Dissertation for the degree of Doctor of Philosophy

Semi-Classical Black Hole Holography

Lukas Schneiderbauer

UNIVERSITY OF ICELAND
FACULTY OF PHYSICAL SCIENCES

School of Engineering and Natural Sciences
Faculty of Physical Sciences
Reykjavík, September 2020

A dissertation presented to the University of Iceland, School of Engineering and Natural Sciences, in candidacy for the degree of Doctor of Philosophy.

Doctoral committee

Prof. Lárus Thorlacius, advisor
Faculty of Physical Sciences, University of Iceland

Prof. Þórður Jónsson
Faculty of Physical Sciences, University of Iceland

Prof. Valentina Giangreco M. Puletti
Faculty of Physical Sciences, University of Iceland

Prof. Bo Sundborg
Department of Physics, Stockholm University

Opponents

Prof. Veronika Hubeny
University of California, Davis

Prof. Mukund Rangamani
University of California, Davis

Semi-classical Black Hole Holography

ISBN: 978-9935-9452-9-7
Author ORCID: 0000-0002-0975-6803
Printed in Iceland by Háskólaprent,
Reykjavík, Iceland, September 2020.

Contents

Abstract

This thesis discusses two aspects of semi-classical black holes.

First, a recently improved semi-classical formula for the entanglement entropy of black hole radiation is examined. This entropy is an indicator of information loss and determines whether black hole evaporation is an information preserving process or destroys quantum information. Assuming information conservation, Page expressed the entanglement entropy as a function of time, which is referred to as the "Page curve." Using the improved formula for evaporating black hole solutions of a gravitational model introduced by Callan, Giddings, Harvey and Strominger (CGHS) and modified by Russo, Susskind and Thorlacius (RST), we find that the entanglement entropy follows the Page curve and thus is consistent with unitary evolution.

Second, the notion of quantum complexity is explored in the context of black holes. The quantum complexity of a quantum state measures how many "simple operations" are required to create that state. Susskind conjectured that the quantum complexity of a black hole state corresponds to a certain volume inside the black hole. A modified conjecture equates the quantum complexity with the gravitational action evaluated for a certain region of spacetime which intersects the black hole interior. We test the complexity conjectures for semi-classical black hole solutions in the CGHS/RST model and find that both conjectures yield the expected behavior.

Ágrip

Í ritgerðinni er fjallað um tvo eiginleika svarthola.

Fyrst skoðum við nýja aðferð við að reikna flækjuóreiðu á milli svarthols og Hawking geislunarinnar sem það sendir frá sér. Þessir útreikningar gefa til kynna hvort skammtaupplýsingar tapist við uppgufun svartholsins eður ei. Ef gert er ráð fyrir að upplýsingarnar varðveitist, þá fylgir flækjuóreiðan svonefndum Page ferli. Við reiknum flækjuóreiðuna, annars vegar í þyngdarfræðilíkani sem kennt er við Callan, Giddings, Harvey og Strominger (CGHS) og hins vegar í líkani kennt við Russo, Susskind og Thorlacius (RST), og sýnum fram á að hún fylgir Page ferlinum í báðum þessum líkönum. Niðurstaðan er í samræmi við það að uppgufun svarthols sé ferli þar sem skammtaupplýsingar tapast ekki.

Við skoðum einnig hugtakið skammtafræðilegt flækjustig í tengslum við svarthol. Flækjustig skammtaástands segir til um hve margar einfaldar aðgerðir eða skref af ákveðinni gerð þarf til þess að mynda ástandið úr tilteknu viðmiðunarástandi. Susskind setti fram þá tilgátu að skammtafræðilegt flækjustig svarthols sé jafngilt tilteknu rúmmáli innan svartholsins. Önnur tilgáta tengir flækjustigið við þyngdarvirkni ákveðins hluta tímarúmsins sem skarast við svartholið. Við prófum þessar tilgátur fyrir svarthol sem gufa upp með Hawking útgeislun í CGHS og RST líkönunum og sýnum að báðar tilgáturnar gefa flækjustig í samræmi við skammtafræði.

Acknowledgements

First and foremost, I would like to express my gratitude to my supervisor Lárus Thorlacius as well as Valentina Puletti for their guidance and help throughout my years as a doctoral candidate.

Special thanks go to the postdocs present during my stay at the University of Iceland, Daniel Fernandez, Friðrik Freyr Gautason, Napat Poovuttikul and Watse Sybesma for lively discussions (about topics not exclusively related to physics) and their valuable feedback. I am especially in debt to Friðrik who helped me translate the abstract into Icelandic in addition to giving me valuable feedback to various parts of this thesis, and Watse who provided his drawing skills to help create some of the figures. It was a delight to work with them as collaborators.

I would also like to thank my doctoral siblings, Juan Angel-Ramelli and Aruna Rajagopal, for all our shared experiences. I wish them all the best for whatever the future may hold for them.

For enduring me throughout as an office mate, I have to thank numerous people who came and went throughout my years in Iceland. I certainly enjoyed the enlightening conversations with Daniel Amankwah, Arnbjörg Soffía Árnadóttir, Hjörtur Björnsson, Jóhann Haraldsson and Auðunn Skúta Snæbjarnarson.

The running group and Bandí team I fortunately joined also deserve mentioning for accompanying me through most of my stay in Iceland. They provided me with a welcome diversion from academic work.

Last but not least, I am grateful to my family and friends. I feel lucky being born to my parents, Grete and Helmut, who accepted and supported my major decisions throughout my life. I want to thank my siblings and their partners, Daniel, David, Eveline, Magdalena and Teresa. They kept my childhood interesting and turned every visit at home into a fun experience. A lot of joy is brought to me by my two nephews, Emil and Oscar, who managed to distract me even while writing these lines.

I am especially thankful to three of my best friends, Dominik, Moritz and Philipp

with whom I spent the majority of my time during my studies. Every visit is a great pleasure to me, and I am lucky to have them in my life.

This work was supported by the Icelandic Research Fund under grants 163422-053 and 195970-051, and by the University of Iceland Research Fund.

1

Introduction

Black holes are objects in space whose mass is so strongly concentrated that their gravitational attraction would not allow anything, not even light, to escape. They are, without doubt, among the most fascinating objects studied by physicists within the last century. Its original discovery as a solution to Einstein's theory of general relativity (GR) by Schwarzschild [1] dates back to 1915. Although initially not taken seriously even by Einstein himself [2], and after a period of confusion about the nature of the event horizon of a black hole [3], there is no doubt to date that black holes are realistic solutions of GR. Not only was it recognized that black holes can be formed within ordinary circumstances, but it has also been realized that the appearance of black holes is actually hard to avoid in nature [4]. As of now, there is also a long list of observational evidence for the existence of black holes. Among the most prominent examples are the detection of gravitational waves [5], which was honored with a nobel prize in 2017, and the popular "Event Horizon Telescope" image [6].

GR is primarily concerned with describing physical phenomena at astronomical scales, that is, distances of order $\gtrsim 10^{11}$ cm. A complementary theory, quantum mechanics, was developed almost in parallel to describe the physical aspects concerning atomic scales $\lesssim 10^{-8}$ cm. With the advent of quantum mechanics into the second half of the 20th century its unification with special relativity resulted in what we know today as quantum field theory (QFT). Although it became clear at an early stage that QFT was the appropriate framework to correctly describe known fundamental particles and their interactions in a unified manner, it does not account for gravity.

This is not a problem for contemporary particle physicists, since gravity is a very weak force and its effect is negligible for usual particle experiments on our planet. However, the gravitational field can in principle become arbitrarily strong, for it only requires a high enough mass density. This leads us back to black holes which represent the archetype of objects with highly concentrated mass. If we want to understand the physics in the neighborhood of a black hole, which undoubtedly exist in nature, we need to understand how quantum mechanics and gravity function together. In other words, we need a unified theory, "quantum gravity."

While finding a full quantum theory of gravity has proven incredibly difficult, one can pursue a more humble path and content oneself with exploring situations where the gravitational effects are stronger than usual but not too strong. In such situations it is sometimes possible to ignore the quantum effects associated with the gravitational degrees of freedom but still incorporate quantum effects of normal matter and their effect on the spacetime geometry. It appears that already in this regime one stumbles across unexpected consequences and interesting puzzles arise.

Around sixty years after the black hole's first discovery, Hawking realized, taking into account quantum effects in the manner discussed above, that a black hole is in fact not black, but it radiates [7]. This radiation is of thermal nature and in line with earlier thermodynamic considerations regarding black holes by Bekenstein [8, 9]. Shortly after, Hawking concluded that the scattering matrix is not unitary and information is lost [10]. At the time this came to a big surprise to the physics community, since unitarity, or loosely speaking, "conservation of information," is a corner stone in the framework of quantum mechanics. Giving up such an important principle can not be easily accepted. This now famous problem was later dubbed the "black hole information paradox," and was since then a primary drive for theoretical physicists to study black holes in all their varieties.

Even though Hawking initially proposed to give up unitarity [11], major problems arose when trying to generalize quantum mechanics to incorporate non-unitary evolution [12, 13]. Strong hints were given by the discovered duality between Anti-de Sitter space (AdS) and conformal quantum field theory (CFT), in short AdS/CFT, in 1997 [14]. AdS/CFT describes a correspondence between two completely different theories. On one hand, we have a theory of gravity, GR including a negative cosmological constant, while, on the other hand, we have a usual quantum mechanical theory, a QFT. The AdS/CFT correspondence states that these theories are equivalent, and one can map a particular problem on one side to another problem on the other side. Now, assuming the AdS/CFT correspondence is correct, since the evolution of the QFT is

manifestly unitary, it also must hold true for the gravitational theory.

Even though AdS/CFT states that black holes in AdS do not destroy information, that does not provide an answer to the question for what is wrong with Hawking's calculation, even in an AdS/CFT setup, nor does it give definite answers about how the required information escapes the black hole.

In the AdS/CFT context, a recent breakthrough was achieved by computing the Page curve [15, 16] using semi-classical methods, by coupling an asymptotically AdS black hole to a CFT reservoir [17, 18]. As we will discuss in more detail in chapter 3, the Page curve quantifies the correlations between the black hole and its radiation and is a good indicator for unitarity. The result uses the Quantum Ryu-Takayanagi (QRT) formula [19–22] and the outcome exactly follows Page's prediction assuming unitarity [15, 16]. It is the first time that it was accomplished to compute the Page curve by semi-classical methods alone. Although the QRT formula is motivated by AdS/CFT technology, there is evidence that it is actually valid even outside the AdS/CFT framework. It is claimed [23–25] that this formula can be derived using the Replica trick [26, 27] without making reference to a holographic dual system. The crucial ingredient is the inclusion of non-trivial topologies, i.e. Euclidean wormholes when evaluating the functional integral over Euclidean manifolds. On a practical level the QRT formula leads one to include the correlations from the radiation degrees of freedom in the interior of the Black Hole after the Page time. From this point of view, it is perhaps not surprising that this formula reproduces the Page curve, since the disappearance of information into the black hole interior was what caused the problem in the first place [7]. However, the point is that this instruction came from a semi-classical computation and the fact that the semi-classical geometry has information about these details *is* surprising.

Even if this calculation still does not reveal how information escapes the black hole, it is at least another indicator that understanding the interior of the black hole is of major importance. Independently of the investigation of the Page curve, Susskind initiated a study of the black hole interior via the notion of quantum complexity [28–30]. Quantum complexity defines a metric on the space of states, which in some sense captures an intuitive understanding of the difficulty to create a state by "simple operations" from another reference state. This is best illustrated by imagining a quantum computer consisting of a set of gates, which are modeled by unitary operators acting on a Hilbert space. These particular set of gates need a certain amount of applications to reach from an initial input state to a final output state. The number of (minimum) applications is what we call complexity, and it quantifies how "hard" it

is for the computer to compute a certain state provided some input state. Susskind conjectured that the growth of complexity of a black hole state should be related to the volume growth of the interior of the black hole. We will review these arguments in chapter 4. One of the initial motivations to study complexity was to argue against a "typical" occurrence of firewalls [31]. The term "firewall" [32, 33] refers to a hypothetical violent region at the event horizon of a black hole for an infalling observer, and thus violating the equivalence principle. It was proposed in [32] so to resolve an apparent paradox within *black hole complementarity* [34]. Susskind reasons [35] that a "typical" state has maximal complexity. Since the complexity of black hole states that are formed by collapse is increasing for an exponentially long time (in its entropy), they are not "typical," and hence there is no reason to expect a firewall.

Outline

In this thesis we expand on the above ideas in the context of asymptotically flat black holes in a two dimensional dilaton gravity theory which we will describe in more detail in chapter 2.

Since the QRT formula was originally formulated in the context of AdS/CFT, it is an interesting question whether it also gives sensible results for asymptotically flat black holes. The two dimensional models we are considering incorporate back-reaction of the Hawking radiation to the geometry. This enables us, at least in principle, to directly compute the entanglement entropy which include effects such as the reduction of the black hole mass over time. In chapter 3 we review Page's argument which leads to the Page curve and elaborate on the proposed entropy formula. We evaluate the entanglement entropy on the semi-classical two-dimensional backgrounds in [36] and find perfect agreement with Page's prediction.

In chapter 4 we revisit the notion of quantum complexity. Although complexity was originally studied for classical black holes in asymptotic AdS spacetime, we analyse Susskind's ideas in the semi-classical regime of asymptotically flat black holes [37, 38]. A part of the problem is to precisely define the supposedly dual quantities. With our definitions we find positive evidence that supports his conjectures even in these cases. That such an identification is also possible in asymptotically flat spacetimes provides further evidence for black hole complementarity [34] which postulates that the black hole can be described by a quantum system with a finite number of degrees of freedom living near (and outside) the event horizon.

This thesis is based on three research articles,

I *Holographic Complexity: Stretching the Horizon of an Evaporating Black Hole* by Lukas Schneiderbauer, Watse Sybesma, and Lárus Thorlacius, published in the Journal of High Energy Physics, volume 3 (2020),

II *Action Complexity for Semi-Classical Black Holes* by Lukas Schneiderbauer, Watse Sybesma, and Lárus Thorlacius, published in the Journal of High Energy Physics, volume 7 (2020),

III *Page Curve for an Evaporating Black Hole* by Friðrik Freyr Gautason, Lukas Schneiderbauer, Watse Sybesma, and Lárus Thorlacius, published in the Journal of High Energy Physics, volume 5 (2020),

The articles are attached at the end of this thesis with copyright permission from JHEP.

2

Black Hole Model

In this thesis we primarily work with a two dimensional dilatonic gravity model introduced by Callan, Giddings, Harvey and Strominger (CGHS) [39] and a variant thereof by Russo, Susskind and Thorlacius (RST) [40]. One of the main advantages coming with the RST model is that the semi-classical equations can be solved explicitly as we show below. The CGHS model can be regarded as an effective description for radial modes of near-extremal dilatonic magnetically charged black holes in four or higher dimensions [41]. While this aspect certainly provides an interesting point of view, in this work it will usually be sufficient to interpret the results from a purely two dimensional perspective.[1] In this section we will review some aspects of the CGHS/RST models while trying to be complementary to the introductions in [36–38] and the already existing review by Thorlacius [42].

The classical CGHS action I_0 is given by

$$I_0 = I_{\text{grav}} + I_{\text{matter}} \tag{2.1}$$

with

$$I_{\text{grav}} = \frac{1}{2\pi}\int d^2x\sqrt{-g}\, e^{-2\phi}\left(R + 4\left(\nabla\phi\right)^2 + 4\lambda^2\right), \tag{2.2}$$

$$I_{\text{matter}} = -\frac{1}{4\pi}\int d^2x\sqrt{-g}\sum_{i=1}^{N}\left(\nabla f_i\right)^2, \tag{2.3}$$

[1] This is going to be true apart from an isolated case, when we define the volume functional which represents complexity, see [37].

and depends on the fields ϕ , the dilaton, the metric $g_{\mu\nu}$, and the N free matter fields f_i minimally coupled to the gravitational system. As usual, $g := \det g_{\mu\nu}$, and R is the Ricci scalar. A natural scale is provided by the parameter λ related to the magnetic charge of the higher dimensional theory. We take length to be measured in units of λ^{-1} and set $\lambda = 1$.

The strength of the classical gravitational coupling is characterized by $G_{(2)} := \frac{1}{8}e^{2\phi}$ which is allowed to vary over spacetime. If the gravitational coupling $G_{(2)}$ is small compared to N, and N much larger than 24,

$$e^{-2\phi} \gg \frac{N}{24} \gg 1, \tag{2.4}$$

we can attempt a semi-classical treatment [39]. These conditions are necessary if we want to reliably ignore quantum effects induced by the dilaton and metric field since they will be dominated by the one-loop effects of N matter fields. Considering that the dilaton ϕ varies over spacetime, it is not clear a priori that these conditions can be satisfied and the solutions of the semi-classical equations of motion have to be checked for consistency. When discussing black hole solutions it turns out that the gravitational coupling $G_{(2)}$ goes to zero in the asymptotic region, but increases as one approaches the black hole singularity. One can, however, make the coupling arbitrary small at its event horizon by considering a black hole of large enough mass. One should then be able to trust the semi-classical solution at least everywhere outside the black hole as long as its mass is still large enough.[2]

We review the inclusion of quantum corrections in section 2.1. From the quantum effective action we derive an expression for the stress tensor in section 2.2 and discuss its properties. Further, in order to describe dynamical black hole creation we shortly examine the concept of coherent states in chapter 2.3. After inclusion of the RST term, discussed in 2.4, we consider the equations of motion corresponding to the quantum effective action in section 2.5. Finally, black hole thermodynamics in the context of the CGHS model is discussed in paragraph 2.6.

2.1 The Polyakov term

To calculate the correction I_1 to the classical action induced by the matter fields f_i, we can use a well known argument [43–45] which goes as follows. The crucial ingredient

[2] A classical black hole retains its initial mass for all times. However, quantum effects lead to Hawking radiation and the black hole slowly evaporates.

is the fact that the free matter fields f_i constitute a CFT on a fixed background. Formally, the correction I_1 is given by the functional integral

$$e^{iI_1} := \int \mathcal{D}f\, e^{iI_{\text{matter}}}, \tag{2.5}$$

where we integrate over all matter fields f_i over a two-dimensional manifold with fixed metric $g_{\mu\nu}$.

By the definition of the stress tensor $T_{\mu\nu}$, its expectation value is given by

$$\langle T_{\mu\nu}\rangle = -\frac{4\pi}{\sqrt{-g}}\frac{\delta I_1}{\delta g^{\mu\nu}}. \tag{2.6}$$

We will exploit this identity to deduce the form of I_1. In a classical CFT the stress tensor is traceless. However, when quantizing the theory in a way compatible with conformal symmetry one encounters that the trace of the stress tensor operator acquires a non-zero expectation value, the so-called trace anomaly. In two dimensions its form is completely fixed by the central charge c of the CFT (in the present case $c = N$) and is given by [46]

$$\langle T^{\mu}{}_{\mu}\rangle = -\frac{4\pi}{\sqrt{-g}}g^{\mu\nu}\frac{\delta I_1}{\delta g^{\mu\nu}} = \frac{N}{12}R, \tag{2.7}$$

with R being the Ricci scalar and N the number of scalar fields f_i which equals the central charge c of the CFT. In conformal coordinates (y^+, y^-), in which the line element reads

$$ds^2 = -e^{2\rho}dy^+dy^-, \tag{2.8}$$

the Ricci scalar can be expressed as $R = 8e^{-2\rho}\partial_+\partial_-\rho$, and so equation (2.7) becomes

$$\frac{\delta I_1}{\delta \rho} = \frac{N}{6\pi}\partial_+\partial_-\rho. \tag{2.9}$$

This equation is solved by[3]

$$I_1 = \frac{N}{12\pi}\int dx^+dx^-\rho\partial_+\partial_-\rho = -\frac{N}{96\pi}\int d^2x\sqrt{-g}R\frac{1}{\nabla^2}R, \tag{2.10}$$

where we write

$$\left(\frac{1}{\nabla^2}R\right)(x) := \int d^2y\sqrt{-g(y)}G(x,y)R(y), \tag{2.11}$$

[3]Given any solution we can of course add terms whose variation is zero as we vary the conformal factor ρ and the result still solves equation (2.9). We will come back to this issue later on.

with $G(x)$ being the Green's function of the Laplacian $\nabla^2 = \nabla_\mu \nabla^\mu$ defined by

$$\sqrt{-g}\nabla^2 G(x) = \delta(x). \tag{2.12}$$

It is a peculiarity of I_1 that the covariant expression in (2.10) is obviously non-local, while in conformal gauge the term reduces to a single integral involving the conformal factor.

For later convenience let us introduce a slightly shorter notation,

$$Z(x) := -\left(\frac{1}{\nabla^2} R\right)(x). \tag{2.13}$$

Equation (2.12) does not define the Green's function G uniquely. It is only determined up to a harmonic function. This is simply a manifestation of the fact that the Laplace operator ∇^2 has a non-trivial kernel on non-compact manifolds. The undetermined harmonic function has to be fixed by imposing boundary conditions.

Let us render these statements more explicit by choosing conformal gauge (2.8). In order to compare Green's function with different boundary conditions, let us evaluate the field Z for a general Green's function $G(y^\pm) = G_0(y^\pm) + h_+(y^+) + h_-(y^-)$. The function $h(y^\pm) = h_+(y^+) + h_-(y^-)$ is harmonic by construction and contains information about the boundary conditions of G. We have

$$Z(y^\pm) = -4 \int dz^+ dz^- G(z^\pm, y^\pm) \partial_+ \partial_- \rho(z^\pm). \tag{2.14}$$

Naively, one would perform repeated integration by parts and conclude that the harmonic function h has no influence on Z. However, this conclusion is not correct, since the function ρ has in general not the required falloff behavior to justify integration by parts. In general, one has

$$Z(y^\pm) = 2\rho(y^\pm) + \eta(y^\pm), \tag{2.15}$$

with $\eta(y^\pm) = \eta_+(y^+) + \eta_-(y^-)$ being another harmonic function, characterized by boundary conditions.

Comparing this to the discussion before, especially equation (2.10), we conclude that the equal sign in (2.10) is only correct in a very specific conformal coordinate system, that is to say, in the coordinate system where $h \equiv \eta \equiv 0$. For reasons becoming transparent shortly, we will refer to this coordinate system as "vacuum coordinates." One might ask, whether this is not at odds with our derivation of the

Polyakov term (2.10) in the first place, since nowhere did we assume a particular coordinate system (apart from using a conformal coordinate system). The answer is no, since when integration equation (2.9) we assumed there are no extra terms in I_1 whose variation is zero. This assumption turns out to be simply incorrect when considering general boundary conditions and the missing piece is automatically restored, when writing the Polyakov term in a covariant way. To see this more explicitly let us consider the Polyakov term with general boundary conditions in conformal gauge (2.8),

$$-\frac{N}{96\pi}\int d^2x\sqrt{-g}R\frac{1}{\nabla^2}R=-\frac{N}{12\pi}\int dx^+dx^-\left(\rho+\frac{1}{2}\eta\right)\partial_+\partial_-\rho. \tag{2.16}$$

The first order variation of the second term is simply zero,

$$\frac{\delta}{\delta\rho(y)}\int dx^+dx^-\eta\partial_+\partial_-\rho=\int dx^+dx^-\eta\partial_+\partial_-\delta(x-y)=0, \tag{2.17}$$

because of harmonicity of the function η. Therefore the more general expression (2.16) still solves equation (2.9). Crucially, the term itself, $\int dx^+dx^-\eta\partial_+\partial_-\rho$, is *not* zero in general.

2.2 The stress tensor

To substantiate our conclusion from previous paragraph, it is instructive to derive the expectation value (2.6) for all components of the stress tensor $T_{\mu\nu}$ directly from the Polyakov term I_1. There are two ways to do this. Either we vary the non-local term in (2.10) directly, or we introduce an auxiliary field in order to render the non-local term effectively local. Both lead to identical results, but the second way is technically simpler, so let us pursue this route. One replaces the Polyakov term by

$$\frac{N}{48\pi}\int d^2x\sqrt{-g}\left(ZR-\frac{1}{2}\left(\nabla Z\right)^2\right), \tag{2.18}$$

where the field Z is now promoted to a dynamical variable. Extremizing this action with respect to Z yields the equation of motion

$$\nabla^2Z+R=0, \tag{2.19}$$

and its solution coincides with our definition of Z in (2.13). That is, by "integrating out" the field Z one recovers the previous non-local Polyakov term (2.10).

It is now possible to calculate the response of the action to the variation of the metric components $g^{\mu\nu}$. One observes that by using the equation of motion (2.19) after carrying out the variation, that the result does not depend on the variation of Z itself,

$$\delta I_1 = \frac{N}{48\pi}\int d^2x \left(\delta\left(\sqrt{-g}R\right) Z - \frac{1}{2}\delta\left(\sqrt{-g}g^{\alpha\beta}\right)\partial_\alpha Z\partial_\beta Z\right), \tag{2.20}$$

so one only needs to be concerned about the variation of the metric determinant $\sqrt{-g}$ and the Ricci scalar R. These are well known [47], and the results we need can be written as

$$\begin{aligned}\delta\left(\sqrt{-g}R\right) &= \sqrt{-g}\left(g_{\mu\nu}g^{\sigma\lambda} - \delta^\sigma_\mu\delta^\lambda_\nu\right)\nabla_\sigma\nabla_\lambda\delta g^{\mu\nu},\\ \delta\left(\sqrt{-g}g^{\alpha\beta}\right) &= \sqrt{-g}\left(-\frac{1}{2}g^{\alpha\beta}g_{\mu\nu} + \delta^\alpha_\mu\delta^\beta_\nu\right)\delta g^{\mu\nu}.\end{aligned} \tag{2.21}$$

Thus the final result in terms of the auxiliary field Z is

$$\begin{aligned}\langle T_{\mu\nu}\rangle &= -\frac{4\pi}{\sqrt{-g}}\frac{\delta I_1}{\delta g^{\mu\nu}}\\ &= \frac{N}{24}\left(-\frac{1}{2}g_{\mu\nu}\left(\nabla Z\right)^2 + \nabla_\mu Z\nabla_\nu Z - 2g_{\mu\nu}\nabla^2 Z + 2\nabla_\mu\nabla_\nu Z\right).\end{aligned} \tag{2.22}$$

As a consistency check, we can calculate the trace to confirm the conformal anomaly,

$$\langle T^\mu_{\ \mu}\rangle = -\frac{N}{12}\nabla^2 Z = \frac{N}{12}R, \tag{2.23}$$

which is reproduced correctly as expected. More interestingly we can evaluate the diagonal components $\langle T_{\pm\pm}\rangle$ of the stress tensor in conformal gauge (2.8), since they correspond to physical energy flux. Using (2.22) and the general solution for Z in conformal gauge (2.15), we get

$$\langle T_{\pm\pm}\rangle = \frac{N}{24}\left(4\partial^2_\pm\rho - 4\left(\partial_\pm\rho\right)^2 + 2\partial^2_\pm\eta + \left(\partial_\pm\eta\right)^2\right). \tag{2.24}$$

This is in line with an alternative derivation of this expression going back to [43], which is also discussed in the original CGHS paper [39]. In [39, 40, 42, 43], the piece $\left(\partial_\pm\eta\right)^2 + 2\partial^2_\pm\eta$ appears as integration "constants" $t_+(y^+)$ and $t_-(y^-)$ when integrating the conservation equations for the stress tensor $T_{\mu\nu}$. Again, these integration constants have to be fixed by imposing boundary conditions. The above analysis shows how the functions $t_\pm$ are related to the function η originating from boundary conditions of the Green's function G in (2.12).

We can now also justify why we called the coordinates with $\eta = 0$ the "vacuum

coordinates." In manifestly flat spacetime the derivatives of the conformal factor are zero, $\partial_\pm \rho = 0$, and the energy momentum flux is purely given in terms of η,

$$\langle T_{\pm\pm} \rangle = t_\pm \equiv \frac{N}{24} \left((\partial_\pm \eta)^2 + 2\partial_\pm^2 \eta \right), \tag{2.25}$$

so that if $\eta = 0$, $\langle T_{\pm\pm} \rangle = 0$, and no energy momentum flux is detected by an inertial observer.

Another interpretation can be given to the function η by realizing that $t_\pm$ can be written as a Schwarzian derivative $\{f(x); x\} := \left(\frac{f''}{f'}\right)' - \frac{1}{2}\left(\frac{f''}{f'}\right)^2$,

$$(\partial_\pm \eta)^2 + 2\partial_\pm^2 \eta = -2 \left\{ z^\pm(x^\pm); x^\pm \right\}, \tag{2.26}$$

with new coordinates $z^\pm(x^\pm)$ given by

$$z^\pm(x^\pm) := \int^{x^\pm} dy^\pm e^{-\eta_\pm(y^\pm)}, \tag{2.27}$$

so that $\ln \frac{\partial x^\pm(z^\pm)}{\partial z^\pm} = \eta_\pm(x^\pm(z^\pm))$. It is then readily seen that under the coordinate transformation $x^\pm \mapsto z^\pm(x^\pm)$ the conformal factor transforms as

$$\rho \mapsto \rho + \frac{1}{2} \ln \frac{\partial x^+(z^+)}{\partial z^+} + \frac{1}{2} \ln \frac{\partial x^-(z^-)}{\partial z^-} = \rho + \frac{1}{2}\eta. \tag{2.28}$$

As such, specifying η in some coordinates is equivalent to specifying the relation to the vacuum coordinates, i.e. the coordinates where the energy momentum flux vanishes asymptotically. This also provides a direct interpretation for the scalar field Z given in (2.15). It represents the value (more precisely, two times the value) of the conformal factor in vacuum coordinates.

We already mentioned before, it is not necessary to introduce the auxiliary variable Z. One can also vary the non-local expression (2.10) directly. Following [48], one can use equation (2.12) to deduce the response of the Green's function G to a change of the metric. Since the Dirac delta function does not depend on the metric, we have

$$\delta \left(\sqrt{-g} \nabla^2 G \right) = 0. \tag{2.29}$$

After some algebraic manipulations of equation (2.29), one concludes that

$$\frac{\delta G(y,z)}{\delta g^{\mu\nu}(x)} = \sqrt{-g(x)} \left(-\frac{1}{2} g^{\alpha\beta}(x) g_{\mu\nu}(x) + \delta^\alpha_\mu \delta^\beta_\nu \right) \nabla^x_\beta G(x,z) \nabla^x_\alpha G(y,x), \tag{2.30}$$

and by using that

$$\begin{aligned}\delta I_1 = &-\frac{N}{48\pi}\int d^2x\int d^2y\sqrt{-g(y)}R(y)G(x,y)\delta\left(\sqrt{-g(x)}R(x)\right)\\ &-\frac{N}{96\pi}\int d^2x\int d^2y\sqrt{-g(x)}R(x)\sqrt{-g(y)}R(y)\delta G(x,y),\end{aligned} \tag{2.31}$$

one recovers the same result (2.22).

2.3 Coherent matter states

To dynamically create large macroscopic black holes, one needs a macroscopic particle beam consisting of a large number of f-quanta. To still have analytic control one models this particle beam as a coherent state built on top of the vacuum $|0\rangle$ corresponding to an inertial observer in flat spacetime. Following [49], a left-moving coherent state can for example be written as

$$|f^c\rangle \propto :\exp\left(\frac{i}{\pi}\sum_{i=1}^{N}\int d\sigma^+\partial_+ f_i^c(\sigma^+)\hat{f}_i(\sigma^+)\right):|0\rangle, \tag{2.32}$$

where $\sigma^\pm$ are the vacuum coordinates and f_i^c is the classical left-moving matter profile. The $\hat{f}_i s$ are the quantum field operators corresponding to left-moving f_i-quanta, and :: denotes normal ordering with respect to the vacuum state $|0\rangle$.

The important properties of such a state are

$$\langle f^c|\hat{f}_i|f^c\rangle = f_i^c, \tag{2.33}$$

and in particular

$$\langle f^c|:\hat{T}_{\mu\nu}:|f^c\rangle = T^c_{\mu\nu} + \langle 0|:\hat{T}_{\mu\nu}:|0\rangle, \tag{2.34}$$

where $T^c_{\mu\nu}$ is the classical energy momentum tensor calculated from the fields f_i^c.

If these quantities are sufficiently large, i.e. much larger than N, then one can easily incorporate the coherent state in our semi-classical analysis by perturbing around a classical background field f_i^c, which obeys the classical equations of motion. Explicitly, the semi-classical quantum effective action is given by

$$I[g,\phi,f^c] = I_{\text{grav}}[g,\phi] + I_{\text{matter}}[g,f^c] + I_1[g]. \tag{2.35}$$

The coordinate dependence of the asymptotic energy flux is still automatically taken

care of in the way it was presented in section 2.2.

The form of the coherent state (2.32) is only meaningful if the scalar fields f_i are free. However, even if the fields f_i were not free, the quantum effective action (2.35) still makes sense by perturbing around a field configuration f_i^c which obeys the (in this case non-linear) classical equations of motion. When calculating the expectation value of the stress tensor $\langle T_{\mu\nu} \rangle$ one could then still interpret the result in the spirit of (2.34) by identifying the first and second term with the leading order and next to leading order outcome in an expansion around $Ne^{2\phi} \to 0$.

2.4 The RST term

Unfortunately, the equations of motion obtained by extremizing the quantum effective action (2.35) are hard to solve and no general solution is known. However, numerical studies of this model exist [50–53]. Another way forward was to slightly modify the model, as for instance was done in [40, 54, 55]. Here we follow RST [40] who modify the gravitational sector of the theory by adding the term

$$I_2 = \frac{N}{48\pi} \int d^2x \sqrt{-g}\, \phi R \tag{2.36}$$

to the effective action. This term is of order N which is of the same order as the Polyakov term I_1 and therefore does not disturb the classical physics taking place when $Ne^{2\phi} \to 0$.

The effect of this term is to restore a classical symmetry which allows to choose a coordinate system in which the conformal factor ρ equals the dilaton, $\phi = \rho$. We will refer to this choice of coordinates as "Kruskal coordinates." This feature allows the equations of motion to be solved explicitly.

2.5 Equations of motion

Finally, the full semi-classical action we are working with can be written down. It is given by

$$I[g, \phi, f^c] = I_{\text{grav}}[g, \phi] + I_2[g, \phi] + I_{\text{matter}}[g, f^c] + I_1[g]. \tag{2.37}$$

By extremizing this action with respect to the metric components, the dilaton, and the matter fields we obtain the full set of equations of motion for the semi-classical

gravitational system,

$$\left(2g_{\mu\nu}\nabla^2 - 2\nabla_\mu\nabla_\nu\right)\left(e^{-2\phi} + \frac{N}{24}\phi\right) + \\ +4e^{-2\phi}\left(2\nabla_\mu\phi\nabla_\nu\phi - g_{\mu\nu}\left((\nabla\phi)^2 + \lambda^2\right)\right) = T^c_{\mu\nu} + \langle T_{\mu\nu}\rangle, \tag{2.38}$$

$$\nabla^2\phi - (\nabla\phi)^2 + \lambda^2 + \frac{1}{4}R = 0, \tag{2.39}$$

$$\nabla^2 f^c_i = 0, \tag{2.40}$$

where $\langle T_{\mu\nu}\rangle$ is given by (2.22) and we temporarily reinstated the scale λ.

To simplify the equations, one usually proceeds in conformal gauge (2.8) and further chooses Kruskal coordinates where the conformal factor ρ equals the dilaton ϕ [39, 40]. To show that this is possible, one takes the trace of equation (2.38),

$$8e^{-2\phi}(\nabla\phi)^2 - 4e^{-2\phi}\nabla^2\phi + \frac{N}{12}\nabla^2\phi - 8e^{-2\phi}\lambda^2 = \frac{N}{12}R, \tag{2.41}$$

and combines it with the equation of motion for the dilaton (2.39) to get

$$\left(e^{-2\phi} + \frac{N}{24}\right)\left(\nabla^2\phi + \frac{1}{2}R\right) = 0. \tag{2.42}$$

Assuming $e^{-2\phi} + \frac{N}{24} \neq 0$, which is always true in the semi-classical regime $e^{-2\phi} \gg \frac{N}{24}$, the equation reads, in conformal gauge,

$$\partial_+\partial_-(\phi - \rho) = 0, \tag{2.43}$$

and is solved by $\phi(y^+, y^-) = \rho(y^+, y^-) + \omega_+(y^+) + \omega_-(y^-)$. One then observes that the functions ω_+ and ω_- can always be eliminated by a conformal coordinate transformation $y^\pm \mapsto x^\pm(y^\pm)$, so that in the coordinates (x^+, x^-) we are left with

$$\phi(x^+, x^-) = \rho(x^+, x^-). \tag{2.44}$$

This leads us to the simple equations of motion in Kruskal gauge,

$$\partial^2_\pm\left(e^{-2\phi} + \frac{N}{24}\phi\right) = \frac{1}{2}\left(T^c_{\pm\pm} + t_\pm\right), \tag{2.45}$$

$$\partial_+\partial_-\left(e^{-2\phi} + \frac{N}{24}\phi\right) = -\lambda^2, \tag{2.46}$$

$$\partial_+\partial_- f^c = 0. \tag{2.47}$$

These equations can be solved for the field $\Omega := e^{-2\phi} + \frac{N}{24}\phi$ provided a general classical matter profile $T^c_{\pm\pm}$ by specifying the asymptotic energy flux $t_\pm$ in Kruskal coordinates.

2.6 Black hole thermodynamics

In order to work out thermodynamic relations to leading order we will concentrate on equilibrium solutions in the classical limit $Ne^{2\phi} \to 0$. It is also possible to calculate corrections in the RST model, but these will not be of major importance here.

We consider a canonical ensemble for the system of the gravitational field $g_{\mu\nu}$, the dilaton ϕ and the matter fields f_i at temperature $T \equiv \beta^{-1}$ within a cavity of large size, represented by the variable V, and prescribed boundary conditions for the dilaton field $\phi\big|_{\text{bdry}} = \phi^B$ and the matter fields $f_i\big|_{\text{bdry}} = f_i^B$. The partition function is then formally given by the Euclidean path integral [56] with Euclidean action I_E, where the imaginary time coordinate is periodically identified with a period β,

$$Z(\beta, V, \phi^B, f_i^B) := \int \mathcal{D}g_{\mu\nu}\, \mathcal{D}\phi\, \mathcal{D}f\, e^{-I_E[g_{\mu\nu},\phi,f]}. \tag{2.48}$$

In the following we will only consider a cavity of infinite size $V \to \infty$. This step involves some subtleties which we will ignore for the sake of brevity. A more detailed analysis is for instance given in [57–59]. Let us from now on suppress the arguments but keep in mind that the expressions are parametrized by β, ϕ^B and f_i^B.

In the classical approximation, valid for $e^{-2\phi} \gg \frac{N}{12}$, the leading order result is just given by the saddle points obeying the given boundary conditions, that is

$$e^{-\beta F} = Z \approx e^{-I_E[g_{\text{cl}},\phi_{\text{cl}},f_{\text{cl}}]} \tag{2.49}$$

where we introduced the free energy F. Using the free energy, one can then utilize the usual thermodynamic relations to calculate the entropy s,

$$s = \beta^2 \partial_\beta F, \tag{2.50}$$

where the derivative is taken keeping the parameters V, ϕ^B and f_i^B fixed.

It is well known, that for arbitrary inverse temperature β, the metric $g_{\mu\nu}$ typically features conical singularities [60]. However, the conical singularities disappear for a special temperature, the Hawking temperature T_H. When using equation (2.50) to compute the entropy of a black hole with Hawking temperature T_H, it is obviously

not enough to know $F(\beta)$ only at $\beta = \beta_H$, and so it is argued in [60] that one should still take contributions from conical singularities into account. We shall follow this approach here.

A one-parameter family of static black hole solutions in the Euclidean classical CGHS model (in units such that $\lambda = 1$) is given by

$$ds^2 = e^{2\phi} dz d\bar{z}, \tag{2.51}$$

$$e^{-2\phi} = M + z\bar{z}, \tag{2.52}$$

$$f_i = 0, \tag{2.53}$$

with the complex coordinate z being related to asymptotic Euclidean time τ as $z = re^{2\pi i \frac{\tau}{\beta}}$, from which the periodic identification $\tau \to \tau + \beta$ is obvious. The connection to the Lorentzian solution is established by interpreting $t = \frac{2\pi}{\beta} i\tau$ as a real-time coordinate.

The ADM mass $\mathcal{M}$ of the black hole is related to the parameter M by $M = \pi\mathcal{M}$ [61]. Hence we can and will refer to M as the black hole mass. The ADM mass coincides with the usual Komar integral in Lorentzian signature,[4]

$$\mathcal{M} = -\lim_{r\to\infty} \frac{1}{16\pi G_{(2)}} \star d\xi, \tag{2.54}$$

if one keeps in mind the relationship between the effective coupling $G_{(2)}$ and the dilaton ϕ, $8G_{(2)} = e^{2\phi}$. The one-form $\xi = -r^2 e^{2\phi} dt$ is dual to a timelike Killing vector normalized such that $\xi^\mu \xi_\mu \to -1$ as $r \to \infty$.

In (r, τ) coordinates the metric reads

$$ds^2 = e^{2\phi} \left(dr^2 + \frac{2\pi}{\beta} r^2 d\tau^2 \right), \tag{2.55}$$

which makes the appearance of a conical singularity for $\beta \neq \beta_H := 2\pi$ at $r = 0$ manifest. A peculiarity of the CGHS model becomes evident. The Hawking temperature does not depend on the mass M. This is atypical, as for a higher dimensional black hole the Hawking temperature always depends on its mass. A prominent example is the Hawking temperature of a four dimensional Schwarzschild black hole with mass M which behaves like $T_H \sim M^{-1}$. The fact that the Hawking temperature does not depend on the mass also implies that no smooth solutions (i.e. without conical

[4] In general the Komar integral is an integral over a co-dimension two manifold. In two dimension this is just a point and so no integral appears in formula (2.54).

singularities) for temperatures other than the Hawking temperature exist.

To calculate the free energy F of this geometry, we need to evaluate the Euclidean on-shell action on a regularized manifold with boundary,

$$-I_E = \frac{1}{2\pi}\int d^2x\sqrt{g}\, e^{-2\phi}\left(R + 4\left(\nabla\phi\right)^2 + 4\right) + \frac{1}{\pi}\int d\Sigma\sqrt{h}\, e^{-2\phi}\left(K-2\right), \quad (2.56)$$

where the bulk action is supplemented by the Gibbons-Hawking-York term [56, 57] on the boundary to render the variational principle on a manifold with boundary well defined. The determinant of the induced metric on the boundary parametrized by Σ is denoted by h while K is the extrinsic curvature of the boundary. The second term $\frac{2}{\pi}\int d\Sigma\sqrt{h}e^{-2\phi}$ is needed to obtain a finite result.

The treatment of the conical singularity is simplified using the analysis in [62]. By regularizing the cone one can evaluate the Ricci scalar R, and after removing the regulator one obtains, in (r,τ) coordinates,

$$R(r) = R_{\text{reg}}(r) + 2\left(1 - \frac{\beta}{\beta_H}\right)\delta(r), \quad (2.57)$$

where $R_{\text{reg}}(r)$ is the regular Ricci scalar of the geometry with $\beta = \beta_H$.

One finds that the Euclidean action I_E, given by (2.56), vanishes when evaluated for the geometry with Ricci scalar R_{reg}, so that the free energy $F(\beta,\phi^B)$ simply reads

$$F(\beta,\phi^B) = \frac{1}{\pi}e^{-2\phi}\Big|_{r=0}\left(\frac{\beta}{\beta_H} - 1\right). \quad (2.58)$$

Notice that we suppressed the dependence on the boundary condition for the matter fields f^B due to the only nontrivial static case being $f^B = 0$, c.p. (2.53). Further, due to (2.52), fixing the value of the dilaton ϕ^B at the boundary essentially fixes the mass M of the black hole solution.

Using (2.50), the thermodynamic entropy of the black hole at the Hawking temperature $T_H = \frac{1}{2\pi}$ yields

$$s = \beta^2\partial_\beta F\Big|_{\beta=\beta_H} = 2e^{-2\phi}\Big|_{r=0} = 2M. \quad (2.59)$$

Remarkably, this result is in line with Bekenstein's expression [8], $s = \frac{A}{4G_{(2)}}$, where A

represents the area of the black hole horizon, since

$$G_{(2)}\Big|_{\text{horizon}} = \frac{1}{8}e^{2\phi}\Big|_{r=0}, \tag{2.60}$$

and $A = 1$. We also note that the vanishing of the free energy $F(\beta_H)$ at the Hawking temperature is consistent with the thermodynamic relation $F(T) = U - Ts$, since the internal energy U is given by

$$U = (1 + \beta\partial_\beta)\, F_M(\beta)\Big|_{\beta_H} = \frac{1}{\pi}e^{-2\phi}\Big|_{r=0} = M/\pi = \mathcal{M} \tag{2.61}$$

and $T_H s = M/\pi$.

To conclude we collect our findings of the thermodynamics variables for the classical eternal CGHS black hole of mass M,

$$T_H = \frac{1}{2\pi}, \tag{2.62}$$

$$U = \frac{1}{\pi}M, \tag{2.63}$$

$$s = 2M, \tag{2.64}$$

with the Hawking temperature T_H, the internal energy U and black hole entropy s.

3

Page Curve

A useful diagnostic to quantify the loss of information due to Hawking radiation turns out to be the entanglement entropy of the radiation S_{rad}. To define this quantity, we imagine a total quantum system consisting of a black hole and its radiation, so that the Hilbert space $\mathcal{H}$ describing the total system is given by $\mathcal{H} = \mathcal{H}_{\text{BH}} \otimes \mathcal{H}_{\text{rad}}$. We assume that such a product structure exists, at least approximately.[1] Given a state ρ on $\mathcal{H}$, the entanglement entropy of radiation can simply be defined to be the von Neumann entropy of the reduced state $\rho_{\text{rad}} = \text{tr}_{\text{BH}}(\rho)$,

$$S_{\text{rad}} := -\text{tr}\left(\rho_{\text{rad}} \ln\left(\rho_{\text{rad}}\right)\right). \tag{3.1}$$

Analogously we can define the von Neumann entropy of the black hole S_{BH}. Performing the calculation in the spirit of Hawking [7], one would find that S_{rad} is a growing and strictly monotonic function of time, at least for the time a semi-classical analysis can be trusted.

This, however, appears to be in contradiction with unitarity, as was argued by Page [15, 65] as follows. Let us assume a unitary time evolution of the state ρ, so that an initial pure state stays pure for all times. In this case the two quantities S_{rad} and S_{BH} are equal to each other, $S_{\text{rad}} = S_{\text{BH}}$. Further, we assume that the initial

[1] That such an assumptions is justified, is not clear by any means. It is known that such a factorization typically fails for gauge theories, see for example [63]. Since gravity itself can be thought of as a gauge theory, it would not be very surprising if that were the case. However, we can still hope that the factorization holds in some approximate sense. One should remark that there are also arguments against the correctness of this assumption, see e.g. [64].

state (which we imagine to be shortly after a black hole has been formed), does not yet carry any radiation. This premise provides us with an estimate of the dimension of the accessible Hilbert space $\mathcal{H}$ if we accept the coarse grained entropy of the initial black hole to be one quarter of the initial area A_0 of its event horizon in Planck units, $s_0 = \frac{A_0}{4}$ [9]. The estimate thus is $\dim \mathcal{H} \sim e^{A_0/4}$. It is then already easy to see that the entanglement entropy S_{rad} must be bounded by the coarse grained black hole entropy, $S_{\text{rad}} \leq s_0$. One gets a stronger bound by using $S_{\text{rad}}(t) = S_{\text{BH}}(t)$ at each time and $S_{\text{BH}}(t) \leq s(t)$, i.e.

$$S_{\text{rad}}(t) \leq s(t), \tag{3.2}$$

where $s(t)$ represents the course grained black hole entropy at a time t. Notably, this bound is strongest, when the black hole entropy s attains its minimum, zero, which corresponds to a complete evaporation of the black hole, and weakest, when its maximal, $s = s_0$, the initial moment after black hole creation. It is immediately apparent, that since $s(t) \to 0$ as the evaporation progresses, sooner or later this will be in contradiction with Hawking's result which states the entanglement entropy keeps growing.

A similar but complementary bound, which is strongest at the initial moment and weakest at the evaporation end point, can be derived. It relies on an argument using thermodynamics of the radiated gas [66]. If we assume the radiation consists of massless particles at temperature T, the relation between it's thermodynamic entropy s_{rad} and its internal energy U is given by

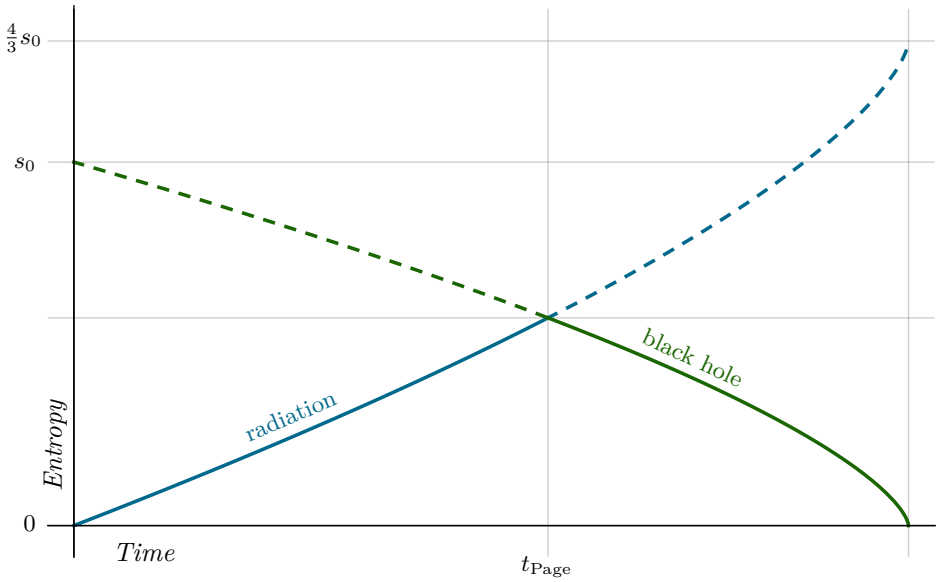
$$s_{\text{rad}} = -\frac{d}{d-1}\frac{U}{T}, \tag{3.3}$$

where d is the spacetime dimension. If the black hole radiates particles of some energy ΔM, this will of course change the internal energy of the radiation gas accordingly $\Delta U = -\Delta M$. By the first law of thermodynamics, $T\Delta s = -\Delta M$, and therefore $\Delta s_{\text{rad}} = -\frac{d}{d-1}\Delta s$, or integrating, $s_{\text{rad}}(t) = \frac{d}{d-1}\left(s_0 - s(t)\right)$. Since the entanglement entropy S_{rad} should again be bounded by the coarse grained entropy s_{rad}, we have

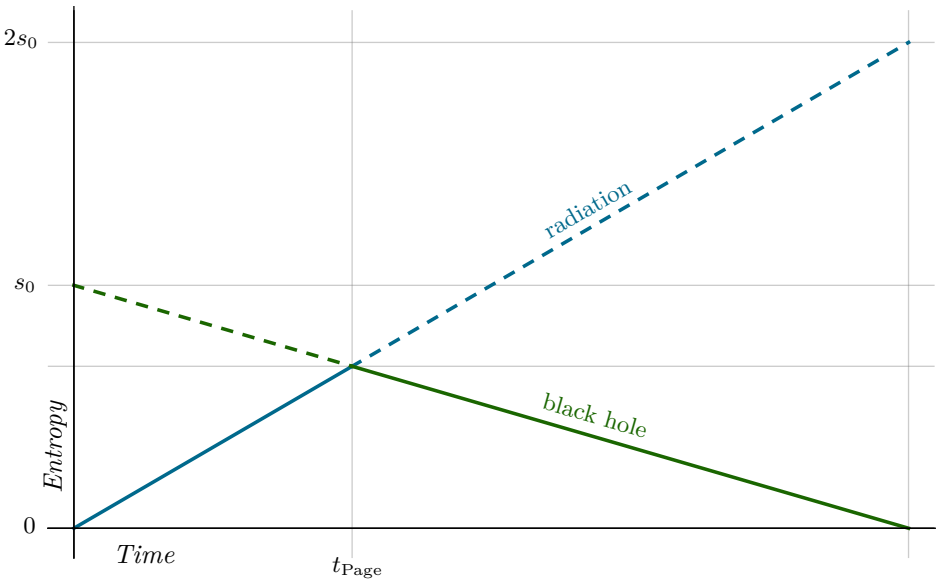
$$S_{\text{rad}}(t) \leq \frac{d}{d-1}\left(s_0 - s(t)\right). \tag{3.4}$$

The time when the two bounds for the coarse grained entropy of the radiation and the black hole meet each other is called the Page time t_{page} [16].

The two bounds are illustrated in Figure 3.1 on page 23 for a black hole emitting massless particles in four dimensions and a two dimensional black hole with a



(a) Four dimensional black hole with $\frac{dM}{dt} \sim -\frac{1}{M^2}$. [16]



(b) Two dimensional black hole with $\frac{dM}{dt} \sim -1$.

Figure 3.1: The coarse grained (thermal) entropy of the black hole $s(t)$ and radiation $s_{\text{rad}}(t)$ respectively as a function of time shown in four and two dimensions. Their intersection determines the Page time t_{Page}.

temperature independent of its mass.

If the reduced density matrix of the radiation is thermal, then the von Neumann entropy S_{rad} equals the coarse grained entropy s_{rad}. This is the behavior expected from Hawking's result [7, 10]. However, we just reasoned that unitarity implies a turnover around the Page time, where the von Neumann entropy S_{rad} should be bounded by the coarse grained black hole entropy s. Page showed [15], that the entanglement entropy of a typical state of a big enough system roughly assumes this maximum, after a time of order, say, half the evaporation time. That is to say, to good approximation, the entanglement entropy of the radiation $S_{\text{rad}}(t)$ should follow the minimum of the two values $s_{\text{rad}}(t)$ and $s(t)$ if unitarity applies.

3.1 Generalized entropy

Up until very recently, it was not known how to "repair" the computation of Hawking in order to reproduce the Page curve, i.e. in order to restore unitarity, by a semi-classical computation. So far it was not even clear, it *can* be restored by a semi-classical computation and it was suspected that the full theory of quantum gravity has to be formulated in order to get the correct behavior. Quite the opposite is claimed by [17, 18, 67] where a unitary Page curve is obtained only by using semi-classical gravity.

Motivated by AdS/CFT and the QRT formula [19–22], it was proposed that the correct semi-classical formula for the entanglement entropy S_{rad} is given by

$$S_{\text{rad}}(A) = \min \operatorname*{ext}_{I} \left(\frac{\text{Area}}{4G}(I) + S_{\text{Bulk}}(\mathcal{S}_{AI}) \right). \tag{3.5}$$

To explain this formula it will be convenient to concentrate on a Penrose diagram of an asymptotically flat evaporating black hole in Figure 3.2, page 25.

Let us imagine dividing spacetime into a region which includes the black hole and its complement which potentially only includes radiation. This is indicated by the blue line in Figure 3.2. This line (which in four dimensions would be a large sphere at each instant, $S^2 \times \mathbb{R}$) should be placed in a region where gravity is very weak. We would then attribute the degrees of freedom carried by the outside region to the Hilbert space $\mathcal{H}_{\text{rad}}$ and the rest being described by $\mathcal{H}_{\text{BH}}$, such that the total Hilbert space is $\mathcal{H} = \mathcal{H}_{\text{BH}} \otimes \mathcal{H}_{\text{rad}}$. We are interested in calculating the entanglement entropy S_{rad} associated to the infinite region outside A, that is the von Neumann entropy of a state in $\mathcal{H}_{\text{rad}}$. The prescription (3.5) then tells us to consider all possible spacetime

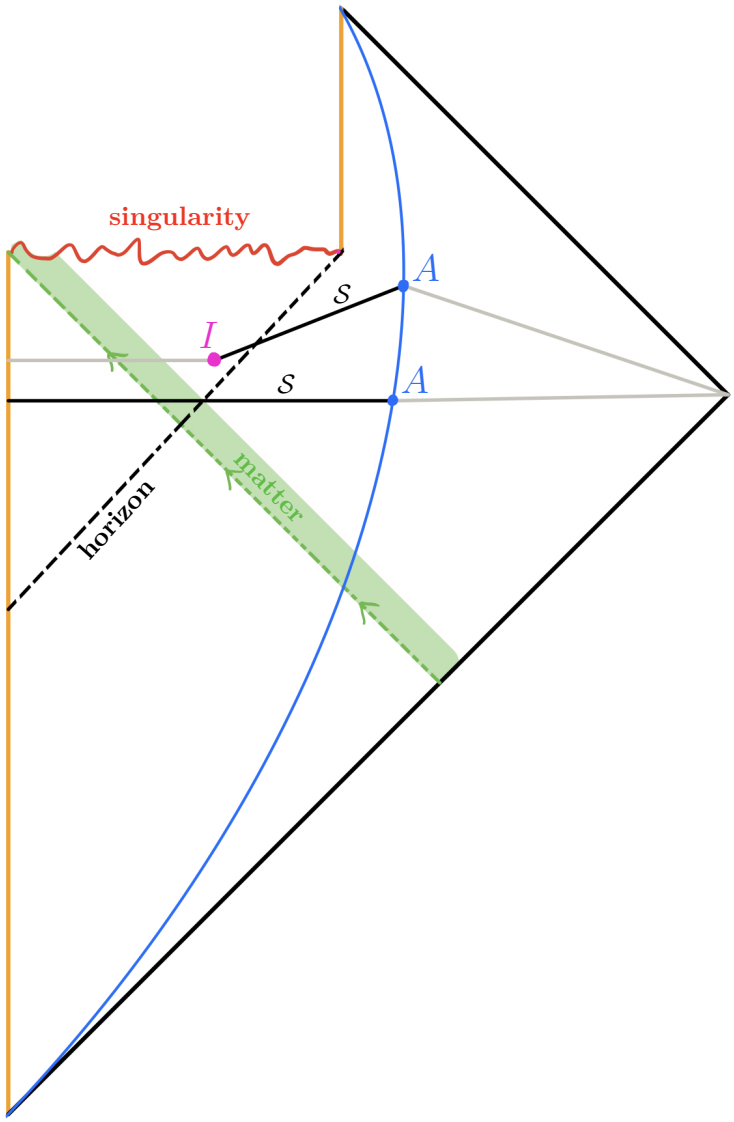


Figure 3.2: Penrose diagram of an asymptotically flat evaporating black hole. The green region represents an infalling matter beam. The event horizon is shown by the dashed line. The blue line separates the inner black hole region from the outer region where the outgoing Hawking radiation is collected. A point on the blue line A corresponds to a particular time for an observer far away. As time moves forward, the point A moves upwards on the blue line. The pink dot I indicates the appearance of an island.

points I to the left of the dividing blue line and extremize the sum of the area divided by Newton's constant G at the point I and the entanglement entropy S_{Bulk} corresponding to a matter state defined on the subregion $\mathcal{S}_{AI}$ connecting the point I with A. If multiple extrema exist one is instructed to take the minimum.

When we apply this prescription to a semi-classical black hole one expects to find two extrema of the expression in formula (3.5). One of them will be found at the (yellow) boundary far inside the black hole while the second one will be found near the horizon, indicated in Figure 3.2 by the pink dot. At early times (meaning, up to the Page time t_{Page} after the black hole was formed) the minimum of these two extrema is given by the first extremum at the boundary while after the Page time the second extremum at I will overtake.

Notice that the first minimum instructs us to evaluate the bulk entanglement entropy S_{bulk} from the anchor point A on the whole spatial slice going inside the black hole where the area term is zero. If one were to consider a pure state on the whole spatial slice divided into two complementary regions R_L and R_R, this would just coincide with the naive statement that $S_L = S_R$. At this point formula (3.5) is telling us nothing new. We get the usual result obtained in the spirit of Hawking (see the blue lines in Figure 3.1), which closely tracks the coarse grained entropy of the radiation.

The novel feature enters at the Page time, when the second extremum located near the horizon will actually overtake. Then (3.5) will be dominated by the area term which near the horizon is nothing else than the coarse grained black hole entropy s and we recover a Page curve that is consistent with unitarity, Figure 3.1. It is also quite interesting to observe that S_{Bulk} only has to be evaluated on a region between a point on the horizon I and the anchor A. In other words we are explicitly instructed to exclude the interior of the black hole in the calculation.

In [36] we explicitly evaluate the generalized entropy (3.5) for an evaporating black hole in the RST model described in chapter 2. We indeed find the exact behavior just described and one recovers the entanglement entropy depicted in Figure 3.1b at leading order. The crucial difference to the original derivation of entanglement entropy curve by Page is that in Page's derivation unitarity was assumed. In [36], however, we just perform the calculation of (3.5) in semi-classical gravity without explicitly assuming unitarity.

4

Quantum Complexity

We can give a precise definition of quantum complexity in the context of Quantum Information Theory (QIT) [68, 69]. In QIT one models a quantum computer in terms of a finite set of *gates* which are unitary operators acting on a finite dimensional Hilbert space $\mathcal{H}$. In order for the gate set to be useful in practice the gates should be universal, i.e. able to produce any quantum state in an approximate sense. However, at the same time, the gate set should not be too large since they represent the building blocks of the quantum computer.

Let $\mathcal{H}$ be a finite dimensional Hilbert space. Consider a finite set of unitary operators on $\mathcal{H}$,

$$\mathcal{G} := \{U_i : U_i \in U(\mathcal{H}),\, i = 1, \ldots, M\} \tag{4.1}$$

for M being a positive integer. Further consider the group generated by the set $\mathcal{G}$ via matrix multiplication, denoted by $\langle \mathcal{G} \rangle$. A typical element $S_{\mathcal{G}} \in \mathcal{G}$ is of the form $S_{\mathcal{G}} = U_1^{k_1} U_2^{k_2} \ldots U_M^{k_M}$, with $U_i^{k_i} \neq \mathbb{I}$, or permutations thereof.

Definition. *(Universal Gate Set)* A gate set $\mathcal{G}$ is called *universal* iff for any $U \in U(\mathcal{H})$ and $\epsilon > 0$ there exists an element $S_{\mathcal{G}} \in \langle \mathcal{G} \rangle$ such that

$$\|U - S_{\mathcal{G}}\| < \epsilon, \tag{4.2}$$

with the matrix norm being defined as usual,

$$\|U\| := \sup_{\psi \in \mathcal{H}, \|\psi\| = 1} \|U\psi\| \,. \tag{4.3}$$

In other words, the set $\langle \mathcal{G} \rangle \subset U(\mathcal{H})$ is dense in $U(\mathcal{H})$ with respect to the matrix norm.

Equivalently, given a reference state $\psi_0 \in \mathcal{H}$ with $\|\psi_0\| = 1$, any state $\psi_1 \in \mathcal{H}$ with $\|\psi_1\| = 1$ can be constructed to arbitrary accuracy by acting on ψ_0 with gates in $\mathcal{G}$. That is, given an $\epsilon > 0$, there exists a sequence of gates $S_{\mathcal{G}} \in \langle \mathcal{G} \rangle$ with

$$\|S_{\mathcal{G}}\psi_0 - \psi_1\| < \epsilon. \tag{4.4}$$

It is perhaps not obvious that this "factorization problem" has a solution, i.e. that a universal gate set exists. But it does, and the rate of convergence in the number of required gates is even remarkably fast [70]. Typically there exist more than one such sequence $S_{\mathcal{G}}$. Let us denote the set of all such sequences for a given ψ_0, ψ_1 and ϵ as $\mathfrak{c}(\psi_0, \psi_1; \epsilon)$. As just stated, this set is always non-empty, $\mathfrak{c}(\psi_0, \psi_1; \epsilon) \neq \emptyset$.

Within QIT the element $S_{\mathcal{G}} \in \langle \mathcal{G} \rangle$ represents a particular computer program. The number of steps of a computer program $S_{\mathcal{G}}$ is then given by the sum of exponents $\#S_{\mathcal{G}} := k_1 + k_2 + \cdots + k_M$. Intuitively, the complexity of a computer program should be the number of its steps. However, it could occur, that there is "better" program (i.e. a program which requires less steps), which performs the same task. In the language of above, this would be represented by another element in $\mathfrak{c}(\psi_0, \psi_1; \epsilon)$ for the same ψ_0, ψ_1 and ϵ. In this case it is clear, that the actual task was really not as "complex" as suggested by the original computer program. Therefore, we require complexity to be the number of steps of the "best" computer program performing a particular computation.

Definition. *(Quantum Complexity)* For a fixed $\epsilon > 0$, $\psi_0 \in \mathcal{H}$, $\psi_1 \in \mathcal{H}$, *Quantum Complexity* $\mathcal{C}_{\mathcal{G}}$ is defined to be the minimum number of steps of all elements in $\mathfrak{c}(\psi_0, \psi_1; \epsilon)$,

$$\mathcal{C}_{\mathcal{G}}(\psi_0, \psi_1; \epsilon) := \min_{S_{\mathcal{G}} \in \mathfrak{c}(\psi_0, \psi_1; \epsilon)} \#S_{\mathcal{G}}. \tag{4.5}$$

As such, for a fixed $\epsilon > 0$, $\mathcal{C}_{\mathcal{G}}$ can be regarded as a function $\mathcal{H} \times \mathcal{H} \to \mathbb{R}_+$. It would be tempting to view $\mathcal{C}_{\mathcal{G}}$ as a metric[1] on the projective Hilbert space $P(\mathcal{H}) \cong \mathbb{C}P^{d-1}$ [30], where d is the dimension of the Hilbert space $\mathcal{H}$. However, strictly speaking, there is a slight subtly. Since there exist two states ψ_0, and ψ_1, which are not equal but within the same ϵ-ball, $\|\psi_0 - \psi_1\| < \epsilon$, we have $\mathcal{C}_{\mathcal{G}}(\psi_0, \psi_1; \epsilon) = 0$ and thus the function $\mathcal{C}_{\mathcal{G}}$ is not positive definite on $P(\mathcal{H})$. This is easily fixed by defining our metric

[1] A *metric* on $\mathcal{H}$ is here understood to be a symmetric, positive definite function $\mathcal{H} \times \mathcal{H} \to \mathbb{R}_+$, which satisfies the triangle inequality.

to assume a value between zero and one, say $\frac{1}{2}$, if evaluated on states $\psi_0 \neq \psi_1$ with $\|\psi_0 - \psi_1\| < \epsilon$. Some elements of $\mathfrak{c}(\psi_0, \psi_1; \epsilon)$ are illustrated in Figure 4.1.

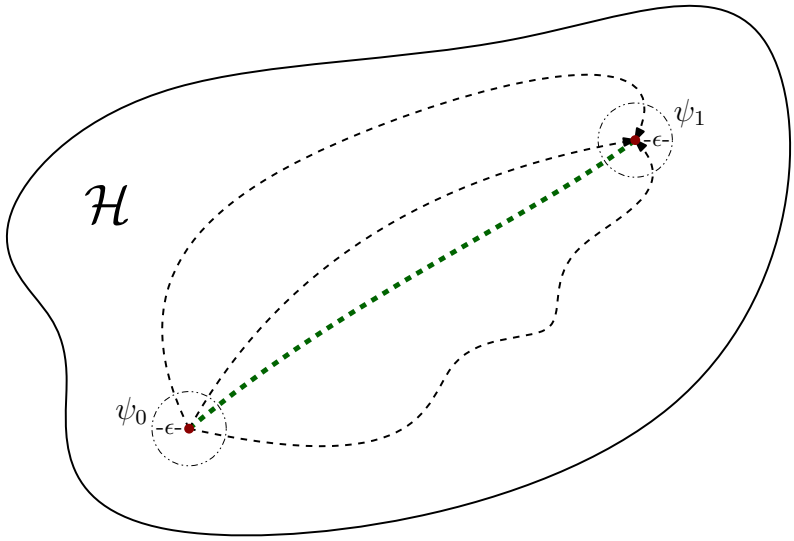


Figure 4.1: Illustration of the Hilbert space $\mathcal{H}$ and two states $\psi_0 \in \mathcal{H}$, $\psi_1 \in \mathcal{H}$ surrounded by ϵ-balls related by different sequences of gates $\mathcal{G}$. Each of the dashed lines represents an element of $\mathfrak{c}(\psi_0, \psi_1; \epsilon)$. Complexity $\mathcal{C}$ is determined by the shortest path.

Note, that equivalently we could've defined the complexity metric on the set of unitary operators $U(\mathcal{H})$, since $U(\mathcal{H})$ acts transitively on $P(\mathcal{H})$. We are then asking the equivalent question, what the minimum number of gates is to construct a particular unitary matrix U provided an accuracy ϵ. Such a geometric approach to complexity appears to be fruitful and is studied for example in [71].

4.1 Complexity of black holes

Motivated by Bekenstein's black hole entropy formula [9], the idea of the *holographic principle* [72–74] was developed, which roughly states that a gravitational system in d spacetime dimensions is described by a quantum system in $d-1$ dimensions. More specifically, a black hole can be described by a quantum system with its number of degrees of freedom dictated by the Bekenstein entropy. This fits nicely with the principle of black hole complementarity [34], which replaces the black hole interior for an asymptotic observer by a quantum system living on the "stretched horizon", a surface with an area of one unit more than the area of the event horizon in Planck units.

The precise dynamics of the quantum system living on the stretched horizon is not known. However, some necessary features of the quantum features can be deduced. One example is the "fast scrambling" property [75, 76], which refers to a short thermalization time scale. Thinking of the quantum system as a system consisting of a finite number of qubits [75], this property implies that the interaction between different qubits are "k-local" and "all-to-all" [30]. That means in particular that the interaction is not local if one thinks of the qubits to be distributed on the horizon surface, but that each interaction involves k qubits or fewer. Further, "all-to-all" implies that every subset of k qubits interact. In the language of QIT, the universal gate set $\mathcal{G}$ includes gates that act on at most k qubits. Given any k qubits, there exists a k-local gate in $\mathcal{G}$ that acts on them. As an example, a small quantum circuit is shown in Figure 4.2.

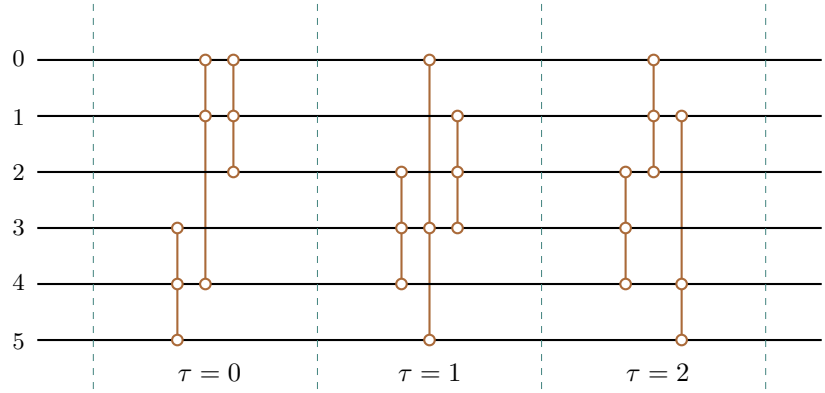


Figure 4.2: Sketch of a small quantum circuit $S_{\mathcal{G}}$ with $\#S_{\mathcal{G}} = 3 \times 3$ and $\dim \mathcal{H} = 2^6$. Each connected brown component represents a (3-local) gate. Each black line corresponds to a qubit, while the circles correspond to the action of gates. This circuit only includes three steps while in principle an arbitrary number is possible.

One can now heuristically argue, that the complexity growth for early times should be linear in time [29], where "time" is proportional to the number of steps in the quantum circuit. Let's consider an initial state ψ_0. If we have k-local gates and at each time step all qubits are involved, then the number of gates acting at each time step is given by

$$\#S_{\mathcal{G}} \sim \frac{s}{k}, \tag{4.6}$$

where s is the number of total qubits in $\mathcal{H}$. At this point the state ψ_0 evolved to the new state ψ_1 via the application of those gates. If the total space of states is large enough (this also requires that the regulator ϵ is chosen small enough), then it is likely

that the minimal amount of gates required to go from ψ_0 to ψ_1 is the same as those which we just acted with,

$$\mathcal{C}(\psi_0, \psi_1; \epsilon) \sim \frac{s}{k}. \tag{4.7}$$

This behavior is expected to roughly last until the space of states is exhausted, so that the minimal amount of gates is not given anymore by the number of gates we acted with. The time when this happens is argued to be exponential in the total number of qubits, $\tau \sim e^s$ [30]. Changing notation a little bit, $\mathcal{C}(\tau) := \mathcal{C}(\psi_0, \psi_\tau; \epsilon)$, we have

$$\frac{d\mathcal{C}(\tau)}{d\tau} \sim \frac{s}{k} \quad \text{for } \tau \ll e^s. \tag{4.8}$$

Now, we shall interpret this result as a result for a quantum system living on the stretched horizon of a black hole, describing its interior. The number of total qubits is dictated by the Bekenstein entropy $s = \frac{A}{4}$ of the black hole, where A is the area of its event horizon in Planck units. Further, we identify τ with the proper time on the stretched horizon. The relation to asymptotic Schwarzschild time t is then simply given by $\tau \sim T_H t$, where T_H is the Hawking temperature associated to the black hole. Keeping k constant, we can conclude that [35]

$$\frac{d\mathcal{C}(t)}{dt} \sim T_H s \quad \text{for } t \ll \frac{e^s}{T_H}. \tag{4.9}$$

Let us remark, that for asymptotically flat black holes, the evaporation time is at worst polynomial in s, so the behavior (4.9) is expected to be relevant for pretty much the whole existence of the black hole. Of course, for an evaporating black hole, the Hawking temperature T_H and the coarse grained entropy s typically depend on time, so that in general we expect a departure from constant growth. For example, a four dimensional Schwarzschild black hole has a Hawking temperature of $T_H \sim \frac{1}{M}$ while the Bekenstein entropy is $s \sim M^2$, so that one expects [30]

$$\frac{d\mathcal{C}(t)}{dt} \sim M(t) \sim \left(M_0^3 - 3t\right)^{1/3}, \tag{4.10}$$

assuming the black hole radiates only massless particles, $\frac{dM}{dt} \sim -\frac{1}{M^2}$. On the other hand, in this thesis, we are primarily interested in two dimensional black holes with constant Hawking temperature $T_H \sim 1$ and entropy $s \sim M$ with $\frac{dM}{dt} \sim -1$. This leads to an expected complexity growth of

$$\frac{d\mathcal{C}(t)}{dt} \sim M(t) \sim M_0 - t. \tag{4.11}$$

4.2 The geometric dual of complexity

Black hole complementarity posits that the full quantum gravitational system should be accurately described by a quantum system on the stretched horizon which in effect replaces the interior of the black hole. However, in the semi-classical limit, the dynamics of the quantum system should also be reflected by the original black hole interior geometry. This begs the question, what the corresponding geometric quantity of complexity is in the black hole interior. Obviously, whatever this quantity is, it should obey equation (4.9). The original conjecture made by Susskind [29] was based on exactly that point. He realized that the growth of the volume of an Einstein-Rosen bridge is in accordance with equation (4.9). Susskind then conjectured that the quantum complexity of a general black hole is given by its volume. In short, **C**omplexity equals **V**olume (CV). A Penrose diagram of the Einstein-Rosen bridge, including their growing volumes, can be seen in Figure 4.3.

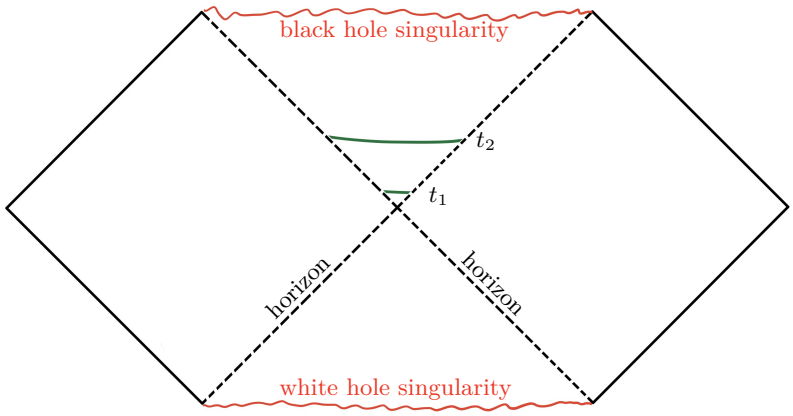


Figure 4.3: Penrose diagram of an Einstein-Rosen bridge including two volume slices (in green) of its interior at different times $t = t_1$ and $t = t_2$.

Some month later, a refined quantity that also behaves like (4.9) was discussed [77, 78]. This quantity is the action I which describes the gravitational model. The claim is, that the action I evaluated on a certain bounded region of the black hole spacetime, called the Wheeler-DeWitt (WdW) patch, yields complexity. We refer to this proposal as **C**omplexity equals **A**ction (CA). For the Einstein-Rosen bridge, we can define the WdW patch by considering the union of all spacelike surfaces connecting two anchor points on each side of the wormhole, see Figure 4.4.

An obvious difference between those two proposals is, that while CV can be defined inherently for any black hole solution, CA requires this black hole to be a solution

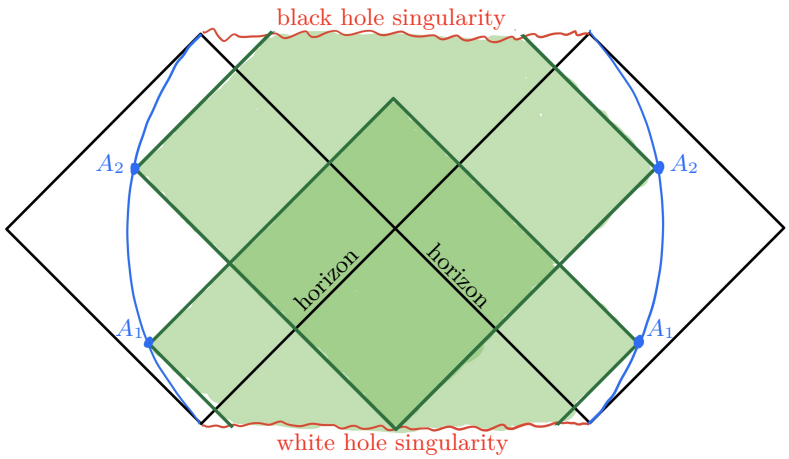


Figure 4.4: Penrose diagram of an Einstein-Rosen bridge including the WdW patch anchored at symmetrical anchor points A_1 and A_2.

of equations of motion derived by an action principle with a given action I. Both these quantities have been extensively studied in literature,[2] and most results show consistent growth behaviors for both of them. To my knowledge, the precise relation between CA and CV is not yet completely understood.

In [37, 38] we precisely define the notion of volume and action respectively for the two dimensional dilaton gravity model discussed in chapter 2. We find that the results are consistent with the growth in (4.11) which we inferred from the more general formula (4.9).

[2] To name a few influential studies in this direction: [79–85].

5

Conclusion

In this dissertation we applied a QRT formula to asymptotically flat semi-classical black holes of a two-dimensional dilaton gravity model to compute the entanglement entropy between the black hole and its radiation. Waiting for a scrambling time $t_{\text{scr}} \sim \log(s)$ after the black hole has been formed, one finds two extrema of the generalized entropy (3.5). The first one is located at the spacetime boundary "inside" the black hole, while a second one is located close to the horizon. At early times, the first extremum is minimal and the entanglement entropy follows the coarse grained entropy of radiation. This behavior persists until the Page time which in the present model is one third of the black hole lifetime. At that time, the second extremum starts to take over. Since this extremum is located near the horizon, its value is dominated by the area term in (3.5) which is nothing else than the coarse grained entropy of the black hole. As a result, we recover the Page curve in Figure 3.1b on page 23, consistent with unitarity. Although originally the QRT formula is motivated by AdS/CFT, this result provides evidence that the applicability of this formula extends beyond asymptotically AdS spacetimes. This claim is substantiated by [86] who provide a derivation of the QRT formula in the CGHS/RST model using the Replica trick which does not rely on any holographic dual theory.

It should be noted that computing a Page curve by semi-classical methods is only a first step in solving Hawking's black hole information paradox in the context of the presently discussed model. It is clear that there is still a need to understand the detailed mechanisms involved that allow information to escape the black hole.

How precisely information is encoded in the Hawking radiation is not evident and an answer requires further research.

In addition to the Page curve, we investigated the CV and CA proposals as potential dual candidates of quantum complexity. The CV proposal instructs us to evaluate the volume of extremal volume slices anchored at certain points which are labeled by time. Although we work with a theory in two dimensions, we were led to consider a four-dimensional volume functional which upon dimensional reduction reduces to a two-dimensional integral weighted with the inverse of the two dimensional gravitational coupling $e^{-2\phi}$. A volume in two dimensions is a curve and curves extremizing their length are simply geodesics. It turned out that geodesics do not even have the right qualitative behavior to be consistent with the behavior of quantum complexity, (4.9). To be able to reproduce the correct time-dependence of quantum complexity, it was further necessary to place the anchor points on the stretched horizon [34]. This supports the interpretation of the black hole interior being described by an ordinary quantum system residing at the stretched horizon, whose chaotic dynamics generates the growth in complexity.

To inspect the CA proposal we had to evaluate the gravitational action on a Wheeler-DeWitt patch. To ensure a well-posed variational principle for a gravitational action, the action has to be supplemented by appropriate boundary terms. These boundary terms are generally not unique which leaves an ambiguity in the CA prescription. The late-time limit of complexity growth is usually not sensitive to these ambiguities. However, since we are dealing with evaporating black holes a late-time limit is not useful, and we have to find a way to resolve these ambiguities. To do that we proposed to restrict the possible boundary terms in the action by a symmetry principle. Using this prescription we recovered a growth rate that is consistent with (4.9). In contrast to the CV computation where we had to rely on numerical methods to compute the volume functional in the semi-classical black hole geometry, the CA calculation provides explicit results without having to resort to numerics.

Bibliography

[1] K. Schwarzschild, *On the gravitational field of a mass point according to Einstein's theory*, *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)* **1916** (1916) 189 [physics/9905030].

[2] A. Einstein, *On a stationary system with spherical symmetry consisting of many gravitating masses*, *Annals Math.* **40** (1939) 922.

[3] D. Finkelstein, *Past-Future Asymmetry of the Gravitational Field of a Point Particle*, *Phys. Rev.* **110** (1958) 965.

[4] R. Penrose, *Gravitational collapse and space-time singularities*, *Phys. Rev. Lett.* **14** (1965) 57.

[5] LIGO Scientific Collaboration and Virgo Collaboration collaboration, *Observation of gravitational waves from a binary black hole merger*, *Phys. Rev. Lett.* **116** (2016) 061102.

[6] Event Horizon Telescope collaboration, *First M87 Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole*, *Astrophys. J.* **875** (2019) L1 [1906.11238].

[7] S. Hawking, *Particle Creation by Black Holes*, in *1st Oxford Conference on Quantum Gravity*, pp. 219–267, 8, 1975.

[8] J. Bekenstein, *Black holes and the second law*, *Lett. Nuovo Cim.* **4** (1972) 737.

[9] J. D. Bekenstein, *Black holes and entropy*, *Phys. Rev. D* **7** (1973) 2333.

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

[11] S. Hawking, *The Unpredictability of Quantum Gravity*, *Commun. Math. Phys.* **87** (1982) 395.

[12] J. R. Ellis, J. Hagelin, D. V. Nanopoulos and M. Srednicki, *Search for Violations of Quantum Mechanics*, .

[13] T. Banks, L. Susskind and M. E. Peskin, *Difficulties for the Evolution of Pure States Into Mixed States*, *Nucl. Phys. B* **244** (1984) 125.

[14] J. M. Maldacena, *The Large N limit of superconformal field theories and supergravity*, *Int. J. Theor. Phys.* **38** (1999) 1113 [hep-th/9711200].

[15] D. N. Page, *Average entropy of a subsystem*, *Phys. Rev. Lett.* **71** (1993) 1291 [gr-qc/9305007].

[16] D. N. Page, *Time Dependence of Hawking Radiation Entropy*, *JCAP* **09** (2013) 028 [1301.4995].

[17] G. Penington, *Entanglement Wedge Reconstruction and the Information Paradox*, 1905.08255.

[18] 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 [1905.08762].

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

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

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

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

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

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

[25] G. Penington, S. H. Shenker, D. Stanford and Z. Yang, *Replica wormholes and the black hole interior*, 1911.11977.

[26] J. Callan, Curtis G. and F. Wilczek, *On geometric entropy*, *Phys. Lett. B* **333** (1994) 55 [hep-th/9401072].

[27] P. Calabrese and J. L. Cardy, *Entanglement entropy and quantum field theory*, *J. Stat. Mech.* **0406** (2004) P06002 [hep-th/0405152].

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

[29] L. Susskind, *Computational Complexity and Black Hole Horizons*, *Fortsch. Phys.* **64** (2016) 24 [1403.5695].

[30] L. Susskind, *Three Lectures on Complexity and Black Holes*, 10, 2018, 1810.11563.

[31] D. Marolf and J. Polchinski, *Violations of the Born rule in cool state-dependent horizons*, *JHEP* **01** (2016) 008 [1506.01337].

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

[33] D. Marolf and J. Polchinski, *Gauge/Gravity Duality and the Black Hole Interior*, *Phys. Rev. Lett.* **111** (2013) 171301 [1307.4706].

[34] L. Susskind, L. Thorlacius and J. Uglum, *The Stretched horizon and black hole complementarity*, *Phys. Rev. D* **48** (1993) 3743 [hep-th/9306069].

[35] L. Susskind, *The Typical-State Paradox: Diagnosing Horizons with Complexity*, *Fortsch. Phys.* **64** (2016) 84 [1507.02287].

[36] F. F. Gautason, L. Schneiderbauer, W. Sybesma and L. Thorlacius, *Page Curve for an Evaporating Black Hole*, *JHEP* **05** (2020) 091 [2004.00598].

[37] L. Schneiderbauer, W. Sybesma and L. Thorlacius, *Holographic Complexity: Stretching the Horizon of an Evaporating Black Hole*, *JHEP* **03** (2020) 069 [1911.06800].

[38] L. Schneiderbauer, W. Sybesma and L. Thorlacius, *Action Complexity for Semi-Classical Black Holes*, *JHEP* **07** (2020) 173 [2001.06453].

[39] J. Callan, Curtis G., S. B. Giddings, J. A. Harvey and A. Strominger, *Evanescent black holes*, *Phys. Rev. D* **45** (1992) 1005 [hep-th/9111056].

[40] J. G. Russo, L. Susskind and L. Thorlacius, *The Endpoint of Hawking radiation*, *Phys. Rev. D* **46** (1992) 3444 [hep-th/9206070].

[41] S. B. Giddings and A. Strominger, *Dynamics of extremal black holes*, *Phys. Rev. D* **46** (1992) 627 [hep-th/9202004].

[42] L. Thorlacius, *Black hole evolution*, *Nucl. Phys. B Proc. Suppl.* **41** (1995) 245 [hep-th/9411020].

[43] S. Christensen and S. Fulling, *Trace Anomalies and the Hawking Effect*, *Phys. Rev. D* **15** (1977) 2088.

[44] A. S. Schwarz, *Instantons and fermions in the field of instanton*, *Commun. Math. Phys.* **64** (1979) 233.

[45] A. Polyakov, *Quantum geometry of bosonic strings*, *Physics Letters B* **103** (1981) 207.

[46] P. Di Francesco, P. Mathieu and D. Senechal, *Conformal Field Theory*, Graduate Texts in Contemporary Physics. Springer-Verlag, New York, 1997, 10.1007/978-1-4612-2256-9.

[47] S. M. Carroll, *Spacetime and Geometry*. Cambridge University Press, 7, 2019.

[48] R. C. Myers, *Black hole entropy in two-dimensions*, *Phys. Rev. D* **50** (1994) 6412 [hep-th/9405162].

[49] T. M. Fiola, J. Preskill, A. Strominger and S. P. Trivedi, *Black hole thermodynamics and information loss in two-dimensions*, *Phys. Rev. D* **50** (1994) 3987 [hep-th/9403137].

[50] D. A. Lowe, *Semiclassical approach to black hole evaporation*, *Phys. Rev. D* **47** (1993) 2446 [hep-th/9209008].

[51] T. Piran and A. Strominger, *Numerical analysis of black hole evaporation*, *Phys. Rev. D* **48** (1993) 4729 [hep-th/9304148].

[52] T. Tada and S. Uehara, *Consequence of Hawking radiation from 2-d dilaton black holes*, *Phys. Rev. D* **51** (1995) 4259 [hep-th/9409039].

[53] A. Ashtekar, F. Pretorius and F. M. Ramazanoglu, *Evaporation of 2-Dimensional Black Holes*, *Phys. Rev. D* **83** (2011) 044040 [1012.0077].

[54] A. Bilal and J. Callan, Curtis G., *Liouville models of black hole evaporation*, *Nucl. Phys. B* **394** (1993) 73 [hep-th/9205089].

[55] S. de Alwis, *Quantization of a theory of 2-d dilaton gravity*, *Phys. Lett. B* **289** (1992) 278 [hep-th/9205069].

[56] G. Gibbons and S. Hawking, *Action Integrals and Partition Functions in Quantum Gravity*, *Phys. Rev. D* **15** (1977) 2752.

[57] J. York, *Boundary terms in the action principles of general relativity*, *Found. Phys.* **16** (1986) 249.

[58] G. Gibbons and M. Perry, *The Physics of 2-D stringy space-times*, *Int. J. Mod. Phys. D* **1** (1992) 335 [hep-th/9204090].

[59] V. P. Frolov, *Two-dimensional black hole physics*, *Phys. Rev. D* **46** (1992) 5383.

[60] S. N. Solodukhin, *The Conical singularity and quantum corrections to entropy of black hole*, *Phys. Rev. D* **51** (1995) 609 [hep-th/9407001].

[61] E. Witten, *On string theory and black holes*, *Phys. Rev. D* **44** (1991) 314.

[62] D. V. Fursaev and S. N. Solodukhin, *On the description of the Riemannian geometry in the presence of conical defects*, *Phys. Rev. D* **52** (1995) 2133 [hep-th/9501127].

[63] H. Casini, M. Huerta and J. A. Rosabal, *Remarks on entanglement entropy for gauge fields*, *Phys. Rev. D* **89** (2014) 085012 [1312.1183].

[64] A. Laddha, S. G. Prabhu, S. Raju and P. Shrivastava, *The Holographic Nature of Null Infinity*, 2002.02448.

[65] D. N. Page, *Black hole information*, in *5th Canadian Conference on General Relativity and Relativistic Astrophysics (5CCGRRA)*, pp. 0001–41, 5, 1993, hep-th/9305040.

[66] W. Zurek, *Entropy Evaporated by a Black Hole*, *Phys. Rev. Lett.* **49** (1982) 1683.

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

[68] R. Cleve, *An introduction to quantum complexity theory*, quant-ph/9906111.

[69] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010, 10.1017/CBO9780511976667.

[70] A. Y. Kitaev, *Quantum computations: algorithms and error correction*, *Russian Mathematical Surveys* **52** (1997) 1191.

[71] M. A. Nielsen, M. R. Dowling, M. Gu and A. C. Doherty, *Quantum computation as geometry*, *Science* **311** (2006) 1133 [https://science.sciencemag.org/content/311/5764/1133.full.pdf].

[72] C. B. Thorn, *Reformulating string theory with the 1/N expansion*, in *The First International A.D. Sakharov Conference on Physics*, pp. 0447–454, 5, 1991, hep-th/9405069.

[73] C. R. Stephens, G. 't Hooft and B. F. Whiting, *Black hole evaporation without information loss*, *Class. Quant. Grav.* **11** (1994) 621 [gr-qc/9310006].

[74] L. Susskind, *The World as a hologram*, *J. Math. Phys.* **36** (1995) 6377 [hep-th/9409089].

[75] P. Hayden and J. Preskill, *Black holes as mirrors: Quantum information in random subsystems*, *JHEP* **09** (2007) 120 [0708.4025].

[76] Y. Sekino and L. Susskind, *Fast Scramblers*, *JHEP* **10** (2008) 065 [0808.2096].

[77] A. R. Brown, D. A. Roberts, L. Susskind, B. Swingle and Y. Zhao, *Holographic Complexity Equals Bulk Action?*, *Phys. Rev. Lett.* **116** (2016) 191301 [1509.07876].

[78] A. R. Brown, D. A. Roberts, L. Susskind, B. Swingle and Y. Zhao, *Complexity, action, and black holes*, *Phys. Rev. D* **93** (2016) 086006 [1512.04993].

[79] M. Alishahiha, *Holographic Complexity*, *Phys. Rev. D* **92** (2015) 126009 [1509.06614].

[80] L. Lehner, R. C. Myers, E. Poisson and R. D. Sorkin, *Gravitational action with null boundaries*, *Phys. Rev. D* **94** (2016) 084046 [1609.00207].

[81] D. Carmi, R. C. Myers and P. Rath, *Comments on Holographic Complexity*, *JHEP* **03** (2017) 118 [1612.00433].

[82] R. Jefferson and R. C. Myers, *Circuit complexity in quantum field theory*, *JHEP* **10** (2017) 107 [1707.08570].

[83] S. Chapman, M. P. Heller, H. Marrochio and F. Pastawski, *Toward a Definition of Complexity for Quantum Field Theory States*, *Phys. Rev. Lett.* **120** (2018) 121602 [1707.08582].

[84] A. R. Brown and L. Susskind, *Second law of quantum complexity*, *Phys. Rev. D* **97** (2018) 086015 [1701.01107].

[85] A. R. Brown, H. Gharibyan, H. W. Lin, L. Susskind, L. Thorlacius and Y. Zhao, *Complexity of Jackiw-Teitelboim gravity*, *Phys. Rev. D* **99** (2019) 046016 [1810.08741].

[86] T. Hartman, E. Shaghoulian and A. Strominger, *Islands in Asymptotically Flat 2D Gravity*, 2004.13857.

Articles

Article I

Published for SISSA by Springer

Received: *December 2, 2019*
Accepted: *February 10, 2020*
Published: *March 12, 2020*

Holographic complexity: stretching the horizon of an evaporating black hole

Lukas Schneiderbauer, Watse Sybesma and Lárus Thorlacius

Science Institute, University of Iceland,
Dunhaga 3, Reykjavík 107, Iceland

E-mail: lukas.schneiderbauer@gmail.com, watse@hi.is, lth@hi.is

Abstract: We obtain the holographic complexity of an evaporating black hole in the semi-classical RST model of two-dimensional dilaton gravity, using a volume prescription that takes into account the higher-dimensional origin of the model. For classical black holes, we recover the expected late time behaviour of the complexity, but new features arise at the semi-classical level. By considering the volume inside the stretched horizon of the evolving black hole, we obtain sensible results for the rate of growth of the complexity, with an early onset of order the black hole scrambling time followed by an extended period where the rate of growth tracks the shrinking area of the stretched horizon as the black hole evaporates.

Keywords: 2D Gravity, Black Holes, Classical Theories of Gravity

ArXiv ePrint: 1911.06800

https://doi.org/10.1007/JHEP03(2020)069

JHEP03(2020)069

Contents

JHEP03(2020)069

1 Introduction

The interior of a black hole is the archetype of an emergent spacetime in the holographic approach to quantum gravity. The principle of black hole complementarity posits that the interior geometry and any matter that enters a black hole can be described in terms of a finite number of quantum mechanical degrees of freedom associated with a stretched horizon located outside the event horizon [1]. In order to reproduce black hole thermodynamics, the number of stretched horizon degrees of freedom should match the Bekenstein-Hawking entropy and the dynamics must be sufficiently chaotic to scramble quantum information on a relatively short timescale, but, beyond that, the precise nature of the stretched horizon dynamics and the holographic encoding of the black hole interior remain elusive. In what follows, we will not make any specific assumptions about the scrambling dynamics but it can be useful to keep in mind a collection of qubits undergoing k-local interactions as a simple model [2].

Quantum complexity has in recent years emerged as an important entry in the holographic dictionary for black holes following Susskind's conjecture that the expanding spatial volume of the Einstein-Rosen bridge of a two-sided eternal black hole reflects the growing complexity of a corresponding quantum state [3]. In the present paper, we explore the relation between quantum complexity and interior black hole geometry in the context of semi-classical black holes that are formed by gravitational collapse and subsequently evaporate due to the emission of Hawking radiation. Our main result, based on explicit calculations in a two-dimensional dilaton gravity model that allows analytic study of semi-classical effects, is that the rate of growth of holographic complexity precisely tracks the shrinking area of the stretched horizon as the black hole evaporates, where the stretched horizon is taken to be a membrane with an area larger than that of the event horizon by order one in the appropriate units of the model.

The dilaton gravity model has explicit classical solutions describing black hole formation from arbitrary incoming matter energy flux. We focus, for simplicity, on black holes

formed by an infalling thin shell and start off by adapting the complexity as volume conjecture to this context. A suitably defined volume functional exhibits precisely the expected linear growth with time at late times and by restricting to the volume inside the stretched horizon of the dynamically formed black hole one finds reasonable early-time behaviour as well. We then consider a semi-classical extension of the model where the field equations remain analytically soluble and numerically evaluate the volume functional in an evaporating black hole background.

The transitory nature of semi-classical black holes highlights certain technical aspects of the identification between complexity and volume, that can often be ignored when considering classical black holes. In the present paper, we only consider the volume representation of the semi-classical black hole complexity, where these issues are relatively easy to address. The alternative formulation of holographic complexity in terms of the action on a Wheeler-DeWitt patch [4, 5] is also of interest for these dilaton gravity models, but it is more subtle to implement at the semi-classical level, and we postpone this to a forthcoming paper [6].

2 Complexity of classical CGHS black holes

We work within a class of two-dimensional dilaton gravity theories first introduced by Callan, Giddings, Harvey, and Strominger (CGHS) [7]. These are simple toy theories for black hole physics that can be systematically studied at the semi-classical level. They have classical solutions that describe black hole geometries with a spacelike singularity inside an event horizon. The black holes include static two-sided black holes and also dynamical black holes formed by the gravitational collapse of matter fields. The quantization of matter fields in a black hole background leads to Hawking radiation and its back-reaction on the geometry causes the black hole to evaporate. The subsequent evolution is particularly simple to track in a variant of the semi-classical model that was introduced by Russo, Susskind, and Thorlacius (RST), where the semi-classical field equations can be solved analytically [8].

The original CGHS model can be viewed as a spherical reduction of a four-dimensional dilaton gravity theory in a near-extremal magnetically charged black hole background [9, 10]. The two-dimensional theory captures the low-energy dynamics of radial modes in the near-horizon region of higher-dimensional geometry. The volume that is to be identified as the quantum complexity is that of a spacelike three-dimensional surface in the original theory, rather than the length of a spacelike curve in two-dimensions, and this will be reflected in our calculations below.

The CGHS model and related semi-classical models were studied extensively in the early 1990's and several reviews were written at that time, including [11–14]. We will be brief and only introduce the minimal ingredients needed for the purposes of this paper.

The classical GGHS action,

$$S_{\rm CGHS} = \int_{\mathcal{M}} d^2x\sqrt{-g}\left[e^{-2\phi}\left(R + 4(\nabla\phi)^2 + 4\lambda^2\right) - \frac{1}{2}\sum_{i=1}^{N}(\nabla f_i)^2\right], \tag{2.1}$$

JHEP03(2020)069

involves the two-dimensional metric, a scalar dilaton field, and matter in the form of N minimally coupled scalar fields f_i. The two-dimensional theory inherits a scale λ from the parent theory set by the magnetic charge of the near-extremal black hole. In the following, we take length to be measured in units of λ^{-1} and thus set $\lambda = 1$.

We find it convenient to work in a conformal gauge,

$$ds^2 = -e^{2\rho} dx^+ dx^- , \tag{2.2}$$

and use a residual conformal reparametrisation to choose coordinates where $\rho = \phi$. These are referred to as Kruskal coordinates for reasons that will become apparent below. In this coordinate system, the classical equations of motion and constraints reduce to

$$\partial_+\partial_- f_i = 0 \,, \qquad \partial_+\partial_- e^{-2\phi} = -1 \,, \qquad \partial_\pm^2 e^{-2\phi} = -T^f_{\pm\pm} \,, \tag{2.3}$$

where $T^f_{\mu\nu}$ is the energy-momentum tensor of the f_i matter fields.

In order to obtain the holographic complexity of a classical CGHS black hole, we introduce a volume functional

$$V = \int ds\, e^{-2\phi} \sqrt{g_{\mu\nu} \dot{y}^\mu \dot{y}^\nu} \,, \tag{2.4}$$

where $y^\mu(s)$ is a spacelike curve in the two-dimensional spacetime and the integrand includes a factor of $e^{-2\phi}$, which is proportional to the area of the local transverse two-sphere S^2 of the near-extremal dilaton black hole (in Einstein frame) in the higher-dimensional parent theory. A corresponding factor was included when defining $\mathcal{C}_V$ for black holes in two-dimensional Jackiw-Teitelboim gravity in [15]. In the case of the Jackiw-Teitelboim black hole, the transverse S^2 has constant area and the calculation of the complexity reduces to calculating a two-dimensional geodesic length. For a CGHS black hole, on the other hand, the transverse area depends on spatial location and curves that maximize (2.4) are *not* geodesics.

2.1 Complexity of a two-sided black hole

The first solution we consider describes a two-sided eternal black hole,

$$e^{-2\phi} = e^{-2\rho} = M - x^+ x^- \,, \qquad f_i = 0 \,. \tag{2.5}$$

From the Ricci scalar,

$$R = -2\nabla^2 \rho = \frac{4M}{M - x^+ x^-} \,, \tag{2.6}$$

it is apparent that, for $M > 0$, the curvature is singular on the spacelike curves $x^+ x^- = M$, corresponding to the white hole and black hole singularity. The Penrose diagram, shown in figure 1, is identical to the one obtained for a Schwarzschild black hole in 3+1 dimensional Einstein gravity. The event horizon is at $x^+ x^- = 0$ and there are two asymptotic regions, $-x^+ x^- \to \infty$, where the curvature goes to zero. One can introduce Schwarzschild-like coordinates (t, σ) in the outside region on the right (where $x^+ > 0$ and $x^- < 0$) via the transformation $x^\pm = \pm e^{\pm t + \sigma}$. The corresponding coordinate transformation on the left

JHEP03(2020)069

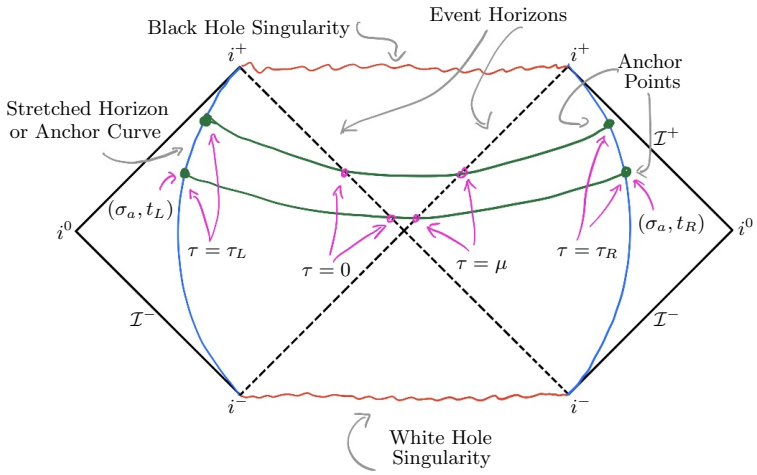


Figure 1. This cartoon depicts the Penrose diagram of a two-sided eternal CGHS black hole.

(where $x^+ < 0$ and $x^- > 0$) is given by $x^\pm = \mp e^{\mp t+\sigma}$. With these conventions, the metric in (t, σ) coordinates approaches the two-dimensional Minkowski metric as $\sigma \to +\infty$ on both sides of the black hole and t propagates to the future in the 'upward' direction on both sides.

The volume in (2.4) is divergent for spacelike curves that extend all the way to spatial infinity. In order to obtain a finite expression for the complexity, we introduce timelike anchor curves outside the black hole where the volume integral is cut off (see figure 1). We find it convenient to use anchor curves on which the dilaton field is constant, $\phi(x_a^+, x_a^-) = \phi_a$, providing a coordinate invariant notion of spatial position outside the black hole, and place the curves symmetrically on the left- and right-hand side of the black hole. The anchor curves take a particularly simple form in the Schwarzschild-like coordinates, where they are curves of constant $\sigma = \sigma_a$, with

$$\sigma_a = \frac{1}{2} \log \left(e^{-2\phi_a} - M \right) , \tag{2.7}$$

where the anchor curve on the left (right) is parametrised by t_L (t_R).

Our prescription for the volume complexity $\mathcal{C}_V(t_R, t_L)$ of a two-sided CGHS black hole is then given by the maximal volume on the set of spacelike curves $(y^+(s), y^-(s))$ with fixed endpoints at t_L and t_R on the left and right anchor curves. This amounts to maximizing the volume functional,

$$V = \int ds \, \sqrt{-\dot{y}^+ \dot{y}^- (M - y^+ y^-)} \, , \tag{2.8}$$

and evaluating the resulting maximal volume.

In order to proceed, we note that the functional is invariant under the transformations

$$\begin{aligned} y^+ &\mapsto e^\epsilon y^+ \, , \\ y^- &\mapsto e^{-\epsilon} y^- \, , \end{aligned} \tag{2.9}$$

JHEP03(2020)069

for $\epsilon \in \mathbb{R}$. The corresponding conserved quantity is given by

$$E = \sqrt{\frac{M - y^+ y^-}{-\dot{y}^+ \dot{y}^-}} \left(\dot{y}^+ y^- - y^+ \dot{y}^- \right) . \tag{2.10}$$

We construct the corresponding maximum volume curve by first focusing on the outside region on the right and rewriting the conserved charge in the (t, σ) coordinates,

$$E = -2e^{\sigma} \frac{dt}{d\sigma} \sqrt{\frac{M + e^{2\sigma}}{1 - \left(\frac{dt}{d\sigma}\right)^2}} . \tag{2.11}$$

This integrates to

$$t - t_0 = -\sigma + \frac{1}{2} \log \left(E^2 + 2Me^{2\sigma} + E\sqrt{E^2 + 4Me^{2\sigma} + 4e^{4\sigma}} \right) . \tag{2.12}$$

We then convert back to Kruskal coordinates by using $y^{\pm} = \pm e^{\pm t + \sigma}$ and obtain the following equation for a maximal volume curve, which is valid over the entire extended spacetime,

$$e^{-2t_0} \left(y^+\right)^2 + 4My^+ y^- + 4e^{2t_0} \left(M^2 - E^2\right) \left(y^-\right)^2 - 2E^2 = 0 . \tag{2.13}$$

Maximal volume curves that extend between the anchor curves correspond to a conserved charge in the range $-M < E < M$, and can be parametrised as

$$y^+ = \sqrt{2} e^{t_0} \epsilon \sinh \tau , \quad y^- = -\frac{e^{-t_0}}{\sqrt{2}} \sinh(\tau - \mu) , \tag{2.14}$$

with $\epsilon = \sqrt{M^2 - E^2}$ and $\tanh \mu = \frac{E}{M}$, while curves satisfying (2.13) with $|E| > M$ run into the curvature singularity. The parameter τ in (2.14) runs from a negative value $\tau_L < 0$ at the endpoint on the left anchor curve to $\tau = 0$, where the curve enters the black hole from the left. At $\tau = \mu$ the curve exits the black hole to the right and reaches the endpoint on the right anchor curve at $\tau = \tau_R$. We illustrate this setup in figure 1.

The maximal volume curve is labelled by E and t_0 in (2.14) but these labels are in one-to-one correspondence with the Schwarzschild times t_L and t_R where the curve meets the anchor curves. To see this, consider the intersection points between the maximal volume curve and the anchor curves. On the one hand we have

$$e^{2\sigma_a} = \epsilon \sinh \tau_R \sinh(\tau_R - \mu) = \epsilon \sinh(-\tau_L) \sinh(\mu - \tau_L) , \tag{2.15}$$

relating curve parameters to the spatial location of the anchor curves, and on the other hand a pair of relations involving the Schwarzschild times at the anchor points,

$$e^{2t_R} = 2\epsilon e^{2t_0} \frac{\sinh(\tau_R)}{\sinh(\tau_R - \mu)} , \quad e^{2t_L} = \frac{e^{-2t_0}}{2\epsilon} \frac{\sinh(\mu - \tau_L)}{\sinh(-\tau_L)} . \tag{2.16}$$

The second equation in (2.15) is satisfied by imposing $\tau_R = \mu - \tau_L$ and the time relations can then be re-expressed as

$$e^{t_R - t_L} = 2\epsilon e^{2t_0} , \quad e^{t_R + t_L} = \frac{\sinh(\tau_R)}{\sinh(\tau_R - \mu)} . \tag{2.17}$$

JHEP03(2020)069

By combining (2.15) with the second equation in (2.17) and doing some algebra one eventually arrives at

$$\epsilon = -e^{2\sigma_a}\cosh(t_R{+}t_L) + \sqrt{e^{4\sigma_a}\cosh^2(t_R{+}t_L) + 2Me^{2\sigma_a} + M^2}\,. \tag{2.18}$$

The parametrisation (2.14) can then be re-expressed in terms of t_R and t_L, as

$$y^+ = e^{\frac{1}{2}(t_R-t_L)}\sqrt{\epsilon}\sinh\tau\,, \quad y^- = -e^{-\frac{1}{2}(t_R-t_L)}\sqrt{\epsilon}\sinh(\tau-\mu)\,. \tag{2.19}$$

with $\epsilon(t_R{+}t_L)$ given by (2.18).

The volume functional (2.8) is easily evaluated in this parametrisation,

$$\begin{aligned} V &= \epsilon\int_{\tau_L}^{\tau_R} d\tau\ \cosh\tau\cosh(\tau-\mu) \\ &= \frac{M}{2}\big(2\tau_R-\mu\big) + \frac{\epsilon}{2}\sinh\big(2\tau_R-\mu\big)\,. \end{aligned} \tag{2.20}$$

The first equation in (2.15) can be combined with second equation in (2.17) to give

$$2\tau_R - \mu = \operatorname{arccosh}\left[\frac{1}{M}\left(e^{2\sigma_a}\cosh(t_R{+}t_L){+}\sqrt{e^{4\sigma_a}\cosh^2(t_R{+}t_L){+}2Me^{2\sigma_a}{+}M^2}\right)\right], \tag{2.21}$$

which can then be inserted in (2.20) to obtain an exact, if somewhat unwieldy, formula for the maximal volume as a function of $t_R{+}t_L$.

The expression for the volume simplifies enormously at late times,

$$V = \frac{M}{2}(t_R{+}t_L) + e^{2\sigma_a} + \frac{M}{2}\left(1 + 2\sigma_a - \log\frac{M}{2}\right) + \mathcal{O}\big(e^{-2(t_R+t_L)}\big)\,, \tag{2.22}$$

with a leading term that grows linearly with time, followed by a constant term that depends on the location of the anchor curve, and subsequent terms that are exponentially suppressed at late times. As expected, the volume diverges in the $\sigma_a \to \infty$ limit, where the anchor curves are moved off to spatial infinity, but the late time rate of growth is unaffected by the location of the anchor curves. The volume prescription for complexity is sometimes taken to only include the volume inside the event horizon of the black hole. This amounts to cutting off the volume integration in (2.20) at the event horizon,

$$\begin{aligned} V_{\rm EH} &= \epsilon\int_0^{\mu} d\tau\ \cosh\tau\cosh(\tau-\mu) \\ &= \frac{M}{2}\big(\mu + \tanh\mu\big)\,, \end{aligned} \tag{2.23}$$

which can be shown to grow at the same rate at late times as the full volume between anchor curves. Later on, when we consider dynamical black holes formed by the gravitational collapse of matter, we will see that the stretched horizon is a natural choice of anchor curve. In the case at hand, we define the stretched horizon to be a membrane outside the black hole, with an area that is one unit larger than the area of the event horizon,

$$e^{-2\phi_{\rm SH}} = e^{-2\phi_{\rm EH}} + 1 = M + 1\,. \tag{2.24}$$

JHEP03(2020)069

This is a curve of constant dilaton field outside the black hole, which is how we defined our anchor curves above. Indeed, with this definition, the stretched horizon is located at $\sigma_{\rm SH} = 0$ in the (t, σ) coordinate system and the volume inside the stretched horizon can be obtained by setting $\sigma_a = 0$ in (2.22). The result differs from the volume inside the event horizon by only a small amount and the late time volume growth is the same. For a static two-sided black hole the location of the anchor curve is unimportant if all we are interested in is the late time rate of growth of the complexity. It is only when we consider dynamical black holes that the advantage of using the stretched horizon as the anchor curve becomes apparent.

We note that the Hawking temperature of a CGHS black hole is $T_{\rm H} = \frac{1}{2\pi}$, independent of the black hole mass [7], and the Bekenstein-Hawking entropy is given by $S_{\rm BH} = 2M$. The late time rate of growth of the complexity given by the volume in (2.22) is thus proportional to $S_{\rm BH}\, T_{\rm H}$, which is precisely in line with the original $\mathcal{C} = V$ proposal [3].

2.2 Complexity of a black hole formed by gravitational collapse

Next, we consider collapsing a thin shell of matter at $x^+ = x_0^+$, mediated by matter fields f_i, into a CGHS vacuum and creating a black hole. This amounts to

$$T^f_{++} = \frac{M}{x_0^+}\delta(x^+ - x_0^+)\,, \tag{2.25}$$

where $\delta(x^+ - x_0^+)$ is a delta function. For the geometry this implies

$$e^{-2\phi} = e^{-2\rho} = \begin{cases} -x^+x^- & \text{if } x^+ < x_0^+\,, \\ -x^+\left(x^- + \frac{M}{x_0^+}\right) + M & \text{if } x^+ \geq x_0^+\,. \end{cases} \tag{2.26}$$

The resulting black hole is one-sided, as illustrated in figure 2.

Just as in the eternal black hole geometry, we need to introduce an anchor curve outside the black hole to obtain a finite volume. The extremal curve now reaches from a point on the anchor curve to a point on $x^-x^+ = 0$. In figure 2, we sketch the setup. In contrast to the eternal black hole where we had two anchor points and therefore a unique extremal curve connecting these two points, there is now only one anchor point and therefore we have to supply a prescription for additional boundary conditions. In [16], for instance, it was argued that in order to obtain a smooth volume at the radial origin, the additional boundary conditions should be $t'(r) = 0$ at $r = 0$. Expressing the corresponding condition in our setup in Kruskal coordinates one finds the relation

$$x^-\dot{x}^+ - x^+\dot{x}^- = 0\,, \tag{2.27}$$

at $x^+x^- = 0$. An alternative prescription is to consider all locally extremal curves originating from the anchor point, and selecting the curve that maximizes the volume inside the black hole. This computation can be done and, interestingly, it turns out that this prescription leads to exactly the same curves as the boundary conditions (2.27).

Curves that maximise the volume functional (2.4) in the one-sided black hole background are obtained by patching together maximal curves across the infalling shockwave.

JHEP03(2020)069

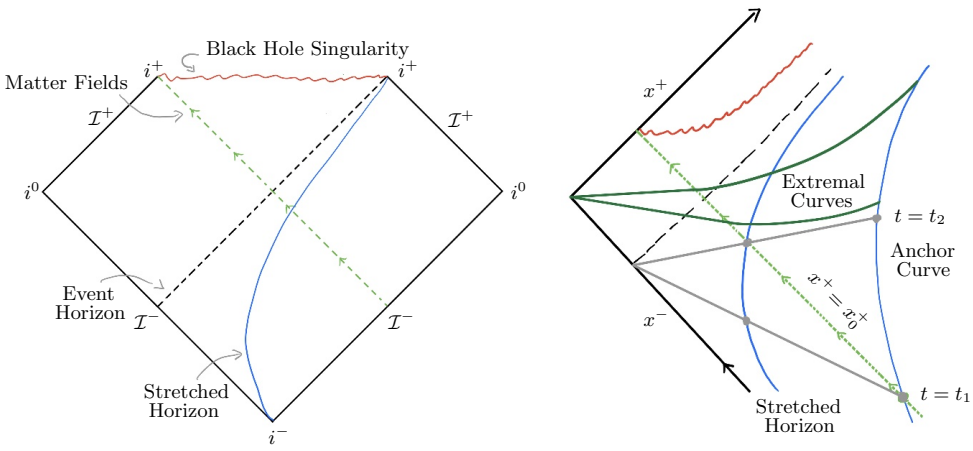


Figure 2. (*Left*) This cartoon depicts the Penrose diagram of a one-sided CGHS black hole formed by gravitational collapse. (*Right*) This is a Kruskal diagram of the Penrose diagram on the left. The color coding coincides. The gray line is an equal time curve.

In the inside region, $0 < x^+ < x_0^+$, the geometry is flat and the volume functional takes the simple form,

$$V = \int ds \, \sqrt{y^+ \dot{y}^+ y^- \dot{y}^-} \, . \tag{2.28}$$

One obtains a two-parameter family of maximal curves,

$$\left(x^+\right)^2 = \alpha \left(x^-\right)^2 + \beta \tag{2.29}$$

where $\alpha > 0$ and $\beta < \sqrt{x_0^-}$ are real valued parameters. The boundary condition (2.27) selects curves with $\beta = 0$, which are simply straight lines emanating from the origin $x^+ = x^- = 0$ in Kruskal coordinates (see figure 2).

In the outside region, $x^+ > x_0^+$, the geometry is that of a static black hole and the volume functional reduces to (2.8) in shifted Kruskal coordinates,

$$y^+ = x^+ \, , \qquad y^- = x^- + \frac{M}{x_0^+} \, . \tag{2.30}$$

There is again a two-parameter family of maximal volume curves satisfying (2.13) and labelled by E and t_0. We use Weierstraß-Erdmann conditions to patch across the shockwave. First of all, the curve itself should be continuous. A second condition comes from viewing the integral in the volume functional as a Lagrange density and requiring that the momenta conjugate to y^+ and y^- be continuous across the shock. Those matching conditions uniquely determine the parameter α and β in terms of E and t_0, and vice versa. One finds, in particular, that curves with $\beta = 0$ in the inside region match onto curves with $E = M$ on the outside leaving us with a one-parameter family

$$y^- = \begin{cases} \frac{M}{2y^+} - \frac{e^{-2t_0}}{4M} y^+ & y^+ > x_0^+ \, , \\ \frac{M}{x_0^+} - \left(\frac{M}{2x_0^+} + \frac{e^{-2t_0}}{4M} \right) y^+ & y^+ < x_0^+ \, . \end{cases} \tag{2.31}$$

JHEP03(2020)069

Subsequently, the remaining parameter t_0 can be uniquely related to the tortoise time t on the anchor curve, see figure 2. As in section 2.1, we define the anchor curves, parametrised by σ_a, so that the dilaton field is constant $\phi = \phi_a$,

$$e^{2\sigma_a} = e^{-2\phi_a} - M = \begin{cases} -y^+ y^- & y^+ > x_0^+ \\ -y^+ \left(y^- - \frac{M}{x_0^+}\right) - M & y^+ < x_0^+ \end{cases} . \tag{2.32}$$

We now obtain a relation between σ_a, t_0 and t by combining (2.31) and (2.32). One is now in a position to evaluate the volume functional (2.4).

The growth of the volume inside the stretched horizon as a function of tortoise time t is given by

$$V'_{\rm SH}(t) = \begin{cases} \frac{1}{2}M & t \geq t_2\,, \\ \frac{1}{2}M\cosh(t)^{-2} + \mathcal{O}\left(\frac{1}{M}\right) & t < t_2\,, \end{cases} \tag{2.33}$$

where t_2 is the moment the black hole is formed. We conclude that the volume growth sets in essentially at the time the black hole is formed, which is consistent with causality.

In contrast to the aforementioned, we could consider the total volume up to an arbitrary anchor line.[1] The result agrees at times $t > t_2$ but disagrees strongly before that time. We have

$$V'_{\rm AC}(t) = \begin{cases} \frac{1}{2}M & t \geq t_1\,, \\ 0 & t < t_1\,, \end{cases} \tag{2.34}$$

where t_1 is the time when the anchor curve crosses the shockwave line, see figure 2. This is long before the black hole is created[2] and it would imply that complexity starts growing at the moment the shockwave is released, see figure 3. For this reason, we prefer to use the stretched horizon as the anchor curve and only consider the volume inside the black hole. The volume integral can be cut off at either the event horizon or at the stretched horizon.

We can compare the complexity growth of the gravitational collapse model to the complexity growth of the eternal black hole, both at late time. The result for the gravitational collapse model is $V' = M/2$ for $t > t_2$, which is exactly half of the eternal black hole result, see (2.23). This factor $\frac{1}{2}$ corresponds to the fact that in the current case we only consider half of the volume slice (since we only have a one-sided black hole), as compared to the eternal black hole case (which is two-sided).

3 The RST model: complexity in a semi-classical black hole

The CGHS model can be extended such that one can study a semi-classical black hole analytically [17, 18]. Here we adopt a particular modification due to RST [8], which is

[1] In particular, we could imagine this anchor line to be asymptotically far away, in analogy to computations done in the AdS/CFT setup.

[2] In fact, as we move the anchor curve infinitely far away, the time difference of black hole creation and t_R^1 also goes to infinity.

JHEP03(2020)069

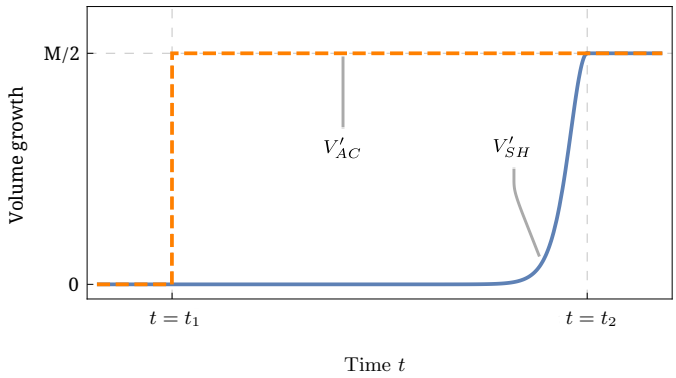


Figure 3. This graph demonstrates the difference between the volume growth $V'_{\rm SH}$ and $V'_{\rm AC}$. $V'_{\rm AC}$ starts growing at the time the shockwave crosses the anchor curve t_1 which is much earlier than the time t_2 the black hole is created.

given by

$$S_{\rm RST} = S_{\rm CGHS} + S_{\rm q} + S_{\rm ct}\,. \tag{3.1}$$

If one takes N, the number of matter fields f_i, to be large, then $S_{\rm q}$ represents the leading order quantum correction due to matter, which in conformal gauge reads

$$S_q = -\kappa \int_{\mathcal{M}} d^2x\, \partial_+\rho\partial_-\rho\,. \tag{3.2}$$

Here $\kappa := N/12$ can be thought of as playing the role of $\hbar$, which is put to unity. The additional RST counter term has the following form in conformal gauge,

$$S_{\rm ct} = -\kappa \int_{\mathcal{M}} d^2x\, \phi\partial_+\partial_-\rho\,. \tag{3.3}$$

This term is allowed by the symmetries of the model and when it is added the semi-classical field equations take a particularly simple form and are easily solved analytically.

The solutions of the equations of motion can be written in compact form if one defines new field variables,

$$\sqrt{\kappa}\Omega := e^{-2\phi} + \frac{\kappa}{2}\phi\,, \quad \sqrt{\kappa}\chi := e^{-2\phi} - \frac{\kappa}{2}\phi + \kappa\rho\,. \tag{3.4}$$

Using the new fields, the semi-classical action reads

$$S_{\rm RST} = 2\int_{\mathcal{M}} d^2x \left[-\partial_+\chi\partial_-\chi + \partial_+\Omega\partial_-\Omega + e^{\frac{2}{\sqrt{\kappa}}(\chi-\Omega)} + \frac{1}{2}\sum_{i=1}^{N}\partial_+ f_i\partial_- f_i\right]. \tag{3.5}$$

The RST model continues to enjoy the symmetry of the classical theory that allowed us to choose Kruskal coordinates, setting $\phi = \rho$ and as a result $\Omega = \chi$.

Again, we study an incoming shockwave of energy M at $x^+ = x_0^+$ of the form

$$T^f_{++} = \frac{M}{x_0^+}\delta(x^+ - x_0^+)\,. \tag{3.6}$$

JHEP03(2020)069

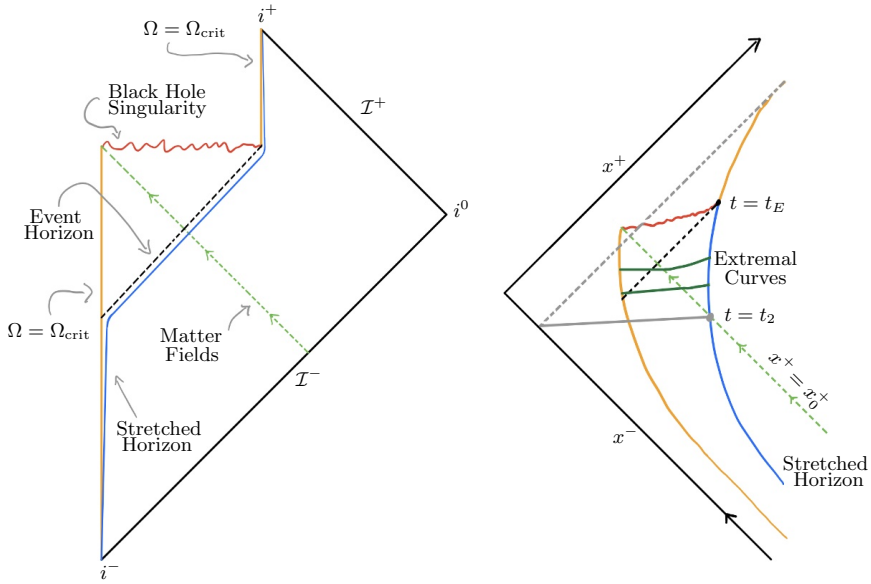


Figure 4. (*Left*) Cartoon of the Penrose diagram of the life cycle of an evaporating black hole formed by collapse. (*Right*) Depicts Kruskal diagram of the scenario illustrated in the left figure. The color coding coincides.

The semi-classical collapse solution of interest, given in Kruskal coordinates, is

$$\sqrt{\kappa}\Omega(x^+,x^-) = -x^+x^- + \left(x_0^+ - x^+\right)\frac{M}{x_0^+}\Theta(x^+ - x_0^+) - \frac{\kappa}{4}\ln\left(-x^+x^-\right) . \tag{3.7}$$

The solution describes flat spacetime for $x^+ \leq x_0^+$ and an evaporating black hole for $x^+ > x_0^+$, as shown in figure 4. It is important to note that in the RST model only those regions of spacetime where $\Omega(x^+,x^-) \geq \Omega_{\rm crit} := \frac{\sqrt{\kappa}}{4}\left(1 - \log\left(\frac{\kappa}{4}\right)\right)$ are considered physical. In the flat spacetime region inside the infalling shell, the boundary of the physical region is a timelike curve that can be interpreted as the origin in spherical coordinates in a higher-dimensional parent theory. For $x^+ > x_0^+$, the curve (x_S^+, x_S^-) for which $\Omega(x_S^+, x_S^-) = \Omega_{\rm crit}$ turns spacelike and defines the location of the black hole singularity. The semi-classical black hole evaporates and eventually the singularity terminates at an endpoint, after which the solution can be extended into a late time flat region where the physical boundary is again timelike. A more detailed description of the semi-classical geometry can for instance be found in [12].

Although we have closed expressions for the solutions of the RST model, we were not able to solve the extremization problem of a spacelike volume analytically and in order to check whether the complexity as volume prescription gives results consistent with general expectations we had to resort to numerical methods.

Since the extremization of

$$V = \int ds\, e^{-2\phi}\sqrt{g_{\mu\nu}\dot{y}^\mu\dot{y}^\nu} \tag{3.8}$$

JHEP03(2020)069

can be performed by solving corresponding Euler-Lagrange equations, the numerical problem is simply to solve a non-linear ordinary differential equation with appropriate boundary conditions, whose solutions provide a parametrisation of the extremized volume. Numerically integrating (3.8) inside the stretched horizon provides us with the volume complexity $V(t)$ at a value of tortoise time t that is determined by the choice of anchor curve in the same way as for a classical dynamical black hole. If we use a curve of constant Ω far from the black hole, then the complexity begins to grow very early, long before the incoming shockwave reaches the stretched horizon. If we instead use the stretched horizon as our anchor curve, then the complexity growth turns on essentially when the black hole is formed and this prescription is used in obtaining the numerical results presented in figure 5. We note, however, that the choice of anchor curve only affects the onset time and the early growth rate of the complexity, but not the slope of the curve (at leading order for large M/κ) showing the decreasing growth rate after a scrambling time has passed from the onset of complexity growth.

Note that, like in the case of classical collapse, it is not a priori clear what the appropriate boundary conditions at the origin are. We have chosen to apply boundary conditions analogous to (2.27) at $\Omega = \Omega_{\text{crit}}$ and then use a shooting algorithm to obtain the corresponding maximal volume curve. We apply the Weierstraß-Erdmann matching conditions to patch across the shockwave and then continue the numerical evaluation outwards.

Further, it could be reasoned that the factor $e^{-2\phi}$ in (3.8), which is interpreted as the area of the transverse two-sphere in the higher-dimensional theory, should be replaced by the quantum corrected area $\Omega - \Omega_{\text{crit}}$. We find that the slope of $V'(t)$ is not particularly sensitive to this replacement, at least not in the parameter range where our numerical evaluation is reliable (see below).

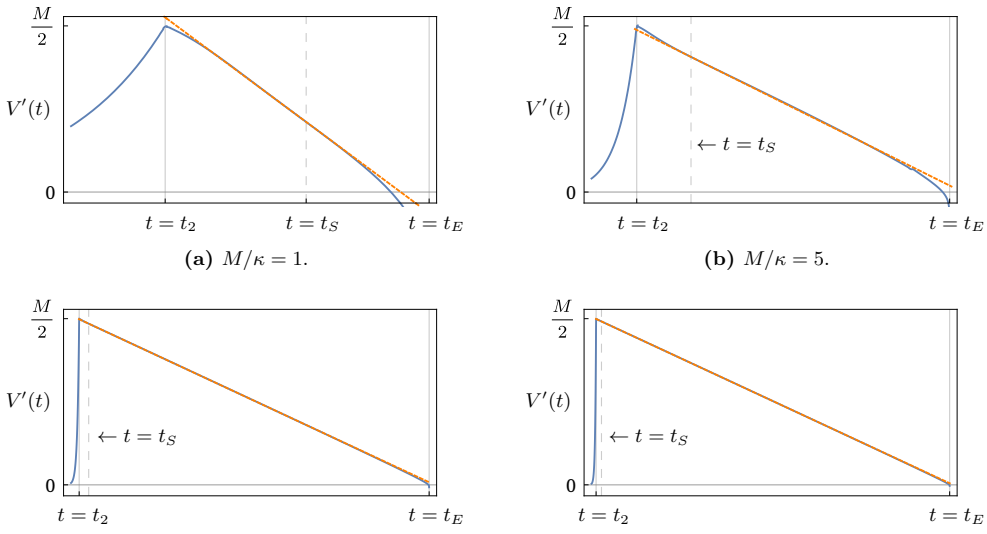
Results for the functions $V'(t)$ for different values of M/κ are plotted in figure 5. We observe that for reasonably high values of M/κ the numerical result is consistent with a linear decrease after the scrambling time $t_S = \log(4M/\kappa)$. We do not expect that the volume growth follows the linear trend forever, since eventually, the black hole is small enough, so that quantum corrections become strong on the stretched horizon. In this model, the coupling strength[3] on the stretched horizon at the scrambling time is approximately given by $\left(M - \frac{\kappa}{4} t_S\right)^{-\frac{1}{2}}$. For instance, a ratio $M/\kappa = 1$ yields a coupling strength of ≈ 1.2 which indicates we should not trust the result at all for this choice of parameters. This is reflected in figure 5a which barely exhibits a linear growth rate. On the other hand a ratio $M/\kappa = 100$ yields a coupling strength of ≈ 0.1 which demonstrates that we should be able to trust the solution for quite some time after the scrambling time. This is confirmed in figure 5c and 5d which shows a linearly falling growth rate over a long period of time.

4 Conclusion and outlook

In this paper we have computed holographic complexity as the volume of Einstein-Rosen bridges inside black holes in two-dimensional dilaton gravity models where we have explic-

[3] In this model the gravitational coupling is given by the value of $\left(e^{-2\phi} - \frac{\kappa}{4}\right)^{-\frac{1}{2}}$.

JHEP03(2020)069



(a) $M/\kappa = 1$. **(b)** $M/\kappa = 5$.

(c) $M/\kappa = 50$. **(d)** $M/\kappa = 100$.

Figure 5. Numerical results of volume growth for different values of M/κ. The black hole creation time is indicated by $t = t_2$ while the evaporation process is completed at time $t = t_E$. The blue curve depicts the numerical result while the dashed orange line is obtained by a linear extrapolation of the curve around the scrambling time t_S.

itly known semi-classical black hole solutions. This allows us to follow the time evolution of the complexity of a black hole that is formed in gravitational collapse and subsequently evaporates by emitting Hawking radiation. Our main results can be summarized in three statements.

First, in order to obtain sensible results, we have to calculate extremal volumes in the four-dimensional parent theory rather than lengths of geodesics in the two-dimensional reduction. The appropriately defined volume functional can be explicitly evaluated in the classical CGHS model and it exhibits the expected linear growth with Schwarzschild time at late times. At the same time it is easy to check that the length of the spacelike geodesics, that will otherwise arise, does not lend itself to a direct interpretation in terms of complexity.

Second, when considering dynamical black holes formed by gravitational collapse, we find it natural to cut off the volume integration at the stretched horizon of the black hole rather than extending the integration range to a distant anchor curve. This distinction is unimportant if all we are interested in is the late time rate of growth of the complexity for a classical black hole but it does affect the onset of complexity growth. If the volume prescription extends to a distant anchor curve then the complexity already starts growing as soon the infalling shockwave passes the anchor point and the complexity growth turns on abruptly. If, on the other hand, we use the stretched horizon to delimit the integration range, then complexity growth turns on smoothly at a time that coincides with the onset of Hawking emission at the semi-classical level.

JHEP03(2020)069

Third, using numerical methods, we obtain the complexity of an evaporating black hole as a function of time using the volume prescription inside the stretched horizon. We find that after the black hole is created, the complexity growth needs a time period of order the scrambling time to settle to a rate of growth proportional to the area of the stretched horizon. The growth rate then reliably tracks the area of the horizon as it shrinks due to black hole evaporation. Towards the end of the black hole lifetime, higher order quantum corrections are expected to become important and semi-classical calculations can no longer be trusted.

Holographic complexity can also be calculated in these models using the Wheeler-DeWitt action formalism. We will present our results on that in a forthcoming companion paper [6], where we find that the numerical results obtained in the present paper for the volume complexity are confirmed by action calculations that can be carried out analytically even at the semi-classical level.

Our results fit very well with Susskind's argument [2], that the rate of complexity growth for an evaporating black hole should at any given time be proportional to the product of the black hole entropy S and the Hawking temperature T at that time. In the models we are considering, the Hawking temperature remains constant and the entropy is proportional to black hole mass. It follows that the black hole loses area at a constant rate determined by N the number of matter channels available for Hawking emission. This translates into a growth rate of complexity that is initially proportional to the initial mass of the black hole and then drops linearly with time until the black hole has completely evaporated. In other words, $\dot{\mathcal{C}} \propto ST$ should decrease linearly in time. This precisely the behaviour we see in our numerical calculations, following an initial onset period of order the scrambling time, which is logarithmic in M in these models.

Our results also support the notion that the holographic complexity corresponds to the quantum complexity of the combined system of black hole and emitted Hawking radiation [2]. The scrambling dynamics that generates the growth in complexity takes place at the stretched horizon and the growth rate is reduced as the area of the horizon shrinks. The outgoing radiation is free streaming and no further complexity is generated by the degrees of freedom that have been emitted from the black hole. At the end of the day, the black hole has disappeared and all that is left is a long train of outgoing radiation in a state of high, but no longer growing, complexity.

Acknowledgments

It is a pleasure to thank Shira Chapman and Nick Poovuttikul for discussions. This research was supported by the Icelandic Research Fund under grants 185371-051 and 195970-051, and by the University of Iceland Research Fund.

JHEP03(2020)069

References

[1] L. Susskind, L. Thorlacius and J. Uglum, *The Stretched horizon and black hole complementarity*, *Phys. Rev.* **D 48** (1993) 3743 [hep-th/9306069] [INSPIRE].

[2] L. Susskind, *Three Lectures on Complexity and Black Holes*, arXiv:1810.11563 [INSPIRE].

[3] L. Susskind, *Computational Complexity and Black Hole Horizons*, *Fortsch. Phys.* **64** (2016) 44 [arXiv:1403.5695] [INSPIRE].

[4] A.R. Brown, D.A. Roberts, L. Susskind, B. Swingle and Y. Zhao, *Holographic Complexity Equals Bulk Action?*, *Phys. Rev. Lett.* **116** (2016) 191301 [arXiv:1509.07876] [INSPIRE].

[5] A.R. Brown, D.A. Roberts, L. Susskind, B. Swingle and Y. Zhao, *Complexity, action and black holes*, *Phys. Rev.* **D 93** (2016) 086006 [arXiv:1512.04993] [INSPIRE].

[6] L. Schneiderbauer, W. Sybesma and L. Thorlacius, *Action Complexity for Semi-Classical Black Holes*, arXiv:2001.06453 [INSPIRE].

[7] C.G. Callan Jr., S.B. Giddings, J.A. Harvey and A. Strominger, *Evanescent black holes*, *Phys. Rev.* **D 45** (1992) R1005 [hep-th/9111056] [INSPIRE].

[8] J.G. Russo, L. Susskind and L. Thorlacius, *Black hole evaporation in* $(1+1)$*-dimensions*, *Phys. Lett.* **B 292** (1992) 13 [hep-th/9201074] [INSPIRE].

[9] S.B. Giddings and A. Strominger, *Dynamics of extremal black holes*, *Phys. Rev.* **D 46** (1992) 627 [hep-th/9202004] [INSPIRE].

[10] T. Banks, A. Dabholkar, M.R. Douglas and M. O'Loughlin, *Are horned particles the climax of Hawking evaporation?*, *Phys. Rev.* **D 45** (1992) 3607 [hep-th/9201061] [INSPIRE].

[11] J.A. Harvey and A. Strominger, *Quantum aspects of black holes*, in proceedings of the *Theoretical Advanced Study Institute (TASI 92): From Black Holes and Strings to Particles*, Boulder, U.S.A., 1–26 June 1992, pp. 537–588, [hep-th/9209055] [INSPIRE].

[12] L. Thorlacius, *Black hole evolution*, *Nucl. Phys. Proc. Suppl.* **41** (1995) 245 [hep-th/9411020] [INSPIRE].

[13] A. Strominger, *Les Houches lectures on black holes*, in proceedings of the *NATO Advanced Study Institute: Les Houches Summer School, Session 62: Fluctuating Geometries in Statistical Mechanics and Field Theory*, Les Houches, France, 2 August–9 September 1994, hep-th/9501071 [INSPIRE].

[14] S.B. Giddings, *Quantum mechanics of black holes*, in proceedings of the *Summer School in High-energy physics and cosmology*, Trieste, Italy, 13 June–29 July 1994, pp. 0530–574 [hep-th/9412138] [INSPIRE].

[15] A.R. Brown, H. Gharibyan, H.W. Lin, L. Susskind, L. Thorlacius and Y. Zhao, *Complexity of Jackiw-Teitelboim gravity*, *Phys. Rev.* **D 99** (2019) 046016 [arXiv:1810.08741] [INSPIRE].

[16] S. Chapman, H. Marrochio and R.C. Myers, *Holographic complexity in Vaidya spacetimes. Part I*, *JHEP* **06** (2018) 046 [arXiv:1804.07410] [INSPIRE].

[17] A. Bilal and C.G. Callan Jr., *Liouville models of black hole evaporation*, *Nucl. Phys.* **B 394** (1993) 73 [hep-th/9205089] [INSPIRE].

[18] S.P. de Alwis, *Quantization of a theory of* $2D$ *dilaton gravity*, *Phys. Lett.* **B 289** (1992) 278 [hep-th/9205069] [INSPIRE].

JHEP03(2020)069

Article II

t

Published for SISSA by Springer

Received: *February 4, 2020*
Accepted: *June 10, 2020*
Published: *July 23, 2020*

JHEP07(2020)173

Action complexity for semi-classical black holes

Lukas Schneiderbauer, Watse Sybesma and Lárus Thorlacius

Science Institute, University of Iceland,
Dunhaga 3, Reykjavík 107, Iceland

E-mail: lukas.schneiderbauer@gmail.com, watse@hi.is, lth@hi.is

Abstract: We adapt the complexity as action prescription (CA) to a semi-classical model of two-dimensional dilaton gravity and determine the rate of increase of holographic complexity for an evaporating black hole. The results are consistent with our previous numerical results for semi-classical black hole complexity using a volume prescription (CV) in the same model, but the CA calculation is fully analytic and provides a non-trivial positive test for the holographic representation of the black hole interior.

Keywords: 2D Gravity, Black Holes

ArXiv ePrint: 2001.06453

https://doi.org/10.1007/JHEP07(2020)173

Contents

JHEP07(2020)173

1 Introduction

The quantum complexity of a black hole is generated by the scrambling dynamics of quantum mechanical degrees of freedom that are enumerated by the black hole entropy. These degrees of freedom can be usefully modelled in terms of a quantum circuit with k-local gates acting on a finite number of qubits.[1] In line with black hole complementarity [2], the qubits can be taken to be located at (or near) a stretched horizon just outside the event horizon and the black hole interior is then viewed as an emergent spacetime region that provides a dual geometric description of the quantum dynamics of the stretched horizon degrees of freedom. In particular, the expanding spatial volume of the Einstein-Rosen bridge of a two-sided eternal black hole is conjectured to reflect the growing complexity of the corresponding quantum state [3]. A refined version of the conjecture instead relates the complexity to the gravitational action evaluated on a specific bounded region of the black hole spacetime, referred to as the Wheeler-DeWitt (WdW) patch, that intersects the black hole interior [4, 5].

The volume (CV) and action (CA) complexity conjectures have been explored for a variety of black hole geometries in Einstein gravity as well as extended to black hole

[1] For a recent review of complexity and black holes, see [1].

solutions in other gravitational theories. Much of the attention and effort has been focused on stationary black hole solutions while there have been fewer studies of dynamical black holes (see e.g. [6–9]). Black holes are stable in classical gravity and at late times, long after its formation by gravitational collapse, the geometry of a dynamical black hole will closely resemble that of a stationary one. Similarly, in the late time limit, the rate of growth of the holographic complexity (both CV and CA) of a classical dynamical black hole reduces to that of a stationary black hole of the same mass and other conserved charges (up to a factor of two accounting for the two-sided nature of the maximally extended stationary solution). This changes, however, when semi-classical effects are taken into account. In asymptotically flat spacetime, black hole evaporation due to Hawking emission results in the steady reduction of the black hole area and ends in a final state where there is an outgoing train of Hawking radiation but no black hole. For a large initial black hole mass, the quantum complexity of this final state will be very large, but finite, and presumably no longer growing as the Hawking radiation free streams outwards [1].

In [10] we initiated the study of holographic complexity in semi-classical gravity with the aim of testing the geometric representation of quantum complexity in the context of black hole evaporation. In order to have analytical control, we considered semi-classical toy models, the Callan, Giddings, Harvey, and Strominger (CGHS) and Russo, Susskind, and Thorlacius (RST) models of two-dimensional dilaton gravity [11, 12], that arise from the near-horizon limit of a near-extremal charged dilaton black hole in higher dimensions. The RST model is particularly well suited for our purposes. The formation of an RST black hole by gravitational collapse, and its subsequent evaporation in asymptotically flat spacetime, is described by an analytic solution of the field equations and this allows us to make precise statements about the complexity of an evaporating black hole. The complexity of two-dimensional black holes has been studied previously in the Jackiw-Teitelboim model in [13, 14], but in this case the two-dimensional spacetime is asymptotically AdS_2 and the black holes do not evaporate.

In [10] we found non-trivial agreement between the volume of certain extremal surfaces and the expected behaviour of holographic complexity of classical CGHS black holes and evaporating RST black holes, respectively. We restricted our attention to volume complexity (CV), using a suitably defined volume functional that corresponds to spatial volume in the higher dimensional parent theory rather than geodesic length in the two-dimensional theory. For classical CGHS black holes, the volume complexity grows at a constant rate which is proportional to the product of Bekenstein-Hawking entropy and the Hawking temperature,

$$\frac{d\mathcal{C}_V}{dt} \propto S\,T\,, \tag{1.1}$$

as is expected on general grounds [1]. Here t is the proper time of a distant fiducial observer and (1.1) holds both for two-sided eternal black holes and at late times for dynamical black holes formed by gravitational collapse. For semi-classical RST black holes, on the other hand, complexity growth slows down as the black hole evaporates and the rate of growth approaches zero at the endpoint of evaporation. While the extremal volume could be obtained analytically for classical CGHS holes, we had to rely on numerical evaluation for

JHEP07(2020)173

semi-classical RST black holes in [10]. Our numerical results confirmed that at leading order in a semi-classical expansion the rate of growth of the complexity, when expressed as a function of the proper time of a distant fiducial observer, is proportional to the product $S\,T$ for most of the black hole lifetime.

In the present paper we extend this work by evaluating the holographic complexity of semi-classical black holes in terms of an action on a Wheeler-DeWitt patch. While this is more technically involved than the volume computations in [10], it has the distinct advantage that the entire semi-classical calculation can be carried out analytically and yields explicit results for the rate of complexity growth of an evaporating RST black hole throughout its evolution. The relation (1.1) carries over to the semi-classical theory with the entropy at time t given to leading order by the time-dependent area of the black hole. The next-to-leading order (logarithmic) term in the Bekenstein-Hawking entropy can also be read off from our analytic expression for the complexity growth and the result agrees with previous semi-classical entropy calculations in the RST model [15, 16]. The numerical results for volume complexity obtained in our previous work [10] are consistent with the new analytic results for action complexity. The fact that our holographic complexity calculations for a semi-classical black hole geometry reproduce the time-dependent rate of growth of complexity expected for the quantum dynamics of an evolving stretched horizon, provides a non-trivial positive test for black hole complementarity and the holographic duality between the stretched horizon and the black hole interior. At the same time, it supports the validity of the holographic complexity conjectures themselves in a new dynamical regime.

A Wheeler-DeWitt patch is bounded by co-dimension one and co-dimension two surfaces in spacetime. The associated gravitational action must include boundary terms in order to make the variational problem well-posed. For time-like and space-like co-dimension one boundaries the appropriate boundary terms in the two-dimensional theory are obtained from the standard Gibbons-Hawking-York term in the higher-dimensional parent theory, while contributions from null boundaries and co-dimension two boundaries require a more careful treatment [17, 18]. By working in so-called Kruskal gauge and arranging the boundary terms in the action to respect the same symmetry that simplifies the bulk field equations of the RST model [12], we are able to eliminate a certain ambiguity in the holographic action complexity and obtain a remarkably simple end result.

The structure of the paper is as follows. In section 2, we carefully develop the boundary terms needed to have a well-posed variational problem for the two-dimensional dilaton gravity models. We then present the holographic action complexity (CA) for a classical CGHS black hole formed in gravitational collapse in section 3, followed by the corresponding semi-classical calculation for an evaporating RST black hole in section 4. We end with a brief discussion and outlook for future work.

2 RST model

The action of the semi-classical RST model consists of three terms,

$$S_{\rm bulk} = S_0 + S_{\rm q} + S_{\rm ct}\,, \tag{2.1}$$

JHEP07(2020)173

where

$$S_0 = \int_{\mathcal{M}} d^2y\sqrt{-g}\left[e^{-2\phi}\left(R + 4(\nabla\phi)^2 + 4\lambda^2\right) - \frac{1}{2}\sum_{i=i}^{N}(\nabla f_i)^2\right] \tag{2.2}$$

is the classical CGHS action involving a two-dimensional metric $g_{\mu\nu}$ along with a dilaton field ϕ and N scalar matter fields f_i. The length scale λ^{-1} is set by the magnetic charge of the higher-dimensional near extremal black hole and from now on we work in units where $\lambda = 1$. The second term,

$$S_{\rm q} = -\frac{\kappa}{4}\int d^2y\sqrt{-g}\left(R\frac{1}{\nabla^2}R\right), \tag{2.3}$$

with $\kappa = N/12$, was introduced by Callan et al. in [11] and captures the one-loop correction to the quantum effective action due to the conformal anomaly of the matter fields. For large N this term dominates over one-loop effects coming from the dilaton gravity sector that can therefore be ignored. The third term in the semi-classical action,

$$S_{\rm ct} = -\frac{\kappa}{2}\int d^2y\sqrt{-g}\phi R\,, \tag{2.4}$$

was introduced by Russo et al. in [12]. This term is allowed by general covariance and does not disrupt the classical physics obtained in the limit $e^{-2\phi} \gg \kappa$. It enters at the same order as $S_{\rm q}$ and serves to preserve the classical symmetry of S_0 generated by the current $\partial^\mu(\rho - \phi)$, where $e^{2\rho}$ is the conformal factor of the metric $g_{\mu\nu}$ with respect to a flat reference metric. We set $\hbar = 1$ throughout but note that when $\hbar$ is retained in the action it accompanies the prefactor κ and thus any expression involving κ will be directly related to quantum corrections in the semi-classical theory.

2.1 A well-posed variational principle

The CA proposal instructs us to evaluate the on-shell action of the model in question on a so-called WdW patch [4, 5]. However, it is well known that the action associated to a given set of equations of motion is not unique. For instance, adding boundary terms does not change the equations of motion but will in general change the value of the action itself. To restrict the set of possible actions, the CA proposal comes with the further requirement that the variational principle on the WdW patch should be well-posed. The equations of motion should follow from a stationary action principle assuming appropriate boundary conditions on the boundary of the WdW patch. A solution to this problem was presented for Einstein-gravity in [18], where a particular set of co-dimension one boundary terms and co-dimension two joint terms were proposed. In general, these terms are still not unique, but the requirements imposed in [18] were enough to ensure a unique answer for the late-time complexity growth rate in well-known classical black hole geometries. This is not the case, however, for the semi-classical model we consider below. Indeed, when we calculate the complexity growth for dynamical solutions that describe evaporating black holes, we find that certain boundary terms can be added that change the value of the action on the WdW patch while leaving the variational principle well-posed. One therefore has to introduce further criteria, beyond those considered in [18], in order to have a definite prescription for

JHEP07(2020)173

the holographic action complexity. A similar issue arises when the CA prescription is used to evaluate the complexity of two-dimensional Jackiw-Teitelboim black holes [13, 14]. In this case, the rate of complexity growth is found to be crucially influenced by boundary terms and additional physical input is required to fully determine the complexity. As explained in detail below, a sufficient criterion for the situation at hand is to impose on the full action, including boundary terms, the same symmetry that allowed the semi-classical field equations of the RST model to be analytically solved in the first place in [12].

To obtain a well-posed variational problem, we adapt the procedure proposed by [18] to a two-dimensional dilaton-gravity theory. However, a direct application is obstructed by the non-local term $S_{\rm q}$. One way to remedy this problem is to introduce an auxiliary scalar field Z and write the action in terms of integrals over local quantities only [19],

$$S_{\rm bulk} = \int_{\mathcal{M}} d^2y\sqrt{-g}\left[R\tilde{\chi} + 4\left((\nabla\phi)^2+1\right)e^{-2\phi} - \frac{\kappa}{4}(\nabla Z)^2 - \frac{1}{2}\sum_{i=1}^{N}(\nabla f_i)^2\right] \tag{2.5}$$

with

$$\tilde{\chi} := e^{-2\phi} - \frac{\kappa}{2}(\phi - Z)\,. \tag{2.6}$$

As one can easily check, integrating out the auxiliary field Z will return the original non-local action, up to boundary terms.

Considering a region with piecewise smooth space-like, time-like or null boundaries, the prescription of [18] gives the following boundary terms involving the combination of fields $\tilde{\chi}$ that multiplies the Ricci scalar R in the bulk action,

$$\begin{aligned} S_{\rm boundary} = {} & 2\sum_{\mathcal{S}\in\mathscr{S}}\sigma_{\mathcal{S}}\int_{\mathcal{S}}\sqrt{|h|}\tilde{\chi}K d\Sigma \\ & + 2\sum_{\mathcal{N}\in\mathscr{N}}\sigma_{\mathcal{N}}\int_{\mathcal{N}}d\lambda\tilde{\chi}\boldsymbol{\kappa} + 2\sum_{\mathcal{N}\in\mathscr{N}}\sigma_{\mathcal{N}}\int_{\mathcal{N}}d\lambda\,\partial_\lambda\tilde{\chi}\log|\partial_\lambda l| \\ & + 2\sum_{j\in\text{joints}}\sigma_j\tilde{\chi}\big|_j a_j\,. \end{aligned} \tag{2.7}$$

The first term on the right hand side is the analogue of the familiar Gibbons-Hawking-York term, where K is the extrinsic curvature of each time/space-like boundary component $\mathcal{S}\in\mathscr{S}$ and h is the determinant of the induced metric on $\mathcal{S}$. The terms on the second line of (2.7) accompany null boundary components $\mathcal{N}\in\mathscr{N}$. The integration variable λ parametrizes the null line $\mathcal{N}$. The failure of λ to be an affine parameter is measured by $\boldsymbol{\kappa}$, defined by the equation[2]

$$k^\alpha\nabla_\alpha k^\beta = \boldsymbol{\kappa}k^\beta \tag{2.8}$$

with $k^\alpha := \frac{\partial\gamma^\alpha}{\partial\lambda}$ and $\gamma^\alpha(\lambda)$ being coordinates of the null curve $\mathcal{N}$ parametrized by λ. The first term on the second line of (2.7) is not invariant under reparametrizations $\lambda\mapsto\lambda'$ by itself and the second term is added to offset this pathological feature.[3] Here l can a priori be an arbitrary function of any scalar field, provided $\partial_\lambda l$ does not vanish anywhere.

[2]Unfortunately, conventions dictate using the Greek letter kappa both in this context and as $\kappa = N/12$. We've opted for using boldface for one of them to reduce the scope for confusion.

[3]See [18] for a detailed analysis.

JHEP07(2020)173

Finally, for each non-smooth joint j, we have to add a term a_j which depends on the type and position of the joint. More explicitly, in the case of joints formed by two curves $\mathcal{S}_1$ and $\mathcal{S}_2$, that are separately either spacelike or timelike, one finds

$$a = \log |(n_1 + p_1)_\mu n_2^\mu| \,, \tag{2.9}$$

where n_i are unit normal vectors to $\mathcal{S}_i$ and p_1 is a tangent vector of $\mathcal{S}_1$ that points outwards from the region of interest.

In case of a joint of two null-lines parametrized by λ and $\bar{\lambda}$ respectively (and corresponding vectors k^α and $\bar{k}^\alpha$), a reads

$$a = \log \left| \frac{1}{2} k_\mu \bar{k}^\mu \right| , \tag{2.10}$$

while in case of a joint between a null and a space- or timelike boundary component, we have

$$a = \log |k^\mu n_\mu| \tag{2.11}$$

where k^μ corresponds to the null boundary in the way explained above and n^μ is the unit normal vector associated to the space- or timelike boundary.

The various terms in (2.7) are accompanied by signs $\sigma_\mathcal{S}$, $\sigma_\mathcal{N}$ and σ_j that are sensitive to the conventions adapted in the procedure. A coherent set of rules is presented in [18].

Now, consider adding to the action a boundary term of the form

$$\int_\mathcal{S} \sqrt{|h|} g(\phi, Z) d\Sigma \tag{2.12}$$

involving an arbitrary function $g(\phi, Z)$. Adding a boundary term does not alter the equations of motion and a term of this particular form will not influence the variational principle if we impose Dirichlet boundary conditions, i.e. take the variation of the induced metric h and the variation of the scalar fields ϕ and Z to vanish at the boundary. However, it is easy to see that such a term can drastically change the result of holographic complexity in our setup. Furthermore, considering regions with null-boundaries, (2.7) depends on an undetermined function l, which again influences the holographic complexity. The above prescription thus needs to be supplemented by additional restrictions, as discussed below.

2.2 RST symmetry

In order to overcome the troublesome arbitrariness in the choice of boundary terms, we propose to restrict the allowed terms by an invariance requirement of the total action S under the symmetry which guided the definition of the RST model in the first place.

In the following, we work in conformal gauge, where the line element takes the form

$$ds^2 = -e^{2\rho} dy^+ dy^- \,. \tag{2.13}$$

Recall that the term $S_{\rm ct}$, given by (2.4), was introduced to preserve the symmetry generated by the current $\partial^\mu(\rho - \phi)$. The corresponding infinitesimal transformation of the fields ϕ and ρ are given by [12, 20]

$$\delta_{\rm RST}\, \phi = \delta_{\rm RST}\, \rho = \frac{1}{2} \delta_{\rm RST}\, Z = \epsilon \, \frac{e^{2\phi}}{1 - \frac{\kappa}{4} e^{2\phi}} \,, \tag{2.14}$$

while the matter fields do not transform.

JHEP07(2020)173

We now impose the additional requirement that the total action S, *including* boundary terms, remains invariant under the RST transformation,

$$\delta_{\rm RST}\, S = 0\,. \tag{2.15}$$

The bulk action (2.5) is invariant under $\delta_{\rm RST}$ up to a boundary term that will have to be cancelled. Going back to the example (2.12), it is evident that, generically, the RST variation of such a term will not vanish. We can use this to our advantage and choose the additional boundary term so that its variation cancels against the variation of the bulk action.

In order to work out the RST variation of the action $S_{\rm bulk} + S_{\rm boundary}$, it is convenient to define the fields[4]

$$\begin{aligned} \Omega &:= e^{-2\phi} + \frac{\kappa}{2}\phi\,, \\ \chi &:= e^{-2\phi} + \kappa\rho - \frac{\kappa}{2}\phi\,, \end{aligned} \tag{2.16}$$

for which the bulk action (2.5) can be written as

$$\begin{aligned} S_{\rm bulk} = &\int d^2y\sqrt{-g}\left[\frac{1}{\kappa}\left(\nabla\chi\right)^2 - \frac{1}{\kappa}\left(\nabla\Omega\right)^2 + \frac{2}{\sqrt{-g}}e^{\frac{2}{\kappa}(\chi-\Omega)} - \frac{1}{2}\sum_{i=1}^{N}\left(\nabla f_i\right)^2\right] \\ &- 2\int d^2y\sqrt{-g}\left[\nabla\left(\tilde{\chi}\nabla\rho\right) + \frac{\kappa}{8}\nabla\left(\eta\nabla\eta\right)\right]\,, \end{aligned} \tag{2.17}$$

where η is a harmonic field, $\nabla^2\eta = 0$, obtained from the auxiliary Z field via $Z = 2\rho + \eta$.

The variations of the new fields Ω and χ can easily be evaluated and yield

$$\delta_{\rm RST}\,\Omega = \delta_{\rm RST}\,\chi = -2\epsilon\,, \tag{2.18}$$

while $\delta_{\rm RST}\,\eta = 0$. It is now evident that the RST variation of the first line of (2.17) vanishes and also the variation of the last term on the second line. The remaining non-vanishing RST variation of the first total derivative term cancels against a contribution coming from the Gibbons-Hawking-York (GHY) term that we consider next.

2.2.1 Gibbons-Hawking-York boundary terms

Let us now consider the GHY term in (2.7) of the form

$$2\sigma_{\mathcal{S}}\int_{\mathcal{S}}\sqrt{|h|}\tilde{\chi}K d\Sigma\,, \tag{2.19}$$

with $K = \nabla_\mu n^\mu$ where n^μ is a unit normal vector to the surface $\mathcal{S}$.

In conformal gauge (2.13), the following identity

$$\sqrt{|h|}\nabla_\mu n^\mu = \sqrt{|h_0|}\partial_\mu n_0^\mu + \sqrt{|h|}n^\mu\partial_\mu\rho \tag{2.20}$$

[4]Note that this definition differs slightly from the one employed in [10].

JHEP07(2020)173

holds, where quantities with subscript 0 are to be evaluated with respect to the flat reference metric $ds^2 = -dy^+dy^-$. This implies that the GHY boundary term can be expressed as

$$2\sigma_{\mathcal{S}} \int_{\mathcal{S}} \sqrt{|h|}\tilde{\chi} K \, d\Sigma = 2\sigma_{\mathcal{S}} \int_{\mathcal{S}} \sqrt{|h_0|}\tilde{\chi} K_0 \, d\Sigma + 2\sigma_{\mathcal{S}} \int_{\mathcal{S}} \sqrt{|h|}\tilde{\chi} n^\mu \partial_\mu \rho \, d\Sigma \,, \tag{2.21}$$

and furthermore, due to Stokes' theorem (assuming for the moment a region without null boundaries), the second term can be written as

$$2 \sum_{\mathcal{S}\in\mathscr{S}} \sigma_{\mathcal{S}} \int_{\mathcal{S}} \sqrt{|h|}\tilde{\chi} n^\mu \partial_\mu \rho \, d\Sigma = 2 \int_{\mathcal{M}} d^2y \sqrt{-g} \nabla(\tilde{\chi} \nabla \rho) \,, \tag{2.22}$$

which precisely cancels the first term in the second line of the bulk action (2.17). This leaves us only with the term involving the flat reference metric in (2.21). This term does not contribute when we obtain field equations using field variations (under which the flat reference metric is fixed). However, its RST variation does not vanish in general,

$$\delta_{\mathrm{RST}} \left(2 \sum_{\mathcal{S}} \sigma_{\mathcal{S}} \int_{\mathcal{S}} \sqrt{|h_0|}\tilde{\chi} K_0 \, d\Sigma \right) = -4\epsilon \sum_{\mathcal{S}} \sigma_{\mathcal{S}} \int_{\mathcal{S}} \sqrt{|h_0|} K_0 \, d\Sigma \neq 0 \,. \tag{2.23}$$

This variation can be cancelled by introducing a suitably chosen additional boundary term.

2.2.2 Time-/spacelike joint contributions

Still assuming no null boundaries, this leaves us with the analysis of the joint contributions to the action S_{boundary} in (2.7), i.e.

$$2 \sum_{j\in\mathrm{joints}} \sigma_j \tilde{\chi}\big|_j \log |(n_1^j + p_1^j)_\mu (n_2^j)^\mu| \,. \tag{2.24}$$

It turns out that in two dimensions, these terms are actually not necessary to obtain a well-posed variational principle. The reason is that the argument of the logarithm does not depend on the conformal factor at all. This is easily seen by writing the unit normal vector as

$$n_\mu = \frac{\partial_\mu \Phi}{\sqrt{g^{\sigma\rho}\partial_\sigma\Phi\partial_\rho\Phi}} = e^\rho \frac{\partial_\mu \Phi}{\sqrt{g_0^{\sigma\rho}\partial_\sigma\Phi\partial_\rho\Phi}} = e^\rho (n_0)_\mu \tag{2.25}$$

where Φ is temporarily introduced as a scalar field whose contour lines describe the surface $\mathcal{S}$ locally. Similarly,

$$n^\mu = e^{-\rho} n_0^\mu \tag{2.26}$$

and the same is true for the tangent vector p_1. It follows that the inner product is independent of the conformal factor ρ.

Since the term is proportional to $\tilde{\chi}$, its RST variation is easily evaluated,

$$\delta_{\mathrm{RST}} \left(2\sigma_j \tilde{\chi}\big|_j a_j \right) = -4\epsilon \, \sigma_j a_j, \tag{2.27}$$

which does not vanish in general. The RST symmetry is easily enforced by simply leaving out joint terms of this form. This is possible, because, as we have just seen, such terms do not influence the variational principle in two dimensions.

JHEP07(2020)173

2.2.3 Null boundary contributions

Let us now include null boundaries in our discussion. It will be beneficial to rewrite the terms in (2.7) corresponding to null boundaries,

$$2\sigma_{\mathcal{N}} \int_{\mathcal{N}} d\lambda \tilde{\chi} \boldsymbol{\kappa} + 2\sigma_{\mathcal{N}} \int_{\mathcal{N}} d\lambda \, \partial_\lambda \tilde{\chi} \log |\partial_\lambda l|, \tag{2.28}$$

in a way that is manifestly reparametrization invariant. This is achieved by integrating the second term by parts which gives

$$2\sigma_{\mathcal{N}} \int_{\mathcal{N}} d\lambda \partial_\lambda \tilde{\chi} \log |\partial_\lambda l| = 2\sigma_{\mathcal{N}} \tilde{\chi} \log |\partial_\lambda l| \Big|_1^2 - 2\sigma_{\mathcal{N}} \int_{\mathcal{N}} d\lambda \tilde{\chi} \partial_\lambda \log |\partial_\lambda l| \tag{2.29}$$

and using that

$$\partial_\lambda \log |\partial_\lambda l| = \boldsymbol{\kappa} + \frac{k_-^\mu k_-^\nu \nabla_\mu \partial_\nu l}{k_-^\sigma \partial_\sigma l}, \tag{2.30}$$

we note, that the term involving $\boldsymbol{\kappa}$ cancels the original $\boldsymbol{\kappa}$-dependent term in (2.28), so that in total we have

$$(2.28) = 2\sigma_{\mathcal{N}} \tilde{\chi} \log |\partial_\lambda l| \Big|_1^2 - 2\sigma_{\mathcal{N}} \int_{\mathcal{N}} d\lambda \tilde{\chi} \frac{k_-^\mu k_-^\nu \nabla_\mu \partial_\nu l}{k_-^\sigma \partial_\sigma l}. \tag{2.31}$$

The second term is now manifestly invariant under a change of parametrization $\lambda \mapsto \lambda' := e^\beta \lambda$, since $k^\alpha \mapsto e^{-\beta} k^\alpha$. When the original joint terms are combined with the new terms obtained from integration by parts in (2.31) the full expression is also reparametrization invariant. Independent of the precise nature of the joints, the original contribution will be of the form

$$2\sigma \tilde{\chi} \log |A \, m_\mu k^\mu|, \tag{2.32}$$

where A is a constant, k^α is the null vector associated with the null boundary in question, and m^μ is a vector that depends on nature of the joint, which we will be agnostic about for this argument. The sign σ depends on conventions, but the relative sign to $\sigma_{\mathcal{N}}$ is fixed by

$$\sigma = \begin{cases} -\sigma_{\mathcal{N}} & \text{if joint lies in the future of } \mathcal{N} \\ +\sigma_{\mathcal{N}} & \text{if joint lies in the past of } \mathcal{N} \end{cases} \tag{2.33}$$

assuming k^μ is future directed. This implies, that for the total joint contributions on either side of the null boundary, we obtain

$$(2.32) - 2\sigma \tilde{\chi} \log |\partial_\lambda l| = 2\sigma \tilde{\chi} \log \left| A \frac{m_\mu k^\mu}{\partial_\lambda l} \right|, \tag{2.34}$$

which is also manifestly invariant under reparametrization $\lambda \mapsto \lambda'$.

We have now successfully rewritten the terms corresponding to a null boundary and its attached joints in a manifestly reparametrization invariant form. Next, in order to obtain the RST variation, it will again be convenient to split those terms into parts which depend on the conformal factor and parts which do not. We have

$$-2\sigma_{\mathcal{N}} \int_{\mathcal{N}} d\lambda \tilde{\chi} \frac{k_-^\mu k_-^\nu \nabla_\mu \partial_\nu l}{k_-^\sigma \partial_\sigma l} = -2\sigma_{\mathcal{N}} \int_{\mathcal{N}} d\lambda \tilde{\chi} \frac{k_-^\mu k_-^\nu \partial_\mu \partial_\nu l}{k_-^\sigma \partial_\sigma l} + 4\sigma_{\mathcal{N}} \int_{\mathcal{N}} d\lambda \tilde{\chi} \partial_\lambda \rho, \tag{2.35}$$

JHEP07(2020)173

where the first term is reparametrization invariant and does not depend on the conformal factor ρ.

Considering the joint terms (2.34), we can write them as

$$\begin{aligned} &-2\sigma_{\mathcal{N}}\tilde{\chi}_2 \log\left|A\,\frac{m_\mu k^\mu}{\partial_\lambda l}\right|_2 + 2\sigma_{\mathcal{N}}\tilde{\chi}_1 \log\left|\bar{A}\,\frac{\bar{m}_\mu k^\mu}{\partial_\lambda l}\right|_1 \\ &\quad = -2\sigma_{\mathcal{N}}\int_{\mathcal{N}} d\lambda\partial_\lambda\left(\tilde{\chi}\log(e^\rho)\right) - 2\sigma_{\mathcal{N}}\tilde{\chi}_2 \log\left|A\,\frac{e^{-\rho}m_\mu k^\mu}{\partial_\lambda l}\right|_2 + 2\sigma_{\mathcal{N}}\tilde{\chi}_1 \log\left|\bar{A}\,\frac{e^{-\rho}\bar{m}_\mu k^\mu}{\partial_\lambda l}\right|_1 . \end{aligned} \tag{2.36}$$

One can easily check that $e^{-\rho}m_\mu$ does not depend on ρ in case of a joint with a space- or timelike curve (since then m^μ is given by a normal vector, see (2.25)). In case of a joint between two null curves, the argument of the logarithm is given by $g_{\mu\nu}\bar{k}^\mu k^\nu = e^{2\rho}\eta_{\mu\nu}\bar{k}^\mu k^\nu$, but now the above procedure is performed twice (once for each null surface), so that the resulting argument is $\eta_{\mu\nu}\bar{k}^\mu k^\nu$. Hence, in all cases, the resulting joint terms will be independent of the conformal factor ρ.

Combining the resulting terms, and using the product rule, we obtain

$$\begin{aligned} -2\sigma_{\mathcal{N}}\int_{\mathcal{N}} d\lambda\tilde{\chi}\frac{k_-^\mu k_-^\nu\partial_\mu\partial_\nu l}{k_-^\sigma\partial_\sigma l} - 2\sigma_{\mathcal{N}}\tilde{\chi}_2 \log\left|A\,\frac{e^{-\rho}m_\mu k^\mu}{\partial_\lambda l}\right|_2 + 2\sigma_{\mathcal{N}}\tilde{\chi}_1 \log\left|\bar{A}\,\frac{e^{-\rho}\bar{m}_\mu k^\mu}{\partial_\lambda l}\right|_1 \\ +2\sigma_{\mathcal{N}}\int_{\mathcal{N}} d\lambda\tilde{\chi}\partial_\lambda\rho - 2\sigma_{\mathcal{N}}\int_{\mathcal{N}} d\lambda\rho\partial_\lambda\tilde{\chi}\,, \end{aligned} \tag{2.37}$$

where each term is now manifestly invariant under reparametrization $\lambda \mapsto \lambda'$ and the first line is independent of the conformal factor ρ. Importantly, the first term in the second line combines with the corresponding terms arising from space- or timelike boundary components, in conjunction with Stokes' theorem,[5]

$$2\sum_{\mathcal{N}\in\mathscr{N}}\sigma_{\mathcal{N}}\int_{\mathcal{N}} d\lambda\tilde{\chi}\partial_\lambda\rho + 2\sum_{\mathcal{S}\in\mathscr{S}}\sigma_{\mathcal{S}}\int_{\mathcal{S}}\sqrt{|h|}\tilde{\chi}n^\mu\partial_\mu\rho\,d\Sigma = 2\int_{\mathcal{M}} d^2y\sqrt{-g}\nabla(\tilde{\chi}\nabla\rho)\,, \tag{2.38}$$

in order to cancel with the total derivative contribution coming from the bulk (2.17).

Since the RST variation of the remaining bulk contribution vanishes, the RST variation of the null contributions has to vanish as well. Because the first line of (2.37) is independent of the conformal factor, these terms are not necessary in order to ensure a well-posed variational problem and we leave them out to implement the RST symmetry. As these terms are manifestly reparametrization invariant, removing them will not spoil overall reparametrization invariance. The RST variation of the last term in the last line of (2.37) does not vanish, but since it is reparametrization invariant and the variation of ρ vanishes at the boundary (there are no derivatives involved), it can be cancelled by adding a boundary term.

[5] See appendix A for a justification of the formula when including null boundaries.

JHEP07(2020)173

2.2.4 Complete action

Let us now collect the results of the above considerations. We obtain a simple expression for the total action S, in conformal gauge,

$$S = \int_{\mathcal{M}} d^2y\sqrt{-g}\left[\frac{1}{\kappa}(\nabla\chi)^2 - \frac{1}{\kappa}(\nabla\Omega)^2 + \frac{2}{\sqrt{-g}}e^{\frac{2}{\kappa}(\chi-\Omega)} - \frac{1}{2}\sum_{i=1}^{N}(\nabla f_i)^2\right] - \frac{\kappa}{4}\int d^2y\sqrt{-g}\left[\nabla\left(\eta\nabla\eta\right)\right] . \tag{2.39}$$

This action has the properties that the variational principle is well-posed on any spacetime region bounded by spacelike, timelike, or null boundaries. Additionally, it is invariant under RST transformations in the sense that $\delta_{\mathrm{RST}}S = 0$. Note that the action does not involve any boundary or joint terms anymore, since they were either consistently removed, or canceled against total derivative contributions from the original bulk action.

As a result, the expression for the holographic complexity does not involve an arbitrary function l anymore. Further, as a side note, the value of holographic complexity on the WdW patch can also be obtained by a limiting procedure, regulating the WdW patch with space- or timelike surfaces only. The resulting limit is finite, and it agrees with the result obtained by the above RST symmetric prescription for null boundaries.

The on-shell action. The equation of motion

$$\nabla^2\chi = \nabla^2\Omega \tag{2.40}$$

allows us to choose Kruskal coordinates (x^+, x^-) where $\rho = \phi$, implying $\Omega = \chi$. In this coordinate system, the on-shell action S is subject to a remarkable simplification,

$$S = \int_{\mathcal{M}} dx^+dx^-\left[2 - \sum_{i=1}^{N}\partial_+f_i\partial_-f_i + \frac{\kappa}{2}\partial_+\eta\partial_-\eta\right] . \tag{2.41}$$

In particular, in this form the action has no explicit dependence on the dilaton field ϕ. This is, of course, somewhat misleading, for the shape of the WdW patch in Kruskal coordinates will indeed depend on the spacetime metric and therefore the dilaton as well.

3 Classical black hole complexity

3.1 Gravitational collapse

Before discussing the semi-classical case, let us analyse the classical gravitational collapse of an infinitely thin shell of incoming matter f with mass M in the CGHS model. The energy momentum tensor associated to the matter field f is given by

$$T^f_{++} = \frac{M}{x_0^+}\delta(x^+ - x_0^+) , \tag{3.1}$$

JHEP07(2020)173

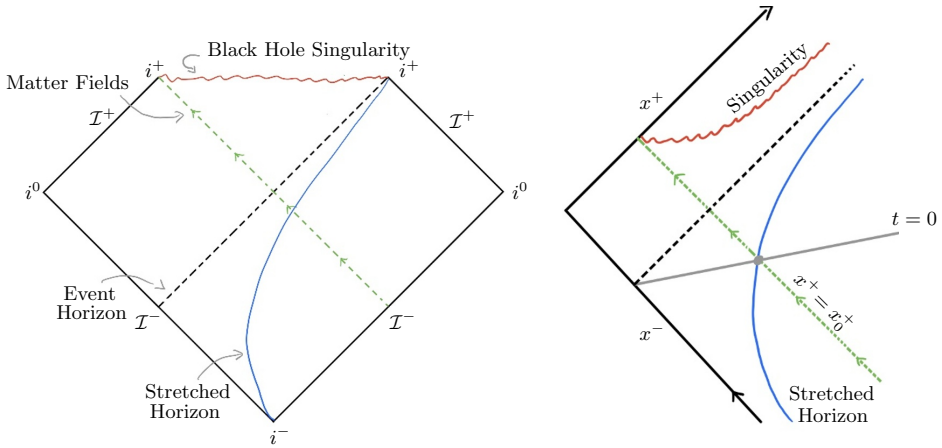


Figure 1. Left panel: Penrose diagram of one-sided CGHS black hole formed by gravitational collapse. Right panel: the corresponding Kruskal diagram with the same color coding. The gray line denotes a curve of equal tortoise time t.

and for the dilaton ϕ and conformal factor ρ this implies

$$e^{-2\phi} = e^{-2\rho} = \begin{cases} -x^+x^- & \text{if } x^+ < x_0^+ \\ -x^+\left(x^- + \frac{M}{x_0^+}\right) + M & \text{if } x^+ \geq x_0^+ \,, \end{cases} \tag{3.2}$$

in Kruskal coordinates. The infalling shell creates a black hole singularity, as shown in figure 1, which depicts a Penrose and Kruskal diagram on the left and right, respectively.

In line with our previous paper [10], we take the WdW patch to be anchored at the stretched horizon, defined as a membrane outside the black hole, with an area that is one unit larger than the area of the event horizon,

$$e^{-2\phi_{\rm SH}} = e^{-2\phi_{\rm EH}} + 1 = M + 1 \,. \tag{3.3}$$

In the classical collapse solution considered here, the stretched horizon is a curve of constant dilaton ϕ outside the black hole. If we instead anchor the WdW patch on a curve far outside the black hole, the main difference is to shift the onset of complexity growth forward in time, to the time in tortoise coordinates when the infalling shock wave passes through the anchor curve on its way to forming the black hole. As was discussed in [10], it seems more physical to place the anchor curve at the stretched horizon and have the onset of complexity growth coincide, at least approximately, with the time of black hole formation (here defined as the tortoise time at which the shock wave meets the stretched horizon).

The WdW patch at a given tortoise time t is defined as the union of all spacelike surfaces originating from the point on the anchor curve that intersects the appropriate constant t curve and extending towards the black hole, see figure 2. The holographic complexity $\mathcal{C}$ at time t is then given by the action evaluated on the WdW patch. The classical action can be obtained by formally setting κ to zero in (2.41). Furthermore, since

JHEP07(2020)173

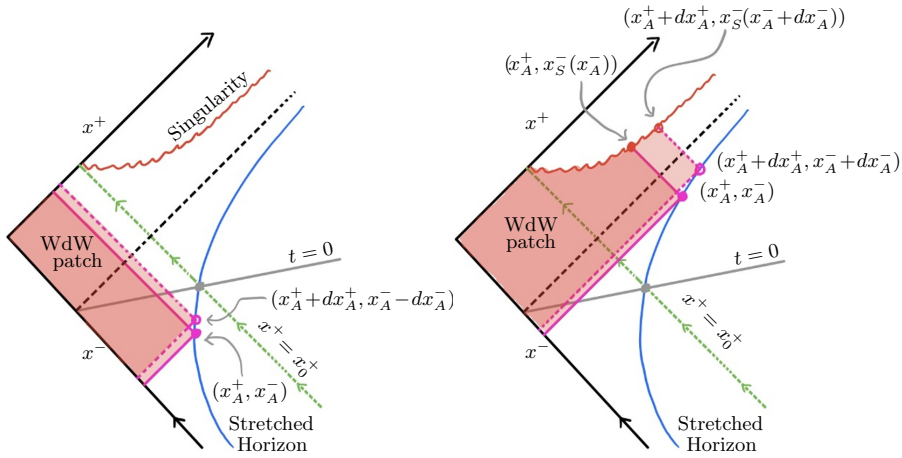


Figure 2. Evolution of the WdW patch for classical gravitational collapse. Color coding coincides with figure 1.

the collapse solution only involves infalling matter, for which $\partial_- f = 0$, the matter term in (2.41) does not contribute to the action. This is true independent of the precise form of the matter profile. In particular, we could have started with a smoothly varying incoming matter flux, instead of a sharp shockwave, and this conclusion would still hold. It follows, that the on-shell action reduces to

$$\mathcal{C} = S_{\rm cl} = 2\int_{\rm WdW} dx^+ dx^- =: 2\mathcal{A}\,, \tag{3.4}$$

where $\mathcal{A}$ can be interpreted as the reference metric "area" of the WdW patch drawn in Kruskal coordinates.

An asymptotic observer would use the tortoise coordinates (t, x), which are related to Kruskal coordinates (x^+, x^-) by the equations

$$\begin{aligned} x^+ &= e^{t+x}\,, \\ x^- + \frac{M}{x_0^+} &= -e^{-t+x}. \end{aligned} \tag{3.5}$$

We are interested in the growth rate of holographic complexity $\frac{d\mathcal{C}}{dt}$, where t is the tortoise time associated to the anchor point of the WdW patch, see figure 1. To this end, we denote the Kruskal coordinates describing the anchor point as (x_A^+, x_A^-), while the singularity curve is described by (x_S^+, x_S^-). It is practical to consider two separate cases: the WdW patch anchored before the shock wave arrives, $x^+ < x_0^+$, and after, $x^+ > x_0^+$.

Before incoming shockwave. It is apparent from figure 2 that the change of the area in Kruskal coordinates, $\mathcal{A}$, before the shockwave arrives, is given by

$$d\mathcal{A} = -x_A^- dx_A^+ - x_A^+ dx_A^-\,. \tag{3.6}$$

JHEP07(2020)173

From the definition of the stretched horizon (3.3), which we identify with the anchor curve, we obtain the equivalent relation

$$-x_A^+\left(x_A^- + \frac{M}{x_0^+}\right) = 1\,, \tag{3.7}$$

which immediately implies that

$$-x_A^- dx_A^+ - x_A^+ dx_A^- = \frac{M}{x_0^+} dx_A^+\,, \tag{3.8}$$

so that

$$d\mathcal{A} = \frac{M}{x_0^+} dx_A^+\,. \tag{3.9}$$

Furthermore, since $dx_A^+ = x_A^+ dt$, if we shift the time variable t, so that $t = 0$ corresponds to where the shock wave meets the stretched horizon, then $d\mathcal{A} = Me^t dt$, or

$$\dot{\mathcal{C}} = 2Me^t \quad \text{for } x^+ < x_0^+\,. \tag{3.10}$$

We observe an exponential onset towards $2M$ at $t = 0$, the black hole creation time. The time scale of the exponential growth is given by the characteristic scale λ of the model, which we have set to 1.

After incoming shockwave. The analogous calculation can be done for times after the black hole creation, $t > 0$. One easily finds

$$d\mathcal{A} = \left(x_S^-(x_A^+) - x_A^-\right) dx_A^+ - x_A^+ dx_A^-\,. \tag{3.11}$$

Using (3.8) in conjunction with the defining relation for the black hole singularity,

$$M = x_S^+\left(x_S^- + \frac{M}{x_0^+}\right)\,, \tag{3.12}$$

one obtains

$$d\mathcal{A} = \left(x_S^-(x_A^+) + \frac{M}{x_0^+}\right) dx_A^+ = \frac{M}{x_A^+} dx_A^+ = M dt\,, \tag{3.13}$$

or

$$\dot{\mathcal{C}} = 2M \quad \text{for } x^+ > x_0^+\,, \tag{3.14}$$

so that the holographic complexity growth $\dot{\mathcal{C}}$ is continuous at $x^+ = x_0^+$ and remains constant for $x^+ > x_0^+$.

Our findings, summarized in figure 3, are consistent with the expectation that the complexity of the quantum state corresponding to a black hole should grow with a rate proportional to the black hole entropy times its temperature.

JHEP07(2020)173

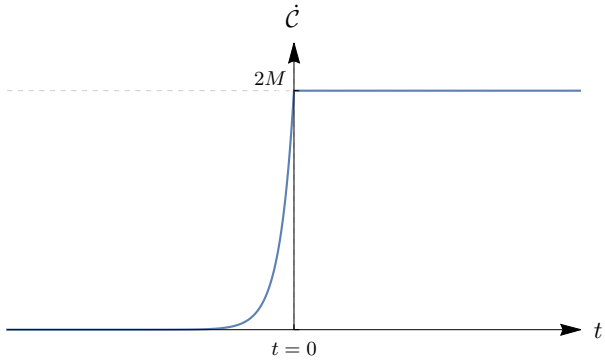


Figure 3. Growth rate $\dot{\mathcal{C}}$ of holographic complexity as a function of tortoise time t, using the CA prescription, for classical gravitational collapse. Following an exponential onset, holographic complexity grows linearly with time.

3.2 Eternal black hole

For completeness we should mention that our prescription also works for the classical eternal black hole for late time. Its solution in terms of the dilaton, in Kruskal coordinates, is given by

$$e^{-2\phi} = e^{-2\rho} = M - x^+x^- \,, \tag{3.15}$$

see e.g. [21]. The black and white hole singularities are located on the curves defined by $M = x_S^+ x_S^-$.

A Kruskal diagram including the WdW patch for late times is given in figure 4. As is usual for the CA prescription in the context of two-sided black holes, we have to provide a second anchor point on a 'left' anchor curve, see e.g. [22, 23]. The result for complexity growth will then be a function of $t_R - t_L$ where t_R (t_L) are the tortoise times w.r.t. the right (left) side associated to the respective anchor point position, see figure 4. For simplicity, we take the two anchor points to move symmetrically as time progresses, i.e. $t_L = -t_R$, so that the result only depends on $t = t_R$. In this case, we expect the complexity growth to have twice the contribution of a one-sided black hole.

The variation of the Kruskal area $\mathcal{A}$ is easily performed and indeed provides the expected holographic complexity growth

$$\dot{\mathcal{C}} = 4M \,, \tag{3.16}$$

in the late time limit.

4 Semi-classical black hole complexity

4.1 Evaporating black hole

Again, we study an incoming leftmoving shockwave pulse of energy M at $x^+ = x_0^+$ of the form

$$T^f_{++} = \frac{M}{x_0^+}\delta(x^+ - x_0^+) \,. \tag{4.1}$$

JHEP07(2020)173

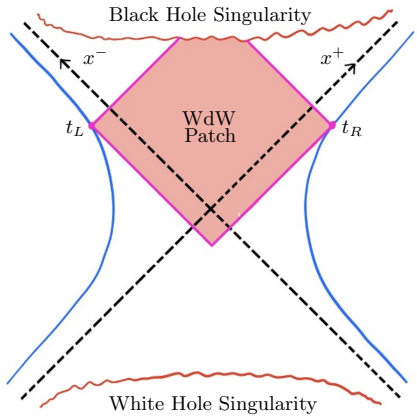


Figure 4. Kruskal diagram of eternal black hole. A symmetric WdW patch is presented. Color coding agrees with previous figures.

The semi-classical collapse solution, in terms of the field Ω defined in (2.16), using Kruskal coordinates, is given by [12]

$$\Omega(x^+, x^-) = -x^+x^- + \left(x_0^+ - x^+\right)\frac{M}{x_0^+}\Theta(x^+ - x_0^+) - \frac{\kappa}{4}\ln\left(-x^+x^-\right) , \tag{4.2}$$

which in turn determines the dilaton ϕ and the metric via its conformal factor $\rho = \phi$. The field η (see (2.17)), needed for the holographic complexity computation, is related to the conformal anomaly of the energy momentum tensor $T^f_{\mu\nu}$, $\langle T^{f\,\mu}_{\mu}\rangle = \frac{\kappa}{2}R$, which fixes the form of the energy momentum tensor in light cone coordinates $y^\pm$, cf. (2.13), as [11, 24]

$$\langle T^f_{\pm\pm}\rangle = -\kappa\left[\partial_\pm\rho\partial_\pm\rho - \partial_\pm^2\rho + t_\pm(y^\pm)\right] , \tag{4.3}$$

where

$$t_\pm = \frac{1}{4}\left(\partial_\pm\eta\partial_\pm\eta + 2\partial_\pm^2\eta\right) \tag{4.4}$$

is determined by boundary conditions imposed at past null infinity $\mathcal{I}^-$, stating that there should be no outgoing energy flux at $\mathcal{I}^-$. In this form, the energy momentum 'tensor', including its quantum corrections, is actually not a tensor anymore. This is related to the fact, that the quantum corrections depend on a choice of vacuum, which makes reference to a specific coordinate system. For the case at hand, this is the coordinate system where the metric is manifestly Minkowskian near $\mathcal{I}^-$.

The form of $t_\pm$ in Kruskal coordinates $x^\pm$ can be determined [11] to be $t_\pm(x^\pm) = -1/(2x^\pm)^2$. The field η is fixed by (4.4) and reads [19]

$$\eta(x^+, x^-) = \log(-x^+x^-) , \tag{4.5}$$

which is needed for the evaluation of holographic complexity, see equation (2.41). Since the field η determines the functions $t_\pm$ (at least in Kruskal coordinates), it can be viewed

JHEP07(2020)173

as encoding the boundary conditions of the energy momentum tensor $\langle T^f \rangle$. This is in line with the fact that the field η only appears within a total derivative term in the action (2.17).

The form of $\langle T^f_{++} \rangle$ implies, that an asymptotic observer near future null infinity $\mathcal{I}^+$ will see a non-vanishing outgoing matter energy flux, i.e. Hawking radiation, which turns on with an exponential onset as the black hole is formed and turns off when the mass of the black hole has been depleted. The endpoint of Hawking emission is abrupt in the RST model and even requires a small adjustment in the form of a negative energy shock wave emanating from the black hole endpoint.[6] This reflects a breakdown of the semi-classical description when the remaining black hole mass approaches the Planck scale and serves as a reminder that our semi-classical calculation of holographic complexity will be subject to similar limitations as the black hole mass is depleted.

Altogether, the solution (4.2) together with (4.5), describes flat spacetime for $x^+ \leq x_0^+$ and an evaporating black hole for $x^+ > x_0^+$ with outgoing Hawking radiation towards future infinity $\mathcal{I}^+$, see figure 5. The location of the black hole singularity is determined by the curve (x_S^+, x_S^-) which satisfies $\Omega(x_S^+, x_S^-) = \Omega_{\text{crit}} = \frac{\kappa}{4}(1 - \ln \frac{\kappa}{4})$ for $x^+ > x_0^+$. A useful parametrization of this curve, which we employ at a later stage, is given by

$$\begin{pmatrix} x_S^+(u) \\ x_S^-(u) \end{pmatrix} = \begin{pmatrix} x_0^+ \left(\frac{\kappa}{4M} \left(e^{\frac{4M}{\kappa} u} - 1 \right) - u + 1 \right) \\ -\frac{\kappa}{4x_0^+} \frac{e^{\frac{4M}{\kappa} u}}{\frac{\kappa}{4M}\left(e^{\frac{4M}{\kappa} u} - 1\right) - u + 1} \end{pmatrix}, \tag{4.6}$$

where the range of the parameter u is the interval $(0, 1)$. The point $(x_S^+(0), x_S^-(0))$ represents the formation of the black hole singularity, while $(x_S^+(1), x_S^-(1))$ describes the point where the black hole has entirely evaporated. This parametrization has the convenient property

$$-x_S^+(u)\, x_S^-(u) = \frac{\kappa}{4} e^{\frac{4M}{\kappa} u}\,. \tag{4.7}$$

The curve $\Omega(x_B^+, x_B^-) = \Omega_{\text{crit}} = \frac{\kappa}{4}\left(1 - \ln \frac{\kappa}{4}\right)$ for $x^+ < x_0^+$ defines the boundary of physical spacetime before the matter shockwave arrives, and is given by

$$-x_B^+ x_B^- = \frac{\kappa}{4}\,, \tag{4.8}$$

in Kruskal coordinates.

As in the classical theory, we take the anchor curve for our WdW patch to be the stretched horizon of the black hole, defined as a membrane outside the black hole event horizon, with an area of order 1, in Planck units, larger than the area of the black hole event horizon. For technical simplicity, we follow [2] and take the stretched horizon of an RST black hole formed by shockwave collapse to coincide with the apparent horizon during the period of evaporation. This determines the anchor curve as

$$-x_A^+ \left(x_A^- + \frac{M}{x_0^+} \right) = \frac{\kappa}{4}\,, \tag{4.9}$$

[6] See e.g. [21] for more details.

JHEP07(2020)173

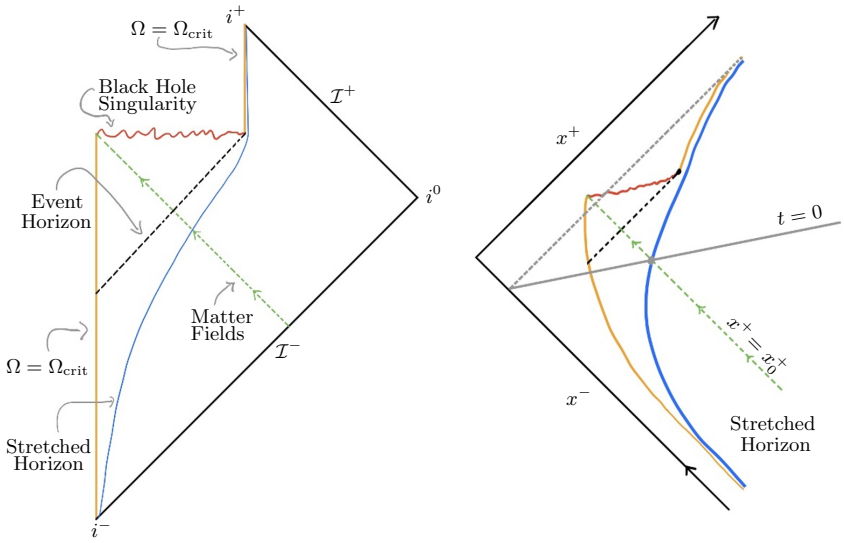


Figure 5. Left panel: Penrose diagram depicting the life cycle of evaporating black hole formed by collapse. Right panel: the corresponding Kruskal diagram with the same color coding. The gray line denotes a curve of equal tortoise time t.

as usual, in Kruskal coordinates. With this convention, the area of the stretched horizon vanishes at the evaporation end point, whereas it should strictly speaking be 1 in Planck units there. However, the error is negligible as long as the black hole remains large compared to the Planck scale, i.e. whenever the semi-classical approximation can be relied on in the first place. With these ingredients, it is now possible to define the WdW patch in a similar fashion as in the classical case, see figure 6.

The on-shell action (2.41), together with the field η given by (4.5), can be formulated as

$$\mathcal{C} = 2\mathcal{A} + \frac{\kappa}{2}\mathcal{B} := \int_{\text{WdW}} dx^+ dx^- + \frac{\kappa}{2}\int_{\text{WdW}} \frac{dx^+}{x^+}\frac{dx^-}{x^-}\,. \tag{4.10}$$

In addition to the 'flat' reference metric area $\mathcal{A}$ in Kruskal coordinates, the semi-classical holographic complexity acquires a correction $\frac{\kappa}{2}\mathcal{B}$. For future evaluation purposes, we note that the correction term can also be given an 'area' interpretation, by changing from Kruskal coordinates to their logarithm,

$$\begin{aligned} \sigma^+ &= \log(x^+) \\ \sigma^- &= \log(-x^-)\,. \end{aligned} \tag{4.11}$$

To evaluate holographic complexity we again consider two cases: the WdW patch anchored in the region before the shock wave arrives, $x^+ < x_0^+$, and after, $x^+ > x_0^+$.

Before incoming shockwave. The evaluation of the change of area $\mathcal{A}$ is completely analogous to the classical case, see figure 6. We have

$$d\mathcal{A} = \left(x_B^-(x_A^+) - x_A^-\right) dx_A^+ - \left(x_A^+ - x_B^+(x_A^-)\right) dx_A^-\,, \tag{4.12}$$

JHEP07(2020)173

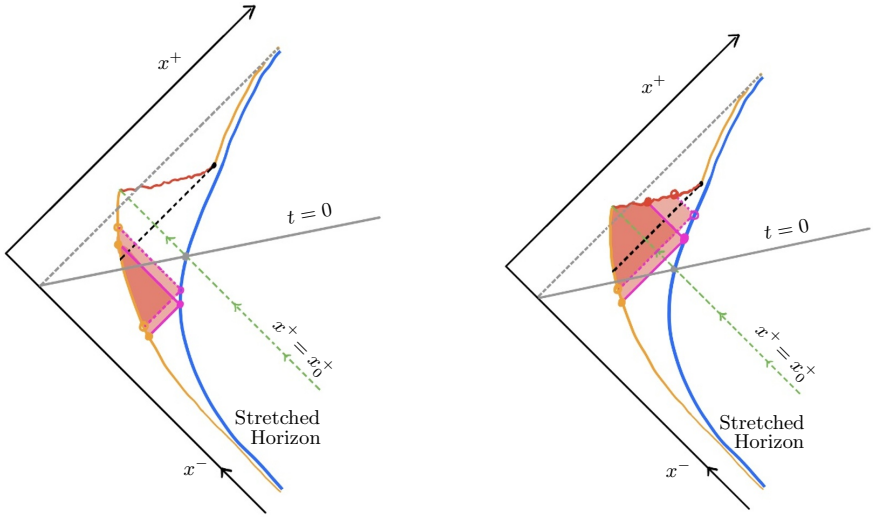


Figure 6. Kruskal diagrams of evolution of a WdW patch of an evaporating black hole. Color coding coincides with figure 5.

which can be evaluated, making use of (4.8) and (4.9), to give

$$d\mathcal{A} = \left(Me^t - \frac{\kappa}{4} + \mathcal{O}\left(M^{-1}\right)\right) dt\,. \tag{4.13}$$

The result deviates from the classical calculation by a constant contribution proportional to κ.

Similarly,

$$d\mathcal{B} = \log\left(\frac{x_B^-(x_A^+)}{x_A^-}\right) d\log(x_A^+) - \log\left(\frac{x_A^+}{x_B^+(x_A^-)}\right) d\log(x_A^-)\,. \tag{4.14}$$

The result is exponentially suppressed as $t \to -\infty$, but gives a non-negligible contribution for times near the black hole creation ($t = 0$) of the form

$$\frac{\kappa}{2} d\mathcal{B} \approx -\frac{\kappa}{2}\left(\log\left(\frac{4M}{\kappa}\right) + t\right) dt\,. \tag{4.15}$$

Combining the results leads to a total complexity growth $\dot{\mathcal{C}}$ before the incoming shock-wave $t < 0$ of

$$\dot{\mathcal{C}} = \begin{cases} 2Me^t - \frac{\kappa}{2}\left(1 + \log\left(\frac{4M}{\kappa}\right) + t + \mathcal{O}\left(\frac{\kappa}{M} e^{-t}\right)\right) & \text{for } -t_S \lesssim t < 0\,, \\ \kappa\,\mathcal{O}\left(\left(\frac{M}{\kappa} e^t\right)^3\right) & \text{for } t \lesssim -t_S\,, \end{cases} \tag{4.16}$$

where $t_S = \log\left(\frac{4M}{\kappa}\right)$ is the scrambling time. A graphical representation of this tiny on-set behaviour can be seen in figure 7 for $t < 0$.

JHEP07(2020)173

After incoming shockwave. Replacing the spacetime boundary by the singularity curve in the region $x^+ > x_0^+$ gives the correct change of Kruskal area,

$$d\mathcal{A} = \left(x_S^-(x_A^+) - x_A^-\right) dx_A^+ - \left(x_A^+ - x_B^+(x_A^-)\right) dx_A^- \,. \tag{4.17}$$

In addition to (4.8) and (4.9), using the parametrization (4.6), we can express the result as

$$d\mathcal{A} = M(1-u)\, dt - \frac{\kappa}{4} dt + \mathcal{O}\left(e^{-\frac{4M}{\kappa}u}\right) , \tag{4.18}$$

where corrections are suppressed after a scrambling time t_S. The parameter u is related to time t via

$$\frac{4M}{\kappa}\left(e^t - 1\right) = e^{\frac{4M}{\kappa}u} - \frac{4M}{\kappa}u - 1 \,, \tag{4.19}$$

which, for times after the scrambling time t_S, can be expressed as

$$u(t \gtrsim t_S) \approx \frac{\kappa}{4M}\left(t + \log\left(\frac{4M}{\kappa}\right)\right) . \tag{4.20}$$

Notably, the growth of the Kruskal area $\mathcal{A}$ with time t is linear after the scrambling time. Moreover, the result for $\frac{d\mathcal{A}}{dt}$ is continuous at the black hole creation time $t = 0$.

That leaves us with the evaluation of the logarithmic area $\mathcal{B}$,

$$d\mathcal{B} = \log\left(\frac{x_S^-(x_A^+)}{x_A^-}\right) d\log(x_A^+) - \log\left(\frac{x_A^+}{x_B^+(x_A^-)}\right) d\log(x_A^-) \,. \tag{4.21}$$

The result is suppressed after the scrambling time t_S, but contributes near the black hole creation time $t = 0$,

$$\frac{\kappa}{2} d\mathcal{B} \approx -\frac{\kappa}{2}\log\left(\frac{4M}{\kappa}\right) dt \tag{4.22}$$

which shows that also $\frac{d\mathcal{B}}{dt}$ is continuous at $t = 0$.

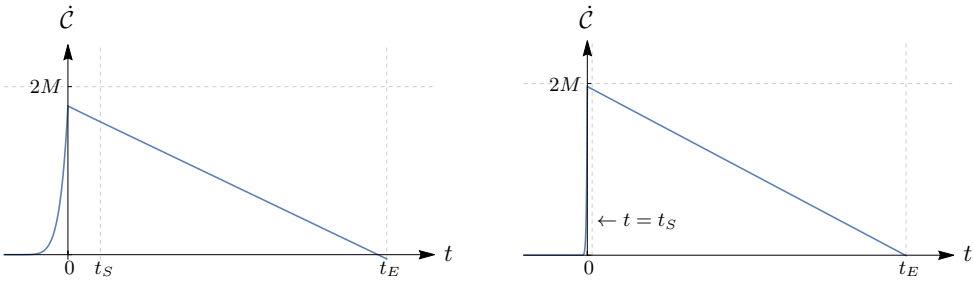
It is interesting to observe, that the contribution of $\mathcal{B}$ before the scrambling time t_S conspires with the non-linear contribution of $\mathcal{A}$ before the scrambling time to provide linear growth up to corrections of order M^{-1} even before the scrambling time. The final exact result for the complexity growth rate after the black hole creation is given by

$$\begin{aligned} \dot{\mathcal{C}}(t) &= 2M\left(1 - \frac{t + 1 + \log\left(\frac{4M}{\kappa} + e^{-t}\right)}{\frac{4M}{\kappa} + e^{-t}}\right) \\ &= 2M - \frac{\kappa}{2}\left(\log\left(\frac{4M}{\kappa}\right) + 1 + t + \mathcal{O}\left(\frac{\kappa}{M}e^{-t}\right)\right) . \end{aligned} \tag{4.23}$$

This result holds until the black hole has fully evaporated, but, as stated above, it becomes unreliable when the remaining mass is of order the Planck scale.

The total result is plotted in figure 7. The plot confirms the continuity and linear falloff of the growth rate of holographic complexity $\dot{\mathcal{C}}$. A short time before the lifetime t_E of the black hole has been reached, the value of the complexity growth rate hits zero and subsequently becomes slightly negative. This is potentially problematic, since there is no reason to believe, that the complexity growth rate of an evaporating black hole should ever be negative. However, at that time, the mass of the black hole has already attained the Planck scale, and, as stated above, our result is not trustworthy anymore.

JHEP07(2020)173



(a) $M/\kappa = 10$. **(b)** $M/\kappa = 100$.

Figure 7. The growth rate $\dot{\mathcal{C}}$ of holographic complexity as a function of tortoise time t, using the CA prescription, for evaporating black holes of different initial mass. The exponential onset at the creation time of the black hole is followed by a linear falloff period until the black hole has evaporated.

4.2 Eternal black hole

We can also study a semi-classical eternal black hole by including a heat bath at spatial infinity, with a temperature equal to the Hawking temperature of the black hole. The heat bath provides a steady incoming energy flux which matches the outgoing flux from the radiating black hole. The solution, in Kruskal coordinates, is given by [21]

$$\Omega(x^+, x^-) = -x^+x^- + M + \frac{\kappa}{4} - \frac{\kappa}{4}\log\left(\frac{\kappa}{4}\right) . \tag{4.24}$$

The spacetime curvature is singular where $\Omega = \Omega_{\text{crit}}$, i.e. on curves satisfying

$$x_S^+ x_S^- = M , \tag{4.25}$$

describing a black hole and white hole singularity. These are the same curves as for the singularities of the classical eternal black hole described by (3.15). The Kruskal diagram for a semi-classical eternal black hole solution is thus identical to that of a classical eternal black hole, shown in figure 4, but the physics described by the semi-classical solution is somewhat different. In contrast to all other solutions considered in this work, the parameter M in (4.24) is *not* proportional to the ADM mass of the black hole. Since the semi-classical solution describes a black hole in equilibrium with a heat bath at infinity there is non-vanishing energy density in the asymptotic region and the ADM mass diverges.[7] The parameter M is characteristic of the black hole size and therefore we will continue to refer to it as the 'mass' of the black hole.

Since the semi-classical Kruskal diagram is unchanged compared to the diagram of a classical black hole, and the Kruskal area $\mathcal{A}$ is only sensitive to the location of the singularity and not the detailed form of the dilaton field, it agrees with the classical calculation,

$$d\mathcal{A} = 2M dt , \tag{4.26}$$

[7]The infinite train of radiation does not lead to a catastrophic back-reaction on the geometry because the gravitational coupling, governed by e^{ϕ}, goes rapidly to zero asymptotically.

JHEP07(2020)173

for late times. Further, it follows from (4.24) that $t_\pm = 0$ in Kruskal coordinates and then equation (4.4) immediately implies that $\eta = 0$. It follows that the semi-classical correction $\mathcal{B}$ vanishes.

We conclude that the complexity growth of the semi-classical eternal black hole for late times agrees with the classical result (3.16),

$$\dot{\mathcal{C}} = 4M\,, \tag{4.27}$$

and does not receive semi-classical corrections.

5 Discussion and outlook

In this paper we have investigated the holographic complexity of evaporating black holes in a toy model where the semi-classical geometry is known explicitly. The CA proposal for black hole complexity can be adapted to this model and we have obtained analytic expressions for the increase in complexity over the lifetime of a semi-classical black hole. This extends our previous work in [10] where we numerically evaluated the semi-classical complexity using a CV prescription for the same model. The analytic CA results presented here provide a much more detailed picture of how complexity evolves as the black hole evaporates compared to the previous numerical CV evaluation. For parameter values where the semi-classical approximation can be trusted, the two approaches are in good agreement, starting from a scrambling time after the black hole is formed and for most of the remainder of the black hole lifetime.

In order to ensure a well-posed variational principle for the action on a Wheeler-DeWitt patch, it is necessary to include appropriate boundary terms in the action. These boundary terms are not unique, something that is true for CA in general, but for a range of black holes in classical Einstein gravity the ambiguity does not affect the late time rate of increase of complexity [18]. In the context of semi-classical black holes, the finite black hole lifetime limits the ability to take a late time limit and the ambiguity involving boundary terms needs to be addressed in order to have a well-defined CA prescription. This can be achieved in a natural way in the RST model by extending a symmetry of the original semi-classical bulk theory to the boundary terms as well. The final analytic result for complexity growth rate during the evaporation process, presented in formula (4.23), has several interesting features.

First, it confirms linear falloff of $\dot{\mathcal{C}}$ with time after the scrambling time t_S, already observed (numerically) in [10] using a CV prescription. In fact, up to small corrections, equation (4.23) exhibits linear behaviour already from $t = 0$, the time of black hole formation, in contrast to CV where the numerics indicates an initial adjustment period of order the scrambling time. The linear falloff is important, as it captures the time evolution of the entropy of the evaporating black hole. Classical black hole entropy is given by $S_0 = 2M$ in this model and at the semi-classical level the black hole radiates mass at a constant rate $\kappa/4$, so that

$$S_0(t) = 2M(t) = 2M - \frac{\kappa}{2}t\,. \tag{5.1}$$

JHEP07(2020)173

The Hawking temperature is independent of black hole mass in this model so the relation

$$\dot{\mathcal{C}}(t) \propto S(t)\ T\,, \tag{5.2}$$

is seen to hold at leading order in a κ/M expansion for large initial black hole mass.

Second, the subleading logarithmic term in the rate of complexity increase in (4.23) can also be given an interpretation in terms of entropy. In [15, 16], it was shown that the leading order quantum-corrected entropy for a semi-classical black hole in equilibrium with a thermal heat bath, is given by

$$S = 2e^{-2\phi_h} + \frac{\kappa}{2}\phi_h - \frac{\kappa}{2} + \frac{\kappa}{4}\log\frac{\kappa}{4}\,, \tag{5.3}$$

where ϕ_h is the value of the dilaton field at the horizon. When evaluated for a dynamical solution of the RST model describing a black hole formed by an incoming shock wave, this gives

$$S = 2M - \frac{\kappa}{2}\log\left(\frac{4M}{\kappa}\right) - \frac{\kappa}{2}\,, \tag{5.4}$$

immediately after the black hole is formed and zero at the evaporation endpoint. Comparing to (4.23) shows that the rate of complexity growth at the onset of black hole evaporation is consistent with the relation (5.2), even including the subleading logarithmic term.[8] If we instead evaluate the entropy formula (5.3) for the static solution (4.24), describing an eternal black hole in equilibrium with a heat bath, we find that the entropy takes its classical value,

$$S = 2M\,. \tag{5.5}$$

The cancellation of the semi-classical corrections can ultimately be traced to the back-reaction on the spacetime geometry due to the matched ingoing and outgoing radiation flux [15]. Interestingly, the corresponding cancellation also takes place in the rate of complexity increase (4.27) for an eternal RST black hole and we once again find that the relation (5.2) holds with $S(t)$ given by the semi-classical entropy.

One may wonder how to interpret our formulas after the black hole has evaporated. One can still define a stretched horizon as the timelike curve where the transverse area is one unit larger than zero. This curve is very close to the boundary at $\phi = \phi_{\rm crit}$ and a WdW patch anchored on it only covers a microscopic spacetime region.[9] Both the Kruskal area $\mathcal{A}$ and the semi-classical correction $\mathcal{B}$ will have minuscule values, which do not change with time. This is consistent with zero growth of the holographic complexity at late times, but having a vanishingly small WdW patch action at late times does not reflect the very large absolute complexity that was built up during the evaporation of the black hole and is carried in the outgoing train of Hawking radiation. For a classical black hole, the growth of the WdW patch action continues indefinitely and this issue does not arise. An obvious way around this is to only use the action prescription to calculate the change in complexity and define the absolute holographic complexity as the integral of $\dot{\mathcal{C}}$ over time. With this

[8] Due to the slow evolution of the logarithm, this remains true for the bulk of the black hole lifetime.

[9] The same is of course true for a WdW patch at very early times, long before the black hole is formed.

JHEP07(2020)173

prescription the complexity is a monotonically growing function of time and does not get discontinuously adjusted to a near-zero value at the endpoint of the evaporation process.

The constant Hawking temperature of CGHS and RST black holes simplifies calculations but is rather unphysical. It would be interesting to study the charged version of the CGHS black holes, see e.g. [25, 26], or the semi-classically corrected Jackiw-Teitelboim (JT) model, see e.g. [27], since black holes in those two-dimensional models have varying temperature.

Acknowledgments

It is a pleasure to thank Shira Chapman and Nick Poovuttikul for discussions. This research is supported by the Icelandic Research Fund under grants 185371-051 and 195970-051, and by the University of Iceland Research Fund.

A Stokes' theorem in two dimensions with null boundaries

Stokes' theorem in the context of Lorentzian manifolds is usually presented for manifolds with spacelike or timelike boundaries. For simplicity, we focus on a single smooth boundary component ∂M. The result for piecewise smooth boundaries is obtained by summing over individual boundary components. The theorem states that

$$\int_M \sqrt{|g|}\nabla_\mu j^\mu = \int_{\partial M} \sqrt{|h|} n^\mu j_\mu + \dots \tag{A.1}$$

where n^μ is a inwards (outwards) pointing unit normal vector if ∂M is spacelike (timelike) and the... indicates contributions from other boundary components. The integrals naturally involve the metric determinant $|g|$ and the determinant of the induced metric $|h|$ at the boundary ∂M.

The expression on the right hand side of (A.1) does not make sense for null boundaries, as the induced metric h and the unit normal vector n^μ are degenerate in this case. Since Stokes' theorem in its general form is a statement involving differential forms, it is oblivious to a metric on a manifold.[10] Thus Stokes' theorem also has to be valid for manifolds with null boundaries. Our goal in this appendix is to find a simple expression to replace (A.1) for null boundary components in a two-dimensional context. This is easily achieved using a limiting procedure where the null boundary curve is approximated by a family of either timelike or spacelike curves.

We work in conformal gauge,

$$ds^2 = -e^{2\rho} dy^+ dy^- \,, \tag{A.2}$$

with light-like coordinates (y^+, y^-) and consider a null boundary component, ∂M, that we take to lie in the future of M and described by a null curve of the form $y^- = y_0^- = \text{const}$.

[10] The need for a metric only arises if one wants to integrate scalar functions over a manifold in a coordinate-invariant way.

JHEP07(2020)173

In a region near the point (y_0^+, y_0^-), the null curve is approached by a family of timelike curves

$$y^- = y_0^- + \epsilon\,(y^+ - y_0^+)\,, \tag{A.3}$$

with $\epsilon > 0$, in the limit $\epsilon \to 0$.

The outward-directed unit normal to the timelike curve in (A.3) is given by

$$n = e^{-\rho}\left(-\frac{1}{\sqrt{\epsilon}}\partial_+ + \sqrt{\epsilon}\,\partial_-\right)\,, \tag{A.4}$$

and the determinant of the induced metric evaluates to

$$\sqrt{|h|} = e^{\rho}\sqrt{\epsilon}\,, \tag{A.5}$$

thus giving

$$\lim_{\epsilon\to 0}\int_{\partial M}\sqrt{|h|}n^\mu j_\mu = -\int_{\partial M} dy^+ j_+\,. \tag{A.6}$$

We could also have approximated the null curve $y^- = y_0^-$ by a family of spacelike curves and considered the past directed normal vector. The resulting limit agrees with (A.6).

Furthermore, a similar procedure can be applied for null curves defined by $y^+ = \text{const.}$ and for null boundaries that lie in the past of M. The general result, for the case of M only having null boundaries defined by $y^\mp = \text{const.}$, and not making reference to a particular coordinate system, can be written as

$$\int_M \sqrt{|g|}\nabla_\mu j^\mu = \sum_{\mathcal{N}\in\mathscr{N}} \sigma_{\mathcal{N}} \int_{\mathcal{N}} d\lambda\, k^\mu_{\mathcal{N}} j_\mu \tag{A.7}$$

where the sum runs over all piecewise smooth null boundary components. The future-directed null vector k^μ, tangential to the null boundary $y^\mp = \text{const.}$, is introduced, such that $\partial_\lambda = k^\mu\partial_\mu$, and $\sigma_{\mathcal{N}}$ are signs determined by

$$\sigma_{\mathcal{N}} = \begin{cases} 1 & \mathcal{N} \text{ lies in the past of } M \\ -1 & \mathcal{N} \text{ lies in the future of } M \end{cases} \tag{A.8}$$

In this form, the expression is manifestly invariant under reparametrizations $\lambda \mapsto \lambda'$.

References

[1] L. Susskind, *Three Lectures on Complexity and Black Holes*, arXiv:1810.11563 [INSPIRE].

[2] L. Susskind, L. Thorlacius and J. Uglum, *The Stretched horizon and black hole complementarity*, *Phys. Rev. D* **48** (1993) 3743 [hep-th/9306069] [INSPIRE].

[3] L. Susskind, *Computational Complexity and Black Hole Horizons*, *Fortsch. Phys.* **64** (2016) 24 [*Addendum ibid.* **64** (2016) 44] [arXiv:1403.5695] [INSPIRE].

JHEP07(2020)173

[4] A.R. Brown, D.A. Roberts, L. Susskind, B. Swingle and Y. Zhao, *Holographic Complexity Equals Bulk Action?*, *Phys. Rev. Lett.* **116** (2016) 191301 [arXiv:1509.07876] [INSPIRE].

[5] A.R. Brown, D.A. Roberts, L. Susskind, B. Swingle and Y. Zhao, *Complexity, action and black holes*, *Phys. Rev. D* **93** (2016) 086006 [arXiv:1512.04993] [INSPIRE].

[6] D. Stanford and L. Susskind, *Complexity and Shock Wave Geometries*, *Phys. Rev. D* **90** (2014) 126007 [arXiv:1406.2678] [INSPIRE].

[7] L. Susskind and Y. Zhao, *Switchbacks and the Bridge to Nowhere*, arXiv:1408.2823 [INSPIRE].

[8] S. Chapman, H. Marrochio and R.C. Myers, *Holographic complexity in Vaidya spacetimes. Part I*, *JHEP* **06** (2018) 046 [arXiv:1804.07410] [INSPIRE].

[9] S. Chapman, H. Marrochio and R.C. Myers, *Holographic complexity in Vaidya spacetimes. Part II*, *JHEP* **06** (2018) 114 [arXiv:1805.07262] [INSPIRE].

[10] L. Schneiderbauer, W. Sybesma and L. Thorlacius, *Holographic Complexity: Stretching the Horizon of an Evaporating Black Hole*, *JHEP* **03** (2020) 069 [arXiv:1911.06800] [INSPIRE].

[11] C.G. Callan Jr., S.B. Giddings, J.A. Harvey and A. Strominger, *Evanescent black holes*, *Phys. Rev. D* **45** (1992) R1005(R) [hep-th/9111056] [INSPIRE].

[12] J.G. Russo, L. Susskind and L. Thorlacius, *The Endpoint of Hawking radiation*, *Phys. Rev. D* **46** (1992) 3444 [hep-th/9206070] [INSPIRE].

[13] A.R. Brown, H. Gharibyan, H.W. Lin, L. Susskind, L. Thorlacius and Y. Zhao, *Complexity of Jackiw-Teitelboim gravity*, *Phys. Rev. D* **99** (2019) 046016 [arXiv:1810.08741] [INSPIRE].

[14] K. Goto, H. Marrochio, R.C. Myers, L. Queimada and B. Yoshida, *Holographic Complexity Equals Which Action?*, *JHEP* **02** (2019) 160 [arXiv:1901.00014] [INSPIRE].

[15] S.N. Solodukhin, *Two-dimensional quantum corrected eternal black hole*, *Phys. Rev. D* **53** (1996) 824 [hep-th/9506206] [INSPIRE].

[16] R.C. Myers, *Black hole entropy in two-dimensions*, *Phys. Rev. D* **50** (1994) 6412 [hep-th/9405162] [INSPIRE].

[17] K. Parattu, S. Chakraborty, B.R. Majhi and T. Padmanabhan, *A Boundary Term for the Gravitational Action with Null Boundaries*, *Gen. Rel. Grav.* **48** (2016) 94 [arXiv:1501.01053] [INSPIRE].

[18] L. Lehner, R.C. Myers, E. Poisson and R.D. Sorkin, *Gravitational action with null boundaries*, *Phys. Rev. D* **94** (2016) 084046 [arXiv:1609.00207] [INSPIRE].

[19] J.D. Hayward, *Entropy in the RST model*, *Phys. Rev. D* **52** (1995) 2239 [gr-qc/9412065] [INSPIRE].

[20] J. Cruz, J. Navarro-Salas, C.F. Talavera and M. Navarro, *Conformal and non-conformal symmetries in 2D dilaton gravity*, *Phys. Lett. B* **402** (1997) 270 [hep-th/9606097] [INSPIRE].

[21] L. Thorlacius, *Black hole evolution*, *Nucl. Phys. B Proc. Suppl.* **41** (1995) 245 [hep-th/9411020] [INSPIRE].

[22] D. Carmi, R.C. Myers and P. Rath, *Comments on Holographic Complexity*, *JHEP* **03** (2017) 118 [arXiv:1612.00433] [INSPIRE].

[23] D. Carmi, S. Chapman, H. Marrochio, R.C. Myers and S. Sugishita, *On the Time Dependence of Holographic Complexity*, *JHEP* **11** (2017) 188 [arXiv:1709.10184] [INSPIRE].

JHEP07(2020)173

[24] S.M. Christensen and S.A. Fulling, *Trace Anomalies and the Hawking Effect*, *Phys. Rev. D* **15** (1977) 2088 [INSPIRE].

[25] A.V. Frolov, K.R. Kristjansson and L. Thorlacius, *Semi-classical geometry of charged black holes*, *Phys. Rev. D* **72** (2005) 021501 [hep-th/0504073] [INSPIRE].

[26] A.V. Frolov, K.R. Kristjansson and L. Thorlacius, *Global geometry of two-dimensional charged black holes*, *Phys. Rev. D* **73** (2006) 124036 [hep-th/0604041] [INSPIRE].

[27] A. Almheiri and J. Polchinski, *Models of AdS_2 backreaction and holography*, *JHEP* **11** (2015) 014 [arXiv:1402.6334] [INSPIRE].

JHEP07(2020)173

Article III

Published for SISSA by Springer

Received: *April 22, 2020*
Accepted: *April 29, 2020*
Published: *May 20, 2020*

Page curve for an evaporating black hole

Friðrik Freyr Gautason,[a,b] **Lukas Schneiderbauer,**[b] **Watse Sybesma**[b] **and Lárus Thorlacius**[b]

[a] *Instituut voor Theoretische Fysica, KU Leuven, Celestijnenlaan 200D, 3001 Leuven, Belgium*
[b] *Science Institute University of Iceland, Dunhaga 3, 107 Reykjavík, Iceland*

E-mail: ffg@kuleuven.be, lukas.schneiderbauer@gmail.com, watse@hi.is, lth@hi.is

Abstract: A Page curve for an evaporating black hole in asymptotically flat spacetime is computed by adapting the Quantum Ryu-Takayanagi (QRT) proposal to an analytically solvable semi-classical two-dimensional dilaton gravity theory. The Page time is found to be one third of the black hole lifetime, at leading order in semi-classical corrections. A Page curve is also obtained for a semi-classical eternal black hole, where energy loss due to Hawking evaporation is balanced by an incoming energy flux.

Keywords: 2D Gravity, Black Holes, AdS-CFT Correspondence

ArXiv ePrint: 2004.00598

https://doi.org/10.1007/JHEP05(2020)091

JHEP05(2020)091

Contents

JHEP05(2020)091

1 Introduction

If black hole evaporation is a unitary process, the entanglement entropy between the outgoing radiation and the quantum state associated to the remaining black hole is expected to follow the so-called Page curve as a function of time [1, 2]. Early on, the entanglement entropy is then a monotonically increasing function of time which closely tracks the coarse grained thermal entropy of the radiation that has been emitted up to that point. This changes when the coarse grained entropy of the radiation exceeds the coarse grained entropy associated to the remaining black hole, at which point the entanglement entropy is limited by the black hole entropy and becomes a decreasing function of time. The time when the entanglement entropy transitions from increasing to decreasing is referred to as the Page time. Reproducing the Page curve without explicitly assuming unitarity is an important step towards resolving Hawking's black hole information paradox [3].

In a recent breakthrough, a Page curve was computed using semi-classical methods by studying black holes in asymptotically anti-de Sitter (AdS) spacetimes coupled to a conformal field theory (CFT) reservoir [4, 5].[1] The result hinges on the use of the Quantum Ryu-Takayanagi (QRT) formula [7–10] and the existence of extremal hypersurfaces terminating on so-called islands behind the event horizon [11]. A version of the Page curve can also be obtained for eternal AdS black holes, but in this case the islands extend outside

[1] See e.g. [6] for further details on such a setup.

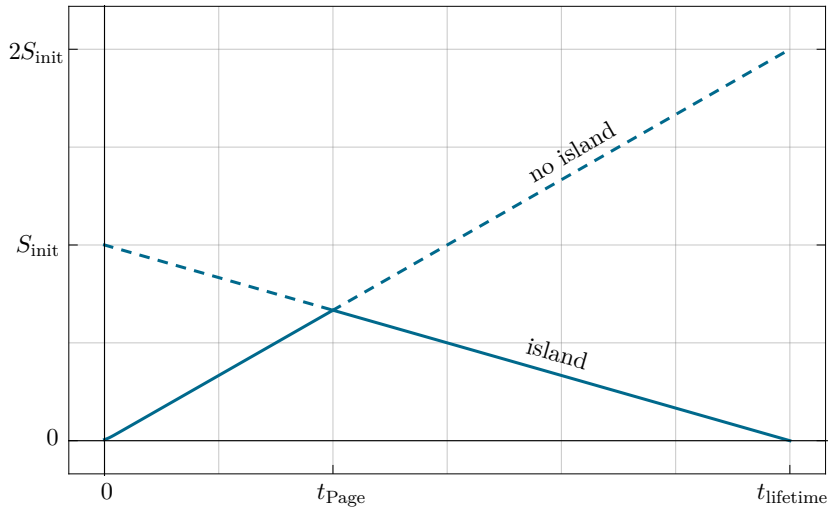


Figure 1. Page curve for an evaporating RST black hole.

the horizon [12]. Explicit computations have for the most part been restricted to two-dimensional Jackiw-Teitelboim gravity [5, 11], but see [13] for a discussion of islands in higher dimensional AdS black hole spacetimes.

In the present paper we demonstrate that the QRT formula can also be applied in the context of an evaporating black hole in asymptotically flat spacetime. At leading semi-classical order in the model that we use, and for a large initial black hole mass, the Page time is found to be one third of the black hole lifetime. This main result is presented in figure 1, where the entanglement entropy of the outgoing Hawking radiation that has passed beyond a distant spatial reference point is plotted as a function of time registered at the reference point.

We work with a two-dimensional dilaton gravity model of the type introduced by Callan, Giddings, Harvey, and Strominger (CGHS) in [14]. More specifically, the dilaton gravity sector is that of the model introduced by Russo, Susskind, and Thorlacius (RST) in [15], which remains analytically solvable at the semi-classical level. For the matter sector, we take a two-dimensional CFT with a large central charge $c \gg 1$, but rather than working with a large number of free scalar fields as in the CGHS-model, we assume that the conformal matter is holographic. This allows us to take advantage of an insight put forward by Almheiri et al. [11] in the context of two-dimensional AdS gravity, and use a three-dimensional gravitational dual description to calculate the contribution of the two-dimensional bulk matter to the generalized entropy in the QRT formula. In this model, the formation of a black hole from collapsing CFT matter into vacuum and its subsequent evaporation can be studied analytically. The Hawking radiation emitted in this process naturally propagates towards future null infinity. This is in contrast to the AdS setup, where the coupling to a heat bath is essential for letting the radiation escape. Our computation in asymptotically flat

JHEP05(2020)091

spacetime thus gives rise to a rather clean physical picture, both from a computational and conceptual point of view, and offers evidence that the QRT prescription applies beyond asymptotically AdS spacetimes.

The QRT prescription can be motivated via a replica trick involving Euclidean wormholes, as explored in [16–20]. Furthermore, the Page curve has been computed from boundary conformal field theories in [21] and the effect of a quench protocol on the Page curve was studied in [22]. Outside the direct scope of black hole physics, Page curves are found to be connected to the eigenstate thermalization hypothesis [23] and the study of chaos [24].

The paper is organised as follows. In section 2 we review the main result of [4, 6, 11] and argue for the validity of the QRT formula in our asymptotically flat spacetime model. Then in section 3 we introduce the two-dimensional dilaton gravity model and establish notation. Following that, in section 4 we derive the QRT formula for our model. Results for eternal black holes are presented in section 5.1 and for a dynamical black hole in section 5.2. In section 6 we present some conclusions and outlook.

2 Page curve from QRT

Our goal is to compute Page curves for black holes in asymptotically flat spacetime. We will do the computation both for an eternal black hole in an asymptotically linear dilaton spacetime and for a dynamical black hole formed by the gravitational collapse of matter into a linear dilaton vacuum. We follow the holographic approach of [4, 5], which uses the quantum Ryu-Takayanagi (QRT) formula

$$S_{\text{gen}} = \frac{\text{Area}(I)}{4G_N} + S_{\text{Bulk}}[\mathcal{S}_{AI}]\,. \tag{2.1}$$

The first term on the right hand side is the standard Ryu-Takayanagi entropy associated to a given subregion A of the spatial manifold on which the CFT in question is defined. Here I denotes a codimension two region that penetrates into the dual bulk spacetime and is homologous to A. The second term is the von Neumann entropy of bulk quantum fields with support inside a spacelike region bounded by $A \cup I$.

In [4] the system consists of a standard holographic CFT, with Hilbert space $\mathcal{H}_{\text{CFT}}$, at finite temperature T so that the dual geometry is an asymptotically AdS black hole. The CFT is assumed to be coupled to an auxiliary system, denoted by $\mathcal{H}_{\text{rad}}$, where Hawking radiation emanating from the black hole is collected. The region A in (2.1) above is taken to be the entire boundary where the CFT is defined, i.e. $\mathcal{H}_A = \mathcal{H}_{\text{CFT}}$ and $\mathcal{H}_{\bar{A}} = \mathcal{H}_{\text{rad}}$, and one looks for regions I, homologous to A, for which the generalized entropy (2.1) takes extremal values. The QRT prescription for entanglement entropy between A and $\bar{A}$, or in this case the entanglement entropy between the black hole and the Hawking radiation, is then given by the smallest extremal value of (2.1). At early times the minimum value corresponds to the "empty" surface $I = \varnothing$ and the generalized entropy S_{gen} is dominated by the von Neumann entropy of the Hawking radiation which grows monotonically with time. Eventually another extremum, involving an "island" I which lies just inside the horizon,

JHEP05(2020)091

takes over.[2] For this latter extremum, $S_{\rm gen}$ is dominated by the area term in (2.1) and is therefore given approximately by the Bekenstein-Hawking entropy of the black hole. The end result is the Page curve,

$$S_{\rm gen} = \min(S_{\rm rad}, S_{\rm BH})\,, \tag{2.2}$$

and a Page time defined as the time when the two extrema trade places providing the smallest extremal value of $S_{\rm gen}$.

In this paper we study black hole geometries in 1+1 dimensional dilaton gravity, which are asymptotically flat with an asymptotically linear dilaton field. Linear dilaton spacetimes are familiar from constructions in string theory where they arise as holographic duals to some non-conformal theories. A prominent example is given by the near-horizon limit of NSNS fivebranes, which is a spacetime of the form

$$\mathbf{R}^{1,5} \times \mathbf{R}_\phi \times S^3\,, \tag{2.3}$$

where $\mathbf{R}_\phi$ denotes the direction along which the string theory dilaton is linear. This background is an α'-exact solution of heterotic string theory [25]. The dual field theory in this case is $\mathcal{N} = (1,1)$ supersymmetric Yang-Mills theory in six dimensions, which does not flow to a conventional QFT in the UV but rather to a non-local theory called little string theory [26]. The dilaton gravity models studied in the present paper are in fact closely related to the above fivebrane background, as explained for example in [27]. However, this will not play an important role in our discussion beyond exemplifying that linear dilaton spacetimes can serve as holographic backgrounds for a class of non-conformal theories.

The holographic dictionary for such linear dilaton backgrounds works in a similar way as in standard AdS/CFT, except the dual variables are defined in the asymptotic linear dilaton region instead of the AdS region in conventional holography (see [26] for details). In our computation we will place a timelike "anchor curve" at a fixed radial position far outside the black hole. In the gravitational theory the Hawking radiation emitted from the black hole will pass through the anchor curve as depicted in figure 2. Hence, we do not need to artificially split our system into a QFT dual to the black hole plus an auxiliary system where the Hawking quanta are collected as in an AdS background. Instead the split is taken care of in a natural way by the anchor curve dividing the system into an "inside" part containing the black hole and an "outside" region containing outgoing Hawking radiation. We will compute the entanglement entropy between the radiation that has passed through the anchor curve and all that remains inside, including the black hole itself, and see explicitly that it follows a Page curve as a function of time experienced by asymptotic observers who remain stationary with respect to the black hole. The challenging aspect of the computation is the evaluation of the second term in (2.1) for any given trial island I. To simplify this task, we follow [11] and use $\mathrm{AdS}_3/\mathrm{CFT}_2$ duality to compute the von Neumann entropy of the bulk fields using a standard Ryu-Takayanagi prescription. We will come back to this in section 4.

[2] More precisely, the island refers to the internal entanglement wedge defined by I [11]. In this paper, we will often neglect this distinction and refer to I itself as the island.

JHEP05(2020)091

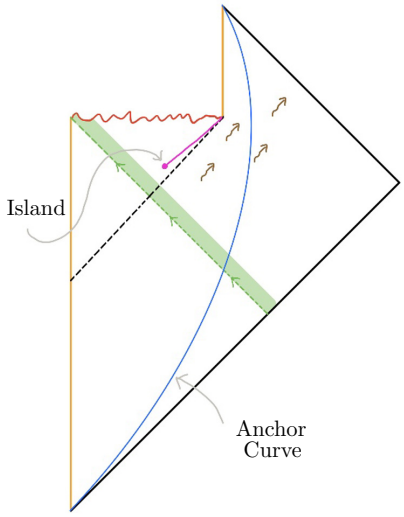


Figure 2. Penrose diagram of an evaporating RST black hole formed from collapsing matter (green). A timelike anchor curve separates the spacetime into interior and exterior regions. As time evolves along this curve, more and more Hawking radiation has passed through it on its way to future null infinity. The island moves with time along the purple curve inside the event horizon.

3 The model

We start with the classical CGHS dilaton gravity action [14],

$$I_{\text{grav}} = \frac{1}{2\pi}\int \mathrm{d}^2x\sqrt{-g}\,\mathrm{e}^{-2\phi}\left\{R + 4(\nabla\phi)^2 + 4\lambda^2\right\}\,, \tag{3.1}$$

where ϕ is the dilaton field and λ is a characteristic length scale that can be set to $\lambda = 1$ by a rescaling of the two-dimensional coordinates. The vacuum solution is given by flat spacetime with a linear dilaton profile,

$$\mathrm{d}s^2 = -\mathrm{d}\sigma^+\mathrm{d}\sigma^-\,, \quad \phi = \phi_0 - \sigma\,, \tag{3.2}$$

where ϕ_0 is an arbitrary constant that can be absorbed by a constant shift of the spatial coordinate $\sigma = \frac{1}{2}(\sigma^+ - \sigma^-)$. The strength of the gravitational coupling is controlled by the dilaton field and becomes large as σ tends to $-\infty$. The two-dimensional model can be viewed as a spherical reduction of a four-dimensional theory. In this case, the scale λ is inherited from the parent theory and $\mathrm{e}^{-2\phi}$ is proportional to the area of the transverse 2-sphere in four-dimensional Planck units.

We will find it useful to employ so-called Kruskal coordinates, for which the metric in conformal gauge takes the form

$$\mathrm{d}s^2 = -\mathrm{e}^{2\rho(x^+,x^-)}\mathrm{d}x^+\mathrm{d}x^-\,, \tag{3.3}$$

JHEP05(2020)091

with the conformal factor equal to the dilaton $\rho = \phi$. In this coordinate system the equations of motion obtained from (3.1) reduce to

$$\partial_+\partial_- \mathrm{e}^{-2\phi} + 1 = \partial_+^2 \mathrm{e}^{-2\phi} = \partial_-^2 \mathrm{e}^{-2\phi} = 0\,. \tag{3.4}$$

In the absence of matter fields, the dilaton gravity theory is non-dynamical and the general solution to the above field equations, up to constant shifts of the $x^\pm$ coordinates, is given by

$$\mathrm{e}^{-2\phi} = M - x^+x^-\,, \tag{3.5}$$

where M an integration constant. For $M = 0$ we get back the vacuum solution (3.2), written in Kruskal coordinates. For $M < 0$ the solution exhibits a naked singularity analogous to the negative mass Schwarzschild solution in four dimensions. For $M > 0$, a rescaling of the coordinates, $x^+ = \sqrt{M}v$ and $x^- = \sqrt{M}u$, gives the well-known two-dimensional "cigar" solution in Lorentzian signature [28, 29],

$$\mathrm{d}s^2 = -\frac{\mathrm{d}v\,\mathrm{d}u}{1 - vu}\,, \tag{3.6}$$

with a curvature singularity at $vu = 1$ and a bifurcate event horizon at $vu = 0$. The integration constant M is proportional to the black hole mass,

$$\mathcal{M} = \frac{\lambda M}{\pi}\,, \tag{3.7}$$

where we have temporarily restored the characteristic mass scale λ. The temperature of a CGHS black hole is independent of its mass,

$$T = \frac{\lambda}{2\pi}\,. \tag{3.8}$$

The Bekenstein-Hawking entropy, given by one quarter of the horizon area in Planck units in the original four dimensional theory, can be expressed in terms of the dilaton field evaluated at the horizon,

$$S = 2\mathrm{e}^{-2\phi}\Big|_{\text{Horizon}} = 2M = \frac{2\pi\mathcal{M}}{\lambda}\,. \tag{3.9}$$

A purely two-dimensional argument leading to the dilaton dependence in (3.9) is that while the area of the horizon is unity, the gravitational coupling constant is controlled by the dilaton as is apparent from (3.1), and this must be taken into account when evaluating the Bekenstein-Hawking entropy $S = \text{Area}/4G_N$.

3.1 Coupling to matter

In the original CGHS model [14], the dilaton gravity sector is coupled to matter in the form of N minimally coupled free scalars, with $N \gg 24$ so that semi-classical corrections are dominated by one-loop effects due to the matter fields. Here we will instead assume a strongly coupled matter sector described by a holographic two-dimensional CFT with large central charge c that has an AdS_3 gravitational dual. This is an important technical assumption which allows us to simply evaluate the von Neumann entropy of the CFT fields

JHEP05(2020)091

on a spacelike section but does not affect the gravitational sector. In particular, the theory still has solutions describing dynamical black holes formed from incoming matter energy-momentum.[3]

Through the holographic dictionary, the two-dimensional central charge is related to three-dimensional gravitational quantities via the Brown-Henneaux formula,

$$c = \frac{3L_3}{2G_{(3)}} \,. \tag{3.10}$$

As discussed in section 5.2 below, we can arrange our computation of the von Neumann entropy of the matter fields in such a way that we only have to deal with pure gravity in AdS_3 spacetime.

We are interested in semi-classical black holes with initial mass M large compared to the scale set by the central charge of the matter CFT, for which there is a natural expansion parameter given by

$$\epsilon \equiv \frac{c}{48M} \ll 1 \,. \tag{3.11}$$

In most of what follows we work to leading non-trivial order in ϵ, but to get started it is useful to consider the $\epsilon \to 0$ limit where semi-classical effects are turned off. In this limit, the gravitational field equation is sourced by the energy-momentum tensor of the two-dimensional matter CFT,[4]

$$4\mathrm{e}^{-2\phi}\Big[\nabla_\mu\nabla_\nu\phi - g_{\mu\nu}\big(\Box\phi - (\nabla\phi)^2 + 1\big)\Big] = T_{\mu\nu} \,, \tag{3.12}$$

while the equation of motion of the dilaton field is unaffected by the coupling to matter. The CFT energy-momentum tensor has two non-trivial components $T_{++}(x^+)$ and $T_{--}(x^-)$, each of which only depends on one of the light cone coordinates. The field equations take a particularly simple form in the Kruskal coordinates (3.3),

$$-\partial_+\partial_-\mathrm{e}^{-2\phi} = 1 \,, \quad -2\partial_\pm^2\mathrm{e}^{-2\phi} = T_{\pm\pm} \,, \tag{3.13}$$

and the response to arbitrary incoming matter energy flux is easily obtained,

$$\mathrm{e}^{-2\phi} = \mathrm{e}^{-2\rho} = F(x^+) - x^+\big[x^- + G(x^+)\big] \,, \tag{3.14}$$

where

$$F'(x^+) = \frac{1}{2}x^+T_{++}(x^+) \,, \qquad G'(x^+) = \frac{1}{2}T_{++}(x^+) \,. \tag{3.15}$$

We take the energy-momentum tensor to have compact support in x^+ corresponding to a thin shell of infalling matter energy incident on the linear dilaton vacuum. For our purposes,

[3] When comparing to black holes in [14, 15] we make the identification $N = c$ and $\kappa = c/12$.

[4] Our conventions match those of [30] except we are dealing with Lorentzian CFT. In particular, the classical energy-momentum tensor is defined as

$$T^{\mu\nu} = -\frac{4\pi}{\sqrt{-g}}\frac{\delta S}{\delta g_{\mu\nu}} \,,$$

and the normalization for the energy-momentum tensor therefore differs from the one used in [14] by a factor of 2.

JHEP05(2020)091

the detailed form of the solution (3.14) is not needed, only the behaviour at early and late times, and we can therefore consider an idealised solution where two static configurations are patched together across an infinitely thin null shock wave,

$$\mathrm{e}^{-2\phi(x^+,x^-)} = \mathrm{e}^{-2\rho(x^+,x^-)} = \begin{cases} -x^+x^- & \text{if } x^+ < x_0^+\,, \\ M - x^+\left(x^- + \frac{M}{x_0^+}\right) & \text{if } x^+ > x_0^+\,. \end{cases} \tag{3.16}$$

A rescaling of the coordinates,

$$x^+ = x_0^+\, v\,, \quad x^- = \frac{M}{x_0^+}\, u\,, \tag{3.17}$$

brings the metric and dilaton into the following simple form,

$$\mathrm{e}^{-2\rho(v,u)} = \frac{1}{M}\,\mathrm{e}^{-2\phi(v,u)} = \begin{cases} -vu & \text{if } v < 1\,, \\ (1 - v(u+1)) & \text{if } v > 1\,. \end{cases} \tag{3.18}$$

In the $v < 1$ linear dilaton region, the change of coordinates,

$$v = e^{\omega^+}\,, \qquad u = -e^{-\omega^-}\,, \tag{3.19}$$

brings the metric into manifestly flat form, $\rho(\omega^+, \omega^-) = 0$, while a set of coordinates, for which the metric is asymptotically Minkowskian in the $v > 1$ region outside the shock wave, is given by

$$v = e^{\sigma^+}\,, \qquad u = -1 - e^{-\sigma^-}\,. \tag{3.20}$$

Time as measured by asymptotic observers at rest with respect to the black hole is

$$t = \frac{1}{2}(\sigma^+ + \sigma^-)\,. \tag{3.21}$$

3.2 Semi-classical black holes

On a curved spacetime background, the energy-momentum tensor of the matter CFT is no longer traceless due to the conformal anomaly,

$$\langle T^\mu_{\ \mu} \rangle = \frac{c}{12} R\,, \tag{3.22}$$

where R is the Ricci scalar of the background metric. In two spacetime dimensions the continuity equation expressing energy-momentum conservation can be integrated using only (3.22) as input [31] leading to the following expressions in conformal coordinates,

$$\langle T_{+-} \rangle = -\frac{c}{6}\partial_+\partial_-\rho\,, \quad \langle T_{\pm\pm} \rangle = \frac{c}{12}\left(2\partial_\pm^2\rho - 2(\partial_\pm\rho)^2 - t_\pm\right)\,, \tag{3.23}$$

where $t_\pm(x^\pm)$ are functions of integration determined by physical boundary conditions that reflect the matter quantum state.

The boundary functions $t_\pm$ are sensitive to the choice of coordinate system. This is to be expected since notions of positive frequency and normal ordering depend on the

JHEP05(2020)091

choice of time variable. Under a conformal reparametrization of the light-cone coordinates, $x^\pm \to y^\pm(x^\pm)$, the conformal factor of the metric transforms as

$$\rho(y^+, y^-) = \rho(x^+, x^-) - \frac{1}{2}\log\frac{\mathrm{d}y^+}{\mathrm{d}x^+}\frac{\mathrm{d}y^-}{\mathrm{d}x^-}\,. \tag{3.24}$$

When inserted in (3.23) this leads to the usual anomalous transformation of the energy-momentum tensor involving a Schwarzian derivative,

$$\left(\frac{dy^\pm}{dx^\pm}\right)^2 T_{\pm\pm}(y^\pm) = T_{\pm\pm}(x^\pm) - \frac{c}{12}\{y^\pm, x^\pm\}\,, \quad \{y, x\} = \frac{y'''}{y'} - \frac{3}{2}\frac{(y'')^2}{(y')^2}\,. \tag{3.25}$$

In order to preserve the form (3.23) for the energy-momentum tensor in the new coordinates, we effectively obtain a new function $t_\pm(y^\pm)$ related to the old one via

$$\left(\frac{dy^\pm}{dx^\pm}\right)^2 t_\pm(y^\pm) = t_\pm(x^\pm) + \{y^\pm, x^\pm\}\,. \tag{3.26}$$

As an example, consider a black hole formed by gravitational collapse as in (3.18). At early advanced time before the arrival of the collapsing matter (i.e. $v < 1$), we have a linear dilaton vacuum and vanishing energy-momentum tensor. The metric is manifestly flat in the (ω^+, ω^-) coordinate system in (3.19) and it follows that $t_-(\omega^-) = 0$. Upon transforming to the (σ^+, σ^-) coordinate system (3.20), in which the metric is manifestly asymptotically flat, one finds non-vanishing outgoing energy flux at $\sigma^+ \to \infty$,

$$T_{--}(\sigma^-) = -\frac{c}{12}t_-(\sigma^-) = \frac{c}{24}\left(1 - \frac{1}{(1+\mathrm{e}^{\sigma^-})^2}\right)\,. \tag{3.27}$$

In [14] this expression was interpreted as the energy flux of Hawking radiation from the black hole as observed by an asymptotic observer. Energy conservation implies that a black hole emitting Hawking radiation loses mass. When the semi-classical expansion parameter ϵ in (3.11) has a small but finite value, the classical solution (3.14) is only valid on timescales that are short compared to the lifetime of the black hole, which is $t_{\text{lifetime}} = 1/\epsilon$ at leading order. The semi-classical back-reaction on the spacetime geometry due to Hawking emission matter can, however, be accounted for by adding to the classical action I_{grav} in (3.1) a non-local Polyakov term induced by matter quantum effects [14],

$$I_Q = -\frac{c}{12\pi}\int \mathrm{d}x^+\mathrm{d}x^-\partial_+\rho\partial_-\rho\,, \tag{3.28}$$

written here in conformal coordinates.[5] If we take $c \gg 24$ then I_Q should be dominant compared to semi-classical contributions from the dilaton gravity sector. Further modifications to the theory are needed in order to find analytic solutions to the semi-classical equations of motion. We will follow the approach of [15] and add the following term,

$$I_{\text{RST}} = \frac{c}{48\pi}\int \mathrm{d}^2x\sqrt{-g}\phi R\,, \tag{3.29}$$

[5] While the non-local nature of I_Q is not immediately apparent in the conformal gauge expression (3.28), it enters the formalism via the boundary functions $t_\pm$ in (3.23).

JHEP05(2020)091

to the semi-classical action, which is allowed by general covariance and does not disturb the classical ($\epsilon \to 0$) limit of the theory. The RST term $I_{\rm RST}$ involves a factor of c and therefore enters at the same order as the Polyakov term I_Q.

The resulting semi-classical field equations simplify dramatically when a new field variable is introduced,

$$\Omega = \mathrm{e}^{-2\phi} + \frac{c}{24}\phi\,, \tag{3.30}$$

but there are subtleties involved. In particular, the new field variable is bounded from below, $\Omega \geq \Omega_{\rm crit} = \frac{c}{48}\left(1 - \log\frac{c}{48}\right)$, and when $\Omega \to \Omega_{\rm crit}$ the gravitational coupling becomes strong in the semi-classical theory [15]. This has a suggestive physical interpretation, where $\Omega \to \Omega_{\rm crit}$ represents a boundary of spacetime, analogous to the boundary at the origin of radial coordinates in the higher-dimensional theory from which the CGHS model is descended.

One benefit of including the RST term (3.29) is that semi-classical solutions of the full theory $I_{\rm grav} + I_Q + I_{\rm RST}$ can be expressed in Kruskal coordinates (3.3), where the field equations reduce to

$$\partial_+\partial_-\Omega + 1 = 0\,, \quad -\partial_\pm^2\Omega = \frac{c}{24}t_\pm\,, \tag{3.31}$$

with $t_\pm$ the same boundary functions as before. The linear dilaton vacuum remains an exact solution of the semi-classical equations and takes the form

$$\Omega = -x^+x^- - \frac{c}{48}\log(-x^+x^-)\,, \tag{3.32}$$

in the new field variable. Notice that $t_\pm(x^\pm) \neq 0$ even if this is the vacuum solution but this is because the metric is not manifestly flat in Kruskal coordinates. Transforming to a manifestly flat coordinate system $\sigma^\pm$ renders the functions $t_\pm(\sigma^\pm) = 0$ as expected.

A two-sided eternal black hole solution is given by

$$\Omega = M(1 - vu) + \Omega_{\rm crit}\,, \tag{3.33}$$

where we have rescaled the coordinates as in (3.6). Here we find that in Kruskal coordinates that $t_\pm(x^\pm) = 0$ but if we transform to coordinates for which the metric is manifestly asymptotically flat,

$$v = e^{\sigma^+}\,, \qquad u = -e^{-\sigma^-}\,, \tag{3.34}$$

we find that $t_\pm(\sigma^\pm) = \frac{1}{2}$. This corresponds to a flat space energy-momentum tensor $T_{\pm\pm}(\sigma^\pm) = \frac{c}{24}$ which is exactly the energy-momentum tensor of a thermal gas of temperature $T = \frac{1}{2\pi}$ which is the temperature of the eternal black hole. The outgoing energy flux carried by the Hawking radiation is matched by an incoming flux of thermal radiation at the same temperature as the Hawking temperature of the black hole.

Finally, consider the formation and subsequent evaporation of a dynamical black hole. As in the classical case without back-reaction, we imagine a situation where a short burst of matter energy is injected into a linear dilaton vacuum described by (3.32). The solution describing the full evolution of such a black hole can be found in [15]. Here we are mainly interested in the geometry outside the collapsing matter shell, i.e. for $v > 1$, where it takes the form

$$\mathrm{d}s^2 = -M\mathrm{e}^{2\phi}\,\mathrm{d}v\mathrm{d}u\,, \quad \Omega = M\big(1 - v(u+1) - \epsilon\log(-Mvu)\big)\,, \tag{3.35}$$

with ϕ and Ω related via (3.30).

JHEP05(2020)091

4 Generalized entropy

In order to derive a Page curve for these semi-classical black holes, we adapt the expression for the generalized entropy,

$$S_{\rm gen} = \frac{\text{Area}(I)}{4G_N} + S_{\rm Bulk}[\mathcal{S}_{AI}]\,, \tag{4.1}$$

to the two-dimensional setting at hand. The first term on the right hand side involves the area of the transverse two-sphere evaluated locally at an island, and gives zero in the absence of an island. Comparing with the black hole entropy in (3.9) yields $2e^{-2\phi(I)}$ as the area contribution of an island in the classical limit. The natural semi-classical extension of this expression, which gives zero in the absence of an island, is given by

$$\frac{\text{Area}(I)}{4G_N} = 2\big(\Omega(I) - \Omega_{\rm crit}\big). \tag{4.2}$$

The second term on the right hand side in (4.1) is universal and is the main focus of this section. It is the von Neumann entropy of the CFT matter fields on a spacelike surface $\mathcal{S}_{AI}$ that is bounded at one end by the island I and at the other end by a point A on a timelike anchor curve. We take the anchor curve to be a constant Ω curve with $\Omega = \Omega_A \gg M$ so that it is located well outside the black hole. For an eternal black hole (3.33) a curve of constant Ω is at a fixed spatial coordinate, $\sigma = \sigma_A$ in the manifestly asymptotically flat coordinate system (3.34). For an evaporating black hole, the corresponding statement is no longer exact due to the log term in (3.35). The spatial location of the anchor curve drifts in the asymptotic coordinates (3.20) but for $\Omega_A \gg M$ the drift is extremely slow and can be ignored on time scales of order the black hole lifetime. The final answer for $S_{\rm Bulk}$ does not depend on which $\mathcal{S}_{AI}$ is chosen as long as it is a spacelike surface that connects A and I. In the absence of an island, the surface $\mathcal{S}_{AI}$ is instead bounded by A at one end and a point on the boundary curve $\Omega = \Omega_{\rm crit}$ at the other.

Following [11], we compute the von Neumann entropy holographically by passing to a three-dimensional gravitational theory and evaluating the geodesic length between the points where A and I are embedded in the dual three-dimensional spacetime,

$$S_{\rm Bulk}[\mathcal{S}_{AI}] \simeq \frac{\text{Length}}{4G_{(3)}}\,. \tag{4.3}$$

The calculation is simplified if we arrange the embedding geometry to be pure AdS_3. This can be achieved in two steps. The first step is to identify a set of light-cone coordinates

$$\mathrm{d}s^2 = -\mathrm{e}^{2\rho}\mathrm{d}y^+\mathrm{d}y^-\,, \tag{4.4}$$

where the integration functions $t_\pm(y^\pm)$ are zero. The second step is to perform a Weyl rescaling of the two-dimensional metric that strips off the conformal factor $e^{2\rho}$. Then both $t_\pm(y^\pm)$ and the gravitational contribution the energy-momentum tensor in (3.23) vanish. In this case, the matter CFT is in a vacuum state and the dual three-dimensional geometry is empty AdS_3 spacetime. In Poincare coordinates the metric is

$$\mathrm{d}s_3^2 = \frac{L_3^2}{z^2}\left(\mathrm{d}z^2 - \mathrm{d}y^+\mathrm{d}y^-\right), \tag{4.5}$$

JHEP05(2020)091

and the geodesics are semi-circles centered on the holographic boundary. The Weyl transformation in step two above can be implemented as a coordinate transformation in three dimensions which maps the regulated holographic boundary to a surface,

$$z = \delta\, \mathrm{e}^{-\rho(y^+,y^-)}\,, \tag{4.6}$$

that depends on dynamical input from the two-dimensional matter theory. Here δ is a UV cutoff parameter.

A standard calculation involving AdS_3 geodesics then leads to the following result for the holographic entropy,

$$S_{\mathrm{Bulk}}[\mathcal{S}_{AI}] = \frac{c}{6}\log\left[d(A,I)^2 \mathrm{e}^{\rho(A)}\mathrm{e}^{\rho(I)}\right]\Big|_{t_\pm=0}\,, \tag{4.7}$$

where $d(A,I)$ is the two-dimensional distance measured between the points A and I in the flat metric $\mathrm{d}s^2_{\mathrm{flat}} = -\mathrm{d}y^+\mathrm{d}y^-$ and the subscript is a reminder that the formula should be evaluated in coordinates for which $t_\pm(y^\pm) = 0$. We have dropped the UV cutoff from the formula as it just contributes an additive constant.

5 Page curves

We now have everything in place to calculate generalized entropy in the RST model using the QRT prescription. Our primary goal is to obtain the Page curve of an evaporating black hole that has a finite lifetime but first we carry out the corresponding calculation for a semi-classical eternal black hole. This provides a first test involving a black hole in asymptotically flat spacetime which turns out to be considerably simpler than the evaporating case.

5.1 Eternal black hole

A semi-classical eternal black hole in asymptotically flat spacetime is supported by a thermal gas of incoming radiation that maintains the mass of the black hole against the energy loss to Hawking radiation. The two-dimensional black holes studied in this paper all have temperature $T = \frac{1}{2\pi}$ and the thermal gas must be at the same temperature. As was noted below (3.33), the energy flux outside an eternal semi-classical RST black hole is

$$T_{\pm\pm}(\sigma^\pm) = \frac{c}{24}\,, \tag{5.1}$$

when evaluated in manifestly asymptotically flat coordinates and this is precisely the energy-momentum tensor of a thermal CFT at a temperature of $T = \frac{1}{2\pi}$. It was also noted that $t_+(v) = t_-(u) = 0$ for an eternal black hole and therefore (v,u) is the appropriate set of coordinates to use when evaluating S_{Bulk} in (4.7). The relevant matter CFT vacuum state is the Hartle-Hawking state where positive frequency modes are determined with respect to time in Kruskal coordinates rather than asymptotic Minkowski time.

The eternal black hole is two sided and we place a timelike anchor curve in each asymptotic region. For simplicity, we assume that our anchor points lie symmetrically on the anchor curves, as shown in figure 3. Then each anchor point has a mirror anchor point

JHEP05(2020)091

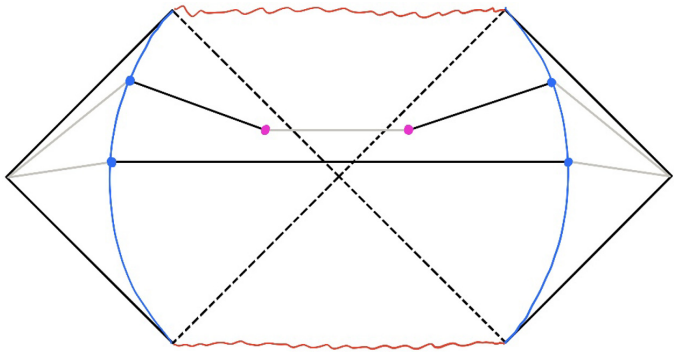


Figure 3. A Penrose diagram of an eternal black hole. A pair of timelike anchor curves (blue curves) separates the spacetime into an interior and two exteriors. The two spatial hypersurfaces intersect the anchor curves at different times. On the late time surface the generalized entropy is dominated by the area term associated to the islands denoted by purple dots.

(denoted by superscript m) in the other exterior region, which is related to the original point by $(v,u)^m = (u,v)$. We also take the black hole mass to be large compared to the scale set by the matter central charge, so that $\epsilon = \frac{c}{48M} \ll 1$, and the anchor curves to be located in the linear dilaton region, so that $\Omega_A \gg M$. With these assumptions in place the semi-classical field variables are well approximated by their classical counterparts in all regions of interest and our calculations simplify.

Inspired by [4, 11], we now perform two calculations: one with no islands, and one with a single island on each side. Consider first the no-island scenario. In this case, the area term of the generalized entropy is by definition zero, as I is empty. The von Neumann entropy of the bulk fields is non-vanishing and given by the length of the geodesic in AdS_3 that connects the two mirrored anchor points. This means that we can directly apply (4.7) but with I replaced by A^m, as indicated in figure 3,

$$S_{\text{bulk}} = \frac{c}{12} \log \left[(v_A - v_{A^m})^2 (u_A - u_{A^m})^2 \, e^{2\rho(v_A,u_A)} e^{2\rho(v_{A^m},u_{A^m})} \right], \tag{5.2}$$

where (v_A, u_A) denotes an anchor point on the curve on the right in the figure. The anchor curves are assumed to be located well outside the black hole where the conformal factor is well approximated by its classical value,

$$e^{2\rho(v,u)} \approx \frac{1}{1-vu} \,. \tag{5.3}$$

The bulk entropy then takes a simple form,

$$S_{\text{bulk}} = \frac{c}{12} \log \frac{(v_A - u_A)^4}{(1 - v_A u_A)^2} \approx \frac{c}{3} t_A \,, \tag{5.4}$$

where t_A is asymptotic time, measured by an observer on the anchor curve, and the asymptotically flat coordinates (t,σ) are related to the (v,u) coordinates via,

$$v = e^{t+\sigma} \,, \qquad u = -e^{-t+\sigma} \,. \tag{5.5}$$

JHEP05(2020)091

Corrections to this result are either exponentially suppressed (by factors of e^{-2t_A} or $e^{-2\sigma_A}$) or subleading in powers of ϵ, or both. Our computation includes, by construction, the entropy of the radiation emitted on both sides of the black hole and we note that the entropy growth rate in (5.4) is precisely twice the rate that was obtained in [32] for the entanglement entropy of radiation emitted to one side.

We now repeat the calculation with symmetrically placed islands at $I = (v_I, u_I)$ and $I^m = (u_I, v_I)$, as indicated in figure 3. In this case, the area term in the generalized entropy (4.1) is non-vanishing and bulk term involves geodesics in AdS_3 that connect the anchor point and corresponding island on each side of the black hole. The contributions from the two sides of the black hole are identical and add up to

$$S_{\rm gen}^{\rm island} = 4M(1 - v_I u_I) + \frac{c}{6} \log \frac{(v_A - v_I)^2 (u_A - u_I)^2}{(1 - v_A u_A)(1 - v_I u_I)} \,, \tag{5.6}$$

where we have used (3.33) for the semi-classical area function $\Omega(I) - \Omega_{\rm crit}$. Both the anchor point and the island are assumed to lie in a region where the classical approximation (5.3) can be used for the conformal factor. This is automatically satisfied for an anchor point outside a large mass black hole and we will check ex post facto that it also holds for the island. Extremizing over (v_I, u_I) and working to leading order in $\epsilon \ll 1$, we obtain

$$\frac{u_I}{v_I} = \frac{u_A}{v_A} \,, \quad v_I \approx -\frac{4\epsilon}{u_A} \,. \tag{5.7}$$

This is a saddle point and not a minimum. However, the QRT prescription instructs us find all extrema and select the one that gives the lowest value for the generalized entropy. Inserting the leading order saddle point values for v_I and u_I into (5.6) gives

$$S_{\rm gen}^{\rm island} = 4M + \frac{c}{3} \sigma_A + \dots \,. \tag{5.8}$$

Comparing to the no-island result in (5.4) shows that for $t_A > \sigma_A + \frac{1}{4\epsilon}$ the generalized entropy is dominated by the island configuration. The σ_A term accounts for the time it takes for the Hawking radiation to travel from the black hole to the anchor curve. Correcting for this, we obtain

$$t_{\rm Page} = \frac{1}{4\epsilon} = \frac{12M}{c} \,, \tag{5.9}$$

for the Page time of an eternal RST black hole. The Page curve is drawn in figure 4.

For a two-sided black hole in AdS_2 gravity, the island and its mirror were found to be outside the event horizon [12]. This remains true here as well. The island saddle point (5.7) is outside the event horizon but inside the stretched horizon, with the proper distance between island and event horizon given by a tiny number,

$$d_I \approx 4\epsilon \sqrt{\frac{M}{\Omega_A}} \,. \tag{5.10}$$

The fact that the island is close to the event horizon justifies using the classical approximation for the conformal factor in (5.6), as promised. For another perspective on the location of the island, consider an observer sitting on the anchor curve who sends an ingoing light

JHEP05(2020)091

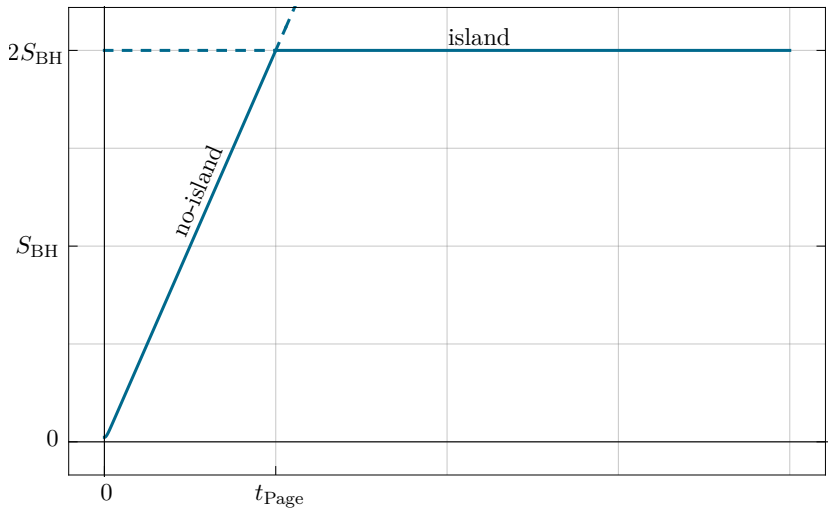


Figure 4. Page curve for the eternal RST black hole with $t_{\rm Page} = 6S_{\rm BH}/c$. The graph plots $S_{\rm gen} - \frac{c}{3}\sigma_A$ as a function of retarded time on the anchor curve.

signal to the island. A straightforward calculation shows that in order to be received at an island at (v_I, u_I), that corresponds to an anchor point at time t_A, the signal must be emitted from the anchor curve at an earlier time $t_A^{\rm obs}$, such that

$$t_A - t_A^{\rm obs} = 2\sigma_A + t_s\,, \tag{5.11}$$

where $t_s = \log\left(\frac{1}{4\epsilon}\right)$ is the scrambling time. This is in line with a similar result in [12]. However, because our black hole is in asymptotically flat spacetime and not AdS_2 the time difference $t_A - t_A^{\rm obs}$ explicitly depends on the location of the anchor curve.

5.2 Dynamical black hole

We now turn our attention to dynamical black holes and compute a Page curve for a black hole that is formed by gravitational collapse of matter and then gradually evaporates due to Hawking emission. The steps in the calculation are the same as before, i.e. to find extrema of the generalized entropy with and without an island and determine which one gives the minimum value. The area term in the generalized entropy can be read off directly from the semi-classical black hole solution but the remaining bulk term requires more work.

In the holographic evaluation of the bulk entropy term in (4.7) we are instructed to identify light-cone coordinates where the $t_{\pm}$ contribution to the two-dimensional matter energy momentum tensor is zero. The correct choice is the (ω^+, ω^-) system in (3.19) where the metric is manifestly flat in the initial linear dilaton region before the matter shell collapses to form the black hole. These coordinates are suitable for the evaluation of (4.7) when calculating the generalized entropy on a trial surface in the linear dilaton vacuum where the CFT is manifestly in its vacuum state and the three-dimensional holographic

JHEP05(2020)091

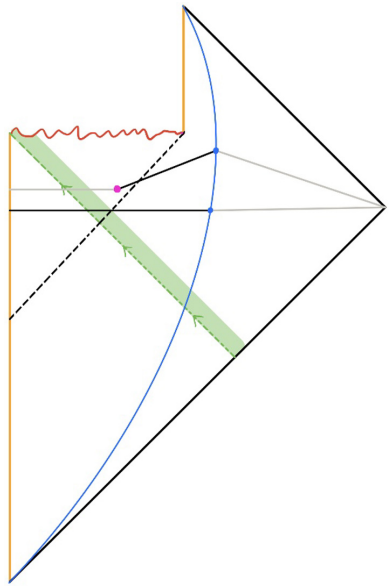


Figure 5. Penrose diagram of a dynamical RST black hole with two spacelike hypersurfaces indicated, one before the Page time and the other after, corresponding to the no-island and island configurations, respectively.

dual is pure AdS_3. Of course, a dynamical black hole is not the linear dilaton vacuum and $t_+(\omega^+)$ is non-vanishing due to the incoming energy flux that forms the black hole. There is, however, a simple way around this problem. Following [32], we take the incoming matter to be described by a coherent state built on the vacuum state of inertial observers at past null infinity. As shown in [32], the von Neumann entropy of such a state is identical to the von Neumann entropy of the vacuum state. As a result, we can use the AdS_3/CFT_2 Ryu-Takayanagi prescription (4.7) to calculate the bulk term in the generalized entropy, provided we use the coordinate system that corresponds to the CFT in its vacuum state. This means in particular, that we are instructed to calculate the two-dimensional distance $d(A, I)$ in (ω^+, ω^-) coordinates.

The generalized entropy is to be computed for the two competing configurations, with and without an island, indicated in the Penrose diagram in figure 5. The final result is the one that gives a smaller value for the entropy.

5.2.1 Island configuration

Let us start by determining the generalized entropy for an island configuration,

$$S_{\text{gen}}^{\text{island}} = 2M\big(1 - v_I(1+u_I) - \epsilon \log\left(-M v_I u_I\right)\big) + \frac{c}{12}\log\left[\left(\log\frac{v_A}{v_I}\log\frac{u_A}{u_I}\right)^2 \frac{v_A u_A}{(1-v_A(1+u_A))}\frac{v_I u_I}{(1-v_I(1+u_I))}\right], \tag{5.12}$$

JHEP05(2020)091

where we have used (3.35) for the area term $2(\Omega(I)-\Omega_{\rm crit})$ and the coordinate distance $d(A,I)=\sqrt{-\Delta\omega^+\Delta\omega^-}$ has been expressed in (v,u) coordinates. We are assuming that the island is located outside the infalling shell of matter and that both the anchor point and the island lie in a region where a classical approximation can be used for the conformal factor of the dynamical black hole metric. The anchor point is by assumption far outside the black hole where the classical approximation is always valid. It turns out to also be valid for the island for much of the lifetime of an evaporating black hole provided it starts out with a large enough mass but it will fail towards the end of the lifetime when the black hole has evaporated down to a small size.

Extremizing (5.13) over (v_I,u_I) yields the following two conditions,

$$0=-2M(1+u_I)+\frac{c}{12v_I}\,\frac{1+u_I}{(1-v_I(1+u_I))}+\frac{c}{24v_I}-\frac{c}{6v_I\log\left(\frac{v_A}{v_I}\right)}\,, \tag{5.13}$$

$$0=-2Mv_I+\frac{c}{12u_I}\,\frac{v_I}{(1-v_I(1+u_I))}+\frac{c}{24u_I}-\frac{c}{6u_I\log\left(\frac{u_A}{u_I}\right)}\,. \tag{5.14}$$

In order to solve for the location of the island we make the simplifying assumption $\log(\frac{v_A}{v_I})\gg 1$, which allows us to drop the last term on the right in the top equation, and later on we verify the self-consistency of this assumption. The resulting equations can be rearranged as

$$(v_I(1+u_I))^2-(1-\epsilon)\,v_I(1+u_I)+\epsilon=0\quad\text{ and }\quad\log\left(\frac{u_A}{u_I}\right)=4(1+u_I)\,. \tag{5.15}$$

One of the two solutions of the quadratic equation for $v_I(1+u_I)$ corresponds to an island in the near horizon region,

$$u_I=-1+\frac{\epsilon}{v_I}+O(\epsilon^2)\,. \tag{5.16}$$

The other solution has the island near the black hole singularity and is unphysical. Inserting the island solution into the remaining equation in (5.15) we find

$$u_A=-1-\frac{3\epsilon}{v_I}+O(\epsilon^2)\,. \tag{5.17}$$

In terms of the asymptotic coordinates (3.20) we have

$$u_A=-1-e^{\sigma_A-t_A}\,,\qquad v_A=e^{\sigma_A+t_A}\,. \tag{5.18}$$

The above relations imply that

$$\log\left(\frac{v_A}{v_I}\right)\approx 2\sigma_A+\log\left(\frac{1}{\epsilon}\right)\gg 1\,, \tag{5.19}$$

so the simplifying assumption that we used to obtain the island solution is indeed justified. We also assumed in the calculation that the island is located at $v_I>1$ and this turns out to be valid when $t_A-\sigma_A\gtrsim\log(1/\epsilon)$. The expression for the outgoing energy flux (3.27) reveals that the first Hawking radiation passes through the anchor curve at $t_A-\sigma_A\approx 0$ and the island solution is already valid within a time of order the scrambling time after that.

JHEP05(2020)091

We can again probe the location of the island by considering an observer sitting on the anchor curve who sends an ingoing light signal. The relation between v_I and v_A in (5.19) implies that in order to be received at an island at (v_I, u_I), the signal must be emitted from the anchor curve at a time $t_A^{\rm obs}$, such that $t_A - t_A^{\rm obs} = 2\sigma_A + t_s$ with t_s the black hole scrambling time. Earlier, we found the same result for an island just outside the event horizon of an eternal black hole. Here the island is inside the black hole but still located very close to the event horizon.

Finally, inserting the leading order saddle point values for v_I and u_I into (5.13) gives

$$S_{\rm gen}^{\rm island} = 2M - \frac{c}{24}(t_A - \sigma_A) + \dots \,, \tag{5.20}$$

as a function of the retarded time at the anchor curve. Although the time-dependent contribution is initially of order ϵ, compared to the leading order area term, it is important to keep in mind that this contribution grows with time, and eventually becomes comparable to the leading order result. This expression for the generalized entropy, which is valid when a time of order the scrambling time has passed after the first Hawking radiation emerges into the outside region beyond the anchor curve, is to be compared to the contribution from a no-island configuration that we now turn our attention to.

5.2.2 No-island configuration

In the absence of an island, the spacelike surface $\mathcal{S}_{AI}$ extends from the anchor curve to the semi-classical boundary at $\Omega = \Omega_{\rm crit}$. The gravitational coupling becomes strong at the semi-classical boundary and it is not a priori clear how to proceed. The validity of the semi-classical black hole solution indeed breaks down near the boundary but from a higher dimensional perspective this has a simple interpretation in terms of the area of the transverse two-sphere going to zero. We are primarily interested in the dependence of the bulk entropy (4.7) on asymptotic time and this will not be greatly affected by the detailed conditions imposed at the origin. This can be seen by adopting a simple prescription for the strong coupling region and then checking that a change in the prescription does not change the leading order result at late times on the anchor curve.

In the following we let the spacelike surface $\mathcal{S}_{AI}$ end at a fixed reference point (v_0, u_0) on the boundary curve for all anchor points. In other words, we will simply ignore any adjustment of the endpoint at the semi-classical boundary in response to changing the anchor point. We take the reference point to be in the $v < 1$ linear dilaton region inside the matter shock wave, i.e. with $u_0 = -\epsilon/v_0$ for some $v_0 < 1$. The generalized entropy is given by

$$\begin{aligned} S_{\rm gen}^{\rm no\text{-}island} &= \frac{c}{12} \log \left[\left(\log \frac{v_A}{v_0} \log \frac{u_A}{u_0} \right)^2 \frac{v_A u_A}{(1 - v_A(1 + u_A))} \right] \\ &= \frac{c}{12}(t_A - \sigma_A) + \dots \,, \end{aligned} \tag{5.21}$$

up to logarithmic correction terms. In particular, all dependence on v_0 is contained in the sub-leading terms that we have dropped. This expression is valid for retarded time of order the scrambling time and onwards.

JHEP05(2020)091

Comparing the island and no-island results in (5.20) and (5.21), respectively, shows that for retarded time $t_A - \sigma_A > \frac{1}{3\epsilon}$ the generalized entropy will be dominated by the island configuration. The Page time for a dynamical RST black hole is one third of the black hole lifetime,

$$t_{\text{Page}} = \frac{1}{3\epsilon} = \frac{16M}{c} \,. \tag{5.22}$$

The corresponding Page curve is drawn in figure 1.

6 Discussion

By assuming a QRT formula we have explicitly obtained Page curves for semi-classical black holes in asymptotically flat spacetime in a two-dimensional dilaton gravity model. This includes both an eternal black hole, supported by an incoming energy flux matching the outgoing Hawking flux, and a black hole formed by gravitational collapse that gradually evaporates. In both cases, the generalized entropy is minimised at early times by the bulk von Neumann entropy of the two-dimensional matter CFT but at a Page time the system crosses over to a configuration where the minimum generalized entropy includes a non-trivial Ryu-Takayanagi area term associated with an island near the black hole horizon. This is consistent with earlier results obtained for AdS_2 black holes but our computation offers a rather clean physical picture and offers evidence that the QRT prescription applies beyond asymptotically AdS spacetimes.

For the eternal black hole we confirm the appearance of an island outside the horizon, whereas for an evaporating black hole the island is always inside the horizon. Towards the end of evaporation the island appears to melt together with both singularity and horizon at the black hole endpoint. Our semi-classical calculation is, however, only reliable as long as the black hole mass remains large compared to the scale set by the central charge of the matter CFT.

For the evaporating black hole the Page time comes out to be 1/3 of the black hole lifetime. This fits with the following simple reasoning [33]. Radiating into a cold surrounding space is an irreversible process and the entropy of the radiation grows at twice the evaporation rate of the black hole. This can be seen more explicitly by considering the relation between internal energy and entropy of the gas emitted by a black body in two dimensions,

$$S = \frac{2U}{T} \,. \tag{6.1}$$

We can compute the ratio of the entropy of the gas that the black hole emits compared to the entropy lost by the black hole. The black hole satisfies the first law $\Delta S_{\text{BH}} = -\Delta M/T$ where ΔM is the mass lost by the black hole in a given time interval. The entropy increase of the gas radiated is $\Delta S_{\text{gas}} = 2\Delta U/T$ where ΔU is the energy of the emitted gas, which must equal ΔM, and we get $\Delta S_{\text{gas}} = -2\Delta S_{\text{BH}}$.

Our results rely on two dimensional conformal methods. On a technical level this is reflected in the fact that we can conveniently obtain the bulk contribution to the entropy via a AdS_3/CFT_2 computation. It is furthermore convenient that in two dimensions there are no grey body factors. On a more fundamental level, we are able to account for radiation

JHEP05(2020)091

and back-reaction in the RST model by working in a conformal gauge and explicitly using the conformal anomaly in two dimensions. Although there is no known analogue of the RST model in higher dimensions, our result suggests that a QRT like prescription may also work for other non-AdS black holes, especially ones that can be embedded into a boundary theory with a higher-dimensional AdS dual.

Acknowledgments

We thank Valentina Giangreco M. Puletti for stimulating discussions. LT would like to thank the Kavli Institute for Theoretical Physics at UC Santa Barbara for hospitality during the completion of this work. This research was supported in part by the Icelandic Research Fund under grants 185371-051 and 195970-051, by the University of Iceland Research Fund, and by the National Science Foundation under grant NSF PHY-1748958. FFG is a Postdoctoral Fellow of the Research Foundation — Flanders (FWO). FFG is also supported by the KU Leuven C1 grant ZKD1118 C16/16/005.

References

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

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

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

[4] G. Penington, *Entanglement Wedge Reconstruction and the Information Paradox*, arXiv:1905.08255 [INSPIRE].

[5] 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] [INSPIRE].

[6] J.V. Rocha, *Evaporation of large black holes in AdS: Coupling to the evaporon*, *JHEP* **08** (2008) 075 [arXiv:0804.0055] [INSPIRE].

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

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

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

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

JHEP05(2020)091

[11] 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] [INSPIRE].

[12] A. Almheiri, R. Mahajan and J. Maldacena, *Islands outside the horizon*, arXiv:1910.11077 [INSPIRE].

[13] A. Almheiri, R. Mahajan and J.E. Santos, *Entanglement islands in higher dimensions*, arXiv:1911.09666 [INSPIRE].

[14] C.G. Callan Jr., S.B. Giddings, J.A. Harvey and A. Strominger, *Evanescent black holes*, *Phys. Rev.* **D 45** (1992) R1005 [hep-th/9111056] [INSPIRE].

[15] J.G. Russo, L. Susskind and L. Thorlacius, *The Endpoint of Hawking radiation*, *Phys. Rev.* **D 46** (1992) 3444 [hep-th/9206070] [INSPIRE].

[16] 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] [INSPIRE].

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

[18] C. Akers, N. Engelhardt and D. Harlow, *Simple holographic models of black hole evaporation*, arXiv:1910.00972 [INSPIRE].

[19] V. Balasubramanian, A. Kar, O. Parrikar, G. Sárosi and T. Ugajin, *Geometric secret sharing in a model of Hawking radiation*, arXiv:2003.05448 [INSPIRE].

[20] A. Bhattacharya, *Multipartite Purification, Multiboundary Wormholes and Islands in AdS_3/CFT_2*, arXiv:2003.11870 [INSPIRE].

[21] M. Rozali, J. Sully, M. Van Raamsdonk, C. Waddell and D. Wakeham, *Information radiation in BCFT models of black holes*, *JHEP* **05** (2020) 004 [arXiv:1910.12836] [INSPIRE].

[22] H.Z. Chen, Z. Fisher, J. Hernandez, R.C. Myers and S.-M. Ruan, *Information Flow in Black Hole Evaporation*, *JHEP* **03** (2020) 152 [arXiv:1911.03402] [INSPIRE].

[23] J. Pollack, M. Rozali, J. Sully and D. Wakeham, *Eigenstate Thermalization and Disorder Averaging in Gravity*, arXiv:2002.02971 [INSPIRE].

[24] H. Liu and S. Vardhan, *A dynamical mechanism for the Page curve from quantum chaos*, arXiv:2002.05734 [INSPIRE].

[25] C.G. Callan Jr., J.A. Harvey and A. Strominger, *Supersymmetric string solitons*, hep-th/9112030 [INSPIRE].

[26] O. Aharony, M. Berkooz, D. Kutasov and N. Seiberg, *Linear dilatons, NS five-branes and holography*, *JHEP* **10** (1998) 004 [hep-th/9808149] [INSPIRE].

[27] J.M. Maldacena and A. Strominger, *Semiclassical decay of near extremal five-branes*, *JHEP* **12** (1997) 008 [hep-th/9710014] [INSPIRE].

[28] G. Mandal, A.M. Sengupta and S.R. Wadia, *Classical solutions of two-dimensional string theory*, *Mod. Phys. Lett.* **A 6** (1991) 1685 [INSPIRE].

[29] E. Witten, *On string theory and black holes*, *Phys. Rev.* **D 44** (1991) 314 [INSPIRE].

[30] J. Polchinski, *String theory. Vol. 1: An introduction to the bosonic string*, Cambridge Monographs on Mathematical Physics, Cambridge University Press (1998) [INSPIRE].

[31] S.M. Christensen and S.A. Fulling, *Trace Anomalies and the Hawking Effect*, *Phys. Rev.* **D 15** (1977) 2088 [INSPIRE].

JHEP05(2020)091

[32] T.M. Fiola, J. Preskill, A. Strominger and S.P. Trivedi, *Black hole thermodynamics and information loss in two-dimensions*, *Phys. Rev.* **D 50** (1994) 3987 [hep-th/9403137] [INSPIRE].

[33] W.H. Zurek, *Entropy Evaporated by a Black Hole*, *Phys. Rev. Lett.* **49** (1982) 1683 [INSPIRE].

JHEP05(2020)091