Islands and Uhlmann phase: explicit recovery of classical information from evaporating black holes

Josh Kirklin

Qubits and Spacetime Unit, Okinawa Institute of Science and Technology,
1919-1 Tancha, Onna-son, Kunigami-gun, Okinawa, 904-0495, Japan

E-mail: joshua.kirklin@oist.jp

ABSTRACT: Recent work has established a route towards the semiclassical validity of the Page curve, and so provided evidence that information escapes an evaporating black hole. However, a protocol to explicitly recover and make practical use of that information in the classical limit has not yet been given. In this paper, we describe such a protocol, showing that an observer may reconstruct the phase space of the black hole interior by measuring the Uhlmann phase of the Hawking radiation. The process of black hole formation and evaporation provides an invertible map between this phase space and the space of initial matter configurations. Thus, all classical information is explicitly recovered. We assume in this paper that replica wormholes contribute to the gravitational path integral.

Contents

1 Introduction

Classical black holes are airtight, but semiclassical ones are leaky, gradually losing energy over time via the Hawking process [1, 2]. In the absence of any infalling matter, they will diminish in size until they reach the Planck scale, at which point higher order quantum gravitational effects take over. Before that point, Hawking's original calculations suggest that the entropy of the radiation emitted by the black hole will steadily grow [3]. But for the combined process of black hole formation and evaporation to be unitary, the entropy must instead follow the Page curve — at a certain point (the Page time) it must stop growing and start decreasing, eventually becoming very small after the black hole is gone [4, 5]. The Page time takes place well before the black hole reaches the Planck scale, so there is tension between Hawking's picture and unitarity. If we believe quantum gravity is unitary,[1] then this constitutes a paradox.

The resolution of this issue has remained beyond our reach for nearly half a century, but a flurry of recent progress indicates that we may be getting closer [8–13]. Apparently, we have been computing the entropy S of the Hawking radiation in the wrong way. We should have been using the 'island formula':

$$S = \min \left\{ \operatorname*{ext}_{I} \left[\frac{\operatorname{Area}(\partial I)}{4G} + S_{\text{bulk}}(I \cup R) \right] \right\}. \tag{1.1}$$

---

[1]Quantum gravity does not have to be unitary, but most of the theoretical physics community seems to believe it is at the moment. One reason for this is the existence of AdS/CFT [6, 7], which shows that unitary theories of quantum gravity exist by construction.

R is a region of space containing all of the Hawking radiation, while I is a separate region called the 'island'. Area(∂I) is the area of the boundary of the island,[2] G is Newton's constant, and $S_{\text{bulk}}(I \cup R)$ is the so-called 'bulk entropy' of the combined region $I \cup R$, i.e. the entropy of perturbative QFT fluctuations (including gravitons and any other field excitations) on top of a fixed background geometry. According to (1.1), we need to find the set of extrema of the functional Area(∂I)/$4G + S_{\text{bulk}}(I \cup R)$ with respect to choices of the island I. The entropy S is then supposed to be given by the minimum value in this set.

Consider the island I for which the minimum extremum is attained.[3] At first, this I is empty, and S reduces to the bulk entropy $S_{\text{bulk}}(R)$ of the Hawking radiation R. However, at the Page time a non-empty island appears in the black hole, and during the subsequent evaporation, I continues to grow in size until it spans almost the entire black hole interior. The papers [8–13] argue that accounting for this in the island formula reproduces the Page curve, thus removing the tension between black hole evaporation and unitarity in quantum gravity. So, in principle, any information inside the black hole eventually escapes and becomes accessible to an external observer.

Despite this progress, it has not been made clear exactly what form the escaped information takes in the classical limit, or where a classical observer should look for it. The problem is that the information has been heavily encrypted by the black hole formation and evaporation process, and no protocol for decrypting it has yet been firmly established. In this paper, we will describe such a protocol.

In particular, we will explain how an observer can use the Hawking radiation to reconstruct the classical phase space, including the symplectic form and therefore all the Poisson brackets, of the fields located in the island I. Thus, at the Page time, the observer starts to learn about the black hole interior. If the observer waits a sufficiently long time, the island will grow large enough that all classical degrees of freedom in the interior of the black hole are accounted for in the island phase space. Then the dynamical map $\mathscr{S}$ giving the evolution from the initial phase space of classical collapsing matter configurations to the island phase space is invertible. The structure of a phase space determines the full set of measurements a classical observer can make, so the invertibility of $\mathscr{S}$ means that a late-time classical observer and an early-time classical observer have access to the exact same set of measurements. In this way all the classical information is explicitly recovered.

The entropy formula (1.1) is inspired by similar formulae in AdS/CFT and holography [16–20], but it is supposed to hold in a general theory of quantum gravity. Similarly, the protocol presented here takes inspiration from the recent identification of the holographic dual of the bulk symplectic form [21–23]. Suppose we are given a closed curve of holographic states $|\lambda\rangle$ in AdS/CFT, and a boundary subregion of interest A. By tracing over degrees of freedom in the complement of A, one obtains a closed curve of reduced states ρ_A in A. For any such curve, one can compute a number γ called the Uhlmann

---

[2]This area term is appropriate for general relativity. In higher derivative theories of gravity, it is replaced by some other functional of the geometry of ∂I [14, 15].

[3]We will interchangably use I to refer to either the free variable over which the extremisation is carried out in (1.1), or the island at the minimum extremum. In each instance, it should be clear from the context to which we are referring.

phase [24–30]. This number depends *only* on the curve of reduced states ρ_A — we will explain later the precise way in which it is defined. The key result of [23] was the formula

$$\gamma = \int_C \Theta, \tag{1.2}$$

where C is the curve of bulk field configurations dual to the states $|\lambda\rangle$, and $\Omega = \mathrm{d}\Theta$ is the symplectic form of the entanglement wedge of A. It is convenient to view Θ as a connection, and Ω as the curvature of this connection. Thus, the entanglement wedge symplectic form is holographically dual to the curvature of the boundary Uhlmann phase.

The analogue of the entanglement wedge for an evaporating black hole is the domain of dependence of the island I, so one might expect that by measuring the curvature of the Uhlmann phase of the Hawking radiation one can reproduce the symplectic form for the fields on the island. This would imply that the space of possible Hawking radiation states is a symplectic manifold, isomorphic to the island phase space. The main point of the present work is to confirm that this expectation is correct by bringing together the insights that led to (1.1) and (1.2). Both of the derivations for these formulae make use the same kind of mathematical tool — namely the replica trick in a gravitational path integral. To get (1.1), one has to account for so-called 'replica wormholes' in the path integral. Our results will follow from similarly accounting for these wormholes. There has been a lot of discussion on whether or not including them is the right thing to do [31–33], but we will just assume that it is and make no further comments.

To summarise, we will show that the curvature of the Hawking radiation Uhlmann phase is equivalent to the symplectic form of the island I, and that at late times semi-classical evolution provides an invertible map $\mathscr{S}$ between the island phase space and the initial phase space. Together these make up a complete decryption protocol for recovering classical information. The rest of the paper starts in section 2 with a description of the set-up, and a brief review of how the contributions of replica wormholes in the replica trick lead to the island formula (1.1). Then, in section 3, we define the Uhlmann phase, and explain how to compute it for Hawking radiation using an extension of the replica trick. In section 4, we show how to use this result to reconstruct the island phase space, and how to use the reconstruction to recover classical information from the black hole interior. We finish in section 5 with a short discussion of the broader consequences of our results.

2 Evaporating black hole and islands

The set-up we consider consists of a spacetime divided into a gravitational region and an asymptotic region, which are coupled together at a 'cutoff surface'. We assume that gravitational effects only play a role in the gravitational region. This could be an exact assumption, so that the theories in the two regions are genuinely different, or it could just be an approximation. We form a black hole in the gravitational region, and then allow the asymptotic region to absorb the resulting Hawking radiation, until the black hole evaporates. This is illustrated in figure 1.[4] The spacetime depicted lies between two

---

[4]Our figures will depict spacetime as if it is asymptotically flat, but the precise behaviour of the fields in the asymptotic region is not particularly important for our purposes.
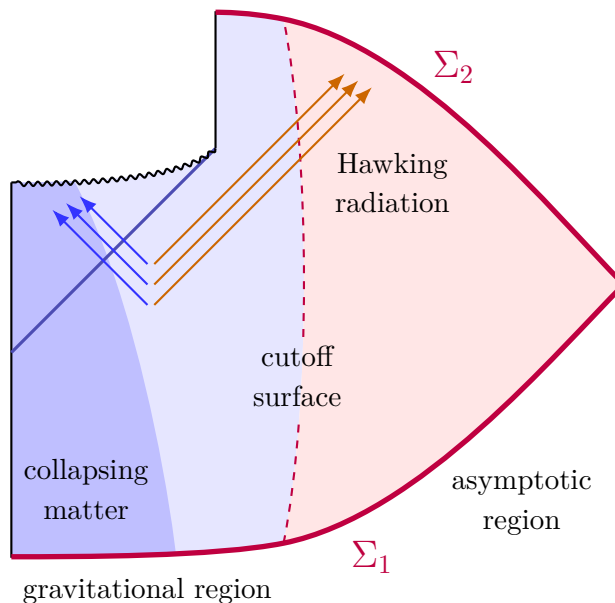
Figure 1. Spacetime is divided into asymptotic and gravitational regions, joined by a cutoff surface. Initially, the system is set up with a matter configuration that collapses to form a black hole. Then, the black hole evaporates, and all of the Hawking radiation is collected by the asymptotic region. (Note that we will continue to use this color scheme, in which the asymptotic region is red and the gravitational region is blue, throughout the paper.)

spacelike surfaces Σ_1 and Σ_2. We would like to think of them as Cauchy surfaces on which the initial and final states live, and indeed Σ_1 is a genuine Cauchy surface. But Σ_2 is not, because its domain of dependence does not include a large chunk of the spacetime, including the black hole region and the collapsing matter. As a consequence, the map representing the evolution from Σ_1 to Σ_2 is not invertible, so from a naive perspective classical information appears to have been lost down the black hole. This is the problem we intend to solve in this paper.

Let $\mathcal{H}_G, \mathcal{H}_R$ denote the Hilbert spaces of the gravitational region and the asymptotic region respectively, so that the total Hilbert space of the full system is $\mathcal{H}_G \otimes \mathcal{H}_R$.[5] During the evolution of the system, the Hawking radiation is collected in the part of the state in $\mathcal{H}_R$. For this reason, one can think of $\mathcal{H}_R$ as actually being the Hilbert space of the Hawking radiation itself.

---

[5]This factorisation comes with two caveats. First, the algebra of observables of a subregion in a continuum QFT or gravitational theory is believed to generically be of Von Neumann type III [34]. Such algebras cannot be associated with tensor factors of the full Hilbert space. In this paper, we will assume that the theory under consideration has been regularised (perhaps on a lattice), so that the subsystems are no longer of type III. Second, even after such regularisation, the presence of gauge symmetries and edge modes means that the total physical Hilbert space is not a tensor product, but rather a direct sum of tensor products. Again, we will assume that the theory has been sufficiently regularised that we can isometrically embed the physical Hilbert space inside a 'kinematic' Hilbert space of the desired form $\mathcal{H}_G \otimes \mathcal{H}_R$. We expect that our results are regularisation-independent, but it would be interesting and important to check this directly. Such a check is beyond the scope of the present paper, and would require an algebraic definition of Uhlmann phase.
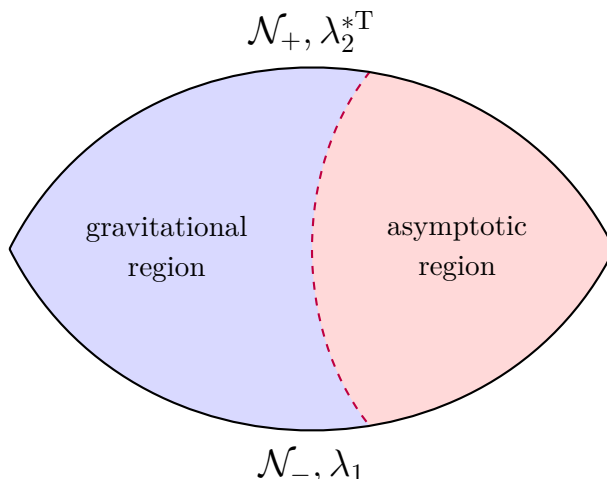
Figure 2. The manifold and fields for the partition function $Z(\lambda_2, \lambda_1)$ must obey the boundary conditions λ_1 in the past $\mathcal{N}_-$, and the complex conjugate and time reversal of λ_2 in the future $\mathcal{N}_+$.

2.1 States and partition functions

The states $|\lambda\rangle$ in $\mathcal{H}_G \otimes \mathcal{H}_R$ which we will consider essentially amount to the specification of a set of boundary conditions λ for the fields on a surface. The inner product of such states is given by

$$\langle \lambda_2 | \lambda_1 \rangle = \frac{Z(\lambda_2, \lambda_1)}{\sqrt{Z(\lambda_2, \lambda_2) Z(\lambda_1, \lambda_1)}}, \tag{2.1}$$

where $Z(\lambda_2, \lambda_1)$ is a Euclidean partition function. The denominator ensures that $|\lambda\rangle$ is normalised. In a semiclassical limit for the gravitational fields, we can approximate the partition function with

$$-\log Z(\lambda_2, \lambda_1) = I_{\mathrm{grav}}[\mathcal{M}, g] - \log Z_{\mathrm{QFT}}[\mathcal{M}, g]. \tag{2.2}$$

Here, $\mathcal{M}$ is a manifold with metric g. I_{grav} is the effective gravitational action of $\mathcal{M}, g$, and Z_{QFT} is a QFT partition function for fields on top of $\mathcal{M}$. The boundary $\mathcal{N} = \partial \mathcal{M}$ splits into two pieces: the 'past' boundary $\mathcal{N}_-$, and the 'future' boundary $\mathcal{N}_+$. All of the fields, including the metric g, have to obey the boundary conditions λ_1 on $\mathcal{N}_-$ and $\lambda_2^{*\mathrm{T}}$ on $\mathcal{N}_+$, where $\lambda_2^{*\mathrm{T}}$ denotes the complex conjugate and time reversal of λ_2. This is shown in figure 2. Because we are doing a saddlepoint calculation, the right hand side is supposed to be evaluated on the $\mathcal{M}, g$ for which it is minimised.

Taking a semiclassical limit for the QFT fields as well, we can write

$$-\log Z_{\mathrm{QFT}}[\mathcal{M}, g] = I_{\mathrm{fields}}[\mathcal{M}, \phi], \tag{2.3}$$

where I_{fields} is the effective action for the fields ϕ (which include the metric g). The full partition function is then given by

$$-\log Z(\lambda_2, \lambda_1) = I[\mathcal{M}, \phi] = I_{\mathrm{grav}}[\mathcal{M}, g] + I_{\mathrm{fields}}[\mathcal{M}, \phi], \tag{2.4}$$

where the total effective action $I[\mathcal{M}, \phi]$ is evaluated on the $\mathcal{M}, \phi$ for which it is minimised, subject to the appropriate boundary conditions. We call this minimal $\mathcal{M}, \phi$ the 'on-shell' configuration.

At leading order in the semiclassical approximation, we assume that the action is local, i.e. it may be written as

$$I[\mathcal{M}, \phi] = \int_{\mathcal{M}} L[\phi], \tag{2.5}$$

where the Lagrangian density $L[\phi]$ is a top form that depends locally on the fields ϕ. Under a variation of the fields $\phi \to \phi + \delta\phi$, the change in the Lagrangian density may in general be written at linear order in $\delta\phi$ as

$$\delta L[\phi] = E[\phi] \cdot \delta\phi + \mathrm{d}(\theta[\phi, \delta\phi]), \tag{2.6}$$

where $E[\phi] = 0$ are the equations of motion, and $\cdot$ denotes a sum over all the fields. If we vary the fields from an on-shell configuration ϕ to an off-shell one $\phi + \delta\phi$, the action therefore changes by

$$\delta I = \int_{\mathcal{N}} \theta[\phi, \delta\phi]. \tag{2.7}$$

The consistency of the saddlepoint approximation requires the variational principle to be well defined. In other words, if the field variation $\delta\phi$ does not change the boundary conditions λ, then we must have $\delta I = 0$, so that ϕ genuinely extremises the action. This implies that the right hand side of (2.7) can only depend on $\delta\phi$ through changes $\delta\lambda_1, \delta\lambda_2$ in the boundary conditions.

In the case $\lambda_1 = \lambda_2$, the boundary conditions for the partition function have a $\mathbb{Z}_2$ symmetry under which they are complex conjugated, and reflected in time such that the past boundary is exchanged with the future boundary. Assuming this symmetry is not broken on-shell, it will fix a codimension 1 surface Σ in $\mathcal{M}$. On this surface, all fields are invariant under the combination of time reversal and complex conjugation, and so can be Wick rotated to real Lorentzian fields. The Lorentzian development of those fields is what was depicted in figure 1. The black hole 'lives' in this Lorentzian spacetime, even though we are using Euclidean partition functions to prepare the state.

The surface Σ divides $\mathcal{M}$ into a past $\mathcal{M}_-$ and a future $\mathcal{M}_+$, such that $\partial\mathcal{M}_\pm = \Sigma \cup \mathcal{N}_\pm$, as shown in figure 3. If we fix some boundary conditions φ for the fields on Σ, then we can write the on-shell action as
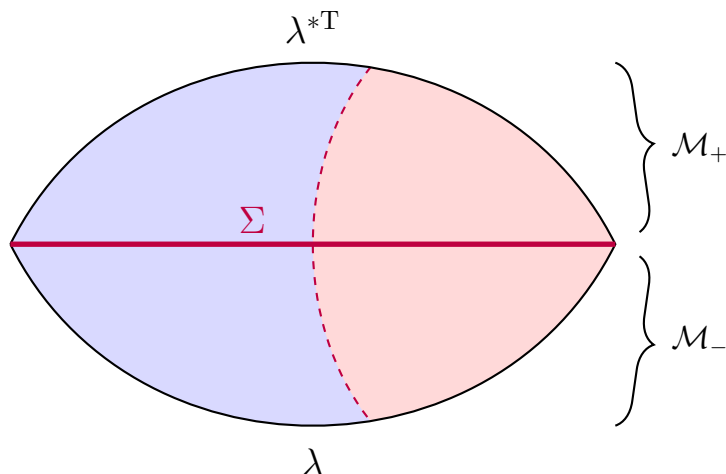
$$I[\mathcal{M}, \phi] = \min_{\varphi} \left(I[\mathcal{M}_-, \phi_-] + I[\mathcal{M}_+, \phi_+] \right), \tag{2.8}$$

where ϕ_-, ϕ_+ are the fields in the past and future respectively, and

$$I[\mathcal{M}_-, \phi_-] = \int_{\mathcal{M}_-} L[\phi_-] + \int_{\Sigma} D[\phi_-], \tag{2.9}$$

$$I[\mathcal{M}_+, \phi_+] = \int_{\mathcal{M}_+} L[\phi_+] - \int_{\Sigma} D[\phi_+] \tag{2.10}$$

are evaluated at their minima. Here, $D[\phi]$ only depends on the fields at Σ. The choice of $D[\phi]$ directly corresponds to the class of boundary conditions φ. To see this, suppose the

Figure 3. When $\lambda_1 = \lambda_2$, all the fields are invariant under time reflection and complex conjugation. This symmetry fixes a surface Σ which divides $\mathcal{M}$ into a past piece $\mathcal{M}_-$ and a future piece $\mathcal{M}_+$.

past and future boundary conditions λ are fixed. Then, under a variation of the fields from an on-shell configuration, we have

$$\delta I[\mathcal{M}_-, \phi_-] = \int_\Sigma \left(\theta[\phi_-, \delta\phi_-] + \delta D[\phi_-] \right), \tag{2.11}$$

$$\delta I[\mathcal{M}_+, \phi_+] = -\int_\Sigma \left(\theta[\phi_+, \delta\phi_+] - \delta D[\phi_+] \right). \tag{2.12}$$

For the variational principle to be well-defined, the right hand sides above should vanish if we keep φ fixed. Thus they can only depend on $\delta\phi_\pm$ via $\delta\varphi$. Different choices of $D[\phi]$ modify the dependence on $\delta\phi$, and so specify different classes of boundary conditions φ. Let us absorb $D[\phi]$ into the Lagrangian, i.e. redefine $L \to L + \mathrm{d}D$.

In terms of the quantum theory, what we have just done amounts to the insertion of a resolution of the identity

$$I = \int \mathrm{D}\varphi \, |\varphi\rangle \, \langle\varphi| \tag{2.13}$$

at Σ. The minimisation over φ in (2.8) then comes from a saddlepoint approximation in the $\mathrm{D}\varphi$ integral. The states $|\varphi\rangle$ specify the boundary conditions at Σ. The time reflection and complex conjugation invariance of the fields at Σ ensures that these states $|\varphi\rangle$ are genuine states in $\mathcal{H}$.

2.2 Entropy from a replica trick

Given a state $|\lambda\rangle \in \mathcal{H}_G \otimes \mathcal{H}_R$, we can compute the reduced state ρ_R of the Hawking radiation by taking a partial trace over $\mathcal{H}_G$. The entropy of the Hawking radiation is then given by the von Neumann entropy of ρ_R:

$$S = -\operatorname{tr}(\rho_R(\lambda) \log \rho_R(\lambda)) \quad \text{where} \quad \rho_R(\lambda) = \operatorname{tr}_G |\lambda\rangle \langle\lambda| . \tag{2.14}$$

According to [8–13], S is given by the island formula (1.1). At early times in the evaporation of a black hole, there is no island, but after the Page time an island appears in the black hole region, as shown in figure 4.

(a) $t < t_{\text{Page}}$ **(b)** $t > t_{\text{Page}}$
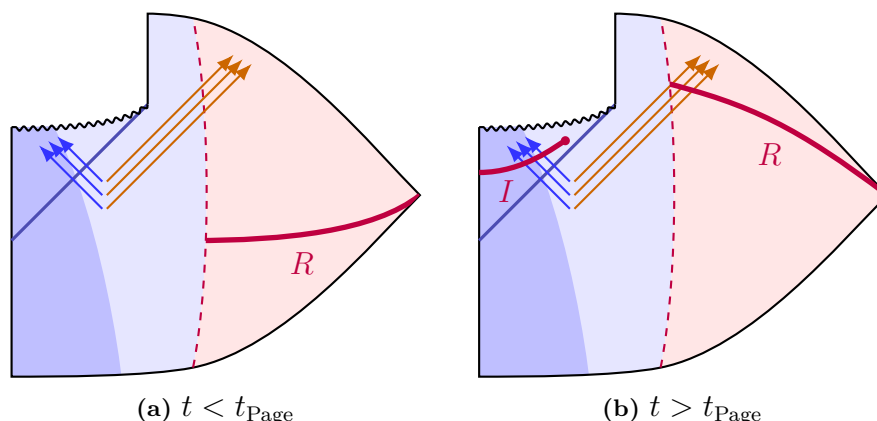
Figure 4. **(a)** Before the Page time, there is no island, and the entropy of the Hawking radiation is just given by the bulk entropy on the surface R. **(b)** After the Page time, enough entanglement has been built up by the Hawking process for an island to appear inside the black hole. Accounting for this in the island formula reproduces the Page curve.

Let us now give an indication of where the island formula comes from.

To compute the inner product $\langle \lambda_2 | \lambda_1 \rangle$ we summed over degrees of freedom in both the asymptotic and gravitational regions — colloquially, this is called 'gluing'. On the other hand, the partial trace in $\rho_R(\lambda) = \text{tr}_G |\lambda\rangle\langle\lambda|$ involves just a sum over the fields in the gravitational region. The composition of several reduced states

$$\rho_R(\lambda_n)\rho_R(\lambda_{n-1})\ldots\rho_R(\lambda_2)\rho_R(\lambda_1) \tag{2.15}$$

involves then gluing the future asymptotic region of $\rho_R(\lambda_i)$ to the past asymptotic region of $\rho_R(\lambda_{i+1})$ for $i = 1, \ldots, n-1$. By taking a trace, we then glue the future asymptotic region of $\rho_R(\lambda_n)$ back to the past asymptotic region of $\rho_R(\lambda_1)$, and the result may be written in terms of partition functions:

$$\text{tr}\left(\rho_R(\lambda_n)\ldots\rho_R(\lambda_1)\right) = \frac{Z_{(n)}[\lambda_n, \ldots, \lambda_1]}{Z_{(1)}[\lambda_n]\ldots Z_{(1)}[\lambda_1]}. \tag{2.16}$$

Here $Z_{(1)}[\lambda] = Z(\lambda, \lambda)$, and $Z_{(n)}[\lambda_n, \ldots, \lambda_1]$ is the 'replica' partition function for the boundary conditions just described. It may also be computed with a semiclassical approximation, using the effective gravitational action. We have

$$-\log Z_{(n)}[\lambda_n, \ldots, \lambda_1] = I_{\text{grav}}[\widetilde{\mathcal{M}}_n, g] - \log Z_{\text{QFT}}[\widetilde{\mathcal{M}}_n, g], \tag{2.17}$$

where $\widetilde{\mathcal{M}}_n$ is a manifold with metric g obeying the appropriate boundary conditions, and the right hand side is evaluated at the $\widetilde{\mathcal{M}}_n$ for which it is minimised. In the asymptotic region, the topology of the n copies is fixed. However, in the gravitational region, we should allow $\widetilde{\mathcal{M}}_n$ to have any topology that fills in the boundary conditions. In particular, this means that there can be 'replica wormholes' in the gravitational region that connect the n copies. This is demonstrated in figure 5.

Before we take the semiclassical limit for the QFT fields as well, we have to insert so-called twist operators $\mathcal{T}_n \ldots \mathcal{T}_1$, which ensure that the various copies of the QFT are
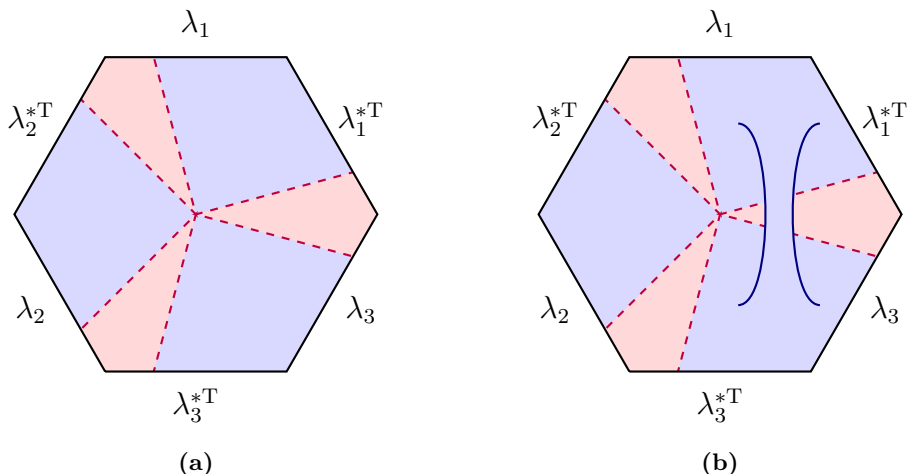
Figure 5. The partition function $Z_{(n)}[\lambda_n, \ldots, \lambda_1]$ should be computed on manifolds $\widetilde{\mathcal{M}}_n$ with the boundary conditions shown (the case shown is for $n = 3$). The n gravitational regions are disconnected in **(a)**, and this is what we would do in an ordinary QFT. However, since this is a theory of gravity, we should allow all topologies in the gravitational regions, including those which connect various copies with wormholes, for example like in **(b)**.

connected in the right way [35, 36]. The semiclassical limit of Z_{QFT} then includes the expectation value of the twist operators, as well as the effective action:

$$- \log Z_{\text{QFT}}[\widetilde{\mathcal{M}}_n, g] = I_{\text{fields}}[\widetilde{\mathcal{M}}_n, \phi] - \log \langle \mathcal{T}_n \ldots \mathcal{T}_1 \rangle. \tag{2.18}$$

Overall, in the full semiclassical limit, the replica partition function is given by

$$- \log Z_{(n)}[\lambda_n, \ldots, \lambda_1] = I[\widetilde{\mathcal{M}}_n, \phi] - \log \langle \mathcal{T}_n \ldots \mathcal{T}_1 \rangle. \tag{2.19}$$

To compute the von Neumann entropy of $\rho_R = \rho_R(\lambda)$ we can use the 'replica trick',

$$S = - \text{tr}(\rho_R \log \rho_R) = - \partial_n \text{tr}((\rho_R)^n)|_{n=1}. \tag{2.20}$$

The aim is to compute $\text{tr}((\rho_R)^n)$ for integer n in terms of the replica partition function. Then we analytically continue in n in order to compute the n derivative. Assuming we have done this, we can write

$$S = - \partial_n \left(\frac{\log Z_{(n)}}{n} \right) \Big|_{n=1}, \tag{2.21}$$

where $Z_{(n)} = Z_{(n)}[\lambda, \lambda, \ldots, \lambda]$. The boundary conditions for this partition function have a $\mathbb{Z}_n$ replica symmetry which cyclically permutes the n copies of the theory. Assuming this symmetry is not spontaneously broken, it will also be obeyed by the on-shell configuration $\widetilde{\mathcal{M}}_n, \phi$, as shown in figure 6. Then we can define the quotient manifold $\mathcal{M}_n = \widetilde{\mathcal{M}}_n / \mathbb{Z}_n$, and the field configuration ϕ reduces unambiguously to $\mathcal{M}_n$. In $\widetilde{\mathcal{M}}_n$ there is a codimension 2 surface $\widetilde{\Upsilon}$ which is fixed by $\mathbb{Z}_n$, and in $\mathcal{M}_n$ this becomes a surface Υ with a conical defect of opening angle $2\pi/n$. Then we can write

$$I[\widetilde{\mathcal{M}}_n, \phi] = I[\widetilde{\mathcal{M}}_n \setminus \widetilde{\Upsilon}, \phi] = nI[\mathcal{M}_n \setminus \Upsilon, \phi] = n\left(I[\mathcal{M}_n, \phi] + (n-1)\frac{\text{Area}(\Upsilon)}{4G} \right). \tag{2.22}$$
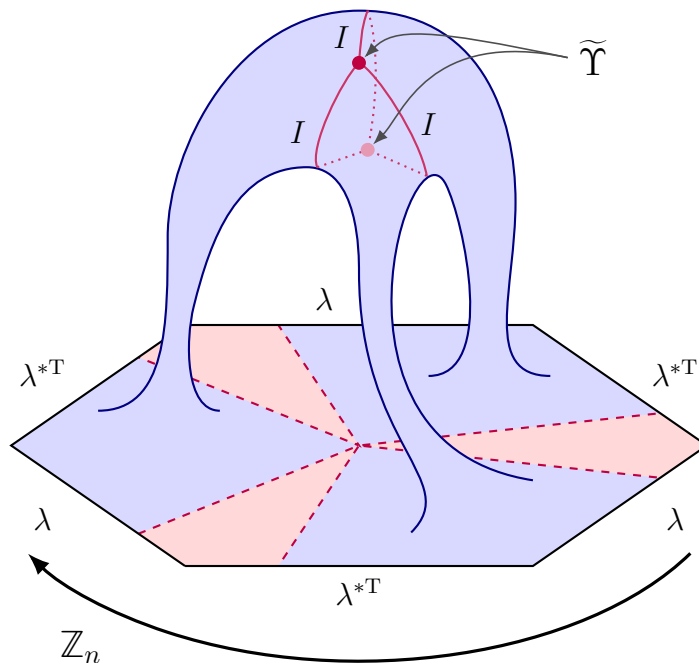
Figure 6. The on-shell configuration for $Z_{(n)} = Z_{(n)}[\lambda, \ldots, \lambda]$ has a $\mathbb{Z}_n$ replica symmetry. In general there will be a codimension 2 surface $\widetilde{\Upsilon}$ in the gravitational region which is fixed by $\mathbb{Z}_n$, here shown as two points. $\widetilde{\Upsilon}$ is the boundary of n copies of a codimension 1 surface I, which are cyclically permuted by $\mathbb{Z}_n$.

The first equality follows from the assumption that the fields in $\widetilde{\mathcal{M}}_n$ are regular near $\widetilde{\Upsilon}$, while the second comes from replica symmetry. The area term on the right hand side cancels out any contributions that come from integrating over the conical defect Υ in the action.[6] We can quite easily analytically continue this action in n, because all the dependence on n comes in the strength of the conical defect, which can be varied continuously.

Note that the fixed surface $\widetilde{\Upsilon}$ is the boundary of n copies of a region I in $\widetilde{\mathcal{M}}_n$ which are cyclically permuted by replica symmetry. These become a single region I in $\mathcal{M}_n$. This I is the island. The overall structure of $\mathcal{M}_n$ is shown in figure 7.

So the entropy may be written

$$S = -\partial_n \left(I[\mathcal{M}_n, \phi] + (n-1)\frac{\text{Area}(\Upsilon)}{4G} + \log \langle \mathcal{T}_n \ldots \mathcal{T}_1 \rangle \right)\Big|_{n=1}. \tag{2.23}$$

Note that at $n = 1$, there is no conical defect in $\mathcal{M}_n$. In fact, at $n = 1$, $\mathcal{M}_n, \phi$ is the exact same on-shell configuration $\mathcal{M}, \phi$ relevant to $\mathcal{Z}_{(1)}[\lambda]$. Computing the derivative in n of $I[\mathcal{M}_n, \phi]$ at $n = 1$ amounts to finding a linear perturbation of an on-shell action. But since the boundary conditions are fixed, by definition the action does not change at linear order, because we are on-shell. So the first term above vanishes. The second term is simple to compute. The third term is what we would get if we were computing an entropy

---

[6]This is the appropriate form for this term in general relativity. In higher derivative theories of gravity the area should be replaced by some other functional of the geometry of Υ [14, 15].
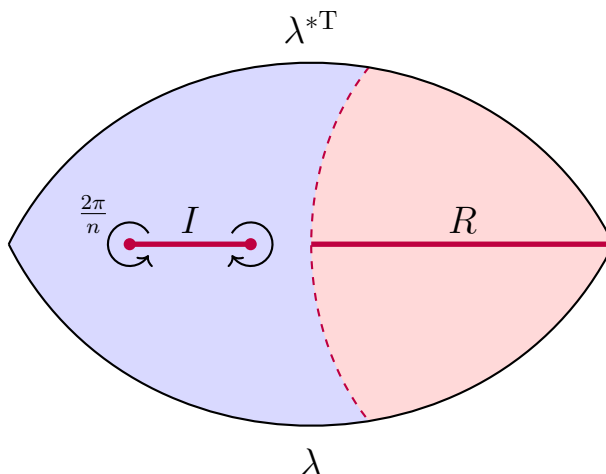
Figure 7. After quotienting by the $\mathbb{Z}_n$ replica symmetry, we get a manifold $\mathcal{M}_n$ with a single copy of the boundary conditions. There is a conical defect at $\Upsilon = \partial I$, and we insert twist operators on $I \cup R$.

using the replica trick in a ordinary perturbative QFT. The twist operators act on both the asymptotic region R *and* the island I. Thus, what the third term is computing is the bulk entropy of the combined region $I \cup R$. Combining all of this, we get

$$S = \frac{\text{Area}(\partial I)}{4G} + S_{\text{bulk}}(I \cup R). \tag{2.24}$$

It remains to find the location of the island, and to do this we should note that $-\log Z_{(n)}$ should be minimised for all n. For n close to 1, we can use the series expansion

$$-\frac{1}{n}\log Z_{(n)} = I[\mathcal{M}, \phi] + (n-1)S + \mathcal{O}\big((n-1)^2\big). \tag{2.25}$$

We should minimise both terms appearing in this expansion individually. Minimising the first just means $\mathcal{M}, \phi$ is on-shell. Note that we should take $n \geq 1$, because this is the regime in which the replica trick works (according to Carlson's theorem). So minimising the second term is the same as minimising S. Therefore, we should compute the entropy S using the island I for which it is minimised, which is exactly the claim of the island formula.

3 Uhlmann phase of Hawking radiation

The quantity we now wish to compute is called the Uhlmann phase. This is not a widely known object, so let us briefly review how it is defined. For a more extensive description see [24–30], or for the holographic context see [23]. The calculation that appears in the present work is essentially a more streamlined and generalised version of the one in that paper.

Given a curve of density matrices $\rho(t)$, $0 \leq t \leq 1$ acting on a Hilbert space $\mathcal{H}$, we can always pick an auxiliary Hilbert space $\mathcal{H}'$, and find a curve of pure states $|\psi(t)\rangle$ in $\mathcal{H} \otimes \mathcal{H}'$

such that

$$\rho(t) = \mathrm{tr}' \, |\psi(t)\rangle \, \langle\psi(t)| \,, \tag{3.1}$$

where tr' denotes a partial trace over $\mathcal{H}'$. Such states $|\psi(t)\rangle$ are said to be purifications of $\rho(t)$. If the curve also maximises the transition probability

$$|\langle\psi(t+\epsilon)|\psi(t)\rangle|^2 \tag{3.2}$$

at leading order in small ϵ, subject to (3.1), then it is said to be 'parallel'.

If $\rho(t)$ is a closed curve of invertible density matrices, purified by $|\psi(t)\rangle$ in a parallel manner,[7] then the Uhlmann phase γ of $\rho(t)$ is defined via

$$e^{i\gamma} = \lim_{m\to\infty} \langle\psi(0)|\psi(1)\rangle \, \langle\psi(1)|\psi\Big(\frac{m-1}{m}\Big)\rangle \ldots \langle\psi\Big(\frac{2}{m}\Big)|\psi\Big(\frac{1}{m}\Big)\rangle \, \langle\psi\Big(\frac{1}{m}\Big)|\psi(0)\rangle \,. \tag{3.3}$$

For any closed curve $\rho(t)$ there are many possible choices for the parallel purifying curve $|\psi(t)\rangle$, but it may be shown that γ does not depend on this choice.

Suppose $\lambda(t)$ is a closed curve of boundary conditions, and let $\rho_R(t) = \mathrm{tr}_G \, |\lambda(t)\rangle \, \langle\lambda(t)|$ be the corresponding Hawking radiation states. Typical density matrices of subsystems in (regularised) quantum field theories and quantum gravity are invertible, and we will assume this is true for ρ_R. To compute the Uhlmann phase of $\rho_R(t)$, we need to find a parallel purifying curve $|\psi(t)\rangle$. We will use Uhlmann's theorem [28], which says that two purifications $|\psi\rangle, |\psi'\rangle$ of density matrices ρ, ρ' have maximal transition probability if and only if

$$|\langle\psi'|\psi\rangle| = \mathrm{tr}\Big(\sqrt{\sqrt{\rho}\rho'\sqrt{\rho}}\Big). \tag{3.4}$$

The right hand side is known as the fidelity of ρ and ρ'.

3.1 Fidelity and parallel states from a replica trick

To employ Uhlmann's theorem, we will first need to compute the fidelity of two Hawking radiation states

$$\rho_R = \mathrm{tr}_G \, |\lambda\rangle \, \langle\lambda| \quad \text{and} \quad \rho'_R = \mathrm{tr}_G \, |\lambda'\rangle \, \langle\lambda'| \,. \tag{3.5}$$

We can do this with a replica trick. First we compute

$$\mathrm{tr}\bigg(\Big(\sqrt{\sqrt{\rho_R}\rho'_R\sqrt{\rho_R}}\Big)^{2k}\bigg) = \mathrm{tr}\Big((\rho_R\rho'_R)^k\Big) \tag{3.6}$$

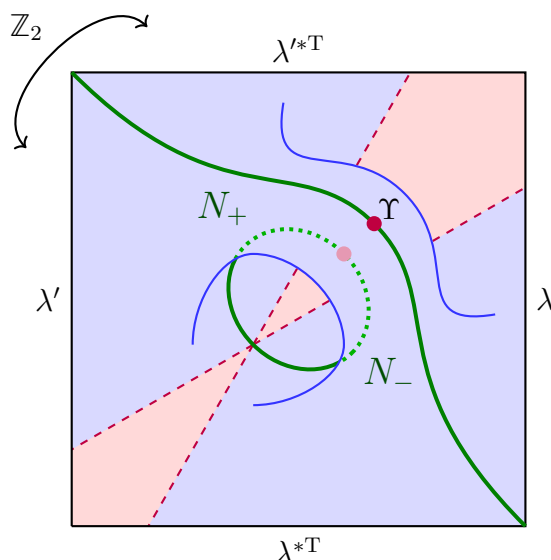for integer k, and then we analytically continue to $k = \frac{1}{2}$.

In terms of partition functions, we have

$$\mathrm{tr}\Big((\rho_R\rho'_R)^k\Big) = \frac{Z_{(2k)}[\lambda, \lambda', \ldots, \lambda, \lambda']}{Z_{(1)}[\lambda]^k Z_{(1)}[\lambda']^k} \,. \tag{3.7}$$

---

[7]Note that in general the curve of pure states $|\psi(t)\rangle$ need not be closed, even if the curve of mixed states $\rho(t)$ is.

Figure 8. The configuration relevant for the partition function $Z_{(2k)}[\lambda', \lambda, \ldots, \lambda', \lambda]$. By quotienting $\widetilde{\mathcal{M}}_{2k}$ by $\mathbb{Z}_k$, we get the manifold $\widehat{\mathcal{M}}_{2k} = \widetilde{\mathcal{M}}_{2k}/\mathbb{Z}_k$ shown here, with a conical defect at Υ of opening angle $2\pi/k$. This manifold, and the fields on it, enjoy a remaining $\mathbb{Z}_2$ symmetry that complex conjugates all the fields, and reflects them across the gravitational region in the way shown. This symmetry fixes a surface $N_+ \cup N_-$, which goes through Υ and meets the boundary of the asymptotic region.

In the partition function in the numerator, the boundary conditions are alternately given by λ and λ' on the $2k$ copies. We may write this in terms of the usual semiclassical approximation,

$$- \log Z_{(2k)}[\lambda, \lambda', \ldots, \lambda, \lambda'] = I[\widetilde{\mathcal{M}}_{2k}, \phi] - \log \langle \mathcal{T}_{2k} \ldots \mathcal{T}_1 \rangle. \qquad (3.8)$$
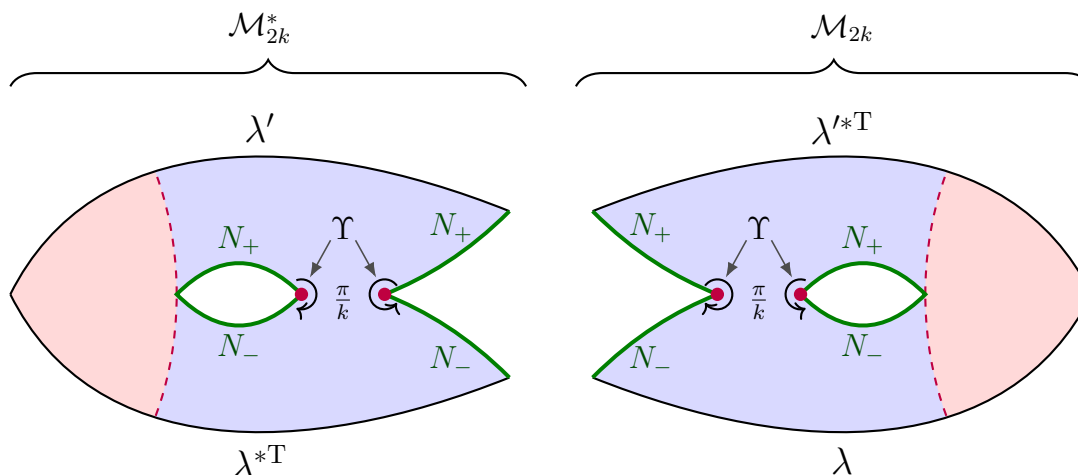
When $\lambda \neq \lambda'$, the previous $\mathbb{Z}_{2k}$ replica symmetry of the boundary conditions is explicitly broken down to a $\mathbb{Z}_k$ subgroup. Quotienting by this remaining symmetry, we can write

$$I[\widetilde{\mathcal{M}}_{2k}, \phi] = I[\widetilde{\mathcal{M}}_{2k} \setminus \widetilde{\Upsilon}, \phi] = I[\widehat{\mathcal{M}}_{2k} \setminus \Upsilon, \phi], \qquad (3.9)$$

where $\widehat{\mathcal{M}}_{2k} = \widetilde{\mathcal{M}}_{2k}/\mathbb{Z}_k$, and Υ is the image of the $\mathbb{Z}_k$ fixed codimension 2 surface $\widetilde{\Upsilon}$. There is a conical defect of opening angle $2\pi/k$ at Υ.

There is a remaining $\mathbb{Z}_2$ symmetry of the boundary conditions for $\widehat{\mathcal{M}}_{2k}, \phi$. This symmetry complex conjugates the boundary conditions, and reflects them across the remaining two copies of the gravitational region. Assuming this symmetry is not broken, it will fix a codimension 1 surface $N_+ \cup N^-$ which divides $\widehat{\mathcal{M}}_{2k}$ into two pieces, as depicted in figure 8. We will call the two pieces $\mathcal{M}_{2k}$ and $\mathcal{M}_{2k}^*$; they are shown separately in figure 9. The two pieces N_+ and N_- of $N_+ \cup N_-$ meet at Υ and the boundary of the asymptotic region. N_+ denotes the piece closest to the boundary conditions λ', and N_- denotes the piece closest to the boundary conditions λ. We can then write

$$I[\widehat{\mathcal{M}}_{2k} \setminus \Upsilon, \phi] = I[\mathcal{M}_{2k} \setminus \Upsilon, \phi] + I[\mathcal{M}_{2k}^* \setminus \Upsilon, \phi] = 2 \operatorname{Re}(I[\mathcal{M}_{2k} \setminus \Upsilon, \phi]), \qquad (3.10)$$

Figure 9. By cutting $\widehat{\mathcal{M}}_{2k}$ along $N_+ \cup N_-$, we get two manifolds $\mathcal{M}_{2k}$ and $\mathcal{M}_{2k}^*$, which are related to each other by reflection and complex conjugation.

where the last line follows from the $\mathbb{Z}_2$ symmetry. Note that, in $\mathcal{M}_{2k}$, the angle around Υ from N_- to N_+ is π/k, because the $\mathbb{Z}_2$ symmetry divides the $2\pi/k$ conical defect in two. Accounting for this with an area term, we have

$$I[\mathcal{M}_{2k} \setminus \Upsilon, \phi] = I[\mathcal{M}_{2k}, \phi] + (2k-1)\frac{\text{Area}[\Upsilon]}{4G}. \tag{3.11}$$

Because the fields at N_+ and N_- are invariant under the $\mathbb{Z}_2$, they define states $|\varphi_+\rangle$ and $|\varphi_-\rangle$. These states depend on k. Let us now analytically continue to $k = \frac{1}{2}$, so that $\mathcal{M}_{2k} \to \mathcal{M}_1$, which has no conical defect at Υ. Let U be a unitary operator that maps $|\varphi_-\rangle$ to $|\varphi_+\rangle$. Note that U only changes the fields in the gravitational region, so it is a unitary operator acting on $\mathcal{H}_G$. We can think of $I[\mathcal{M}_1, \phi]$ as the effective action in the presence of an insertion of U on a surface N. N_+ lies just to the 'future' of N, while N_- lies just to the 'past' of N. To be more precise, consider the correlator

$$\langle \lambda' | U \otimes I | \lambda \rangle = \frac{Z^U(\lambda', \lambda)}{\sqrt{Z(\lambda, \lambda)Z(\lambda', \lambda')}}, \tag{3.12}$$

where I is the identity acting on $\mathcal{H}_R$, and Z^U is the partition function with U inserted. We can compute this with an effective action

$$-\log Z^U(\lambda', \lambda) = I^U[\mathcal{M}, \phi]. \tag{3.13}$$

The right hand side is the effective action for boundary conditions λ' in the future, and λ in the past. Additionally, the fields in this action are acted on by U at N. We must compute this action for the configuration $\mathcal{M}, \phi$ which minimises it. But note that these boundary conditions, and the action of U, are exactly the same conditions that apply to $I[\mathcal{M}_1, \phi]$. Thus, we can write

$$-\log Z^U(\lambda', \lambda) = I[\mathcal{M}_1, \phi]. \tag{3.14}$$

The contribution of the twist operators in (3.8) is a symmetric function of λ and λ',

$$F(\lambda, \lambda') = -\frac{1}{2k} \log \langle \mathcal{T}_{2k} \dots \mathcal{T}_1 \rangle. \tag{3.15}$$

We know that this function is real, because the effective action is real, and so is (3.6) (since it is the trace of a Hermitian operator). From the calculations of the previous section, we also know that at $\lambda = \lambda'$ we can expand this function near $2k = 1$ in terms of the bulk entropy for those boundary conditions:

$$F(\lambda, \lambda) = (2k - 1)S_{\text{bulk}}(I \cup R)|_\lambda + \mathcal{O}\big((2k-1)^2\big) = \mathcal{O}(2k-1). \tag{3.16}$$

(We know that the $\mathcal{O}(1)$ piece vanishes because at $k = 1$ we have $\langle \mathcal{T}_{2k} \dots \mathcal{T}_1 \rangle = \langle 1 \rangle = 1$.) At this point, let us assume that λ' is only a small perturbation of λ. Assuming that $F(\lambda, \lambda')$ is sufficiently smooth near $\lambda = \lambda'$, for a small enough perturbation we can bound $F(\lambda, \lambda')$ with either

$$F(\lambda, \lambda) \leq F(\lambda, \lambda') \leq F(\lambda', \lambda') \tag{3.17}$$

or

$$F(\lambda, \lambda) \geq F(\lambda, \lambda') \geq F(\lambda', \lambda') \tag{3.18}$$

But since in both cases the upper and lower bounds are $\mathcal{O}(2k - 1)$, we must have

$$F(\lambda, \lambda') = \mathcal{O}(2k - 1). \tag{3.19}$$

Therefore, when we analytically continue to $k = \frac{1}{2}$ we will get

$$-\log \langle \mathcal{T}_{2k} \dots \mathcal{T}_1 \rangle \big|_{k=\frac{1}{2}} = 0. \tag{3.20}$$

Using (3.7) and continuing to $k = \frac{1}{2}$, we therefore get the formula for the fidelity when λ' is a small perturbation of λ:

$$\text{tr}\left(\sqrt{\sqrt{\rho_R} \rho'_R \sqrt{\rho_R}} \right) = \frac{1}{\sqrt{Z_{(1)}[\lambda] Z_{(1)}[\lambda']}} \exp\left(-\text{Re}(I[\mathcal{M}_1, \phi])\right). \tag{3.21}$$

The location of Υ is constrained as before by considering the behaviour of the partition function for small $2k - 1$. From this point on, let us write $\lambda' = \lambda + \epsilon \delta \lambda$, where $|\epsilon| \ll 1$. We have

$$-\frac{1}{2k} \log Z_{(2k)}[\lambda, \lambda + \epsilon \delta \lambda, \dots] = (I[\mathcal{M}, \bar{\phi}] + \mathcal{O}(\epsilon)) + (2k-1)(S + \mathcal{O}(\epsilon)) + \mathcal{O}\big((2k-1)^2\big), \tag{3.22}$$

where $\bar{\phi}$ is the field configuration at $\epsilon = 0$. We will need to minimise this for arbitray small ϵ. This means we need to minimise each of the terms in this expansion, so, up to $\mathcal{O}(\epsilon)$ corrections, Υ is located at the boundary of the island I.

At this point we can immediately write down a pair of states that satisfy Uhlmann's theorem for ρ_R and ρ'_R. Suppose V is any unitary operator acting on $\mathcal{H}_G$, and let

$$|\psi\rangle = (V \otimes I)|\lambda\rangle, \qquad |\psi'\rangle = (VU^\dagger \otimes I)|\lambda + \epsilon \delta \lambda\rangle, \tag{3.23}$$
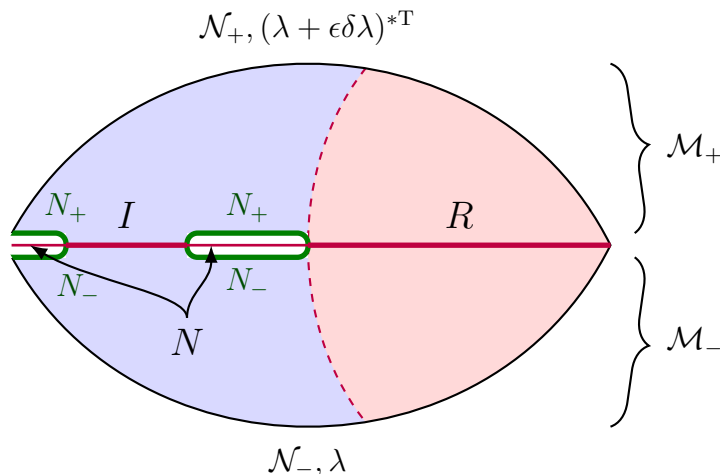
Figure 10. At small ϵ, computing the fidelity of ρ_R and ρ'_R is equivalent to computing the absolute value of the inner product of two states $|\lambda\rangle$ and $|\lambda'\rangle$, with the insertion of a unitary operator U acting at N. We have $N \cup I \cup R = \Sigma$, where Σ was shown in figure 3.

where U was defined above. Clearly these states are purifications of ρ_R, ρ'_R respectively. Also, we have

$$\langle\psi'|\psi\rangle = \langle\lambda + \epsilon\delta\lambda|U \otimes I|\lambda\rangle = \frac{1}{\sqrt{Z_{(1)}[\lambda]Z_{(1)}[\lambda + \epsilon\delta\lambda]}}\exp\left(-I[\mathcal{M}_1, \phi]\right), \tag{3.24}$$

where the second equality follows from (3.12) and (3.14). Taking the absolute value of this, and using (3.21), we observe that

$$\left|\langle\psi'|\psi\rangle\right| = \text{tr}\left(\sqrt{\sqrt{\rho_R}\rho'_R\sqrt{\rho_R}}\right), \tag{3.25}$$

as desired. This is illustrated in figure 10.

Overlap of parallel states. Let us now consider the behaviour of the fields ϕ for $\lambda' = \lambda + \epsilon\delta\lambda$, for small ϵ. At $\epsilon = 0$, there is an additional $\mathbb{Z}_2$ symmetry on $\mathcal{M}_1$ that complex conjugates the fields, and reflects everything in time. At $\epsilon = 0$, this symmetry implies that the fields on N_- are equal to those on N_+. This implies that the unitary operator U at $\epsilon = 0$ is just the identity. When ϵ is small but non-zero, this $\mathbb{Z}_2$ symmetry is broken, but we can still use it to analyse the resulting perturbation to the fields at linear order. We will write

$$\phi = \bar{\phi} + \frac{1}{2}\epsilon(\delta\phi + \tilde{\delta}\phi), \tag{3.26}$$

where $\bar{\phi}$ is the background field configuration at $\epsilon = 0$, and under the broken $\mathbb{Z}_2$ symmetry we have

$$\delta\phi \to \delta\phi, \qquad \tilde{\delta}\phi \to -\tilde{\delta}\phi. \tag{3.27}$$

The variation $\delta\phi$ corresponds to a variation in the boundary conditions of $\delta\lambda$ at $\mathcal{N}_-$ and $\delta\lambda^{*\text{T}}$ at $\mathcal{N}_+$. It is continuous as we go from N_- to N_+. The other variation $\tilde{\delta}\phi$ corresponds

to a variation in the boundary conditions of $-\delta\lambda$ at $\mathcal{N}_-$ and $\delta\lambda^{*\mathrm{T}}$ at $\mathcal{N}_+$. It is discontinuous as we go from N_- to N_+. In particular it changes sign.

Note that $\bar\phi$ is the dominant field configuration for the partition function $Z_{(1)}[\lambda]$, and $\bar\phi + \epsilon\delta\phi$ is the dominant field configuration for the partition function $Z_{(1)}[\lambda + \epsilon\delta\lambda]$. Thus, we may write (3.24) as

$$\langle\psi'|\psi\rangle = \exp\left(\frac{1}{2}I[\mathcal{M}, \bar\phi] + \frac{1}{2}I[\mathcal{M}, \bar\phi + \epsilon\delta\phi] - I[\mathcal{M}_1, \bar\phi + \frac{1}{2}\epsilon(\delta\phi + \tilde\delta\phi)] + \mathcal{O}\!\left(\epsilon^2\right)\right) \quad (3.28)$$

$$= \exp\left(-\frac{1}{2}\epsilon\tilde\delta I[\mathcal{M}_1, \bar\phi] + \mathcal{O}\!\left(\epsilon^2\right)\right). \quad (3.29)$$

Here, $\epsilon\tilde\delta I$ denotes the variation of the action due to $\bar\phi \to \bar\phi + \epsilon\tilde\delta\phi$. In terms of $\theta[\bar\phi, \tilde\delta\phi]$, this can be written

$$\frac{1}{2}\tilde\delta I[\mathcal{M}_1, \bar\phi] = \frac{1}{2}\left(\int_{\partial\mathcal{M}_1} \theta[\bar\phi, \tilde\delta\phi]\right) \quad (3.30)$$

$$= \frac{1}{2}\left(\int_{\mathcal{N}_-} \theta[\bar\phi, \tilde\delta\phi] + \int_{\mathcal{N}_+} \theta[\bar\phi, \tilde\delta\phi] + \int_{N_-} \theta[\bar\phi, \tilde\delta\phi] + \int_{N_+} \theta[\bar\phi, \tilde\delta\phi]\right) \quad (3.31)$$

$$= \int_{\mathcal{N}_+} \theta[\bar\phi, \tilde\delta\phi] + \int_{N_+} \theta[\bar\phi, \tilde\delta\phi]. \quad (3.32)$$

The manifold $\mathcal{M}_1$ is shown in figure 10, from which we can see that $\partial\mathcal{M}_1 = \mathcal{N}_+ \cup \mathcal{N}_- \cup N_+ \cup N_-$, and so the second line holds. The third line follows from (3.27). The change in boundary conditions for $\tilde\delta\phi$ and $\delta\phi$ are the same at $\mathcal{N}_+$, and $\theta|_{\mathcal{N}_+}$ only depends on the change in boundary conditions, so we can write

$$\int_{\mathcal{N}_+} \theta[\bar\phi, \tilde\delta\phi] = \int_{\mathcal{N}_+} \theta[\bar\phi, \delta\phi]. \quad (3.33)$$

Also, it will be convenient to use a basis for the states on N such that

$$\int_{N_+} \theta[\bar\phi, \tilde\delta\phi - \delta\phi] = 0. \quad (3.34)$$

As explained in the previous section, this can be done by adding appropriate boundary terms to the action. Then we have

$$\langle\psi'|\psi\rangle = \exp\left(-\epsilon\int_{\partial\mathcal{M}^+\cup N} \theta[\bar\phi, \delta\phi] + \mathcal{O}\!\left(\epsilon^2\right)\right). \quad (3.35)$$

3.2 Uhlmann phase

At this point, we are ready to compute the Uhlmann phase of a curve $\rho_R(t)$ of density matrices corresponding to a curve of boundary conditions $\lambda(t)$, $0 \le t \le 1$. At large m, for each pair of density matrices $\rho(i/m)$ and $\rho_R((i+1)/m)$, we can use the derivation above to obtain a unitary operator U_i such that

$$|\langle\lambda((i+1)/m)|U_i \otimes I|\lambda(i/m)\rangle| = \mathrm{tr}\left(\sqrt{\sqrt{\rho_R(i/m)}\rho_R((i+1)/m)\sqrt{\rho_R(i/m)}}\right). \quad (3.36)$$

Each of the operators U_i maps a state $|\varphi_-^i\rangle$ on N_- to a state $|\varphi_+^i\rangle$ on N_+. These states come from the boundary conditions at $N_\pm$ corresponding to a field configuration $\phi_i = \bar{\phi}_i + \frac{1}{2m}\left(\delta_i\phi + \tilde{\delta}_i\phi\right)$, where $\bar{\phi}_i$ is the dominant field configuration relevant to $Z_{(1)}[\lambda(i/m)]$, and

$$\bar{\phi}_{i+1} - \bar{\phi}_i = \frac{1}{m}\delta_i\phi + \mathcal{O}\left(\frac{1}{m^2}\right). \tag{3.37}$$

To be precise, we have

$$\varphi_\pm^i = \left.\left(\bar{\phi}_i + \frac{1}{2m}\left(\delta_i\phi + \tilde{\delta}_i\phi\right)\right)\right|_{\text{b.c.s at } N_\pm}. \tag{3.38}$$

By adding the appropriate boundary terms to the action, we can ensure that the states on $N_\pm$ are in a basis such that

$$\int_{N_+} \theta[\bar{\phi}_i, \tilde{\delta}_i\phi - \delta_i\phi] = 0, \tag{3.39}$$

or equivalently

$$\int_{N_-} \theta[\bar{\phi}_i, \tilde{\delta}_i\phi + \delta_i\phi] = 0. \tag{3.40}$$

It is possible to find a boundary term that enforces this condition for all i, because in field space the vector $(\tilde{\delta}_i\phi - \delta_i\phi)|_{N_+}$ is never tangent to the curve $\bar{\phi}(t)$. Since this condition determines the class of boundary conditions φ, we can use it to write

$$\varphi_+^i = \left.\left(\bar{\phi}_i + \frac{1}{m}\delta_i\phi\right)\right|_{\text{b.c.s at } N}, \tag{3.41}$$

$$\varphi_-^i = \bar{\phi}_i\big|_{\text{b.c.s at } N}. \tag{3.42}$$

Therefore, $|\varphi_+^i\rangle = |\varphi_-^{i+1}\rangle$.

Then we define the sequence of states

$$|\psi(0)\rangle = |\lambda(0)\rangle, \tag{3.43}$$

$$|\psi(1/m)\rangle = (U_0^\dagger \otimes I)|\lambda(1/m)\rangle, \tag{3.44}$$

$$\cdots \tag{3.45}$$

$$|\psi(i/m)\rangle = (U_0^\dagger U_1^\dagger \ldots U_{i-1}^\dagger \otimes I)|\lambda(i/m)\rangle, \tag{3.46}$$

$$\cdots \tag{3.47}$$

$$|\psi(1)\rangle = (U_0^\dagger U_1^\dagger \ldots U_{m-1}^\dagger \otimes I)|\lambda(1)\rangle. \tag{3.48}$$

These states purify $\rho(0), \rho(1/n), \ldots$, and satisfy (3.4), so as $m \to \infty$ we get a parallel purifying curve $|\psi(t)\rangle$.

To find the Uhlmann phase we need to compute the overlaps of these states. First, from (3.35) we know that

$$\langle\psi((i+1)/m)|\psi(i/m)\rangle = \exp\left(-\frac{1}{m}\int_{\mathcal{N}_+\cup N}\theta[\bar{\phi}_i, \delta_i\phi] + \mathcal{O}\left(\frac{1}{m^2}\right)\right). \tag{3.49}$$

We also need to compute

$$\langle\psi(0)|\psi(1)\rangle = \langle\lambda(0)|U_0^\dagger U_1^\dagger \ldots U_{m-1}^\dagger|\lambda(0)\rangle. \tag{3.50}$$

But note that

$$\langle \varphi_0^- | U_0^\dagger U_1^\dagger \ldots U_{m-1}^\dagger = \langle \varphi_0^+ | U_1^\dagger \ldots U_{m-1}^\dagger \tag{3.51}$$

$$= \langle \varphi_1^- | U_1^\dagger \ldots U_{m-1}^\dagger \tag{3.52}$$

$$= \ldots \tag{3.53}$$

$$= \langle \varphi_m^- | = \langle \varphi_0^- |, \tag{3.54}$$

where the last equality follows because the curve of states is closed. So $U_0^\dagger U_1^\dagger \ldots U_{m-1}^\dagger$ maps $|\varphi_0^-\rangle$ to itself. φ_0^- are just the boundary conditions for $\bar{\phi}_0$ at N, so inserting this operator in between $\langle \lambda(0)|$ and $|\lambda(0)\rangle$ actually has no effect at the semiclassical level, because the dominant field configuration is unchanged. Thus, we just have

$$\langle \lambda(0)|U_0^\dagger U_1^\dagger \ldots U_{m-1}^\dagger|\lambda(0)\rangle = \exp\Big(\mathcal{O}\Big(\frac{1}{m^2}\Big)\Big). \tag{3.55}$$

Finally, the Uhlmann phase is given by

$$e^{i\gamma} = \lim_{m\to\infty} \langle \psi(0)|\psi(1)\rangle \, \langle \psi(1)|\psi\Big(\frac{m-1}{m}\Big)\rangle \ldots \langle \psi\Big(\frac{2}{m}\Big)|\psi\Big(\frac{1}{m}\Big)\rangle \, \langle \psi\Big(\frac{1}{m}\Big)|\psi(0)\rangle \tag{3.56}$$

$$= \lim_{m\to\infty} \prod_{i=0}^{m-1} \exp\Big(-\frac{1}{m}\int_{\mathcal{N}^+ \cup N}\theta(\bar{\phi}_i, \delta_i\phi) + \mathcal{O}\Big(\frac{1}{m^2}\Big)\Big) \tag{3.57}$$

$$= \lim_{m\to\infty} \exp\Big(-\frac{1}{m}\sum_{i=0}^{m-1}\int_{\mathcal{N}^+ \cup N}\theta(\bar{\phi}_i, \delta_i\phi) + \mathcal{O}\Big(\frac{1}{m^2}\Big)\Big). \tag{3.58}$$

Taking the $m \to \infty$ limit converts the sum to an integral, and we are left with

$$\gamma = \int_C \Theta, \tag{3.59}$$

where C is the curve of field configurations $\bar{\phi}(t) = \bar{\phi}(\lambda(t))$ relevant to the partition function $Z_{(1)}[\lambda(t)]$, and Θ is the one-form on field space whose component in the direction $\delta\phi$ at a field configuration ϕ is given by

$$\Theta[\phi, \delta\phi] = i\int_{\mathcal{N}_+ \cup N}\theta[\phi, \delta\phi]. \tag{3.60}$$

4 Classical information recovery

We now wish to show how to use the Uhlmann phase of the Hawking radiation to recover classical information contained in the black hole interior. The main claim is that the field space 1-form Θ defined in (3.60) is in fact a symplectic potential for $I \cup R$, i.e. the island and the radiation. This means its exterior derivative is equal to the symplectic form for the fields in $I \cup R$. To justify this claim, let us take the exterior derivative:

$$\delta\Theta = i\int_{\mathcal{N}_+ \cup N}\omega. \tag{4.1}$$

Here, $\omega = \delta\theta$, and δ denotes an exterior derivative on field space. Note that $d\omega = \delta(\delta L) = 0$ on-shell. All field configurations we are considering are on-shell. Thus we can use Stokes' theorem to write

$$\delta\Theta = -i \int_{I \cup R} \omega, \tag{4.2}$$

because $\mathcal{N}_+ \cup N \cup I \cup R = \partial\mathcal{M}_+$, as can be verified from figure 10. The above expression is in terms of Euclidean fields. The fields at $I \cup R$ are invariant under time reflection and complex conjugation, so we can Wick rotate to real Lorentzian fields. After doing so, the factor of i goes away because there are an odd number of time derivatives, and we end up with

$$\delta\Theta = \int_{I \cup R} \omega_{\text{Lorentzian}} = \Omega_{I \cup R}. \tag{4.3}$$

The right hand side here is exactly the definition of the symplectic form of $I \cup R$, according to the covariant phase space formalism [37–52]. So we can conclude that Θ is indeed a symplectic potential for $I \cup R$. Therefore, the curvature of the Uhlmann phase of the Hawking radiation is the symplectic form of $I \cup R$.

4.1 General information recovery

Before explaining how to apply our results to the recovery of classical information from black holes, it is important to first describe what we mean by information recovery. There are various ways to formulate this — we will use one that is reminiscent of error correction, and which is directly applicable to the black hole case.

We start with a system whose initial state takes values in some space $\mathcal{P}_1$. We also have another space $\mathcal{I}$, each element of which describes the state of some information in the system. To embed this information in the physical system we assume the existence of a map

$$f: \quad \mathcal{I} \to \mathcal{P}_1, \quad i \mapsto x_i. \tag{4.4}$$

For each possible value taken by the information $i \in \mathcal{I}$, $f(i) = x_i \in \mathcal{P}_1$ is the physical state containing that information. We assume that there exists some map $R_1 : \mathcal{P}_1 \to \mathcal{I}$ obeying $R_1(x_i) = i$. This means that, after embedding the information in the system, we can easily get it back using R_1.

However, we are interested in a process which could potentially erase the information. In particular, we will assume the existence of a space of final states $\mathcal{P}_2$, and a map $\mathscr{S} : \mathcal{P}_1 \to \mathcal{P}_2$ which represents the evolution of the system. After this evolution, the state of the system is $\bar{x}_i = \mathscr{S}(x_i)$. The information is recoverable if there exists a map $R_2 : \mathcal{P}_2 \to \mathcal{I}$ which obeys $R_2(\bar{x}_i) = i$, because we can use this map to read the information from the final state.

The typical example is that of a diary on fire. The initial space $\mathcal{P}_1$ consists of all of the possible initial states of the diary, while each $i \in I$ specifies the diary's contents, i.e. the state x_i has the property that i is written in the diary. We then set the diary on fire and let the fire burn until it has exhausted its fuel — this burning happens via the map $\mathscr{S} : \mathcal{P}_1 \to \mathcal{P}_2$. The space $\mathcal{P}_2$ consists of all of the possible final states of the radiation emitted from the fire, as well as the leftover ash. In principle, this is quantum mechanically

a unitary process, so the map $\mathscr{S}$ is invertible, and we can just write $R_2 = R_1 \circ \mathscr{S}^{-1}$. Thus, the information (i.e. the contents of the diary) is recoverable. On the other hand, burning is a highly chaotic process, and in practice we can't measure the final state of the ash and radiation with perfect accuracy. To account for this we can modify $\mathscr{S}$ so that it implements a coarse graining of the final state, and redefine $\mathcal{P}_2$ to contain coarse grained states. This renders $\mathscr{S}$ uninvertible — so we can no longer use the previous trick to recover information. Indeed, information will typically be lost in such a process.

For what follows it is useful to have:

Lemma. *Information is recoverable if and only if, given any function $F : \mathcal{P}_1 \to \mathbb{R}$, one can construct a function $\bar{F} : \mathcal{P}_2 \to \mathbb{R}$ obeying $F(x_i) - F(x_j) = \bar{F}(\bar{x}_i) - \bar{F}(\bar{x}_j)$ for all $i, j \in \mathcal{I}$.*

Proof. Suppose information is recoverable. Then we can set $\bar{F} = F \circ f \circ R_2$, and for all $i \in \mathcal{I}$ we have

$$\bar{F}(\bar{x}_i) = F(f(R_2(\bar{x}_i))) = F(f(i)) = F(x_i). \tag{4.5}$$

Thus $F(x_i) - F(x_j) = \bar{F}(\bar{x}_i) - \bar{F}(\bar{x}_j)$ as required.

For the opposite direction, let us pick a set of functions $a_n : \mathcal{I} \to \mathbb{R}$, $n = 1, 2, \ldots$ such that the map $A : i \mapsto (a_1(i), a_2(i), \ldots)$ is injective (so that knowing the values of $a_1, a_2, \ldots$ fully determines the state of the information). Moreover, let us fix a $j \in \mathcal{I}$, and pick a_n such that $a_n(j) = 0$ for all n. Let $F_n = a_n \circ R_1$; then by assumption there exists an $\bar{F}_n : \mathcal{P}_2 \to \mathbb{R}$ obeying

$$\bar{F}_n(\bar{x}_i) - \bar{F}_n(\bar{x}_j) = F_n(x_i) - \underbrace{F_n(x_j)}_{=a_n(j)=0} = a_n(i). \tag{4.6}$$

Note that we can set $\bar{F}_n(\bar{x}_j) = 0$ by redefining $\bar{F}_n(\bar{x}) \to \bar{F}_n(\bar{x}) - \bar{F}_n(\bar{x}_j)$. Then we just have $\bar{F}_n(\bar{x}_i) = a_n(i)$, so by measuring $\bar{F}_n$ in the final state we get the values of a_n for all n, and thus we can determine i. This yields the map[8]

$$R_2 : \quad \bar{x} \mapsto A^{-1}(\bar{F}_1(\bar{x}), \bar{F}_2(\bar{x}), \ldots), \tag{4.7}$$

which obeys $R_2(\bar{x}_i) = i$ and so renders the information recoverable. $\qquad\square$

4.2 Black hole classical information recovery

Let us now see how we can use the Uhlmann phase to recover classical information from the interior of a black hole. In particular, we will show how one can recover any function of the classical information by measuring the Uhlmann phase around an appropriate curve of states.

Let $\mathcal{P}_1 = \mathcal{P}_{\text{classical}}$ be a phase space of all classical field configurations containing a black hole. Let us suppose for simplicity that after a certain time T the classical field configuration outside of the black hole is the same for all configurations in $\mathcal{P}_{\text{classical}}$. This

---

[8]Note this map is only guaranteed to be defined on $\bar{x} = \bar{x}_i$ for $i \in \mathcal{I}$. But for other values of $\bar{x}$ we can just set $R_2(\bar{x})$ to be anything at all, since its behaviour on such final states is irrelevant to the question of information recovery.

means that an observer who can only read the classical values of the fields outside the black hole at time T has no way of knowing which state in $\mathcal{P}_{\text{classical}}$ is the true one. To model the information recovery capabilities of such an observer, we would set $\mathcal{P}_2$ to contain the possible classical field configurations outside of the black hole at time T, and $\mathscr{S}$ would be the restriction map from the total field configuration to this exterior field configuration. Since $\mathcal{P}_2$ only contains one configuration, and $\mathscr{S}$ maps everything to that configuration, the observer is incapable of recovering any non-trivial information from the black hole.

Let us therefore give our observer an upgrade: we now will also allow them access to the quantum state ρ of the Hawking radiation it emits. If Hawking's original calculations were valid, this state would have no knowledge of the black hole interior, and so the observer would be no more capable of recovering information than they were before their upgrade. However, the recent breakthroughs [8–13] show that, past the Page time, the state of the radiation *does* contain information about the interior.

For the observer with access to the Hawking radiation, $\mathcal{P}_2 = \mathcal{P}_{\text{Hawking}}$ is the space of possible states of Hawking radiation at time T arising from classical field configurations in $\mathcal{P}_{\text{classical}}$. Given a field configuration $\phi \in \mathcal{P}_{\text{initial}}$, we get a unique Hawking radiation state $\rho \in \mathcal{P}_{\text{Hawking}}$; let this define the map

$$\mathscr{S}: \quad \mathcal{P}_{\text{classical}} \to \mathcal{P}_{\text{Hawking}}, \quad \phi \mapsto \rho. \tag{4.8}$$

As in the previous section, we can imagine that the field configuration ϕ contains a diary. We can write some classical information $i \in I$ into that diary; let the corresponding field configuration be denoted ϕ_i, and let the corresponding Hawking radiation be denoted $\rho_i = \mathscr{S}(\phi_i)$. Let us take T to be greater than the Page time, so that an island has formed inside of the black hole. We will assume that the diary falls into the black hole and is at time contained T within the island.[9] We will further assume that the fields away from the diary are (to a good approximation) independent of i — in other words, the information is localised to the diary.

We will show that the information in the diary is recoverable by using the lemma from the previous subsection. Suppose $F : \mathcal{P}_{\text{classical}} \to \mathbb{R}$ is a function of the classical field configuration. Note that if we restrict to $\phi = \phi_i$, then by our assumptions $F(\phi_i)$ only depends on the fields in the vicinity of the diary. Thus, we can define a new function $\tilde{F} : \mathcal{P}_{\text{classical}} \to \mathbb{R}$ such that

$$\tilde{F}(\phi_i) = F(\phi_i), \tag{4.9}$$

and such that $\tilde{F}(\phi)$ only depends on the fields in the vicinity of the diary for all $\phi \in \mathcal{P}_{\text{classical}}$.

Since $\mathcal{P}_{\text{classical}}$ is a phase space, it comes equipped with a symplectic structure Ω, which according to the covariant phase space formalism can be written as the integral of the previously mentioned spacetime 2-form $\omega_{\text{Lorentzian}}$ over any Cauchy surface Σ:

$$\Omega = \int_{\Sigma} \omega_{\text{Lorentzian}}. \tag{4.10}$$

---

[9]Otherwise it would not possible to recover the information in the diary. Eventually the island spans the entire black hole interior, so if the diary is not yet in the island we can just increase T until it is.

It is convenient to pick a Σ which contains the island I. Let V be the vector field on $\mathcal{P}_{\text{classical}}$ obeying

$$\iota_V \Omega = -\delta \tilde{F}, \tag{4.11}$$

i.e. V is the Hamiltonian vector field associated with $\tilde{F}$. Let $\delta_V \phi$ denote the variation of the field configuration associated with this vector field. Note that since $\tilde{F}$ only depends on the fields in the vicinity of the diary, $\delta_V \phi$ must vanish away from the diary, which means we can write

$$\iota_V \Omega = \iota_V \Omega_{I \cup R}, \tag{4.12}$$

where $\Omega_{I \cup R}$ was defined in (4.3).

Now let $\bar{V} = \mathscr{S}_* V$ be the pushforward of V through the dynamical map $\mathscr{S}$. We can use $\bar{V}$ and the Uhlmann phase to construct a certain function $\bar{F} : \mathcal{P}_{\text{Hawking}} \to \mathbb{R}$, in the following way. Let us fix a $\rho_0 \in \mathcal{P}_{\text{Hawking}}$. Suppose $\rho \in \mathcal{P}_{\text{Hawking}}$ is connected to ρ_0 by a curve $D_1 \subset \mathscr{S}$. Let $\bar{\psi}_t$ be the flow generated by the vector field $\bar{V}$, let

$$D_2 = \{\bar{\psi}_t(\rho) : t \in [0, \epsilon]\}, \quad D_3 = \bar{\psi}_\epsilon(D_1), \quad D_4 = \{\bar{\psi}_t(\rho_0) : t \in [0, \epsilon]\}, \tag{4.13}$$

where $0 < \epsilon \ll 1$, and let $D = D_1 \cup D_2 \cup D_3 \cup D_4$ be the closed curve obtained by taking the union of these four curves. Finally, let

$$\bar{F}(\rho) = \lim_{\epsilon \to 0} \frac{\gamma(D)}{\epsilon}, \tag{4.14}$$

where $\gamma(D)$ is the Uhlmann phase around the curve D.

Note that D is the image under $\mathscr{S}$ of a curve of field configurations $C \subset \mathcal{P}_{\text{classical}}$, and using (3.59) and (4.3) we can write

$$\bar{F}(\rho) = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_C \Theta \tag{4.15}$$

where Θ is the symplectic potential of the fields in the island and asymptotic regions. Similarly, each of the four curves D_1, D_2, D_3, D_4 are the images under $\mathscr{S}$ of four curves C_1, C_2, C_3, C_4 respectively, where C_1 connects ϕ_0 to ϕ (which obey $\mathscr{S}(\phi) = \rho$, $\mathscr{S}(\phi_0) = \rho_0$), and we have
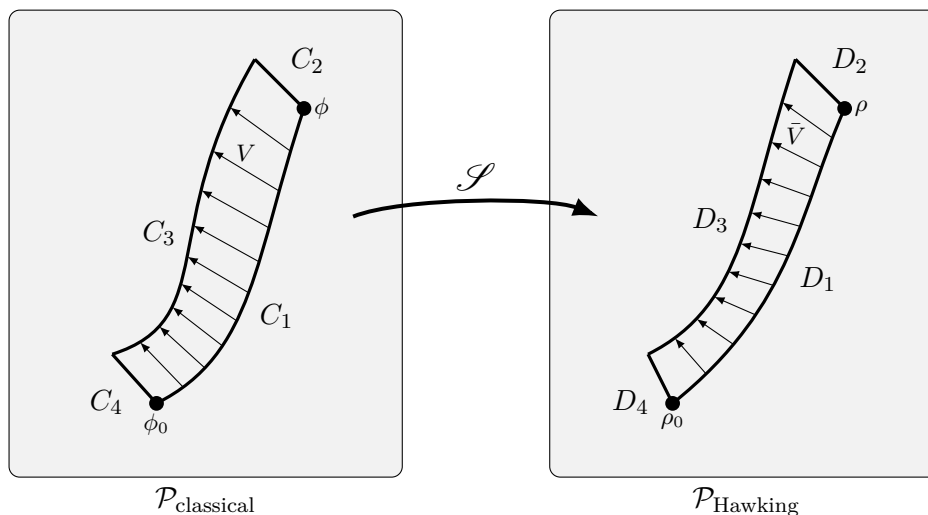
$$C_2 = \{\psi_t(\phi) : t \in [0, \epsilon]\}, \quad C_3 = \psi_\epsilon(C_1), \quad C_4 = \{\psi_t(\phi_0) : t \in [0, \epsilon]\}, \tag{4.16}$$

where ψ_t is the flow generated by the vector field V. Both of the curves C and D are depicted in figure 11.

Taking into account the correct orientations of the components of C, we have

$$\begin{aligned}
\bar{F}(\rho) &= \lim_{\epsilon \to 0} \frac{1}{\epsilon} \left[\int_{C_1} \Theta + \int_{C_2} \Theta + \int_{C_3} \Theta + \int_{C_4} \Theta \right] \\
&= \partial_\epsilon \left[\int_0^\epsilon \mathrm{d}s \, (\iota_V \Theta)(\psi_s(\rho)) - \int_{\psi_\epsilon(C_1)} \Theta - \int_0^\epsilon \mathrm{d}s \, (\iota_V \Theta)(\psi_s(\rho_0)) \right]_{\epsilon=0} \\
&= (\iota_V \Theta)(\rho) - \underbrace{\int_{C_1} \mathcal{L}_V \Theta}_{= \delta(\iota_V \Theta) + \iota_V \delta \Theta} - (\iota_V \Theta)(\rho_0) \\
&= - \int_{C_1} \iota_V \Omega_{I \cup R},
\end{aligned} \tag{4.17}$$

Figure 11. The curves $C \subset \mathcal{P}_{\text{classical}}$ and $D = \mathscr{S}(C) \subset \mathcal{P}_{\text{Hawking}}$.

where $\Omega_{I \cup R} = \delta \Theta$. But now, using (4.11) and (4.12), we may write

$$\bar{F}(\rho) = \int_{D_1} \delta \tilde{F} = \tilde{F}(\phi) - \tilde{F}(\phi_0). \tag{4.18}$$

Thus

$$\bar{F}(\rho_i) - \bar{F}(\rho_j) = \tilde{F}(\phi_i) - \tilde{F}(\phi_j) = F(\phi_i) - F(\phi_j). \tag{4.19}$$

Therefore, by the lemma in the previous section, the information in the diary is recoverable from the Hawking radiation. Moreover, we have described the protocol by which one recovers the information: one picks a function F of the classical information, one finds the Hamiltonian vector V that it generates, one pushes this vector forward to the space of Hawking states, one uses it to construct a curve of states, and then one computes the Uhlmann phase around that curve. The result is a reconstruction of the value of the function F, in the way just described.

4.3 Island phase space reconstruction

Our result also makes it possible to recover the full structure of the island phase space — so in particular we can compute Poisson brackets of all the functions of the information in the black hole. By measuring the Uhlmann phases of various curves in $\mathcal{P}_{\text{Hawking}}$, one can completely recover all of the components of the symplectic form $\Omega_{I \cup R}$. Actually, because all the configurations in $\mathcal{P}_{\text{classical}}$ are assumed to lead to the same classical exterior, in this case $\Omega_{I \cup R}$ is just the symplectic form for the fields on the island I (because ϕ will be constant in R, so $\delta\phi|_R = 0$).

In this way, although $\mathcal{P}_{\text{Hawking}}$ is a space of quantum states, in the semiclassical limit it aquires a symplectic structure. Thus, it becomes a symplectic manifold, and so (in the semiclassical limit) can be thought of as a classical phase space. As a phase space it is isomorphic to the phase space of the fields on the island. So we have reconstructed the classical phase space of the island: it is just isomorphic to $\mathcal{P}_{\text{Hawking}}$ with the symplectic

form $\mathscr{S}^*\Omega_{I\cup R}$. This means we are able to compute arbitrary Poisson brackets of fields on the island by measuring the Hawking radiation Uhlmann phase.

Note that we do not need to know anything about the map $\mathscr{S}$ in order to construct Poisson brackets of island observables. In particular, suppose $\bar{F}_1, \bar{F}_2$ are two functions on $\mathcal{P}_{\text{Hawking}}$. These can be viewed as classical observables, and indeed this is clear if we pull them back through $\mathscr{S}$:

$$F_1 = \bar{F}_1 \circ \mathscr{S}, \quad F_2 = \bar{F}_2 \circ \mathscr{S}. \tag{4.20}$$

If we don't know about $\mathscr{S}$, then we can't directly access F_1, F_2. However, using the Uhlmann phase, we can deduce all the components of the symplectic form on $\mathcal{P}_{\text{Hawking}}$, and we can do this *without* needing to use $\mathscr{S}$. We can use this symplectic form to compute the Poisson bracket of $\bar{F}_1$ and $\bar{F}_2$

$$\left\{\bar{F}_1, \bar{F}_2\right\}_{\text{Hawking}} = -(\mathscr{S}^*\Omega_{I\cup R})^{-1}(\delta\bar{F}_1, \delta\bar{F}_2). \tag{4.21}$$

As in (4.20), this corresponds to the classical observable

$$\begin{aligned}
\left\{\bar{F}_1, \bar{F}_2\right\}_{\text{Hawking}} \circ \mathscr{S} &= -(\mathscr{S}^*\Omega_{I\cup R})^{-1}(\delta\bar{F}_1, \delta\bar{F}_2) \circ S \\
&= -\Omega_{I\cup R}^{-1}(\mathscr{S}^*\delta\bar{F}_1, \mathscr{S}^*\delta\bar{F}_2) \\
&= -\Omega_{I\cup R}^{-1}(\delta F_1, \delta F_2) \\
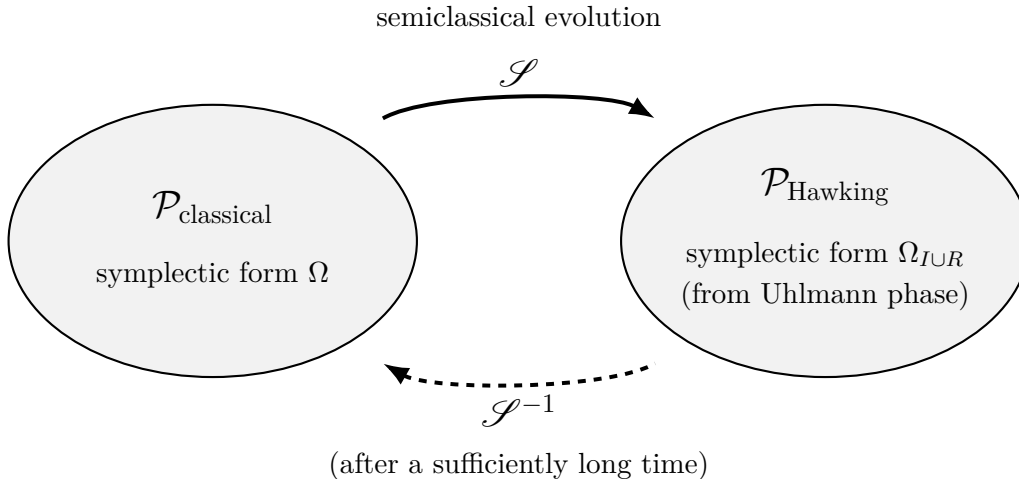&= \{F_1, F_2\}_{\text{classical}}.
\end{aligned} \tag{4.22}$$

Thus, we can construct classical Poisson brackets using the Uhlmann phase alone, without knowledge of $\mathscr{S}$. We only need to know about $\mathscr{S}$ when we want to actually evaluate these Poisson brackets on a given Hawking radiation state.

Finally, if we wait long enough after the Page time, the island will grow to span the entire black hole interior. At that point, we can use our results to recover all information that fell inside the black hole, and to reconstruct the entire classical phase space of the interior.

5 Conclusion

In this paper, we have shown that the curvature of the Uhlmann phase of Hawking radiation is equivalent to the symplectic form of the island and the radiation. This calculation depended in a crucial way on the contributions of replica wormholes to the gravitational path integral. Our results have provided a protocol for reconstructing the phase space of the black hole interior. All classical information is explicitly recovered after enough Hawking radiation is emitted. This is depicted in figure 12.

We have so far restrained from discussing the operational viability of our information recovery protocol. In particular, we have assumed that as an observer we have access to some experimental apparatus that allows us to move the state of the Hawking radiation around certain curves in $\mathcal{P}_{\text{Hawking}}$, and to measure the resulting Uhlmann phase. It is clear that the apparatus only needs access to the state of the Hawking radiation, so in principle interior information can be recovered from the outside of the black hole, but we have not

semiclassical evolution

Figure 12. We reconstruct the island phase space by measuring the Uhlmann phase of the Hawking radiation. After a sufficiently long time, the map $\mathscr{S}$ representing the evolution between the space of classical fields and the space of Hawking radiation states becomes invertible, and all classical information is recovered.

considered what kind of computational resources we would need in order to construct and use the apparatus. We will leave this important question to future work, but it is worth making two comments.

First, any process which recovers information from Hawking radiation is believed to be highly algorithmically complex [53], and this should apply to the results of this paper. There are two stages to our protocol: the construction of the curve of radiation states, and the measurement of the Uhlmann phase around that curve. It seems highly likely that for our protocol the complexity comes in at the first stage, because the construction of the curve uses the dynamical map $\mathscr{S} : \mathcal{P}_{\text{classical}} \to \mathcal{P}_{\text{Hawking}}$. The most obvious way to experimentally implement this map would be to simulate the evolution of the black hole on a quantum computer, which is a very complex procedure. On the other hand, there is no obvious reason why the measurement of the Uhlmann phase itself should be complex, once we know which curve of states to measure it around. As shown in (4.22), we do not need $\mathscr{S}$ to access the Poisson brackets of classical observables. So if we are only interested in the algebraic relations of observables in the black hole interior, we don't need to simulate the evolution of the black hole, i.e. we can avoid the complex stage of the protocol.

Second, there is an important distinction between 'asymptotic' and 'one-shot' information recovery. In the former case, we are tasked with recovering information using an arbitary number of copies of a final state. In the latter case, we only have a single copy of the final state from which we need to be able to recover the information with high probability. One-shot information recovery is strictly more difficult than asymptotic information recovery, but it is also strictly more desirable in a practical setting. There exists in the literature a description of an experiment for measuring Uhlmann phase [54, 55], but this measurement is an asymptotic one, because it requires many copies of the state. The distinction between the asymptotic and one-shot types of information recovery is par-

ticularly relevant to quantum systems, because the no-cloning theorem implies that there is no way of producing an arbitary number of copies of the final state, if we start with just a single copy. However, even though we must do quantum measurements to find the Uhlmann phase, our end goal in the present work is only the recovery of classical information. There is no problem with copying classical information an arbitrary number of times with arbitrary accuracy in the classical limit. Our expectation therefore is that if there is an asymptotic implementation of the protocol described here, then there is also a one-shot implementation. It would be very interesting to see whether this expectation is correct.

We should mention that another method for recovering black hole interior information from the Hawking radiation uses the so-called 'Petz map', as described in [10]. This map provides a way to reconstruct operators acting on the fields in the island as operators acting on the Hawking radiation. In this way it provides a way to recover *quantum* information (in the form of expectation values of operators acting inside the black hole), which is in contrast to the *classical* information recovery described in this work. It is clear that the protocol described here is quite different to the one involving the Petz map. Note that not all classical observables can be represented as the expectation values of quantum operators — so it is plausible that our protocol provides access to some information that the Petz map does not. On the other hand, in the classical limit, the two protocols may become more closely aligned, and this would be interesting to check. It would also be interesting to understand how the Petz map and Uhlmann phase are related in a more general setting.

Before ending the paper, we want to just point out an interesting consequence of our results. We have been fairly vague about the type of theory we are considering in the asymptotic region. Really, the only condition it has to satisfy is that it can be consistently coupled to the gravitational theory at the cutoff surface, in such a way that it absorbs the Hawking radiation.

But what we want to claim is that *any* theory which can be consistently coupled to gravity in this way *is itself a gravitational theory*, at least in some kind of classical limit.

To see this, let us take a state in such a theory, and couple it to gravity. We then form a black hole in the gravitational region, and let it evaporate. Once the asymptotic theory has absorbed the Hawking radiation, we throw away the gravitational theory — we don't care about it any more. We are only interested in the state of the asymptotic theory ρ_R.

As we have shown, the curvature of the Uhlmann phase of such states is a symplectic form for gravitational fields. So even though we have discarded the gravitational theory, we still have a class of states in the asymptotic theory with a gravitational interpretation. If we are interested in a classical limit in which the Uhlmann phase is important, then this necessarily *implies* that the classical limit of the asymptotic theory is a theory of gravity.

We think this observation has two possible explanations. Either: the only theories in the asymptotic region which can be consistently coupled to the gravitational region must already have had gravity in them. Or: there is a 'gravitational sector' of density matrices in a wide class of theories that don't a priori contain gravity in a traditional sense. Both of these possibilities seem interesting, and we leave their exploration to future work.

Acknowledgments

I wish to thank Philipp Hoehn, Isha Kotecha and Fabio Mele for helpful discussions and comments. This work was supported in part by the Okinawa Institute of Science and Technology. I am also grateful for the continued hospitality of the University of Cambridge while I am stuck in the UK.

References

[1] S.W. Hawking, *Particle Creation by Black Holes*, *Commun. Math. Phys.* **43** (1975) 199 [*Erratum ibid.* **46** (1976) 206] [ɪɴSPIRE].

[2] S.W. Hawking, *Black hole explosions*, *Nature* **248** (1974) 30 [ɪɴSPIRE].

[3] S.W. Hawking, *Breakdown of Predictability in Gravitational Collapse*, *Phys. Rev. D* **14** (1976) 2460 [ɪɴSPIRE].

[4] D.N. Page, *Information in black hole radiation*, *Phys. Rev. Lett.* **71** (1993) 3743 [`hep-th/9306083`] [ɪɴSPIRE].

[5] D.N. Page, *Time Dependence of Hawking Radiation Entropy*, *JCAP* **09** (2013) 028 [`arXiv:1301.4995`] [ɪɴSPIRE].

[6] J.M. Maldacena, *The Large N limit of superconformal field theories and supergravity*, *Adv. Theor. Math. Phys.* **2** (1998) 231 [`hep-th/9711200`] [ɪɴSPIRE].

[7] E. Witten, *Anti-de Sitter space and holography*, *Adv. Theor. Math. Phys.* **2** (1998) 253 [`hep-th/9802150`] [ɪɴSPIRE].

[8] A. Almheiri, N. Engelhardt, D. Marolf and H. Maxfield, *The entropy of bulk quantum fields and the entanglement wedge of an evaporating black hole*, *JHEP* **12** (2019) 063 [`arXiv:1905.08762`] [ɪɴSPIRE].

[9] A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian and A. Tajdini, *The entropy of Hawking radiation*, *Rev. Mod. Phys.* **93** (2021) 035002 [`arXiv:2006.06872`] [ɪɴSPIRE].

[10] G. Penington, S.H. Shenker, D. Stanford and Z. Yang, *Replica wormholes and the black hole interior*, `arXiv:1911.11977` [ɪɴSPIRE].

[11] G. Penington, *Entanglement Wedge Reconstruction and the Information Paradox*, *JHEP* **09** (2020) 002 [`arXiv:1905.08255`] [ɪɴSPIRE].

[12] A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian and A. Tajdini, *Replica Wormholes and the Entropy of Hawking Radiation*, *JHEP* **05** (2020) 013 [`arXiv:1911.12333`] [ɪɴSPIRE].

[13] A. Almheiri, R. Mahajan, J. Maldacena and Y. Zhao, *The Page curve of Hawking radiation from semiclassical geometry*, *JHEP* **03** (2020) 149 [`arXiv:1908.10996`] [ɪɴSPIRE].

[14] J. Camps, *Generalized entropy and higher derivative Gravity*, *JHEP* **03** (2014) 070 [`arXiv:1310.6659`] [ɪɴSPIRE].

[15] X. Dong, *Holographic Entanglement Entropy for General Higher Derivative Gravity*, *JHEP* **01** (2014) 044 [`arXiv:1310.5713`] [ɪɴSPIRE].

[16] S. Ryu and T. Takayanagi, *Holographic derivation of entanglement entropy from AdS/CFT*, *Phys. Rev. Lett.* **96** (2006) 181602 [hep-th/0603001] [ɪɴSPIRE].

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

[18] A. Lewkowycz and J. Maldacena, *Generalized gravitational entropy*, *JHEP* **08** (2013) 090 [arXiv:1304.4926] [ɪɴSPIRE].

[19] T. Faulkner, A. Lewkowycz and J. Maldacena, *Quantum corrections to holographic entanglement entropy*, *JHEP* **11** (2013) 074 [arXiv:1307.2892] [ɪɴSPIRE].

[20] N. Engelhardt and A.C. Wall, *Quantum Extremal Surfaces: Holographic Entanglement Entropy beyond the Classical Regime*, *JHEP* **01** (2015) 073 [arXiv:1408.3203] [ɪɴSPIRE].

[21] A. Belin, A. Lewkowycz and G. Sárosi, *The boundary dual of the bulk symplectic form*, *Phys. Lett. B* **789** (2019) 71 [arXiv:1806.10144] [ɪɴSPIRE].

[22] A. Belin, A. Lewkowycz and G. Sárosi, *Complexity and the bulk volume, a new York time story*, *JHEP* **03** (2019) 044 [arXiv:1811.03097] [ɪɴSPIRE].

[23] J. Kirklin, *The Holographic Dual of the Entanglement Wedge Symplectic Form*, *JHEP* **01** (2020) 071 [arXiv:1910.00457] [ɪɴSPIRE].

[24] A. Uhlmann, *The Metric of Bures and the Geometric Phase*, in *Groups and Related Topics: Proceedings of the First Max Born Symposium*, R. Gielerak, J. Lukierski and Z. Popowicz eds., Springer, Dordrecht The Netherlands (1992), pg. 267 [ISBN:978-94-011-2801-8].

[25] A. Uhlmann, *Gauge field governing parallel transport along mixed states*, *Lett. Math. Phys.* **21** (1991) 229 [ɪɴSPIRE].

[26] A. Uhlmann, *The "transition probability" in the state space of a*-algebra*, *Rep. Math. Phys.* **9** (1976) 273.

[27] S.L. Braunstein and C.M. Caves, *Statistical distance and the geometry of quantum states*, *Phys. Rev. Lett.* **72** (1994) 3439 [ɪɴSPIRE].

[28] R. Jozsa, *Fidelity for Mixed Quantum States*, *J. Mod. Opt.* **41** (1994) 2315.

[29] D. Bures, *An Extension of Kakutani's Theorem on Infinite Product Measures to the Tensor Product of Semifinite w*-Algebras*, *Trans. Am. Math. Soc.* **135** (1969) 199.

[30] C.W. Helstrom, *Minimum mean-squared error of estimates in quantum statistics*, *Phys. Lett. A* **25** (1967) 101.

[31] S.B. Giddings and G.J. Turiaci, *Wormhole calculus, replicas, and entropies*, *JHEP* **09** (2020) 194 [arXiv:2004.02900] [ɪɴSPIRE].

[32] D. Marolf and H. Maxfield, *Transcending the ensemble: baby universes, spacetime wormholes, and the order and disorder of black hole information*, *JHEP* **08** (2020) 044 [arXiv:2002.08950] [ɪɴSPIRE].

[33] D. Marolf and H. Maxfield, *Observations of Hawking radiation: the Page curve and baby universes*, *JHEP* **04** (2021) 272 [arXiv:2010.06602] [ɪɴSPIRE].

[34] J. Yngvason, *The Role of type-III factors in quantum field theory*, *Rept. Math. Phys.* **55** (2005) 135 [math-ph/0411058] [ɪɴSPIRE].

[35] C.G. Callan, Jr. and F. Wilczek, *On geometric entropy*, *Phys. Lett. B* **333** (1994) 55 [hep-th/9401072] [ɪɴSPIRE].

[36] H. Casini and M. Huerta, *Entanglement entropy in free quantum field theory*, *J. Phys. A* **42** (2009) 504007 [arXiv:0905.2562] [ɪɴSPIRE].

[37] R.E. Peierls, *The Commutation laws of relativistic field theory*, *Proc. Roy. Soc. Lond. A* **214** (1952) 143 [INSPIRE].

[38] B.S. DeWitt, *The spacetime approach to quantum field theory*, in *Relativity, groups and topology: Proceedings of 40th Summer School of Theoretical Physics - Session 40. Vol. 2*, Les Houches France (1983), pg. 381.

[39] B.S. DeWitt, *The global approach to quantum field theory. Vol. 1, 2*, *Int. Ser. Monogr. Phys.* **114** (2003) 1.

[40] P.G. Bergmann and R. Schiller, *Classical and Quantum Field Theories in the Lagrangian Formalism*, *Phys. Rev.* **89** (1953) 4 [INSPIRE].

[41] C. Crnković and E. Witten, *Covariant description of canonical formalism in geometrical theories*, in: *Three Hundred Years of Gravitation*, S.W. Hawking and W. Israel eds., Cambridge University Press, Cambridge U.K. (1989), pg. 676.

[42] C. Crnković, *Symplectic Geometry of the Covariant Phase Space, Superstrings and Superspace*, *Class. Quant. Grav.* **5** (1988) 1557 [INSPIRE].

[43] G.J. Zuckerman, *Action principles and global geometry*, *Conf. Proc. C* **8607214** (1986) 259 [INSPIRE].

[44] A. Ashtekar and A. Magnon-Ashtekar, *On the symplectic structure of general relativity*, *Commun. Math. Phys.* **86** (1982) 55.

[45] J. Lee and R.M. Wald, *Local symmetries and constraints*, *J. Math. Phys.* **31** (1990) 725 [INSPIRE].

[46] J.D. Brown and J.W. York, Jr., *Quasilocal energy and conserved charges derived from the gravitational action*, *Phys. Rev. D* **47** (1993) 1407 [gr-qc/9209012] [INSPIRE].

[47] D.M. Marolf, *Poisson brackets on the space of histories*, *Annals Phys.* **236** (1994) 374 [hep-th/9308141] [INSPIRE].

[48] V. Iyer and R.M. Wald, *Some properties of Noether charge and a proposal for dynamical black hole entropy*, *Phys. Rev. D* **50** (1994) 846 [gr-qc/9403028] [INSPIRE].

[49] R.M. Wald and A. Zoupas, *A General definition of 'conserved quantities' in general relativity and other theories of gravity*, *Phys. Rev. D* **61** (2000) 084027 [gr-qc/9911095] [INSPIRE].

[50] G. Barnich and F. Brandt, *Covariant theory of asymptotic symmetries, conservation laws and central charges*, *Nucl. Phys. B* **633** (2002) 3 [hep-th/0111246] [INSPIRE].

[51] S. Hollands and D. Marolf, *Asymptotic generators of fermionic charges and boundary conditions preserving supersymmetry*, *Class. Quant. Grav.* **24** (2007) 2301 [gr-qc/0611044] [INSPIRE].

[52] D. Harlow and J.-Q. Wu, *Covariant phase space with boundaries*, *JHEP* **10** (2020) 146 [arXiv:1906.08616] [INSPIRE].

[53] D. Harlow and P. Hayden, *Quantum Computation vs. Firewalls*, *JHEP* **06** (2013) 085 [arXiv:1301.4504] [INSPIRE].

[54] J. Åberg, D. Kult, E. Sjöqvist and D.K.L. Oi, *Operational approach to the Uhlmann holonomy*, *Phys. Rev. A* **75** (2007) 032106 [quant-ph/0608185].

[55] O. Viyuela, A. Rivas, S. Gasparinetti, A. Wallraff, S. Filipp and M.A. Martin-Delgado, *Observation of topological Uhlmann phases with superconducting qubits*, [arXiv:1607.08778].