Quantum Information and String Theory 2019, Yukawa Institute for Theoretical Physics, Kyoto University, 24 June 2019.

[73] A. Almheiri, R. Mahajan, J. Maldacena and Y. Zhao, "The Page curve of Hawking radiation from semiclassical geometry", 2019, [arXiv:1908.10996 [hep-th]].

[74] P.K. Townsend, "Black Holes", 1997, [arXiv:gr-qc/9707012v1].

[75] W. Israel, "Thermofield dynamics of black holes", *Phys.Lett.* **A57**, 107–110, 1976.

[76] J. Maldacena and L. Susskind, "Cool horizons for entangled black holes", *Fortsch.Phys.* **61** 81–811, 2013. [arXiv:1306.0533].

[77] J. M. Maldacena, "Eternal black holes in anti-de Sitter", *JHEP* **04**, 021, 2003, [arXiv:hep-th/0106112 [hep-th]].

[78] A. Almheiri, D. Marolf, J. Polchinski, and J. Sully, "Black Holes: Complementarity or Firewalls?", *JHEP* **02**, 062, 2013, [arXiv:1207.3123].

[79] A. Almheiri, "Holographic Quantum Error Correction and the Projected Black Hole Interior", 2018

[arXiv:1810.02055v2 [hep-th]].

[80] P. Saad, "Late Time Correlation Functions, Baby Universes, and ETH in JT Gravity", 2019, [arXiv:1910.10311 [hep-th]].

[81] A. Almheiri, D. Marolf, J. Polchinski, D. Stanford, and J. Sully, "An Apologia for Firewalls", *JHEP* **09**, 018, 2013, [arXiv:1304.6483].

[82] K.Papadodimas and S. Raju, "An Infalling Observer in AdS/CFT", *JHEP* **10**, 212, 2013, [arXiv:1211.6767 [hep-th]].

[83] K.Papadodimas and S. Raju, "State-dependent bulk-boundary maps and black hole complementarity", *Phys. Rev.* **D89**, 8, 086010, 2014.

[84] E. Verlinde and H. Verlinde, "Black Hole Entanglement and Quantum Error Correction", *JHEP* **10**, 107, 2013, [arXiv:1211.6913].

[85] Y. Nomura, F. Sanches, and S. J. Weinberg, "Black Hole Interior in Quantum Gravity", *Phys. Rev. Lett.* **114**, 201301, 2015, [arXiv:1412.7539].

[86] A. Almheiri, X. Dong and D. Harlow, "Bulk Locality and Quantum Error Correction in AdS/CFT", *JHEP* **04**, 163, 2015, [arXiv:1411.7041 [hep-th]].

[87] D. L. Jafferis, A. Lewkowycz, J. Maldacena and S. J. Suh, "Relative entropy equals bulk relative entropy", *JHEP* **06**, 004, 2016, [arXiv:1512.06431 [hep-th]].

[88] X. Dong, D. Harlow and A. C. Wall, "Reconstruction of Bulk Operators within the Entanglement Wedge in Gauge-Gravity Duality", *Phys. Rev. Lett.* **117**, 021601, 2016, [arXiv:1601.05416 [hep-th]].

[89] P. Gao, D. L. Jafferis and A. Wall, "Traversable Wormholes via a Double Trace Deformation", 2016, [arXiv:1608.05687 [hep-th]]

[90] J. Cotler, P. Hayden, G. Penington, G. Salton, B. Swingle, and M. Walter, "Entanglement Wedge Reconstruction via Universal Recovery Channels", *Phys. Rev.* **X9**, 3, 031011, 2019, [arXiv:1704.05839 [hep-th]].

[91] C. F. Chen, G. Penington, and G. Salton, "Entanglement Wedge Reconstruction using the Petz Map", *JHEP* **01**, 168, 2020, [arXiv:1902.02844 [hep-th]].

[92] D. Petz, "Sufficient subalgebras and the relative entropy of states of a von neumann algebra", *Communications in mathematical physics* **105**, no. 1, 123–131, 1986.

[93] D. Petz, "Sufficiency of channels over von neumann algebras", *The Quarterly Journal of Mathematics*, **39**, no. 1, 97–108, 1988.

[94] P. Hayden and J. Preskill, 'Black holes as mirrors: quantum information in random subsystems", *JHEP* **09**, 120, 2007.

[95] A. Almheiri, R. Mahajan and J. E. Santos, "Entanglement islands in higher dimensions", 2019, [arXiv:1911.09666 [hep-th]]].

[96] Y. Xiao, "Microscopic derivation of the Bekenstein-Hawking entropy for Schwarzschild black holes", *Phys. Rev.* **D101**, 046020, 2020.

[97] K. Sfetsos and K. Skenderis "Microscopic derivation of the Bekenstein-Hawking entropy formula for non-extremal black holes", *Nucl.Phys.* **B517**, 179–204, 1998, [arXiv:hep-th/9711138].

[98] N.Bao and A. C. Davies "The Complexity of Identifying Ryu-Takayanagi Surfaces in AdS_3/CFT_2", *JHEP* **11**, 034, 2016, [arXiv:1609.01727 [hep-th]].

[99] J. Polchinski, 'The Black Hole Information Problem", 2016, [arXiv:1609.04036 [hep-th]].

[100] D. Harlow, "Aspects of the Papadodimas-Raju Proposal for the Black Hole Interior", *JHEP* **11**, 055, 2014, [arXiv:1405.1995].

[101] D. Marolf and J. Polchinski, "Violations of the Born rule in cool state-dependent horizons", *JHEP* **01**, 008, 2016, [arXiv:1506.0133].

[102] J. Maldacena, "Black holes and quantum information", *Nature Reviews Physics* **2**, 123–125, 2020.