Imperial College London

IMPERIAL COLLEGE LONDON

DEPARTMENT OF PHYSICS

Resolving a paradox: AdS/CFT and black hole information loss

Author:
Andrew King

Supervisor:
Daniel Waldram

A thesis submitted in partial fullfillment for the degree

MSc Quantum Fields and Fundamental Forces

23rd September 2022

In this report, the recent developments towards understanding black hole dynamics and resolving the black hole information paradox through the lens of the AdS/CFT correspondence are reviewed. The origin of the paradox from consideration of black hole evaporation and information is presented. The main tenets of the AdS/CFT correspondence are reviewed, motivated by the origins of the duality from string theory, and by scalar dynamics in AdS spacetime. Applying the holographic principle of the AdS/CFT duality to translate a non-gravitational quantum theory to a gravity theory in one higher dimension, the entanglement entropy of the Hawking radiation of an evaporating black hole is calculated to show whether the unitary Page curve is followed. Finally, recent insight into the evolution of quantum extremal surfaces is used to test whether a black hole-radiation system in AdS_2 obeys unitarity.

Contents

Chapter 1

Introduction

For decades, efforts have been made to understand the dynamics of the beginning of the Universe where the effects of quantum gravity (QG) are important. QG effects come into play when considering QFTs at short-length scales of the order of the Plank length, l_p. It is currently unknown how to quantise gravity using standard perturbative methods.

Progress in the last few years has been made by considering a similar physical process: the dynamics of black holes (BHs). BHs contain a singularity within the event horizon that requires QG to describe. However, as BHs are embedded within the universe, they can be studied by an exterior observer at large distances from the center. This allows a semi-classical approximation of QG to be used, where quantum (matter) fields propagate on the classical background geometry.

Black holes were previously thought to possess, through classical arguments, zero temperature (as no particles are able to escape the event horizon) and, by the no hair theorem [1], zero entropy. However, in 1974, Berkenstein argued that if a BH's entropy was zero, then the second law of thermodynamics,

$$\delta S \geq 0, \tag{1.1}$$

would be violated as an arbitrary amount of matter could fall into the black hole and decrease the overall entropy of the Universe as seen by an outside observer. To resolve this, a generalised entropy composed of the entropy of the two regions, the interior of the black hole and the exterior Universe, was proposed, $S_{gen} = S_{ext} + S_{BH}$, as well as a generalised second law of thermodynamics for black holes,

$$\delta S_{gen} = \delta \left(S_{ext} + S_{BH} \right) \geq 0, \tag{1.2}$$

where S_{ext} has contributions from the matter, gravitons outside the BH and vacuum contributions from quantum fields. Later in 1974, Hawking used the semi-classical approximation to quantum gravity to show that BHs radiate particles, called *Hawking radiation*, due to the instability of the QFT vacuum coupled to the classical black hole background geometry. Heuristically, this Hawking radiation is formed by entangled particle-antiparticle pairs close to the BH event horizon, one of which escapes to infinity and one which crosses the event horizon and falls into the black hole interior. The radiation emitted by black holes as measured by an observer at infinity has been shown to follow the black-body spectrum corresponding to a finite

temperature. Hence, black holes are thermal objects with a temperature called the Hawking temperature [2], T_H,

$$T_H = \frac{\hbar\kappa}{2\pi}, \tag{1.3}$$

and an entropy proportional to the area of the BH event horizon, called the Berkenstein-Hawking (B-H) entropy, S_{B-H},

$$S_{B-H} = \frac{Area}{4\hbar G_N}, \tag{1.4}$$

where G_N is Newton's constant. This was the first realisation of the *holographic principle*, relating the degrees of freedom of a region in d dimensions to its boundary in one lower dimension. Berkenstein also later proposed an entropy bound for arbitrary regions of spacetime that precisely saturates the B-H entropy of a black hole (1.4) [3]. This bound places a maximum limit on the amount of entropy that a region can contain as the area of the region's boundary in Planck units,

$$S_{bound} \leq \frac{Area}{4l_p^2}. \tag{1.5}$$

The bound prevents an arbitrarily large amount of information accumulating in a finite region, which would require an infinite amount of energy.

Inspired by the Berkenstein bound (1.5), 't Hooft [4] and Susskind [5] developed the holographic principle, a conjecture that a theory of QG in d dimensions is physically equivalent, or dual, to a non-gravitational QFT living on the boundary of the space in one lower dimension. In section 2, the AdS/CFT correspondence will be developed as a concrete example of the holographic principle, proposed by Maldacena in 1997 [6]. A result of this correspondence states that black holes in Anti-de Sitter (AdS) spacetime are dual to non-gravitational quantum systems (with boundary conditions). As the dual quantum theory manifestly obeys unitarity, the AdS/CFT correspondence provides evidence for (AdS) black holes obeying unitarity. Hence, extending to any spactime geometry, the holographic principle implies BHs in general are unitary and conserve (quantum) information and entropy.

In 1976, Hawking proposed, in a landmark paper, the presence of an apparent paradox involving the quantum information associated to a system with a black hole, known as the *black hole information loss paradox* [7]. The paradox is a consequence of black hole evaporation: as a black hole of mass M emits radiation, its mass will reduce until a time $t_{evap} \sim M^3$, at which point the black hole will evaporate, leaving behind only the emitted thermal radiation. During the evaporation, the Berkenstein-Hawking entropy associated with the black hole's area (1.4) will decrease as the mass and event horizon area shrink, $S_{BH} \to 0$ as $t \to t_{evap}$. Conversely, the entanglement entropy of the radiation escaping to $r \to \infty$, $S_{rad} = -Tr_{BH}\left(\rho_{rad} log\left(\rho_{rad}\right)\right)$[1], increases monotonically with time as the black hole emits radiation, following the "Hawking curve". This process violates information conservation and unitarity: in general, the final entropy of emitted radiation can be larger than the initial B-H entropy, $S^i_{B-H} < S^f_{rad}$. Hence,

[1] Tr_{BH} traces out the BH degrees of freedom, and ρ_{rad} is the density matrix for the radiation.

the total entropy (a measure of the disorder of a system) increases over time – the Universe appears to lose information during black hole formation and evaporation, violating unitarity and the second law of thermodynamics.

Furthermore, black hole evaporation implies that the entanglement, or von Neumann, entropy[2] associated to the Hawking radiation isn't conserved, violating unitarity. The time at which the two entropies coincide, $S_{B-H} = S_{rad}$, is called the *Page time*, denoted t_{Page}. The fields comprising the Hawking radiation form a pure state with their partner particles in the BH interior. The entanglement entropy of the two-particle system is initially zero. Before t_{Page}, when $S_{B-H} > S_{rad}$, the degrees of freedom of the particles contained in the interior of the BH can be completely described by the BH degrees of freedom $\propto Area_{BH}$, such that the interior and exterior particles remain in a pure, entangled state. However, after t_{Page}, when $S_{B-H} < S_{rad}$, the BH area doesn't contain enough degrees of freedom to describe the particles in the interior. Hence, after t_{Page}, the radiation-BH system can no longer be fully entangled, with the pure state reducing to a mixed state with non-zero entanglement entropy.

On general grounds, it was later argued that unitarity of the radiation-BH system can be enforced if, instead of the entropy of radiation following the Hawking curve, it follows the *Page curve*: increasing monotonically up until the Page time, and then decreasing to zero as $t \to t_{evap}$ [8, 9]. However, lacking a full theory of QG to describe the precise state of the radiation-BH system, determining the entropy of the system is difficult. Furthermore, for several years, it was unclear how to formulate the system such that its evolution obeys the Page curve, and whether, in fact, BHs evolve unitarily (Hawking famously denounced this). The proposal of the AdS/CFT correspondence in 1997, as well as providing concrete evidence for BH unitarity, provided a means to calculate the unitary evolution of the radiation-BH system.

The AdS/CFT correspondence states that a theory of QG in a $(d+1)$-dimensional spacetime asymptotic to AdS space is dual to a conformal field theory (CFT) in d dimensions. In 2006, Ryu and Takayanagi (RT) used the AdS/CFT correspondence to calculate the entanglement entropy of a CFT in $(d+1)$-dimensions (coupled to a gravity theory) via. the area of d-dim. minimal extremal surfaces in (asymptotically) AdS_{d+2} bulk spacetime [10]. Using the AdS/CFT correspondence makes calculating the generalised entropy S_{gen} of the BH-radiation system far easier, as it translates the problem of calculating entropy into a geometrical problem where the entropy corresponds to an extremised and minimised surface. The RT prescription has since been extended to covariant theories [11], and to the formulation of quantum extremal surfaces (QESs) on which the entanglement entropy of an evaporating black hole over time is calculated [12–14]. Recent work has shown that in order to recreate the Page curve of unitary black hole evolution, a second, non-trivial QES must be considered after the Page time, when the state of the system transitions from pure to mixed. Before the Page time, the QES is trivial, with the entanglement region of the BH covering all of the interior region. At later times, the QES transitions to a non-trivial surface located close to the event horizon, with the phase transition occurring at the Page time. This excludes most of the BH interior from the BH entanglement wedge at later times, such that the entanglement entropy associated with the wedge shrinks to zero as the BH evaporates.

This prescription has been applied to determine the evolution of Hawking radiation for a black hole in a toy model 2d JT gravity theory with holographic matter coupled to a non-gravitational bath [15]. Translating the 2-dimensional gravity theory to its dual 3-dimensional description, the quantum extremal surfaces reduce to those found in the (covariant general-

[2] A concept that will be defined in Section 2.4

isation of the) RT prescription, and reproduces the unitary evolution of Hawking radiation predicted by the Page curve. This result rests on the inclusion of an "island" contribution to the full minimal surface over which the entanglement entropy is calculated, which reduces the entropy at late times by "purifying" the entangled modes contained in the island entanglement wedge in the BH interior with those contained in the exterior region.

As such, the aim of this dissertation is to show how key developments in recent years has lead towards resolving the Hawking information paradox.

In section 2, we will discuss the origin of BH entropy, the BH laws of thermodynamics and motivate the description of black holes as quantum systems. We will then define the entanglement entropy of quantum systems, and use this to describe the information paradox and the unitary Page curve.

In section 3, the AdS/CFT correspondence will be reviewed. By considering the case of coincident D3-branes in string theory, the AdS/CFT correspondence will be motivated through the prototypical example of $\mathcal{N} = 4$ SYM $\leftrightarrow AdS_5 \times S^5$. After defining CFTs and AdS space and discussing (scalar) dynamics in AdS, the AdS/CFT correspondence will be stated with the links between physical observables in the dual theories. Finally, we will describe how the holographic principle is a generalisation of the AdS/CFT correspondence, taking the entropy of black holes in AdS_3 as an example.

In section 4, the AdS/CFT correspondence will be applied to the calculation of entanglement entropy via. the Ryu-Takayanagi prescription, and used to define a generalised entropy in gravitational systems. Then, following recent work in [13, 14, 16], we will use the holographic principle to obtain the Page curve and resolve the black hole information paradox. Finally, in section 5, we will consider the toy model of a black hole coupled to a non-gravitational bath in simplified 2d JT gravity as a concrete example of the RT/HRT prescription producing unitary Hawking radiation evolution.

Chapter 2

Hawking radiation and black hole information

2.1 Hawking radiation

In classical General Relativity (GR), BHs have zero temperature and entropy, as no particles or radiation is able to pass from the interior, through the event horizon, to the exterior of a BH. However, when considering a classical gravity geometry coupled to a QFT, with quantum fields propagating on the classical background, i.e. the semi-classical gravity approximation, it was shown by Hawking in 1974 that BHs do possess a finite temperature, called the Hawking temperature (1.3) [2]. The Hawking temperature is a consequence of Hawking radiation emitted from the black hole region, which is required by the Unruh effect and the equivalence principle of relativity.

The Unruh effect is a kinematic prediction for a QFT, where an accelerated frame observes a thermal bath whilst an inertial observer doesn't. Heuristically, close to the BH event horizon, a local observer must accelerate to counter strong gravitational forces and prevent falling in. This accelerating observer will observe a thermal bath of particles that come out of the locally accelerating horizon, and then fall freely back in. Local thermal equilibrium and the equivalence principle requires the consistent extension of the thermal bath to a distance $r \to \infty$ from the BH. Hence, a finite temperature can be observed by an observer at large distance from the BH due to some of the emitted particles not being reabsorbed and forming Hawking radiation.

The result for the Hawking temperature (1.3) can be derived explicitly following [16]: considering Schwarzschild spacetime with metric,

$$ds^2 = -\left(1 - \frac{r_s}{r}\right) dt^2 + \frac{dr^2}{1 - \frac{r_s}{r}} + r^2 d\Omega_2^2. \tag{2.1}$$

Taking a change of coordinates, $t \to 4M\tau$ and $r \to 2M + \frac{\rho^2}{8M}$, and fixing the angular directions, the metric (2.1) expanded around the event horizon $r = 2M$ for $\rho << 2M$ becomes,

$$ds^2 \approx -\rho^2 d\tau^2 + d\rho^2. \tag{2.2}$$

The metric above is in Rindler coordinates for an accelerated observer with local acceleration $a = 1/\rho$. Physically, an observer close to the event horizon must have a constant acceleration to remain at a fixed distance from the event horizon, in order to counter the strong gravitational attraction. For $\rho \to 0$, $a = 1/\rho \to \infty$, an infinite acceleration is required to escape the BH at the event horizon.

Performing another change of coordinates, taking $(\rho, \tau) \to (x_0, x_1)$ and applying the chain rule, the metric (2.2) reduces to Minkowski locally,

$$ds^2 \approx -\left(dx^0\right)^2 + \left(dx^1\right)^2. \tag{2.3}$$

Hence, a local observer falling into the BH observers flat space locally, with the geometry extending smoothly past the event horizon at $r = 2M$. This contrasts the observations made by an exterior observer at $r >> 2M$, for which no signals can reach from within the BH interior.

Applying the equivalence principle, non-inertial, free-falling observers described by (2.2) are equivalently described by an accelerated observer in Minkowski space (2.3) (see Fig. 2.1). For ρ fixed, an observer is following a uniformly accelerated geodesic in Minkowski space.

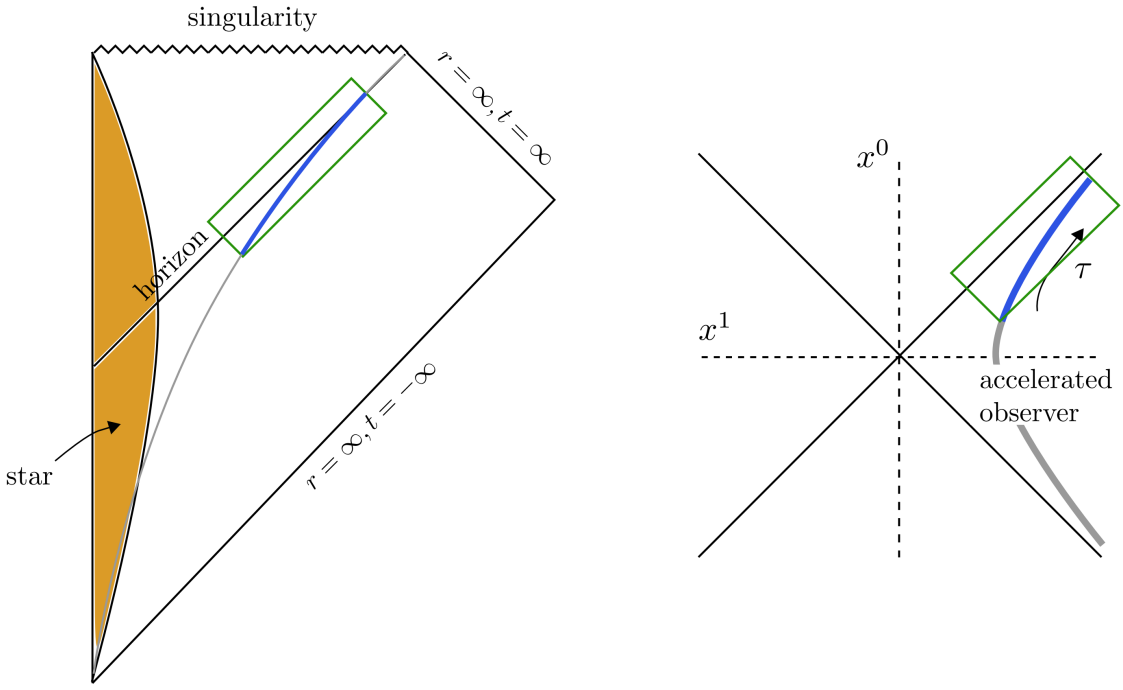


Figure 2.1: (a) Penrose diagram for matter collapsing into a black hole; (b) "zoomed-in" near-horizon region in x_0, x_1 coordinates, showing a uniformly accelerated observer at $a = \rho^{-1}$. [16]

The Unruh effect states that for an accelerating frame observing a QFT, the fields are excited at a local temperature, which for an observer at a fixed distance close to the horizon is given by,

$$T = \frac{a}{4\pi} = \frac{1}{\pi\rho} = \frac{1}{4\pi\sqrt{2Mr\left(1 - \frac{2M}{r}\right)}}. \tag{2.4}$$

Extending this to observers at large distance $r >> 2M$ from the event horizon, gravitational redshift of the temperature by a factor $\sqrt{g_{00}}$ takes place such that there exists a thermal background everywhere, with local temperature $T(r')$ at distance r' [17],

$$T(r') = \frac{1}{4\pi\sqrt{2Mr\left(1-\frac{2M}{r}\right)}}\sqrt{\frac{1-\frac{2M}{r}}{1-\frac{2M}{r'}}} = \frac{1}{4\pi\sqrt{2Mr\left(1-\frac{2M}{r'}\right)}}. \tag{2.5}$$

For $r' \to \infty$, $r' >> 2M$,

$$T(\infty) = T_H = \frac{1}{4\pi\sqrt{2Mr}} \simeq \frac{1}{8\pi M}, \tag{2.6}$$

where we have used Planck units throughout, and taken $r \simeq 2M$ to go from the second to third equality. Hence, a field theory defined on a BH background is a thermal state with temperature at large distance from the event horizon given by the Hawking temperature, T_H.

2.2 Black hole thermodynamics and entropy bound

Using the first law of BH thermodynamics and the Hawking temperature of the field theory in a BH geometry, the Berkenstein-Hawking entropy, which acts as a minimum bound for the entropy contained within a region, can be calculated.

In 1973, the laws of BH thermodynamics were published by Bardeen, Carter and Hawking [18], which extended the existing classical thermodynamic laws to black holes. The first law of thermodynamics is a statement of conservation of energy with respect to a change in mass M, charge Q and angular momentum J of a BH and governs the response of the area of a rotating BH to a change in these three quantities,

$$dM = \frac{1}{8\pi}\kappa dA + \Omega_H dJ + \Phi_H dQ, \tag{2.7}$$

where Ω_H is the angular velocity of the BH and Φ_H is a constant of the BH. This is in analogy with the classical first law of thermodynamics,

$$dE = TdS + \sum_i \mu_i N_i. \tag{2.8}$$

Applying the Hawking temperature (1.3) to the first law of thermodynamics (2.8), and noting that the internal energy of a black hole is determined by its mass, $E = M_{bh}$, and the horizon area is $A = 4\pi r_s^2 = 4\pi\left(2M\right)^2$, then

$$dS = \frac{dQ}{T} = 8\pi M dQ, \tag{2.9}$$

and performing integration on both sides and setting the integration constant to zero by the initial condition of a BH of $S = 0$ for $r = 0$,

$$S_{BH} = \pi r_s^2 = \frac{Area}{4} = S_{B-H}. \tag{2.10}$$

This is the Berkenstein-Hawking entropy formula, with S_{B-H} an upper bound that places a maximum on the amount of information contained within an arbitrary region of space. Hence, the entropy of a black hole maximally saturates the information content within its region of spacetime.

The Berkenstein-Hawking entropy was the first realisation of the holographic principle, where the physical description of a gravity theory is encoded onto the bounding surface of the region.

2.3 Black hole as a quantum system

Whilst we now have an expression for the entropy of a black hole (2.10), dependent only on the area of the horizon, this quantity lacks a physical meaning: how can the entropy (and the associated classical statistical degrees of freedom) be interpreted?

One interpretation of black hole entropy and information is to view the interior of the black hole as a quantum system. Namely, as seen by an outside observer, BHs can be viewed as a quantum system with $\frac{A}{4G_N}$ number of degrees of freedom evolving unitarily, with a "cut-off" surface separating the quantum system from the exterior defined within a Planck length of the event horizon and defining the black hole region. This has been coined the *central dogma* of black hole quantum information, inspired from its use in molecular biology where genetic information can be viewed increasingly abstractly as DNA, then RNA and finally proteins [16]. Notably, Hawking strongly opposed this hypothesis.

The central dogma hypothesis is supported by the counting of microstates of supersymmetric extremal BHs in string theory [19]. By counting the degeneracy of BPS soliton bound states, the entropy reproduces the area formula (2.10) plus corrections.

Furthermore, a consequence of the AdS/CFT correspondence (introduced in Chp. 3) is that, in all known cases, BHs in a gravity theory region are related to solutions of QFTs at finite temperature on the boundary of the region. Hence, as QFTs are unitary, this supports the claim that black holes are unitary and have no loss of information. Also, the thermal radiation of a blackbody (which is unitary) isn't equal to the Hawking radiation of BHs, implying Hawking radiation isn't unitary. Hence, the AdS/CFT correspondence implies that BH thermal radiation is Hawking radiation with corrections.

2.4 Entanglement entropy

In order to quantify the information and entropy of the radiation-BH system, it is necessary to define the entanglement, or fine-grained, entropy of a quantum system.

In quantum mechanics (QM), a quantum state is described by a vector ϕ in the Hilbert space $\mathcal{H}$. A central object in QM is the density operator (or matrix) ρ, defined for a pure state as $\rho = |\psi\rangle\langle\psi|$. The density operator is Hermitian ($\rho = \rho^\dagger$) and positive definite ($|\rho|^2 \geq 0$). The expectation value of an operator $\hat{O}$ is expressed as,

$$\langle\psi|\hat{O}|\psi\rangle = Tr(\hat{O}\rho). \tag{2.11}$$

For a pure state, it follows that $Tr(\rho) = \langle\psi|\psi\rangle = 1$ and $Tr(\rho^2) = \langle\psi|\psi\rangle^2 = 1$. A mixed state, i.e. where the state can be expressed as $|\psi\rangle = \sum_i \sqrt{p_i}|\psi_i\rangle$ for $p_i \geq 0$ and $|\psi_i\rangle$ an orthonormal basis, has a density operator $\rho = \sum_i p_i|\psi_i\rangle\langle\psi_i|$. It follows that $Tr\rho = \sum_i p_i = 1$ as for the pure state, but, from [20, 21],

$$Tr\rho^2 = \sum_i \left(p_i\left(p_i - 1\right) + 1\right) = \sum_i (p_i(p_i - 1) + 1) < 1. \tag{2.12}$$

Dividing a quantum system into two subsystems, A and B, the density matrix ρ lives on the Hilbert space $\mathcal{H} = \mathcal{H}_A + \mathcal{H}_B$ [22]. Restricting an observer's access to only one of the subsystems, i.e. $\mathcal{H}_A$, the reduced density matrix is given as the trace over the complementary, i.e. $\mathcal{H}_B$, space of the total density matrix,

$$\rho_A = Tr_{\mathcal{H}_B}\rho, \tag{2.13}$$

as measured by an observer in subsystem A. Then, using the von-Neumman entropy,

$$S_{vN} = -Tr(\rho) \tag{2.14}$$

The entanglement, or fine-grained, entropy for a subsystem A with reduced density matrix ρ_A, is defined as,

$$S_A = -Tr_{\mathcal{H}_A}[\rho_A ln \rho_A] \tag{2.15}$$

Physically, the entanglement entropy is a measure of how much a given quantum system is quantum mechanically entangled, or how much (quantum) information is omitted through the exclusion of a region of the space.

In the example of a pure state ϕ on a Hilbert space split into 2 sub-spaces, $\mathcal{H} = \mathcal{H}_A + \mathcal{H}_B$, the total entanglement entropy of the system is zero,

$$\begin{aligned} S_{tot} = S_{A\cup B} &= -Tr_{\mathcal{H}_{A\cup B}}[\rho] \\ &= -Tr_{\mathcal{H}_{A\cup B}}\left[|\Psi\rangle\langle\Psi|\sum_{n=1}^{\infty} c_n(|\Psi\rangle\langle\Psi| - 1)^n\right] \\ &= -\langle\Psi|\Psi\rangle\langle\Psi|\sum_{n=1}^{\infty} c_n(|\Psi\rangle\langle\Psi| - 1)^n|\Psi\rangle \\ &= \sum_{n=1}^{\infty} c_n\langle\Psi|(|\Psi\rangle| - 1)|\Psi\rangle = 0, \end{aligned} \tag{2.16}$$

whilst each subsystem considered individually (with the complementary subsystem hidden from view) has non-zero entanglement entropy. This can be seen explicitly for a bipartite (2 subsystem) quantum system, where we consider a mixed Bell state,

$$|\psi\rangle_{AB} = \frac{1}{2}(|0,1\rangle + |1,0\rangle). \tag{2.17}$$

The corresponding density matrix has the form,

$$\begin{aligned}\rho = |\psi\rangle_{AB}\langle\psi|_{AB} &= \frac{1}{2}(|0,1\rangle + |1,0\rangle)(\langle 0,1| + \langle 1,0|) \\ &= \frac{1}{2}\Big((|0\rangle\langle 0|)_A(|0\rangle\langle 0|)_B + (|0\langle 1|)_A(|0\rangle\langle 1|)_B + ... \\ &\qquad ... + (|1\rangle\langle 0|)_A(|1\rangle\langle 0|)_B + (|1\rangle\langle 1|)_A(|1\rangle\langle 1|)_B\Big),\end{aligned} \tag{2.18}$$

which means that tracing out the total density matrix with respect to sub-system B will give,

$$\rho_A = Tr_B\rho = \frac{1}{2}\Big((|0_A\langle 0|)_A + (|1\rangle_A\langle 1|)_A\Big) \equiv \frac{1}{2}1_A \tag{2.19}$$

This gives an entanglement entropy of,

$$S_A = ln(2) \neq 0 \tag{2.20}$$

The Bell state (2.17) is the maximally mixed, maximally entangled state of this quantum system, with (2.20) the highest deviation that a state can obtain from a pure state with $S = 0$. More generally, maximally-mixed states have an entanglement entropy $S = ln(N)$, where N is the dimensionality of the Hilbert space. Hence, the entanglement entropy measures the degree of divergence of a quantum system from the pure state, and corresponds to the amount of information lost by assuming that the subsystem B is not visible.

The entanglement (von Neumann) entropy has several important properties:

1. $0 \leq S(\rho) \leq ln(N)$, where the lower bound is for a pure state with no entanglement, and the upper bound is for a maximally mixed state with maximal entanglement;
2. For a pure state, the density matrix is idempotent, $\rho = \rho^2$;
3. $S(\rho)$ is invariant under unitary time evolution, obeying unitary: $S(\rho) \to S(U^\dagger\rho U) = S(\rho)$ for $\rho(0) \to \rho(t) = U^\dagger(t)\rho(0)U(t)$. This means, a pure (mixed) state will remain pure (mixed);
4. For a pure state at zero temperature with ρ, then $S_A = S_B$: the entanglement entropy of the observed region is the same as the complementary region. Hence, the entanglement entropy is *not* an extensive property, independent of the system scale (this property does not hold for finite temperature systems);
5. Strong subadditivity (SSA): for 3 subsystems, $\mathcal{H} = \mathcal{H}_A + \mathcal{H}_B + \mathcal{H}_C$, with any (mixed or pure) total density matrix ρ_{tot},

$$S_{A+B+C} + S_B \leq S_{A+B} = S_{B+C} \quad \text{and} \quad S_A + S_C \leq S_{A+B} + S_{B+C}. \tag{2.21}$$

6. Entanglement entropy can be defined for quantum fields on a surface, σ, at a fixed time through the corresponding density matrix $\rho_\Sigma, S(\Sigma) \equiv S(\rho_\Sigma)$. As a consequence of unitarity, this quantity is the same for any Cauchy surface (at any time) that possesses the same causal diamond (see 2.2).

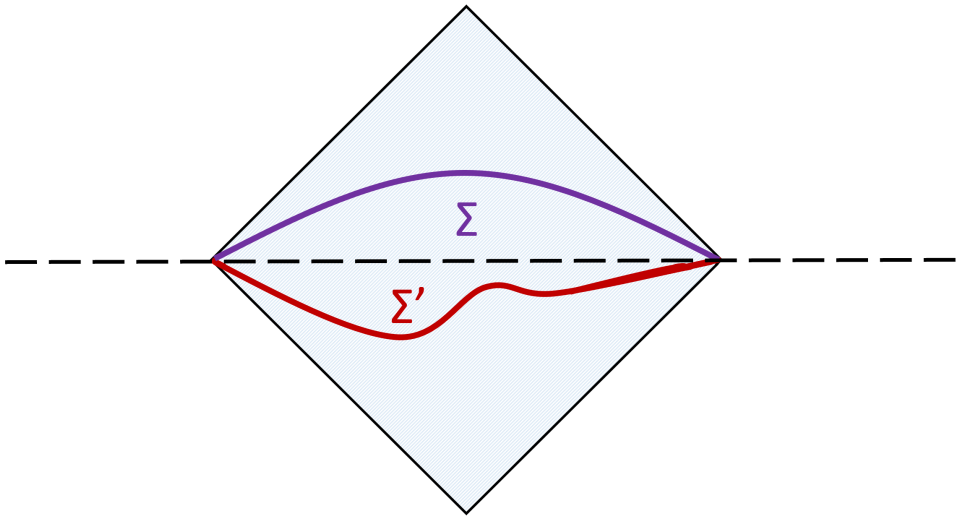


Figure 2.2: Casual diamond shared by Cauchy surfaces Σ and $\tilde{\Sigma}$. S_{vN} is the same on both slices. Adapted from [16].

There also exists a second definition of the entropy associated to a quantum system: the coarse-grained, or thermodynamic, entropy. Considering a subset of macroscopic (simple) observables O_i, with $\langle O_i \rangle = Tr(\rho O_i)$ for density matrix ρ, vary a different density matrix $\tilde{\rho}$ to consider all possible matrices that produce the same observable expectation values,

$$\langle O_i \rangle = Tr(\rho O_i) = Tr(\tilde{\rho} O_i) = \langle \tilde{O}_i \rangle, \tag{2.22}$$

for all O_i. Then, of the set of possible entries for $\tilde{\rho}$, the one that produces the maximum value of the corresponding entanglement entropy defines the course-grained entropy, S_c,

$$S_c = max_{\tilde{\rho}} \big(- Tr(\tilde{\rho} log \tilde{\rho}) \big) \tag{2.23}$$

Hence, the coarse-grained, or thermodynamic, entropy is computed by maximising the fine-grained entropy over all possible states described by $\tilde{\rho}$ that satisfy the condition (2.22). From its definition, the coarse-grained entropy satisfies,

$$S_{vN} \leq S_c, \tag{2.24}$$

such that the S_c acts as as an upper bound on the number of degrees of freedom in the system, and the amount of entanglement that states can have. Also, the coarse-grained entropy satisfies the second law of thermodynamics, $\delta S_c \geq 0$.

2.5 The information paradox

Having introduced the entropies describing the information of a quantum system, the information paradox, formulated by Hawking in 1976 [2], can be fully introduced.

The thermal properties of black holes arise from the splitting of a vacuum state near the event horizon into 2 entangled particles, one of which falls into the BH and the other which escapes to infinity and becomes Hawking radiation [16]. If the vacuum state in the QFT is a pure state initially, the splitting of the particles creates 2 mixed states, which are entangled at short distances and possess finite entanglement entropy (entangled degrees of freedom) individually. Also, from the central dogma, the BH can be viewed as a quantum system with a finite number of degrees of freedom. Hence, the BH and Hawking radiation are each subsystems of a bipartite quantum system, pure when combined, but mixed when one subsystem is considered in isolation.

The BH will emit Hawking radiation as it evaporates, with the entanglement degrees of freedom of the infalling (outgoing) particles contained within the BH interior (exterior). As Hawking radiation continues to be emitted, the entangled degrees of freedom corresponding to the particles in the interior (entangled with those escaping to infinity) become a larger share of the total degrees of freedom contained within the BH region. Also, as the BH shrinks via. Hawking radiation, the coarse-grained (thermodynamic) entropy of the BH corresponding to the Berkenstein-Hawking entropy (1.3) reduces proportional to the area of the horizon.

At early times, the (entanglement) entropy of Hawking radiation, S_{rad}, is low, whilst the thermodynamic (Berkenstein-Hawking) entropy of the BH, S_{BH}, remains high. However, as the BH evaporates, S_{BH} falls whilst S_{rad} continues to rise (see Fig. 2.3). At a finite time, the Page time t_{Page}, S_{rad} will exceed S_{BH}, such that the BH no longer contains enough degrees of freedom in the interior to remain entangled to the emitted radiation. Also, the degrees of freedom contained within the BH region will exceed the maximum amount allowed by the Berkenstein bound (1.4). Hence, following the Hawking curve proposed in [2], the BH-radiation system transitions from a pure state initially to a mixed state at late times $t > t_{Page}$, violating unitarity and subsequently the central dogma.

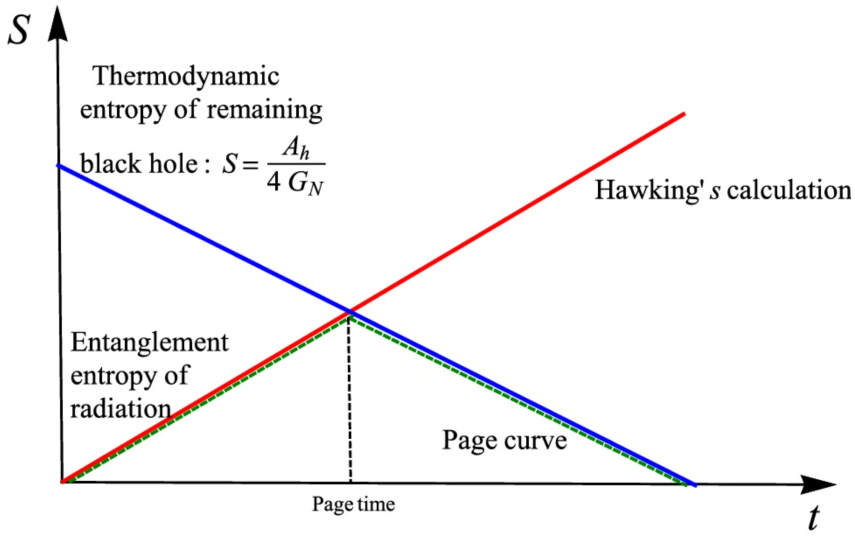


Figure 2.3: Evolution of entropy for: the Hawking curve (red), thermodynamic BH entropy (blue), and the Page curve (dashed green). [23]

Physically, this means that as the BH evaporates, information is lost from the BH-radiation system when the BH vanishes. This is illustrated by the "baby universe" description of the evaporating black hole (Fig. 2.4), where the branching off of an interior "baby universe" (containing a singularity) from the exterior spacetime means that unitarity is violated as modes in the baby universe can't evolve past the singularity. This highlights that although the information of the infalling particles is no longer observed by an observer outside of the black hole region, it is still

present in the system and the particles in the interior and exterior remain entangled. However, after the BH evaporates and vanishes at late times, the information in the BH interior is lost, and the system transitions from a pure to mixed state, violating unitarity.

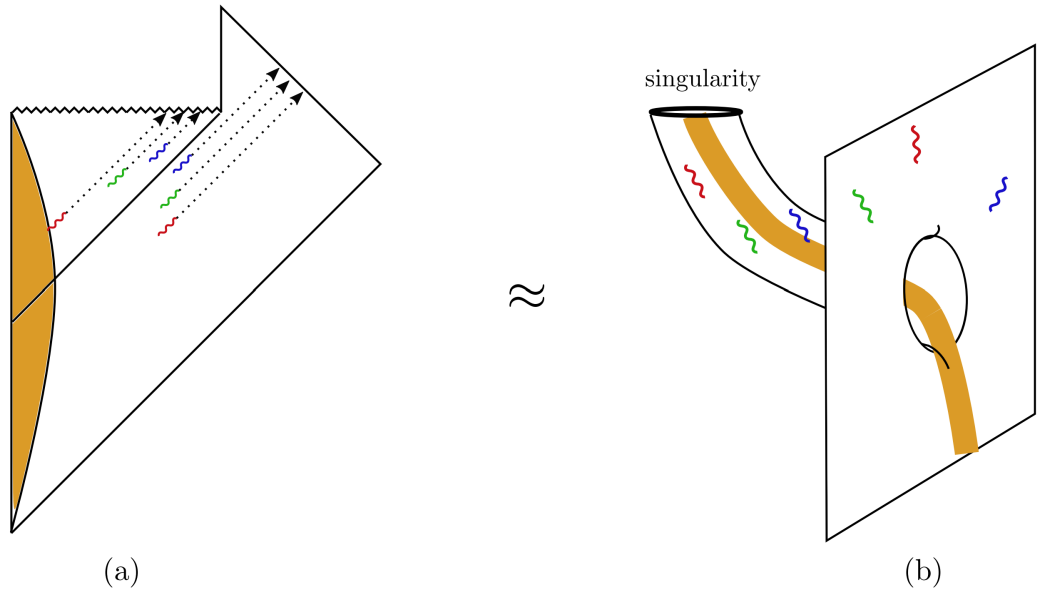


Figure 2.4: (a) Penrose diagram for an evaporating black hole, with pairs of entangled modes split between the interior and exterior of the BH; (b) the equivalent "baby universe" picture of the black hole evaporation, where the BH interior is represented by the smaller branched universe containing the singularity. [16]

However, this is just one aspect of the information paradox involving BH entropy. Other important elements include:

1. The full Schwarzschild solution containing a black hole (Fig. 2.5(a)) has 2 exterior regions (I and II) and 2 interior regions (III and IV). Hence, hoes does the Berkenstein-Hawking entropy arise and what does it physically correspond to?
2. "Bags of gold" paradox, formulated by Wheeler [24], shows that some classical geometries exist that look like a black hole from an asymptotic region, but can have arbitrarily large entropy past the "throat" larger than the area of the horizon (Fig. 2.5(b)). This violates the B-H entropy and central dogma, where the entropy of a BH region is proportional to its area.

The former was resolved by Maldacena and Susskind [26] through the ER=EPR conjecture, which states that 2 BHs connected by a wormhole (or Einstein-Rosen bridge) in the interior is equivalent to 2 entangled particles (or Einstein-Podolsky-Rosen pair) contained within each BH. Therefore, through the ER=EPR conjecture, the 2 BHs and wormhole system form a pure state with zero entropy. However, taking each BH separately creates mixed states with finite entropy due to loss of information of the entangled particles.

The "bags of gold" paradox is another violation of the central dogma that is also caused by the Hawking curve of radiation entropy, and is resolved in a similar way. Assuming that the central dogma is true, the resolution of this paradox requires the entanglement entropy of emitted radiation to follow the Page curve (green dashed line in Fig. 2.3) as opposed to the Hawking curve (red line): increasing monotonically up to t_{Page}, and then decreasing to track

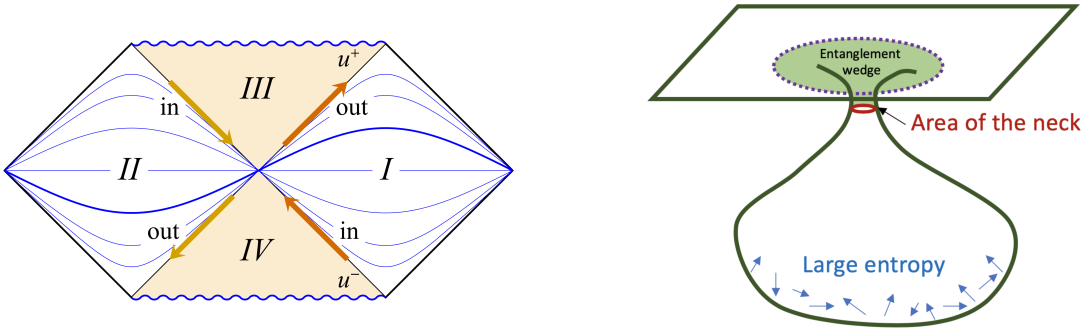


Figure 2.5: (a) the Penrose diagram for the full Schwarzschild solution, with exterior regions I, II and interior regions III and IV [25]; (b) classical geometry with the area of the "neck" independent of the number of degrees of freedom contained in the interior: giving the "bags of gold" paradox [16].

S_{BH} such that all radiation degrees of freedom remain entangled with the BH interior degrees of freedom.

The Page curve was first calculated by Don Page in 1993 [8] using semi-classical gravity and based on simple properties of the entanglement entropy, but is also valid for the BH area $A \to 0$ as $t \to t_{evap}$, where the radius of the BH is of the order of the Planck length [9].

Although the Page curve has been successfully determined, a lack of physical motivation and interpretation has persisted for four decades, with a full theory of QG previously thought to be required to shed light on the origins of the Page curve and unitarity of BHs through study of the BH interior close to the singularity. With the discovery of the AdS/CFT correspondence in the late '90s, the proposal of a holographic entanglement entropy using the holographic principle has lead to, in addition to development of quantum extremal surfaces, the reproduction of the Page curve in black hole systems and a step towards resolving the black hole information paradox.

Chapter 3

AdS/CFT correspondence

In this section, the key developments and dictionary of the AdS/CFT, or gauge-gravity, duality will be reviewed. It emerged from the world sheet duality in string theory and the observation of the duality between D-branes and black branes in string theory.

In the search for a theory of quantum gravity (QG), string theory has emerged as a consistent and desirable candidate in the last 30 years. The central premise of string theory is that the point-like particles of QFTs are replaced by extended objects called strings which oscillate at different "frequencies", subject to boundary conditions, to recreate the full spectrum of particles observed at low energies.

A glimpse of how the AdS/CFT correspondence emerges from studying string theory can already be seen through the original motivation for its development: to offer an explanation for the large number of hadrons and mesons being experimentally discovered in the 1960's. The original goal was to describe this spectrum of particles, and although it was partially successful at predicting the spin–mass relation $m^2 \approx TJ^2 + const.$ for light hadrons (modelling the hadrons as a rotating relativistic string with mass, angular momentum and tension), it was soon superseded by the discovery of quarks and the development of QCD.

QCD is a gauge theory with gauge group $SU(3)$ ($N = 3$ colours), which at low energies becomes strongly coupled such that perturbative calculations become difficult. 't Hooft [27] developed a method to simplify the calculation when $N \to \infty$. The idea was to perform an expansion in $1/N$ when N is large to get the exact spectrum, then do a $1/N = 1/3$ expansion. As string theory gives the correct relation between mass and angular momentum, it is expected that $N = 3$ and $N \to \infty$ are similar – hence, the large N limit connects string theories with gauge theories. This is true generally: different gauge theories will correspond to different types of string theory. As string theory is a valid description of QG, this analysis motivates a more general duality between theories of QG and gauge theories in the large N limit. The most famous example of this is $\mathcal{N} = 4$ Super Yangs-Mills gauge theory dual to Type IIB string theory in $AdS_5 \times S^5$.

3.1 Black branes/D-branes duality

The AdS/CFT correspondence emerged from the study of p-branes in string theory, and in particular the observation that D-branes are equivalent to black p-branes, which are classical solutions to supergravity (SUGRA) at the low-energy limit of string theory ($\alpha' \to 0$).

In perturbative string theory, Dp-branes are $(p+1)$-dimensional hypersurfaces upon which p-branes can end (see Fig. 3.2). The massless modes of the open string describe the oscillations of the brane, gauge fields living on the brane, and fermionic partners [28]. A Dp-brane can be charged by a $(p+1)$-form gauge potential, $A_{(p+1)}$, in extension of a 0-brane (particle) charged under a 1-form gauge potential (the electromagnetic potential A_μ), with an associated $(p+2)$-dim. field strength tensor $F_{(p+2)}$. Introducing N coincident Dp-branes, open strings will have N^2 possible endpoints and a flux of the $F_{(p+2)}$ field strength between branes will contribute to the stress-energy tensor, inducing a curved geometry.

The N coincident Dp-branes are low-energy, SUGRA solutions that resemble those of extremal, charged black holes in ordinary general relativity generalised to black *branes* in p spatial dimensions. As an observer approaches the black p-brane event horizon, energy is redshifted (as measured by an observer at ∞) due to large gravitational potential and becomes very small. Low-energy, classical SUGRA is valid when the curvature of the extremal p-brane geometry (characterised by the radius of the event horizon r_+) is much smaller than the string scale, such that quantum effects can be neglected: $r_+ \propto L_{geom} >> l_s$.

Alternatively, the extremal p-brane can be reformulated in terms of D-branes using the worldsheet duality. In string theory, the *worldsheet duality* states that a worldsheet with boundaries on a D-brane can be thought of as either a closed string or open string scattering process depending on the boundary conditions chosen [28, 29]. Regardless, the scattering amplitudes for either open or closed strings are the same, forming a duality between open and closed string scattering amplitudes. For example, a cylindrical worldsheet stretched between 2 D-brane worldvolumes can be represented by 2 different but equivalent processes (Fig. 3.1): either (a) a closed string (red) emitted by one of the branes and absorbed by the other; or (b) two open strings (blue), bounded by the two branes, propagating around the worldvolume in a 1-loop vacuum diagram [30, 31].

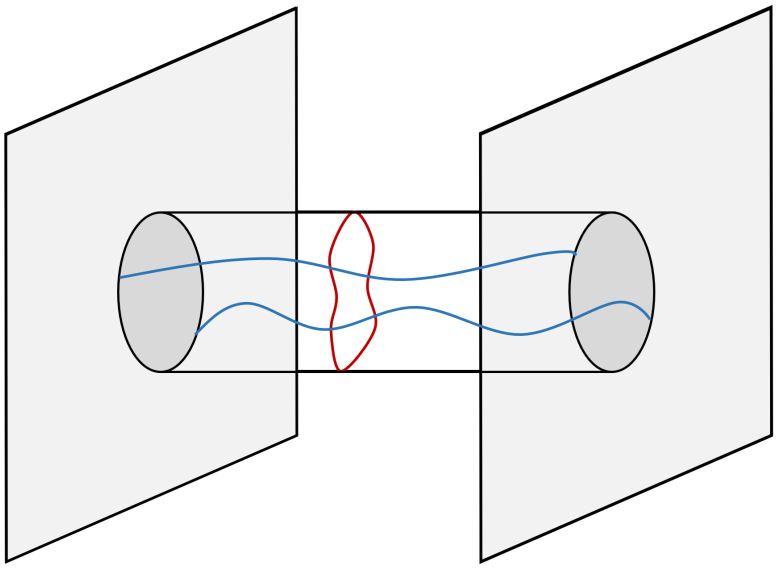


Figure 3.1: Open-closed string (worldsheet) duality. Inspired by [30].

Due to the worldsheet duality, the D-brane can be a source of closed strings as well as a hypersurface for open strings (see Fig. 3.2)). In particular, for N coincident Dp-branes, the R-R charges carried by each individual D-brane sum to produce a $(p+1)$-form charge of N units on

the $(p+1)$-dimensional D-brane hypersurfaces [32]. In the low-energy regime, the dynamics on the N coincident Dp-branes reduces to a gauge theory (with conformal invariance) in general.

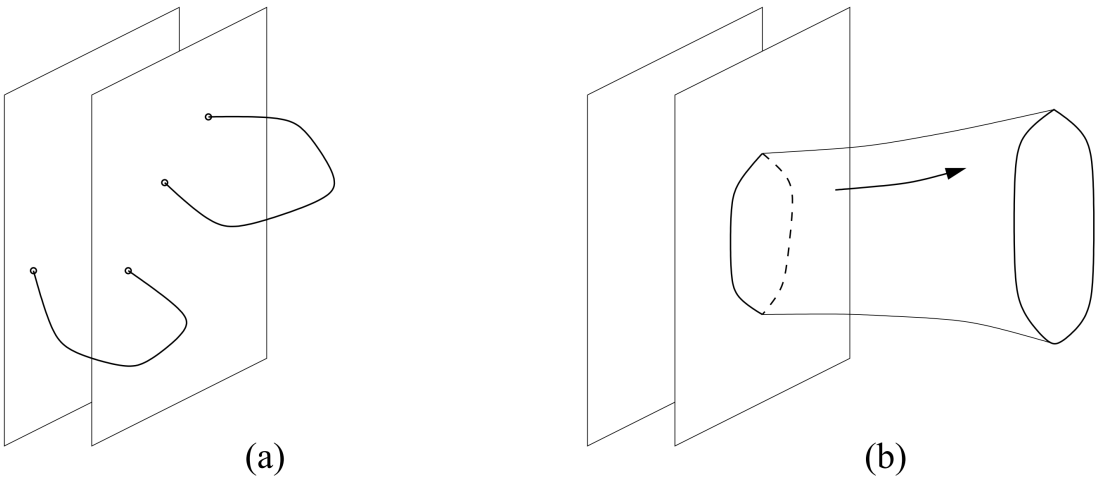


Figure 3.2: The D-brane as: (a) an endpoint for open strings; (b) a source of closed strings. [28]

As the Dp-brane description involves the string worldsheet, it will be well-described in string perturbation theory where the effective loop expansion for N coincident D-branes is in the weak limit, $\lambda \propto g_s N << 1$.

Hence, from the worldsheet duality, there exist two descriptions of the low-energy regime of N coincident D-branes: the near-horizon geometry of a black p-brane for strong effective coupling ($\lambda >> 1$); and a conformal gauge theory (or SUSY CFT) at strong coupling ($g_{gauge} >> 1$) and weak effective coupling ($\lambda << 1$).

This highlights one of the main powers of the AdS/CFT correspondence: although gauge theories at strong coupling are difficult to evaluate exactly, there exists a dual gravity theory at weak effective coupling where calculations of physical quantities can be performed. The example of N parallel $D3$-branes provides a concrete realisation of the above gauge-gravity duality sketch, and is the prototypical example explored in Maldacena's original paper that proposed the AdS/CFT conjecture as a novel approach to computing large N gauge theories [6].

3.1.1 Worked example: N coincident $D3$-branes $\leftrightarrow$ black extremal $p = 3$-brane

First, consider the simple example of a charged extremal black hole in 4 dimensions with a metric and the 2-form field strength $F_{\mu\nu}$ ($p = 1$). An extremal ($M = |Q|$) Reissner-Nordstrom (charged but non-rotating, $J = 0$) BH has the near horizon geometry of $AdS_2 \times S^2$ for $r \to M$ (for event horizon at $r_\pm = M$). Rewriting the RN metric in isotropic coordinates, $r = \rho + M$, with $\frac{M^2-Q^2}{4\rho} = 0$ as $M^2 - Q^2 = 0$,

$$ds^2 = \frac{-\rho^2}{(\rho+M)^2}dt^2 + \frac{(\rho+M)^2}{\rho^2}\big(d\rho^2 + \rho^2 d\Omega^2\big). \tag{3.1}$$

Now, defining $r \equiv M(1+\lambda) \Leftrightarrow \rho \equiv 2\lambda$, near the horizon $r \to M$, $\lambda \to 0$, the metric (3.1) becomes,

$$\begin{aligned} ds^2 &= \frac{-\lambda^2}{(1+\lambda)^2}dt^2 + \frac{M^2(1+\lambda)^2}{\lambda^2}d\lambda^2 + M^2(1+\lambda)^2 d\Omega^2) \\ &\to \underbrace{(-\lambda^2 dt^2 + M^2\frac{d\lambda^2}{\lambda^2})}_{AdS_2} + \underbrace{M^2 d\Omega^2}_{S^2}, \end{aligned} \tag{3.2}$$

where it can be seen that (3.2) describes an $AdS_2 \times S^2$ geometry in the near-horizon limit. AdS_2 is two-dimensional Anti de Sitter (AdS) space, embedding a hyperboloid of radius R into $\mathbb{R}^{2,1}$.[1] This can be shown diagrammatically by "zooming-in" on the near-horizon region of the extremal RN Penrose diagram to produce the Penrose diagram of AdS_2 (see Fig. 3.3).

Asymptotically, for $\rho \to \infty$, $r \to \infty$, (3.1) tends to asymptotically flat Minkowski space. In the asymptotically flat region, the field strength tensor $F = dA$ becomes,

$$\begin{aligned} F = dA &= \frac{Q}{r^2}dr \wedge dt \\ &= \frac{Q}{M^2(1+\lambda)^2}d\lambda \wedge dt \\ &= \frac{1}{M(1+\lambda)^2}d\lambda \wedge dt \\ &\to d\lambda \wedge dt, \end{aligned} \tag{3.3}$$

where we have used $M = |Q|$ in the third line, and have taken the asymptotic limit $r \to \infty, \lambda \to \infty$ in the final line. Hence, in the asymptotic region for a 4d extremal Reissner-Nordstrom black hole, the EM field strength is constant.

Extending the above case to 10-dimensions with a Type-II SUSY string theory, we can similarly solve the equation of motion (EoM) in the near-horizon limit to determine the near-horizon geometry. For the $d = 10$ case with general p, there will be both electric and magnetic fields w.r.t. the R-R charges $A_{(p+1)}$, due to electrically charged Dp-branes and magnetically charged $D(6-p)$-branes, electrically charged under the dual potential,

$$\begin{aligned} dA_{7-p} &= \star dA_{p+1} \\ \Rightarrow F_{8-p} &= \star F_{p+2}, \end{aligned} \tag{3.4}$$

with the corresponding $(p+2)$-form field strength $F_{p+2} = d(A_{p+1})$. To find a black hole solution, start with the (low-energy) effective action in the string frame,

$$S = \frac{1}{(2\pi)^7 l_p^8}\int d^{10}x\sqrt{-g}\big(e^{-2\phi}(R + 4(\nabla\phi)^2) - \frac{2}{(8-p)!}F_{p+2}^2\big), \tag{3.5}$$

[1] More precisely, AdS_2 is the pull-back of $\mathbb{R}^{2,1}$ metric onto the two-dimensional surface of a hyperboloid, $t^2 + \omega^2 - x^2 = L^2$.

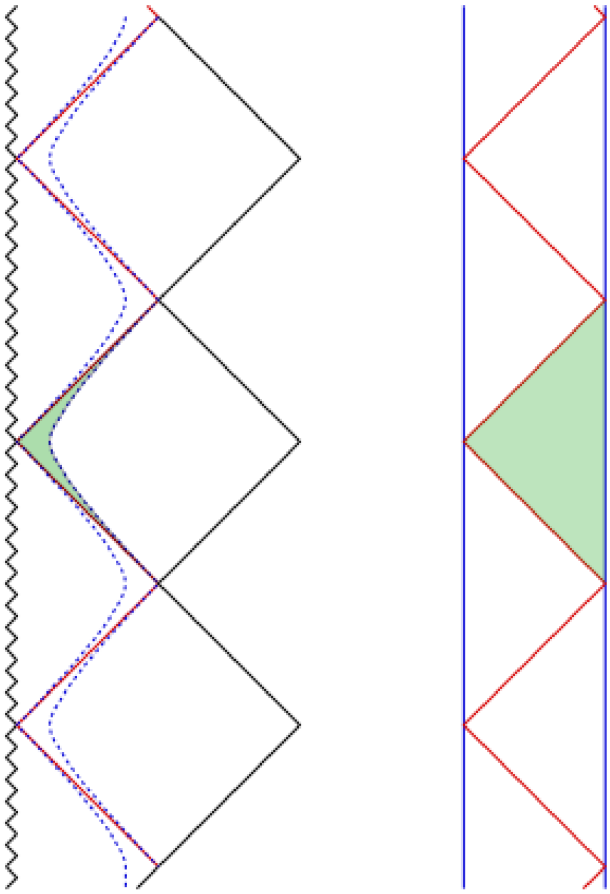


Figure 3.3: Penrose diagrams for an extremal Reissner-Norstrom black hole (left) and AdS_2 space (right) [32]. The AdS_2 is a "zoomed-in" version of the full RN solution, corresponding to the near-horizon region denoted by the dashed blue line. The shaded green region in the RN diagram corresponds to the area covered by the Poincare chart in the AdS_2 space, with the RN BH horizon at $r = r_+$ corresponding to the null surface denoting the horizon of the Poincare chart. The boundaries of the near-horizon region form the 2 conformal boundaries for the AdS_2 space. Note that angular coordinates have been suppressed in both diagrams, meaning that each point represents a 2-sphere.

where l_s is the string length, F_{p+2} is the $(p+2)$-dimensional field strength, $F_{p+2} = dA_{p+1}$. Now, consider the $p = 3$ case with the self-duality constraint (from (3.4)),

$$F_5 = \star F_5. \tag{3.6}$$

The Dirac quantisation condition is applied to the R-R charges by considering a solution with Euclidean symmetry $ISO(p)$ in p dimensions,

$$ds^2 = ds^2_{10-p} + e^{\alpha} \sum_{i=1}^{p} dx^i dx^i, \tag{3.7}$$

so that the electric source for A_{p+1} is quantised to have charge N. Assuming that the ds^2_{10-p} metric is spherically symmetric in $(10-p)$ dimensions and, placing the R-R source at the origin, we have that,

$$\int_{S^{8-p}} \star F_{p+2} = N. \tag{3.8}$$

For $p = 3$ and with the self-duality constraint (3.4), (3.8) gives the usual form of the Dirac quantisation condition,

$$\int_{S^5} F_5 = N. \tag{3.9}$$

Hence, N appears on the gravity side of the duality as the flux of the 5-form R-R field strength on the 5-sphere.

Using the Euclidean symmetry $ISO(p)$, the search for a black hole solution can be reduced to finding a spherically symmetric charged black hole solution in $(10-p)$-dimensions [33–35]. In the string frame, the resulting metric for $p = 3$ is,

$$ds^2 = -\frac{f_+(\rho)}{\sqrt{f_-(\rho)}}dt^2 + \sqrt{f_-(\rho)}\sum_{i=1}^{p} dx^i dx^i + \frac{f_-(\rho)^{-\frac{1}{2}-\frac{5-p}{7-p}}}{f_+(\rho)}d\rho^2 + r^2 f_-(\rho)^{\frac{1}{2}-\frac{5-p}{7-p}}, d\Omega^2_{8-p} \tag{3.10}$$

where

$$f(\rho) = 1 - \left(\frac{r}{\rho}\right)^{7-p}. \tag{3.11}$$

The mass and R-R charge N are related to the radii $r_\pm$ (for $p = 3$),

$$M \propto \left(5r_+^4 - r_-^4\right) \quad , \quad N \propto (r_+ r_-)^2. \tag{3.12}$$

In the Einstein frame metric, there exists an event horizon at $r = r_+$, and a curvature singularity at $r = r_-$ for $p \leq 6$. Hence, where $r_+ > r_-$, the solution (3.10) describes a black hole with curvature singularity at r_-. For $r_+ < r_-$, there is a timelike *naked* singularity which can be neglected via. Penrose's singularity theorem. For $r_+ = r_- \equiv R$, we have an extremal p-brane solution.

As the mass (per unit volume) M is a function of r_+ and r_- (3.12), the cloaked singularity condition places a bound on the radii and M,

$$r_+ \geq r_- \quad \Rightarrow \quad M \geq \frac{N}{(2\pi)^3 g_s l_s^{p+1}}, \tag{3.13}$$

where g_s is the string coupling constant. The lower bounds of r_+ and M correspond to an extremal p-brane, and the strictly greater than to a non-extremal black p-brane (with event horizon at $r_+ > r_-$).

Taking $r^4 \to r^4 - R^4$, at the extremal limit $R = r_+ = r_-$, the solution (3.10) gives,

$$ds^2 = f^{-\frac{1}{2}}(-dt^2 + dx_1^2 + dx_2^2 + dx_3^2) + f^{\frac{1}{2}}(dr^2 + r^2 d\Omega_5^2), \tag{3.14}$$

with $f = 1 + \frac{R^4}{r^4}$ and $R^4 \equiv 4_s \alpha'^2 N$. Note that as g_{tt} is non-constant, an observer approaching the horizon at $r \to 0$ with energy E_p is increasingly redshifted w.r.t. to the energy as seen by an observer at infinity, E, by a factor,

$$E = f^{-\frac{1}{4}} E_p. \tag{3.15}$$

Hence, the near-horizon geometry is a low-energy region as seen by an observer at infinity. For the horizon at $r = 0$, in the near-horizon region where $r << R$, $f \approx \frac{R^4}{r^4}$ such that the geometry (3.10) becomes,

$$ds^2 = \frac{r^2}{R^2}(-dt^2 + dx_1^2 + dx_2^2 + dx_3^2) + R^2 \frac{dr^2}{r^2} + R^2 d\Omega_5^2, \tag{3.16}$$

which is the geometry of $AdS_5 \times S^5$.[2] This is valid for an extremal p-brane in the classical supergravity limit, where the curvature of the AdS_5 and S^5 geometry, R, is large compared to the string scale, l_s,

$$\frac{R^4}{l_s^4} < \frac{R^4}{l_p^4} << 1. \tag{3.17}$$

Using the expression for the radius (3.12), the region of validity for the near-horizon geometry is,

$$\frac{R^4}{l_s^4} \propto g_s N \propto g_{YM}^2 N >> 1. \tag{3.18}$$

Next, we can consider the D-brane description of the extremal p-brane as N coincident $D3$-branes on which open string perturbations end. In the low energy regime (for energies smaller than the scale $1/l_s$), the effective theory is $\mathcal{N} = 4 \quad U(N)$ SYM on the (3+1)-dimensional D-branes [36, 37]. This description of the low energy $D3$-branes is valid when the effective loop expansion parameter $g_s N$ is small such that the perturbative expansion can be trusted,

$$g_{YM}^2 N \propto g_s N \propto \frac{R^4}{l_s^4} << 1. \tag{3.19}$$

Hence, from the worldsheet duality, we have been able to find two different theories describing N parallel $D3$-branes at the low-energy limit which can be naturally identified. This leads to the conjecture that: $\mathcal{N} = 4 \quad U(N)$ SYM theory in $(3 + 1)$ dimensions is dual to Type IIB string theory in $AdS_5 \times S^5$ [6]. The first evidence of this duality came from the calculation of low energy graviton absorption cross sections [38–40]. This example also highlights the utility of the conjecture. As the two sides of the duality are valid in completely distinct regimes yet give the same results, calculations involving CFTs at strong coupling (which are difficult to evaluate) can instead be re-expressed in terms of a classical SUGRA calculation at weak coupling (which is easier to calculate), and vice versa.

This is the most famous example of the more general conjecture that any (super-)string theory on a $AdS \times n$-sphere near horizon geometry is dual to a super-CFT worldvolume theory on the branes, which will be explored further in the next section.

[2] This will become apparent in section 3.2.1, where the geometry of AdS space is reviewed.

3.2 The AdS/CFT duality

Following the discussion taken in the original proposal of the AdS/CFT duality by Maldacena [6] in the previous section, the conjecture can now be stated more formally: certain CFTs in d-dimensional space are *dual* (or physically equivalent) to certain theories of (quantum) gravity living in asymptotic AdS space in at least one higher dimension. The most established example of the duality is for N coincident $D3$-branes, for Type-IIB supergravity on $AdS_5 \times S^5$ is dual to $\mathcal{N} = 4 \;\; d = (3,1) \;\; U(N)$ SYM theory. More generally, it has been proposed that the AdS/CFT correspondence is a specific example of a more general gauge-gravity duality: a QFT in d-dimensions is dual to a theory of QG in a $(d+1)$-dimensional space (with d-dim. asymptotic boundary) [41]. This is in agreement with proposals pre-dating Maldacena's paper [4, 5] that suggested that theories of QG obey the *holographic principle*: physics in a region can be described by a theory at the boundary. The holographic principle follows from the Berkenstein bound (1.5) that places a maximum on the amount of entropy within a region. As discussed later in section 3.5, the AdS/CFT correspondence offers a concrete example of the holographic principle.

The power of the AdS/CFT correspondence comes from its ability to re-package a theory's degrees of freedom in order to more easily solve it. The duality has lead to rigorous definitions of theories of quantum gravity (QG) through solving the gauge field theory side, and similarly, strongly-coupled CFTs can be described and solved in terms of their corresponding semi-classical supergravity theories.

The AdS/CFT duality remains a conjecture because examples have so far only been found in the limit where quantum gravity (string theory) reduces to classical supergravity i.e. where the gravity side has $g_s N \propto \lambda >> 1$ and $N \to \infty$ (see Fig. 3.4). A stronger test would be to prove the correspondence for finite $g_s N$, $N \to \infty$, in the regime of perturbative string theory. A complete proof would involve checking that both sides of the duality are equal *in general*, solving both strongly coupled quantum gauge field theories and full M-theory, when N is finite instead of large, and $g_s N$ is finite (full M-theory).

Before exploring the full properties and consequences of the AdS/CFT duality, it is first necessary to review the main features of AdS spaces and CFTs.

3.2.1 Anti-de Sitter spaces

Anti-de Sitter (AdS) spaces are the negative cosmological constant, maximally symmetric solutions of the (vacuum) Einstein equations. Geometrically, a $(d+1)$-dim. AdS space embeds as a hyperboloid, radius R, in $(2,d)$-dimensional Minkowski space. In other words, a $(d+1)$-dimensional AdS space is the intrinsic geometry of the pull-back of the 2-time Minkowski metric (the embedding space),

$$ds^2 = \sum_{i=1}^{d} dX_i^2 - du^2 - dv^2, \tag{3.20}$$

onto the surface of the $(d+1)$-dim. hyperboloid in Minkowski space, radius R,

$$\sum_{i=1}^{d} X_i^2 - u^2 - v^2 = -R^2. \tag{3.21}$$

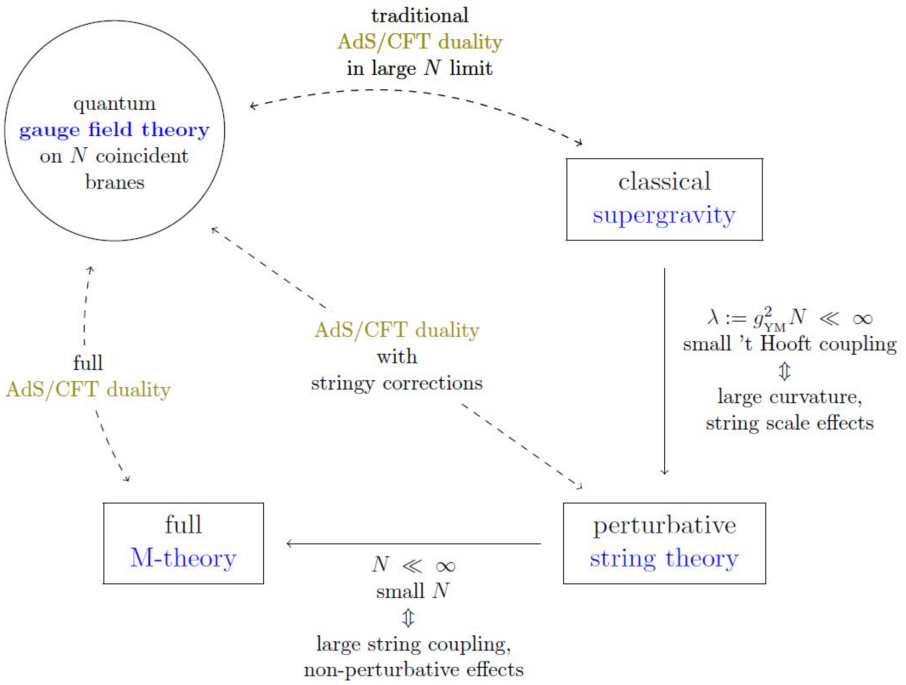


Figure 3.4: The different coupling regimes of field theory and string/M- theory, and different iterations of the AdS/CFT duality linking them [5].

As the geodesics of constant X^i give closed timelike curves (CTCs) generated by the Killing vectors of u and v time coordinates, for a dynamical spacetime, want to "unwrap" these to recover the *universal cover* of AdS space. First, we take a change of variables on the surface of the hyperboloid to *global* coordinates, $(u, v, X^i) \to (t, \rho, \Omega_i)$,

$$\begin{aligned} u &= R\, cosh\rho\, sin\tau \\ v &= R\, cosh\rho\, cos\tau \\ X_i &= R\, sinh\rho\, \Omega_i, \end{aligned} \tag{3.22}$$

where Ω_i (for $i = 1, ..., d-1$) parameterises the unit sphere in $(d-2)$ dimensions with $\sum_i \Omega_i^2 = 1$. The CTCs correspond to rotations of $\tau \in [0, 2\pi]$. Then, the unwrapping of AdS to recover the universal cover is done by taking $\tau \in \mathbb{R}$.

Then, the (universal cover of the) AdS_{d+1} metric (3.20) in global coordinates (3.22) is,

$$ds^2 = R^2\big(-cosh^2\rho\, d\tau^2 + d\rho^2 + sinh^2\rho\, d\Omega^2_{d-1}\big). \tag{3.23}$$

This is a global chart that covers the entire AdS_{d+1} space, and hence is called the "global AdS" metric. By performing another change of variables,

$$r = Rsinh(\rho) \quad , \quad t = R\tau, \tag{3.24}$$

the global AdS metric can be rewritten in a simpler and more manipulatable form for later discussion of black hole entropy in AdS (in analogy to (3.14)),

$$ds^2 = -f dt^2 + \frac{dr^2}{f} + r^2 d\Omega_{(d-1)}, \tag{3.25}$$

where $f = \frac{r^2}{R^2} + 1$. Some features of AdS space can now be discussed. The isometries of AdS_{d+1} are associated to the Killing vectors that generate the symmetries in the theory. For a generalised vector containing the u, v and X^i cords of the metric (3.20) onto the hyperboloid (3.21), the invariance of this line element is satisfied by transformations belonging to the $SO(2, d)$ symmetry group. Hence, the isometries of a $(d+1)$-dimensional AdS space live in the $SO(2, d)$ symmetry group.

From (3.25), can sketch the asymptotic behaviour of AdS space at small and large r. For $r << 1$, AdS space tends to flat Minkowski space (in spherical cords). For $r >> 1$, AdS deviates from flat space as the g_{00} component and spherical metric tend to ∞ as $r \to \infty$. For a massive particle, the behaviour of $g_{00} = -f$ acts as a harmonic potential, $V \sim \sqrt{-g_{00}} \approx R + \frac{R}{2} r$, preventing a massive particle from going out to $r \to \infty$.

By conformally compactifying the AdS space, the asymptotic behaviour can be analysed and illustrated in a "Penrose"-like diagram representation. In general, a metric g defined on a manifold $\mathcal{M}$ with asymptotically divergent behaviour can be written as,

$$ds^2 = \frac{1}{\tilde{\Omega}^2} \tilde{g}_{\mu\nu} dx^\mu dx^\nu, \tag{3.26}$$

where $\tilde{g}_{\mu\nu}$ is a regular metric within the asymptotic region, and $\tilde{\Omega}$ is a defining function which goes to 0 "linearly" (such that $d\tilde{\Omega} \neq 0$). Then, taking a Weyl transformation, $g \to \tilde{g} = \Omega^2 g$, the new space $(\mathcal{M}, \tilde{g})$ is conformally compactified by including the set of points where $\tilde{\Omega} \to 0$. For (3.23), in the asymptotic limit $\rho \to \infty$,

$$ds^2 \to R^2 \big(e^{2\rho} (-d\tau^2 + d\Omega^2) + d\rho^2 \big). \tag{3.27}$$

Defining $e^\rho = 1/z$, such that in the limit $\rho \to \infty$, $z \to 0$, the conformally compactified AdS metric takes the form,

$$ds^2 = \frac{R^2}{z^2} (-d\tau^2 + dz^2 + d\Omega_{d-1}^2), \tag{3.28}$$

where points at infinity, $\rho \to \infty$, are now included on the boundary of the space at finite affine parameter, $z = 0$. Fig. 3.5 displays the geometry of the conformally compactified AdS_{d+1} space on a cylinder, with each point representing a S^{d-1} sphere and with the boundary geometry that of the Einstein static universe, $\mathbb{R} \times S^{d-1}$ (a stationary sphere). For null geodesics, it will take finite time to travel to the boundary at finite affine parameter and reflect back inwards. Assuming reflective boundary conditions that allows radiation to bounce off the boundary instead of escaping to infinity, the geometry of AdS space acts like a box. The boundary is purely time-like, with massive particles taking an infinite time to reach it due to the harmonic-like nature of the potential at large distance from the centre.

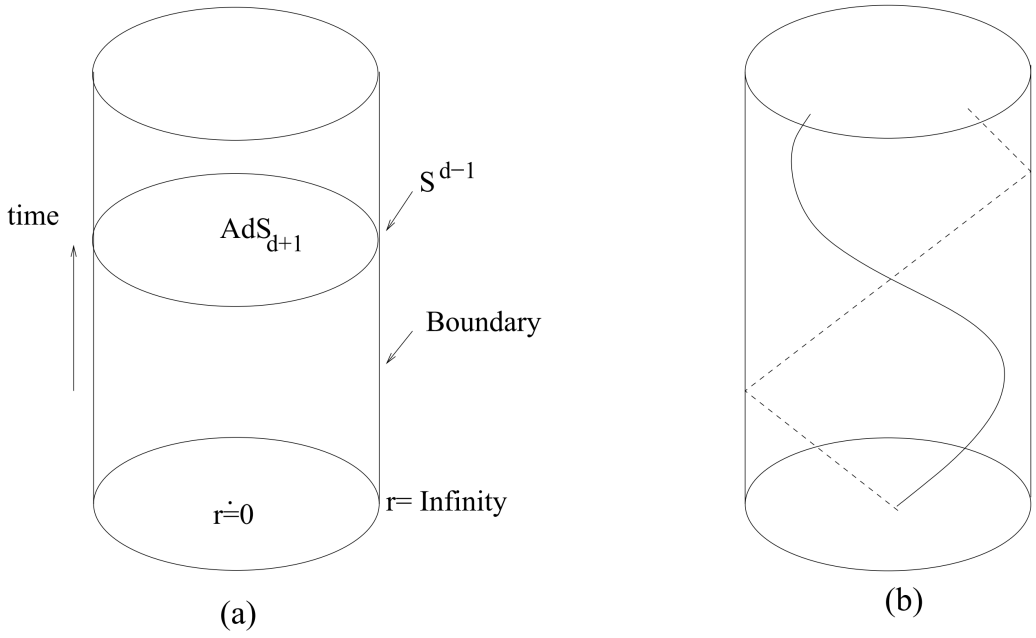


Figure 3.5: (a) Penrose diagram for conformally compactified AdS_{d+1}, with a conformal boundary at $z = 0$ with geometry S^{d-1} for a given time slice; (b) paths taken by massive (solid line) and massless (dashed) particles in the AdS_{d+1} bulk. [42]

In later discussions of conformal field theories "living" on the boundary of AdS spaces, it will be useful to consider a local patch of the global AdS space, called the *Poincare chart*. Taking a coordinate transformation of (3.20) $(u, v, X^i) \to (t, z, x^i)$,

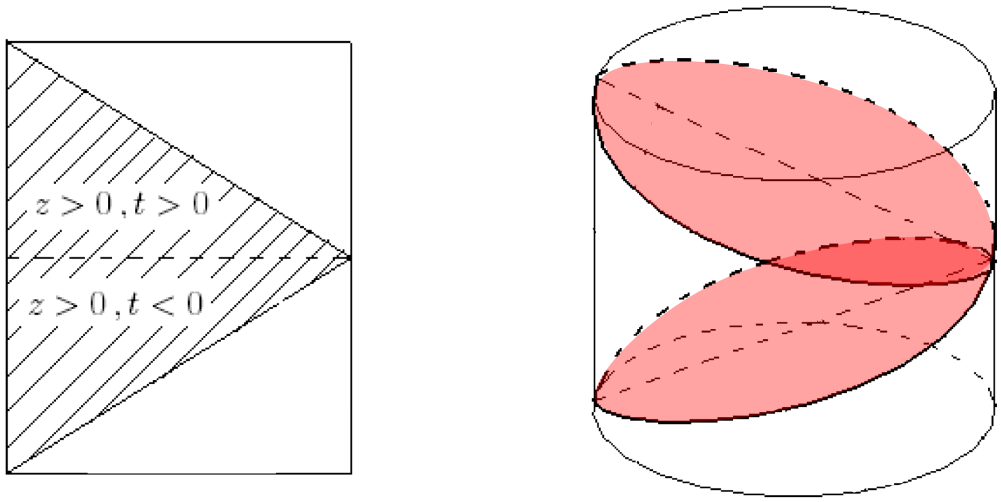
$$\begin{aligned} u &= \frac{z}{2}\big(1 + \frac{1}{z^2}(R^2 + \sum_{i=1}^{d-1} x_i - t)\big), \\ X^i &= \frac{Rx^i}{z}, \\ X_d &= \frac{z}{2}\big(1 - \frac{1}{z^2}(R^2 - \sum_{i=1}^{d-1} x_i + t)\big), \\ v &= \frac{Rt}{z} \end{aligned} \tag{3.29}$$

such that the metric becomes,

$$ds^2 = \frac{R^2}{z^2}(-dt^2 + dz^2 + \sum_{i=1}^{d-1} dx_i^2). \tag{3.30}$$

This describes a $(d+1)$-dimensional Minkowski spacetime (see Fig. 3.6), with (3.30) reducing to a boundary metric describing d-dim. Minkowski at the conformal boundary $z \to 0$,

$$ds^2 = -dt^2 + \sum_{i=1}^{d-1} dx_i^2. \tag{3.31}$$



(a) (b)

Figure 3.6: (a) the Penrose diagram for the Poincare chart, with the conformal boundary at z = 0 and Cauchy "horizons" associated null surfaces at z $\to \infty$ bounding the patch; (b) the local Poincare patch embedded in the global AdS Penrose diagram, with the red null "leaves" corresponding to the horizon at z $\to \infty$. Adapted from [43].

This chart is particularly useful for considering CFTs in d-dim. Minkowski space due to the symmetries of the space. The Poincare chart manifests the Poincare subgroup of $SO(2,d)$ explicitly,

$$\text{Poincaré:} \quad z \to z, \quad x^\mu \to \Lambda^\mu_\nu x^\nu - a^\mu. \tag{3.32}$$

The AdS boundary is also invariant under dilations,

$$\text{Dilation:} \quad z \to \lambda z, \quad x^\mu \to \lambda x^\mu. \tag{3.33}$$

These symmetry transformations, in addition to "special conformal" transformations, form the conformal group in d dimensions, which means that isometries of AdS_{d+1} on the d-dim. conformal boundary (which maps points on the boundary $\to$ boundary) are equivalent to conformal symmetries in d dimensions. Hence, we can say that a CFT on Minkowski space in d dimensions "lives on" the conformal boundary ($z = 0$) of an *asymptotically* AdS space (AdS on the boundary, $z = 0$) in $(d+1)$ dimensions. This is the central relation of the AdS/CFT correspondence, and will be motivated further in section 3.4.

3.2.2 Conformal Field Theories

Conformal field theories (CFTs) are quantum field theories (QFTs) which are invariant under transformations belonging to the conformal group, isomorphic to $SO(2,d)$. The conformal group

is a larger group than the Poincare group, which also contains scale transformations (dilations) and, for $d > 2$, special conformal transformations,

$$x^\mu \to \frac{x^\mu + a^\mu x^2}{1 + 2x \cdot a + a^2 x^2}. \tag{3.34}$$

More abstractly, the conformal group is the set of transformations that preserve angles, but not necessarily lengths, in Minkowski space [21]. Scale invariance implies that CFTs are marginal, where the couplings are dimensionless, and the theory is valid for all energy/length scales.

The generators, P_μ and $M_{\mu\nu}$, that form the Poincare algebra are of the usual form, with their actions on functions and fields given by,

$$\begin{aligned} P_\mu f &= i_\mu f, \\ M_{\mu\nu} f &= i(x_{\mu\nu} - x_{\nu\mu}) f, \\ [P_\mu, \Phi] &= i_\mu \phi, \\ [M_{\mu\nu}, \Phi] &= [i(x_{\mu\nu} - x_{\nu\mu}) + \epsilon_{\mu\nu}]\Phi. \end{aligned} \tag{3.35}$$

The scale symmetry $x^\mu \to \lambda x^\mu$ extends the Poincare algebra to include an additional scale generator $\mathcal{D}$, with an action on functions and fields given by,

$$\begin{aligned} \mathcal{D} f &= i x^\mu_\mu f, \\ [\mathcal{D}, \Phi] &= i(x^\mu \partial_\mu + \Delta)\Phi, \end{aligned} \tag{3.36}$$

where the second line implies that, infinitesimally, fields $\Phi(x)$ transform as $\Phi(x) \to \lambda^\Delta \Phi(\lambda x) = \Phi(x)$. Δ is the (classical) mass, or engineering, dimension characterising the rescaling properties of operators under dilations. When taking classical field theories to QFTs, the mass dimension is promoted to the conformal dimension, and is real and positive for unitary CFTs.

The simplest example of a CFT is for a massless scalar field ϕ in (3+1)-dimensions, with dimensionless coupling α,

$$S(\phi) = -\frac{1}{2} \int d^4x \sqrt{-g} g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi + \alpha \phi^4. \tag{3.37}$$

The action is invariant under scaling of the form,

$$\begin{aligned} x^\mu &\to x^\mu, \\ \phi(x) &\to \lambda^{-1} \phi(x). \end{aligned} \tag{3.38}$$

Although the theory is scale invariant classically, when quantising the field theory, the introduction of a regulator at a length scale to deal with UV divergences may break the scale invariance of the action. This follows from an additional anomalous dimension added to the classical engineering dimension, $\Delta \to \Delta + \gamma$. In the case above, directly quantising the action (3.37) leads to Yang Mills theory in 4 dimensions, which is not scale invariant. Hence, the

natural extension to quantum field theory for ϕ^4 theory is maximally supersymmetric $\mathcal{N}$=4 SYM theory, which posseses enough symmetry to remain invariant under scale transformations.

Furthermore, the scaling symmetry $(x, t, z) \to \lambda(x, t, z)$ in AdS space correspond to dilations on the boundary, such that the AdS boundary theory must be scale invariant (i.e. no dimensionful parameters). As all known cases of scale invariant theories are also conformally invariant, this further motivates the theory living on the boundary of asymptotic AdS to be a CFT.

Using the conformal dimension, a special class of operators belonging to the CFT can be defined: primary operators. These operators transform simply under conformal transformations, and under dilations, transform as,

$$\mathcal{O}(x) \to \lambda^{-\Delta}\mathcal{O}(x). \tag{3.39}$$

A special property of primary operators is that time-ordered correlation functions involving them take a simple form. This is allows physical quantities such as scattering amplitudes to be defined easily. From Poincare invariance alone, the form of the 2-point correlation function for 2 operators at positions x and y is,

$$\langle\Omega|T\mathcal{O}(x)\mathcal{O}(y)|\Omega\rangle = f(|x-y|). \tag{3.40}$$

Once the full conformal invariance is applied, the correlator becomes further constrained by the additional symmetry and takes the form,

$$\langle\Omega|T\mathcal{O}(x)\mathcal{O}(y)|\Omega\rangle = \Big(\frac{1}{-t^2+|x-y|^2+i\epsilon}\Big)^{\Delta}. \tag{3.41}$$

Similarly, for 2 different operators $\mathcal{O}_i$ and $\mathcal{O}_j$ with conformal dimensions Δ_i and Δ_j respectively, the 2 point correlation function is constrained as,

$$\langle\mathcal{O}_i(x_i)\mathcal{O}_j(x_j)\rangle = \frac{i^n}{Z}\frac{\delta}{\delta J(x_j)}\frac{\delta}{\delta J(x_i)}Z(J)|_{J=0}, \tag{3.42}$$

where J is the source for the operator $\mathcal{O}$, and Z is the generating function associated with the CFT.

Finally, an important general property of CFTs is the state-operator correspondence. For any CFT, regardless of there being a gravity dual, there exists a map between (eigen)states on a cylinder $\mathbb{R}\times S^{d-1}$ and operators on the flat plane $\mathbb{R}^d$. Starting with the $\mathbb{R}\times S_{d-1}$ cylindrical boundary metric for AdS_{d+1} (upon which the CFTs dual to the quantum gravity in the bulk live),

$$ds^2 = -d\tau^2 + d\Omega_{d-1}^2. \tag{3.43}$$

Taking a Wick rotation to the Euclidean plane (Euclidean analytic continuation), $\tau_E \equiv i\tau$, the boundary metric becomes,

$$ds^2 = d\tau_E^2 + d\Omega_{d-1}^2. \tag{3.44}$$

Then this Euclidean cylinder can be mapped to a flat Euclidean plane via. the change of variables $r = e^{\tau_E} \Leftrightarrow \tau_E = lnr$, with the Euclidean metric (3.44) becoming that of $\mathbb{R}^d$,

$$ds^2 \rightarrow \ d(lnr)^2 + d\Omega_{d-1}^2 = \frac{1}{r^2}[dr^2 + r^2 d\Omega_{d-1}^2] \simeq dr^2 + r^2 d\Omega_{d-1}^2, \tag{3.45}$$

where CFTs are invariant under conformal transformations such that the divergent factor in the metric can be removed via. a Weyl rescaling.

Hence, this relation maps states of the CFT living on $\mathbb{R} \times S^{d-1}$ cylinder to (primary) operators inserted at different values of r from the origin of the flat Euclidean plane $\mathbb{R}^d$, see Fig. 3.7.

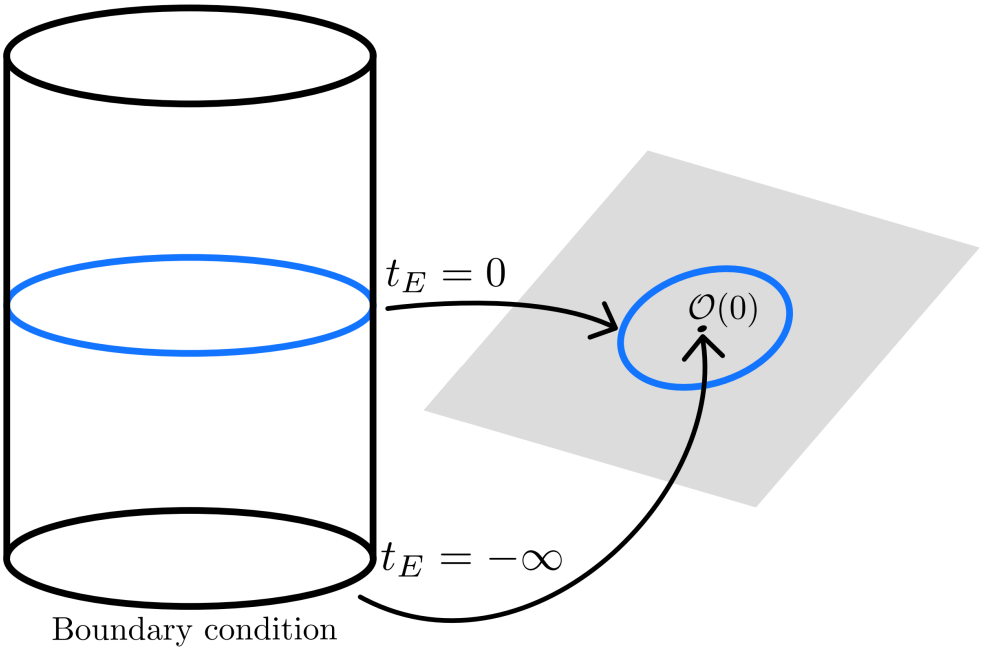


Figure 3.7: State-operator correspondence [44]. A state at time $t_E = -\infty$ on the cylinder will be mapped to an operator at $r = 0$ on the Euclidean plane, whilst a state at time t_E corresponds to an operator inserted at r on the plane.

3.3 Scalar field dynamics in AdS

Finally, before providing the fully-fledged AdS/CFT dictionary relating the two sides of the duality, the relation of certain observables in AdS space to the dual CFT's results can be illustrated through consideration of dynamics of fields in AdS [45]. The simplest example is of a massive scalar field with EoM,

$$\Delta^2 \phi = m^2 \phi. \tag{3.46}$$

Considering the Poincare chart of AdS_d, the scalar field has an action,

$$S = -\frac{1}{2}\int d^d x\sqrt{g}\big(g^{ab}\partial_a\phi\partial_b\phi + m^2\phi^2\big). \tag{3.47}$$

The goal is to compute the path integral of the scalar field theory with fixed boundary conditions at the Cauchy horizon, $z \to \infty$. The scalar field ϕ can be Fourier decomposed into a sum of modes as,

$$\phi(x,z) = \int d^d k e^{ik\cdot x} f_k(z) a_k. \tag{3.48}$$

Focussing on a single mode, we can take a Fourier space basis for a general solution as $\phi(x) = e^{ik}f(z)$, which has translation symmetry along the (conformal) boundary directions. By filling in (3.48) into the EoM (3.46), can determine the behaviour of the function $f(z)$,

$$f''(z) - \Big(\frac{d-1}{z}\Big)f'(z) - \Big(k^2 + \frac{m^2R^2}{z^2}\Big)f(z) = 0. \tag{3.49}$$

The solution of this differential equation are Bessel functions,

$$\begin{aligned} f(z) &= az^{\frac{d}{2}}h(kz) \\ &= az^{\frac{d}{2}}K_\nu(kz) + bz^{\frac{d}{2}}I_\nu(kz), \end{aligned} \tag{3.50}$$

where $\nu^2 = R^2m^2 + \frac{d^2}{4}$, $k = \sqrt{k_\mu k^\mu}$, and a, b are integration constants.

Near the conformal boundary, $z \sim 0$, expanding $f(z)$ in powers of z gives 2 independent solutions: one that goes like $f(z) \sim z^\Delta$, and the other that goes like $f(z) \sim z^{d-\Delta}$. Hence, (3.50) can be split into 2 parts,

$$f(z) = \underbrace{az^{d-\Delta}(1+...)}_{\text{leading}} + \underbrace{\tilde{b}z^\Delta(1+...)}_{\text{sub-leading}}, \tag{3.51}$$

where $\tilde{b} = \tilde{b}(a,b)$ and the ... denote higher-order contributions. Δ obeys $\Delta(\Delta - d) = R^2m^2$, with a bound placed on Δ that follows from the BF bound on the mass squared, $R^2m^2_{BF} = -\frac{d^2}{4}$. The m_{BF} bound combined with the physical requirement that $m^2 > 0$ leads to a bound on the mass dimensions, $\Delta \geq \frac{d}{2}$.

In AdS space, the Δ function is the energy of the fields, $\omega = \Delta$, whilst on the CFT side, Δ is the scaling dimension for operators. Due to the harmonic-like potential in AdS space discussed previously, the fields are localised near the centre of AdS for non-zero mass, with large mass ($mL >> 1$) corresponding to a sharply localised field ϕ around the centre of AdS at $r = 0$.

In order to completely describe the function $f(z)$, and hence the scalar fields at the boundary, the constants a and b need to fixed through boundary conditions. At the Cauchy horizon, $z \to \infty$, we require that $b = 0$ – this B.C. prevents any incoming radiation or matter from outside the Poincare wedge to enter through the null surface. Note that outgoing radiation is still allowed to *escape* via. the Cauchy horizon.

Fixing $b = 0$ at $z \to \infty$, $f(z)$ (3.51) reduces to,

$$f(z \to \infty) \to a z^{d-\Delta}(1 + ...). \tag{3.52}$$

Going back to real space, the scalar field is similarly expanded in terms of 2 independent solutions, with leading coefficient ϕ_0 and sub-leading coefficient ϕ_d in the expansion of $\phi(x, z)$ in powers of z,

$$\phi(x,z) = \underbrace{\phi_0(x) z^{d-\Delta} + \phi_2(x) z^{d-\delta+2} + ...}_{\text{leading}} + \underbrace{\phi_d(x) z^{\Delta} + ...}_{\text{sub-leading}}, \tag{3.53}$$

where $\phi_n = \partial^n \phi_0$. Approaching the conformal boundary, $z \to 0$, the leading part of $f(z)$ is the b-component expansion, which, with $b = 0$, means $f(z) \to 0$ at the boundary. Hence, the two functions ϕ_0 and ϕ_d become related,

$$\phi_d(x) \sim \phi_0(x) \sim \int d^d y \frac{1}{|x-y|^{2\Delta}} \phi_0(y), \tag{3.54}$$

where the x and y cords are the boundary coordinates that appear in the boundary metric for the Poincare chart at $z = 0$,

$$ds^2 = \frac{R^2}{z^2}\big(\eta_{\mu\nu} dx^\mu dx^\nu\big). \tag{3.55}$$

The action (3.47) can now be rewritten to include the boundary contributions of the space,

$$S = \frac{1}{2}\int_{\mathcal{M}} d^d x dz \sqrt{|g|}\phi(\nabla^2\phi - m^2\phi) - \frac{1}{2}\int_{\partial\mathcal{M}} dS^A \phi \partial_A \phi, \tag{3.56}$$

where $\mathcal{M}$ denotes the region of AdS covered by the Poincare chart, and $\partial\mathcal{M}$ denotes the conformal boundary and null surfaces (Cauchy horizons) that bound the region. As the $b = 0$ B.C. ensures no contributions to the action from the null surface "boundaries", only the conformal boundary at $z = 0$ contributes to the boundary term.

On-shell, the first term in the action (3.56) vanishes, and the action is reduced to,

$$S \to S_{on-shell} = -\frac{1}{2}\int_{\partial M} dS^A \phi \partial_A \phi. \tag{3.57}$$

In the limit $z \to 0$, the field expansion (3.53) means the action (3.57) becomes,

$$\begin{aligned}
\lim_{z\to 0} S_{on-shell} &= -\frac{1}{2}\lim_{z\to 0}\int_{\partial M} dS^A \phi\partial_A\phi, \\
&= -\frac{1}{2}\lim_{z\to 0}\int d^d x \Big(\frac{R}{z}\Big)^{d-1} \phi \partial_z \phi, \\
&= -\frac{1}{2}\lim_{z\to 0}\int d^d x\, R^{d-1}\big(\underbrace{[(d-\Delta)\phi_0^2 z^{d-2\Delta} + ...]}_{\text{divergent}} + \underbrace{[d\,\phi_0(x)\phi_d(x) z^0 + ...]}_{\text{finite}}\big).
\end{aligned} \tag{3.58}$$

From this expression, it can be seen that ϕ_0 is like fixing the value of the field at the boundary $z = 0$, and ϕ_d is like the normal derivative of ϕ_0. Hence, the second boundary condition for the region $\mathcal{M}$ can be imposed on the conformal boundary $z = 0$ by setting the field at the boundary to tend to the leading component of the field expansion,

$$\phi(x, z)|_{z=\epsilon} = \phi_0(x)\epsilon^{(d-\Delta)}, \tag{3.59}$$

with $\epsilon \to 0$ as $z \to 0$.

As $\Delta \geq d/2$, all the leading terms in $S_{on-shell}$ (3.58) up to the z^0 term have divergence of order $\mathcal{O}(1/z)$ as $z \to 0$. This is resolved by a process called *holographic renormalisation*, in analogy to ordinary renormalisation in QFTs. In order to cancel the divergences, a regulator cutoff at $z = \epsilon$ is introduced that excludes the conformal boundary from the integration over the Poincare chart region $\mathcal{M}$. Then, additional boundary terms are introduced to the action, that are functionals of the scalar field and induced metric on the regulated surface $\mathcal{M}_\epsilon$, which act as "counter-terms" to cancel the finite number of divergent terms, contained within $S_{c.t.}$. Finally, the regulated action and counter-terms are combined and the limit $\epsilon \to 0$ is taken to produce the renormalised on-shell action,

$$S^{renorm}_{on-shell} = \lim_{\epsilon \to 0} \big[S_{on-shell} + S_{c.t.} \big]|_{\phi(x,\epsilon)=\phi_0(x)\epsilon^{d-\Delta}}. \tag{3.60}$$

The divergences in $S_{on-shell}$ are absorbed by the counter-term action $S_{c.t.}$, and taking into account the boundary conditions $\phi_0 \sim \phi_d$ (3.54) and (3.59), the finite, renormalised action takes the form,

$$\begin{aligned} S^{renorm}_{on-shell} &\sim \int d^d x \phi_0(x)\phi_d(x) \\ &\sim \int d^d x \int d^d y \frac{\phi_0(x)\phi_0(y)}{|x-y|^{2\Delta}}, \end{aligned} \tag{3.61}$$

up to a constant coefficient. Hence, specifying the B.C. $J(x)$ at the conformal boundary, the scalar action can be completely determined.

3.4 AdS/CFT dictionary

We are now able to state the AdS/CFT correspondence precisely as:

> A d-dim. CFT on $\mathbb{R} \times S^{d-1}$ is dual to a theory of quantum gravity living in asymptotically $AdS_{d+1} \times \mathcal{M}$ spacetime, where $\mathcal{M}$ is a non-trivial, compact manifold.

An alternative statement is that a CFT in d dimensions lives on the (conformal) boundary of an asymptotically AdS space containing (quantum) gravity. As CFTs are comparatively well-understood, the duality allows theories of quantum gravity to be formulated.

An asymptotically AdS space is one where the spacetime tends to AdS at the boundary $r \to \infty$, $z \to 0$. Hence, the interior of the asymptotic AdS space contains non-trivial pertubations of the spacetime due to matter or radiation in general. The gravity theory contained within the asymptotically AdS space is called the "bulk", and the CFT lives on the "boundary" of the asymptotic AdS space.

This conjecture, although thought to hold in general for quantum gravity, has only been shown to exist for specific CFTs in the semi-classical gravity limit (see Fig. 3.4) [6, 41]. For the conjecture to be proven, it would need to be shown for a full quantum theory of gravity, which currently remains beyond our grasp.

The key consequence of the AdS/CFT correspondence is that every observable quantity in the bulk theory will have a dual in the boundary theory, and vice versa. This allows a powerful "AdS/CFT dictionary" to be built relating the physical quantities on both sides.

Following section (3.3), the AdS/CFT correspondence states that fields in the bulk gravity theory with energy $\omega = \Delta$ are dual to a primary operator $\mathcal{O}$ on $\mathbb{R}^d$ in the boundary theory, with scaling dimension Δ. This relation uses the state-operator map (Fig. 3.7) to take states of the CFT on the cylindrical metric (3.43) to operators inserted on the Euclidean plane $\mathbb{R}^d$. From section (3.3), the scalar field ϕ with mass squared m^2 and energy $\omega = \Delta$ is dual to scalar operator Φ with scaling dimension Δ. Furthermore, the boundary condition on the asymptotic AdS space at $z \to 0$, $J(x) \equiv \phi_0(x)$ (with mass dimension $(d - \Delta)$) is dual to a source $J(x)$ for the operator $\mathcal{O}$ in the CFT, with dimension $(d - \Delta)$.

By taking the generating functionals for the bulk and boundary theories, it is seen that they are equal,

$$Z_{grav}(J(x)) = Z_{CFT}(J(x)) = \int \mathcal{D}[grav] e^{iS_{grav}} \quad + \text{B.C.s}, \tag{3.62}$$

where the integral is performed over all gravity fields. $J(x) = \phi_0$ is the boundary condition for the scalar field on the gravity theory side, but $J(x)$ is the source for the operator $\mathcal{O}$ on the CFT side.

This allows physical quantities such as correlation functions to be directly compared. Explicitly, the 2-pt. correlation function for a scalar field with source J on the bulk can be computed and compared to the result obtained from the dual CFT (3.41).

Taking the semi-classical limit of the gravity theory, the generating function can be approximated using a classical, saddle point expansion,

$$\begin{aligned} Z_{grav}(J(x)) = Z_{CFT}(J(x)) &= \langle e^{\int d^d x \phi_0(x) \mathcal{O}(x)} \rangle \\ &\simeq e^{iS^{renorm}_{on-shell}(\phi_0 = J)}. \end{aligned} \tag{3.63}$$

Using (3.42), we can calculate the 2-pt. function from the general expression,

$$\begin{aligned} \langle \Phi(x)\Phi(y) \rangle &= -\frac{1}{Z_{CFT}} \frac{\partial^2}{\partial J(x) \partial J(y)} Z_{CFT}[J]|_{J=0} \\ &= -\frac{1}{Z_{CFT}} \frac{\partial^2}{\partial \phi_0(x) \partial \phi_0(y)} Z_{grav}[\phi_0]|_{\phi_0=0} \end{aligned} \tag{3.64}$$

Filling in for $S^{renorm}_{on-shell}$ (3.61) into (3.64), we obtain,

$$\langle \Phi(x)\Phi(y)\rangle = \frac{c}{|x-y|^{2\Delta}} \tag{3.65}$$

which is precisely the CFT result (3.41) for the dual operators $\mathcal{O}$ and source J. Hence, to calculate all correlation functions in a strongly-coupled CFT in general, all that needs to be known is the dual gravitational generating functional in the semiclassical limit.

Another example of the duality is how the isometries of the $AdS \times S^n$ space are isomorphic to the symmetries of the corresponding CFT on the boundary. Returning to the example of maximally supersymmetric $\mathcal{N} = 4$ SUSY $U(N)$ gauge theory in 4 dimensions dual to a Type IIB string theory in $AdS^5 \times S^5$ (section 3.1.1), the duality can be shown heuristically through a symmetry analysis [6, 28, 41]. The gauge theory contains a global $SU(4)$ R-symmetry that maps the 4 fermions and 6 scalar fields in the Lagrangian onto each other. Also, SUSY $SU(N)$ in $d = 4, \mathcal{N} = 4$, is conformally invariant such that its coupling doesn't run, $g \neq g(E)$ [3]. Hence, it contains a 4-dim. conformal symmetry with group $SO(2, 4)$. The isometries of $SO(2, 4)$ live naturally in 5-dim. AdS space, AdS_5. As SUSY string theory naturally lives in 10 dimensions to ensure conformal invariance, there must be 5 additional dimensions. The global R-symmetry $SU(4) \simeq SO(6)$ implies that the remaining dimensions can be added by a 5-sphere, S^5. Hence, it is motivated that $\mathcal{N} = 4$ SYM theory is dual to a ten-dim. string theory on $\mathrm{AdS}_5 \times S^5$. This holds in general, with a general $\mathrm{AdS}_{d+1}/\mathrm{CFT}_d$ sharing the same $SO(2, d)$ isometry on both sides.

Lastly, and most importantly for addressing the black hole information paradox, the AdS/CFT duality states that the conformal boundary encodes the same amount of degrees of freedom as the bulk theory in AdS space. As the number of degrees of freedom are encoded by the entropy of a system, the entropy on both sides of the duality are equal. The relation of holography and entropy will be outlined in the next section, with a generalised formula for the entropy of black holes introduced in section 2.

3.5 Holography and entropy

As introduced in the section above, the AdS/CFT correspondence is a specialised case of the more general gauge-gravity duality [41] which states that; any QFT in d-dim is dual to a theory of QG in $(d+1)$-dimensions with an asymptotic boundary in d-dim. (plus boundary conditions).

Prior to the development of AdS/CFT, it had been suggested by 't Hooft [4] and Susskind [5] that theories of QG should exhibit holography: the physics of a theory in a region is equivalently described by a theory on the boundary, with no more than one degree of freedom per Planck area. This was motivated, and first realised, by the Berkenstein-Hawking entropy (1.4) for black hole entropy, with the degrees of freedom solely dependent on the area of the event horizon. Berkenstein first formulated the holographic principle in terms of a minimum entropy bound, stating that the minimum entropy of some region is given by the area of the region in Planck units [3]. Otherwise, a finite region of spacetime could contain an arbitrary amount of information, and hence, infinite energy.

[3] Although scale invariance of a theory doesn't imply conformal invariance in general, it holds for all known cases.

At first glance, the foundational premise of the AdS/CFT correspondence raises questions: how can a $(d+1)$-dim. bulk theory be related to a d-dim. boundary theory? Naively, it would appear that there are more degrees of freedom in the bulk theory than the boundary theory due to the additional dimension. However, by calculating the entropy on both sides of the correspondence as a function of each theories' effective temperature following [41], we will see that the result is the same, up to a constant coefficient.

First, for the CFT, introduce an effective temperature T for the gas of particles contained on the boundary. If a theory has only massless fields with no scale dimensions, then the entropy scales as $S \sim V_{d-1}T^{d-1}$. Then, as scalar invariance nearly always implies conformal invariance, and for the effective temperature $T >> R$,

$$S_{CFT} \sim cT^{d-1}, \tag{3.66}$$

where R is the radius of curvature of S^{d-1}, and c is a constant dependent on the number of fields in the theory.

For the bulk theory, the particles in the AdS region can be modelled as a gas of massless gravitons that mediate the gravitational interactions, with the entropy scaling as $S_{grav} \sim V_d T^d$. Considering a finite region with $r \sim 1$, we get a lower bound on the entropy, $S_{grav} > T^d$, as other fields have been excluded.

Hence, for large enough T, the entropy of the CFT on the boundary, (3.66), will exceed the lower bound on the entropy of the gravitons in AdS, $S_{CFT} < S_{grav}$. Hence, we have an apparent contradiction of the AdS/CFT correspondence as the two values for the entropy are incompatible in general.

However, this contradiction is easily resolved by noting that, as the bulk theory contains gravity, the formation of BHs must be accounted for which will give a lower bound on the entropy in the AdS space. BHs in AdS take the form [46–48],

$$ds^2_{AdS_{d+1}} = R^2\Big[-\big(r^2+1-r_+^d r^{d-2}\big)d\tau^2 + \frac{dr^2}{\big(r^2+1-\frac{r_+^d}{r^{d-2}}\big)} + r^2 d\Omega^2_{d-1}\Big], \tag{3.67}$$

where the event horizon is at r_+, with $r_+^d = 2gm$, and g is a constant proportional to Newton's constant in units of the AdS radius of curvature, $g \sim \frac{G_N^{d+1}}{R^{d-1}}$. We define the radius that the gas of gravitons extends as $r_z \sim T$, and the mass of the graviton gas as $m \sim T^{d+1}$. Then rewriting the AdS black hole metric (3.67) in terms of $f = r^2 + 1 - \frac{2gm}{r^{d-2}}$, for large $r >> 1$, $f \to r^2 - \frac{2gm}{r^{d-2}}$. For the Schwarzschild radius, r_s, defined for $f = 0$, it is then seen that,

$$r_s^d \sim gm \sim gT^{d+1}. \tag{3.68}$$

Hence, $r_s > r_z$ for large enough temperatures, $T > 1/g$. Hence, the earlier entropy calculation of S_{grav} isn't valid in the limit of large temperature as the gas of gravitons would condense into a black hole, with Schwarzschild radius r_s. Hence, in the large T limit, for $r_s > r_z$, need to consider the black hole entropy as opposed to the gas of gravitons which, in analogy to the B-H entropy (1.4) takes the form,

$$S_{bulk} \sim \frac{r_s^{d-1}}{g}. \tag{3.69}$$

The Hawking temperature for large black holes $T \sim r_s$ can also be generalised, such that the entropy for the AdS bulk is given by,

$$S_{bulk} \sim \frac{T^{d-1}}{g}. \tag{3.70}$$

This is consistent with the entropy calculated for the CFT at the boundary (3.66), fulfilling the AdS/CFT correspondence. Enforcing equality of the two sides of the duality gives a relation for the constant c,

$$c \sim \frac{1}{g} \sim \frac{R^{d-1}}{G_N^{(d+1)}}. \tag{3.71}$$

Hence, the AdS/CFT correspondence connects the entropy of black holes in the AdS bulk (that dominate at high energies) with the entropy of a thermal CFT on the boundary. Several important implications follow from this connection. Firstly, this provides evidence that the AdS/CFT correspondence realises the holographic principle, with the degrees of freedom of the bulk internal space described by the entropy of the AdS region equivalent to an ordinary thermal gas in a CFT. Also, as the BH dynamics is dual to an ordinary thermal state in a unitary CFT, this shows that BHs are consistent with quantum mechanics and unitarity.

Chapter 4

Holographic Entanglement Entropy and the Page Curve

Having introduced the AdS/CFT correspondence as a realisation of the general holographic principle for theories of QG in spacetime, and demonstrated that the entropy of a black hole in AdS space matches that of the dual CFT living in one lower dimension, it is natural to attempt to apply AdS/CFT to the calculation of black hole entropy evolution introduced in Chapter 2.

Previously, the entropy of a region of space containing a BH was given by a generalised entropy made up of the entropy of the BH region, S_{B-H}, and the entropy of the quantum fields outside of the BH, $S_{outside}$,

$$S_{gen} = \frac{Area}{4G_N} + S_{outside}, \tag{4.1}$$

where S_{gen} is the coarse-grained entropy and obeys the generalised second law of thermodynamics (1.1), which states that S_{gen} is non-decreasing with time.

In the semi-classical limit of quantum gravity, where the spacetime is reduced to quantum fields living on a classical (GR) curved geometry, the generalised entropy formula (4.1) can be made more precise by taking into account the contribution of the quantum fields explicitly through the fine-grained entropy, S_{vN},

$$S_{gen} = \frac{Area}{4G_N} + S_{vN} + ... \tag{4.2}$$

where ... denotes additional counter-terms at higher order, $\mathcal{O}(\hbar)$. (4.2) is correct to leading order in G_N, and S_{gen} again obeys the second law of thermodynamics.

In order to describe the dynamics of the black hole and fully analyse the information paradox, we also need to formulate the fine-grained entropy for the black hole. However, as the fine-grained entropy requires knowledge of the interior of the black hole in order to define the density matrix, and hence must take into account quantum gravity effects present close to the singularity, it is not yet known how to calculate the fine-grained entropy directly using (2.15).

However, in 2006, Ryu and Takayanagi [10] developed a method in which the AdS/CFT correspondence is utilised to translate the calculation of the fine-grained entropy on the boundary of the black hole to a dual bulk theory in one higher dimension, in which the entropy is coined the holographic entanglement entropy (HEE).

4.1 Holographic Entanglement Entropy

Inspired by the microscopic derivation of the Berkenstein-Hawking entropy using string theory on BPS black holes, which recovers AdS/CFT in the near-horizon limit of the geometry, Ryu and Takayanagi sought to extend the correspondence of coarse-grained entropy in AdS/CFT to entanglement entropy. In light of the AdS/CFT correspondence and earlier work [49, 50], they proposed an entanglement entropy duality in 2006 [10]. This is stated as:

> The entanglement entropy of a subsystem A, S_A, in a CFT on $\mathbb{R}^{1,d}$ (or $\mathbb{R} \times S^d$) with a d-dimensional boundary $\partial A \in \mathbb{R}^d$ (or S^d) is given by an area law,
>
> $$S_A = \frac{Area(\gamma_A)}{4G_N^{(d+2)}}, \tag{4.3}$$
>
> where γ_A is a d-dim. (co-dimension 2) static, minimal area surface in AdS_{d+2} with associated d-dimensional boundary $\partial A = \partial\gamma_A$ and $\gamma_A \sim A$ (homologous); and $G_N^{(d+2)}$ is Newton's constant in a $(d+2)$-dimensional gravity theory.

This area law is called the RT formula, and S_A is commonly called the holographic entanglement entropy as it manifests the holographic principle.

The RT formula takes a form similar to the B-H entropy, with γ_A acting as a "holographic screen" for an observer in A, blocking all subsystems complementary to A in the same way the event horizon for a black hole blocks the interior for an outside observer. This similarity formed original motivation for the holographic entanglement entropy formula. As the RT formula takes into account quantum corrections, it can be viewed as a generalisation of the black hole entropy given by the Berkenstein-Hawking formula, with the B-H formula recovered in the presence of an event horizon i.e. AdS Schwarzschild black hole solutions [51].

Furthermore, the intuitive interpretation of the entanglement entropy S_A, which "smears out" region B, as the entropy for an observer in A unable to access B informed the proposal of (4.3). In AdS space, the inaccessibility of the observer to region B corresponds to a region of the bulk space AdS_{d+2} hidden by a "horizon" given by a surface γ_A. In analogy with the B-H formula, the holographic formula places an entropy bound on the region contained by the surface proportional to its area [52, 53]. Hence, by minismising the surface γ_A, S_A saturates the entropy bound [4, 5, 52, 53].

The RT formula exhibits the key properties of entanglement entropy, namely $S_A = S_B$, where B is the complementary subsystem to A, as γ_A is shared by both regions; and SSA (2.21), where for 3 regions A, B and C: (a) $S_{A\cup B} + S_{B\cup C} \geq S_{A\cup B\cup C} + S_B$; and (b) $S_A + S_C \leq S_A + S_B$. The first condition of strong subadditivity can be proven geometrically [22]: in the bulk theory, the minimal surfaces $\gamma_{A\cup B}$ and $\gamma_{B\cup C}$ bound an area equal to that contained by the (non-minimal) surfaces $\gamma'_{A\cup B\cup C}$ and γ'_B, see Fig. 4.1(a). However, by considering the alternative surfaces $\gamma_{A\cup B\cup C}$ and γ_B, the area contained by $\gamma'_{A\cup B\cup C}$ and γ'_B is minimised. As the area of minimal surfaces of a region are proportional to the region's entanglement entropy through the RT formula,

$$\begin{aligned} Area(\gamma_{A\cup B}) + Area(\gamma_{B\cup C}) = Area(\gamma'_{A\cup B\cup C}) + Area(\gamma'_B) \geq Area(\gamma_{A\cup B\cup C}) + Area(\gamma_B) \\ \Rightarrow S_{A\cup B} + S_{B\cup C} \geq S_{A\cup B\cup C} + S_B \end{aligned} \tag{4.4}$$

where the second line follows by dividing through by $G_N^{(d+2)}$. A similar process can be performed for the second relation of SSA (see Fig. 4.1(b)).

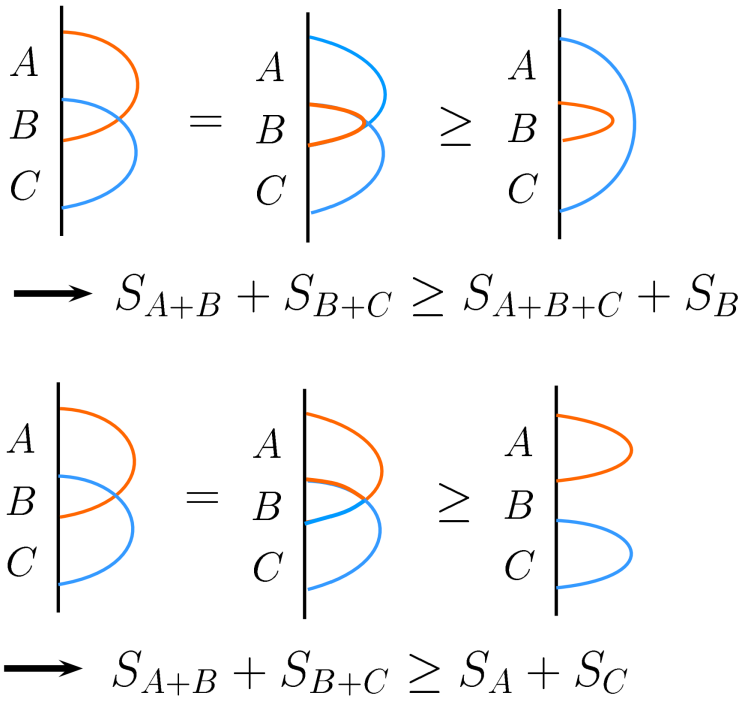


Figure 4.1: Digaram showing a holographic proof of the strong subadditivity conditions of entanglement entropy for each relation (a) (upper) and (b) (lower). The diagrams are simplified by projecting the AdS_{d+2} space onto a two-dim. plane. [51]

The RT formula can be proven heuristically through a technique called the *replica trick*. The replica trick is a method to calculate the entanglement entropy in QFTs using a new type of entropy that takes the form of a path integral over n sheets,

$$S_A^{(n)} = \frac{1}{1-n} log Tr_A(\rho_A^n). \tag{4.5}$$

This is called the Re'nyi entropy. ρ_A is the reduced density matrix for the subsystem A and n is analytically continued from $\mathbb{Z}^+ \to \mathbb{R}^+$. In the limit of $n \to 1$, the Renyi entropy reduces to the entanglement entropy,

$$\begin{aligned} S_A = \lim_{n\to 1} S_A^{(n)} &= \lim_{n\to 1} \frac{Tr_A(\rho_A)^n - 1}{1-n} \\ &= -\frac{\partial}{\partial n} Tr_A(\rho_A)^n \\ &= -\frac{\partial}{\partial n} log\big(Tr_A(\rho_A)^n\big)|_{n=1}. \end{aligned} \tag{4.6}$$

Hence, in order to calculate the entanglement entropy of a QFT, we have to evaluate $Tr(\rho_A^n)$ [51]. This can be done by using a path integral formalism to compute ρ_A, then taking a functional integral over the n-sheeted Riemannian surface to compute the Renyi entropy.

The subsystem A is defined at a constant (Euclidean) time slice $t_E = 0$ in flat Euclidean coordinates (t_E, $\mathbb{R}^{d-1}$) as a finite interval in the $(d-1)$ spatial coordinates $x \in [u, v]$, which is represented as a "cut" in the path integral representation of the reduced density matrix (see Fig. 4.2(a)). For a scalar field $\phi(x)$ of the QFT in the region A, we impose boundary conditions,

$$\phi_A(x)|_{t=0^\pm} = \phi_\pm(x), \tag{4.7}$$

for boundaries $t_E = \pm 0$ of the branch cut, such that the result of the path integral is projected onto definite field values $\phi_\pm$. The ground state wave functional $\Psi(\phi_\pm, t) = \langle \phi_\pm(x),\, \phi(t) \rangle$ is defined by the path integral from $t_E = -\infty$ to $t_E = 0$,

$$\Psi(\phi_0(x)) = \int_{t_E=-\infty}^{\phi(t_E=0,x)=\phi_0(x)} D\phi e^{-S(\phi)}. \tag{4.8}$$

This leads to the reduced density matrix of the region A (with boundary conditions $\phi_\pm$), $[\rho_A]_{\phi_\pm} = [Tr_B(|\Psi\rangle\langle\Psi|)]_{\phi_\pm} = \Psi(\phi_+)\bar{\Psi}(\phi_-)$, where the complex conjugate $\bar{\Psi}$ is obtained by path integrating from $t_E = \infty$ to $t_E = 0$. Integrating ϕ on $x \in$ B with the condition $\phi_\pm(x) = \phi_\mp(x)$,

$$[\rho_A]_{\phi_\pm} = \frac{1}{Z_1} \int_{t_e=-\infty}^{t_E=\infty} D\phi e^{-S(\phi)} \prod_{x \in A} \delta(\phi(0^+, x) - \phi_+(x)) \cdot \delta(\phi(0^-, x) - \phi_-(x)), \tag{4.9}$$

where Z_1 is the vacuum partition function on the d-dimensional Euclidean geometry, necessary to normalise ρ_A such that $Tr(\rho_A) = 1$. To calculate $Tr(\rho_A)^n$, the path integral in (4.9) is extended to n copies,

$$Tr_A(\rho_A)^n = [\rho_A]_{\phi_{1\pm}} [\rho_A]_{\phi_{2\pm}} ... [\rho_A]_{\phi_{n\pm}}. \tag{4.10}$$

This can be viewed in the path-integral formalism as gluing each of the lower boundaries of the i^{th} copy to the upper boundaries of the $(i+1)^{th}$-copy, $\phi_i(x) = \phi_{(i+1)}(x)$ for $(i = 1, 2, ..., n)$, and integrating ϕ_i over the n-sheeted Reimann surface, $\mathcal{R}_n$ (see Fig. 4.2(b)). This $\mathcal{R}_n$ space has a deficit angle $\delta = 2\pi(1-n)$ on the surface ∂A. Hence, $Tr_A(\rho_A)^n$ is given by the path integral over the n-sheeted Reimann surface,

$$Tr_A(\rho_A)^n = (Z_1)^{-n} \int_{(t_E,x) \in R_n} D\phi e^{-S(\phi)} \equiv \frac{Z_n}{(Z_1)^n}, \tag{4.11}$$

where Z_n is the partition function on the n-sheeted Reimann surface.

Calculating the entanglement entropy analytically from (4.5) is possible at low dimensions, and has been performed for 2-dimensional CFTs to recover the expected entanglement entropy [54–57]. However, analytical calculations of S_A quickly become complicated for higher dimensions, although numerical analysis is possible for certain cases [57, 58].

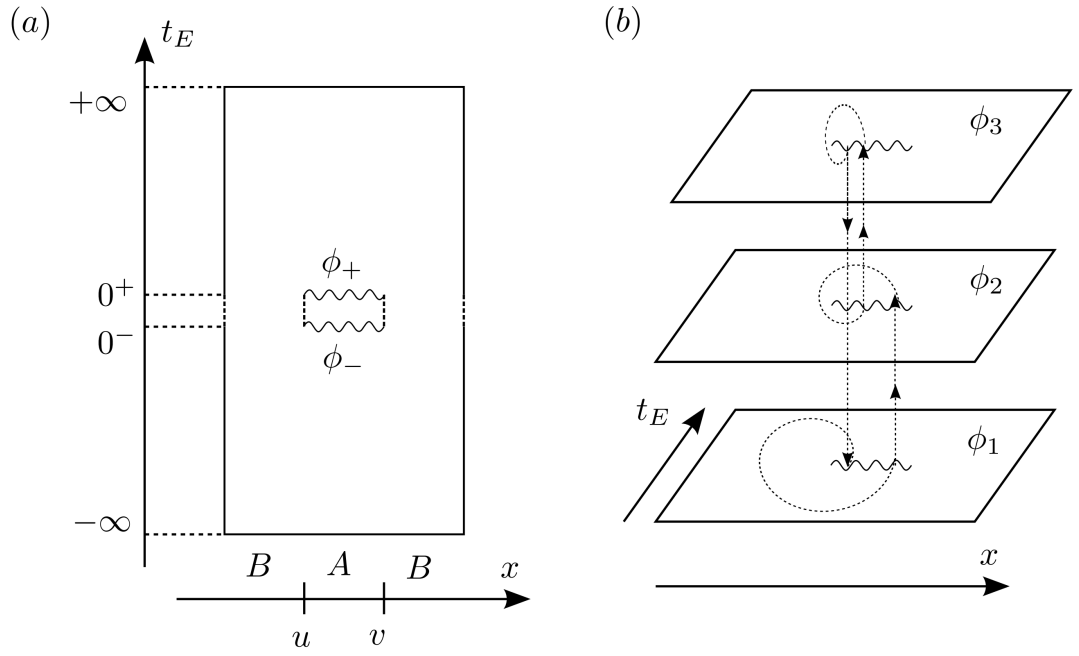


Figure 4.2: (a) the path integral representation of $[\rho_{\mathsf{A}}]_{\phi_\pm}$, with boundary conditions ϕ_+ and ϕ_- on the cut defined within region A; (b) The n-sheeted Riemann surface $\mathcal{R}_n$ over which the ϕ_{i}'s are integrated over. [51]

The calculation of $S_A^{(n)}$ is simplified by considering the AdS/CFT dictionary relation for the partition functions on either side of the duality, $Z_n = Z_{CFT} = Z_{AdS}^{(n)} = e^{-iS_{grav}}$, where the final equality follows from the semi-classical gravity limit. In order to define the partition function on the AdS_{d+1} bulk space, a $(d+1)$-dimensional back-reacted geometry $\mathcal{S}_n$ is needed, found by solving the Einstein equation with a negative cosmological constant such that the metric approaches the n-sheeted space $\mathcal{R}_n$ at the boundary $z \to 0$ [51]. In general this is technically complicated, so to simplify, a natural assumption can be taken that the back-reacted geometry $\mathcal{S}_n$ is given by an n-sheeted AdS_{d+1}, with the deficit angle δ localised on a co-dimension 2 surface γ_A [59]. This leads to the Ricci scalar in the bulk, $R = 4\pi(1-n)\delta(\gamma_A) + R^{(0)}$, where $R^{(0)}$ is the Ricci scalar of the pure AdS_{d+1} gravity theory (no matter fields) and $\delta(\gamma_A)$ is a delta function localised around γ_A, such that $\delta(\gamma_A) = \infty$ for $x \in \gamma_A$ and is zero otherwise. Filling the Ricci scalar into the supergravity action (with corrections to the bulk Einstein-Hilbert action cancelling in 4.11),

$$\begin{aligned} log Z_{AdS}^{(n)} = S_{AdS} &= -\frac{1}{16\pi G_N^{(d+2)}} \int_{\mathcal{M}} dx^{d+2} \sqrt{g}(R + \Lambda) + ... \\ &= \frac{4\pi(1-n)Area(\gamma_A)}{16\pi G_N^{d+1}} - \frac{1}{16\pi G_N^{d+1}} \int d^{d+1}x \sqrt{-g}(R^{(0)} + \Lambda) + ..., \end{aligned} \tag{4.12}$$

where $\mathcal{M}$ is the time-slice of the bulk AdS_{d+1} spacetime, and the last equality follows from the AdS/CFT duality. Then, using (4.6), and substituting the expression for $log(Z_n) = \frac{(1-n)Area(\gamma_A)}{16\pi G_N^{d+1}}$ into (4.6),

$$S_A = -\frac{\partial}{\partial n} log\big(Tr_A(\rho_A)^n\big)|_{n=1} = -\frac{\partial}{\partial n}\left[\frac{(1-n)Area(\gamma_A)}{4G_N^{d+2}}\right]\Big|_{n=1} = \frac{Area(\gamma_A)}{4G_N^{d+2}}. \tag{4.13}$$

Thus, we have reproduced the holographic RT formula for entanglement entropy, (4.3). Note that this derivation relies on the assumption of the back-reacted geometry $\mathcal{S}_n$, which has been shown for AdS_3 but not higher dimensional AdS/CFT [51].

4.2 Calculating entanglement entropy from the RT formula

The holographic RT formula can now be applied to CFTs for which their holographic dual is known. One of the simplest examples of the correspondence is AdS_3/CFT_2 [6], where entanglement entropy can be calculated following [10].

4.2.1 Entanglement entropy of CFT_2: Zero Temperature

In global coordinates, the metric for AdS_3 in global coordinates is,

$$ds^2 = R^2\big(-cosh^2(\rho)d\tau^2 + d\rho^2 + sinh^2(\rho)d\theta^2\big), \tag{4.14}$$

which is divergent at the conformal boundary $\rho = \infty$. In order to regulate physical quantities, a cut-off $\rho \leq \rho_0$ is introduced to restrict the space. The introduction of the AdS cut-off is related to the dimensionless UV cut-off of the dual CFT, $\delta^{-1} = \frac{L}{a} \sim e^{\rho_0}$, where a is the lattice spacing (short-distance/UV cutoff) and L is the total length of the system. As discussed in section 3.4, a two-dimensional CFT lives at the (regularised) boundary of the AdS_3 space at $\rho = \rho_0$, with the boundary geometry described by coordinates (τ, θ) and hence taking the form of a cylinder, $\mathbb{R} \times S^1$, see Fig. 4.3.

Using the metric (4.14) for a fixed time slice ($d\tau = 0$) with the cut-off constraint $\rho \leq \rho_0$, the length of the geodesic can be determined as,

$$\int ds = \int d\lambda = \lambda_*, \tag{4.15}$$

where L_{AdS} is defined by,

$$cosh(\lambda_*/R) = 1 + 2sinh^2\rho_0\ sin^2\left(\frac{\pi l}{L_{AdS}}\right). \tag{4.16}$$

For a large UV cutoff, $a << 1$, $e^{\rho_0} >> 1$,

$$Area(\gamma_A) = R\ log\left(e^{2\rho_0} sin^2\left(\frac{\pi l}{L_{AdS}}\right)\right) \tag{4.17}$$

Hence, the entanglement entropy given by the holographic formula (4.6) can be found in the cut-off limit as,

$$S_A = \frac{R}{4G_N^{(3)}} log\left(e^{2\rho} sin^2\left(\frac{\pi l}{L_{AdS}}\right)\right) = \frac{c}{3} log\left(e^{\rho} sin\left(\frac{\pi l}{L_{AdS}}\right)\right). \tag{4.18}$$

Taking $e^{\rho_0} \sim L/a$, this is the same form as the known 2d CFT result [55, 60], including coefficients once the central charge, $c = 3R_{AdS}/2G_N^{(3)}$ [61], is recalled,

$$S_A = \frac{c}{3}\left(\frac{L}{\pi a} sin\left(\frac{\pi l}{L}\right)\right). \tag{4.19}$$

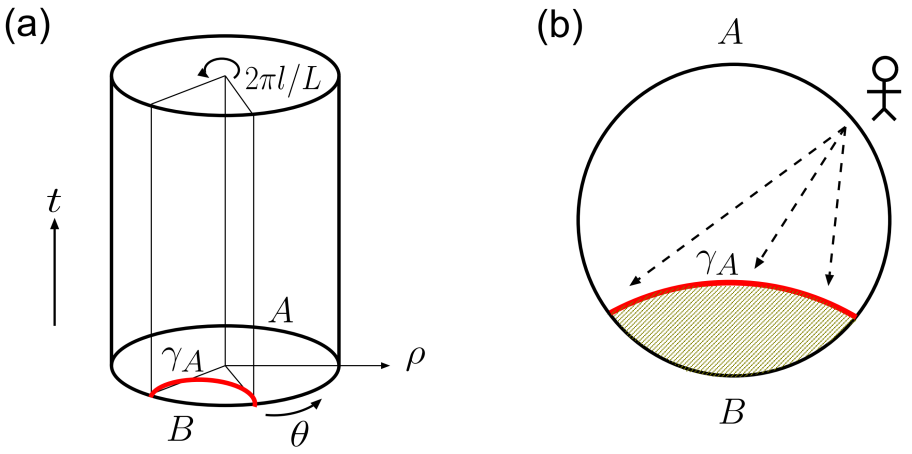


Figure 4.3: (a) Penrose diagram of AdS_3 space, with region B bounded by the conformal boundary and the minimal surface γ_A; (b) γ_A acting as a holographic screen for an observer in region A. [10]

The same analysis can also be performed using Poincare coordinates of AdS, defined by the metric,

$$ds^2 = \frac{L^2}{z^2}(-dt^2 + dx^2 + dz^2) \tag{4.20}$$

As before, a UV cut-off $z \sim \epsilon$ is imposed. In this case, the (conformal) boundary metric at $z \sim \epsilon$ describes a 2-dim. Minkowski spacetime, topologically $\mathbb{R}^{1,1}$. Taking a time-slice at the boundary, the region A can be defined by a line element of length l, $A = \{r, r \in (-l/2, +l/2)\}$. In this case, it follows from the boundary metric that the minimal geodesic line γ_A is a half circle in the rz-plane, $(r, z) = l/2(cos(\phi), sin(\phi))$ for $\epsilon \leq \phi \leq \pi - \epsilon$, where $\epsilon \sim 2a/l << 1$ is the UV cutoff. As before, the area of the one-dimensional geodesic is its length L_{γ_A},

$$L_{\gamma_A} = 2R \int_{\epsilon}^{\pi/2} \frac{d\phi}{sin\phi} = -2Rlog\left(\frac{\epsilon}{2}\right) = 2R\, log\left(\frac{l}{a}\right). \tag{4.21}$$

This leads to the entanglement entropy S_A obtained from (4.3),

$$S_A = \frac{L_{\gamma_A}}{4G_N^{(3)}} = \frac{c}{3} log\left(\frac{l}{a}\right). \tag{4.22}$$

This reproduces the expected small l limit of (4.19), corresponding to the region of AdS covered by the Poincare chart.

4.2.2 Entanglement entropy of CFT_2: Finite Temperature

The entanglement entropy can also be considered in $\mathrm{CFT}_2/\mathrm{AdS}_3$ for finite temperature, $T = 1/\beta \neq 0$. Assuming the size of the system, characterised by its spatial length L, is infinite (such that $\beta/L << 1$), at high temperatures, $T >> 1$, the gravity dual of the two-dimensional CFT is the Euclidean BTZ black hole [62], with metric given in global coordinates (τ, ρ, ϕ) as,

$$ds^2 = (r^2 - r_+^2)d\tau^2 + \frac{R^2}{r^2 - r_+^2}dr^2 + r^2 d\phi^2. \tag{4.23}$$

The Euclidean time is compactified to ensure a smooth geometry, $\tau \sim \tau + \frac{2}{r_+}$, where $r = r_+$ denotes the event horizon of the black hole. Also, periodicity of ϕ is imposed, $\phi \sim \phi + 2\pi$. Taking the boundary limit $r \to \infty$, the boundary CFT and the asymptotically AdS geometry of the metric (4.23) are found to be related through,

$$\frac{\beta}{L} = \frac{R}{r_+} << 1. \tag{4.24}$$

The subsystem A is defined as the region $0 \leq \phi \leq 2\pi l/L$ at the boundary of the bulk. Extending the RT formula (4.3) to include minimal surfaces in asymptotically AdS spaces, the geodesic length of the minimal surface γ_A connecting the points on the boundary at $\phi = 0$, $2\pi l/L$ gives the area of the minimal surface as before. The geodesic line can be found by recalling that the Euclidean BTZ black hole at temperature T_{BTZ} is equivalent to thermal AdS_3 at temperature T_{BTZ} via. a modular transformation in the CFT [63]. By taking a change of coordinates in the metric (4.23),

$$r = r_+ cosh(\rho), \quad \tau = \frac{R}{r_+}\theta, \quad \phi = \frac{R}{r_+}t, \tag{4.25}$$

the Euclidean BTZ black holes metric goes to the Euclidean Poincare coordinates metric. From this new form of the metric, the geodesic distance can be found in the same way as the 2 prior cases, with the length of the geodesic given by,

$$cosh\left(\frac{\lambda_*}{R}\right) = 1 + 2cosh^2\rho_0 \; sinh^2\left(\frac{\pi l}{\beta}\right), \tag{4.26}$$

where the UV cut-off is of the form $e^{\rho_0} \sim \beta/a$. Finally, from the holographic formula (4.3), the entanglement entropy for an observer in region A reproduces the known CFT result [54],

$$S_A = \frac{c}{3} \cdot log\left(\frac{\beta}{\pi a} sinh\left(\frac{\pi l}{\beta}\right)\right). \tag{4.27}$$

Note that as a consequence of the finite temperature of the system, the entanglement entropy of subsystem A is not always equal to that of subsystem B, $S_A \neq S_B$. This is expected from the properties of entanglement entropy, as states with finite temperature are mixed. This can be

understood geometrically from how the minimal surfaces of each of the regions evolves as the size of A, l, changes (see Fig. 4.4(a)). When A is small, the geodesic line is almost the same as for ordinary, zero temperature AdS_3. As the size of A increases, a turning point of the geodesic line approaches and covers part of the horizon. At this point, γ_A and γ_B are no longer equal, and hence $S_A \neq S_B$ in the presence of a horizon (see Fig. 4.4(b)). This is the origin behind the thermal behaviour of entropy when $l/\beta >> 1$.

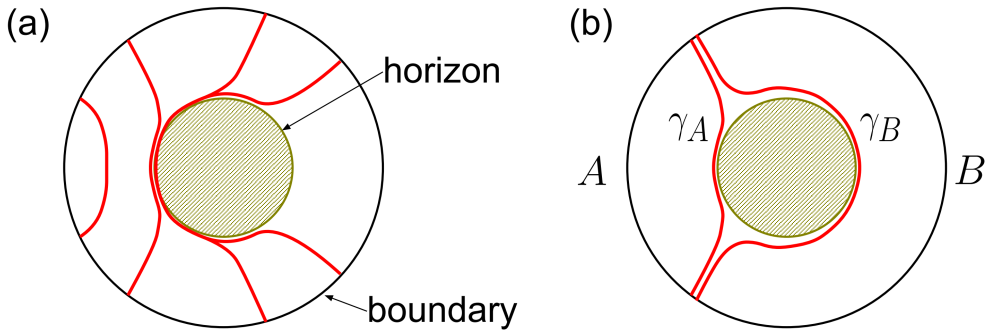


Figure 4.4: (a) minimal surfaces γ_A for different sizes of A in the BTZ BH system;(b) γ_A and γ_B wrap different parts of the BH horizon. [58]

Furthermore, the single interval covered by A can be extended to a subsystem consisting of multiple disconnected intervals,

$$A = \{x | x \in [r_1, s_1] \cup [r_2, s_2] \cup ... \cup [r_N, s_N]\}. \tag{4.28}$$

Then, each interval will have an individual minimal surface, with the total entanglement entropy calculated over of all of these contributions,

$$S_A = min\left(\frac{c}{3}\sum_{i,j} log\frac{|s_i - r_j|}{\epsilon}\right). \tag{4.29}$$

4.3 The Hubeny-Rangamani-Takayanagi formula

The holographic RT formula for entanglement entropy (4.3) explored in the previous section provides a simple, geometric way to calculate the entanglement entropy in static spacetimes with no time evolution. The symmetry under time translations of the AdS spacetime studied by RT is due to the selection of a preferred Cauchy slice whose geodesics lack time dependence. However, for a dynamical quantum system, i.e. black hole evaporation, we want to consider how the entropy evolves in time. Hence, for dynamical, covariant situations, an extension of the existing definition of the holographic dual of the entanglement entropy is needed, involving a fully covariant generalisation of the minimal surface and the RT formula.

In 2007, Hubeny, Rangamani and Takayanagi (HRT) proposed a generalisation to the RT formula [11]: the entanglement entropy for a region on the CFT boundary of an asymptotically

AdS space is determined by the area of a co-dimension two surface in the bulk, where the surface is now defined on a covariant Cauchy surface rather than a constant time slice.

The definition for the covariant surface defining the entanglement entropy in dynamic spacetimes can be motivated by first considering a naïve generalisation of the minimal surface given in the RT prescription. First, consider a time-dependent version of the AdS/CFT correspondence where the region A on the boundary CFT is in a time-evolving state on a fixed background $\partial\mathcal{M}$, such that the bulk geometry $\mathcal{M}$ has explicit time-dependence [11]. As the boundary geometry is non-dynamical, a foliation in equal time-slices, $\partial M = \partial\mathcal{N}_t \times \mathbb{R}_t$, can be chosen. Choosing a region $A_t \in \partial\mathcal{N}_t$ living on a given time-slice at time t, the entanglement entropy can be calculated using the path integral formulation in section (4.1).

In order to generalise the holographic formula (4.3) to dynamic spacetimes, we need to consider how the entropy can be calculated in the bulk theory. The key issue is how to define the generalisation of the minimal surface to covariant situations. In static spacetimes, the minimal surfaces used to calculate the entanglement entropy are typically associated with Euclidean, rather than Lorentzian, geometries. Minimal surfaces in Lorentzian spacetimes are difficult to define as the area can be made arbitrarily small by compressing a spacelike surface in the time direction, leading to indefinite metric signatures. This issue was dealt with in the above static spacetime examples by restricting the system to a constant time slice. However, in a dynamic Lorentzian setting, equal-time foliation on the boundary $\partial\mathcal{M}$ doesn't automatically lead to a canonical (naturally symmetric) foliation of the bulk $\mathcal{M}$.

Assuming that a natural foliation of $\mathcal{M}$ has been selected, the entanglement entropy can be computed by extending the time-slices $\partial\mathcal{N}_t$ from $\partial\mathcal{M}$ to a preferred spacelike slice $\mathcal{N}_t$ of $\mathcal{M}$. As the metric on $\mathcal{N}_t$ is spacelike, the "minimal surface" is well-defined. Then, applying the holographic principle, a minimal surface $S \in \mathcal{N}$ can be found that satisfies $\partial S|_{\partial\mathcal{M}} = \partial A$. Hence, we need to search for a covariantly-defined spacelike slice of the bulk $\mathcal{N}_t$ upon which to define a minimal surface in the dynamic bulk theory, which is "anchored" at $\partial\mathcal{N}_t$ and which reduces to a constant time-slice for a static bulk.

Although it is expected that there is no preferred time-slicing of $\mathcal{M}$, asymptotically AdS spacetimes permit a nature foliation by zero mean curvature slices, which are slices with vanishing trace of extrinsic curvature [11]. Physically, these are maximal area spacelike slices through $\mathcal{M}$, anchored at $\partial\mathcal{N}$, and denoted by Σ_t. These maximal-area slices are well defined, as the slices have co-dimension 1 to the bulk theory, which means no "crumpling" of the surface in spatial directions can take place as the slices extend over all spatial directions. Also, the area of the slices, which are naively divergent due to the infinitely asymptotic AdS space, are regulated in the UV limit.

Hence, the Σ_t slices allow the construction of minimal area surfaces anchored at ∂A_t. From this the entanglement entropy for the region A can be calculated by: finding a maximal slice in the bulk agreeing with the spacelike foliation of $\partial\mathcal{M}$; and then finding a minimal surface X living on the maximal slice. Finally, the HRT proposal for the entanglement entropy S_A of a system A living on the CFT_d is given by,

$$S_A = min\left(\frac{Area(X_{ext})}{4G_N^{d+1}}\right), \tag{4.30}$$

where X_{ext} is a co-dimension 2 extremal (with zero null geodesic expansion) surface in the bulk $\mathcal{M}$. X satisfies the 3 conditions for being dual to the entanglement entropy of A: (1) covariantly

well-defined; (2) anchored by ∂A, $\partial X|_{\partial \mathcal{M}} = \partial A$; and (3) reduces to the minimal surface when the spacetime is static.

4.4 A generalised gravitational entanglement entropy formula

We are now in a position to apply the RT/HRT prescription to define a generalised entanglement entropy for a gravitational system which will lead to the unitary Page curve when applied to evaporating black holes and Hawking radiation.

Using the holographic principle, the generalised entanglement entropy for a gravitational system containing quantum corrections can be refined from the section introduction (4.2) such that in a semi-classical setting, to order $\mathcal{O}(\hbar^0)$ in the bulk,

$$S_R = \frac{\langle A(X) \rangle}{4G_N\hbar} + S_{bulk}(X) + \text{counter-terms} = S_{gen}(X), \tag{4.31}$$

where X is the codimension 2 extremal surface in the bulk found through the HRT prescription above, S_R is the entanglement entropy of the region R on the boundary CFT, and S_{bulk} is the bulk entanglement entropy across the surface X. This was proposed for static spacetimes, and valid only to leading order in quantum corrections, by Faulkner, Lewkowycz and Maldacena (FLM) [64]. This has been proved to order $\mathcal{O}(\hbar^0)$ by considering quantum corrections to the Euclidean gravitational path integral [65].

The FLM result was further generalised by Engelhardt and Wall (EW) in 2015 [12] to dynamic spacetimes and higher order quantum corrections through the introduction of a quantum extremal surfaces (QES). The idea is that in the EW prescription, the area $Area(X)$ should be extremised with respect to the entire generalised entropy (4.31) (including the quantum corrections), as opposed to only the area being extremised following the RT/HRT prescription, with the quantum correction term added after, as proposed by FLM. The EW approach gives rise to QESs that are defined such that their area extremises the whole generalised entropy (4.31). Hence, the generalised entanglement entropy of the region A on the CFT boundary, correct to any order in $\mathcal{O}(\hbar)$, is expressed as,

$$S_A = min_{X_A}\left\{ext_X\left[\frac{A(X)}{4G_N\hbar} + S_{vN}(\Sigma_X)\right]\right\}, \tag{4.32}$$

where X_A is the co-dimension two quantum extremal surface in the bulk which extremises the generalised entropy, with $\partial X_A = \partial A$ and $X_A \sim A$; Σ_X is the region bounded by X_A and the cut-off surface separating the regions in the bulk (i.e. the boundary separating the black hole region and the radiation region in a BH-radiation system); and $S_{vN}(\Sigma_X)$ is the von-Neumann entropy of the quantum fields on the Σ_X surface. The extremisation of the generalised entropy $S_{gen} = Area(X_A)/4G_N + S_{vN}(\Sigma_X)$ is performed by starting with an extremal surface outside the BH and bringing it in past the event horizon to minimise S_{gen}.

4.5 Computing the entropy of a black hole

By assuming the central dogma to formulate a black hole system as a quantum system with the number of degrees of freedom proportional to the region's area, it is natural to apply the

entanglement entropy formula (4.31) to the BH-radiation system described in section 1. In this case, the entropy (4.32) is that measured by an observer in the exterior of the black hole region, who is only able to access the Hawking radiation degrees of freedom. As such, the surface Σ_X can be understood in the BH-radiation system as the region spanning the surface X and the black hole region boundary, close to the event horizon. Hence, the formula (4.32) can be applied to the BH-radiation system to compute how the entropy evolves under black hole evaporation, and determine whether the entropy follows the unitary Page curve [16]. In this section, we will follow the method taken in the review [16] alongside the original papers [13, 14].

Consider a unitary process that forms a BH in a pure state. After the BH forms, at early times before any Hawking radiation is emitted, no extremal surfaces are found by deforming X inwards, such that the entropy is minimised for a trivial surface at the centre $r = 0$ (see Fig. 4.5 (left)). This means that the area term in the generalised entropy vanishes and, as the BH remains in a pure state, the entanglement entropy S_{vN} is zero. Hence, the total entanglement entropy of the initial BH system (prior to entanglement) is zero. Ignoring the effects of Hawking radiation, this result is invariant under time evolution, in contrast to the B-H entropy which goes from zero initially to $4\pi r_s^2$ once the black hole forms. This highlights the nature of the B-H entropy as a bound on the entropy within an arbitrary region of spacetime, rather than the entropy followed by physical (unitary) processes.

Once the BH starts emitting Hawking radiation and evaporating, S_{vN} becomes finite as the BH transitions from a pure state to mixed with the emission and entanglement of radiation modes with degrees of freedom in the BH interior. As the black hole evaporates, S_{vN} continues to increase in line with the entropy of the emitted Hawking radiation due to increased interior degrees of freedom. If the trivial extremal surface is maintained over the BH evaporation, the entropy of the BH $S_{BH} = S_{vN}$ can obtain an arbitrary entropy that exceeds the Berkenstein bound as for the coarse-grained entropy, violating unitarity and the central dogma (see Fig. 4.5 (right)).

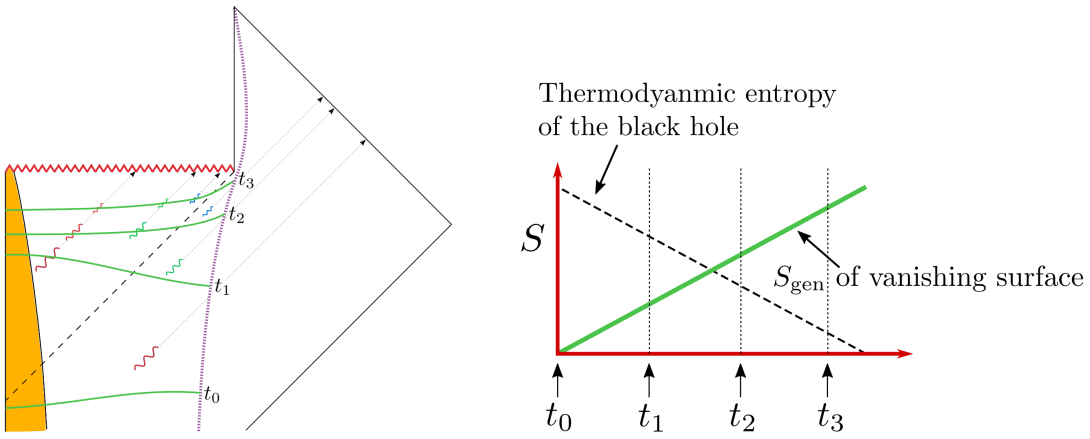


Figure 4.5: For trivial X, the entanglement entropy S_{gen} (green) calculated over the vanishing surface (left) increases monotonically due to the accumulation of entangled modes in the BH interior, whilst the thermodynamic entropy (dashed) declines as the BH evaporates (right). [16]

However, as Hawking radiation is emitted and the entanglement entropy rises, a new, non-trivial extremal surface emerges that replaces the trivial surface at $r = 0$ to minimise the entropy of the BH (see Fig. 4.6 (left)). As entangled degrees of freedom in the interior begin to build up, the trivial extremal surface shifts to one close to the event horizon of the black

hole. This means that some of the entangled modes within the interior no longer live within the region Σ_X bounded by the new non-trivial surface and the cut-off surface of the BH region. This is equivalent to "purifying" some of the interior modes by no longer considering them in isolation as mixed states contributing to the entanglement entropy. Hence, although the area term in the generalised entropy increases for the non-trivial surface, this is countered by a decrease in the von-Neumann entropy, such that the extremal surface X is switched from the trivial to non-trivial surface. Following the formation of the non-trivial surface at a scrambling time $t_s = r_s log S_{B-H}$, the area term in the generalised entropy dominates due to X being located close to the event. Hence, at late times after $t \sim t_s$, the generalised entropy follows the evolution of the thermodynamic entropy of the BH, saturating the Berkenstein bound and decreasing proportional to the area of the BH horizon (see Fig. 4.6 (right)).

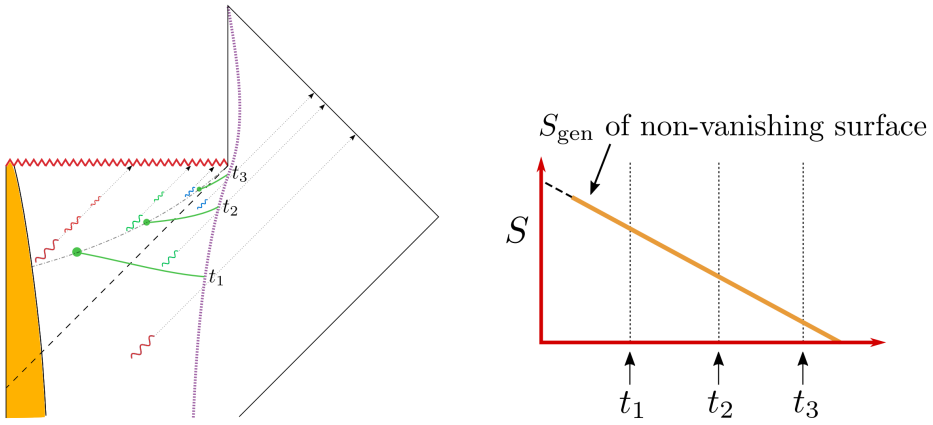


Figure 4.6: For non-trivial X forming a non-vanishing surface after the scrambling time, the entanglement entropy S_{gen} decreases (right) as the black hole evaporates and the surface evolves (left). [16]

Hence, by applying the entanglement entropy formula (4.32) and maxi-minimisation procedure, the fine-grained entropy of the BH will follow an increasing phase initially due to the vanishing extremal surface, and then a decreasing phase due to a non-vanishing surface (see Fig. 4.7). The time of the transition between extremal surfaces is called, suggestively, the *Page time* and takes the form of the Page curve (Fig. 2.3) required from unitary BH evaporation. Therefore, the BH entanglement, or fine-grained, entropy follows the unitarity Page curve.

4.6 Entropy of Hawking radiation

Whilst we have now confirmed that the evolution of a black hole follows the unitarity Page curve, this doesn't directly address the black hole information paradox. The entropy measured by an outside observer is that of the Hawking radiation, hence, in order to resolve the information paradox, we need to verify independently that the entropy of Hawking radiation also follows the Page curve [16].

First, in order to apply the generalised entropy formula to the radiation modes, the region Σ_{rad} is defined between the cut-off surface and asymptotic infinity (see Fug. 4.8). As radiation is emitted into the radiation region (complementary to the black hole region and bounded by the cut-off surface) as the black hole evaporates, the entanglement entropy $S_{vN}(\Sigma_{rad})$ will

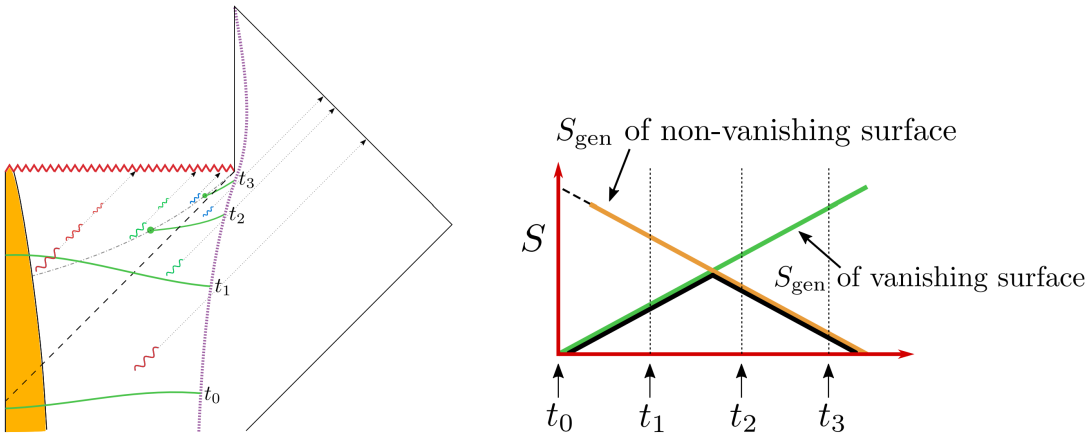


Figure 4.7: The Page curve (black) for the black hole entropy: the surface switches from vanishing to non-vanishing at the Page time (left), which corresponds to a transition from the increasing (green) to decreasing (orange) contributions at a maximum at the Page time (right). [16]

increase. Next, we need to consider the extremal surface that will minimise the entanglement entropy of the radiation. Namely, referring to the general formula (4.31) that follows from the HRT prescription, how should the region Σ_{X_A} be defined? It is found that the entropy for the radiation can be minimised by splitting Σ_{X_A} into disconnected regions, as a generalisation of the connected regions so far considered.

However, by considering Σ_{X_A} as disconnected, the area of the boundary is increased leading to an increased entanglement entropy. This increase in entropy can be countered in a similar way as for the black hole entropy – include more pairs of entangled modes in order to purify the system and decrease $S_{vN}(\Sigma_X)$. The radiation entropy, S_{rad}, can be decreased by considering 2 regions, the radiation region Σ_{rad} and a new "island" region Σ_{Island} defined between the center of the black hole and the surface X, in the same Cauchy slice at time t (see Fig. 4.8). For late times, the inclusion of the island region decreases the generalised entropy, forming a new minimal extremal surface $\Sigma_{rad} \cup \Sigma_{Island}$.

Note that the regions Σ_{rad} and Σ_{Island} over which the entropy of radiation is determined is the exact complement of the corresponding region for black holes for all times. This could be predicted as the total system was initially formed from a pure state, so the entropy of the black hole region and its complement are the same through the property of entanglement entropy and sum to zero when the regions are combined.

Hence, for the radiation, the full generalised entanglement entropy is given by the "island formula",

$$S_{rad} = min_X \Big\{ ext_X \Big[\frac{Area(X)}{4G_N} + S_{vN}[\Sigma_{rad} \cup \Sigma_{island}] \Big] \Big\}. \tag{4.33}$$

The island formula is implemented in the same way as before: the right hand-side is extremised with respect to the position of X, and then minimised with respect to all possible extremal surfaces and choices of islands.

Initially, there are no island contributions, such that the surface X is vanishing ($X = \emptyset$) and the only contribution to S_{rad} is the fine-grained entropy $S_{vN}(\Sigma_{rad})$. Σ_{rad} is the complement of

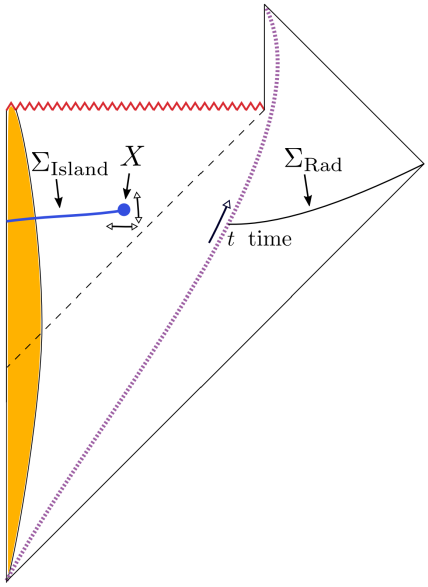


Figure 4.8: Penrose diagram showing an "island" contribution to the entanglement entropy of Hawking radiation from the region Σ_{Island} within the BH, as well as the exterior contribution over Σ_{Rad} and area term from the non-trivial surface X. [16]

the BH region Σ_X at early times, which covers the entire interior of the black hole region. As the radiation is emitted, $S_{vN}(\Sigma_{rad})$ increases (see Fig. 4.9 (right)).

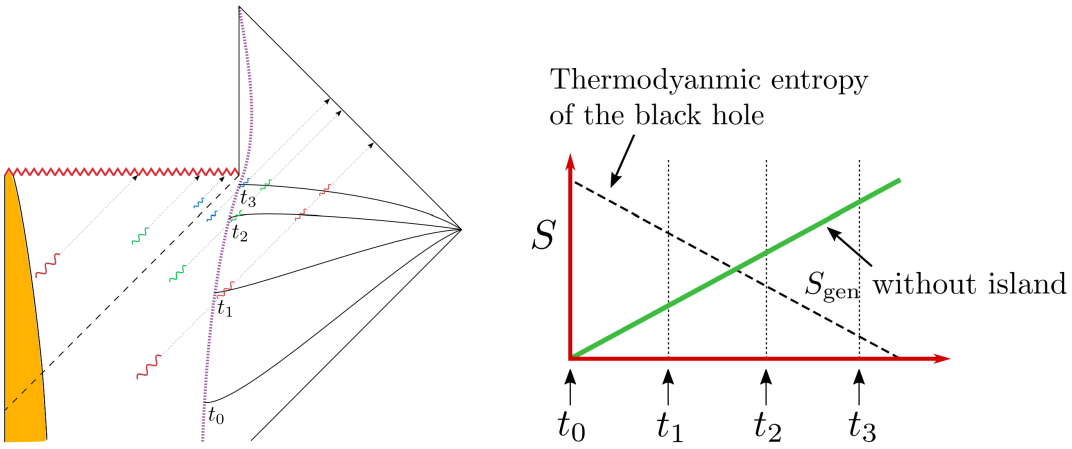


Figure 4.9: The evolution of the radiation entanglement entropy in the absence of an island contribution. Entropy calculated over Σ_{Rad} increases monotonically (green) as the black hole evaporates. [16]

At later times, the island region Σ_{island} emerges close to the event horizon as the complement to the non-vanishing surface for the BH entropy in the black hole region (see Fig. 4.10 (left)). As the non-trivial surface for the black hole appears at the scrambling time $t_s = r_s log(S_{BH})$, the island forms after time $\mathcal{O}(t_s)$. For late times and non-trivial surface X, the area term in (4.33) dominates over the entanglement entropy S_{vN} of the interior-radiation modes, such that the entropy of the Hawking radiation follows the thermodynamic BH entropy proportional to the area of the horizon and saturates the Berkenstein bound (see Fig. 4.10 (right)).

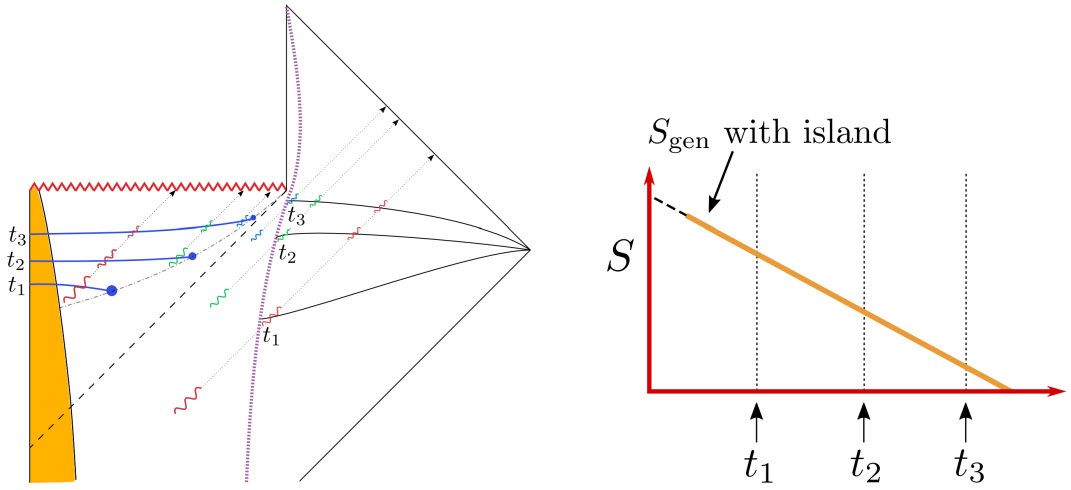


Figure 4.10: After a scrambling time, the island region (blue) forms in the interior. The entanglement entropy (orange) calculated over the union of Σ_{Island} and Σ_{Rad} decreases over time, following the thermodynamic BH entropy (right). [16]

Hence, the entanglement entropy of the Hawking radiation is the minimum of the 2 contributions, with the transition occurring at the Page time t_{Page}. This results in the entropy following the Page curve: Hawking radiation emitted in black hole evaporation evolves unitarily (Fig. 4.11).

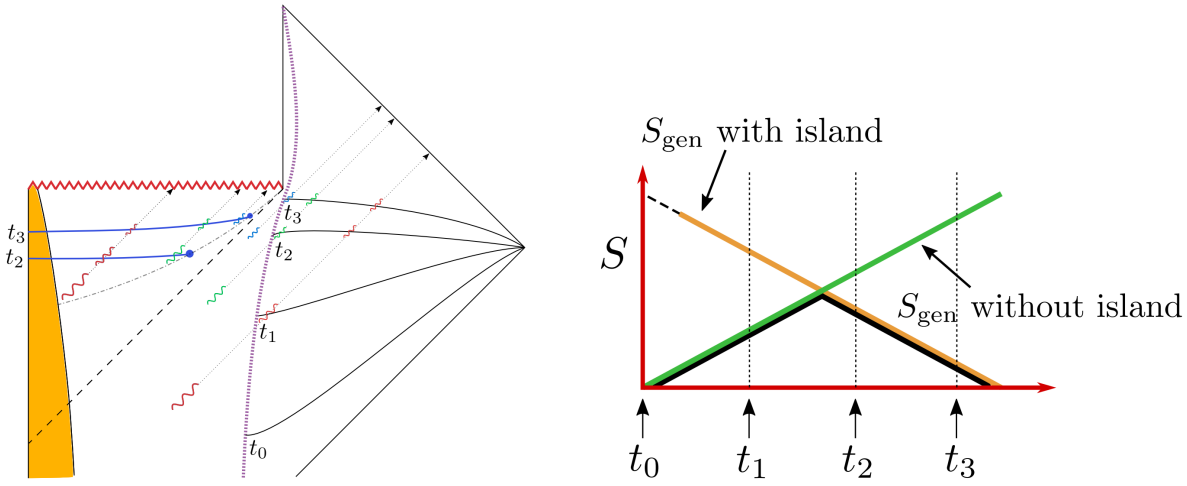


Figure 4.11: The Page curve (black) for Hawking radiation: minimising the two contributions, with and without the island region, leads to an increasing initial phase (green), maximum at the Page time, and a decreasing later phase (orange), as was found for the BH entanglement entropy. [16]

Also, we can see that the BH and Hawking radiation follow the same curve, having the same entropy for all stages of black hole evaporation. This is as expected, as forming the initial BH from a pure state, it is expected that $S_{BH} = S_{rad}$ form the properties of entanglement entropy. This is a result of the same surface X present in both systems and a matter state which is pure on the whole Cauchy slice having entropy $S = S_{vN}(\Sigma_X)$ - $S_{vN}(\Sigma_{rad} \cup \Sigma_{island}) = 0$.

An outstanding question of the information paradox is: why it is possible to simplify a QG calculation in this way, so as to only depend on a semi-classical gravity approximation for input without knowledge of the BH interior? The result of the Page curve emerges from gravitational path integrals. The holographic analogy to non-gravitational entropy calculated using the replica trick in the RT/HRT prescriptions is the calculation of the gravitational entanglement entropy directly using gravitational path integrals in AdS/CFT [65]. The gravitational path integral justifies the switch from trivial to non-trivial QESs which produces unitary evolution. Outstanding elements of the information paradox concern what precisely the entropy calculated through the gravitational path integral and QES corresponds to, and how the entanglement entropy of the radiation can be calculated directly from the entanglement entropy formula (2.15).

4.7 Entanglement wedge of the black hole and radiation

Through the calculation of the entanglement entropy for the Hawking radiation and black hole via. QESs in the previous sections, it is seen that the entropy has an explicit dependence on the geometry of the black hole interior. However, it is unclear how the entanglement entropy that follows from the central dogma is related to the degrees of freedom in the BH interior, and how the dependence of degrees of freedom on regions of the BH-radiation system change as the system evolves. The key to understanding this relationship between degrees of freedom and regions of dependence lies with the entanglement wedge: the causal domain of dependence (causal diamond) of the extremal surface used to calculate the entanglement entropy [16]. The causal diamond is defined by using past and future (null) light cones extended from the extremal surface to form a region containing all causally-connected points that can be determined by the boundary conditions on the extremal surface (see Fig. 2.2). In the case of disconnected extremal surfaces for the radiation sub-system at late times, the entanglement wedge will be disconnected.

Hence, the degrees of freedom described by the entanglement entropy for the BH and radiation depend only on a region of the interior bounded by the extremal surface for each time slice (time dependent). This means that knowing the entanglement entropy for each sub-system, the causal domain of dependence of the interior geometry between the cut-off and extremal surface for the black hole, and complementary region for the emitted radiation, is known.

As the extremal surfaces are time dependent, the entanglement wedges of the BH and radiation degrees of freedom also evolve in time. In particular, three stages of the black hole's evaporation can be considered in more detail.

At initial stages, where $t < t_{Page}$, the extremal surface X is located at $r = 0$, such that the black hole region Σ_X stretches from across the interior to the cut-off surface outside the horizon, with the BH entanglement wedge (green) covering a large portion of the interior's spacetime, see Fig. 4.12(a). Complementary to this, the region Σ_{rad} over which the radiation is considered goes from the cut-off surface to $r \to \infty$, with the radiation entanglement wedge (blue) covering a part of the spacetime outside the cut-off.

For later time after the Page time, $t > t_{Page}$, and before the black hole evaporates at t_{evap}, the shift to the non-trivial QES creates a smaller region over which the black hole entanglement entropy is calculated, such that the entanglement wedge covers a smaller path of the spacetime on either side of the horizon (see Fig. 4.12(b)). For the radiation, the entanglement wedge now extends to the interior of the black hole, through an "island" causal diamond of the Σ_{island}

region, as well as the exterior. This means that for the interval $t_{Page} < t < t_{evap}$, interior states living in the island part of the entanglement wedge contribute to the radiation entanglement entropy.

Finally, when the black hole evaporates and vanishes for $t > t_{evap}$, the radiation entanglement wedge covers the entire black hole "interior", with all degrees of freedom of the initial black hole state encoded in the Hawking radiation modes.

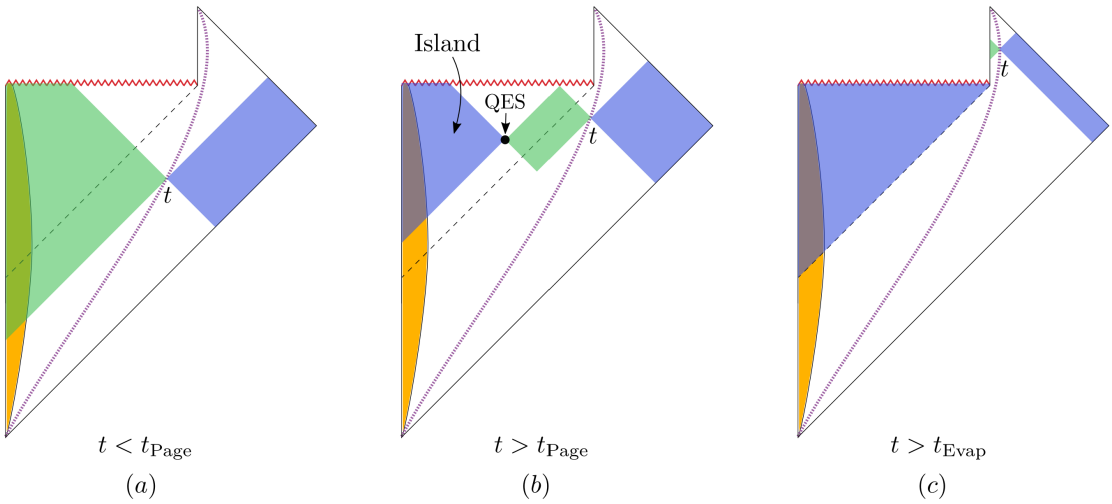


Figure 4.12: The evolution of the entanglement wedge for the radiation (blue) and the BH interior (green) modes: (a) at early times, the BH wedge occupies most of the interior to a point past the event horizon, and the raidation wedge occupies the exterior region; (b) after the Page time, the island region forms, constituting a part of the radiation entanglement wedge, with the BH wedge occupying a smaller region of the interior; and (c) at times after the black hole has evaporated, the radiation wedge occupies all of the interior and the exterior. [16]

The introduction of the entanglement wedge also helps resolve an earlier form of the information paradox in section 2.5: the "Bag of gold" geometry introduced by Wheeler [24]. Although it is possible to create classical geometry which resemble BHs from outside a "horizon" but have arbitrarily large entropy within (violating the central dogma and unitarity), the entanglement wedge will only cover part of the interior when the entropy in the interior exceeds the area of the horizon. This ensures that the "BH" entanglement entropy doesn't violate the central dogma. These classical geometries take the form of the Fig. 4.12(b) above.

Chapter 5

Gravity with holographic matter

In the following, we will seek to demonstrate a concrete example using the holographic techniques of the RT/HRT prescription to determine whether a black hole-radiation system is unitary, i.e. evolving according to the Page curve.

As discussed above, a simple argument to arrive at unitary evolution for Hawking radiation is to first compute the black hole entanglement entropy using the RT/HRT procedure to find the QES and compute the entropy on them. Then, considering an initially pure BH-radiation system, $S_{BH} = S_{rad}$, and as the S_{BH} follows the Page curve, then the radiation is unitary.

However, in order to prove that the radiation is unitary directly, it is necessary to show that the QESs for the radiation modes coincide with the QESs for the black hole modes, as discussed in [14]. This is most easily shown for an evaporating BH with holographic matter, so that the machinery of the AdS/CFT correspondence can be employed to compute the entropy holographically. The example explored in the following sub-sections is a two dimensional gravity-matter theory coupled to a two-dimensional (holographic) CFT for which a simple dual gravity theory exists in one higher dimension, following the procedure laid out by Almheiri, Mahajan, Maldacena and Zhao in 2019 [15]. It is expected that the results arising from the $d = 2$ case generalise to higher dimensions.

5.1 Two-dimensional gravity with holographic matter coupled to $\mathbf{CFT}_2$

In two dimensions, a general classical gravity theory described by the Einstein-Hilbert term,

$$I_{E-H} = \int d^d x \sqrt{-g} \mathcal{R}^{(2)}, \tag{5.1}$$

where $\mathcal{R}^{(2)}$ is the Ricci curvature scalar for a two dimensional spacetime. This term is purely topological, and doesn't contain any local dynamics, with the contribution to the entropy of the system a constant. However, for our discussion, we are interested in a dynamical, interacting spacetime.

The simplest dynamical spacetime in two dimensions is formed by coupling classical gravity to a dilaton field ϕ, described by the general action,

$$I_{grav}[g^{(2)}_{\mu\nu},\phi] = \frac{1}{16\pi G^{(2)}_N}\int_{\mathcal{M}} d^2x\sqrt{-g}\phi\mathcal{R}^{(2)} + U(\phi), \tag{5.2}$$

where $G^{(2)}_N$ is Newton's constant in two dimensions and $g^{(2)}_{\mu\nu}$ is the two-dimensional fixed metric which the gravity theory lives on. The pure Einstein-Hilbert term (5.1) has been absorbed by a shift in ϕ. This 2d dilaton gravity theory is equivalent to the Jackiw–Teitelboim (JT) gravity model, which describes the evaporation of near-extremal black holes [66, 67].

In order to model black hole dynamics, matter needs to be added to the gravity theory. Adding holographic matter described by CFT_2 with matter field χ and action $I_{CFT}[g^{(2)}_{\mu\nu},\chi]$, the total action of the theory becomes,

$$I[g^{(2)}_{\mu\nu},\phi,\chi] = I_{grav}[g^{(2)}_{\mu\nu},\phi] + I_{CFT}[g^{(2)}_{\mu\nu},\chi]. \tag{5.3}$$

The condition of holography means that there exists a dual three-dimensional gravity theory in (asymptotically) AdS_3 space to CFT_2, where the matter fields χ live on the two-dimensional conformal boundary. From the AdS/CFT dictionary, the metric for the three-dimensional dual theory, $g^{(3)}_{\mu\nu}$, has the boundary condition,

$$g^{(3)}_{ij}|_{\partial\mathcal{M}} = \frac{1}{\epsilon^2}g^{(2)}_{ij}, \tag{5.4}$$

where i,j denote the boundary indices, $g^{(2)}_{ij}$ is the fixed background metric of the CFT_2, $\partial\mathcal{M}$ denotes the conformal boundary of the asymptotically AdS_3 space, and ϵ acts as a short-distance (UV) cut-off. Furthermore, the curvature scalar for the 3-dimensional gravity at the boundary, $R_{AdS^{(3)}_3}|_{\partial\mathcal{M}}$, is dual to the 2 dimensional stress-energy tensor for the CFT_2.

To justify the semi-classical limit in the two-dimensional theory, and the large radius of curvature of the dual theory in three-dimensions, it is required that the central charge c of the CFT_2 satisfies,

$$1 << c << \frac{\phi}{4G^{(2)}_N}. \tag{5.5}$$

Also, to have an Einstein (weakly coupled) dual gravity theory, the CFT_2 must be strongly-coupled (see Fig. 3.4).

In order to apply the RT/HRT prescription to describe the QESs of this coupled gravity-matter system, the three-dimensional dual to the full theory described by the action (5.3) is needed. Starting with the geometry $g^{(3)}_{\mu\nu}$ and the boundary metric $g^{(2)}_{ij}$, the dilaton scalar ϕ is added to the boundary $\partial\mathcal{M}$ of the AdS_3 space (with the associated action (5.2)), and ϕ and $g^{(2)}_{ij}$ are integrated over. This gives a three-dimensional bulk metric which looks locally like AdS_3, with a non-conformal boundary, upon which the dilaton gravity action (5.2) lives, at a finite distance in the space.

This procedure differs from the AdS/CFT correspondence, as the fixed boundary metric $g^{(2)}_{ij}$ is also integrated over. This results in a geometry which is the same as that of the Randall-Sundrum (RS) model, an alternative to dimensional compactification in string theory [68]. In

the RS model, the dynamical boundary brane (when considering higher dimensions) is called the *Planck brane*. Hence, the CFT_2 with rigid metric $g_{ij}^{(2)}$ coupled to two-dimensional dilaton gravity theory (described by (5.3), see Fig. 5.1 (left)) has a 3-dimensional geometry description where the matter CFT_2 is replaced by the 3-dimensional dual and the dilaton-gravity action lives on the two-dimensional Planck brane (Fig. 5.1 (right)).

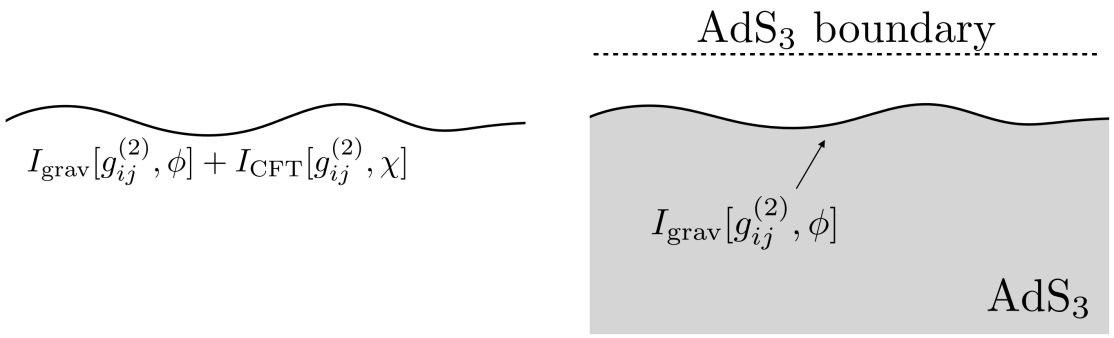


Figure 5.1: The left diagram shows a 2d dilaton-gravity theory coupled to a holographic matter CFT_2 (composed of matter fields χ), containing the full action (5.3). The right diagram shows the 3d geometry with AdS_3 bulk dual to the matter CFT_2, where the dilaton-gravity action (5.2) lives on a dynamic Planck brane. [15]

5.1.1 Embedding of the Planck brane in AdS_3

Having identified the action (5.2) living on the Planck brane, the location of the Planck brane, i.e. how the two-dimensional surface is embedded in AdS_3, can be determined for use in the RT/HRT minimisation procedure later. The embedding of the Planck brane in the 3-dimensional bulk is computed by using the two-dimensional metric and stress tensor profile that follows from the solutions of the full action (5.3). Considering the two-dimensional gravity-matter theory living in an asymptotically AdS_2 geometry, the metric and stress-energy tensor can be expressed as,

$$ds^2 = -\frac{dy^+ dy^-}{(y^+ - y^-)^2} = -e^{2\rho(y)} dy^+ dy^-, \tag{5.6}$$

and

$$T_{y^+y^+}(y^+) \quad \text{and} \quad T_{y^-y^-}(y^-), \tag{5.7}$$

having applied a change of coordinates $t = \frac{y^+ - y^-}{2}$, $z = \frac{y^+ + y^-}{2}$ to (3.30). The stress-energy tensor of the CFT_2 (5.7) is measured in the flat metric $ds^2 = -dy^+ dy^-$. In general, the full stress tensor is related to the curvature of the three-dimensional gravity geometry, which follows from corrections due to the derivatives of $\rho(y)$ forming a conformal anomaly related to the non-vanishing Ricci scalar.

Performing the coordinate transformation $\omega^+(y^+), \omega^-(y^-)$ such that the stress-energy tensor vanishes (locally), under a general diffeomorphism the stress-energy tensor will transform as,

$$\Big(\frac{d\omega}{dy}\Big)^2 T_{\omega^\pm\omega^\pm} = T_{y^\pm y^\pm} + \frac{c}{24\pi}\{\omega^\pm(y), y^\pm\}, \tag{5.8}$$

where $\{\omega(y), y\} = \frac{\omega''}{\omega'} - \frac{3}{2}\big(\frac{\omega''}{\omega'}\big)^2$ is the Schwarzian derivative. Then, taking a Weyl transformation of the metric (5.6) to a flat metric, the $\omega^\pm$ coordinates ensure that the stress tensor vanishes,

$$ds^2 = -d\omega^+ d\omega^- \quad , \quad T_{\omega^+\omega^+} = T_{\omega^-\omega^-} = 0. \tag{5.9}$$

Hence, (5.6) and (5.7) are related to a vacuum solution on flat space, (5.9), through a coordinate transformation and diffeomorphism. The location of the Planck brane can be determined in $\omega^\pm$ coordinates in the following way [69]. The vacuum of the holographic CFT_2 has a dual gravity theory in the bulk of pure AdS_3, described by,

$$ds^2 = \frac{-d\omega^+ d\omega^- + dz_\omega^2}{z_\omega^2}, \tag{5.10}$$

which reduces to a flat space geometry with vanishing $T_{\omega^\pm\omega^\pm}$ for a surface of constant z_ω. Hence, recalling (5.9), we can equate the near-geometry of the Planck brane to the metric (5.10). Then, the boundary condition on the three-dimensional boundary metric (5.4) can be applied to relate the metric of the two-dimensional gravity (5.6) to the AdS metric (5.10) for $z_\omega =$ constant,

$$-\frac{d\omega^+ d\omega^-}{z_\omega^2} = \frac{1}{\epsilon^2} e^{2\rho(y)} dy^+ dy^-. \tag{5.11}$$

Solving this leads to the Planck brane location at,

$$z_\omega = \epsilon e^{-\rho(y)} \sqrt{\frac{d\omega^+}{dy^+}\frac{d\omega^-}{dy^-}}. \tag{5.12}$$

Hence, once the geometry of the two-dimensional geometry of (5.9) is known, the embedding of the two-dimensional geometry into the three-dimensional geometry follows easily. This result can be checked by starting from the stress tensor in terms of the extrinsic curvature to recover the form found in (5.7).

Following this, the computation of the RT/HRT surfaces is simple in the $(z_\omega, \omega^\pm)$ coordinates.[1]

5.2 Two-dimensional black hole coupled to a holographic bath

Considering a black hole in the two-dimensional geometry described by (5.3), we now attach an external (holographic) CFT_2 "bath" of constant zero temperature such that radiation can

[1] Note that the state of the CFT is encoded in more complex geometry deeper in the interior, upon which the area of the embedded RT/HRT surfaces will depend on.

be exchanged between the gravity-matter theory (black hole) and the bath. As the black hole evaporates, the Hawking radiation emitted will accumulate in the bath, in analogy with the radiation region in the BH-radiation system discussed in section 2.5. As before, the two-dimensional geometry is taken to be asymptotic to AdS_2. For simplicity, a toy model of the dilaton-gravity theory is considered where $\phi >> 1$, so that the effects of backreactions may be neglected and matter lives on a fixed, non-dynamical background.

As the bath is described by the same CFT as the matter sector, I_{CFT}, of the full black hole action (5.3), the coupling of the bath and matter CFTs in the two-dimensional theory amounts to joining them at their boundaries such that they can exchange stress energy freely. Defining the coordinate $\sigma_y = \frac{y^+ - y^-}{2}$, points in the bath and asymptotic AdS_2 systems are given by positive and negative values of σ_y respectively, with the boundary of the black hole and bath at $\sigma_y = 0$ (see Fig. 5.2 (left)).

Also, making use of the equivalence between the two-dimensional dilaton-gravity coupled to the holographic matter CFT and the three-dimensional geometry on the dual AdS space (see Fig. 5.1), the coupling of the two-dimensional gravity-matter theory to the bath can be described in the dual AdS space, where the rigid bath CFT_2 lives on the conformal boundary of AdS_3 and is joined to the dynamic Planck brane at $\sigma_y = 0$ (see Fig. 5.2 (centre)).

A third, QM description of the black hole-bath system also exists: as the matter CFT_2 lives in (asymptotically) AdS_2 space, the gravity-matter system can be holographically translated to its dual CFT_1, which is simply QM with additional symmetry constraints. Hence, the two-dimensional black hole-bath system is realised as a dual QM system living at $\sigma_y = 0$ coupled to the bath CFT_2 on the half-line $\sigma_y > 0$ (see Fig. 5.2 (right)). This can also be viewed from the point of view of the central dogma, with the black hole in AdS_2 space equivalent to a QM system with $Area/4G_N$ degrees of freedom. Hence, there are three equivalent descriptions of the gravity-matter theory coupled to a holographic bath, displayed in Fig. 5.2.

5.3 The entanglement entropy of the 2-dimensional theory

We are now in a position to compute the entanglement entropy of the two-dim. gravity-matter theory by extremising the generalised entropy as described in section 4.4.

First, as prescribed by EW [12], a generalised entropy needs to be constructed in analogy to (4.31),

$$S_{gen}(y) = \frac{\phi(y)}{4G_N^{(2)}} + S_{\text{2d-bulk}}(\mathcal{I}_y), \tag{5.13}$$

where the coordinate y is a point in the bulk of the two-dimensional theory; and $\mathcal{I}_y$ is the interval between y and the conformal boundary of the AdS_2 space, to which the gravity + matter theory is asymptotic to (or equivalently $\phi << 1$ to ensure the gravity + matter theory is weakly coupled). $S_{\text{2d-bulk}}(\mathcal{I}_y)$ is the bulk entanglement entropy of the interval $\mathcal{I}_y$, which contains (dominant) contributions from the bulk matter fields χ as well as (sub-dominant) quantum fluctuations from the boundary fields, ϕ and $g_{ij}^{(2)}$. Comparing (4.31) and (5.13), in two-dimensions $\phi(y) = Area^{(2)}$ as the area of a Ricci scalar point is the coefficient of the curvature term. These surfaces and the quantities on them are shown in Fig. 5.3.

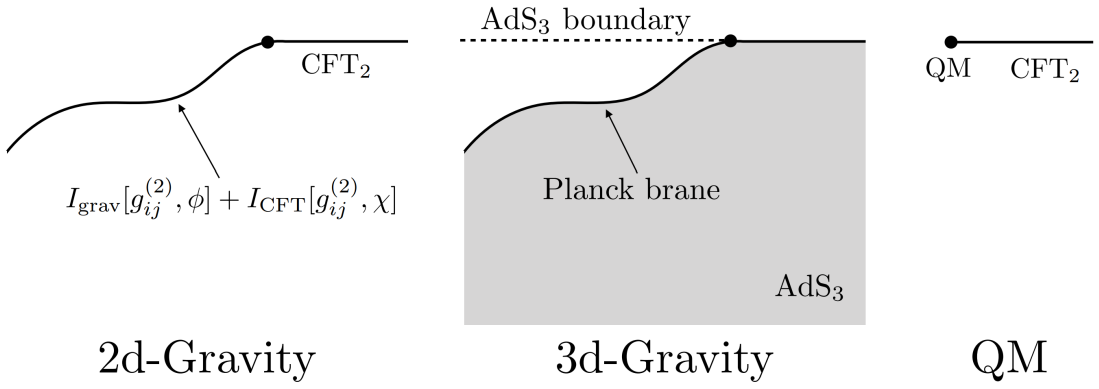


Figure 5.2: The 3 alternative descriptions of the gravity-matter theory coupled to a bath (left to right): 2d dilaton-gravity plus matter CFT$_2$ coupled to a CFT$_2$ bath; 3d gravity theory, with the holographic CFT$_2$ of the matter and bath replaced with the dual theory; and a fully QM description, where the 2d gravity-matter theory (in asymptotic AdS$_2$) is replaced by its QM dual at the boundary of the bath. The 3d gravity description has a dynamical Planck brane, in addition to the conformal boundary. [15]

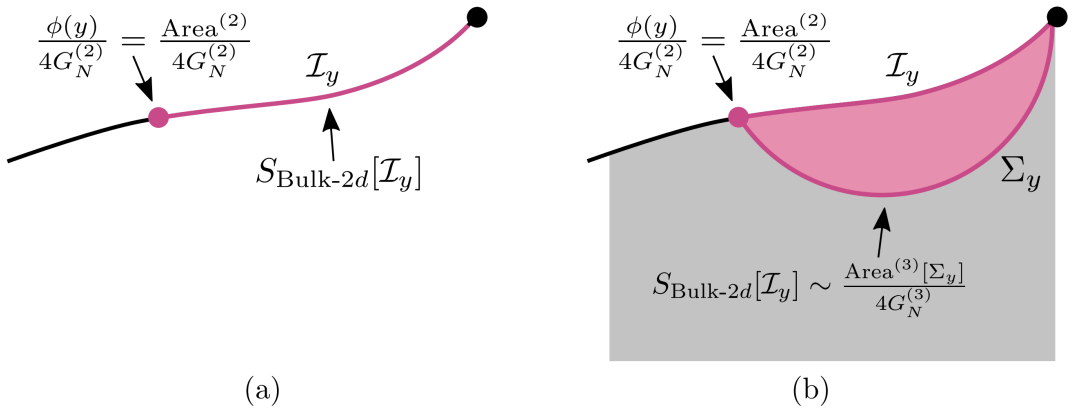


Figure 5.3: (a) The area and bulk contributions to the entropy $S_{gen}(y)$ in the 2d gravity theory, (5.13); (b) transforming to the dual 3d theory, $S_{2d-bulk}$ can be computed using the extremal surface Σ_y through the 3d RT/HRT formula (5.14). [15]

Now, following EW [12], (5.13) is extremised over all possible points of y, and then the minimum over all these extremums with points (y_e^+, y_e^-) chosen to obtain the QES. As the matter fields χ of the CFT$_2$ dominate in $S_{2d-bulk}$, this contribution to the total entanglement entropy can be computed by considering the dual theory to the holographic CFT$_2$ and computing the entropy over the minimal/extremal surfaces of this 3-dim. dual theory via. the RT/HRT prescription [10, 11]. The extremal surface Σ_y, bounded by y and the endpoint of the Planck

brane is an interval in this case (see Fig. 5.3 (b)). Hence, applying the HRT prescription (section 4.3), the generalised entropy (5.13) is approximated as,

$$\begin{aligned} S_{gen}(y) &= \frac{\phi}{4G_N^{(2)}} + S_{2dbulk}(\mathcal{I}_y) \\ &\approx \frac{\phi}{4G_N^{(2)}} + \frac{Area^{(3)}(\Sigma_y)}{4G_N^{(3)}}, \end{aligned} \tag{5.14}$$

where we the $\approx$ comes from neglecting the subdominant contributions from quantum fluctuations of the boundary fields, and dropping the sub-leading 3d bulk entanglement entropy.

Now, extremising the generalised entropy (5.13) in two dimensions can be translated into extremisng the usual RT/HRT surfaces in the dual three-dimensional theory with endpoints on a dynamical Planck brane. As before for the extremisation of (4.31) for the BH-radiation system, extremising Σ_y involves the area contribution $Area(\Sigma_y)$ along the interval (y_e^+, y_e^-) in the Planck brane, as well as a contribution from the dilaton ϕ at the dynamic boundary. As such, at leading order, the extremization of QESs reduces to following the RT/HRT prescription in three dimensions [16].

5.4 Entanglement wedges for an evaporating JT black hole

Having applied the RT/HRT prescription to JT gravity, the entanglement entropy can be determined by extremization of (5.14) via. the method above. Using the known results of JT gravity [13], the evolution of the QESs and entanglement wedges for a black holes in JT gravity evaporating into non-gravitational bath can be reviewed.

5.4.1 Early times

Initially, consider a low temperature black hole decoupled from a non-gravitational bath, where the bath and black hole matter CFTs have holographic duals consisting of two disconnected geometries in AdS_3. The holographic dual AdS_3 of the Cardy boundary condition imposed in the 2d CFTs [70] is called the *Cardy brane*, located at $\sigma_\omega = (\omega^+ - \omega^-)/2 = \pm 0$ on the boundary between the Planck brane and conformal boundary, and descending into the 3d bulk from $\sigma_\omega = 0$ (see Fig. 5.4).

At t = 0, the BH and bath become coupled over an interval Δt. The impulse of energy transmitted between the systems during the coupling is inversely proportional to the coupling interval, $E \sim \frac{c}{\Delta t}$, such that an instantaneous coupling corresponds to an infinite pulse of energy, $E \to \infty$ as $\Delta t \to 0$. Hence, the coupling is carried out over a finite interval.

For a BH with initial temperature T_i following the energy pulse at coupling, the BH temperature declines as Hawking radiation is emitted into the bath, following the relation,

$$T(t) \sim T_i e^{-\frac{\kappa}{2}t}, \tag{5.15}$$

where κ is proportional to a constant c.

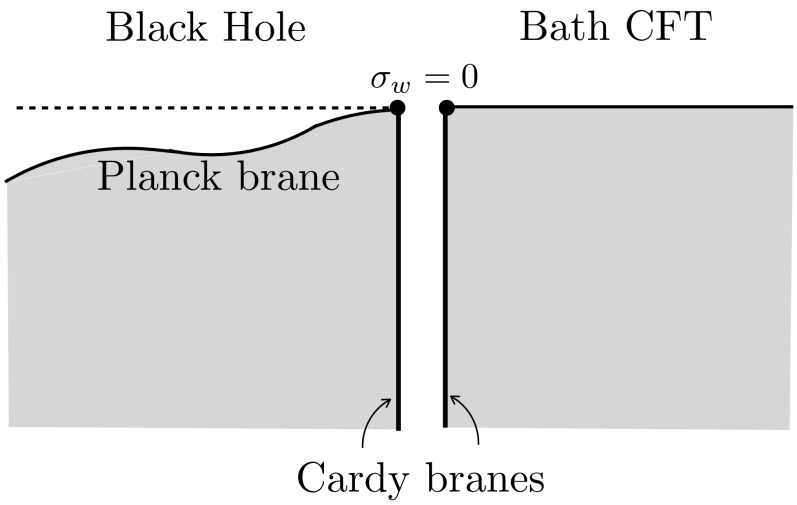


Figure 5.4: The initial decoupled bath (right) and black hole (left) system in the 3d dual theory, with the conformal B.C.s. in the 2d description producing Cardy branes anchored at $\sigma_\omega = \pm 0$ on the conformal boundary in the 3d dual theory. [15]

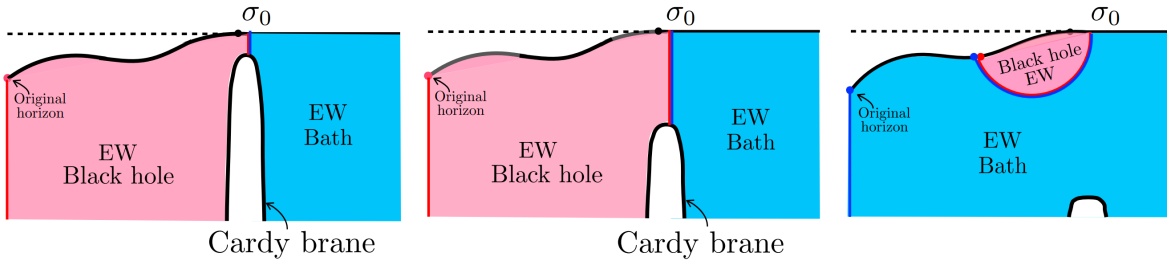
Initially, the BH entanglement wedge occupies all of the black hole region up to a bound close to the event horizon prior to coupling (Fig. 5.5(a)). This is analogous to the trivial QES at early times for the Page curve of the BH-radiation system in section 4.7. Similarly, the radiation entanglement wedge occupies almost all of the bath region.

After coupling at $t = 0$, the $3d$ geometry becomes Lorentzian, with the Cardy brane shifting away from the conformal boundary into the three-dim. bulk. For later times before the Page time, $0 < t < t_{Page}$, the Cardy brane continues to fall further into the bulk whilst the entanglement wedges and topology of the three-dim. theory remain the same (Fig. 5.5(b), (c)). The increasing distance between the conformal boundary and Cardy brane leads to a growing entanglement entropy, which physically corresponds to the accumulation of pairs of entanglement modes separated between the BH and bath.

This entropy is equal to the entropy of the Hawking radiation entering the bath, and evolves according to,

$$\begin{aligned} S_{BH}(t) \ \sim S_{rad}(t) &= \frac{\pi c}{6} \int_0^t dt' \, T(t') \\ &= 2S^i_{B-H}(1 - e^{-\frac{\kappa}{2}t}), \end{aligned} \tag{5.16}$$

where $T(t')$ is the temperature at time t' (which evolves according (5.15)), S^i_{B-H} is the initial (coarse-grained) Berkenstein-Hawking entropy following the coupling, and c is a constant. As the temperature of the BH declines over time, $S_{BH}(t) \sim S_{rad}(t)$ (5.16) increases until it saturates twice the initial coarse-grained entropy of the BH (with the factor of 2 in (5.16) due to Hawking radiation not being adiabatic [9, 14]).



(a) $t = 0$ (b) $0 < t < t_{\text{Page}}$ (c) $t > t_{\text{Page}}$

Figure 5.5: The evolution of the Cardy brane and entanglement wedges of the black hole (pink) and bath (blue): at t = 0, the Cardy brane shifts away from the conformal boundary, falling deeper into the bulk as time progresses. At late times, $t > t_{\text{Page}}$, the entanglement wedges shift due to a transition in the extremal surfaces. [15]

5.4.2 Late times

At late times, for $t > t_{Page}$, a new BH QES forms [13,14]. In the QM description of the BH-bath system, the QES is defined as the interval $\sigma_\omega \in [0, \sigma_0]$, where σ_0 is a small extension of the surface into the CFT bath. This means that all the degrees of freedom at the CFT boundary are encompassed, corresponding to the data from the asymptotically AdS_2 gravity-matter theory in the dual two-dimensional description (see Fig. 5.6 (right)).

As for the BH-radiation system in section 4.7, the new QES is located close to the horizon of the BH, on the interior side at a point (y_e^+, y_e^-). In the two-dim. gravity description, the location of the QES can be defined by a past-directed light ray that reaches the AdS_2 boundary at a scrambling time before the entropy is evaluated at time t,

$$y_e^+ = t - \frac{1}{2\pi T(t)} log\left(\frac{S_{B-H}(T(t)) - S_0}{c}\right) + \ldots, \tag{5.17}$$

where S_0 is the extremal entropy of the BH before coupling, and $S_{B-H}(T)$ is the B-H entropy for a BH at temperature T, with $S_0 << S_{B-H}$. Hence, the BH entanglement entropy at late times is calculated over the RT/HRT surface in the interval $[y_e, \sigma_0]$ at time t. In the three-dimensional case, the entanglement wedge of the black hole is then the causal domain of this interval (see Fig. 5.5(c) and Fig. 5.6(b)).

Then, the non-trivial extremal surface leads to an extropy for the BH,

$$S_{BH}(t) = S_{B-H}(T(t)) + \mathcal{O}(log(S_0)), \tag{5.18}$$

where the log terms are subdominant as $S_{B-H} >> S_0$. As the BH evaporates and the temperature decreases, the BH entropy (5.18) decreases for late times $t > t_{Page}$.

Similarly, the entanglement wedge for radiation can be considered at late time. Up until the Page time, t_{Page}, the entanglement entropy of the bath increases as the distance between the Cardy brane and conformal boundary increases (as for the black hole), see Fig. 5.5 (left, centre). At later times, $t > t_{Page}$, the entropy of the state in the bath CFT is computed outside

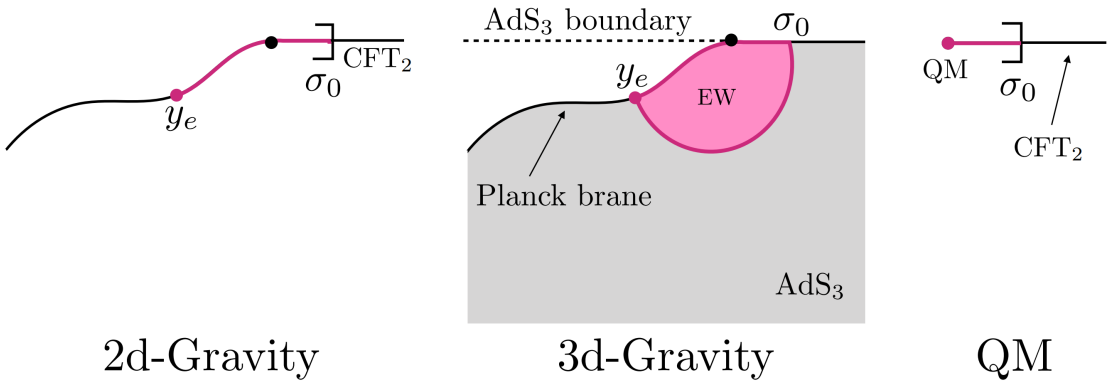


Figure 5.6: At late times, the entanglement wedge for the black hole (pink) is determined by the RT/HRT surface Σ_y in the 3d dual theory (middle panel). In the 2d geometry (left), the entropy is calculated over an interval $[y_e, \sigma_0]$. In the QM picture, this interval reduces to the σ_0 upper bound in the bath containing the degrees of freedom of the dual QM theory. [15]

(to the right of) the point σ_0 in a finite interval $[\sigma_0, \sigma_{IR}]$, where $\sigma_{IR} > t_{evap}$ is large enough to contain all the Hawking radiation in the bath (see Fig. 5.7 (right)). This creates another contribution to the RT/HRT surface, bounding the radiation entanglement wedge in the bath between the σ_{IR} endpoint and the Cardy brane (see Fig. 5.7 (centre)). The entropy this surface contributes is relatively small, $S \sim \frac{c}{6} log \sigma_{IR} << S^i_{B-H}$. Naively, from this picture, it seems that the entanglement entropy continues to grow past the Page time, as initially calculated by Hawking [2].

However, applying insight from [14,16], as the BH entanglement wedge only covers a portion of the BH interior, it can be predicted that the radiation entanglement wedge must cover the remaining interior region. This is not immediately apparent from the two-dim. gravity description: the bath region $[\sigma_0, \sigma_{IR}]$ and the BH interior are disconnected, forming an "island" region inside the BH (see Fig. 5.7 (left)).

However, translating this into the three-dimensional dual gravity theory, the interior and exterior regions are connected through the additional dimension in the bulk (see Fig. 5.7 (centre)). This provides a physical realisation of the resolution of the ER=EPR paradox proposed by Maldacena and Susskind [26], with entangled modes connected via. the "bridge" between the BH "island" and CFT bath "mainland", formed by the extra dimension. Although this is described for the $d = 2$ case, it is expected that this extra-dimension connection extends to any BH coupled to holographic matter with $d > 2$. Finally, to accurately describe the evolution of the radiation entanglement wedge, we also need to take into account the dependence on the initial state of the system. The initial state has an entropy S_0, and therefore additional part of the total RT/HRT surface, associated to the original horizon of the low temperature BH (see Fig. 5.7 (left, centre)).

Hence, it follows that, up to IR corrections and additional boundary contributions dependent on the initial state of the system, the RT/HRT surface for the radiation is the same as that for the BH, with the radiation entanglement wedge complement to the black hole's, as is expected for a pure initial state.

The complete RT/HRT extremal surface for the radiation entanglement wedge is made up of 3 parts (Fig. 5.7): the surface located at the original horizon; the extremal surface between σ_0 in the bath and the Planck brane; and the surface in the bath between σ_{IR} and the Cardy brane.

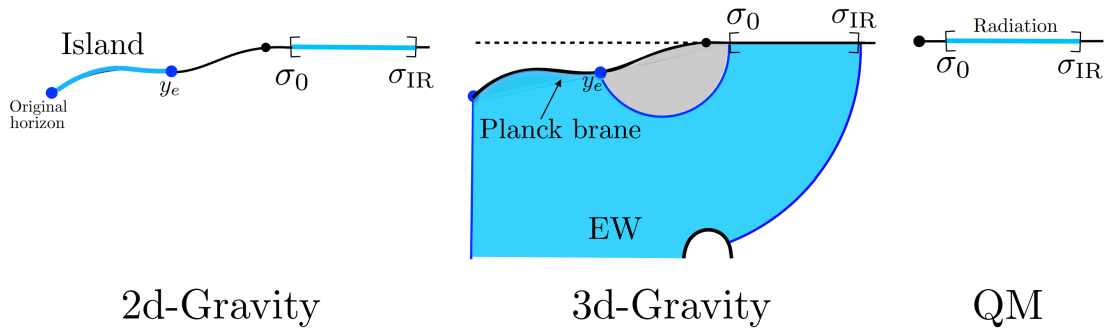


Figure 5.7: The entanglement wedge for radiation (blue) at late times, including the IR and original BH horizon contributions to the RT/HRT surface and entanglement entropy, as well as the surface shared with the BH. [15]

5.5 The Page curve

Now that the extremal surfaces and entanglement wedges have been described, the evolution of the Hawking radiation can be described. At initial times, the entropy obeys (5.16), increasing to a maximum at the Page time (defined as the time at which (5.16) and (5.18) are equal). At the Page time, the extremal surface shifts according to the minimisation condition of the RT/HRT procedure. Hence, after t_{Page}, the entropy of the radiation declines according to (5.18). This produces the transition in Fig. 5.8 that takes the form of the Page curve described generally in section 2.5.

In conclusion, for a 2d JT gravity + matter theory coupled to a (zero temperature) CFT bath, the Hawking radiation emitted by an evaporating black hole has been shown to evolve unitarily according to the Page curve, with the maximum entropy within the Berkenstein bound, S^i_{B-H}.

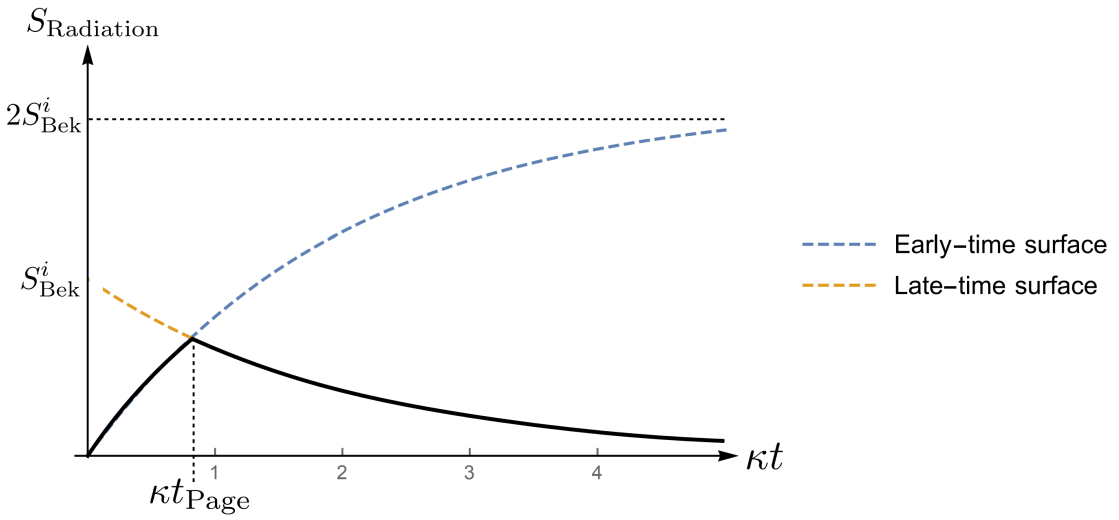


Figure 5.8: A diagram showing the Page curve (black) for the 2d JT BH coupled to a holographic bath system: at the Page time, the minimisation procedure enforces the transition from the early-time surface (blue dashed) to the late-time surface (yellow dashed). κ is a constant proportional to c $\propto$ 2d effective gravitational coupling. [15]

Chapter 6

Conclusion

In summary, the AdS/CFT correspondence has been applied to calculate the entanglement entropy of Hawking radiation for an evaporating black hole in 2d JT gravity coupled to a holographic bath. From the evolution of the radiation's entanglement wedge, and notably the inclusion of the "island" region, the Page curve for unitary evolution has been recovered. This result has recently been extended to a thermal holographic bath [71], and a gravity-matter theory living in an asymptotically flat background [72]. As this 2-dimensional case is believed to be generalisable, this result should also hold for higher dimensional black holes, providing evidence that black holes are unitary and preserve information in a system more generally. Current work has been done on directly showing the Page curve for higher-dimensional black holes, inspired by the method taken for the 2d dilaton toy model [73].

Although this provides an answer to one aspect of the black hole information paradox, it has also raised several new questions. Namely, what precisely does the entropy arising from the gravitational path integral and QES correspond to, and how can this entropy be directly calculated from the defining equation for entanglement entropy? Recent work to answer these questions have focussed on justifying the use of the QES in the entropy calculation [74, 75].

Acknowledgments

I am very grateful to the Thomas Roe Foundation, the Dorothy Johnson Charitable Trust, the Sidney Perry Foundation, the Vegetarian Charity, and the Humanitarian Trust for their financial support over the past academic year. I would also like to thank those at Imperial who have provided inspiration and guidance during the QFFF course.

Bibliography

[1] Remo Ruffini and John A Wheeler. Introducing the Black Hole. *Physics Today*, 1971.

[2] Stephen W Hawking. Breakdown of predictability in gravitational collapse. *Phys. Rev. D*, 14(10):2460, 1976.

[3] Jacob D. Bekenstein. Universal upper bound on the entropy-to-energy ratio for bounded systems. *Phys. Rev. D*, 23:287–298, 1981.

[4] Gerard't Hooft. Dimensional reduction in quantum gravity. *arXiv preprint gr-qc/9310026*, 1993.

[5] Leonard Susskind. The world as a hologram. *Journal of Mathematical Physics*, 36(11):6377–6396, 1995.

[6] Juan Maldacena. The large-N limit of superconformal field theories and supergravity. *International journal of theoretical physics*, 38(4):1113–1133, 1999.

[7] S. W. Hawking. Breakdown of predictability in gravitational collapse. *Phys. Rev. D*, 14:2460–2473, Nov 1976.

[8] Don N Page. Information in black hole radiation. *Physical Review Letters*, 71(23):3743, 1993.

[9] Don N Page. Time dependence of Hawking radiation entropy. *Journal of Cosmology and Astroparticle Physics*, 2013(09):028, 2013.

[10] Shinsei Ryu and Tadashi Takayanagi. Holographic derivation of entanglement entropy from the Anti–de Sitter space/conformal field theory correspondence. *Physical Review Letters*, 96(18):181602, 2006.

[11] Veronika E Hubeny, Mukund Rangamani, and Tadashi Takayanagi. A covariant holographic entanglement entropy proposal. *Journal of High Energy Physics*, 2007(07):062, 2007.

[12] Netta Engelhardt and Aron C Wall. Quantum extremal surfaces: holographic entanglement entropy beyond the classical regime. *Journal of High Energy Physics*, 2015(1):1–27, 2015.

[13] Ahmed Almheiri, Netta Engelhardt, Donald Marolf, and Henry Maxfield. The entropy of bulk quantum fields and the entanglement wedge of an evaporating black hole. *Journal of High Energy Physics*, 2019(12):1–47, 2019.

[14] Geoffrey Penington. Entanglement wedge reconstruction and the information paradox. *Journal of High Energy Physics*, 2020(9):1–84, 2020.

[15] Ahmed Almheiri, Raghu Mahajan, Juan Maldacena, and Ying Zhao. The Page curve of Hawking radiation from semiclassical geometry. *Journal of High Energy Physics*, 2020(3):1–24, 2020.

[16] Ahmed Almheiri, Thomas Hartman, Juan Maldacena, Edgar Shaghoulian, and Amirhossein Tajdini. The entropy of Hawking radiation. *Reviews of Modern Physics*, 93(3):035002, 2021.

[17] Richard C Tolman. On the weight of heat and thermal equilibrium in general relativity. *Physical Review*, 35(8):904, 1930.

[18] James M Bardeen, Brandon Carter, and Stephen W Hawking. The four laws of black hole mechanics. *Communications in mathematical physics*, 31(2):161–170, 1973.

[19] Andrew Strominger and Cumrun Vafa. Microscopic origin of the Berkenstein-Hawking entropy. *Physics Letters B*, 379(1-4):99–104, 1996.

[20] Andrew Tolley. Lecture notes in Foundations of QM. *Imperial College London*, 2021.

[21] Daniel Harlow. Jerusalem lectures on black holes and quantum information. *Reviews of Modern Physics*, 88(1):015002, 2016.

[22] Tadashi Takayanagi. Lectures on Entanglement Entropy and AdS/CFT. *Arnold Sommerfeld School, LMU Munich*, 2013.

[23] Cao H Nam. Entanglement entropy and Page curve of black holes with island in massive gravity. *The European Physical Journal C*, 82(4):1–18, 2022.

[24] John A Wheeler. Relativity, groups and topology. *Edited by C. De-Witt and B. DeWitt, Gordon and Breach, New York*, page 316, 1964.

[25] Gerard t Hooft. The quantum black hole as a hydrogen atom: microstates without strings attached. *arXiv preprint arXiv:1605.05119*, 2016.

[26] Juan Maldacena and Leonard Susskind. Cool horizons for entangled black holes. *Fortschritte der Physik*, 61(9):781–811, 2013.

[27] Gerard 't Hooft. A planar diagram theory for strong interactions. In *The Large N Expansion In Quantum Field Theory And Statistical Physics: From Spin Systems to 2-Dimensional Gravity*, pages 80–92. World Scientific, 1993.

[28] Ofer Aharony, Steven S Gubser, Juan Maldacena, Hirosi Ooguri, and Yaron Oz. Large N field theories, string theory and gravity. *Physics Reports*, 323(3-4):183–386, 2000.

[29] Leonardo Rastelli. Lectures on Open/Closed Duality. *PiTP II, IAS, Princeton*, 2004.

[30] Robi Peschanski. Gauge-gravity duality and high energy collisions. *Nuclear Physics B-Proceedings Supplements*, 191:320–330, 2009.

[31] Volker Schomerus. Strings for quantum chromodynamics. *International Journal of Modern Physics A*, 22(30):5561–5571, 2007.

[32] GW Gibbons. Supersymmetry, supergravity and related topics, 1985.

[33] Gary W Gibbons and Kei-ichi Maeda. Black holes and membranes in higher-dimensional theories with dilaton fields. *Nuclear Physics B*, 298(4):741–775, 1988.

[34] David Garfinkle, Gary T Horowitz, and Andrew Strominger. Charged black holes in string theory. *Physical Review D*, 43(10):3140, 1991.

[35] Gary T Horowitz and Andrew Strominger. Black strings and P-branes. *Nuclear Physics B*, 360(1):197–209, 1991.

[36] Joseph Gerard Polchinski. *String theory, volume I: An introduction to the bosonic string.* Cambridge University Press, Cambridge, 1998.

[37] Edward Witten. Bound states of strings and p-branes. *Nuclear Physics B*, 460(2):335–350, 1996.

[38] Igor R Klebanov. World-volume approach to absorption by non-dilatonic branes. *Nuclear Physics B*, 496(1-2):231–242, 1997.

[39] Steven S Gubser, Igor R Klebanov, and Arkady A Tseytlin. String theory and classical absorption by three-branes. *Nuclear Physics B*, 499(1-2):217–240, 1997.

[40] Steven S Gubser and Igor R Klebanov. Absorption by branes and Schwinger terms in the world volume theory. *Physics Letters B*, 413(1-2):41–48, 1997.

[41] Juan Maldacena. The gauge/gravity duality. *arXiv preprint arXiv:1106.6073*, 2011.

[42] Hisham Sati and Urs Schreiber. Anyonic Topological Order in Twisted Equivariant Differential (TED) K-theory. *arXiv preprint arXiv:2206.13563*, 2022.

[43] Carlos Alfonso Bayona and Nelson RF Braga. Anti-de Sitter boundary in Poincaré coordinates. *General Relativity and Gravitation*, 39(9):1367–1379, 2007.

[44] Alexandre Belin, Jan De Boer, and Jorrit Kruthoff. Comments on a state-operator correspondence for the torus. *SciPost Physics*, 5(6):060, 2018.

[45] Toby Wiseman. Lectures on AdS/CFT. *Special Topics Lectures, Imperial College London*, 2022.

[46] MJ Duff. TASI lectures on Branes, Black Holes and Anti-de Sitter space. In *Strings, Branes And Gravity*, pages 3–125. World Scientific, 2001.

[47] Miguel Socolovsky. Schwarzschild Black hole in Anti-de Sitter space. *arXiv preprint arXiv:1711.02744*, 2017.

[48] Peng Zhao. Black holes in Anti-de Sitter spacetime. In *Lent term Part III Seminar Series, Essay*, 2008.

[49] Stephen Hawking, Juan Maldacena, and Andrew Strominger. De Sitter entropy, quantum entanglement and AdS/CFT. *Journal of High Energy Physics*, 2001(05):001, 2001.

[50] Juan Maldacena. Eternal black holes in Anti-de Sitter. *Journal of High Energy Physics*, 2003(04):021, 2003.

[51] Tatsuma Nishioka, Shinsei Ryu, and Tadashi Takayanagi. Holographic entanglement entropy: an overview. *Journal of Physics A: Mathematical and Theoretical*, 42(50):504008, 2009.

[52] Raphael Bousso. A covariant entropy conjecture. *Journal of High Energy Physics*, 1999(07):004, 1999.

[53] Raphael Bousso. Holography in general space-times. *Journal of High Energy Physics*, 1999(06):028, 1999.

[54] Pasquale Calabrese and John Cardy. Entanglement entropy and quantum field theory. *Journal of statistical mechanics: theory and experiment*, 2004(06):P06002, 2004.

[55] Christoph Holzhey, Finn Larsen, and Frank Wilczek. Geometric and renormalized entropy in conformal field theory. *Nuclear Physics B*, 424(3):443–467, 1994.

[56] H Casini and M Huerta. Analytic results on the geometric entropy for free fields. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(01):P01012, 2008.

[57] H Casini and M Huerta. Entanglement and alpha entropies for a massive scalar field in two dimensions. *Journal of Statistical Mechanics: Theory and Experiment*, 2005(12):P12012, 2005.

[58] Shinsei Ryu and Tadashi Takayanagi. Aspects of holographic entanglement entropy. *Journal of High Energy Physics*, 2006(08):045, 2006.

[59] Dmitri V Fursaev. Proof of the holographic formula for entanglement entropy. *Journal of High Energy Physics*, 2006(09):018, 2006.

[60] Pasquale Calabrese and John Cardy. Entanglement entropy and quantum field theory: a non-technical introduction. *International Journal of Quantum Information*, 4(03):429–438, 2006.

[61] J David Brown and Marc Henneaux. Central charges in the canonical realization of asymptotic symmetries: an example from three dimensional gravity. *Communications in Mathematical Physics*, 104(2):207–226, 1986.

[62] Maximo Banados, Claudio Teitelboim, and Jorge Zanelli. Black hole in three-dimensional spacetime. *Physical Review Letters*, 69(13):1849, 1992.

[63] Igor R Klebanov and Matthew J Strassler. Supergravity and a confining gauge theory: Duality cascades and χSB-resolution of naked singularities. *Journal of High Energy Physics*, 2000(08):052, 2000.

[64] Thomas Faulkner, Aitor Lewkowycz, and Juan Maldacena. Quantum corrections to holographic entanglement entropy. *Journal of High Energy Physics*, 2013(11):1–18, 2013.

[65] Aitor Lewkowycz and Juan Maldacena. Generalized gravitational entropy. *Journal of High Energy Physics*, 2013(8):1–29, 2013.

[66] Claudio Teitelboim. Gravitation and hamiltonian structure in two spacetime dimensions. *Physics Letters B*, 126(1-2):41–45, 1983.

[67] Roman Jackiw. Lower dimensional gravity. *Nuclear Physics B*, 252:343–356, 1985.

[68] Lisa Randall and Raman Sundrum. An alternative to compactification. *Physical Review Letters*, 83(23):4690, 1999.

[69] Maximo Banados. Three-dimensional quantum geometry and black holes. In *AIP Conference Proceedings*, volume 484, pages 147–169. American Institute of Physics, 1999.

[70] John L Cardy. Boundary conditions, fusion rules and the Verlinde formula. *Nuclear Physics B*, 324(3):581–596, 1989.

[71] Hong Zhe Chen, Zachary Fisher, Juan Hernandez, Robert C Myers, and Shan-Ming Ruan. Evaporating black holes coupled to a thermal bath. *Journal of High Energy Physics*, 2021(1):1–71, 2021.

[72] Fririk Freyr Gautason, Lukas Schneiderbauer, Watse Sybesma, and Lárus Thorlacius. Page curve for an evaporating black hole. *Journal of High Energy Physics*, 2020(5):1–23, 2020.

[73] Ahmed Almheiri, Raghu Mahajan, and Jorge Santos. Entanglement islands in higher dimensions. *SciPost Physics*, 9(1):001, 2020.

[74] Ahmed Almheiri, Thomas Hartman, Juan Maldacena, Edgar Shaghoulian, and Amirhossein Tajdini. Replica wormholes and the entropy of Hawking radiation. *Journal of High Energy Physics*, 2020(5):1–42, 2020.

[75] Kanato Goto, Thomas Hartman, and Amirhossein Tajdini. Replica wormholes for an evaporating 2d black hole. *Journal of High Energy Physics*, 2021(4):1–57, 2021.