

## Holographic complexity in Vaidya spacetimes. Part II

---

Shira Chapman,<sup>a</sup> Hugo Marrochio<sup>a,b</sup> and Robert C. Myers<sup>a</sup>

<sup>a</sup>*Perimeter Institute for Theoretical Physics,  
31 Caroline St N, Waterloo, ON N2L 2Y5, Canada*

<sup>b</sup>*Department of Physics & Astronomy, University of Waterloo,  
200 University Avenue West, Waterloo, ON N2L 3G1, Canada*

*E-mail:* [schapman@perimeterinstitute.ca](mailto:schapman@perimeterinstitute.ca),  
[hmarrochio@perimeterinstitute.ca](mailto:hmarrochio@perimeterinstitute.ca), [rmyers@perimeterinstitute.ca](mailto:rmyers@perimeterinstitute.ca)

**ABSTRACT:** In this second part of the study initiated in [1], we investigate holographic complexity for eternal black hole backgrounds perturbed by shock waves, with both the complexity=action (CA) and complexity=volume (CV) proposals. In particular, we consider Vaidya geometries describing a thin shell of null fluid with arbitrary energy falling in from one of the boundaries of a two-sided AdS-Schwarzschild spacetime. We demonstrate how known properties of complexity, such as the switchback effect for light shocks, as well as analogous properties for heavy ones, are imprinted in the complexity of formation and in the full time evolution of complexity. Following our discussion in [1], we find that in order to obtain the expected properties of the complexity, the inclusion of a particular counterterm on the null boundaries of the Wheeler-DeWitt patch is required for the CA proposal.

**KEYWORDS:** AdS-CFT Correspondence, Gauge-gravity correspondence, Black Holes

**ARXIV EPRINT:** [1805.07262](https://arxiv.org/abs/1805.07262)

---

**Contents**

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Background geometry</b>	<b>3</b>
<b>3</b>	<b>Complexity = action</b>	<b>8</b>
3.1	Time evolution	9
3.1.1	Time dependence of complexity	15
3.1.2	A case study: BTZ black holes	19
3.2	Complexity of formation	24
3.3	‘Complexity’ without the counterterm	31
3.3.1	Time evolution	31
3.3.2	Complexity of formation	34
<b>4</b>	<b>Complexity = volume</b>	<b>36</b>
4.1	Time evolution	40
4.2	Complexity of formation	43
4.3	Numerical analysis	45
4.3.1	$d = 2$	45
4.3.2	$d = 4$	48
<b>5</b>	<b>Discussion</b>	<b>49</b>
<b>A</b>	<b>Counterterm for the null boundaries</b>	<b>70</b>
<b>B</b>	<b>Complexity = action in higher dimensions</b>	<b>75</b>
<b>C</b>	<b>Details for complexity = volume</b>	<b>88</b>

---

**1 Introduction**

Using perspectives and techniques from quantum information has brought many surprising insights into properties of the AdS/CFT correspondence [2, 3] in recent years. The best studied example has been the role of holographic entanglement entropy [4–7] in understanding the emergence of a semi-classical bulk spacetime [8–10]. A new suggestion in this research program is that the quantum complexity of states in the boundary theory is encoded in a set of gravitational observables, which probe spacetime properties which are inaccessible from the perspective of holographic entanglement entropy [11].

In the holographic setting, circuit complexity provides a measure of how difficult it is to construct a particular target state from a (simple) reference state by applying a set of

(simple) elementary gates, for a review see, e.g., [12, 13]. There are two complementary proposals for the holographic description of the complexity of the boundary states: the complexity=volume (CV) [14, 15] and the complexity=action (CA) [16, 17] conjectures. The CV conjecture states that the complexity is dual to the volume of an extremal codimension-one bulk surface anchored at the time slice  $\Sigma$  in the boundary on which the state is defined,

$$\mathcal{C}_V(\Sigma) = \max_{\Sigma=\partial\mathcal{B}} \left[ \frac{\mathcal{V}(\mathcal{B})}{G_N L} \right], \tag{1.1}$$

with  $\mathcal{B}$  corresponding to the bulk surface of interest, and  $G_N$  and  $L$  denoting Newton’s constant and the AdS curvature scale, respectively, in the gravitational theory. On the other hand, the CA proposal states that the complexity is given by evaluating the gravitational action on a region of spacetime, known as the Wheeler-DeWitt (WDW) patch, which can be regarded as the causal development of a space-like bulk surface, i.e., a Cauchy surface, anchored on the boundary time slice  $\Sigma$ . The CA proposal then suggests

$$\mathcal{C}_A(\Sigma) = \frac{I_{\text{WDW}}}{\pi \hbar}. \tag{1.2}$$

These conjectures have stimulated a wide variety of recent research efforts investigating the properties of both holographic complexity and circuit complexity in quantum field theories, e.g., [18–43]. In [1], we investigated the CA and CV proposals for Vaidya spacetimes [44–46] describing a thin shell of null fluid collapsing into the AdS vacuum to form a (one-sided) black hole. A surprising result we found was that the standard definition of the WDW action (e.g., [21–23]) was inappropriate for these dynamical spacetimes, in that eq. (1.2) did not reproduce the desired properties of complexity. However, the situation was rectified by adding an additional surface term [21] on the null boundaries, which also ensured that the action was independent of the parametrization of the null generators for these boundaries.

The present paper is a direct continuation of [1] where we examine the holographic complexity for Vaidya geometries in which a thin null shell collapses into an eternal black hole — see figure 1. Such shock wave geometries have already been extensively studied in the context of holographic complexity, e.g., [15, 17–19], however, these studies focused on the case where the energy in the shock was small. Using the formalism developed in [1], we will not need to restrict our attention to this regime of light shocks here. Further, we will investigate the full time evolution of the holographic complexity, i.e., including the transient regime, and this will allow us to identify several critical times that arise as the WDW patch (or the maximal volume surface) evolves forward in the background geometry. As well as the time evolution, we will investigate the complexity of formation in these shock wave geometries. As we found in [1], we will argue that the inclusion of the null surface counterterm is crucial in these dynamical spacetimes in order for the CA proposal (1.2) to properly produce the expected properties of complexity, such as the ‘switchback’ effect [15, 19, 47]. In the CV calculations, we develop a geometric understanding of certain limits of the rate of change in complexity and relate them to having the extremal surfaces wrapping and/or unwrapping certain critical surfaces behind the past and/or future horizons of the black holes. We summarize the main results in some detail at the beginning of section 5.

The rest of the paper is organized as follows: in section 2, we review the Vaidya background geometries in the context of two-sided black holes. We restrict our attention to thin shells of null fluid for which the action vanishes when the thickness shrinks to zero, as we showed in [1]. Next, we investigate the holographic complexity in these background geometries using the CA proposal in 3. We evaluate the time evolution and complexity of formation in the presence of light and heavy shock waves, and also examine the consequences of not including the null surface counterterm. In section 4, we evaluate the time evolution and complexity of formation using the CV proposal, for both light and heavy shock waves. We review our main results and discuss their physical implications in section 5, where we also present some future directions. We leave some technical details to the appendices: in appendix A, we evaluate the counterterm contributions to the WDW patch and review its implications for the UV structure of complexity. In appendix B, we present some numerical results for the holographic complexity in higher dimensions using the CA conjecture. Finally, appendix C presents some of the relevant details for various intermediate results used in applying the CV conjecture in section 4.

## 2 Background geometry

Recall that the (unperturbed) eternal black hole geometry is dual to a thermofield double (TFD) state [48], which is a pure state in which the degrees of freedom of two identical copies of the boundary CFT are entangled,

$$|TFD\rangle \equiv Z^{-1/2} \sum_{n=0}^{\infty} e^{-\frac{1}{2}\beta E_n} |E_n\rangle_L |E_n\rangle_R, \quad (2.1)$$

where the two copies are denoted as left (L) and right (R), in analogy to the left and right boundaries of the eternal geometry. Tracing out either the left or right CFT leaves a thermal density matrix with inverse temperature  $\beta$ . While this density matrix is invariant under time translations, we can time evolve the two sets of degrees of freedom in the TFD state independently to produce

$$\begin{aligned} |TFD(t_L, t_R)\rangle &= U_L(t_L) U_R(t_R) |TFD\rangle, \\ &= Z^{-1/2} \sum_{n=0}^{\infty} e^{-\frac{1}{2}\beta E_n - iE_n(t_L + t_R)} |E_n\rangle_L |E_n\rangle_R, \end{aligned} \quad (2.2)$$

where  $U_{L,R}$  are the usual time evolution operators for the corresponding CFTs, i.e.,  $U_L(t_L) = e^{-iH_L t_L}$  and  $U_R(t_R) = e^{-iH_R t_R}$ . One immediate observation is that the state is invariant when we shift

$$t_L \rightarrow t_L + \Delta t, \quad t_R \rightarrow t_R - \Delta t, \quad (2.3)$$

i.e., the TFD state (2.1) is invariant if we time evolve with the combined Hamiltonian  $H_L - H_R$ . Of course, this invariance is reflected in the ‘boost symmetry’ of the dual black hole geometry. As a result, the holographic complexity remains unchanged by the above shifts (2.3), i.e., it only depends on the combination  $t_L + t_R$  (e.g., see [15, 17, 23]).

In the following, we study Vaidya geometries describing a thin shell of null fluid (or shock wave) injected into an eternal black hole background. Following [47, 49], these Vaidya geometries describe<sup>1</sup> a perturbation of the TFD state (2.1),

$$|TFD\rangle_{\text{pert}} = \mathcal{O}_R(-t_w) |TFD\rangle = U_R(t_w) \mathcal{O}_R U_R^\dagger(t_w) |TFD\rangle, \quad (2.4)$$

where  $\mathcal{O}_R(-t_w)$  is operator inserted in the right CFT at a time  $-t_w$ .<sup>2</sup> In the second expression, we are describing this precursor as  $\mathcal{O}_R(-t_w) = U_R(t_w) \mathcal{O}_R U_R^\dagger(t_w)$ , i.e.,  $U_R^\dagger(t_w)$  evolves the right degrees of freedom backwards by a time  $t_w$ ,  $\mathcal{O}_R$  is inserted and then the right CFT is evolved forward by  $t_w$ . In the following, we will use the complexity of formation [22] (in the Vaidya geometry) to evaluate the complexity of the precursor, i.e., to compare the complexities of  $|TFD\rangle_{\text{pert}}$  and  $|TFD\rangle$ . The nontrivial cancellations in the complexity of the precursor are connected to the switchback effect [15, 19, 47].

We will also examine the complexity of the time evolved state

$$\begin{aligned} |TFD(t_L, t_R)\rangle_{\text{pert}} &= U_L(t_L) U_R(t_R) |TFD\rangle_{\text{pert}} \\ &= U_L(t_L) U_R(t_R + t_w) \mathcal{O}_R U_R^\dagger(t_w) |TFD\rangle \\ &= U_R(t_R + t_w) \mathcal{O}_R U_R(t_L - t_w) |TFD\rangle, \end{aligned} \quad (2.5)$$

where in the last line, we use the boost symmetry of the TFD state, i.e.,  $U_L(t_L) |TFD\rangle = U_R(t_R) |TFD\rangle$ , and that  $U_L$  commutes with all operators in the right CFT [15]. In this case, inserting  $\mathcal{O}_R$  at a fixed time  $-t_w$  breaks the shift symmetry (2.3). However, from the above expression, it is clear that if we combine the previous translations of the left and right times with a shift the insertion time,

$$t_w \rightarrow t_w + \Delta t, \quad (2.6)$$

then the time-evolved state in eq. (2.5) is invariant. We will refer to the combination of eqs. (2.3) and (2.6) together as the *time-shift symmetry* of the problem. Of course, this will also produce a symmetry for the holographic complexity and as a result, we will find that the holographic complexity only depends on two combinations of the boundary times,  $t_R + t_w$  and  $t_L - t_w$ , which appear in eq. (2.5).

Let us now turn to the dual geometry in the bulk. As noted above, we consider the AdS-Vaidya spacetimes [46], sourced by the collapse of a spherically symmetric shell of null fluid,

$$\begin{aligned} ds^2 &= -F(r, v)dv^2 + 2drdv + r^2 d\Sigma_{k, d-1}^2, \\ \text{with } F(r, v) &= \frac{r^2}{L^2} + k - \frac{f_p(v)}{r^{d-2}}, \end{aligned} \quad (2.7)$$

---

<sup>1</sup>Our geometries are more properly interpreted in terms of a thermal quench, e.g., [68, 69], where some boundary coupling is rapidly varied at  $t_R = -t_w$ . Instead, eq. (2.4) corresponds to an excited state in which the excitation becomes coherent at  $t_R = -t_w$  (but with no variations of the couplings). The corresponding bulk geometry involves a null shell which emerges from the white hole singularity and reflects off of the right asymptotic boundary at  $t_R = -t_w$  to become a collapsing shell, e.g., see [15, 47, 49]. Since our evaluations of the holographic complexity always involve  $t_R > -t_w$ , our results would be the same for either geometry.

<sup>2</sup>The details of the operator will not be important for our analysis, however, for the special case of  $d = 2$ , [50] provides a detailed description of the dual of the Vaidya-AdS<sub>3</sub> geometry.

where  $d$  is the dimension of the boundary spacetime and  $k$  denotes the geometry of the horizon — see [1] for further details.<sup>3</sup> In particular, we consider the profile,

$$f_p(v) = \omega_1^{d-2} (1 - \mathcal{H}(v - v_s)) + \omega_2^{d-2} \mathcal{H}(v - v_s), \quad (2.8)$$

where  $\mathcal{H}(v)$  is the Heaviside step function.<sup>4</sup> With this profile, the Vaidya geometry describes the collapse of an infinitely thin shell of null fluid, which raises the mass of the black hole from  $M_1$  to  $M_2$  where<sup>5</sup>

$$M_i = \frac{(d-1)\Omega_{k,d-1}}{16\pi G_N} \omega_i^{d-2}. \quad (2.9)$$

Further, in eq. (2.8), we implicitly consider the shock wave as coming from the right boundary at some early time  $v_s = -t_w$  (with  $t_w > 0$ ), in accord with our description of the boundary state (2.4). Hence we have

$$v_R < -t_w : \quad F(r, v) = f_1(r) = \frac{r^2}{L^2} + k - \frac{\omega_1^{d-2}}{r^{d-2}}, \quad (2.10)$$

$$v_R > -t_w : \quad F(r, v) = f_2(r) = \frac{r^2}{L^2} + k - \frac{\omega_2^{d-2}}{r^{d-2}}. \quad (2.11)$$

where the coordinates  $r$  and  $v_R$  cover the right exterior region and the future black hole interior, labeled I and II in figure 1. There is a corresponding set of coordinates  $r$  and  $v_L$  covering the left exterior region and the past white hole interior, labeled III and IV in the figure. The shock wave does not enter either of the latter regions and so we have

$$\text{for all } v_L : \quad F(r, v) = f_1(r) = \frac{r^2}{L^2} + k - \frac{\omega_1^{d-2}}{r^{d-2}}. \quad (2.12)$$

On either side of the shock wave, the geometry (2.7) corresponds to that of a (static) AdS black hole, whose horizon radius  $r_{h,i}$  determined by<sup>6</sup>

$$\omega_i^{d-2} = r_{h,i}^{d-2} \left( \frac{r_{h,i}^2}{L^2} + k \right). \quad (2.13)$$

As noted above, in each of these regions, the mass is given by eq. (2.9) and further, the temperature and entropy become

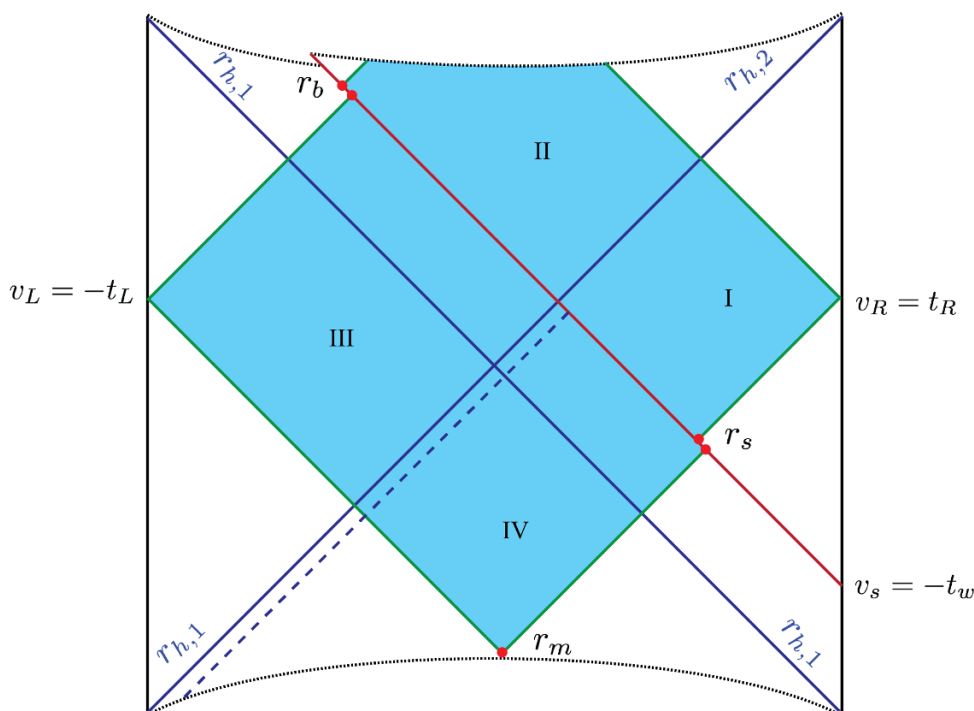
$$T_i = \frac{1}{4\pi} \left. \frac{\partial f}{\partial r} \right|_{r=r_{h,i}} = \frac{1}{4\pi r_{h,i}} \left( d \frac{r_{h,i}^2}{L^2} + (d-2)k \right), \quad S_i = \frac{\Omega_{k,d-1}}{4G_N} r_{h,i}^{d-1}. \quad (2.14)$$

<sup>3</sup>The parameter  $k$  can be  $\{+1, 0, -1\}$ , corresponding to spherical, planar, and hyperbolic horizon geometries. We denote by  $\Omega_{k,d-1}$  the dimensionless volume, see for instance [22, 23]. Of course, for spherical geometries it becomes simply  $\Omega_{1,d-1} = 2\pi^{d/2}/\Gamma(d/2)$ , and for hyperbolic and planar black holes, it should be understood as a dimensionless infrared regulated quantity.

<sup>4</sup>It is a simple task to generalize the present discussion to BTZ black holes with  $d = 2$ , and we will treat this case separately in the following.

<sup>5</sup>Again,  $\Omega_{k,d-1}$  denotes the (dimensionless) volume of the spatial geometry — see footnote 3.

<sup>6</sup>Let us note here that for  $k = -1$ , we will only consider ‘large’ hyperbolic black holes with  $r_h > L$ , which have the casual structure illustrated in figure 1.



**Figure 1.** Penrose-like diagram for one shock wave on an eternal black hole geometry. At  $v_s = -t_w$  a thin shock is injected at the right boundary which raises the mass of the black hole from  $M_1$  to  $M_2$ . We identify three points in the geometry that depend on time,  $r_m$ , where the boundaries of the WDW patch cross behind the past horizon,  $r_s$ , where the boundary of the WDW patch crosses the collapsing shell in the right exterior, and  $r_b$  where the boundary of the WDW patch crosses the shock wave inside the future black hole.

We define the tortoise coordinates with respect to eqs. (2.10)–(2.12) as

$$\text{for all } v_L \text{ \& for } v_R < -t_w : \quad r_1^*(r) = - \int_r^\infty \frac{dr}{f_1(r)}, \quad (2.15)$$

$$v_R > -t_w : \quad r_2^*(r) = - \int_r^\infty \frac{dr}{f_2(r)}, \quad (2.16)$$

where again we chose the range of integration such that both expressions satisfy  $\lim_{r \rightarrow \infty} r_{1,2}^*(r) \rightarrow 0$ . Using the tortoise coordinates, we can define an outgoing null coordinate  $u$  and an auxiliary time coordinate  $t$  as

$$u_{1,2} \equiv v - 2r_{1,2}^*(r), \quad t_{1,2} \equiv v - r_{1,2}^*(r). \quad (2.17)$$

Again these coordinates are discontinuous across the shell because  $f(r)$  jumps from eq. (2.10) to eq. (2.11) at  $v_R = -t_w$ . In analogy to the diagrams in [1], we represent the shock wave geometries with Penrose-like diagrams (e.g., see figure 1), which can be smoothly ruled with lines of constant  $u$  and  $v$ . As before, since the coordinate  $u$  is discontinuous, this introduces a(n unphysical) jump as the outgoing null rays cross the shock

wave. Of course, the spacetime is continuous along this surface and the outgoing null rays are smooth, as can be seen by regulating the thin shell to have a finite thickness — see section 2 in [1].

Before we proceed further, let us comment on synchronizing the times between the left and right boundaries. In principle, the left boundary time  $t_L$  is completely independent of the right boundary time  $t_R$  in the eternal black hole geometry. However, implicitly, they are synchronized by considering a geometric construction, e.g., where an extremal codimension-one surface that runs from one boundary to the other through the bifurcation surface. This surface connects the time slice  $t_L$  on the left boundary to the time slice  $t_R = -t_L$  on the right boundary. The minus sign arises here because of our convention that the boundary times increase upwards on both boundaries in figure 1. Of course, such extremal surfaces have a more interesting profile in the shock wave geometry, as we will discuss extensively in section 4. Hence one might worry that the two boundary times cannot be synchronized in a natural way. However, we observe that the above geometric construction is unaffected for times  $t_L > t_w$ , for which the extremal surface will reach the right boundary at times  $t_R < -t_w$ . That is, for late (early) times on the left (right) boundary, the desired extremal surfaces do not meet the shock wave and remain entirely within the portion of the spacetime where  $F(v, r) = f_1(r)$ . Once the boundary times are synchronized in these regions, this synchronization is straightforwardly extended to the entire boundaries. Implicitly, this is how we match the left and right boundary times in the following.

It will be useful to define some dimensionless quantities in order to express the evolution of complexity, as well as the complexity of formation, in the following. We define,<sup>7</sup>

$$w \equiv \frac{r_{h,2}}{r_{h,1}}, \quad z \equiv \frac{L}{r_{h,2}}, \quad x \equiv \frac{r}{r_{h,2}}, \quad (2.18)$$

which for positive-energy shock waves, yields  $w > 1$ . Also note that for planar (and BTZ) black holes,  $w$  is proportional to the ratio of temperatures, i.e.,  $w = T_2/T_1$ , for  $k = 0$  (or  $d = 2$ ) which can be seen by using eq. (2.14). The ratio between the masses and the entropies reads

$$\frac{M_2}{M_1} = w^d \frac{(1 + kz^2)}{(1 + kz^2w^2)} \quad \text{and} \quad \frac{S_2}{S_1} = w^{d-1}. \quad (2.19)$$

It is also useful to rescale the blackening factor such that

$$f_2(r, L) = \frac{1}{z^2} \tilde{f}(x, z), \quad f_1(r, L) = \frac{1}{w^2 z^2} \tilde{f}(wx, wz), \quad (2.20)$$

where

$$\tilde{f}(x, z) \equiv x^2 + kz^2 - \frac{1}{x^{d-2}} (1 + kz^2). \quad (2.21)$$

Note that  $\tilde{f}(x, z)$  is not a function of  $z$  for planar holes (i.e.,  $k = 0$ ). Finally, if a physical quantity depends on  $z$ , we can use eq. (2.14) to express  $z$  as a function of  $LT_2$ ,

$$z = \frac{d}{\sqrt{4\pi^2(LT_2)^2 - (d-2)dk + 2\pi LT_2}}. \quad (2.22)$$

---

<sup>7</sup>Note that here,  $z$  and  $x$  are defined as the ratio of the AdS scale  $L$  or the radius  $r$  with the final horizon size  $r_{h,2}$ . Using the ratio  $w$  one can easily construct quantities normalized by  $r_{h,1}$  instead.



### 3 Complexity = action

In order to evaluate the holographic complexity using the complexity=action proposal (1.2), we begin by writing the total action as

$$I = I_{\text{grav}} + I_{\text{ct}} + I_{\text{fluid}}, \quad (3.1)$$

where  $I_{\text{fluid}}$  is the null fluid action, as constructed in section 2 of [1]. There we showed that  $I_{\text{fluid}}$  vanishes on shell, and so the imprint of the shock wave in the CA calculations comes only from its backreaction on the metric. Further, we showed in [1] that as the width of the null shell shrinks to zero, the action of the spacetime region occupied by the shell itself vanishes. Hence for an infinitely thin shell, as introduced in eq. (2.8), the action can be evaluated by separately evaluating the action on the portion of the WDW patch above the shell and that below the shell.

The gravitational action can be written as [21]

$$\begin{aligned} I_{\text{grav}} = & \frac{1}{16\pi G_N} \int_{\mathcal{M}} d^{d+1}x \sqrt{-g} \left( \mathcal{R} + \frac{d(d-1)}{L^2} \right) \\ & + \frac{1}{8\pi G_N} \int_{\mathcal{B}} d^d x \sqrt{|h|} K + \frac{1}{8\pi G_N} \int_{\Sigma} d^{d-1}x \sqrt{\sigma} \eta \\ & + \frac{1}{8\pi G_N} \int_{\mathcal{B}'} d\lambda d^{d-1}\theta \sqrt{\gamma} \kappa + \frac{1}{8\pi G_N} \int_{\Sigma'} d^{d-1}x \sqrt{\sigma} a. \end{aligned} \quad (3.2)$$

The bulk integral in the manifold  $\mathcal{M}$  is proportional to the Einstein-Hilbert action with the Ricci scalar  $\mathcal{R}$  and the negative cosmological constant  $\Lambda = -d(d-1)/(2L^2)$ . In addition, the integral along the boundary  $\mathcal{B}$  is the usual Gibbons-Hawking-York (GHY) boundary term for smooth spacelike and timelike boundaries, proportional to the trace of the extrinsic curvature  $K$ . When the boundary includes null hypersurfaces, there is a similar surface term which integrates  $\kappa$  along these portions of the boundary  $\mathcal{B}'$ . The latter indicates how the null coordinate  $\lambda$  along the boundary deviates from affine parametrization. Finally, there are the codimension-two joint terms: if the intersection involves timelike or spacelike boundaries, the Hayward joint terms [51, 52] at  $\Sigma$  are defined by the “boost angle”  $\eta$  between the normal vectors of these hypersurfaces. If either of the boundaries forming the joint is null, there is an equivalent contribution at  $\Sigma'$  given in terms of an analogous “angle”  $a$ , which depends on the null normals involved, as constructed in [21].

In eq. (3.1), we have also included an additional surface term for the null boundaries [21],

$$I_{\text{ct}} = \frac{1}{8\pi G_N} \int_{\mathcal{B}'} d\lambda d^{d-1}\theta \sqrt{\gamma} \Theta \log(\ell_{\text{ct}} \Theta), \quad (3.3)$$

where  $\Theta = \partial_\lambda \log \sqrt{\gamma}$  is the expansion scalar, and  $\ell_{\text{ct}}$  is an arbitrary constant length scale. This counterterm is not needed to produce a well defined variational principle for the gravitational action, i.e.,  $I_{\text{ct}}$  only depends on intrinsic boundary data. Instead it was introduced in [21] to ensure that the action is reparametrization invariant. In [1], we found that including the counterterm was essential if the WDW action was to reproduce certain properties expected of the complexity in dynamical spacetimes and hence it is included in (most of) our CA calculations here.

However, we will also expand our arguments indicating that the counterterm (3.3) is an essential part of the gravitational action by considering our results after we remove the counterterm contributions. In this case, we must come to grips with the various ambiguities arising from the surface and joint terms associated with the null boundaries, e.g., see discussion in [21]. We follow the prescription in [21] where we set  $\kappa = 0$  by choosing an affine parametrization for the null normals. Further, we fix the overall normalization of these null vectors by their inner product with the asymptotic timelike Killing vector at the boundary,  $\hat{t} = \partial_t$ , i.e., we set  $\hat{t} \cdot k = \pm\alpha$ .

In section 3.1, we evaluate the time evolution (including the null counterterm) for the Vaidya spacetimes described in section 2, and then we examine their complexity of formation in section 3.2. In section 3.3, we (mostly) focus on the BTZ black hole (i.e.,  $d = 2$ ) and demonstrate that the CA calculations without the inclusion of the counterterm fail to produce the expected behaviour of holographic complexity. In appendix A, we discuss some further details on the influence of the null counterterm in Vaidya spacetimes. In appendix B we discuss the complexity evolution for higher ( $d > 2$ ) dimensional black holes, focusing mostly on  $d = 4$ .

### 3.1 Time evolution

Consider the shock wave spacetime represented in figure 1, with the Penrose-like diagram describing the geometry in eqs. (2.10)–(2.12). The null shell is injected at the right boundary at  $v_s = -t_w$  (with  $t_w > 0$ ), raising the mass of the black hole from  $M_1$  to  $M_2$ . We can study the time evolution of the holographic complexity in many different ways. However, for simplicity, we will focus on a symmetric time evolution with  $t_L = t/2 = t_R$  starting at  $t_L = t_R = 0$ , in analogy to the analysis in [23].

We identify the three positions which are important in defining the WDW patch, depending on the time:  $r_b$  is where the boundary of the WDW patch originating from the left boundary meets the shock wave inside the future black hole;  $r_s$  is the surface where the WDW boundary in the right exterior meets the shock wave; and  $r_m$  is where the past null boundary segments of the WDW patch meet inside the white hole region — see figure 1. Of course, depending on the parameters of the problem,  $r_b$  and  $r_m$  could be behind the singularities. In particular, if  $r_m < 0$ , the WDW patch has a spacelike boundary segment running along the past singularity. In this section, we carefully evaluate all these possibilities in the shock wave black hole geometry, and show how the critical times where  $r_b$  and  $r_m$  cross  $r = 0$  produce transitions between different behaviours of the holographic complexity.

We calculate the bulk action given by eq. (3.2) by using the same prescription discussed in [1], i.e., implicitly we evaluate the total gravitational action as the sum of the action evaluated on the regions comprising the WDW patch to the future and the past of the shock wave. We start here by identifying the three positions introduced above (i.e.,  $r_b$ ,  $r_s$  and  $r_m$ ) as functions of the times<sup>8</sup>  $v_L = -t_L$  and  $v_R = t_R$  at which the WDW patch is

---

<sup>8</sup>The minus sign arises for  $t_L$  because our convention is that the left boundary time increases as we run upwards in figure 1, while  $v_L$  increases as we move down diagonally towards the bottom left corner.

anchored on the left and right boundaries,

$$\begin{aligned} t_L - t_w &= 2r_1^*(r_b), \\ t_R + t_w &= -2r_2^*(r_s), \\ t_L - t_w &= 2r_1^*(r_s) - 2r_1^*(r_m). \end{aligned} \tag{3.4}$$

Recall that  $t_w$  is defined to be positive. Given eqs. (2.15) and (2.16), the time evolution of these position is relatively simple. For example, if  $t_L$  is held fixed, eq. (3.4) implies that

$$\frac{dr_b}{dt_R} = 0, \quad \frac{dr_s}{dt_R} = -\frac{f_2(r_s)}{2}, \quad \frac{dr_m}{dt_R} = -\frac{f_1(r_m)}{2} \frac{f_2(r_s)}{f_1(r_s)}. \tag{3.5}$$

Similarly, when  $t_R$  is held constant, the evolution with the left boundary time becomes

$$\frac{dr_b}{dt_L} = \frac{f_1(r_b)}{2}, \quad \frac{dr_s}{dt_L} = 0, \quad \frac{dr_m}{dt_L} = -\frac{f_1(r_m)}{2}. \tag{3.6}$$

Recall that we will be interested in  $t_L = t/2 = t_R$  in the following, and so when required we can combine the above results in the appropriate linear combination.

**Bulk contribution**

We start by evaluating the bulk action for the WDW patch represented in figure 1. As before, the Einstein-Hilbert contribution to the action is  $\mathcal{R} - 2\Lambda = -\frac{2d}{L^2}$  in  $d + 1$  bulk dimensions. The total bulk action reads

$$\begin{aligned} I_{\text{bulk}} &= \left( \frac{\Omega_{k,d-1}}{16\pi G_N} \right) \left( -\frac{2d}{L^2} \right) \left[ \int_{r_s}^{r_{\text{max}}} dr r^{d-1} (-2r_2^*(r)) + \int_{r_b}^{r_s} dr r^{d-1} (t_R + t_w) \right. \\ &\quad \times \int_{r_{h,1}}^{r_s} dr r^{d-1} (2r_1^*(r_s) - 2r_1^*(r)) + \int_{r_m}^{r_{h,1}} dr r^{d-1} (-t_L + t_w - 2r_1^*(r) + 2r_1^*(r_s)) + \\ &\quad + \int_{r_{h,1}}^{r_{\text{max}}} dr r^{d-1} (-2r_1^*(r)) + \int_{r_b}^{r_{h,1}} dr r^{d-1} (-t_w + t_L - 2r_1^*(r)) + \\ &\quad \left. + \int_0^{r_b} dr r^{d-1} (t_R + t_w - 2r_2^*(r) + 2r_2^*(r_b)) \right]. \end{aligned} \tag{3.7}$$

Now, we fix the left boundary time  $t_L$  and vary  $t_R$  in the right boundary, as in eq. (3.5). The time derivative of the bulk action with respect to the right boundary reads

$$\frac{dI_{\text{bulk}}}{dt_R} = -\frac{\Omega_{k,d-1}}{8\pi G_N L^2} \left[ r_s^d \left( 1 - \frac{f_2(r_s)}{f_1(r_s)} \right) + r_m^d \frac{f_2(r_s)}{f_1(r_s)} \right]. \tag{3.8}$$

In addition, we can write the time derivative with respect to the left boundary evolution  $t_L$  in eq. (3.6),

$$\frac{dI_{\text{bulk}}}{dt_L} = -\frac{\Omega_{k,d-1}}{8\pi G_N L^2} \left[ r_m^d - r_b^d \left( 1 - \frac{f_1(r_b)}{f_2(r_b)} \right) \right]. \tag{3.9}$$

With respect to a symmetric time evolution  $t_L = t_R = t/2$ , we sum the linear combination of eqs. (3.8) and (3.9),

$$\frac{dI_{\text{bulk}}}{dt} = -\frac{\Omega_{k,d-1}}{16\pi G_N L^2} \left[ r_m^d \left( 1 + \frac{f_2(r_s)}{f_1(r_s)} \right) - r_b^d \left( 1 - \frac{f_1(r_b)}{f_2(r_b)} \right) + r_s^d \left( 1 - \frac{f_2(r_s)}{f_1(r_s)} \right) \right]. \tag{3.10}$$

### Boundary surface contributions

We now turn our attention to the boundary surface contributions in the action in eq. (3.2). As suggested in [21], we choose the normals to the null boundaries to be affinely parametrized (before and after the shock wave — see discussion in section 2 of [1]). Therefore the parameter  $\kappa$  and the corresponding boundary term vanishes for all of these null boundary segments. There are, however, two possible boundary contributions to the action, namely, evaluating the Gibbons-Hawking-York (GHY) term on a spacelike (regulator) surface right before the future singularity, and also in the regime that  $r_m$  is behind the singularity, a similar contribution arises from the past singularity.<sup>9</sup>

We will denote the critical times at which  $r_m$  leaves the past singularity (i.e.,  $r_m$  becomes positive) as  $t_{L,c1}$  and  $t_{R,c1}$ . From eq. (3.4), we have

$$t_{R,c1} = -t_w - 2r_2^*(r_s), \quad t_{L,c1} = t_w + 2r_1^*(r_s) - 2r_1^*(0). \quad (3.11)$$

We can apply this result as follows: if we choose a value for  $t_{R,c1}$ , then the first equation determines a particular value of  $r_s$  and the second equation determines the value  $t_{L,c1}$  at which the WDW patch lifts off the past singularity. Similarly, if instead we choose a value for  $t_{L,c1}$ , we can apply the equations in the opposite order to determine the value of  $t_{R,c1}$  at which the WDW patch lifts off the past singularity. There is also a second critical time for the left boundary, which we will denote  $t_{L,c2}$  and it is the time at which the crossing point  $r_b$  touches the singularity, i.e.,

$$t_{L,c2} = t_w + 2r_1^*(0). \quad (3.12)$$

Of course, whether there are critical times in the range of time evolution that we are studying depends on how early and how energetic the shock wave was. However, once the latter parameters are chosen, one can determine with eqs. (3.11) and (3.12) whether there are critical times and find their respective values.

Now, we first investigate the GHY term at the future singularity. As usual e.g., [17, 22], we introduce a regulator surface at  $r = \varepsilon$  and after evaluating the GHY term on this surface, we take the limit  $\varepsilon \rightarrow 0$ . Since  $t_L > 0$  and  $v_s = -t_w < 0$ , there are two possibilities:  $t_L < t_{L,c2}$  for which the crossing point  $r_b$  arises in the black hole interior region; and  $t_L > t_{L,c2}$  for which the future null boundary of WDW patch from the left boundary reaches the singularity without crossing the shock wave. For  $t_L < t_{L,c2}$ , we have

$$I_{GHY}^{(f)} = -\frac{\Omega_{k,d-1}}{8\pi G_N} \left( \frac{-d\omega_2^{d-2}}{2} \right) (t_R + t_w + 2r_2^*(r_b) - 2r_2^*(0)), \quad (3.13)$$

which leads to

$$\frac{dI_{GHY}^{(f)}}{dt_R} = \frac{\Omega_{k,d-1}}{8\pi G_N} \frac{d\omega_2^{d-2}}{2} \quad \text{and} \quad \frac{dI_{GHY}^{(f)}}{dt_L} = \frac{\Omega_{k,d-1}}{8\pi G_N} \frac{d\omega_2^{d-2}}{2} \frac{f_1(r_b)}{f_2(r_b)}. \quad (3.14)$$

---

<sup>9</sup>We do not consider possible surface terms at the UV regulator surfaces because these contributions will be independent of time, e.g., see [23, 26].

In contrast, for  $t_L > t_{L,c2}$ , we have

$$I_{GHY}^{(f)} = -\frac{\Omega_{k,d-1}}{8\pi G_N} \left( \frac{-d\omega_2^{d-2}}{2} \right) (t_R + t_w) - \frac{\Omega_{k,d-1}}{8\pi G_N} \left( \frac{-d\omega_1^{d-2}}{2} \right) (-t_w + t_L - 2r_1^*(0)) , \quad (3.15)$$

which results in

$$\frac{dI_{GHY}^{(f)}}{dt_R} = \frac{\Omega_{k,d-1}}{8\pi G_N} \frac{d\omega_2^{d-2}}{2} \quad \text{and} \quad \frac{dI_{GHY}^{(f)}}{dt_L} = \frac{\Omega_{k,d-1}}{8\pi G_N} \frac{d\omega_1^{d-2}}{2} . \quad (3.16)$$

Now, the GHY contribution from the past singularity follows a similar analysis. Whenever  $t_R < t_{R,c1}$  or  $t_L < t_{L,c1}$ , the WDW patch intersects the past singularity and one finds the following GHY contribution

$$I_{GHY}^{(p)} = \frac{\Omega_{k,d-1}}{8\pi G_N} \left( \frac{d\omega_1^{d-2}}{2} \right) (-t_L + t_w + 2r_1^*(r_s) - 2r_1^*(0)) . \quad (3.17)$$

The time derivatives of this result then become

$$\frac{dI_{GHY}^{(p)}}{dt_R} = -\frac{\Omega_{k,d-1}}{8\pi G_N} \frac{d\omega_1^{d-2}}{2} \frac{f_2(r_s)}{f_1(r_s)} \quad \text{and} \quad \frac{dI_{GHY}^{(p)}}{dt_L} = -\frac{\Omega_{k,d-1}}{8\pi G_N} \frac{d\omega_1^{d-2}}{2} . \quad (3.18)$$

### Joint contributions

We now focus on the joint contributions to the action (3.2) evaluated on the WDW patch. In principle, such contributions arise where the null boundaries intersect the UV regulator surfaces near the asymptotic boundary. However, these contributions are time independent and so we ignore them in the following. Similarly, there are joint contributions where the null boundaries intersect the regulator surfaces  $r = \epsilon$  near the singularities but these vanish in the limit  $\epsilon \rightarrow 0$ . This leaves three possible different contributions coming from joints at  $r = r_b$ ,  $r_s$  and  $r_m$ , as shown in figure 1.<sup>10</sup> The joint contributions at  $r_s$  and  $r_b$  are analogous to the ones discussed in the one sided geometry in [1], while the contribution from  $r_m$  is similar to the joint action found in unperturbed eternal black holes [23].

We start by evaluating the sum of joint contributions where the past null boundary of the WDW patch crosses the shock wave, i.e., at  $r = r_s$ . The relevant null normals on the past boundary are

$$k_\mu^p dx^\mu = \begin{cases} \alpha \left( -dv + \frac{2}{f_2(r)} dr \right) & \text{for } r > r_s , \\ \tilde{\alpha} \left( -dv + \frac{2}{f_1(r)} dr \right) & \text{for } r < r_s . \end{cases} \quad (3.19)$$

Further we introduce the two normals along the collapsing shock wave,

$$\begin{aligned} v > -t_w : & \quad k_\mu^{s+} dx^\mu = -\beta dv , \\ v < -t_w : & \quad k_\mu^{s-} dx^\mu = \beta dv . \end{aligned} \quad (3.20)$$

---

<sup>10</sup>In the case where  $r_b < 0$ , there would be additional joints where the shock wave hits the future singularity, i.e., on the regulator surface  $r = \epsilon$  at  $v = -t_w$ . However, these again yield a vanishing contribution in the limit  $\epsilon \rightarrow 0$ .

The sum of the two joint contributions then reads

$$I_{\text{joint}}^{(\text{I})} = \frac{\Omega_{k,d-1} r_s^{d-1}}{8\pi G_N} \log \left( \frac{\alpha f_1(r_s)}{\tilde{\alpha} f_2(r_s)} \right). \quad (3.21)$$

We note that in eq. (3.19), the normalization constant  $\alpha$  was fixed with the usual asymptotic condition  $k^p \cdot \hat{t} = -\alpha$  [21]. However, to fix the normalization constant  $\tilde{\alpha}$  below the shell, we demand that the null boundary is affinely parametrized across the shock wave, i.e.,  $\kappa = 0$ , following [1]. The latter constraint imposes

$$\frac{\tilde{\alpha}}{\alpha} = \frac{f_1(r_s)}{f_2(r_s)}. \quad (3.22)$$

As a consequence, the corner contributions at  $r_s$  vanish, as was discussed for the one-sided collapse in [1], i.e.,

$$I_{\text{joint}}^{(\text{I})} = 0. \quad (3.23)$$

Next, we evaluate the sum of joint terms where the future null boundary of the WDW patch crosses the shock wave, i.e., at  $r = r_b$ . The (outward-directed) null normal to this future boundary is

$$k_{\mu}^f dx^{\mu} = \begin{cases} \alpha \left( -dv + \frac{2}{f_1(r)} dr \right) & \text{for } r > r_b, \\ \hat{\alpha} \left( -dv + \frac{2}{f_2(r)} dr \right) & \text{for } r < r_b. \end{cases} \quad (3.24)$$

Using the null normals along the shock wave in eq. (3.20), the total contribution to the action is

$$I_{\text{joint}}^{(\text{II})} = \frac{\Omega_{k,d-1} r_b^{d-1}}{8\pi G_N} \log \left( \frac{\alpha f_2(r_b)}{\hat{\alpha} f_1(r_b)} \right). \quad (3.25)$$

Once again, the condition of affine parametrization across the shock wave fixes the ratio between the normalization constants in eq. (3.24) with

$$\frac{\hat{\alpha}}{\alpha} = \frac{f_2(r_b)}{f_1(r_b)}. \quad (3.26)$$

Therefore, the joint contributions at  $r_b$  also vanishes ,

$$I_{\text{joint}}^{(\text{II})} = 0. \quad (3.27)$$

Finally, we turn to the possible contribution from the joint where the two past null boundaries of the WDW patch meet inside the white hole region, i.e., at  $r_m$ . This joint contribution is evaluated with  $k^p$  in eq. (3.19) on the right boundary (with  $r < r_s$ ), and

$$k_{\mu}^L dx^{\mu} = \alpha dv \quad (3.28)$$

for the normal to the left null boundary.<sup>11</sup> The resulting joint contribution then reads, with the affine parametrization condition (3.22),

$$I_{\text{joint}}^{(\text{III})} = -\frac{\Omega_{k,d-1} r_m^{d-1}}{8\pi G_N} \log \left( \frac{|f_1(r_m)|}{\alpha \tilde{\alpha}} \right) = -\frac{\Omega_{k,d-1} r_m^{d-1}}{8\pi G_N} \log \left( \frac{|f_1(r_m)| f_2(r_s)}{\alpha^2 f_1(r_s)} \right). \quad (3.29)$$

---

<sup>11</sup>Without loss of generality, we are assuming here that the null normals are normalized at both the left and right boundaries with the same constant. In fact, when we add the counterterm (3.3) which ensures reparametrization invariance to the null boundaries, the total action will be independent of  $\alpha$ .

The time derivatives of this joint contribution then become

$$\begin{aligned}
 \frac{dI_{\text{joint}}^{(\text{III})}}{dt_{\text{R}}} &= \frac{(d-1)\Omega_{k,d-1}}{16\pi G_N} r_m^{d-2} f_1(r_m) \frac{f_2(r_s)}{f_1(r_s)} \log \left[ \frac{|f_1(r_m)|f_2(r_s)}{\alpha^2 f_1(r_s)} \right] \\
 &\quad + \frac{\Omega_{k,d-1}}{16\pi G_N} \frac{f_2(r_s)}{f_1(r_s)} \left[ \frac{2r_m^d}{L^2} + (d-2)\omega_1^{d-2} \right] + \frac{\Omega_{k,d-1} r_m^{d-1}}{16\pi G_N} \left[ f_2'(r_s) - f_1'(r_s) \frac{f_2(r_s)}{f_1(r_s)} \right], \\
 \frac{dI_{\text{joint}}^{(\text{III})}}{dt_{\text{L}}} &= \frac{(d-1)\Omega_{k,d-1}}{16\pi G_N} r_m^{d-2} f_1(r_m) \log \left[ \frac{|f_1(r_m)|f_2(r_s)}{\alpha^2 f_1(r_s)} \right] \\
 &\quad + \frac{\Omega_{k,d-1}}{16\pi G_N} \left[ \frac{2r_m^d}{L^2} + (d-2)\omega_1^{d-2} \right]. \tag{3.30}
 \end{aligned}$$

### Counterterm contributions

Next we examine the contributions of the counterterm (3.3) to the time derivative of the holographic complexity. The counterterm is evaluated on each of the four null boundaries of the WDW patch in appendix A, but only three of these contribute to the growth rate. First, for the right past boundary, we have

$$\begin{aligned}
 I_{\text{ct}}^{(\text{I})} &= \text{“UV terms”} - \frac{\Omega_{k,d-1}}{8\pi G_N} r_m^{d-1} \left[ \log \left( \frac{(d-1)\alpha \ell_{\text{ct}}}{r_m} \right) + \frac{1}{d-1} \right] \\
 &\quad + \frac{\Omega_{k,d-1}}{8\pi G_N} (r_s^{d-1} - r_m^{d-1}) \log \left( \frac{f_1(r_s)}{f_2(r_s)} \right). \tag{3.31}
 \end{aligned}$$

The above expression corresponds to eq. (A.7) after we have substituted the affine parametrization condition (3.22). We have also left implicit the terms coming from the UV regulator surface at  $r = r_{\text{max}}$ , since they are time independent and so do not contribute to the growth rate. Of course, if we are considering early times (i.e.,  $t_{\text{R}} < t_{\text{R},c1}$  or  $t_{\text{L}} < t_{\text{L},c1}$ ) when this boundary ends on the past singularity, we simply set  $r_m = 0$  in the above expression leaving only the contribution for the crossing point  $r = r_s$ .

For the left future boundary, we find

$$I_{\text{ct}}^{(\text{II})} = \text{“UV terms”} + \frac{\Omega_{k,d-1}}{8\pi G_N} r_b^{d-1} \log \left( \frac{f_2(r_b)}{f_1(r_b)} \right), \tag{3.32}$$

by substituting eq. (3.26) for  $\hat{\alpha}$  into eq. (A.8). Here we are implicitly assuming that this boundary always terminates on the future singularity and for late times (i.e.,  $t_{\text{L}} > t_{\text{L},c2}$ ) when this boundary does not cross the shock wave, we simply set  $r_b = 0$  above. We also consider the left past boundary, for which eq. (A.9) yields

$$I_{\text{ct}}^{(\text{III})} = \text{“UV terms”} - \frac{\Omega_{k,d-1}}{8\pi G_N} r_m^{d-1} \left[ \log \left( \frac{(d-1)\alpha \ell_{\text{ct}}}{r_m} \right) + \frac{1}{d-1} \right]. \tag{3.33}$$

We might note that this contribution is identical to the first line in eq. (3.31). Again, we set  $r_m = 0$  above for early times (i.e.,  $t_{\text{R}} < t_{\text{R},c1}$  or  $t_{\text{L}} < t_{\text{L},c1}$ ) when this boundary ends on the past singularity. Finally, we also have the counterterm contribution for the right future boundary in eq. (A.10) but as noted above, it will not contribute to the complexity

growth rate, since we only consider the regime when this surface terminates at the future singularity at  $r = 0$ .

We now evaluate the time derivative of these three contributions in turn by using eqs. (3.5) and (3.6) to evaluate the time derivatives of  $r_s$ ,  $r_b$  and  $r_m$ . Let us begin with eq. (3.32) and consider the regime  $t_L < t_{L,c2}$ , which yields

$$\begin{aligned} \frac{dI_{\text{ct}}^{(\text{II})}}{dt_{\text{R}}} &= 0, \\ \frac{dI_{\text{ct}}^{(\text{II})}}{dt_{\text{L}}} &= -\frac{\Omega_{k,d-1}}{16\pi G_N} \left[ \frac{2r_b^d}{L^2} \left( 1 - \frac{f_1(r_b)}{f_2(r_b)} \right) + (d-2) \left( \omega_1^{d-2} - \omega_2^{d-2} \frac{f_1(r_b)}{f_2(r_b)} \right) \right] \\ &\quad + \frac{(d-1)\Omega_{k,d-1}}{16\pi G_N} r_b^{d-2} f_1(r_b) \log \left( \frac{f_2(r_b)}{f_1(r_b)} \right). \end{aligned} \quad (3.34)$$

Of course, for later times  $t_L > t_{L,c2}$ , both of these time derivatives vanish since  $r_b = 0$ .

Given the similarities between eqs. (3.31) and (3.33), we combine  $I_{\text{ct}}^{(\text{I})}$  and  $I_{\text{ct}}^{(\text{III})}$  in evaluating the time derivatives. For early times (i.e.,  $t_{\text{R}} < t_{\text{R},c1}$  or  $t_L < t_{L,c1}$ ), we set  $r_m = 0$ , and the time derivatives only act on  $r_s$  producing

$$\begin{aligned} \frac{d}{dt_{\text{R}}} \left( I_{\text{ct}}^{(\text{I})} + I_{\text{ct}}^{(\text{III})} \right) \Big|_{r_m=0} &= \frac{\Omega_{k,d-1}}{16\pi G_N} \left[ \frac{2r_s^d}{L^2} \left( 1 - \frac{f_2(r_s)}{f_1(r_s)} \right) + (d-2) \left( \omega_2^{d-2} - \omega_1^{d-2} \frac{f_2(r_s)}{f_1(r_s)} \right) \right] \\ &\quad - \frac{(d-1)\Omega_{k,d-1}}{16\pi G_N} r_s^{d-2} f_2(r_s) \log \left[ \frac{f_1(r_s)}{f_2(r_s)} \right], \end{aligned} \quad (3.35)$$

$$\frac{d}{dt_{\text{R}}} \left( I_{\text{ct}}^{(\text{I})} + I_{\text{ct}}^{(\text{III})} \right) \Big|_{r_m=0} = 0.$$

At later times (i.e.,  $t_{\text{R}} < t_{\text{R},c1}$  or  $t_L < t_{L,c1}$ ),  $r_m$  becomes a dynamical variable and so there are additional contributions to the above time derivatives

$$\begin{aligned} \frac{d}{dt_{\text{R}}} \left( I_{\text{ct}}^{(\text{I})} + I_{\text{ct}}^{(\text{III})} \right) &= \frac{d}{dt_{\text{R}}} \left( I_{\text{ct}}^{(\text{I})} + I_{\text{ct}}^{(\text{III})} \right) \Big|_{r_m=0} - \frac{\Omega_{k,d-1}(d-1)}{16\pi G_N} r_m^{d-2} \left[ f_2'(r_s) - \frac{f_2(r_s)}{f_1(r_2)} f_1'(r_s) \right] \\ &\quad + \frac{\Omega_{k,d-1}(d-1)}{8\pi G_N} r_m^{d-2} f_1(r_m) \frac{f_2(r_s)}{f_1(r_s)} \log \left[ \frac{(d-1)}{r_m} \frac{f_1(r_s)}{f_2(r_s)} \ell_{\text{ct}} \alpha \right] \\ \frac{d}{dt_{\text{L}}} \left( I_{\text{ct}}^{(\text{I})} + I_{\text{ct}}^{(\text{III})} \right) &= \frac{\Omega_{k,d-1}(d-1)}{8\pi G_N} r_m^{d-2} f_1(r_m) \log \left[ \frac{(d-1)}{r_m} \frac{f_1(r_s)}{f_2(r_s)} \ell_{\text{ct}} \alpha \right]. \end{aligned} \quad (3.36)$$

Of course, we will see below that when these counterterm contributions are combined with those from the rest of the gravitational action, the dependence on the normalization constant  $\alpha$  is completely eliminated — see also eq. (A.12).

### 3.1.1 Time dependence of complexity

We can now evaluate the time dependence of the holographic complexity by summing the various expressions above. However, we can consider many different forms for the time evolution, e.g., varying  $t_L$  alone or  $t_R$  alone. For simplicity, we will focus on the symmetric case where we vary  $t = t_L + t_R$  while fixing  $t_L - t_R = 0$ . This approach is closely related to the time evolution studied in [23] for an unperturbed eternal black hole (without any shock



waves). In principle though, the results above would easily allow one to study the evolution of the holographic complexity according to any other linear combination  $t' = a t_L + b t_R$ .

Further, in analogy with [23], we will focus on the evolution for  $t > 0$ .<sup>12</sup> However, there remains two important factors in determining how the holographic complexity grows, namely, the time at which the shock wave is sent in from the right boundary and its mass, i.e., the values of  $t_w$  and  $M_2 - M_1$ . In particular, these will determine the geometry of the WDW patch as discussed around eqs. (3.11) and (3.12). That is, as seen in studying the CA conjecture in the unperturbed eternal black hole, e.g., [17, 23], we are generally in a regime where  $r_m < 0$  at  $t = 0$  and the WDW patch touches some interval on the past singularity. Hence in the present situation, the shock wave parameters effect the critical time  $t_{c1}$  when  $r_m$  becomes positive and the WDW patch terminates above the past singularity. This critical time is determined by setting  $t_{L,c1} = t_{c1}/2 = t_{R,c1}$  in eq. (3.11), which then yields

$$t_{c1} = 2t_w - 4r_1^*(0) + 4r_1^*(r_s), \quad r_1^*(r_s) + r_2^*(r_s) = r_1^*(0) - t_w. \quad (3.37)$$

Here the second equation determines the value of  $r_s$  which should be substituted into the first to determine  $t_{c1}$ . Generally, increasing  $t_w$  or the mass of the shock, i.e., increasing  $M_2 - M_1$ , increases the value of  $t_{c1}$ .<sup>13</sup>

Now similarly, we are generally in a regime at  $t = 0$  where the WDW patch touches some interval on the future singularity. If we evolve forward in time, this interval simply expands but there is another critical time  $t_{c2}$  where the interval includes the point where the shock wave hits the singularity. That is,  $t_{c2}$  is the time when  $r_b$  vanishes (and then becomes negative for larger values of  $t$ ). Substituting  $t_{L,c2} = t_{c2}/2$  into eq. (3.4), we find this critical time to be

$$t_{c2} = 2t_w + 4r_1^*(0). \quad (3.40)$$

Here again, the effect of increasing  $t_w$  is to increase  $t_{c2}$ , while varying  $M_2 - M_1$  has no effect on the value of  $t_{c2}$  (if we assume that  $M_1$  is fixed).

We would like to add one more critical time to this list, in analogy with the evolution of the holographic complexity for the unperturbed eternal black hole in [23]. In that instance, there was actually an interval  $-t_{c1} \leq t \leq t_{c1}$  in which the complexity did not change. We will find a similar plateau in the case of the shock wave geometries where the rate of change is small but since shock wave breaks the time-shift symmetry, we introduce a new critical time  $t_{c0}$  to denote the beginning of this period,  $-t_{c0} \leq t \leq t_{c1}$ . Geometrically, this

<sup>12</sup>Although we also consider some times slightly before  $t = 0$ .

<sup>13</sup>These statements can be confirmed as follows: for a fixed ratio  $w = r_{h,2}/r_{h,1}$ , the following identity holds

$$\frac{dt_{c1}}{dt_w} = \frac{2(\omega_2^{d-2} - \omega_1^{d-2})}{r_s^{d-2}(f_1(r_s) + f_2(r_s))}. \quad (3.38)$$

Assuming  $M_2 > M_1$ , the above derivative is positive and so the effect of increasing  $t_w$  here is to increase the value of  $t_{c1}$ . Similarly, we have

$$\frac{dt_{c1}}{dw} = - \frac{4f_2(r_2)}{f_2(r_2) + f_1(r_2)} \frac{dr_2^*}{dw} \Big|_{r_s}, \quad (3.39)$$

and increasing  $M_2$  for a fixed  $r_s$  decreases  $r_2^*(r_s)$ . Hence increasing the mass of the second black hole also increases  $t_{c1}$ .

time is the time at which the WDW patch lifts off of the future singularity if we push  $t$  to sufficiently negative values. This critical time can be determined in a similar way to finding  $t_{c1}$  and the result is

$$t_{c0} = 2t_w - 4r_2^*(0) + 4r_2^*(r_b), \quad r_1^*(r_b) + r_2^*(r_b) = r_2^*(0) - t_w. \quad (3.41)$$

Here again, we determine  $r_b$  from the second equation and then substitute this value into the first equation to determine  $t_{c0}$ . We may note that  $r_b < r_{h,1}$ , i.e., the (future) null boundary of the WDW (on the left) crosses the shock wave behind the black hole horizon, and so one can easily show that  $t_{c0} < 2t_w$ . In some sense,  $t = -2t_w$  is the next critical time since at this point the right boundary time slice coincides with the point on the boundary surface where the shock wave originates.

Comparing eqs. (3.37) and (3.40), we find

$$t_{c2} - t_{c1} = 8r_1^*(0) - 4r_1^*(r_s(t_{c1})). \quad (3.42)$$

Now with our conventions  $r_1^*(0)$  and  $r_1^*(r_s)$  will be negative quantities — see eq. (2.15) — and so there is a competition to determine the sign of this difference.<sup>14</sup> However, at least if  $t_w$  and/or  $(M_2 - M_1)$  are sufficiently large, we expect that  $t_{c2} - t_{c1} > 0$ . In this scenario, there are three regimes of the WDW patch geometry to be considered,

$$\begin{array}{lll} \text{I :} & -t_{c0} < t < t_{c1} & r_b, r_s \text{ exist; } r_m < 0 \\ \text{II :} & t_{c1} < t < t_{c2} & r_b, r_s, r_m, \text{ exist} \\ \text{III :} & t > t_{c2} & r_s, r_m \text{ exist; } r_b < 0. \end{array} \quad (3.43)$$

For the regime I in eq. (3.43), the total rate of change of complexity consists of the bulk contribution in eq. (3.10) (with  $r_m = 0$ ), eqs. (3.14) and (3.18) for the GHY contributions from the past and future singularities, respectively, and the two counterterm contributions from eqs. (3.35) and (3.34). The final result then becomes

$$\begin{aligned} \frac{d\mathcal{C}_A^{(\text{I})}}{dt} &= \frac{M_2}{\pi} \left( 1 + \frac{f_1(r_b)}{f_2(r_b)} \right) - \frac{M_1}{\pi} \left( 1 + \frac{f_2(r_s)}{f_1(r_s)} \right) \\ &+ \frac{M_1}{2\pi} \frac{r_b^{d-2}}{\omega_1^{d-2}} f_1(r_b) \log \frac{f_2(r_b)}{f_1(r_b)} - \frac{M_2}{2\pi} \frac{r_s^{d-2}}{\omega_2^{d-2}} f_2(r_s) \log \frac{f_1(r_s)}{f_2(r_s)}. \end{aligned} \quad (3.44)$$

In regime II, we only need the GHY contribution in eq. (3.14) from the future singularity, the counterterm in the left future boundary of the WDW patch in eq. (3.34) and the sum of the joint contribution at  $r_m$  in eq. (3.30) with the two past counterterm contributions, given by eq. (3.36). In this case, the total reads

$$\begin{aligned} \frac{d\mathcal{C}_A^{(\text{II})}}{dt} &= \frac{M_2}{\pi} \left( 1 + \frac{f_1(r_b)}{f_2(r_b)} \right) + \frac{M_1}{2\pi} \frac{r_m^{d-2}}{\omega_1^{d-2}} f_1(r_m) \left( 1 + \frac{f_2(r_s)}{f_1(r_s)} \right) \log \left( \frac{|f_1(r_m)|(d-1)^2 \ell_{\text{ct}}^2}{r_m^2} \right) \\ &+ \frac{M_1}{2\pi} \frac{r_b^{d-2}}{\omega_1^{d-2}} f_1(r_b) \log \frac{f_2(r_b)}{f_1(r_b)} - \frac{M_2}{2\pi} \frac{r_s^{d-2}}{\omega_2^{d-2}} f_2(r_s) \log \frac{f_1(r_s)}{f_2(r_s)}. \end{aligned} \quad (3.45)$$

<sup>14</sup>Again,  $d = 2$  is a special case with  $r_1^*(0) = 0$  — see eq. (3.56) below.

In the last regime, the GHY contribution from the future singularity is given by eq. (3.16) and the sum of the joint at  $r_m$  in eq. (3.30) with the counterterm contributions from the past boundaries in eq. (3.36). Hence the rate of growth in this regime is

$$\begin{aligned} \frac{d\mathcal{C}_A^{(\text{III})}}{dt} &= \frac{M_1 + M_2}{\pi} + \frac{M_1}{2\pi} \frac{r_m^{d-2}}{\omega_1^{d-2}} f_1(r_m) \left( 1 + \frac{f_2(r_s)}{f_1(r_s)} \right) \log \left( \frac{|f_1(r_m)|(d-1)^2 \ell_{\text{ct}}^2}{r_m^2} \right) \\ &\quad - \frac{M_2}{2\pi} \frac{r_s^{d-2}}{\omega_2^{d-2}} f_2(r_s) \log \frac{f_1(r_s)}{f_2(r_s)}. \end{aligned} \quad (3.46)$$

**Dimensionless variables.** In addition to the dimensionless variables in eq. (2.18), it is also useful to define dimensionless coordinates corresponding to the three positions  $r_b$ ,  $r_s$  and  $r_m$ ,

$$x_s \equiv \frac{r_s}{r_{h,2}}, \quad x_m \equiv \frac{r_m}{r_{h,1}}, \quad x_b \equiv \frac{r_b}{r_{h,1}}. \quad (3.47)$$

Note that these definitions are not completely harmonious with the definition  $x = r/r_{h,2}$  in eq. (2.18). In the definition above, we have chosen to normalize  $r_s$  with  $r_{h,2}$  such that the minimum value of  $x_s$  is one, but  $r_b$  and  $r_m$  are normalized with  $r_{h,1}$  such that the maximum value of  $x_b$  and  $x_m$  is 1 as well. The addition of the counterterm (3.3) introduces one more length scale  $\ell_{\text{ct}}$  and so it will be useful to define the dimensionless quantity

$$\tilde{\ell}_{\text{ct}} \equiv \frac{\ell_{\text{ct}}}{L}. \quad (3.48)$$

Using the rescaled blackening factors in eq. (2.20), we can rewrite the rates of change of complexity in eqs. (3.44), (3.45), (3.46) and (3.48) as

$$\begin{aligned} \frac{d\mathcal{C}_A^{(\text{I})}}{dt} &= \frac{M_2}{\pi} \left( 1 + \frac{\tilde{f}(x_b, wz)}{w^2 \tilde{f}(x_b/w, z)} \right) - \frac{M_1}{\pi} \left( 1 + \frac{w^2 \tilde{f}(x_s, z)}{\tilde{f}(wx_s, wz)} \right) \\ &\quad + \frac{M_1}{2\pi} \frac{x_b^{d-2}}{1 + kz^2 w^2} \tilde{f}(x_b, wz) \log \frac{w^2 \tilde{f}(x_b/w, z)}{\tilde{f}(x_b, wz)} - \frac{M_2}{2\pi} \frac{x_s^{d-2}}{1 + kz^2} \tilde{f}(x_s, z) \log \frac{\tilde{f}(wx_s, wz)}{w^2 \tilde{f}(x_s, z)}, \end{aligned} \quad (3.49)$$

$$\begin{aligned} \frac{d\mathcal{C}_A^{(\text{II})}}{dt} &= \frac{M_2}{\pi} \left( 1 + \frac{\tilde{f}(x_b, wz)}{w^2 \tilde{f}(x_b/w, z)} \right) \\ &\quad + \frac{M_1}{2\pi} \frac{x_m^{d-2}}{1 + kz^2 w^2} \left( 1 + \frac{w^2 \tilde{f}(x_s, z)}{\tilde{f}(wx_s, wz)} \right) \tilde{f}(x_m, wz) \log \frac{|\tilde{f}(x_m, wz)|(d-1)^2 \tilde{\ell}_{\text{ct}}^2}{x_m^2} \\ &\quad + \frac{M_1}{2\pi} \frac{x_b^{d-2}}{1 + kz^2 w^2} \tilde{f}(x_b, wz) \log \frac{w^2 \tilde{f}(x_b/w, z)}{\tilde{f}(x_b, wz)} - \frac{M_2}{2\pi} \frac{x_s^{d-2}}{1 + kz^2} \tilde{f}(x_s, z) \log \frac{\tilde{f}(wx_s, wz)}{w^2 \tilde{f}(x_s, z)}, \end{aligned} \quad (3.50)$$

$$\begin{aligned} \frac{d\mathcal{C}_A^{(\text{III})}}{dt} &= \frac{M_1 + M_2}{\pi} - \frac{M_2}{2\pi} \frac{x_s^{d-2}}{1 + kz^2} \tilde{f}(x_s, z) \log \frac{\tilde{f}(wx_s, wz)}{w^2 \tilde{f}(x_s, z)} \\ &\quad + \frac{M_1}{2\pi} \frac{x_m^{d-2}}{1 + kz^2 w^2} \left( 1 + \frac{w^2 \tilde{f}(x_s, z)}{\tilde{f}(wx_s, wz)} \right) \tilde{f}(x_m, wz) \log \frac{|\tilde{f}(x_m, wz)|(d-1)^2 \tilde{\ell}_{\text{ct}}^2}{x_m^2}. \end{aligned} \quad (3.51)$$

**Early and late time behaviours.** We now turn our attention to two simple limits for the rate of change of complexity. First, let us consider early times which means that we

should consider the growth rate given in eq. (3.44). Now, if  $t_w$  is sufficiently large, then  $r_s$  approaches  $r_{h,2}$  and  $r_b$  approaches  $r_{h,1}$ , i.e.,  $f_2(r_s), f_1(r_b) \rightarrow 0$ . In this limit, the growth rate in eq. (3.49) simplifies to

$$\left. \frac{d\mathcal{C}_A^{(I)}}{dt} \right|_{t_w \rightarrow \infty} = \frac{M_2 - M_1}{\pi} + \mathcal{O}\left(T_1(2t_w - t)e^{-\pi T_1(2t_w - t)}\right), \quad (3.52)$$

i.e., it is simply proportional to the difference of masses.

Another simple limit occurs at late times, when the growth rate is given by eq. (3.51). In this case, irrespective of the value of  $t_w$ ,  $r_m$  and  $r_s$  approach  $r_{h,1}$  and  $r_{h,2}$ , respectively, i.e.,  $f_1(r_m), f_2(r_s) \rightarrow 0$ . In this case, the growth rate of the holographic complexity is given by the sum of the black hole masses, i.e.,

$$\left. \frac{d\mathcal{C}_A^{(III)}}{dt} \right|_{t \rightarrow \infty} = \frac{M_1 + M_2}{\pi} + \frac{M_1}{2} \left( \frac{d + (d-2)kz^2}{1 + kz^2} \right) T_1 t e^{-\pi T_1(t - 2t_w - 4r_1^*(r_{h,2}))} + \mathcal{O}(e^{-\pi T_1 t}), \quad (3.53)$$

as was previously argued in [17]. Further, we note that the second term in eq. (3.53) is always positive, and therefore  $d\mathcal{C}_A/dt$  approaches the previous late time limit from above. Similar behaviour was found for the unperturbed eternal black holes in [23]. In analogy to these earlier results, we also note that the next correction term, of order  $e^{-\pi T_1 t}$ , depends on the normalization factor  $\tilde{\ell}_{ct}$ . We add that more generally, the dependence on  $\tilde{\ell}_{ct}$  is more pronounced at early times.

### 3.1.2 A case study: BTZ black holes

It is instructive to analyze the particular case of BTZ black holes (i.e.,  $d = 2$ ), since the positions  $r_s, r_m$  and  $r_b$  can be determined analytically. First, the two blackening factors are given by

$$f_1(r) = \frac{r^2 - r_{h,1}^2}{L^2} \quad \text{and} \quad f_2(r) = \frac{r^2 - r_{h,2}^2}{L^2}. \quad (3.54)$$

For each black hole (so  $r_h$  can be either  $r_{h,1}$  and  $r_{h,2}$ ), the physical quantities characterizing the black hole solutions are

$$M = \frac{\Omega_{k,1} r_h^2}{16\pi G_N L^2}, \quad T = \frac{r_h}{2\pi L^2}, \quad S = \frac{\Omega_{k,1} r_h}{4G_N} = \frac{\pi}{6} c \Omega_{k,1} L T, \quad (3.55)$$

and  $c = 3L/(2G_N)$  is the central charge of the boundary CFT. In the above formulas,  $k = 0$  and  $1$  correspond to the Ramond and Neveu-Schwarz vacuum, respectively, of the boundary theory [53].

With the blackening factors in eq. (3.54), we can evaluate the tortoise coordinates in eqs. (2.15) and (2.16) as

$$r_1^*(r) = \frac{L^2}{2r_{h,1}} \log\left(\frac{|r - r_{h,1}|}{r + r_{h,1}}\right), \quad r_2^*(r) = \frac{L^2}{2r_{h,2}} \log\left(\frac{|r - r_{h,2}|}{r + r_{h,2}}\right). \quad (3.56)$$

We choose to normalize the time coordinates by the temperature of the final (more massive) black hole, which reads

$$T_2 = \frac{r_{h,2}}{2\pi L^2}. \quad (3.57)$$

For the BTZ geometry,  $w$  which is the ratio of the horizon sizes in eq. (2.18) is just the ratio of the temperatures,

$$w = T_2/T_1. \quad (3.58)$$

Further, we note that  $M_2/M_1 = w^2$  and  $S_2/S_1 = w$ .

Now combining eqs. (3.4), (3.47) and (3.56), as well as the above ratio (3.58), yields the following:

$$\begin{aligned} x_s &= \frac{1 + e^{-2\pi T_2(t_R+t_w)}}{1 - e^{-2\pi T_2(t_R+t_w)}}, & x_b &= \frac{e^{-2\pi T_1(t_L-t_w)} - 1}{e^{-2\pi T_1(t_L-t_w)} + 1}, \\ x_m &= \frac{wx_s + 1 - (wx_s - 1)e^{-2\pi T_1(t_L-t_w)}}{wx_s + 1 + (wx_s - 1)e^{-2\pi T_1(t_L-t_w)}}. \end{aligned} \quad (3.59)$$

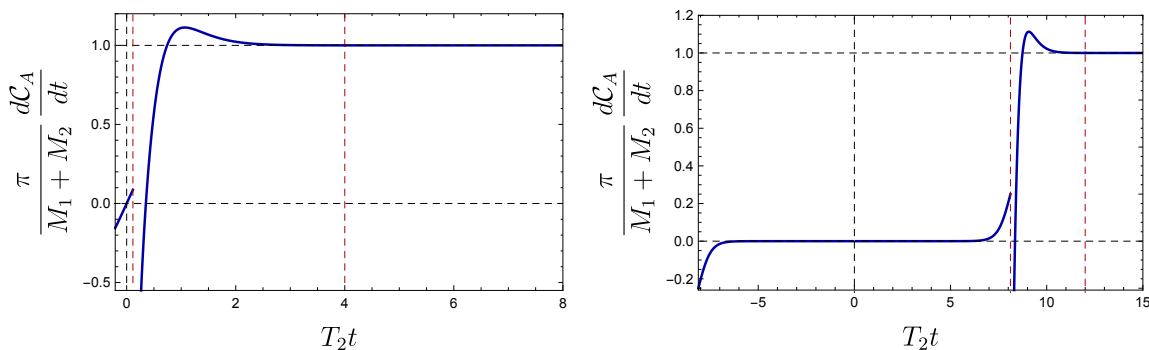
With these three expressions, the growth rates in eqs. (3.49)–(3.51) are implicitly expressed entirely in terms of boundary quantities. Further, from eq. (3.54), we see that  $r_1^*(0) = 0$  and hence the critical times in eqs. (3.37)–(3.41) simplify to

$$t_{c2} = 2t_w, \quad t_{c1} = 2t_w - 4|r_1^*(r_s)| \quad \text{and} \quad t_{c0} = 2t_w - 4|r_2^*(r_b)|. \quad (3.60)$$

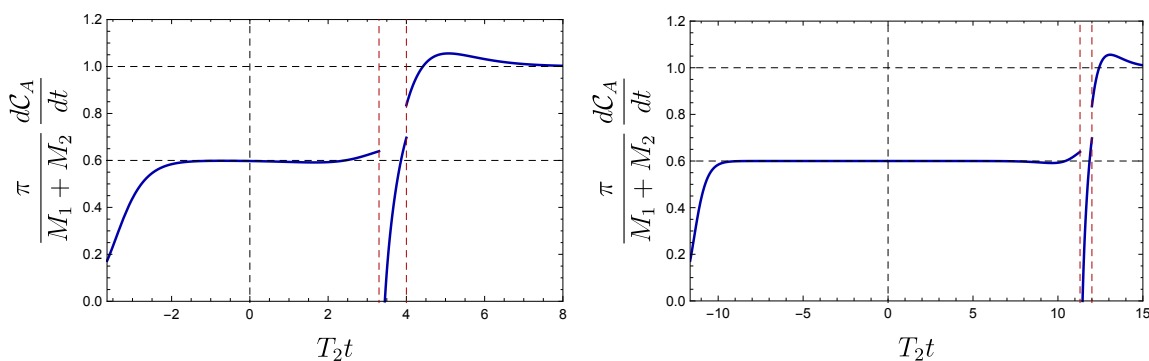
While we do not have an analytic expression for  $r_s$ , it is easily determined numerically by combining the expressions in eqs. (3.37) and (3.54), and similarly for  $r_b$ . We return to examine the critical times  $t_{c1}$  and  $t_{c0}$  in more detail in a moment. In any event, we see that we are in the situation with  $t_{c1} < t_{c2}$  and so the evolution of the holographic complexity is described by the scenario in eq. (3.43) and so let us explicitly examine  $d\mathcal{C}_A/dt$  in a few examples.

In figure 2, we show  $d\mathcal{C}_A/dt$  for a very light shock wave where  $w = 1 + 10^{-5}$ . In the left panel, we show the results for  $T_2 t_w = 2$ , and in the right, for  $T_2 t_w = 6$ . For both cases, the growth rate is essentially zero over the period  $-t_{c0} \lesssim t \lesssim t_{c1}$ , however, this is a longer time period for a larger value of  $t_w$ . Immediately after  $t_{c1}$ , there is a negative spike in the rate of growth, which is similar to the one found for the eternal BTZ black hole with the inclusion of the counterterm in appendix A of [23]. Note that this very small initial growth rate is consistent with eq. (3.52) since the difference  $(M_2 - M_1) = (w^2 - 1)M_1$ . Further, the separation  $t_{c2} - t_{c1}$  (as well as  $2t_w - t_{c0}$ ) appears to be independent of  $t_w$ . We will examine these observations further in the following.

In figure 3, we show  $d\mathcal{C}_A/dt$  for a heavier shock wave where  $w = 2$  (i.e., the temperature doubles or the black hole mass increases by a factor of four) and  $z = 1/w$ , such that the smaller black hole is at the Hawking-Page transition. In the left panel, we show the results for  $T_2 t_w = 2$ , and in the right, for  $T_2 t_w = 6$ . For both cases, the growth rate is significantly lower (than the final rate) in the period  $-t_{c0} \lesssim t \lesssim t_{c1}$ . This plateau is more evident in the case with a larger value of  $t_w$ . Rather than vanishing in this period,  $d\mathcal{C}_A/dt$  is given by the difference  $M_2 - M_1$ , as in eq. (3.52). Note that the separation  $t_{c2} - t_{c1}$  (as well as  $2t_w - t_{c0}$ ) again appears to be independent of  $t_w$ , but is a smaller interval (when normalized by  $T_2$ ) than with a very light shock wave, as in figure 2.



**Figure 2.** Time derivative of complexity, evolving both boundaries as  $t_L = t_R = \frac{t}{2}$  with  $T_2 t_w = 2$  (left) and  $T_2 t_w = 6$  (right). We have set  $w = 1 + 10^{-5}$  and  $\tilde{\ell}_{ct} = 1$ . The condition on  $z$  implies that the smaller black hole is at the Hawking-Page transition, for both cases. The lower horizontal dashed line (near zero) corresponds to  $(M_2 - M_1)/(M_2 + M_1)$ , and by construction, the late time limit approaches 1 at the higher horizontal line. The horizontal axis in both figures starts from the respective  $t_{c0}$  in eq. (3.60). The first vertical black dashed line appears at  $t = 0$ , while the vertical red dashed lines appear at  $t_{c1}$  (left) and  $t_{c2} = 2t_w$  (right), see also eq. (3.60). There is a negative spike right after  $t_{c1}$ , where  $x_m$  is close to the past singularity. For the earlier shock wave in the right figure, there is a long regime where the rate of change is close to zero. In both cases, the late time limit is approached from above.



**Figure 3.** Time derivative of complexity, evolving both boundaries as  $t_L = t_R = \frac{t}{2}$ , with  $T_2 t_w = 2$  (left) and  $T_2 t_w = 6$  (right). In both cases, we have set  $w = 2$  and  $\tilde{\ell}_{ct} = 1$ . The lower horizontal black dashed line corresponds to the time derivative at early times, i.e.,  $(M_2 - M_1)/\pi$  in eq. (3.52), and the higher line to the late time limit, i.e.,  $(M_2 + M_1)/\pi$  in eq. (3.53). The horizontal axis starts at  $t_{c0}$ , and the critical times  $t_{c1}$  and  $t_{c2}$  are shown by the left and right vertical dashed red lines, respectively. There is a negative spike right after  $t_{c1}$ , where  $x_m$  is close to the past singularity. Pushing the shock wave to the past increases the plateau where the time derivative is given roughly by the difference of the masses.

## Critical times in BTZ

Here, we examine the critical times for the special case  $d = 2$  in more detail. Recall that in this case, eq. (3.54) yields  $r_1^*(0) = 0$  and hence the critical times in eqs. (3.37)–(3.41) simplify to the expressions given in (3.60). Hence the critical time where the endpoint of the shock wave on the future singularity enters the WDW patch is simply given by  $t_{c2} = 2t_w$ .<sup>15</sup> However, the critical times  $t_{c1}$  and  $-t_{c0}$  where the WDW patch lifts off of the past singularity and first impinges on the future singularity, respectively, have a more interesting structure. From eq. (3.52), we found that during the period  $-t_{c0} \lesssim t \lesssim t_{c1}$ ,<sup>16</sup> the growth rate of the holographic complexity is roughly proportional to the difference of the masses, at least when  $t_w$  is sufficiently large. This plateau with  $d\mathcal{C}_A/dt \simeq (M_2 - M_1)/\pi$  is clearly shown in figures 2 and 3.<sup>17</sup>

From eq. (3.60), we have  $t_{c1} = 2t_w - 4|r_1^*(r_s)|$  for the  $d = 2$  shock wave geometries. We would like to understand this result in terms of boundary quantities and this is most simply done by considering various limits. First, suppose that the shock wave is very heavy, i.e.,  $w$  in eq. (3.58) is a large parameter. Recalling that  $t_{c1}$  is the critical time when  $x_m$  becomes positive, we may use eq. (3.59) to find

$$t_{c1} = 2t_w - \frac{2}{\pi T_2} + \mathcal{O}\left(\frac{1}{w^2 T_2}\right), \quad (3.61)$$

for large  $w$ . For very high temperatures, the above expression implies that this critical time approaches  $t_{c2}$ , i.e.,  $t_{c1} \rightarrow 2t_w = t_{c2}$ .

We also consider the case of a very light shock for which  $w$  can be parametrized as  $w = 1 + \epsilon$ . Using eqs. (3.55) and (3.58), the ratio of the masses is given by

$$\frac{M_2}{M_1} = w^2 = 1 + 2\epsilon + \epsilon^2, \quad (3.62)$$

and hence the energy of the shock  $E$  is given by

$$\frac{E}{M_1} = 2\epsilon + \epsilon^2 \simeq 2\epsilon. \quad (3.63)$$

Now again using eq. (3.59), we have in the limit  $e^{-2\pi T_2 t_w} \ll \epsilon \ll 1$

$$t_{c1} = 2t_w + \frac{1}{\pi T_1} \log \frac{\epsilon}{2} - \frac{\epsilon}{2\pi T_1} + \mathcal{O}\left(\frac{\epsilon^2}{T_1}\right). \quad (3.64)$$

<sup>15</sup>Note that this corresponds to  $t_L = t_w$  and this simple result arises from the special property that the singularity is a straight horizontal line in a Penrose diagram describing the BTZ black hole [54].

<sup>16</sup>This was regime I in eq. (3.43).

<sup>17</sup>In these figures, we can see that there are significant tails on the plateau in the interval  $-t_{c0} \leq t \leq t_{c1}$ . A better estimate of the length of the plateau can be determined from the analytic expressions of  $x_b$  and  $x_s$  in eq. (3.59), as follows: the plateau is in the regime where both  $x_s$  and  $x_b$  are close to 1. Let us define the “boundaries” of the plateau with the conditions that  $x_s \approx 1 + 2e^{-\gamma\pi} + \mathcal{O}(e^{-2\gamma\pi})$  and  $x_b \approx 1 - 2e^{-\gamma\pi} + \mathcal{O}(e^{-2\gamma\pi})$ , where  $\gamma$  is a number of order 1. It then follows that the length of the plateau is approximately  $T_2 \Delta t_{\text{pl}} \approx 4T_2 t_w - \gamma(w + 1)$ .

Following [49], we can relate the first correction to the scrambling time [55]. If one considers  $E$  to be of the order of the energy of a few thermal quanta of energy, then we may use eqs. (3.55) and (3.63) to write<sup>18</sup>

$$\frac{2}{\epsilon} \simeq \frac{4M_1}{E} \simeq \frac{4M_1}{2T_1} \simeq S_1. \quad (3.65)$$

Hence eq. (3.64) becomes

$$t_{c1} = 2(t_w - t_{scr}^*) + \mathcal{O}(\epsilon) \quad (3.66)$$

where

$$t_{scr}^* \equiv \frac{1}{2\pi T_1} \log \frac{2}{\epsilon} = \frac{1}{2\pi T_1} \log S_1. \quad (3.67)$$

Having evaluated the behaviour of the critical time  $t_{c1}$  for heavy and light shocks in eqs. (3.61) and (3.64), respectively, we plot the numerical solution from eq. (3.60) in figure 4. In the left panel, we show the behaviour of  $t_{c1}$  as a function of  $\log(\epsilon/2)$  for early shock waves, i.e., for which  $x_s - 1 \simeq 2e^{-2\pi T_2 t_w} \ll 1$ . In the figure, we clearly see the transition between the light shock behaviour (where  $2t_w - t_{c1}$  depends linearly on  $\log(\epsilon/2)$ ) and the heavy shock behaviour (where  $2t_w - t_{c1}$  is constant) and that the value of  $w$  where the transition occurs is independent of  $t_w$ . Recall that an essential assumption in deriving eq. (3.64) was that the order of limits  $e^{-2\pi T_2 t_w} \ll \epsilon \ll 1$  held. The geometrical interpretation of this limit is that  $x_s$  is exponentially close to 1, and therefore corrections of order  $e^{-2\pi T_2 t_w}$  are much smaller than corrections to the energy from the shock wave, which are of order  $\epsilon$ . Therefore in figure 4 where the value of  $t_w$  is fixed for each curve, we see there is a regime of very small  $\epsilon$  where  $\epsilon \lesssim e^{-2\pi T_2 t_w}$  where the difference  $2t_w - t_{c1}$  again saturates at some constant value. In the right panel of figure 4, we show the behaviour when the shock waves are not sent very early. In this regime,  $t_{c1}$  just transitions between two different constant regimes without much of a linear regime in between. Further, increasing the mass ratio decreases the critical time. For large  $w$ , notice that most curves (with big enough  $t_w$ ) saturate to 1, which is consistent with the large  $w$  expansion in eq. (3.61). Note that the plots produced here in figure 4 also represent the difference  $t_{c2} - t_{c1}$ , since for  $d = 2$  from eq. (3.60) we have  $t_{c2} = 2t_w$ . Hence in the early shock regime, the right panel shows that the separation between these two critical times is independent of  $t_w$  (except for very small  $\epsilon$ ), as observed in figures 2 and 3.

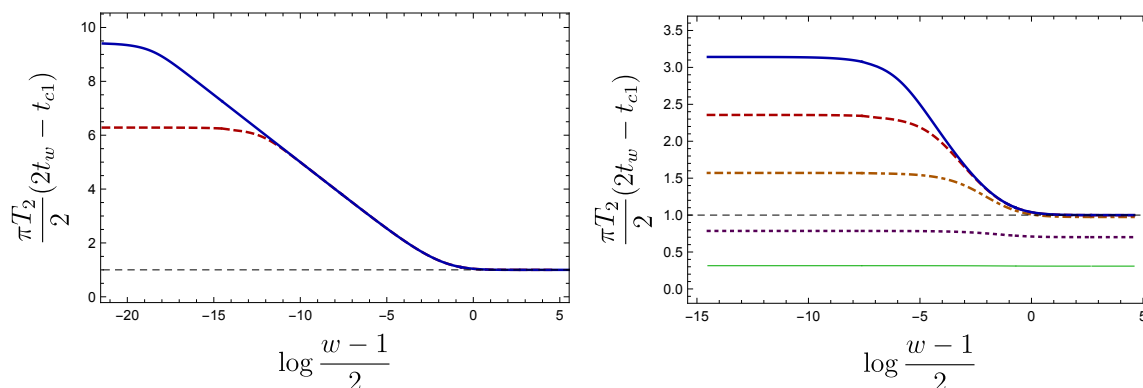
Next, we turn our attention to  $t_{c0} = 2t_w - 4|r_2^*(r_b)|$  from eq. (3.60), which for early shock waves (i.e.,  $t_w$  large) represents the beginning (i.e., at  $t = -t_{c0}$ ) of the plateau where the rate of growth is approximately  $(M_2 - M_1)/\pi$ . We begin by considering light shock waves in the limit with  $e^{-2\pi T_1 t_w} \ll \epsilon \ll 1$ . In this scenario,  $r_b \rightarrow r_{h,1}$  as the ratio of temperatures approaches one (i.e.,  $w \rightarrow 1$ ), and  $r_b \rightarrow 0$  as  $w$  increases. Therefore, we can expand eq. (3.41) for  $w = 1 + \epsilon$  with  $\epsilon$  small,<sup>19</sup>

$$t_{c0} = 2(t_w - t_{scr}^*) + \frac{\epsilon}{\pi T_1} \left( \log \frac{2}{\epsilon} - \frac{1}{2} \right) + \mathcal{O}(\epsilon^2 \log \epsilon). \quad (3.68)$$

<sup>18</sup>Note that we chose  $E \simeq 2T_1$  to simplify the subsequent expressions.

<sup>19</sup>Note that in this case, the expression would be simpler if we had defined the scrambling time with  $T_2$ . That is, using  $t_{scr}^* = \frac{1}{\pi T_2} \log(2/\epsilon)$ , rather than the definition in eq. (3.67), would remove the  $\epsilon \log \epsilon$  correction in eq. (3.68).





**Figure 4.** Dependence of the critical time  $t_{c1}$  on  $\log \frac{w-1}{2} = \log(\epsilon/2)$ , which parametrizes the energy in the shock wave. In the left panel, we show the behaviour of  $t_{c1}$  for early shock waves:  $T_2 t_w = 3$  in solid blue, and  $T_2 t_w = 2$  in dashed red. In this case, we see the transition between the light shock behaviour (3.64) and the heavy shock behavior (3.61). In the right panel, we show the behaviour of  $t_{c1}$  when the shock wave is not sent very early, i.e., with  $T_2 t_w = 0.1, 0.25, 0.5, 0.75, 1$  from bottom to top. For these parameters, the range of  $w$  that has approximately a log dependence starts appearing as the shock wave is sent earlier (larger  $t_w$ ). The horizontal thin dashed black line is just 1 (for both panels).

For heavy shock waves (i.e., large  $w$ ),  $t_{c0}$  scales as

$$t_{c0} = t_w \left( 2 - \frac{4}{w^2} + \mathcal{O}(w^{-4}) \right). \tag{3.69}$$

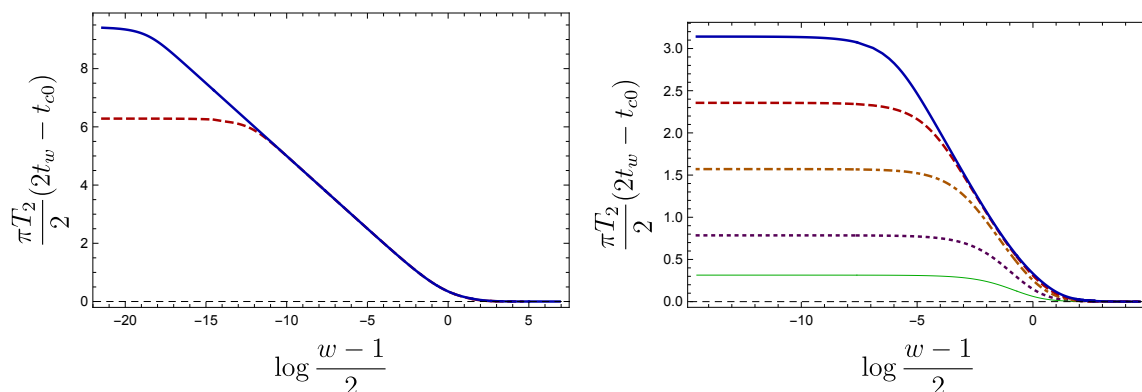
In figure 5, we show the numerical solution of eq. (3.41) for  $T_2 t_w = 2$  in the left panel, and  $T_2 t_w = 0.25$  in the right panel. For the early shock wave and small  $\epsilon$ , we see that  $t_{c0}$  depends linearly on  $\log \epsilon$ . As a result, the plateau (where the derivative is close to zero) will extend far into the past. If the shock wave is not sent early enough the range with this log dependence is much shorter, similar to the behaviour found for  $t_{c1}$ .

We focused our analysis of the critical times here on the special case of  $d = 2$  because many features, such as the dependence of  $x_s$  and  $x_b$  on  $t_w$ , were analytic. In addition, since eq. (3.56) yields  $r^*(0) = 0$  for  $d = 2$ ,  $t_{c2} - t_{c1}$  was always positive and  $t_{c2}$  was simply given by  $2t_w$ . We investigate higher dimensions (in particular  $d = 4$ ) in appendix B. There, the fact that  $r^*(0) \neq 0$  leads to some modifications for shock waves not inserted early enough, i.e., for small  $t_w$ , we may find that  $t_{c2} - t_{c1}$  is negative. On the other hand, if the shock wave is sent early enough, it is also true that in higher dimensions, there is a plateau of rate of change  $(M_2 - M_1)/\pi$  that extends for a length of time of approximately  $4t_w$ .

### 3.2 Complexity of formation

In this section, we consider the complexity of formation, as previously discussed in [22]. There, we compared the complexity of preparing two copies of the boundary CFT in the thermofield double state (TFD) at  $t_L = t_R = 0$  to the complexity of preparing each of the CFT's in its vacuum state,

$$\Delta\mathcal{C} = \mathcal{C}(|TFD\rangle) - \mathcal{C}(|0\rangle \otimes |0\rangle). \tag{3.70}$$



**Figure 5.** Dependence of the critical time  $t_{c0}$  on the energy of the shock wave, parametrized by the temperature ratio  $w$  for BTZ. In the left, we show the behaviour of  $t_{c0}$  with respect to early shock waves with  $T_2 t_w = 3$  in solid blue,  $T_2 t_w = 2$  in dashed red. Similarly to  $t_{c1}$  in figure 4, as  $w$  approaches one, there is a stretched range of  $w$  such that  $t_{c1}$  grows as a logarithm, and the earlier the shock wave the longer this log regime. Also, we see that it approaches  $2t_w$  in the large  $w$  regime. In the right, we show the behaviour of  $t_{c0}$  when the shock wave is not sent early enough, with  $T_2 t_w = 0.1, 0.25, 0.5, 0.75, 1$  from bottom to top. As the shock wave is sent earlier, the region with log dependence becomes more pronounced.

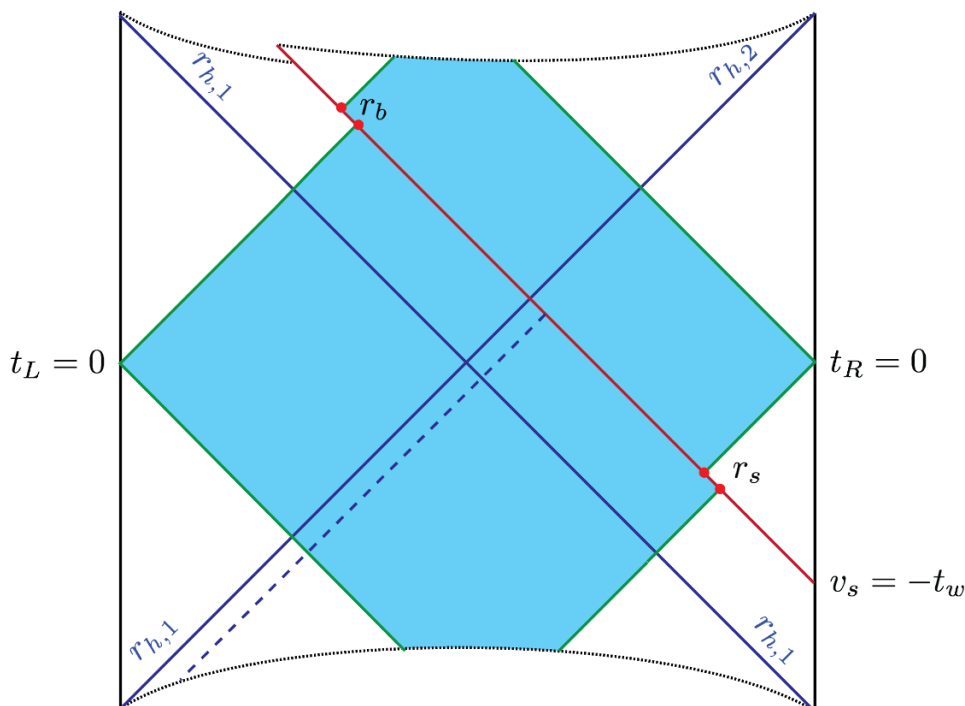
Using the CA conjecture (1.2), the holographic calculation was to evaluate the WDW action for  $t_L = t_R = 0$  in an eternal black hole background and subtract that for two copies of the AdS vacuum geometry. This difference removed the UV divergences leaving a finite quantity. For neutral black holes,<sup>20</sup> we found that at high temperatures generally  $\Delta\mathcal{C}$  was proportional to the entropy of the black hole or alternatively, the entanglement entropy in the TFD state, plus small curvature corrections. However,  $d = 2$  was a special case where for the BTZ black hole,  $\Delta\mathcal{C}$  was a constant proportional to the central charge.

In the following, we aim to evaluate the complexity of formation for the perturbed state dual to the shock wave geometry, again at  $t_L = t_R = 0$ .<sup>21</sup> The resulting  $\Delta\mathcal{C}_A$  can be studied as a function of  $t_w$  and  $M_2 - M_1$ . We illustrate this setup with the Penrose diagram of figure 6. The calculation follows straightforwardly from the considerations of the previous subsection. For instance, the bulk integral is obtained with  $r_m = 0$  and also by setting  $t_L = t_R = 0$  in eq. (3.7). Also, we have to subtract two copies of the vacuum, which was discussed in detail in [22]. We have then

$$\Delta I_{\text{bulk}} = \left( \frac{\Omega_{k,d-1}}{16\pi G_N} \right) \left( -\frac{2d}{L^2} \right) \left[ \int_{r_s}^{r_{\text{max}}} dr r^{d-1} (-2r_2^*(r)) + \int_{r_b}^{r_{\text{max}}} dr r^{d-1} (-2r_1^*(r)) \right. \\ \left. + \int_0^{r_s} dr r^{d-1} (t_w + 2r_1^*(r_s) - 2r_1^*(r)) + \int_0^{r_b} dr r^{d-1} (t_w + 2r_2^*(r_b) - 2r_2^*(r)) \right] - 2I_{\text{bulk,vac}}, \quad (3.71)$$

<sup>20</sup>These calculations were extended to charged black holes in [23].

<sup>21</sup>The boundary times are synchronized according to the procedure outlined at the beginning of this section — see the discussion above eq. (2.18).



**Figure 6.** Penrose-like diagram for one shock wave sent from the right boundary at  $v_s = -t_w$  on an eternal black hole geometry, with the Wheeler-DeWitt patch anchored at  $t_L = t_R = 0$ , which we will call the complexity of formation, in analogy to the case studied in [22] of the unperturbed eternal black hole geometry. There are two coordinates  $r_b$  and  $r_s$  that are usually given by a transcendental equation as functions of  $w$  and  $t_w$ , as shown in eq. (3.4). The effect of crossing the collapsing shell from the right boundary is to increase the surface of the WDW patch above the past singularity.

where  $r_b$  and  $r_s$  are given by eq. (3.4) with  $v_L = v_R = 0$ , and  $I_{\text{bulk,vac}}$  is the appropriate vacuum bulk integral given by

$$2I_{\text{bulk,vac}} = \left( \frac{\Omega_{k,d-1}}{16\pi G_N} \right) \left( -\frac{2d}{L^2} \right) \int_0^{r_{\text{max}}} dr r^{d-1} (-4r_{\text{vac}}^*(r)), \quad (3.72)$$

with

$$r_{\text{vac}}^*(r) = - \int_r^\infty \frac{dr}{k + r^2/L^2}, \quad \lim_{r \rightarrow \infty} r_{\text{vac}}^*(r) = 0. \quad (3.73)$$

The only nonvanishing contributions from the boundary surfaces are the two GHY contributions at the past and future singularities, given by eqs. (3.13) and (3.17) with  $v_L = v_R = 0$ , which results in

$$\Delta I_{GHY} = \frac{d\Omega_{k,d-1}}{16\pi G_N} \left[ \omega_1^{d-2} (t_w + 2r_1^*(r_s) - 2r_1^*(0)) + \omega_2^{d-2} (t_w + 2r_2^*(r_b) - 2r_2^*(0)) \right]. \quad (3.74)$$

Finally, we need to add the contribution of the two counterterms in eqs. (3.31) and (3.32) with  $r_m = 0$ . The UV contributions cancel when subtracting the vacuum, so as a result

we have

$$\Delta I_{\text{ct}} = \frac{\Omega_{k,d-1}}{8\pi G_N} \left[ r_s^{d-1} \log \left( \frac{f_1(r_s)}{f_2(r_s)} \right) + r_b^{d-1} \log \left( \frac{f_2(r_b)}{f_1(r_b)} \right) \right]. \quad (3.75)$$

As argued in [22], the joint contributions at the UV regulator surface precisely cancel the same contributions from the vacuum geometries.

Combining all of these contributions then yields the desired complexity of formation,

$$\Delta \mathcal{C}_A = \frac{\Delta I_{\text{bulk}} + \Delta I_{\text{GHY}} + \Delta I_{\text{ct}}}{\pi}. \quad (3.76)$$

This result is more complicated than the complexity of formation for the unperturbed BH geometry, since the points  $r_s$  and  $r_b$  are determined by inverting a transcendental equation (for higher dimensional black holes). However, it can be studied analytically for  $d = 2$  and we consider this special case in the following. We will also consider  $\Delta \mathcal{C}_A$  for planar AdS<sub>5</sub> black holes in appendix B.

It is straightforward to evaluate eq. (3.76) for  $d = 2$ , and it is instructive to compare the result to the complexity of formation for the unperturbed BTZ black hole. The latter was evaluated in [22], where we found  $\Delta \mathcal{C}_{\text{NS}} = -c/3$  when subtracting the complexity of the Neveu-Schwarz vacuum (i.e.,  $k = +1$ ).<sup>22</sup> Comparing the result for the perturbed state to  $\Delta \mathcal{C}_{\text{NS}}$  then yields

$$\begin{aligned} \frac{\Delta \mathcal{C}_A - \Delta \mathcal{C}_{\text{NS}}}{|\Delta \mathcal{C}_{\text{NS}}|} = LT_1 & \left[ \frac{w^2 - 1}{w} \log \left( \frac{x_s - 1}{x_s + 1} \right) + wx_s \log \left( \frac{w^2 x_s^2 - 1}{w^2 (x_s^2 - 1)} \right) \right. \\ & + (w^2 - x_b^2) \log \left( \frac{1 + x_b}{1 - x_b} \right) - \frac{w^2 - x_b^2}{w} \log \left( \frac{w + x_b}{w - x_b} \right) \\ & \left. + x_b \log \left( \frac{w^2 - x_b^2}{1 - x_b^2} \right) \right]. \end{aligned} \quad (3.77)$$

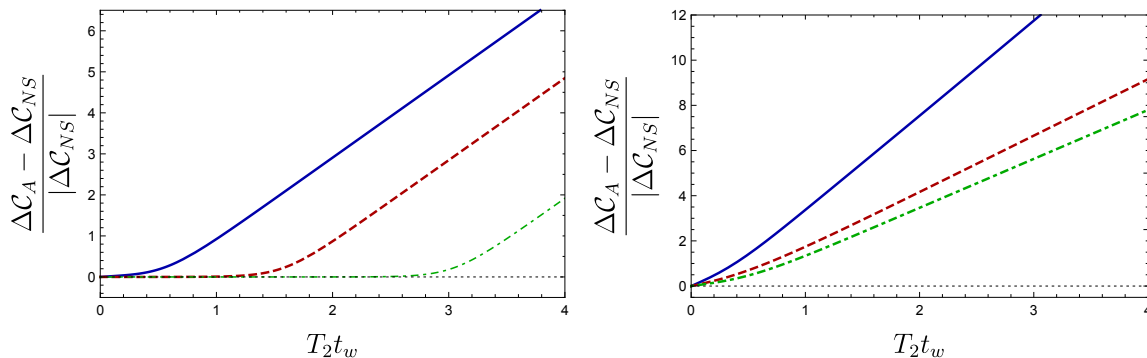
Here the coordinates  $x_s$  and  $x_b$  are given by eq. (3.59) with  $t_L = t_R = 0$ ,

$$x_s = \frac{1 + e^{-2\pi T_2 t_w}}{1 - e^{-2\pi T_2 t_w}}, \quad x_b = \frac{1 - e^{-2\pi T_1 t_w}}{1 + e^{-2\pi T_1 t_w}}. \quad (3.78)$$

In the left panel of figure 7, we show the effect of a light shock wave on the complexity of formation as a function of  $t_w$ . Initially there is a period where  $\Delta \mathcal{C}_A = \Delta \mathcal{C}_{\text{NS}}$  after which  $\Delta \mathcal{C}_A$  begins to grow linearly. As the shock is made lighter (i.e., as  $w$  is brought closer to one), this period over which the complexity of formation is essentially unchanged grows longer. In the period of linear growth, the slope seems more or less the same independent of  $w$ . In the right panel, we show the effect of heavier shock waves. In this regime, the complexity of formation starts changing immediately, even for small  $t_w$ , and  $\Delta \mathcal{C}_A$  rapidly enters a regime of linear growth with increasing  $t_w$ . In appendix B, similar features are found with shock wave geometries which are inserted into a planar AdS<sub>5</sub> black hole spacetime.

We want to investigate the behaviour of figure 7 analytically in the case of a very light shock wave with  $w = 1 + \epsilon$ . We start by analyzing eq. (3.77) in the limit where the shock

<sup>22</sup>With  $c$  being the central charge of the boundary CFT, which is given by  $c = 3L/(2G_N)$ .



**Figure 7.** The complexity of formation for BTZ black holes for  $z = 1/w$ , such that the smaller black hole is at the Hawking-Page transition, or alternatively we could normalize the complexity of formation by the entropy, which would remove the overall multiplicative factor of  $z$ , cf. eq. (3.77). In the left panel, we evaluate the complexity of formation for light shock waves as a function of  $T_2 t_w$ . The energies of the shock waves are parametrized by the temperature ratio  $w$ , with  $w = 1 + 10^{-1}$  (solid blue),  $w = 1 + 10^{-4}$  (dashed red) and  $w = 1 + 10^{-8}$  (dot-dashed green). For a period of time of the order of the scrambling time (3.67), the complexity of formation is approximately the same as the unperturbed state. For early shocks (i.e., larger  $t_w$ ), the complexity of formation grows linearly with  $t_w$ . In the right panel, we show the complexity of formation for heavier shocks,  $w = 4$  (solid blue),  $w = 2.5$  (dashed red) and  $w = 1.5$  (dot-dashed green). For these parameters, we see that the complexity of formation starts changing immediately and rapidly approaches a regime of linear growth with increasing  $t_w$ .

wave enters at a very early time, i.e.,  $T_2 t_w \gg 1$ . In eq. (3.78), the coordinates  $x_s$  and  $x_b$  become

$$x_s = 1 + 2e^{-2\pi T_1 t_w} + \mathcal{O}(\epsilon e^{-2\pi T_1 t_w}, e^{-4\pi T_1 t_w}), \quad x_b = 1 - 2e^{-2\pi T_1 t_w} + \mathcal{O}(e^{-4\pi T_1 t_w}). \quad (3.79)$$

In this limit, the leading order behaviour of eq. (3.77) reduces to

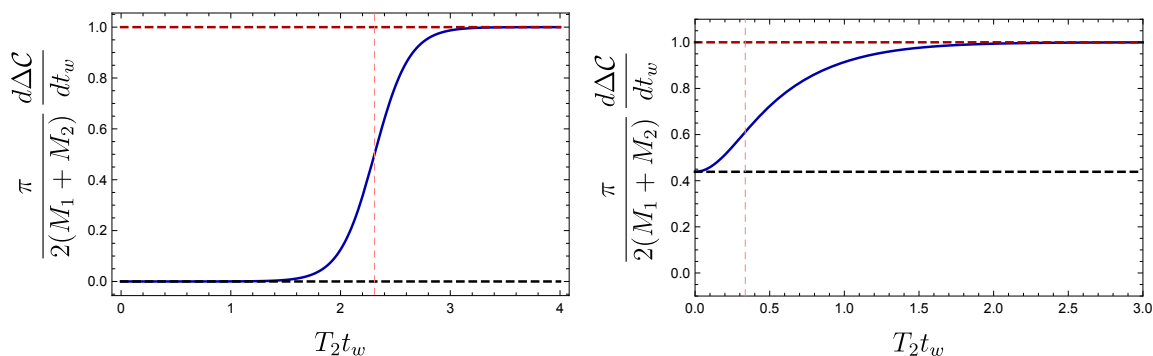
$$\frac{\Delta\mathcal{C}_A - \Delta\mathcal{C}_{NS}}{|\Delta\mathcal{C}_{NS}|} = LT_1 \log\left(\frac{(wx_s - 1)(w - x_b)}{(x_s - 1)(1 - x_b)}\right) + \mathcal{O}(\epsilon, e^{-2\pi T_2 t_w}). \quad (3.80)$$

Now there are two interesting regimes to consider:  $\epsilon \ll 2e^{-2\pi T_2 t_w}$  and  $\epsilon \gg 2e^{-2\pi T_2 t_w}$ . Of course, the transition between these two regimes occurs when  $\epsilon \approx 2e^{-2\pi T_1 t_w}$ , i.e., when  $t_w \approx \frac{1}{2\pi T_1} \log(2/\epsilon) = t_{\text{scr}}^*$  using eq. (3.67). That is, the transition occurs when the perturbation of the thermofield double state is made approximately one scrambling time before the complexity of formation is evaluated!

In the first regime, we can simply approximate  $w \approx 1$  in the argument of the log in eq. (3.80), and as a consequence, the latter becomes

$$\frac{\Delta\mathcal{C}_A - \Delta\mathcal{C}_{NS}}{|\Delta\mathcal{C}_{NS}|} = \mathcal{O}(e^{-2\pi T_2 t_w}) \quad (3.81)$$

where we have omitted order  $\epsilon$  corrections because by assumption they were smaller than the exponential. This is the regime where the complexity of formation is essentially the same as the unperturbed geometry in figure 7.



**Figure 8.** The derivative of the complexity of formation with respect to  $t_w$ . The black dashed line is the expected slope at  $t_w = 0$ , and the dashed red line is the slope at large  $t_w$  (normalized in the plot to approach 1). For both panels, we adopt  $z = 1/w$ . The left panel illustrates the behaviour for a light shock wave, with  $w = 1 + 10^{-6}$ . In this regime, described by eq. (3.84), the slope is approximately zero until  $t \simeq t_{\text{scr}}^*$  (vertical dashed line), at which point it rapidly rises to the final constant value,  $4M_1/\pi$ . The right panel illustrates the behaviour for a heavy shock wave, with  $w = 2$ . In this regime, the slope starts at  $\gamma_0$  in eq. (3.85) and rapidly rises to the final constant value  $2(M_1 + M_2)/\pi$ . In this case, the vertical dashed line indicates  $t_w = t_{\text{del}}$  from eq. (3.87).

In the second regime with  $\epsilon \gg e^{-2\pi T_2 t_w}$ , the denominator of the log in eq. (3.80) becomes the dominant part, with

$$\frac{\Delta\mathcal{C}_A - \Delta\mathcal{C}_{\text{NS}}}{|\Delta\mathcal{C}_{\text{NS}}|} = 2LT_1 \left[ 2\pi T_1 t_w + \log\left(\frac{\epsilon}{2}\right) \right] + \mathcal{O}(\epsilon). \quad (3.82)$$

Hence this second regime is where  $\Delta\mathcal{C}_A$  grows linearly with  $t_w$  in figure 7. Using the expressions in eq. (3.55) (with  $\Omega_{1,1} = 2\pi$  and  $2M_1 = S_1 T_1$ ) and the scrambling time in eq. (3.67), as well as  $|\Delta\mathcal{C}_{\text{NS}}| = c/3$ , we can rewrite the last result as

$$\Delta\mathcal{C}_A = \Delta\mathcal{C}_{\text{NS}} + \frac{4M_1}{\pi} (t_w - t_{\text{scr}}^*) + \mathcal{O}(\epsilon). \quad (3.83)$$

Hence we can approximate the complexity of formation in both regimes with the following simple expression:

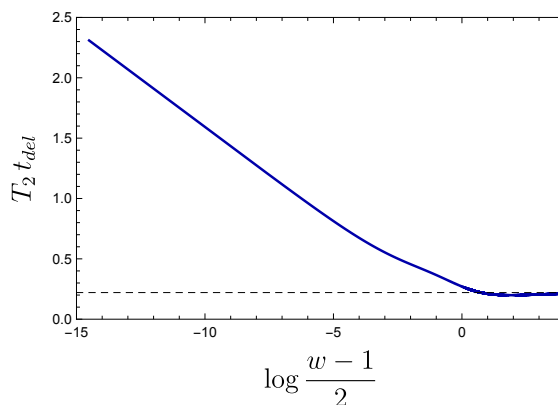
$$\Delta\mathcal{C}_A \simeq \Delta\mathcal{C}_{\text{NS}} + \Theta(t_w - t_{\text{scr}}^*) \frac{4M_1}{\pi} (t_w - t_{\text{scr}}^*). \quad (3.84)$$

For the heavy shock waves, i.e., (relatively) large  $w$ , we see from the right panel of figure 7 that  $\Delta\mathcal{C}_A$  begins increasing immediately as  $t_w$  increases from zero. From eq. (3.77), we can evaluate the slope of this increase,

$$\left. \frac{d\Delta\mathcal{C}_A}{dt_w} \right|_{t_w \rightarrow 0^+} = \frac{1}{\pi} \left( M_2 - M_1 + M_1 \log\left[\frac{M_2}{M_1}\right] \right) \equiv \gamma_0. \quad (3.85)$$

Of course, this result vanishes for light shock waves, with  $M_2 \approx M_1$ , and we recover the result in eq. (3.84). We can also use eq. (3.77) to determine the linear growth for larger values of  $t_w$ ,

$$\Delta\mathcal{C}_A \simeq \Delta\mathcal{C}_{\text{NS}} + \frac{2}{\pi} (M_2 + M_1) (t_w - t_{\text{del}}) + \mathcal{O}(t_w e^{-t_w}) \quad (3.86)$$



**Figure 9.** The ‘delay’ time (3.87) as a function of the energy in the shock wave. For  $w \sim 1$ , we have a line of slope  $-1$ , which is characteristic of the scrambling time in eq. (3.67), as expected from eq. (3.88). For heavy shock waves,  $t_{\text{del}}$  approaches a constant proportional to  $1/T_2$ , as shown in eq. (3.89).

where the generalized ‘delay’ time can be written as

$$t_{\text{del}} = \frac{w+1}{2\pi T_2 (w^2+1)} \left[ (2w-1) \log\left(\frac{2}{w-1}\right) + \frac{2w^2}{w+1} \log w + \log\left(\frac{2}{w+1}\right) \right]. \quad (3.87)$$

It is straightforward to see that this expression reduces to the scrambling time (3.67) when  $w \rightarrow 1$ . More precisely, with  $w = 1 + \epsilon$  and  $\epsilon \ll 1$ ,

$$t_{\text{del}} = t_{\text{scr}}^* + \frac{1}{4\pi T_1} \left( 1 + 5 \log \frac{2}{\epsilon} \right) \epsilon + \mathcal{O}(\epsilon^2 \log \epsilon). \quad (3.88)$$

Eq. (3.86) provides an extension to general values of  $w$  of eq. (3.83), which applies only for light shock waves (i.e.,  $w \simeq 1$ ), and in this sense,  $t_{\text{del}}$  replaces  $t_{\text{scr}}^*$  for general shocks. Roughly, we can think of this time as characterizing when there is a transition between the early time behaviour given in eq. (3.85) and the late time behaviour given in eq. (3.86). As shown in figure 7, this is a sharp transition between two distinct regimes for light shock waves, but not for the heavy shock waves. In the latter case, there is not an extended period of time where eq. (3.85) applies. In any event, considering  $t_{\text{del}}$  for large values of  $w$ , we find to leading order

$$t_{\text{del}} = \frac{1}{\pi T_2} \left[ \log 2 - \frac{1}{w} \log w + \frac{1 + \log 2}{w} + \mathcal{O}(w^{-2}) \right], \quad (3.89)$$

that is, the delay time is simply a constant proportional to the inverse temperature of the final black hole, which is then a small time in the limit of large  $w$ .

Figure 8 shows the variation of  $d\Delta\mathcal{C}_A/dt_w$  for a light and a heavy shock wave. We see that for the light shock wave, the slope vanishes initially but then rapidly rises to the final constant value at  $t \simeq t_{\text{scr}}^*$ , corresponding to the two regimes shown in the left panel of figure 7 — see also eq. (3.84). Instead, for the heavy shock wave, the slope is initially nonvanishing and proportional to  $\gamma_0$  in equation (3.85) and rises quickly to the

final constant value, again in agreement with the results shown in the right panel of figure 7. Figure 9 shows  $t_{\text{del}}$  as a function of  $\log\left(\frac{w-1}{2}\right)$ . For  $w \sim 1$ , a line of slope  $-1$  appears since  $t_{\text{del}} \simeq t_{\text{scr}}^*$ , as shown in eq. (3.88). For heavy shock waves,  $t_{\text{del}}$  approaches a constant proportional to  $1/T_2$ , as shown in eq. (3.89).

### 3.3 ‘Complexity’ without the counterterm

In this section, we turn our attention to the effects of dropping the counterterm (3.3) from the bulk action (3.1). For stationary spacetimes, the WDW action does not seem to be effected in an important way if this surface term is not included. However, in studying holographic complexity for the formation of a black hole [1], we recently found that the counterterm is an essential ingredient for the CA proposal. The most dramatic effect of dropping the counterterm was found for  $d = 2$  (and  $k = +1$ ) where, without the counterterm, the holographic complexity actually decreased throughout the black hole formation process and the rate of change only approached zero for asymptotically late times. In the following, we show that without the counterterm, the holographic calculations fail to reproduce the expected late time growth rate and that the complexity of formation in  $d = 2$  does not exhibit the behaviour that is characteristic of the switchback effect.

Note that without the counterterm (3.3), we must deal with the ambiguities associated with the surface and joint terms on the null boundaries of the WDW patch. We follow the standard prescription proposed in [21] where we set  $\kappa = 0$  by choosing affine parametrization for the null normals. Further, we fix the overall normalization of these null vectors with  $\hat{t} \cdot k = \pm\alpha$  (where  $\hat{t}$  is the asymptotic Killing vector producing time flow in the boundary). Of course, we have already adopted these conventions in the previous sections and so it is straightforward to simply drop the counterterm contributions in eqs. (3.31)–(3.33) (and implicitly, also eq. (A.10)) from the previous analysis. We tentatively denote the resulting quantity as ‘complexity’ i.e.,

$$\tilde{\mathcal{C}}_A = \frac{I_{\text{WDW}} - I_{\text{ct}}}{\pi}, \quad (3.90)$$

but as in [1], we will find that this gravitational observable fails to behave in the manner expected of complexity.

#### 3.3.1 Time evolution

Here, we evaluate the growth rate of  $\tilde{\mathcal{C}}_A$  for the three different regimes described in eq. (3.43). First, in regime I (i.e.,  $-t_{c0} < t < t_{c1}$ ), the total rate of change of complexity only receives contributions from the bulk term in eq. (3.10) (with  $r_m = 0$ ), and from the GHY surface terms in eqs. (3.14) and (3.18) at the past and future singularities, respectively. The growth rate then becomes

$$\begin{aligned} \frac{d\tilde{\mathcal{C}}_A^{(1)}}{dt} &= -\frac{M_2}{(d-1)\pi} \frac{r_s^d}{\omega_2^{d-2} L^2} \left(1 - \frac{f_2(r_s)}{f_1(r_s)}\right) + \frac{M_1}{(d-1)\pi} \frac{r_b^d}{\omega_1^{d-2} L^2} \left(1 - \frac{f_1(r_b)}{f_2(r_b)}\right) \\ &+ \frac{dM_2}{2(d-1)\pi} \left(1 + \frac{f_1(r_b)}{f_2(r_b)}\right) - \frac{dM_1}{2(d-1)\pi} \left(1 + \frac{f_2(r_s)}{f_1(r_s)}\right). \end{aligned} \quad (3.91)$$



In regime II (i.e.,  $t_{c1} < t < t_{c2}$ ), the WDW patch has lifted off of the past singularity and so in addition to the bulk contribution (3.10) and the GHY contribution (3.14) from the future singularity, we also have the joint contribution at  $r_m$  in eq. (3.30). Combining these then yields

$$\begin{aligned} \frac{d\tilde{\mathcal{C}}_A^{(\text{II})}}{dt} = & -\frac{M_2}{(d-1)\pi} \frac{r_s^d}{\omega_2^{d-2} L^2} \left(1 - \frac{f_2(r_s)}{f_1(r_s)}\right) + \frac{M_1}{(d-1)\pi} \frac{r_b^d}{\omega_1^{d-2} L^2} \left(1 - \frac{f_1(r_b)}{f_2(r_b)}\right) \\ & + \frac{dM_2}{2(d-1)\pi} \left(1 + \frac{f_1(r_b)}{f_2(r_b)}\right) + \frac{(d-2)M_1}{2(d-1)\pi} \left(1 + \frac{f_2(r_s)}{f_1(r_s)}\right) \\ & + \frac{M_1}{\pi(d-1)} \frac{r_m^{d-1} r_s}{\omega_1^{d-2} L^2} \left[ \left(1 - \frac{f_2(r_s)}{f_1(r_s)}\right) + \frac{(d-1)L^2}{2r_s^d} \left(\omega_2^{d-2} - \omega_1^{d-2} \frac{f_2(r_s)}{f_1(r_s)}\right) \right] \\ & + \frac{M_1}{2\pi} \frac{r_m^{d-2}}{\omega_1^{d-2}} f_1(r_m) \left(1 + \frac{f_2(r_s)}{f_1(r_s)}\right) \log \left[ \frac{|f_1(r_m)|}{\alpha^2} \frac{f_2(r_s)}{f_1(r_s)} \right]. \end{aligned} \quad (3.92)$$

In the final regime III (i.e.,  $t > t_{c2}$ ), the relevant contributions are: the bulk term given by eq. (3.10) (with  $r_b = 0$ ), the GHY contribution from the future singularity given by eq. (3.16) and the joint term at  $r_m$  given by eq. (3.30). The rate of change of the complexity in this regime is

$$\begin{aligned} \frac{d\tilde{\mathcal{C}}_A^{(\text{III})}}{dt} = & -\frac{M_2}{(d-1)\pi} \frac{r_s^d}{\omega_2^{d-2} L^2} \left(1 - \frac{f_2(r_s)}{f_1(r_s)}\right) + \frac{d(M_2 + M_1)}{2(d-1)\pi} + \frac{(d-2)M_1}{2(d-1)\pi} \left(1 + \frac{f_2(r_s)}{f_1(r_s)}\right) \\ & + \frac{M_1}{\pi(d-1)} \frac{r_m^{d-1} r_s}{\omega_1^{d-2} L^2} \left[ \left(1 - \frac{f_2(r_s)}{f_1(r_s)}\right) + \frac{(d-1)L^2}{2r_s^d} \left(\omega_2^{d-2} - \omega_1^{d-2} \frac{f_2(r_s)}{f_1(r_s)}\right) \right] \\ & + \frac{M_1}{2\pi} \frac{r_m^{d-2}}{\omega_1^{d-2}} f_1(r_m) \left(1 + \frac{f_2(r_s)}{f_1(r_s)}\right) \log \left[ \frac{|f_1(r_m)|}{\alpha^2} \frac{f_2(r_s)}{f_1(r_s)} \right]. \end{aligned} \quad (3.93)$$

**Early and late time behaviours.** Of course, the critical times in the time evolution depend only on the background geometry and so these basic features in the time evolution remain unchanged if we choose to study  $\tilde{\mathcal{C}}_A$ , without the counterterm contributions. However, if the shock wave was injected early enough, there were two clear plateaus in  $d\mathcal{C}_A/dt$  (which included the counterterm contributions), given by eqs. (3.52) and (3.53). So we examine to see to what extent these plateaus arise for  $d\tilde{\mathcal{C}}_A/dt$ .

The first plateau is found in the regime of large  $t_w$ , such that  $r_s$  is very close to  $r_{h,2}$  and  $r_b$  to  $r_{h,1}$ . In this limit, the growth rate in eq. (3.91) becomes

$$\begin{aligned} \left. \frac{d\tilde{\mathcal{C}}_A^{(1)}}{dt} \right|_{t_w \rightarrow \infty} = & \frac{d-2}{2(d-1)\pi} (M_2 - M_1) + \frac{kz^2}{(d-1)\pi} \left( \frac{M_2}{1+kz^2} - \frac{w^2 M_1}{1+kw^2z^2} \right) \\ & + \mathcal{O} \left( e^{-\pi T_1(2t_w-t)} \right). \end{aligned} \quad (3.94)$$

Comparing to eq. (3.52), we see that here we also have a similar plateau with the rate being proportional to  $(M_2 - M_1)$ , at least for  $k = 0$ , but in general there are curvature corrections to this result. Further note that for the BTZ black hole (i.e.,  $d = 2$ ), the time derivative is always zero, irrespective of the shock wave energy

$$\left. \frac{d\tilde{\mathcal{C}}_A^{(1)}}{dt} \right|_{d=2} = 0, \quad (3.95)$$

because  $k$  does not play a role in the BTZ geometries.

The late time limit, analogously to that in eq. (3.53), is approached as  $r_s$  is close to  $r_{h,2}$  and  $r_m$  close to  $r_{h,1}$ . However, without the inclusion of the counterterm, there are further considerations if the shock wave is light. If we denote the ratio of horizons as  $w = 1 + \epsilon$ , with  $\epsilon$  small, there are two regimes to consider, that depend on a time scale related to the scrambling time in eq. (3.67), defined as

$$\hat{t}^* = \frac{1}{\pi T_1} \log \frac{2}{\epsilon} - 2t_w = 2t_{\text{scr}}^* - 2t_w. \quad (3.96)$$

If the late time regime is such that  $t > \hat{t}^*$ , then we can evaluate eq. (3.93) for  $x_m$  and  $x_s$  approaching 1, which yields

$$\begin{aligned} \left. \frac{d\tilde{\mathcal{C}}_A^{(\text{III})}}{dt} \right|_{t \rightarrow \infty, t > \hat{t}^*} &= \frac{M_1}{\pi} \left( 1 + \frac{w(d+1)}{2(d-1)} \right) + \frac{d-2}{2(d-1)} \frac{M_2}{\pi} + \frac{kz^2}{(d-1)(1+kz^2)} \frac{M_2}{\pi} \\ &\quad - \frac{kz^2 w}{2} \frac{(d+1)w^2 - (d-1)}{(d-1)(1+kw^2z^2)} \frac{M_1}{\pi} + \mathcal{O} \left( T_2 t e^{-\pi T_1(t-2t_w)} \right). \end{aligned} \quad (3.97)$$

In contrast to eq. (3.53), the late time rate here is not proportional to the expected sum of the masses, even for the planar horizons ( $k = 0$ ) or in the limit of light but still non-zero shocks ( $w \sim 1$ ). For simplicity, let's rewrite eq. (3.97) for planar black holes ( $k = 0$ ) and light shocks, such that  $M_2 \approx M_1$ . The late time limit then reads

$$\left. \frac{d\tilde{\mathcal{C}}_A^{(\text{III})}}{dt} \right|_{t \rightarrow \infty, t > \hat{t}^*} = \frac{2M_1}{\pi} \left( 1 + \frac{1}{4(d-1)} \right) + \mathcal{O} \left( \frac{T_2 t}{w-1} e^{-\pi T_1(t-2t_w)} \right). \quad (3.98)$$

If one wants to consider a shock wave with exactly zero energy, such that  $w = 1$ , then  $\hat{t}^*$  given by eq. (3.96) goes to infinity, which is equivalent to a regime where  $t < \hat{t}^*$ . This is equivalent to setting  $w = 1$  in eq. (3.93), which simplifies to

$$\left. \frac{d\tilde{\mathcal{C}}_A^{(\text{III})}}{dt} \right|_{w=1} = \frac{2M_1}{\pi} + \frac{M_1}{2\pi} \frac{r_m^{d-2}}{\omega_1^{d-2}} f_1(r_m) \log \left[ \frac{|f_1(r_m)|}{\alpha^2} \right], \quad (3.99)$$

and is simply the rate of change of the eternal black hole discussed in [23]. This demonstrates that the order of limits does not commute.

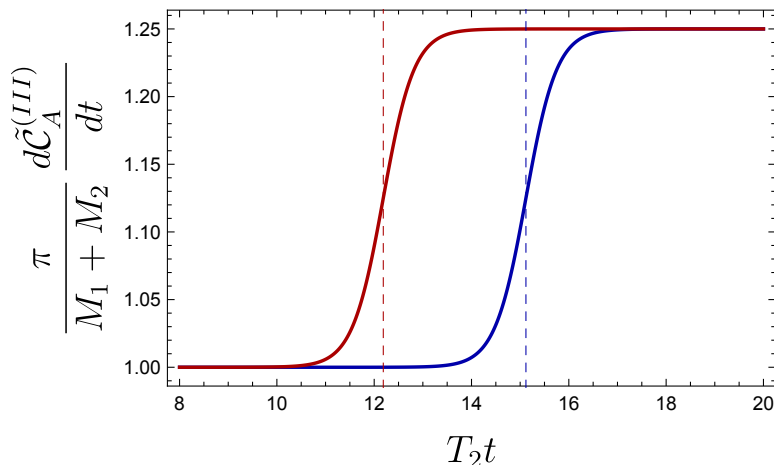
In addition, the heavy shock wave regime of the rate of change given by eq. (3.97) can be calculated by considering the limit  $w \rightarrow \infty$ . The rate of change becomes then, for  $k = 0$  for simplicity,

$$\left. \frac{d\tilde{\mathcal{C}}_A^{(\text{III})}}{dt} \right|_{w \rightarrow \infty, t \rightarrow \infty} = \frac{M_2}{2\pi} \left( \frac{(d-2)}{(d-1)} + \mathcal{O} \left( \frac{T_2 t}{w^{d-1}} e^{-\pi T_1(t-2t_w)} \right) \right), \quad (3.100)$$

which as expected is half of the one sided collapse value without the inclusion of the counterterm in eq. (3.29) of [1], and it is vanishing for BTZ ( $d=2$ ).

Consider as an example the BTZ black hole, with  $k = 0$  and  $d = 2$ . The late time regime for  $t > \hat{t}^*$  reads

$$\left. \frac{d\tilde{\mathcal{C}}_A^{(\text{III})}}{dt} \right|_{d=2, t \rightarrow \infty} = \frac{1}{\pi} \left( M_1 + \frac{3}{2} \sqrt{M_2 M_1} \right) + \mathcal{O} \left( \frac{T_2 t}{w-1} e^{-\pi T_1(t-2t_w)} \right), \quad (3.101)$$



**Figure 10.** Rate of change of complexity for BTZ ( $d = 2$ ) black hole without the addition of the counterterm, given by eq. (3.93). We study the transition between late time rates for very light shock waves, with  $w = 1 + 10^{-33}$  (red) and  $w = 1 + 10^{-37}$  (blue). The vertical lines represent the characteristic transition times  $\hat{t}^*$ , given by eq. (3.96). For light but non-zero shock waves, the late time limit is similar to the eternal black hole for  $t < \hat{t}^*$ , but for  $t > \hat{t}^*$ , it becomes the rate in eq. (3.101) (with  $M_1 \approx M_2$ ).

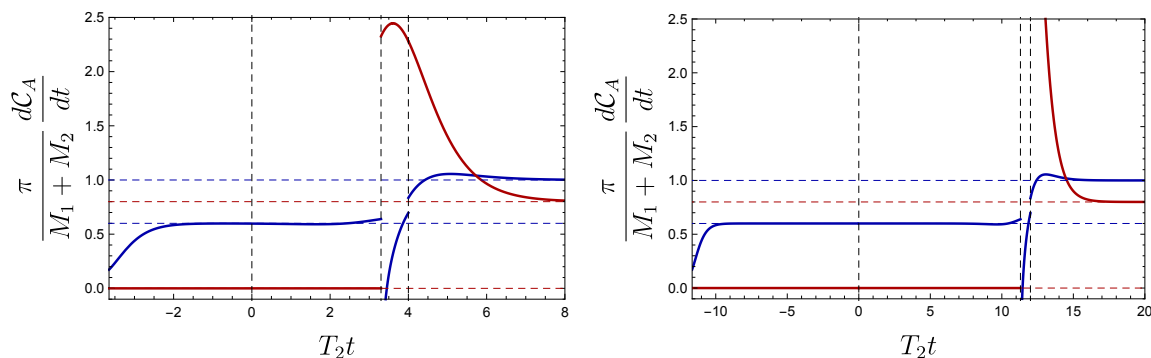
where again we substituted  $w^2 = M_2/M_1$  for  $d = 2$ . This expression again fails to produce the expected late time limit but we also see an unusual nonlinear dependence on the masses, i.e.,  $\sqrt{M_2 M_1}$ .

In figure 10, we investigate these limits for a very light but non-zero shock wave. We show the growth rate for BTZ black holes without the addition of the counterterm given by eq. (3.93), with  $w = 1 + 10^{-33}$  (red) and  $w = 1 + 10^{-37}$  (blue) and the vertical line representing  $\hat{t}^*$  as in eq. (3.96). For times which are large but smaller than  $\hat{t}^*$ , the effective late time limit is the same as the eternal black hole, but at  $t \simeq \hat{t}^*$ , there is a sharp transition to the late time limit of eq. (3.101) (with  $M_1 \approx M_2$ ).

We compare the rate of change for BTZ with and without the inclusion of the counterterm in figure 11. We focus on heavy shock waves ( $w = 2$ ), since for early times (before  $t_{c1}$ ) there is a bigger discrepancy of rates, i.e., vanishing without the counterterm or proportional to  $M_2 - M_1$  with the counterterm. In addition, immediately after  $t_{c1}$  there is a large positive peak in the rate without counterterm, due to a factor of  $f_2(r_s)$  in eq. (3.92), which approaches zero much faster than  $f_1(r_m)$  for early shocks, since the exponent of  $r_s$  approaching  $r_{h,2}$  at late times is proportional to  $T_2(t + 2t_w)$ , while  $r_m$  approaching  $r_{h,1}$  is controlled by an exponent proportional to  $T_1(t - 2t_w)$  for late times. Finally, the late time limit of the rate of change without the inclusion of the counterterm, as discussed previously, approaches the late time limit given by eq. (3.101), which does not reduce to the eternal black hole result for light shock waves.

### 3.3.2 Complexity of formation

Turning now to the complexity of formation (3.70) but evaluated with eq. (3.90), i.e., without the counterterm contribution (3.75). We consider the BTZ black hole (i.e.,  $d = 2$ )



**Figure 11.** Rate of change of complexity evolving both boundaries as  $t_L = t_R = \frac{t}{2}$ ,  $w = 2$ ,  $z = 1/w$ ,  $\tilde{\ell}_{ct} = 1$  and  $\alpha = 1$ . In the left we evaluate  $T_2 t_w = 2$  and in the right  $T_2 t_w = 6$ , and the blue curve is the rate of change with the inclusion of counterterm, while the red line is the rate of change without it. Despite being a shock wave that doubles the temperature, the rate of change is exactly zero without the inclusion of the counterterm for  $t_{c0} < t < t_{c1}$ , as opposed to being proportional to the difference of masses. In addition, there is a large positive peak after  $t_{c1}$  for the red curve, in contrast to the (short) negative spike of the blue curve. The peak in the red curve in the right figure is similar to the one in the left, but is sharper and reaches higher values the earlier the shock wave is sent. Finally, the late time limit is given by eq. (3.101), in contrast to  $(M_1 + M_2)/\pi$ , as discussed in section 3.1.

as a simple example. In this case, our previous result (3.77) is replaced with

$$\frac{\Delta\tilde{\mathcal{C}}_A - \Delta\mathcal{C}_{NS}}{|\Delta\mathcal{C}_{NS}|} = LT_1 \left[ 2 \coth^{-1}(wx_s) - \frac{2}{w} (w^2 - 1) \coth^{-1}(x_s) + (w^2 - x_b^2) \log\left(\frac{1 + x_b}{1 - x_b}\right) - w \log\left(\frac{w + x_b}{w - x_b}\right) + \frac{2x_b^2}{w} \tanh^{-1}\left(\frac{x_b}{w}\right) - \log\left(\frac{wx_s + 1}{wx_s - 1}\right) \right]. \quad (3.102)$$

We numerically evaluate eq. (3.102) and plot the complexity of formation as a function of the insertion time for light and heavy shock waves in figure 12. For both light and heavy shock waves, the complexity of formation just approaches a constant value for large  $t_w$ . Further, the latter value is less than the original  $\Delta\mathcal{C}_{NS} = -c/3$  found with  $t_w = 0$  (or alternatively, with no shock wave). Note that the transition to the final  $\Delta\tilde{\mathcal{C}}_A$  essentially saturates after  $T_1 t_w \sim 1$ . We can evaluate the large  $t_w$  limit of eq. (3.102) analytically to find

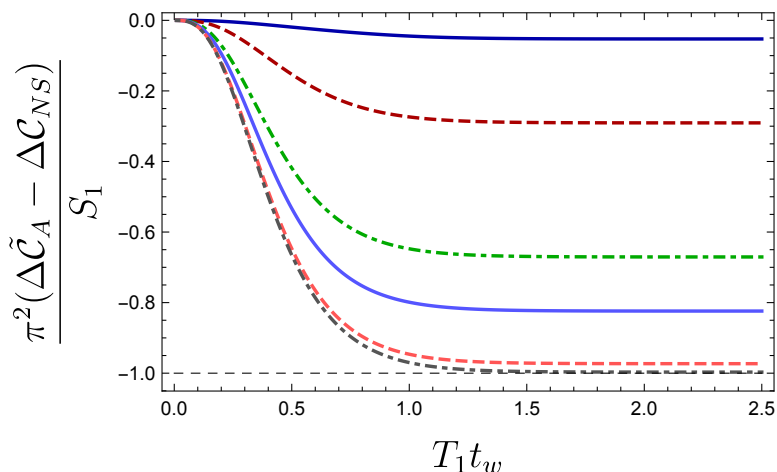
$$\frac{\Delta\tilde{\mathcal{C}}_A - \Delta\mathcal{C}_{NS}}{|\Delta\mathcal{C}_{NS}|} \Big|_{t_w \rightarrow \infty} = -LT_1 \frac{w^2 - 1}{w} \log\left[\frac{w + 1}{w - 1}\right] + \mathcal{O}(T_2 t_w e^{-2\pi T_1 t_w}). \quad (3.103)$$

For light shock waves, the above difference vanishes as  $w \rightarrow 1$ , i.e.,  $\Delta\tilde{\mathcal{C}}_A \rightarrow \Delta\mathcal{C}_{NS}$ . On the other hand, for very heavy shocks (i.e.,  $w \rightarrow \infty$ ), we find

$$\Delta\tilde{\mathcal{C}}_A \Big|_{w \rightarrow \infty} = \Delta\mathcal{C}_{NS} - S_1/\pi^2. \quad (3.104)$$

That is, for heavy shocks injected at early times, the complexity of formation decreases and it does so in a way that only depends on the initial black hole and not on the final black hole.

Of course, these results for  $\Delta\tilde{\mathcal{C}}_A$  contrast with the previous results (including the counterterm) in figure 7, where the complexity of formation began to grow linearly for

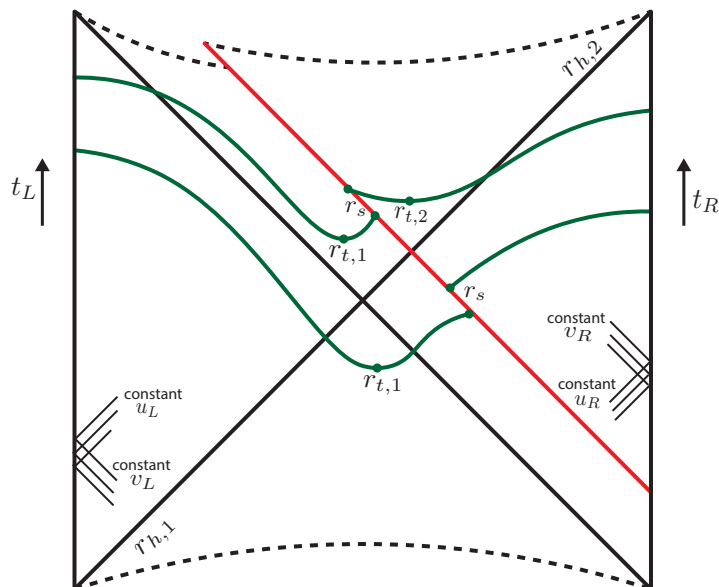


**Figure 12.** Complexity of Formation for BTZ black holes with no counterterm added to the null boundaries (and  $\kappa = 0$ ). From top to bottom, we consider light to heavy shock waves, with  $w = 1 + 10^{-2}$ ,  $w = 1 + 10^{-1}$ ,  $w = 1.5$ ,  $w = 2$ ,  $w = 5$  and  $w = 15$ . For  $T_1 t_w$  of order 1, the complexity of formation with respect to the unperturbed one saturates to a constant, being close to zero for light shocks and  $-S_1/\pi^2$  for heavy ones, represented by the horizontal black dashed line, as given by eq. (3.104).

large  $t_w$  for both light and heavy shock waves. Therefore, if one studies the complexity of formation without the inclusion of the counterterm, there is no dependence on the scrambling time or on how early the shock wave was inserted. As we discuss in section 5, this means that  $\tilde{C}_A$  fails to exhibit the switchback effect for  $d = 2$ . Of course, this only strengthens the argument that this gravitational observable cannot be interpreted in term of complexity in the boundary theory.

#### 4 Complexity = volume

In this section, we apply the complexity=volume conjecture (1.1) to evaluate the rate of change of complexity and also the complexity of formation for the shock wave backgrounds described in section 2. These geometries describe a transition between a black hole of mass  $M_1$  and radius  $r_{h,1}$  and a black hole of mass  $M_2$  and radius  $r_{h,2}$  caused by an incoming shock wave from the right boundary at  $v_R = v_s = -t_w$ . The relevant setup is illustrated in figure 13, which is again a ‘Penrose-like’ diagram built in such a way that lines of constant  $v$ ,  $u$  and  $t$  look continuous in the diagram while lines of constant  $r$  do not. To simplify the discussion, we will denote the (black hole) region before the transition by  $BH_1$  and the one after the transition by  $BH_2$  in what follows. Note that in this section, we use  $r_s$  to denote the point in which the maximal surfaces cross the shell. We will be interested in surfaces that are anchored on the asymptotic boundaries at times  $t_L$  and  $t_R$ . The synchronization of the two boundary times was already explained in the paragraph above eq. (2.18). Much of the following discussion borrows from our analysis of CV complexity in [1] and we refer the interested reader there for further details.



**Figure 13.** Illustration of maximal volume surfaces in the shock wave geometry. Some special points are indicated in the drawing:  $r_{t,1}$  the turning point in  $BH_1$  (if it exists),  $r_{t,2}$  the turning point in  $BH_2$  (if it exists) and  $r_s$  the point in which the maximal volume surface crosses the shell. We have also indicated in the figure lines of constant  $u$  and  $v$  induced from either boundary as well as the direction of the time coordinates on the left and right boundaries.

To describe the full geometry, it will be sometimes useful to replace the coordinate  $v$  with  $u = v - 2r^*(r)$  (as in eq. (2.17)).<sup>23</sup> In each of the two spacetime regions, the metric (2.7) reads

$$ds^2 = -f_i(r)dv^2 + 2drdv + r^2 d\Sigma_{k,d-1}^2 = -f_i(r)du^2 - 2drdu + r^2 d\Sigma_{k,d-1}^2 \quad (4.1)$$

with  $f_i(r)$ , the appropriate blackening factor given in eqs. (2.10) and (2.11). The maximal volume surfaces then extremize the following functional

$$\mathcal{V} = \Omega_{k,d-1} \int d\lambda r^{d-1} \sqrt{-f\dot{v}^2 + 2\dot{v}\dot{r}} = \Omega_{k,d-1} \int d\lambda r^{d-1} \sqrt{-f\dot{u}^2 - 2\dot{u}\dot{r}}, \quad (4.2)$$

where taking advantage of the symmetry of boundary, these directions have been integrated out. Further,  $\lambda$  is the radial coordinate intrinsic to the surface, which will increase along the surface moving from the left asymptotic AdS boundary to the right boundary, i.e., from left to right in figure 13. We fix the reparametrization symmetry of the volume functional with the convenient gauge choice [1]:

$$\sqrt{-f\dot{v}^2 + 2\dot{v}\dot{r}} = \sqrt{-f\dot{u}^2 - 2\dot{u}\dot{r}} = r^{d-1}. \quad (4.3)$$

In addition, to cover the full spacetime, we will need to work with the coordinates  $v_R, u_R$  and  $v_L, u_L$  induced from the right/left boundaries, respectively (see figure 13). Recall

<sup>23</sup>It will also be convenient to distinguish the Eddington-Finkelstein coordinates,  $v_R$  and  $u_R$ , specifying null rays ending on the right AdS boundary from  $v_L$  and  $u_L$ , anchored to the left AdS boundary.

from eqs. (2.15) and (2.16), we choose the constant of integration in the definition of the tortoise coordinate such that  $r_\infty^* = 0$ . The boundary times  $t_L$  and  $t_R$  run upwards on their respective boundaries according to our conventions and hence on the left boundary, we have  $v_L = u_L = -t_L$  while on the right boundary,  $v_R = u_R = t_R$ .

Of course, the metric profile (2.8) makes a sharp transition at  $v_R = v_s$  but otherwise the geometry is independent of  $v$ . Hence in the  $BH_1$  and  $BH_2$  regions separately, there is a conserved ‘ $v$ -momentum’  $P_i$ ,

$$P_i = \frac{\partial \mathcal{L}}{\partial \dot{v}} = \frac{r^{d-1}(\dot{r} - f_i(r)\dot{v})}{\sqrt{-f_i(r)\dot{v}^2 + 2\dot{v}\dot{r}}} = \frac{r^{d-1}(-\dot{r} - f_i(r)\dot{u})}{\sqrt{-f_i(r)\dot{u}^2 - 2\dot{u}\dot{r}}}, \quad (4.4)$$

which with the above gauge choice takes the form

$$P_i = \dot{r} - f_i(r)\dot{v} = -\dot{r} - f_i(r)\dot{u}. \quad (4.5)$$

Using eqs. (4.3) and (4.5) to solve for  $\dot{r}$  and  $\dot{v}$  (as well as  $\dot{u}$ ), we find

$$\begin{aligned} \dot{r}_\pm [P_i, r] &= \pm \sqrt{f_i(r)r^{2(d-1)} + P_i^2}, \\ \dot{v}_\pm [P_i, r] &= \frac{\dot{r} - P_i}{f_i(r)} = \frac{1}{f_i(r)} \left( -P_i \pm \sqrt{f_i(r)r^{2(d-1)} + P_i^2} \right), \\ \dot{u}_\pm [P_i, r] &= -\frac{\dot{r} + P_i}{f_i(r)} = \frac{1}{f_i(r)} \left( -P_i \mp \sqrt{f_i(r)r^{2(d-1)} + P_i^2} \right). \end{aligned} \quad (4.6)$$

Here, we have added the  $\pm$  subscripts to indicate whether  $r$  is increasing/decreasing as we move along the surface from left to right.

An intuitive picture of the dynamics of the extremal surfaces is given by recasting the  $\dot{r}$  equation as a Hamiltonian constraint in each black hole region,

$$\dot{r}^2 + U_i(r) = P_i^2, \quad (4.7)$$

where the effective potential is given by

$$U_i(r) = -f_i(r)r^{2(d-1)}. \quad (4.8)$$

In this framework,  $P_i^2$  plays the role of the conserved energy. We will come back to this classical Hamiltonian picture when studying specific examples below.

Now the maximal surface is determined by fixing two boundary conditions, namely, the times  $t_L$  and  $t_R$  at which it is anchored on the two asymptotic AdS boundaries. However, it will be more convenient to trade these for fixing  $P_1$ , the conserved momentum in the  $BH_1$  region, and  $r_s$ , the radius where the surface crosses the shock wave.<sup>24</sup> For given values of  $P_1$  and  $r_s$ , we can obtain  $\dot{v}(r_s)$  and  $\dot{r}_1(r_s)$  by evaluating eq. (4.6) at  $r = r_s$  and with  $i = 1$ :

$$\begin{aligned} \dot{r}_1(r_s) &= \pm \sqrt{f_1(r_s)r_s^{2(d-1)} + P_1^2}, \\ \dot{v}_1(r_s) &= \frac{1}{f_1(r_s)} \left( -P_1 \pm \sqrt{f_1(r_s)r_s^{2(d-1)} + P_1^2} \right). \end{aligned} \quad (4.9)$$

<sup>24</sup>To be precise, we also have to specify the sign of  $\dot{r}_1(r_s)$  when the surface reaches the shell.

The second order equations of motion derived from eq. (4.2) can be integrated in the vicinity of the shock wave to conclude that  $\dot{v}$  is continuous across the shell (i.e.,  $\dot{v}_2(r_s) = \dot{v}_1(r_s)$ ) while  $\dot{r}$  undergoes a jump given by

$$\dot{r}_2(r_s) = \dot{r}_1(r_s) + \frac{\dot{v}_1(r_s)}{2} (f_2(r_s) - f_1(r_s)). \quad (4.10)$$

Of course, from eq. (4.5), the conserved momentum must also jump when crossing the shell and we find

$$P_2 = \dot{r}_2(r_s) - f_2(r_s) \dot{v}_2(r_s). \quad (4.11)$$

There are a number of different scenarios to be considered for the shape of the surface, namely the surface may pass into the black/white hole region of  $\text{BH}_1$  and it may or may not admit a *turning point* (a point where  $\dot{r}$  changes sign) in  $\text{BH}_1$  and/or in  $\text{BH}_2$ . The conserved momentum  $P_1$  in  $\text{BH}_1$  is positive when the surfaces pass into the black hole part of spacetime and negative when the surfaces pass into the white hole region. The reason for this becomes obvious when expressing the conserved momentum in terms of an auxiliary time coordinate behind the horizon since in the black/white hole part this coordinate has to increase/decrease along the surface respectively. Of course, when  $P_1 = 0$  the maximal surface in the  $\text{BH}_1$  side is a line of constant time going through the bifurcation surface. When turning points exist (where  $\dot{r} = 0$ ), their positions can be obtained from eq. (4.7) by solving

$$P_i^2 + f_i(r_{t,i}) r_{t,i}^{2(d-1)} = 0, \quad (4.12)$$

where  $r_{t,1}$  denotes the turning point in the  $\text{BH}_1$  region (if it exists) and  $r_{t,2}$ , the turning point in  $\text{BH}_2$  (if it exists). Of course when passing in the white hole region of  $\text{BH}_1$ , we must at least admit a turning point in  $\text{BH}_1$ .

Given our boundary conditions  $r_s$  and  $P_1$ , we are able to integrate out to the left boundary to determine  $t_L$ , while with  $P_2$  fixed by eq. (4.11), we can similarly determine  $t_R$ . Here we will simply present the expressions for the left and right boundary times, as well as the volume and its time derivative, for the various cases described above. The interested reader is referred to appendix C for the full derivation. To simplify these expressions, we also define

$$\tau[P, r] \equiv \frac{1}{f(r)} - \frac{P}{f(r)\sqrt{f(r)r^{2(d-1)} + P^2}}, \quad R[P, r] \equiv \frac{r^{2(d-1)}}{\sqrt{f(r)r^{2(d-1)} + P^2}}. \quad (4.13)$$

and we will add subscripts ( $\tau_{1,2}$ ,  $R_{1,2}$ ) to specify which blackening factor (alternatively, which horizon radius) is being used. For the right boundary time  $t_R$ , we obtain

$$t_R + t_w = \begin{cases} \int_{r_{t,2}}^{\infty} \tau_2[P_2, r] dr - \int_{r_{t,2}}^{r_s} \tau_2[-P_2, r] dr & \text{turning point in BH}_2 \\ \int_{r_s}^{\infty} \tau_2[P_2, r] dr & \text{otherwise.} \end{cases} \quad (4.14)$$

For the left boundary time  $t_L$ , we obtain

$$t_L - t_w = \begin{cases} 2r_1^*(r_{t,1}) + \int_{r_{t,1}}^{\infty} \tau_1[P_1, r] dr + \int_{r_{t,1}}^{r_s} \tau_1[P_1, r] dr & P_1 > 0 \text{ (black hole),} \\ & \text{turning point in BH}_1 \\ 2r_1^*(r_s) + \int_{r_s}^{\infty} \tau_1[P_1, r] dr & P_1 > 0 \text{ (black hole),} \\ & \text{no turning point in BH}_1 \\ 2r_1^*(r_s) - 2r_1^*(r_{t,1}) - \int_{r_{t,1}}^{\infty} \tau_1[-P_1, r] dr - \int_{r_{t,1}}^{r_s} \tau_1[-P_1, r] dr & P_1 < 0 \text{ (white hole)} \\ r_1^*(r_s) & P_1 = 0. \end{cases} \quad (4.15)$$



We note that the integrals are well behaved near the black hole horizons due to a cancellation of logarithmic divergences coming from integrating separately the two parts of the expression for  $\tau[P, r]$ . We also note that the only combinations of times which appear here are  $t_R + t_w$  and  $t_L - t_w$ . This is a consequence of the time-shift symmetry that was discussed in the introductory remarks (i.e., see discussion after eq. (2.5)) and hence the same behaviour occurred in the CA subsection (e.g., see eq. (3.4)).

The volume of the maximal surface will in general be given by the sum of the relevant volumes in  $\text{BH}_1$  and  $\text{BH}_2$ , which we denote by  $\mathcal{V}_1$  and  $\mathcal{V}_2$

$$\mathcal{V} = \mathcal{V}_1 + \mathcal{V}_2 \tag{4.16}$$

where

$$\mathcal{V}_1 = \Omega_{k,d-1} \begin{cases} \int_{r_{t,1}}^{r_{\max}} R_1 [P_1, r] dr + \int_{r_{t,1}}^{r_s} R_1 [P_1, r] dr & \text{turning point in BH}_1 \\ \int_{r_s}^{r_{\max}} R_1 [P_1, r] dr & \text{otherwise} \end{cases} \tag{4.17}$$

$$\mathcal{V}_2 = \Omega_{k,d-1} \begin{cases} \int_{r_{t,2}}^{r_{\max}} R_2 [P_2, r] dr + \int_{r_{t,2}}^{r_s} R_2 [P_2, r] dr & \text{turning point in BH}_2 \\ \int_{r_s}^{r_{\max}} R_2 [P_2, r] dr & \text{otherwise} \end{cases}$$

and where  $r_{\max}$  indicates the position of the UV regulator surface.

#### 4.1 Time evolution

The time derivative of the volume admits a very simple form, common to all the cases above (see appendix C for the derivation)

$$\frac{1}{\Omega_{k,d-1}} \frac{d\mathcal{V}}{dt} = P_1 \frac{dt_L}{dt} + P_2 \frac{dt_R}{dt} \tag{4.18}$$

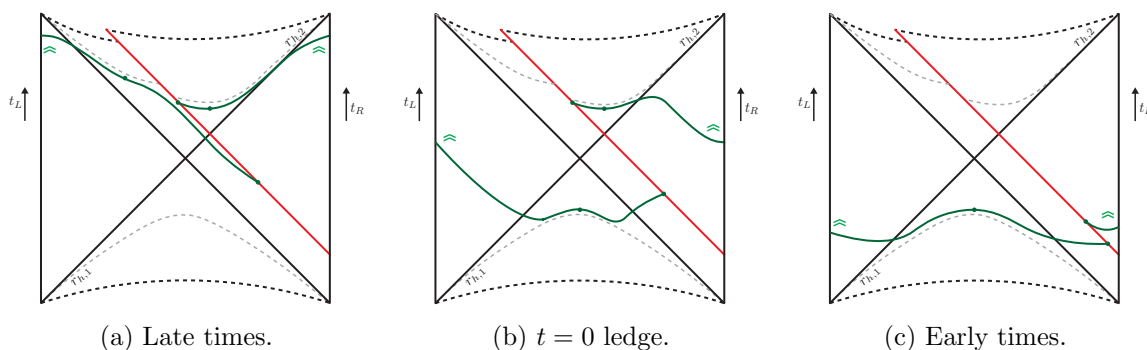
where  $t$  is a time parameter specifying the time evolution of the two boundary times according to  $t_R(t)$ ,  $t_L(t)$ . In this way we can study general patterns for the time evolution, e.g., symmetric  $t_L = t_R$  and antisymmetric  $t_L = -t_R$ . We will focus mainly on the case of symmetric boundary times with  $t_R = t_L = t/2$  in which case we obtain for the complexity using eq. (1.1) with  $\ell = L$

$$\frac{d\mathcal{C}_V}{dt} = \frac{\Omega_{k,d-1}}{2G_N L} (P_1 + P_2). \tag{4.19}$$

Before we proceed, let us note that all the above results continue to be valid for  $d = 2$  when substituting the blackening factors with those of BTZ black holes.

**Early and late time limits.** The relation between the boundary times and the momentum parameters  $P_1$  and  $P_2$  is not given by a closed analytic expression and it requires some numerical treatment as we are about to describe below. However, we were able to make the following general observations for planar black holes, as well as BTZ black holes in  $d = 2$ . In the late time limit, the momenta  $P_1$  and  $P_2$  correspond to the maximal values of the potentials (4.8) according to

$$\begin{aligned} \partial_r \left[ f_1(r_{1,m}) r_{1,m}^{2(d-1)} \right] &= 0, & f_1(r_{1,m}) r_{1,m}^{2(d-1)} + P_{1,m}^2 &= 0, \\ \partial_r \left[ f_2(r_{2,m}) r_{2,m}^{2(d-1)} \right] &= 0, & f_2(r_{2,m}) r_{2,m}^{2(d-1)} + P_{2,m}^2 &= 0, \end{aligned} \tag{4.20}$$



**Figure 14.** The shape of the surface in the various limits discussed in eqs. (4.22), (4.23), (4.25) for symmetric time evolution  $t_L = t_R = t/2$  after an early shock.

and this means that the surface extends along the critical surfaces of constant  $r = r_{1,m}$  in  $BH_1$  and  $r = r_{2,m}$  in  $BH_2$  (see figure 14a). The reason is that the expressions in eqs. (4.14)–(4.15) for the boundary times become divergent in this limit since the limit of integration  $r_{t,i}$  also approaches  $r_{i,m}$  and the expression in the denominator of eq. (4.13) becomes approximately linear in  $r - r_{i,m}$ . Hence the integrals become logarithmically divergent at this point, which leads to very large times  $t_R + t_w$ ,  $t_L - t_w$ . This argument holds both for planar and for spherical geometries. For the planar geometry these equations can be solved for analytically and yield

$$r_{1,m} = \frac{r_{h,1}}{2^{1/d}}, \quad r_{2,m} = \frac{r_{h,2}}{2^{1/d}}, \quad P_{1,m} = \frac{r_{h,1}^d}{2L}, \quad P_{2,m} = \frac{r_{h,2}^d}{2L}. \quad (4.21)$$

Substituting this back into eq. (4.19) and using the definition of the mass in eq. (2.9) we obtain

$$\frac{d\mathcal{C}_V}{dt} = \frac{4\pi(M_1 + M_2)}{d-1}. \quad (4.22)$$

And of course, for light shocks this reduces to the known result for eternal black holes upon setting  $M_1 = M_2 = M$ . In our numerical studies, we have always observed that this limit is approached from below.

If the perturbation is sent early enough we observe a period of time during which the growth rate of the complexity, centered around  $t = 0$ , is constant, cf. figure 16. The width of this ledge is approximately  $4t_w$  and it appears due to the fact that in this region we have approximately  $P_1 \approx -P_{1,m}$  and  $P_2 \approx P_{2,m}$ . The geometric picture indicates that now the surface is unwrapping the critical surface  $r = r_{1,m}$  behind the past horizon of  $BH_1$  while wrapping onto the critical surface  $r = r_{2,m}$  in  $BH_2$  (see figure 14b). The reason is again that for  $t_w$  very large and for  $t_L = t_R = 0$ , the time integrals in eqs. (4.14)–(4.15) should be very large and of opposite signs. This is indeed the case for the momenta above since the lower limit of integration approaches  $r_{i,m}$  and the denominator is approximately linear in  $r - r_{i,m}$ . Substituting these values for the momenta into eq. (4.19) and using again the definition of the mass in eq. (2.9) leads to

$$\frac{d\mathcal{C}_V}{dt} = \frac{4\pi(M_2 - M_1)}{d-1}. \quad (4.23)$$

We note that for light shock this indicates a rate of computation which is close to zero.

One last interesting limit is the one where the boundary times are only slightly above the shock, i.e.,  $t_R \approx -t_w$  (and hence  $t \approx -2t_w$ ). This scenario is demonstrated in figure 14c. In this case we still have  $P_1 \approx -P_{1,m}$ , and in addition we know that  $x_s \rightarrow \infty$ . We can now use the relations (4.9)–(4.11) to obtain the value of  $P_2$  in this limit

$$P_2 \approx P_{2,m} - 2P_{1,m}, \tag{4.24}$$

which using eqs. (4.19), (4.21) and (2.9) yields

$$\frac{d\mathcal{C}_V}{dt} = \frac{4\pi(M_1 - 3M_2)}{d - 1}. \tag{4.25}$$

This limit satisfies the consistency check that when  $M_1 = M_2 = M$  (i.e., a light shock) we recover a negative rate of

$$\frac{d\mathcal{C}_V}{dt} = -\frac{8\pi M}{d - 1}, \tag{4.26}$$

as expected for eternal black holes of mass  $M$  at early (negative) times.

For BTZ black holes, the same conclusions hold upon replacing the blackening factor with that in eq. (3.54) and the masses with those given by eq. (3.55) and of course setting  $d = 2$  in all the above equations. In particular the maximal points of the potential still correspond to those given in eq. (4.21) with  $d = 2$  and the three limits (4.22), (4.23) and (4.25) are still valid.

**Critical time.** As mentioned above, around  $t = 0$  we observe a long regime in which the rate of computation is approximately constant and proportional to the difference of the masses  $M_2 - M_1$ . We would like to estimate the duration of this regime. In our numerical solutions, we observed that this regime extends more or less equally in positive and negative times and that one can approximately identify the end of it with the transition of the surfaces from the white hole part to the black hole part of  $\text{BH}_1$  (cf. figure 16). In other words the symmetric time configuration for which  $P_1 = 0$  defines a critical time  $t_{c,v}$  in which this regime ends.<sup>25</sup> This critical time can be evaluated for planar ( $d \geq 3$ ) and BTZ geometries in closed form in the limit of early shocks by using the fact that it is associated with the following values of the surface momenta —  $P_1 \approx 0$  and  $P_2 \approx P_{2,m}$ . The geometric picture lies in between those illustrated in figures 14a and 14b. The surface inside  $\text{BH}_1$  is a straight line going through the bifurcation surface while it still wraps around the surface of constant  $r = r_{m,2}$  inside  $\text{BH}_2$ . Solving the matching condition (4.11) with those values of the surface momenta, we are able to extract the value of  $r_s$  at the critical time

$$r_{s,c} = \frac{r_{h,2}^2}{(2r_{h,2}^d - r_{h,1}^d)^{1/d}}. \tag{4.27}$$

Substituting this into the expression for  $t_L$  in eq. (4.15) with  $P_1 = 0$ , we obtain

$$t_{L,c} = t_{c,v}/2 = -v_s + r_1^*(r_{s,c}). \tag{4.28}$$

---

<sup>25</sup>Recall that with  $P_1 = 0$ , the extremal surface passes through the bifurcation surface, separating the black hole and the white hole regions.

For instance, in BTZ, using the relevant tortoise coordinate in eq. (3.56), as well as  $v_s = -t_w$  and  $w = T_2/T_1 = \sqrt{M_2/M_1}$  (see eqs. (2.18)–(2.19)), we obtain

$$t_{c,v} = 2t_w + \frac{L^2}{r_{h,1}} \log \left( \frac{r_{h,2}^2 - r_{h,1} \sqrt{2r_{h,2}^2 - r_{h,1}^2}}{r_{h,2}^2 + r_{h,1} \sqrt{2r_{h,2}^2 - r_{h,1}^2}} \right) = 2t_w + \frac{1}{2\pi T_1} \log \left( \frac{\frac{w^2}{\sqrt{2w^2-1}} - 1}{\frac{w^2}{\sqrt{2w^2-1}} + 1} \right). \quad (4.29)$$

For light shocks, this critical time can be expanded as

$$t_{c,v} = 2t_w + \frac{1}{\pi T_1} \log(w-1) + O(w-1) = 2(t_w - t_{scr}^*) + \frac{\log 2}{\pi T_1} + O(w-1) \quad (4.30)$$

where the scrambling time (3.67) (with  $\epsilon \equiv w-1$ ) appears in a manner similar to eq. (3.64) for the CA approach. For heavy shocks, we can expand eq. (4.29) for large  $w$ , which yields

$$t_{c,v} = 2t_w - \frac{\sqrt{2}}{\pi T_2} + \mathcal{O}\left(\frac{1}{w^2 T_2}\right) \quad (4.31)$$

where we have used the relation  $wT_1 = T_2$ . Again this expression is similar to the critical time in eq. (3.61) found for the CA calculations.

Since this regime in which the rate of computation is proportional to the difference of the masses extends more or less equally in negative times we expect it to last approximately for  $\Delta t = 2t_c$ . For instance, if we consider very early shocks of a fixed non-vanishing energy ( $t_w$  very large and  $w$  fixed and finite) we expect that  $\Delta t \approx 4t_w$ . This duration will be (significantly) shortened for light shocks by the scrambling time. We will confirm these predictions as well as the results of eqs. (4.28)–(4.29) with numerical examples below.

**A comment on dimensionless quantities.** It is possible to express all our results in eqs. (4.9)–(4.12), (4.13)–(4.15) and (4.19), in terms of the dimensionless quantities in eq. (2.18), as well as the dimensionless surface momenta

$$p_1 \equiv \frac{LP_1}{r_{h,1}^d}, \quad p_2 \equiv \frac{LP_2}{r_{h,2}^d}. \quad (4.32)$$

We have found this representation particularly useful for the case of planar black holes in  $d \geq 3$  and for BTZ black holes for which the dependence on  $z$  cancels and one is able to demonstrate that the profile of the rate of change in complexity normalized by the sum of masses as a function of  $T_2 t$ , depends on  $w$  and  $T_2 t_w$ , but not on  $z$ .

## 4.2 Complexity of formation

As discussed in eq. (3.70), the complexity of formation was originally defined as the difference between the complexity of the TFD state and that of two copies of the vacuum. Here we evaluate the complexity of formation for the shock wave geometries using the complexity=volume proposal, by evaluating the complexity at  $t_R = t_L = 0$  and subtracting two copies of the equivalent result for empty AdS. For this purpose, it will be convenient to define

$$R_0[r] = \frac{r^{(d-1)}}{\sqrt{f_{\text{vac}}(r)}} \quad (4.33)$$

where  $f_{\text{vac}}(r) = \frac{r^2}{L^2} + k$  is the corresponding ‘blackening’ factor for the AdS vacuum. As we have mentioned before, there are different scenarios for the shape of the surface. In particular, it can have turning points in  $\text{BH}_1$  and/or in  $\text{BH}_2$  (see figure 34 in appendix C). These different cases will have different expressions for the complexity of formation as we detail below. We have reorganized the integrals in such a way that they are all finite in the limit  $r_{\text{max}} \rightarrow \infty$  and so we may replace the relevant integration limits by  $\infty$ . The complexity of formation is a sum of two contributions

$$\Delta\mathcal{C}_V = \Delta\mathcal{C}_{V1} + \Delta\mathcal{C}_{V2} \tag{4.34}$$

where

$$\Delta\mathcal{C}_{V1} = \frac{\Omega_{k,d-1}}{GL} \begin{cases} \int_{r_{t,1}}^{\infty} (R_1[P_1, r] - R_0[r]) dr & \text{turning point in BH}_1 \\ + \int_{r_{t,1}}^{r_s} R_1[P_1, r] dr - \int_0^{r_{t,1}} R_0[r] & \\ \int_{r_s}^{\infty} (R_1[P_1, r] - R_0[r]) dr - \int_0^{r_s} R_0[r] & \text{otherwise} \end{cases} \tag{4.35}$$

$$\Delta\mathcal{C}_{V2} = \frac{\Omega_{k,d-1}}{GL} \begin{cases} \int_{r_{t,2}}^{\infty} (R_2[P_2, r] - R_0[r]) dr & \text{turning point in BH}_2 \\ + \int_{r_{t,2}}^{r_s} R_2[P_2, r] dr - \int_0^{r_{t,2}} R_0[r] & \\ \int_{r_s}^{\infty} (R_2[P_2, r] - R_0[r]) - \int_0^{r_s} R_0[r] dr & \text{otherwise.} \end{cases}$$

We will be able to study the complexity of formation numerically as a function of  $t_w$  and the masses  $M_1, M_2$ . When studying BTZ black holes, we will take  $f_{\text{vac}}(r)$  to correspond to the Neveu-Schwarz vacuum, i.e.,  $k = +1$ .

When plotting our numerical results, we found it convenient to subtract from the complexity of formation in the shock wave geometry the equivalent complexity for forming an eternal AdS black hole of radius  $r_{h,1}$ . In this way, we expect that near  $t_w = 0$  the result will vanish. We have used the following values for the complexity of formation of eternal black holes with  $d = 2$  starting from the Neveu-Schwarz vacuum

$$\Delta\mathcal{C}_{V,NS} = \frac{4\pi L}{G_N}, \tag{4.36}$$

and for planar black holes in  $d \geq 3$

$$\Delta\mathcal{C}_{V,1}^{\text{eternal}} = \frac{\sqrt{\pi}\Omega_{0,d-1}}{G_N} \frac{(d-2)\Gamma\left(1 + \frac{1}{d}\right)}{(d-1)\Gamma\left(\frac{1}{2} + \frac{1}{d}\right)} r_{h,1}^{d-1}, \tag{4.37}$$

which were previously evaluated in [22]. Equivalently, after the subtraction we are evaluating the difference of complexities between the shock wave geometry and an eternal black hole of radius  $r_{h,1}$  and this amounts to substituting in eq. (4.17)  $R_0[r]$  by  $R_1[0, r]$  and replacing the 0 in the integration limit by  $r_{h,1}$  (perhaps with some reorganization of the integrals).

**Early and late time limits of the complexity of formation.** The symmetry of the problem under time shifts (recall eqs. (2.3)–(2.6)), which includes shifting the time of the shock wave, implies that the effect on the complexity of pushing the shock wave to the past by an amount of time  $dt_w$  is the same as that obtained by studying the time evolution of the

complexity under anti-symmetric time evolution with  $dt_R = -dt_L = dt_w$ . We can use this understanding to make some general statements about the slope of the complexity of formation with respect to the time of the shock  $t_w$  in certain limits for planar and BTZ black holes.

The first limit we consider is that of shocks injected at times  $t_w \approx 0$ . In this case we expect that  $x_s \rightarrow \infty$ . We also expect that the configuration is extremely similar to studying the complexity of formation in  $BH_1$  without the shock, which implies that  $P_1 \approx 0$ . Using the matching condition (4.11), this leads to  $P_2 \approx P_{2,m} - P_{1,m}$  and using eq. (4.18) with  $dt_R = -dt_L = dt_w$ , we obtain

$$\frac{d\Delta\mathcal{C}_V}{dt_w} = \frac{8\pi(M_2 - M_1)}{d - 1}. \tag{4.38}$$

In fact, we will see that this approximation of the slope holds for a range of  $t_w$  which increases logarithmically with the inverse of the energy of the shock. This range is another indicator of the scrambling time of the system and we will come back to this point in our numerical analysis.

For very early shocks (large  $t_w$ ), the  $t_L = t_R = 0$  configuration of the complexity of formation is very similar to the one that appears in figure 14b. This implies that  $P_1 \approx -P_{1,m}$  and  $P_2 \approx P_{2,m}$ . Substituting this together with  $dt_R = -dt_L = dt_w$  into eq. (4.18), we obtain

$$\frac{d\Delta\mathcal{C}_V}{dt_w} = \frac{8\pi(M_1 + M_2)}{d - 1}. \tag{4.39}$$

One comment is in order before we proceed: the linear growth with the slope in eq. (4.38) turns out to be superimposed with a tiny exponential growth, which governs the transition to the regime in (4.39). The coefficient in this exponential growth is none other than the Lyapunov exponent, which characterizes the scrambling of information in the system. We will come back to this point in detail in section 5.

**A comment on the complexity of formation and dimensionless variables.** When expressing our results in terms of the dimensionless coordinates (4.32) for planar geometries (with  $d \geq 3$ ), it is possible to demonstrate that the combination

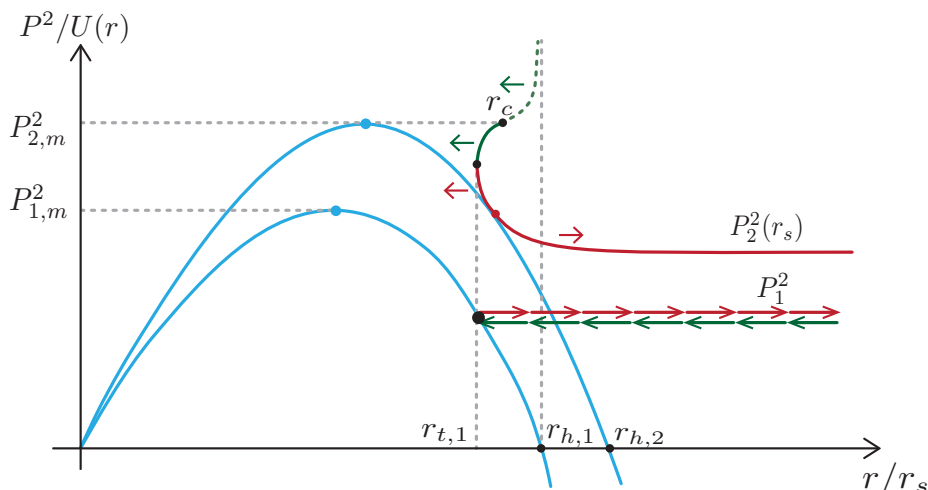
$$\frac{\Delta\mathcal{C}_V - \Delta\mathcal{C}_{V,1}^{\text{eternal}}}{S_1} \tag{4.40}$$

where the entropy  $S_1$  of  $BH_1$  was defined in eq. (2.14), is independent of  $z$ . For BTZ black holes in  $d = 2$ , the same statement holds when replacing  $\Delta\mathcal{C}_{V,1}^{\text{eternal}}$  by  $\Delta\mathcal{C}_{V,NS}$  and the entropy by that for BTZ, see eq. (3.55). In this case, the  $z$  independence implies that the extra complexity generated by the shock wave grows linearly with the temperature (just as  $S_1$  does). This contrasts with the original complexity of formation  $\Delta\mathcal{C}_{V,NS}$  in eq. (4.36), which was a fixed constant. We will plot the combination (4.40) in our numerical analysis below.

### 4.3 Numerical analysis

#### 4.3.1 $d = 2$

In this section, we collect some numerical results for shock waves in BTZ black holes, i.e., with the boundary dimension  $d = 2$ . In obtaining these results, we found it very useful

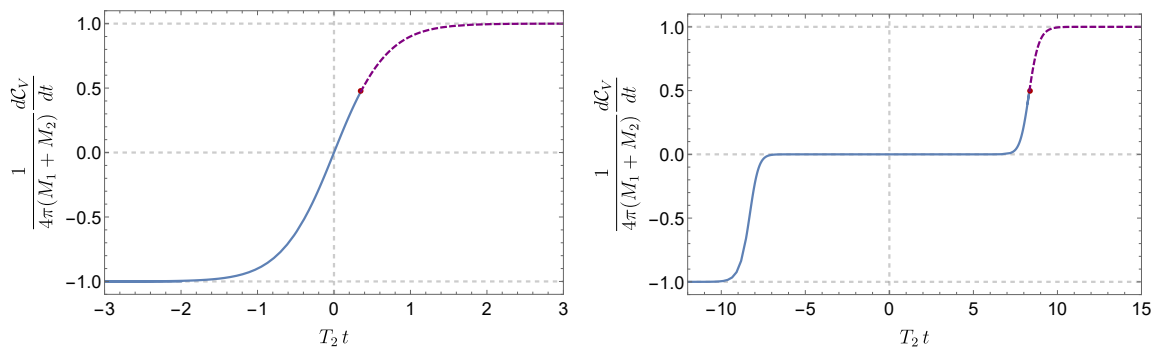


**Figure 15.** Illustration of the potentials  $U_{1,2}(r)$  (see eq. (4.8)), superimposed by a plot of  $P_2^2(r_s, P_1)$  (see eq. (4.11)) for a specific value of  $P_1^2$ . We have indicated by arrows the sign of  $\dot{r}$  before hitting the shell (on the  $P_1^2$  lines) and right after hitting the shell (next to the  $P_2^2(r_s)$  lines). Red lines are associated with surfaces that have a turning point in  $\text{BH}_1$  and green lines with those that do not. The plot was made for  $P_1 > 0$  but a similar plot can be made for  $P_1 < 0$ . The special point  $r_c$  where  $P_2$  becomes equal to the maximal value  $P_{2,m}^2$  was very important in our analysis since in many of the cases studied (early shocks, symmetric boundary times) the solution for  $r_s$  was found to be extremely close to it. The dashed part of the green curve  $P_2^2(r_s)$  is associated with surfaces that will end on the singularity after crossing the shell and are therefore irrelevant for our analysis.

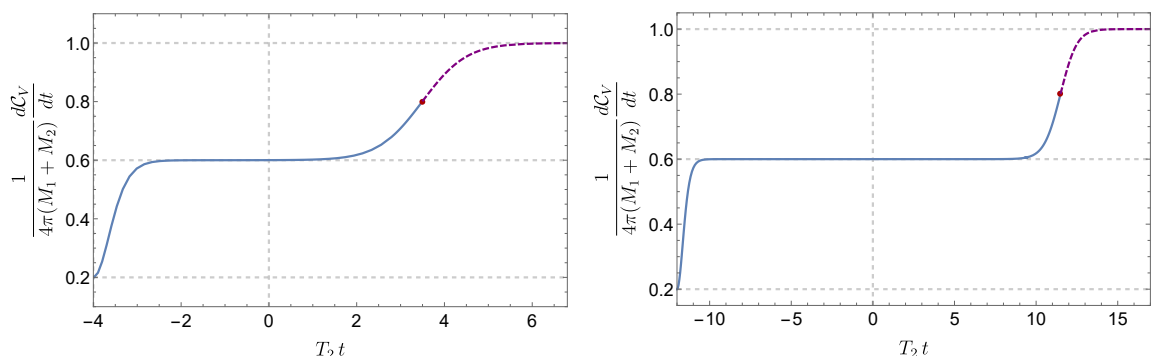
to draw plots of the potentials in eq. (4.8), superimposed by the value of  $(P_2)^2$  predicted from the jump condition (4.11) as a function of  $r_s$  — see figure 15. These plots for a given value of  $P_1$  helped us develop an intuition for the numerical range in which we expect to find the crossing point  $r_s$  for the solution of symmetric boundary times for early shocks.

Figures 16 and 17 contain the time derivatives of the complexity for a symmetric boundary time evolution  $t_L = t_R = t/2$  as a function of  $t$  for light and heavy shocks, respectively. The figures demonstrate that for early shocks, a plateau develops in which the rate of computation is proportional to the difference of the masses, i.e., vanishes for light shocks. As explained earlier, we may want to define a critical time as the point in the symmetric time evolution for which  $P_1 = 0$  and which indicates the end of this plateau. We have plotted this critical time as a function of the time of the shock for a heavy shock  $w = 2$  and checked its agreement with the (early shock) prediction in eq. (4.29) in figure 18. We have also pointed out in eq. (4.30) that the behavior of this critical time for light shocks is an indicator of the scrambling time. We verify this statement by plotting the critical time as a function of  $w$  in figure 18.

We find it insightful to describe the evolution of the extremal surface that came into play in these numerical solutions. Evolving from past to future for times  $t_L = t_R > -t_w$ , we have: 1. The surface passes behind the past horizon in  $\text{BH}_1$  and has a turning point only in  $\text{BH}_1$ ; 2. Part of the surface moves into the future horizon in  $\text{BH}_2$  and a turning point in  $\text{BH}_2$



**Figure 16.** Time derivative of the complexity evolved with symmetric boundary times  $t_L = t_R = t/2$  as a function of  $T_2t$  normalized by the sum of masses for light shocks with  $w = 1 + 10^{-5}$  for BTZ black holes (i.e.,  $d = 2$ ) for  $T_2t_w = 2$  (left) and  $T_2t_w = 6$  (right). The horizontal dashed lines indicate the limits (4.22) and (4.25) which are very close to those of the configuration without the shock, i.e.,  $\pm 1$  in our normalization. We see that the late time limit is approached from below. For early shock times, a plateau develops near  $t = 0$  in which the rate of computation is proportional to the difference of the masses and hence is nearly vanishing (see eq. (4.23)).

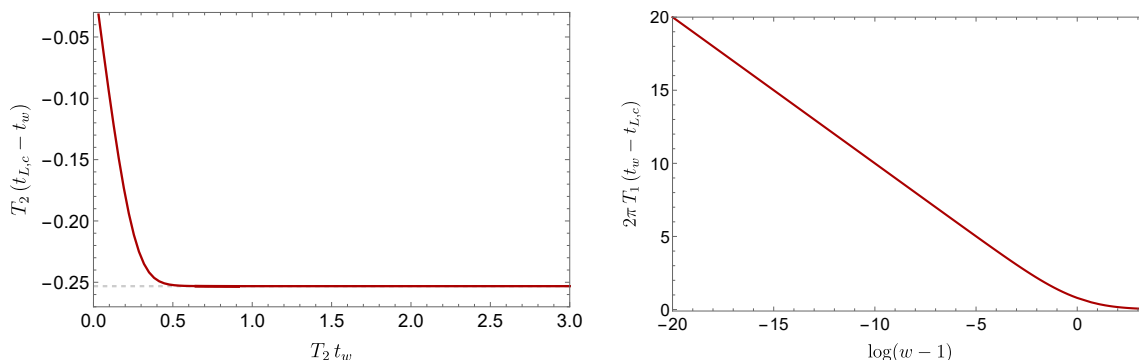


**Figure 17.** Time derivative of the complexity evolved with symmetric boundary times  $t_L = t_R = t/2$  as a function of  $T_2t$  normalized by the sum of masses for finite size shocks with  $w = 2$  for BTZ black holes (i.e.,  $d = 2$ ) for  $T_2t_w = 2$  (left) and  $T_2t_w = 6$  (right). The horizontal dashed lines indicate the limits (4.22), (4.23) and (4.25). We see that the late time limit is approached from below. For early shock times, a plateau develops near  $t = 0$  in which the rate of computation is proportional to the difference of the masses (see eq. (4.23)).

develops; 3. The surface crosses the bifurcation surface and moves behind the future horizon in  $BH_1$ , we still have two turning points; 4. If the shock is light, the turning point in  $BH_1$  disappears for a certain range of surface momenta, then reappears in the late time limit. This regime only takes place when there exists a range of energies for which the crossing point between the red and green curves in figure 15 moves below  $P_2^m$ , which happens if  $w < \sqrt{\frac{1}{2} + \frac{1}{\sqrt{2}}}$  for energies  $|4p_1 + 1 - 2w^2| < \sqrt{1 + 4w^2 - 4w^4}$  where  $p_1$  was defined in eq. (4.32).

We have also studied numerically the complexity of formation for various energies of the shock as a function of the shock time  $t_w$ . The results are extremely similar to those obtained using the complexity=action prescription and can be found in figure 19 (cf. figure 7 in the CA subsection). We see that for light shocks, a long plateau develops in which the rate of





**Figure 18.** Left: critical time in the symmetric boundary time evolution  $t_R = t_L = t/2$  of the extremal surface passing through the bifurcation surface in  $BH_1$  ( $P_1 = 0$ ) in shocks of  $d = 2$  BTZ black holes with  $w = 2$  as a function of the time of the shock  $t_w$ . The dashed gray line indicates the prediction of eq. (4.29) for early shocks. Right: critical time obtained for early shocks with  $T_2 t_w = 6$  for various values of  $w$ . As indicated by eq. (4.30), the logarithmic dependence on  $w - 1$  is an indicator of the scrambling time for light shocks. The critical times in these two plots are related to the ending of the plateau in which the rate of computation is proportional to the difference of the masses, see figures 16–17, where the relevant point was indicated by a red dot.

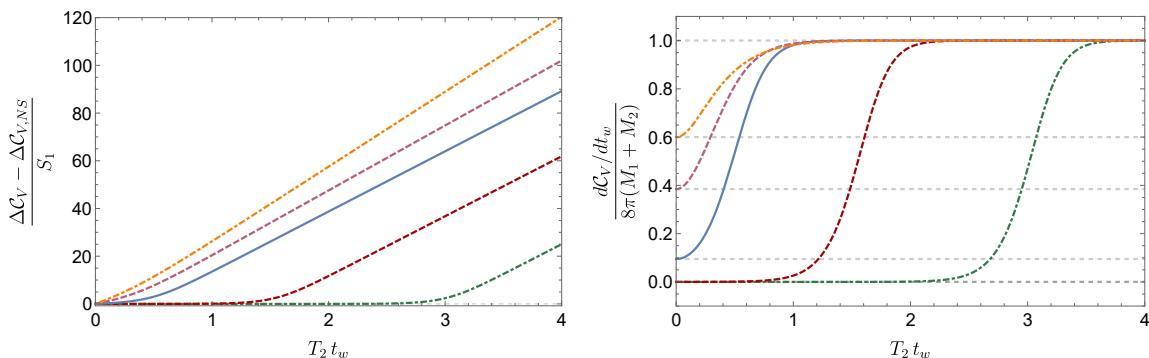
computation is proportional to the difference of the masses. The length of this plateau is logarithmic in the inverse energy of the shock. To confirm that we have plotted the length of this plateau as a function of  $w$  for light shocks and checked that it agrees with the scrambling time in eq. (3.67), up to a small constant shift. We have defined the length of the plateau as the time for which the two straight line asymptotic regimes corresponding to the limits in eqs. (4.38) and (4.39) intersect. The small constant shift is a consequence of the fact that the constant in eq. (3.67) was fixed arbitrarily there in order to simplify the expressions in the CA subsection. If we fit the plot in figure 20, we find a very close agreement with

$$t_{\text{form}}^* = \frac{1}{2\pi T_1} \log\left(\frac{\gamma_2}{2(w-1)}\right), \quad \gamma_2 = 3.1412 \pm 0.0008, \quad (4.41)$$

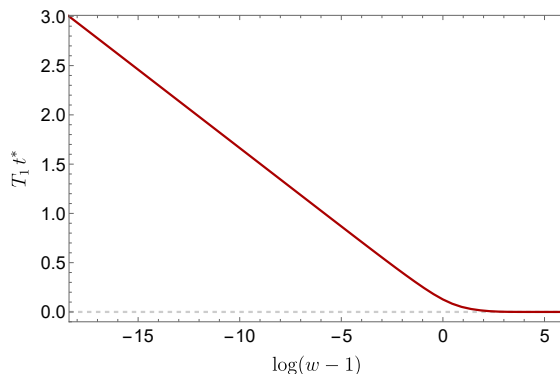
i.e., the factor 2 in the argument of the logarithm in eq. (3.67) is replaced here by  $\pi/2$ .

### 4.3.2 $d = 4$

We have repeated the analysis above for the case of planar black holes in  $d = 4$ . The results for the time derivative of complexity under symmetric time evolution and for the complexity of formation can be found in figure 21 and they are in agreement with the predicted limits in eqs. (4.22)–(4.23), (4.25) and (4.38)–(4.39). We have not studied the case of spherical black holes, however, we would like to suggest that the various limits in this case are still related to the maximal points of the potentials. Of course, this will not give an exact proportionality to the masses but rather the late time limit will be modified by curvature corrections proportional to  $1/(LT)^2$ . There are interesting questions to be asked, e.g., is the late time limit for the rate of computation approached from below like in the time dependent eternal BH or from above like in Vaidya? Is the limit above or below the planar one? We leave these questions for future study.



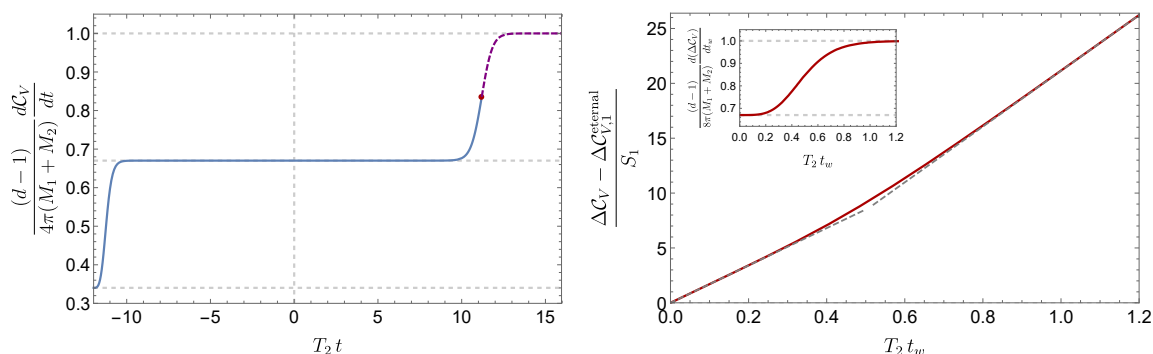
**Figure 19.** Complexity of Formation as a function of the shock time  $t_w$  (left) and its derivative (right) for various energies of the shocks —  $w = 2$  (orange, dot-dashed),  $w = 1.5$  (pink, dashed),  $w = 1.1$  (blue, solid),  $w = 1 + 10^{-4}$  (red, dashed) and  $w = 1 + 10^{-8}$  (green, dot-dashed). The slope of this plot is proportional to the difference of the masses for small  $t_w$  (not so early shocks) and to the sum of the masses for large  $t_w$  (early shocks) and matches the predictions in eqs. (4.38)–(4.39), indicated by dashed gray lines on the right plot.



**Figure 20.** Characteristic time of crossing between the two regimes in the complexity of formation in which the slopes are proportional to the difference of the masses, see eq. (4.38), and the sum of the masses, see eq. (4.39). The logarithmic dependence on the energy of the shock in this plot is another diagnostic of the scrambling time (cf. figure 18, right panel) and matches closely the fit in eq. (4.41).

## 5 Discussion

Using the framework established in [1], we studied holographic complexity in Vaidya geometries (2.7) describing a shock wave propagating into an eternal black hole. Of course, this situation has already been well studied not only for a single shock wave but also for many shock waves as well, e.g., [15–18, 47, 49]. New directions investigated here were to study the full time dependence of the holographic complexity for both light and heavy shocks in sections 3.1 and 4.1, and to evaluate the complexity of formation in sections 3.2 and 4.2. In both cases, we examined both the CA and CV approaches. In the following, we review our results in these calculations, and also consider their physical implications.



**Figure 21.** Numerical results for the complexity in shock wave backgrounds in planar ( $k = 0$ ) black holes in  $d = 4$  with  $w \equiv r_{h,2}/r_{h,1} = 1.5$ . Left: rate of change of complexity as a function of time for symmetric evolution with  $t_L = t_R = t/2$  with an early shock sent at  $T_2 t_w = 6$ . We have started the plot at  $t = -2t_w$  and indicated by horizontal dashed lines three different limits (in the relevant normalization as indicated on the axis) — the limit of times close to  $t_L = t_R = -t_w$  (see eq. (4.25)), the plateau near  $t = 0$  (see eq. (4.23)), and the late time limit (see eq. (4.22)). The red dot indicates the point in which  $P_1 = 0$  which signals the transition of the surface from the white hole to the black hole region in  $BH_1$ . Right: complexity of formation as a function of the shock time  $t_w$  shifted and normalized as in eq. (4.40). We have indicated by dashed lines the predicted slopes (see eqs. (4.38)–(4.39)). The inset shows the derivative and its agreement with the predicted slopes.

### Complexity of formation

The complexity of formation for the shock wave geometries was examined in sections 3.2 and 4.2 using the CA and CV approaches, respectively. Recall that as originally studied in [22],<sup>26</sup> the complexity of formation was defined as the difference between the complexity of preparing two copies of the boundary CFT in the thermofield double state (TFD) at  $t_L = t_R = 0$  and the complexity of preparing each of the CFT’s in its vacuum state, as shown in eq. (3.70). A key feature of this quantity is that the difference of the complexities is UV finite.

In extending these calculations to the perturbed black holes, we first considered light shocks, i.e.,  $w = r_{h,2}/r_{h,1} \simeq 1$ , sent from the right boundary at some early time  $t_R = -t_w$  — see left panel in figures 7 and 30 for  $d = 2$  and  $d = 4$ , respectively, with the CA approach and the green and red curves in figure 19 for  $d = 2$ , with the CV approach. In this case,  $\Delta\mathcal{C}$  was essentially unchanged for a wide range of  $t_w$ . But then beyond some critical  $t_w$ , i.e., for earlier shock waves,  $\Delta\mathcal{C}$  grew linearly as<sup>27</sup>

$$\frac{d\Delta\mathcal{C}_A}{dt_w} = \frac{4M_1}{\pi} \quad \text{and} \quad \frac{d\Delta\mathcal{C}_V}{dt_w} = \frac{16\pi M_1}{d-1}. \tag{5.1}$$

The critical injection time where this transition occurred was precisely given by the scrambling time, as defined in eqs. (3.67) and (4.41) for the CA and CV approaches,

<sup>26</sup>These calculations were extended to charged AdS black holes in [23].

<sup>27</sup>The results for  $\Delta\mathcal{C}_V$  in the present equation, as well as in eqs. (5.2) and (5.3), only apply for planar black holes. For spherical geometries, one can still express the various limits of  $d\Delta\mathcal{C}_V/dt_w$  and  $d\mathcal{C}_V/dt$  using the extremal points of the potentials (4.8). However, in the spherical case these values are no longer simply related to the masses.

respectively. Hence as a rough approximation, the complexity of formation can be described by two linear regimes, as shown in eq. (3.84) for  $\Delta\mathcal{C}_A$ . Of course, this is a manifestation of the switchback effect [15]. That is, we can think that we begin with the unperturbed TFD state at  $t_L = 0 = t_R$ . We then evolve the state backwards in time to  $t_R = -t_w$ , where we make the perturbation dual to the insertion of the shock wave, and then we evolve forward in time again to the initial time, i.e.,  $t_R = 0$ . The perturbation has a minor effect on the final state for  $t_w < t_{\text{scr}}^*$ , with the backward and forward time evolution essentially canceling out, and hence the complexity of formation remains unchanged for these perturbations. However, the perturbation begins to have a dramatic effect in modifying the final state for  $t_w > t_{\text{scr}}^*$  and hence we see that the complexity of formation begins to grow at this point — see further discussion below.

For higher energy shock waves on the other hand, the perturbation brings in an appreciable energy and accesses new degrees of freedom. Hence the state is modified even for small  $t_w$  and the regime in which the  $\Delta\mathcal{C}$  is unchanged is absent. Instead, the complexity of formation as a function of  $t_w$  starts increasing right away with initial rate given by

$$\left. \frac{d\Delta\mathcal{C}_A}{dt_w} \right|_{t_w=0} = \frac{(M_2 - M_1)}{\pi} \quad \text{and} \quad \left. \frac{d\Delta\mathcal{C}_V}{dt_w} \right|_{t_w=0} = \frac{8\pi(M_2 - M_1)}{d - 1}, \quad (5.2)$$

as in eqs. (B.25) and (4.38), respectively.<sup>28</sup> That is, this initial growth rate is driven by the energy in the shock wave, i.e., it is proportional to the difference of the masses of the two black holes. As the injection time  $t_w$  continues to increase,  $d\Delta\mathcal{C}/dt_w$  increases and soon saturates to a constant rate proportional to the sum of the masses

$$\frac{d\Delta\mathcal{C}_A}{dt_w} = \frac{2(M_1 + M_2)}{\pi} \quad \text{and} \quad \frac{d\Delta\mathcal{C}_V}{dt_w} = \frac{8\pi(M_1 + M_2)}{d - 1}. \quad (5.3)$$

Of course, the latter matches eq. (5.1) in the limit where  $M_2 \rightarrow M_1$ .

Let us observe that we can connect these results to the time evolution of holographic complexity for one-sided black holes studied in [1]. As discussed above eq. (4.38), the derivative of the complexity of formation with respect to  $t_w$  can be related via the time shift symmetry to the antisymmetric time evolution of the complexity, i.e., using eqs. (2.3) and (2.6),

$$\frac{d\Delta\mathcal{C}}{dt_w} = \frac{d\mathcal{C}}{dt_R} - \frac{d\mathcal{C}}{dt_L}, \quad (5.4)$$

where we have used the fact that the complexity of preparing each CFT in its vacuum state is independent of the time and so we could replace  $\Delta\mathcal{C}$  with  $\mathcal{C}$  on the right hand side. The limit of one-sided black holes is obtained by setting  $M_1 = 0$ ,  $M_2 = M$  and studying the dependence on the time  $t_0 = t_R + t_w$ , i.e., the time after the perturbation is inserted. The second time interval  $t_L - t_w$  appearing in our expressions will not play a role after setting  $M_1 = 0$ . Hence in eqs. (5.2) and (5.3), we set  $M_1 = 0$  and  $M_2 = M$  and trade the time evolution for an evolution in  $t_0$ . Then taking the limit  $t_w \rightarrow 0$  corresponds to the early time limit  $t_0 \rightarrow 0$ , for which eq. (5.2) yields  $d\mathcal{C}_A/dt_0 = M/\pi$  and  $d\mathcal{C}_V/dt_0 = 8\pi M/(d - 1)$ . Similarly taking

---

<sup>28</sup>Note that we have ignored the logarithmic term for  $\Delta\mathcal{C}_A$  in eq. (3.85), which only appears for  $d = 2$  but not higher boundary dimensions.

$t_w \rightarrow \infty$  corresponds to the late time limit  $t_0 \rightarrow \infty$ , for which eq. (5.2) yields  $d\mathcal{C}_A/dt_0 = 2M/\pi$  and  $d\mathcal{C}_V/dt_0 = 8\pi M/(d-1)$ . Indeed these limits precisely match eqs. (3.47), (3.48) and (3.77) in [1]. In particular, the relative factor of 2 between the early and late time limits of the CA results in eqs. (5.2) and (5.3) is the same ratio observed between the early and late time limits of one sided black holes, see e.g., figure 4 in [1]. Similarly, the equal early and late time limits from the CV results match the constant rate of change in complexity obtained for planar one sided black holes using the CV conjecture. Note that in taking this limit, the details of the left boundary time become unimportant and so one can also extract the same limit from the symmetric time evolution, as we explain below.

**Time evolution**

We have also extended the previous studies of holographic complexity in shock wave geometries [15, 18] by studying the full time evolution of the holographic complexity for both light and finite energy shocks and using both the CA and the CV approaches. For simplicity, we focused on the symmetric time evolution  $t_L = t_R = t/2$ , which is readily compared with the time evolution in the unperturbed black hole backgrounds studied in [23].

Let us begin by discussing the light shocks. Using both conjectures, we observed that if the shock was sent earlier than the scrambling time, the rate of change in complexity was approximately vanishing for a long period of time, centered around  $t = 0$ . At later times, the rate of growth of the complexity rapidly approaches the growth rate found in the unperturbed geometry. For the CA conjecture, we defined a number of critical times which characterized the transitions between the different regimes of the complexity growth. These were:  $-t_{c,0}$ , the time in the past where the WDW patch enters the future singularity;  $t_{c1}$ , the time in which the point  $r_m$  leaves the past singularity; and  $t_{c,2}$ , the time in which the crossing point  $r_b$  enters the singularity, see figure 1. The plateau we have mentioned where the rate of computation vanishes appears for  $-t_{c0} \lesssim t \lesssim t_{c1}$ , see the right panel in figure 2. Using the CV conjecture, we observed a similar behaviour of the time derivative of the complexity, as shown in the right panel of figure 16. We could characterize the period over which the rate of change in complexity was nearly vanishing by another critical time  $t_{c,v}$  defined as the characteristic time in which the extremal surfaces pass through the bifurcation surface in  $BH_1$ . In the limit of early and light shocks in BTZ black holes, we were able to derive analytic expressions for the various critical times

$$t_{c1} = t_{c0} = 2t_w - \frac{1}{\pi T_1} \log \frac{2}{\epsilon} + O(\epsilon \log \epsilon), \tag{5.5}$$

and  $t_{c2} = 2t_w$  in the CA approach, while

$$t_{c,v} = 2t_w - \frac{1}{\pi T_1} \log \frac{1}{\epsilon} + O(\epsilon) \tag{5.6}$$

in the CV approach. Hence, we see the appearance of the scrambling time

$$t_{scr}^* = \frac{1}{2\pi T_1} \log \frac{2}{\epsilon} \tag{5.7}$$

in shortening of the plateau of constant complexity from  $2t_w$ . This is another manifestation of the switchback effect [15], as we explain below.

For heavier shocks, the regime of vanishing computation rate was replaced by a regime in which the rate of computation was approximately constant and proportional to the difference of the masses, i.e., the energy carried by the shock wave,<sup>29</sup>

$$\frac{d\mathcal{C}_A}{dt} = \frac{(M_2 - M_1)}{\pi} \quad \text{and} \quad \frac{d\mathcal{C}_V}{dt} = \frac{4\pi(M_2 - M_1)}{d - 1}. \quad (5.8)$$

There is then a rapid transition where the rate of computation approaches the late-time limit, which is proportional to the sum of the masses,

$$\frac{d\mathcal{C}_A}{dt} = \frac{(M_1 + M_2)}{\pi} \quad \text{and} \quad \frac{d\mathcal{C}_V}{dt} = \frac{4\pi(M_1 + M_2)}{d - 1}. \quad (5.9)$$

This late time limit was approached from above using the CA conjecture and from below using the CV conjecture in all the cases analyzed.

As before, we can relate these results to analogous rates found for one-sided black holes in [1]. In particular, in eq. (5.9), we set  $M_1 = 0$  and replace  $t_R + t_w = t/2 + t_w = t_0$ ,  $M_2 = M$ , which yields  $d\mathcal{C}_A/dt_0 = 2M/\pi$  and  $d\mathcal{C}_V/dt_0 = \frac{8\pi M}{d-1}$ . These then match the late time limits in eqs. (3.48) and (3.77) of [1]. In all these cases, we have assumed that the value of  $t_w$  was large and therefore they correspond to the  $t_0 \rightarrow \infty$  limit of the one-sided black holes.

The CV conjecture suggests a simple geometric picture to explain these rates in terms of surfaces wrapping around constant  $r$  surfaces behind the future and past horizons, see figure 14. A similar interpretation was suggested in [56] starting from boost symmetry principles in terms of two ‘tapes’ storing the forward and backward Hamiltonian evolution in the past and future interiors of the black hole. Our geometric picture makes it clear that evolving only  $t_L$  starting at  $t_L = t_R = 0$  will result in decreasing complexity. In fact, this situation is very similar to the one described in figure 14b, where evolving the right boundary time upwards caused the surface to wrap around the critical surface in BH<sub>2</sub>, while evolving the left boundary time upwards caused the surface to unwrap the critical surface in BH<sub>1</sub>, resulting in a rate of change in complexity proportional to  $M_2 - M_1$  at early times. Of course, if we only evolve  $t_L$  while holding  $t_R = 0$  fixed, we expect to be left with a negative rate of change of the complexity proportional to  $-M_1$ . We will return to this point below.

### Null surface counterterm

Our calculations using the complexity=action proposal in section 3 included the counterterm (3.3) on the null boundaries of the WDW patch. Adding this surface term does not modify many key results for the CA proposal for eternal black holes, e.g., the complexity of formation [22] or the late-time rate of growth [23]. But it does modify the details of the transient behaviour in the time evolution [23]. However, these comments are limited to the behaviour of holographic complexity in stationary spacetimes. In studying holographic complexity in Vaidya spacetimes [1], we found that the counterterm is an essential ingredient for the CA proposal. In particular, we showed that for geometries describing black hole formation, one does not recover the expected late time growth for general  $d$ .

---

<sup>29</sup>Again, for  $\mathcal{C}_V$ , the result here and in the following equation only apply for planar black holes — see footnote 27.

This effect was most dramatic for  $d = 2$  (and  $k = +1$ ) where the growth rate was actually negative throughout the process, i.e., without the counterterm, the complexity appeared to decrease. In section 3.3, we also considered dropping the counterterm in our present calculations. There we found that without the counterterm, the holographic calculations again fail to reproduce the expected late time growth rate and that the complexity of formation does not exhibit the behaviour that is characteristic of the switchback effect. Hence the gravitational observable associated with  $I_{\text{WDW}} - I_{\text{ct}}$  simply does not behave like complexity of the boundary state, and the results in section 3.3 reinforce our previous arguments that the counterterm should be regarded as an essential ingredient for the CA proposal.

One interesting aspect of the counterterm is that the structure of the UV divergences in the holographic complexity is modified, as was first noted in [79], and as is discussed in appendix A. Without the counterterm, the leading UV divergence takes the form (see eq. (A.14))

$$\tilde{\mathcal{C}}_{\text{A}}^{\text{UV}} \sim \frac{\mathcal{V}(\Sigma)}{\delta^{d-1}} \log\left(\frac{L}{\alpha\delta}\right), \tag{5.10}$$

where  $\mathcal{V}(\Sigma)$  is the (total) volume of the time slice  $\Sigma$  on which the boundary state resides. To remove the AdS scale from  $\tilde{\mathcal{C}}_{\text{A}}^{\text{UV}}$ , [26] suggested that one should choose

$$\alpha = L/\ell, \tag{5.11}$$

where  $\ell$  might be some other length scale associated with the microscopic rules used to define the complexity in the boundary theory. This choice then yields

$$\mathcal{C}^{\text{UV}} \sim \frac{\mathcal{V}(\Sigma)}{\delta^{d-1}} \log\left(\frac{\ell}{\delta}\right). \tag{5.12}$$

When the counterterm contributions are included, the  $\alpha$  dependence in eq. (5.10) is eliminated and the leading UV divergence takes the form (see eq. (A.16))

$$\mathcal{C}_{\text{A}}^{\text{UV}} \sim \frac{\mathcal{V}(\Sigma)}{\delta^{d-1}} \log\left(\frac{(d-1)\ell_{\text{ct}}}{L}\right). \tag{5.13}$$

Of course, this expression suffers from the same deficiencies as eq. (5.10), i.e., it contains the AdS scale which has no interpretation in the boundary theory, and it is ambiguous because the counterterm scale is undetermined. However, as before, we can use the latter ambiguity to eliminate the AdS scale. In particular, if we choose

$$\ell_{\text{ct}} = \frac{L}{d-1} \frac{\ell}{\delta}, \tag{5.14}$$

then the leading UV divergence takes the same form as in eq. (5.12). Of course, as before, one is left with the ambiguity of fixing the scale  $\ell$ .

Now in comparing holographic complexity with calculations of complexity in a (free) scalar field theory [33, 34], it was noted that the leading contribution to the complexity took precisely the form given in eq. (5.12).<sup>30</sup> In this case, the scale  $\ell$  corresponds to

---

<sup>30</sup>We note that the logarithmic factor does not seem to arise for a free fermion [35, 57]. That is, the leading singularity takes the form  $\mathcal{C}^{\text{UV}} \sim \mathcal{V}(\Sigma)/\delta^{d-1}$ , which is similar to the holographic result found using the complexity=volume.

the width of the unentangled Gaussian reference state appearing in the evaluation of the complexity. Hence it was suggested that the freedom of choosing this scale in the field theory calculations of complexity could be associated with the ambiguity of fixing  $\alpha$  in the complexity=action proposal. Since we are now advocating that the latter proposal must include the null surface counterterm (3.3), we must instead associate this freedom with the ambiguity in fixing the counterterm scale, e.g., as in eq. (5.14).

However, we would like to point out a difference in these two possibilities. This difference is highlighted by first choosing  $\ell$  to be a UV scale. For example, with  $\ell = e^\sigma \delta$ , the logarithmic factor in eq. (5.12) simply provides a numerical factor<sup>31</sup> and the leading UV divergence reduces to  $\mathcal{C}^{\text{UV}} \sim \sigma \mathcal{V}(\Sigma)/\delta^{d-1}$ . However, with this choice, eq. (5.11) yields  $\alpha = e^{-\sigma} L/\delta$  while eq. (5.14) yields  $\ell_{\text{ct}} = e^\sigma L/(d-1)$ . Hence in previous discussions without the counterterm (e.g., [23]), the UV cutoff  $\delta$  appears in the transient behaviour of  $d\tilde{\mathcal{C}}_A/dt$ , while  $d\mathcal{C}_A/dt$  is completely independent of  $\delta$  after the counterterm is included in the gravitational action. In contrast, if  $\ell$  is chosen to be an IR scale, the leading UV divergence (5.12) is enhanced by the extra logarithmic factor  $\log(\ell/\delta)$ , and  $d\tilde{\mathcal{C}}_A/dt$  is independent of  $\delta$  while this UV cutoff explicitly appears in  $d\mathcal{C}_A/dt$ .

Of course, an interesting question is if either of these two behaviours is reflected in the QFT calculations of complexity. The effect of the reference scale on the complexity of the thermofield double state in a free scalar field theory was recently studied in [72]. In this case, the transient behaviour in the time evolution does exhibit a nontrivial dependence on the reference scale. However, there is no potentially divergent behaviour found either in the case that  $\ell \sim \delta$  or that  $\ell$  remains an arbitrary IR scale. We might add however that this mismatch may not be very surprising. In particular, we note that the spectrum of the free scalar is not ‘chaotic’ enough to produce the linear growth found for holographic complexity.

Another interesting comparison that one might make is with the results of the covariant regulator used for the BTZ black hole in [17]. In this case, the boundary of the WDW patch is defined by two timelike geodesics that originate at the past joint and reach out to  $r = r_{\text{max}}$  before falling back to the future singularity. While the leading UV divergent term takes the form  $\mathcal{V}(\Sigma)/\delta^{d-1}$ , there is an explicit  $\log \delta$  term appearing in the transient contributions to the rate of growth. Hence this regulator produces a result that is similar to the standard action calculations without the null surface term.

### Integrated complexity

We can also consider the behaviour of the integrated complexity, as shown for early and light shocks in BTZ black holes in figure 22. Comparing to the vacuum complexity, we start with the complexity of formation at  $t = 0$  but with an early shock wave. This is much larger than the complexity of formation of the unperturbed black hole, i.e., we are in the linearly rising regime in, e.g., the left panel of figure 7. The complexity remains constant up to the critical time given by eq. (5.5) or eq. (5.6). At (more or less) this time, the complexity matches that of the unperturbed black hole (of mass  $M_1$ ) and it begins to grow

---

<sup>31</sup>In the QFT calculations [33, 34], this choice could be motivated by the fact that it renders the unitary connecting the (unentangled) reference state and the vacuum state continuous in momentum. That is, the unitary approaches the identity when the momentum approaches the cutoff.



such that the evolution is indistinguishable from the unperturbed evolution of an eternal black hole. That is, the effect of inserting these early (but light) shock waves is to lift the value of the initial complexity and then it remains fixed for a (long) initial period. For later times, the complexity not only grows in the same manner as, but is also essentially equal to, that of the unperturbed TFD. We note that this is a feature of the symmetric time evolution (i.e.,  $t_L = t_R = t/2$ ), and we will discuss further how this is in accordance to quantum circuit models in the spirit of [15].

It is possible to interpret this behaviour in terms of summing two independent evolutions for the left and right boundary times.<sup>32</sup> For early enough shocks, when evolving the right boundary time while holding the left time fixed, the complexity begins increasing immediately at the late time limit, i.e.,  $d\mathcal{C}_A/dt_R \simeq 2M_2/\pi$ . In contrast, when evolving  $t_L$  with  $t_R$  fixed, the complexity decreases until  $t_L \sim t_w - t_{scr}^*$  and then makes a rapid transition to increasing with the same late-time growth rate. However, in both of these periods, the rate is governed by the mass of the past black hole, i.e.,  $d\mathcal{C}_A/dt_L \simeq \pm 2M_1/\pi$ .<sup>33</sup> For a light shock wave with  $M_2 \simeq M_1$ , summing these two behaviours together produces the initial period where the complexity is constant in the symmetric time evolution. We show this behaviour for the CA proposal in the BTZ geometry in figure 23. In this example, we set  $w = 1 + 10^{-5}$  and  $t_w T_2 = 6$ , and we also fix the boundary time which is not evolved at zero. We see in these examples that for times smaller than  $2(t_w - t_{scr}^*)$ , the left evolution is negative and opposite to the right evolution, such that it has vanishing rate in the symmetric case, as shown in figure 2. In addition, the behaviour for heavy shock waves follows from this discussion. For times smaller than  $2t_w$ , the left evolution contributes with  $\approx -2M_1/\pi$ , while the right one with  $\approx +2M_2/\pi$ . Hence the rate is proportional to the difference of masses in this initial phase of the symmetric evolution, as in figure 3. We will discuss below a quantum circuit model that explains this behaviour.

To close this discussion, we recast various results from section 3 in order to develop some analytic understanding of the fact that the integrated holographic complexities match in the perturbed and unperturbed black holes (when working with light shocks and symmetric time evolution). We have already noted that for light shocks, the complexity of formation as a function of  $t_w$  remains constant until the critical injection time  $t_{w,c}$  given by eqs. (3.67) and (4.41) for the CA and CV approaches, respectively. That is,

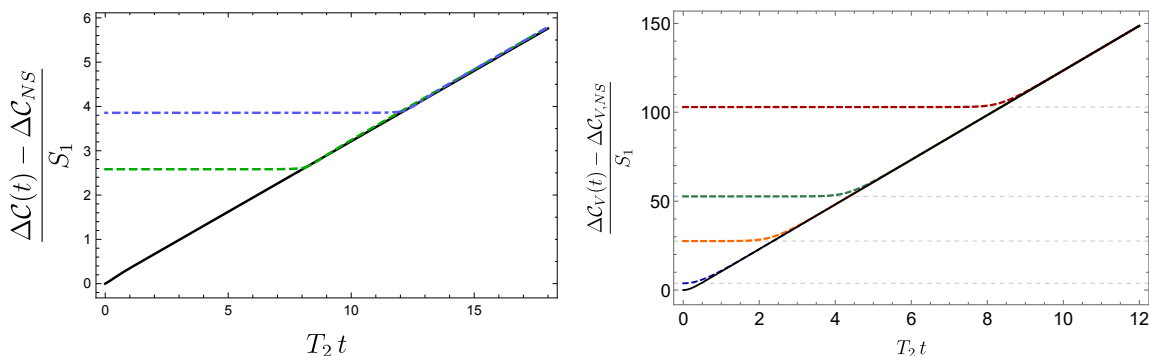
$$t_{w,c}^A = t_{scr}^*, \quad \text{and} \quad t_{w,c}^V = t_{scr}^* + \frac{1}{2\pi T_1} \log \frac{\gamma_2}{4}, \quad (5.15)$$

where  $\gamma_2 \simeq \pi$  (and further, we might note that  $\log(\pi/4) \simeq -0.24$ ). We therefore expect that for  $t_w > t_{w,c}$  the complexity of formation will be increased compared to the unperturbed complexity of formation by

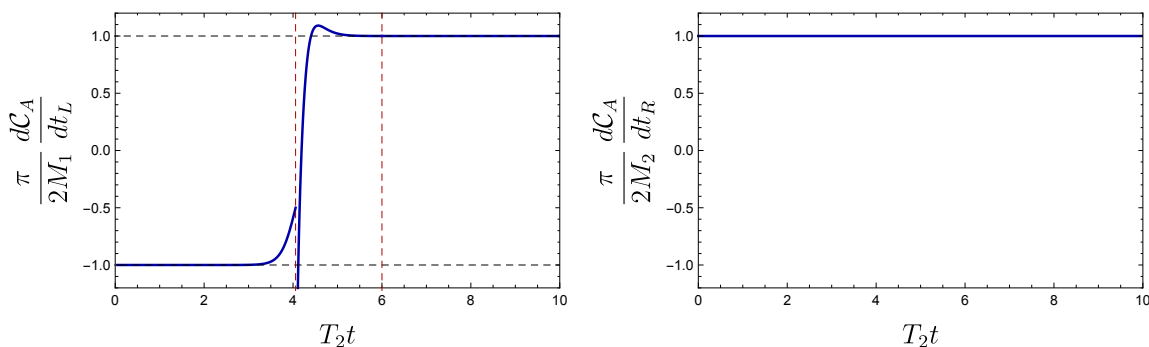
$$\Delta\mathcal{C}_A \simeq \Delta\mathcal{C}_{\text{unp}} + \frac{4M_1}{\pi} (t_w - t_{w,c}^A) \quad \text{and} \quad \Delta\mathcal{C}_V \simeq \Delta\mathcal{C}_{\text{unp}} + \frac{16\pi M_1}{(d-1)} (t_w - t_{w,c}^V). \quad (5.16)$$

<sup>32</sup>We thank Adam Brown for suggesting this explanation.

<sup>33</sup>Of course, similar behaviour occurs with the CV perspective, and can be explained using the scenario in figure 14b, where evolving only the right boundary time results in wrapping around the critical surface inside BH<sub>2</sub>, while evolving the left boundary time results in unwrapping the critical surface behind the past horizon of BH<sub>1</sub>.



**Figure 22.** Complexity evolution in  $d = 2$  for light shock waves. Left: using the CA proposal for  $w = 1 + 10^{-5}$  and  $\tilde{\ell}_{ct} = 1$ . The unperturbed evolution is indicated by a solid blue line. We also plot the complexity evolution in the presence of shock waves — with  $T_2 t_w = 6$  (dashed green) and 8 (dot-dashed light-blue). The initial values for those curves was fixed according to the complexity of formation. We see that the complexity does not change for a long period of time, and at late times, the complexity follows that of the unperturbed evolution. Right: time evolution of complexity using the volume conjecture for  $w = 1 + 10^{-5}$ . The unperturbed evolution is plotted in black and the evolution in the presence of the shock is in dashed colored lines —  $T_2 t_w = 6$  (red), 4 (green), 3 (orange) and 2 (blue). The evolution again begins with a plateau at early times and then joins the unperturbed evolution.



**Figure 23.** The time evolution of complexity using the CA proposal for  $d = 2$ . In the left panel, we fix  $t_R = 0$  and evolve  $t_L$ , while in the right panel, we fix  $t_L = 0$  and evolve  $t_R$ . In both cases, we have used  $w = 1 + 10^{-5}$ ,  $t_w T_2 = 6$  and in the left panel  $\tilde{\ell}_{ct} = 1$ . The vertical red lines in the left panel indicate the critical times at which  $r_m$  leaves the past singularity and later, when  $r_b$  enters the future one. For the right boundary time evolution, since the shock wave is sent early enough,  $r_s$  is already very close to  $r_{h,2}$  and the flat profile is due to the fact that fixing  $t_L = 0$  means that  $r_m$  stays behind the past singularity.

Similarly, we can also approximate the time derivative of the complexity as a step function, which changes at the critical time  $t_{c1}$  (CA) and  $t_{c,v}$  (CV), from a rate proportional to the difference of the masses (i.e., nearly vanishing for light shocks) to a rate proportional to the sum of the masses. Hence for times larger than this critical time, we have

$$\mathcal{C}_A(t) \simeq \Delta\mathcal{C}_A + \frac{2M_1}{\pi} (t - t_{c1}) \quad \text{and} \quad \mathcal{C}_V(t) \simeq \Delta\mathcal{C}_V + \frac{8\pi M_1}{(d-1)} (t - t_{c,v}). \quad (5.17)$$

Combining these two equations together yields

$$\mathcal{C}_A(t) \approx \Delta\mathcal{C}_{\text{unp}} + \frac{2M_1}{\pi} t, \quad \text{and} \quad \mathcal{C}_V(t) \approx \Delta\mathcal{C}_{\text{unp}} + \frac{8\pi M_1}{(d-1)} \left( t - \frac{1}{\pi T_1} \log \frac{\gamma_2}{2} \right), \quad (5.18)$$

again for times larger than the critical time. Note that this expression does not depend on  $t_w$  and as a consequence the time evolutions for different  $t_w$  all unify after a certain point. Further, the expression on the right hand side in both instances is approximately the time evolution of the unperturbed thermofield double. Hence the complexity of formation and the rate of growth are modified in such a way that at large times not only the rate of change of complexity matches the unperturbed result, but also the complexity itself. The extra  $\gamma_2$  factor in the CV result in fact helps increase the accuracy of this argument due to the initial transient regime of the unperturbed TFD.

A similar cancellation does *not* occur for heavier shocks (see figure 24), but one can use a similar reasoning to that above (neglecting the scrambling time) to show that the relative shift between the holographic complexities at late times is approximately given by

$$\begin{aligned} \Delta\mathcal{C}_A(t) - \mathcal{C}_{A,NS} &\approx \frac{M_1 + M_2}{\pi} t + \gamma_3 \\ \Delta\mathcal{C}_V(t) - \mathcal{C}_{V,NS} &\approx \frac{4\pi}{(d-1)} (M_1 + M_2)t + \frac{8\pi}{(d-1)} (M_2 - M_1)t_w, \end{aligned} \quad (5.19)$$

where the shift in the CA result is more complicated,

$$\gamma_3 \equiv \frac{2(M_1 + M_2)}{\pi} (t_w - t_{\text{del}}) - 2\frac{M_1}{\pi} t_{c1}, \quad (5.20)$$

where  $t_{c1}$  and  $t_{\text{del}}$  are given by eqs. (3.37) and (3.87), respectively. For heavy and early shock waves, using eqs. (3.61) and (3.89) for the BTZ black hole, the shift simplifies to

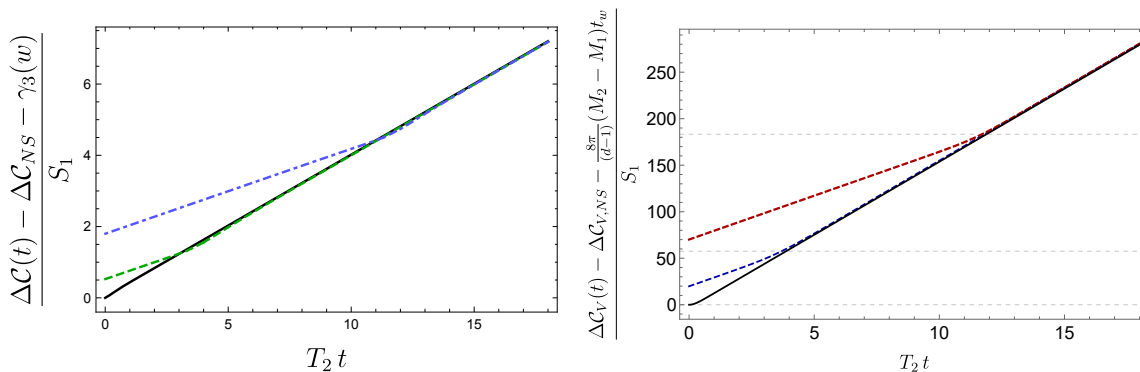
$$\gamma_3 \approx \frac{2(M_2 - M_1)}{\pi} t_w + \frac{4M_1}{\pi^2 T_2} - \frac{2(M_1 + M_2) \log 2}{\pi^2 T_2} + \mathcal{O}\left(\frac{\log w}{w}\right). \quad (5.21)$$

In figure 24, we explicitly show that the integrated complexities line up very closely after shifting the curves by the term proportional to  $(M_2 - M_1)t_w$  for both the CA and CV results.

### Complexity = (spacetime volume)

Recently it was also suggested that the holographic description of boundary complexity might be simply given by the spacetime volume of the WDW patch [58]. Hence this CSV proposal defines

$$\mathcal{C}_{\text{STV}} = \frac{\nu}{16\pi G_N L^2} \int_{\text{WDW}} d^{d+1}x \sqrt{-g}, \quad (5.22)$$



**Figure 24.** Complexity evolution in  $d = 2$  for heavy shock waves. Left: using the CA proposal for  $w = 2$  and  $\tilde{\ell}_{ct} = 1$ . We show the full complexity profile for two different times,  $T_2 t_w = 2$  (dashed green),  $T_2 t_w = 6$  (dot-dashed light blue). Right: time evolution of complexity using the volume conjecture for  $w = 2$ . The unperturbed evolution of an eternal BH of mass  $\frac{1}{2}(M_1 + M_2)$  is plotted in black and the evolution in the presence of the shock is in dashed colored lines for  $T_2 t_w = 6$  (red),  $T_2 t_w = 2$  (blue). The shift between the curves is proportional to  $(M_2 - M_1)t_w$  and matches the prediction of eq. (5.19).

where  $\nu$  is some numerical constant. It is straightforward to test the behaviour of this proposal in the present situation since the integrand is simply a constant in the bulk integral of the action, i.e.,  $\mathcal{R} - 2\Lambda = -2d/L^2$ . Therefore we can use our previous results for the bulk integrals to evaluate

$$\mathcal{C}_{\text{STV}} = -\frac{\nu}{2d} I_{\text{bulk}}. \tag{5.23}$$

For the symmetric time evolution, i.e.,  $t_R = t_L = t/2$ , we examine the growth rate for  $t < t_{c1}$  in the case of early shock waves, which can be read from eqs. (3.10) and (5.23), with  $r_m = 0$ ,  $r_b \rightarrow r_{h,1}$  and  $r_s \rightarrow r_{h,2}$ ,

$$\left. \frac{d\mathcal{C}_{\text{STV}}}{dt} \right|_{t_w \rightarrow \infty} = \frac{\nu}{2d(d-1)(1+kw^2z^2)} (M_2 - M_1), \tag{5.24}$$

which we can compare to the results for the CA and CV proposals in eq. (5.8). Further, for the late time limit, when  $r_m$  approaches  $r_{h,1}$  and  $r_s$  approaches  $r_{h,2}$ , we find

$$\left. \frac{d\mathcal{C}_{\text{STV}}}{dt} \right|_{t \rightarrow \infty} = \frac{\nu}{2d(d-1)(1+kw^2z^2)} (M_2 + M_1), \tag{5.25}$$

which we can compare to eq. (5.9).

It is straightforward to calculate the complexity of formation for planar black holes ( $k = 0$ ) in higher dimensions, following the analysis at the end of appendix B. In this case, we simply consider the large  $t_w$  regime of eq. (3.71) rescaled as in eq. (5.23) and use the tortoise coordinates in eq. (B.38) which results in the following simple expression

$$\left. \frac{d\Delta\mathcal{C}_{\text{STV}}}{dt_w} \right|_{t_w \rightarrow \infty} = \frac{\nu}{d(d-1)} (M_2 + M_1). \tag{5.26}$$

Next, we evaluate the dependence of the complexity of formation on  $t_w$  when the latter is close to zero, in which case  $r_s$  is close to the boundary. In this regime, we have from eq. (B.17)

$$\left. \frac{d\Delta\mathcal{C}_{\text{STV}}}{dt_w} \right|_{t_w \rightarrow 0} = \frac{\nu}{d(d-1)} (M_2 - M_1) . \quad (5.27)$$

In the limit where  $M_2 \gg M_1$ , both results can be related to the time evolution for the one-sided collapse, as discussed above for the CA and CV results, i.e., these rates can be matched with eq. (3.21) in [1], rescaled by the factor in eq. (5.23).

Despite having a different overall multiplicative constant, the general properties of complexity seem to be satisfied by the CSV proposal (5.22). In particular, with this approach, the holographic complexity exhibits the switchback effect for any boundary dimension (including  $d = 2$ ), and the late time rate of change has a smooth limit for light shock waves. Of course, given the simple relation of the spacetime volume to the bulk integral in the CA calculations, one can suggest another simple possibility. Namely, that the holographic complexity is described by the surface and joint terms in the gravitational action alone evaluated on the boundaries of the WDW patch. Our present calculations suggest that if we drop the bulk integral from eq. (3.2), the sum of the remaining surface and joint terms obey the expected properties of complexity, up to an overall normalization. Of course, to better understand this possibility and the CSV proposal more generally, it would be interesting to examine the results for background spacetimes in which matter fields deform the geometry in an interesting way. Of course, a simple example would be to compare the results of these new proposals to the results of the CA and CV proposals for charged black holes given in [23].

Let us also add that [58] suggested a connection between the CSV proposal (5.22) and using the ‘thermodynamic volume’ to define the complexity,<sup>34</sup> which may further hint at connections to the black hole chemistry program, e.g., see [60] for a review. Since the late time limit of geometries with two horizons, such as Reissner-Nordstrom black holes, reduces to a simple expression of differences of ‘internal energies’,<sup>35</sup> the authors of [58] suggested recasting complexity as a function of such extended thermodynamics variables. It would be an interesting future research direction to examine the physical consequences of these proposals for holographic complexity, in particular in the presence of shock waves.

### Circuit model

Next we would like to consider the connections of the behaviours observed in our holographic results to the switchback effect in more detail. Following the discussion of [15], we can interpret our results with some general considerations about quantum circuit models. As discussed in section 2, the boundary state of interest is the perturbed thermofield double state (2.4), in which the precursor

$$\mathcal{O}_R(-t_w) = U_R(t_w) \mathcal{O}_R U_R^\dagger(t_w) \quad (5.28)$$

---

<sup>34</sup>However, we must add that this connection was recently called into question by [59]. In particular, the simple calculations of [58] were shown to not apply for black holes with scalar hair.

<sup>35</sup>Strictly speaking the quantity associated to the inner Cauchy horizon does not have the usual thermodynamic interpretation, but nonetheless it is a useful identification to simplify the formulas. For example, see the early suggestions in [61, 62], and recent results in the context of Lovelock theories in [63].

is inserted in the right CFT (where  $U_{L,R} = \exp[-iH_{L,R}t]$  are the usual time evolution operators). Of course, if  $\mathcal{O}_R$  was the identity operator, nothing would change in the state since the unitaries  $U_R(t_w)$  and  $U_R^\dagger(t_w)$  would simply cancel in eq. (5.28). On the other hand, if  $\mathcal{O}_R$  is a localized simple operator,  $U_R(t_w)$  and  $U_R^\dagger(t_w)$  would still approximately cancel until times of the order of the scrambling time  $t_{\text{scr}}^*$ , when the effect of the perturbation  $\mathcal{O}_R$  has propagated throughout the system. However, the behaviour will be somewhat different for ‘heavy’ operators which inject a finite amount of energy into the system and allow the circuit to access many new degrees of freedom. Therefore we begin with a discussion of the simple operators and return to consider the heavy operators afterwards.

As discussed in section 2, evolving the perturbed state independently in the left and right times yields the expression in eq. (2.5),

$$|TFD(t_L, t_R)\rangle_{\text{pert}} = U_R(t_R + t_w) \mathcal{O}_R U_R(t_L - t_w) |TFD\rangle. \quad (5.29)$$

One immediate observation is that there are two time scales appearing here:  $t_R + t_w$  and  $t_L - t_w$ , precisely matching the holographic results in section 3. Of course, these are the invariant combinations that were naturally picked out by the time-shift symmetry described by eqs. (2.3) and (2.6). However, we would like to understand whether this perspective of the circuit models provides a deeper explanation of the behaviour of the holographic complexity.

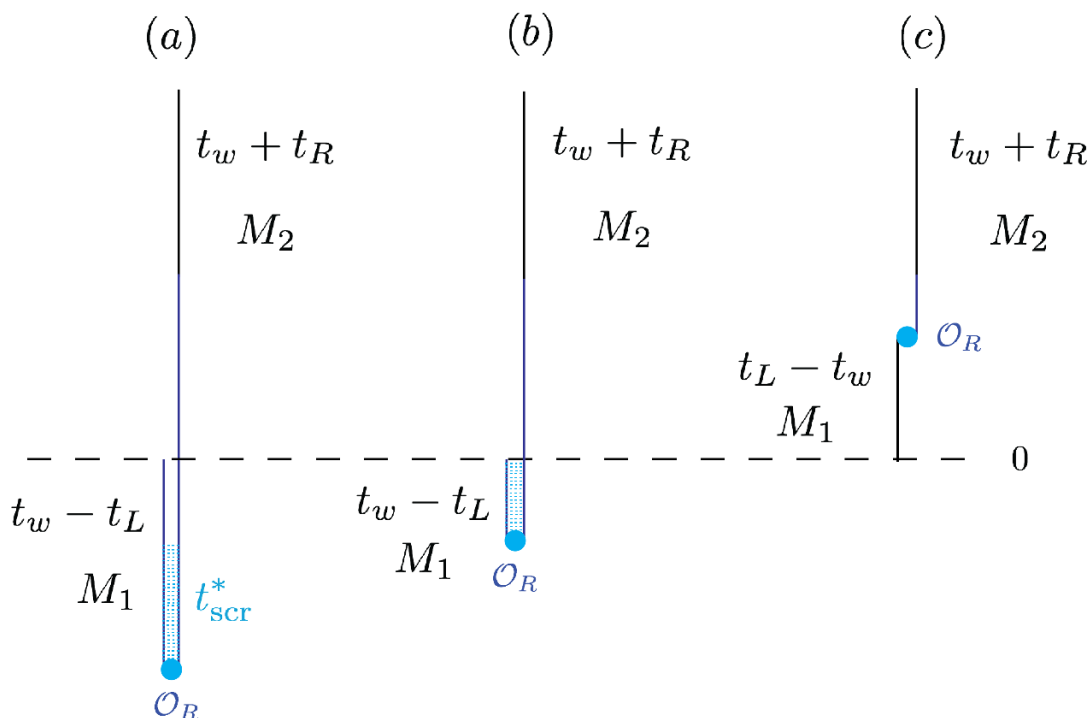
The time evolution of the TFD state perturbed by a simple operator  $\mathcal{O}_R$  is schematically portrayed in figure 25. Along each leg of these sketches, we assume that new gates are being laid out at a fixed rate in the circuit preparing the desired state [15]. Further, as is evident from the holographic results or can be argued on more general grounds [11, 14], the rate is expected to be proportional to the energy of the system.<sup>36</sup> However, after the operator  $\mathcal{O}_R$  is inserted, the evolution ‘folds back’ in circuit space and the circuit complexity experiences the switchback effect, as illustrated in the figure. That is, most of the gates laid out in (the final stages of) the initial evolution are canceled out as the second stage of the evolution begins. This cancellation of gates is only effective for the scrambling time  $t_{\text{scr}}^*$ . Hence as illustrated in the figure, there are three regimes of interest which are distinguished by the value of  $t_L - t_w$ .

In figure 25 (a), if  $t_w$  is very large with respect to  $t_L$ , the initial (i.e., furthest-to-the-right) operator evolves the state backwards by  $|t_L - t_w|$ , which we assume is bigger than the scrambling time. Then the second  $U_R$  carries the state forward again (assuming  $t_R$  is positive). The switchback effect comes into play and while the complexity grows on both legs of the evolution, the two time-evolution operators (at least partially) cancel out by an amount proportional to the scrambling time, as illustrated by the blue shaded region in the first panel of figure 25. That is, the complexity for the perturbation created by a simple operator grows as

$$\begin{aligned} t_L - t_w < -t_{\text{scr}}^* : \quad \mathcal{C}_{\text{pert}} &\approx 2M_1|t_L - t_w| + 2M_2(t_R + t_w) - 4M_1t_{\text{scr}}^* \\ &\approx 4M_1(t_w - t_{\text{scr}}^*) + \Delta M(t + 2t_w - 2t_{\text{scr}}^*), \end{aligned} \quad (5.30)$$

---

<sup>36</sup>More precisely, one would argue that the rate is proportional to  $TS$ , the product of the temperature and the entropy. However, for a CFT as in the holographic framework, this product is proportional to the energy. In the following, we set this rate to be  $2M$ , twice the mass of the dual black hole.



**Figure 25.** A representation of the insertion of a simple perturbation at an early time  $-t_w$  for the thermofield double state as in (5.29), in analogy to the construction in figure 6 of [15]. (a) The regime where  $t_L < (t_w - t_{\text{scr}}^*)$ , which corresponds to rate of change under symmetric time evolution proportional to the difference of masses ( $M_2 - M_1$ ). There is a cancellation in the time fold only during the scrambling time, which has to be accounted for in both sides of the evolution. (b) Transient regime that still represents a late time regime for light shock waves, such that  $t_L > (t_w - t_{\text{scr}}^*)$  but  $t_L < t_w$ , so there is some folding to an earlier time. However, if such folding is smaller than the scrambling time, there is an effective cancellation of the gates, and because  $M_2 \approx M_1$ , the complexity matches that in regime (c). (c) The late time behaviour where  $t_L > (t_w - t_{\text{scr}}^*)$  and  $t_L > t_w$ , so there is no folding backwards in the insertion of the operator.

where we substituted  $t_L = t_R = t/2$  for the symmetric time evolution studied in section 3 in the second line. Here, we have also kept small corrections of order  $\Delta M \equiv M_2 - M_1$ . Hence we see that  $d\mathcal{C}_{\text{pert}}/dt \sim \Delta M$  is proportional to the difference of masses, as found for the holographic complexity, e.g., in eq. (3.52). However, since  $M_2 \approx M_1$ , this rate is very close to zero and the complexity effectively remains constant.

The next regime corresponds to  $-t_{\text{scr}}^* < t_L - t_w < 0$ , as illustrated in part (b) of figure 25. In this range,  $U_R(t_L - t_w)$  evolves the state backward and  $U_R(t_R + t_w)$  evolves forward again. However, the switchback effect produces a cancellation for the duration of the first segment because it is less than the scrambling time. Hence the effective growth of the complexity is simply given by

$$\begin{aligned}
 -t_{\text{scr}}^* < t_L - t_w < 0 : \quad \mathcal{C}_{\text{pert}} &\approx 2M_2(t_R + t_w) - 2M_1|t_L - t_w|, & (5.31) \\
 &\approx 2M_1 t + 2\Delta M t_w,
 \end{aligned}$$

where again we substituted  $t_L = t_R = t/2$  and kept the corrections of order  $\Delta M$ . Therefore in this second regime, the rate of growth already matches that in the unperturbed thermofield double state, i.e.,  $d\mathcal{C}_{\text{pert}}/dt = d\mathcal{C}/dt$ .

Of course, the final regime is when  $t_L - t_w$  is positive, as sketched in part (c) of figure 25. In this case, both segments of the evolution move forward in time, and there is no opportunity for the switchback effect to modify the complexity and so the complexity simply grows as

$$\begin{aligned}
 0 < t_L - t_w : \quad \mathcal{C}_{\text{pert}} &\approx 2M_2(t_R + t_w) + 2M_1(t_L - t_w) & (5.32) \\
 &\approx 2M_1 t + \Delta M (t + 2t_w),
 \end{aligned}$$

where the second line corresponds to the symmetric time evolution. Again, in this third regime, the growth rate matches that of the unperturbed state.

Hence this simple model identifies two critical times for the symmetric time evolution after a simple perturbation, namely  $t_{c1} = 2(t_w - t_{\text{scr}}^*)$  and  $t_{c2} = 2t_w$ . Comparing to eqs. (5.5)–(5.7), we see that these times are in good agreement with our holographic results for light shocks in BTZ. Looking at the growth rate suggested by the circuit model, we see that there are essentially two phases. Initially, the growth rate is almost zero and at  $t = t_{c1}$ , the complexity begins to grow with the same rate of the unperturbed state. Of course, this behaviour is in good agreement with the holographic results where we can see a rapid rise from zero to  $2M_1$  after  $t = t_{c1}$ , as shown in the right panels of figures 2 and 16.

We can also compare the circuit model to the holographic results with a light shock for the complexity of formation by simply setting  $t_L = t_R = 0 = t$ . In this case, if we increase  $t_w$  from zero, we start in the (second) regime described by eq. (5.31). The switchback effect (almost) completely cancels the forward and backward evolution and hence the complexity and the complexity of formation are the same as in the unperturbed state. However, upon reaching  $t_w \simeq t_{\text{scr}}^*$ , we enter the (final) regime described by eq. (5.30). Hence the complexity of formation grows linearly with  $t_w$  for  $t_w \gtrsim t_{\text{scr}}^*$ . Again, this behaviour is in agreement with our holographic results discussed above, and e.g., as shown in eq. (3.84) for  $\Delta\mathcal{C}_A$ .

Our holographic calculations also considered heavy shock waves but in these cases, the perturbation is no longer dual to a simple operator. Rather the dual description would involve ‘heavy’ operators  $\mathcal{O}_R$ , which inject a finite amount of energy and allow the circuit to access new degrees of freedom. In this case, one does not expect a cancellation of the gates when the time evolution reverses. In particular, we can approximate the number of degrees of freedom before and after the perturbation as  $S_1 \sim M_1/T_1$  and  $S_2 \sim M_2/T_2$ . Following [19, 64], we might analyze the circuit after the time reversal in terms of an epidemic model, however, the size of the initial infection is now of order  $S_2 - S_1$ . Hence if  $S_2$  exceeds  $S_1$  by some finite factor, we expect that the infection rapidly spreads through all of the degrees of freedom, i.e., in a single time step — see further comments in the next subsection. In other words, the scrambling time in the above discussion is replaced by a much shorter delay time with

$$t_{\text{del}} \sim 1/T_2, \tag{5.33}$$



which matches our holographic results for heavy shocks, e.g., as in eqs. (3.61) and (4.31) for the CA and CV approaches, respectively.

Hence for heavy operators, the transition between regimes essentially occurs when  $t_L - t_w$  changes sign, i.e.,  $t = 2t_w$ . Following the above analysis of the circuit model, initially the complexity grows as

$$t_L < t_w : \quad \mathcal{C}_{\text{pert}} \approx 2M_1|t_L - t_w| + 2M_2(t_R + t_w) \tag{5.34}$$

$$\approx 2(M_2 + M_1)t_w + (M_2 - M_1)t,$$

where the second line describes the symmetric time evolution,  $t_L = t_R = t/2$ . Similarly in the second regime, the growth is instead described by

$$t_L > t_w : \quad \mathcal{C}_{\text{pert}} \approx 2M_1(t_L - t_w) + 2M_2(t_R + t_w) \tag{5.35}$$

$$\approx 2(M_2 - M_1)t_w + (M_2 + M_1)t.$$

Hence the rate of growth begins with  $d\mathcal{C}_{\text{pert}}/dt \propto (M_2 - M_1)$ , but then makes a transition to  $d\mathcal{C}_{\text{pert}}/dt \propto (M_2 + M_1)$  for  $t \gtrsim 2t_w$ . Of course, this agrees with the behaviour of the holographic complexity with heavy shocks, e.g., as shown in figures 3 and 17.

Setting  $t = 0$ , we can compare the complexity of formation in our holographic calculations for heavy shocks and in the simple circuit model. In particular, for very small injection times, i.e.,  $t_w \lesssim t_{\text{del}}$ , eq. (5.35) would apply, yielding  $d\Delta\mathcal{C}/dt_w \propto M_2 - M_1$ . However, the complexity rapidly transitions to the behaviour in eq. (5.34) where  $d\Delta\mathcal{C}/dt_w \propto M_2 + M_1$ . Of course, holography yields precisely this behaviour, as discussed in eqs. (5.2) and (5.3).

### Simple models

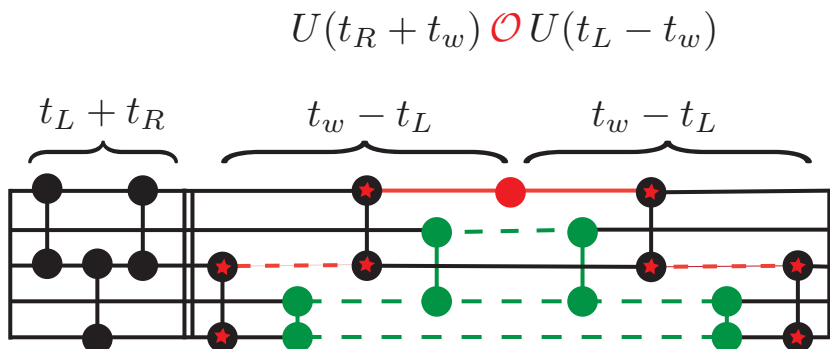
A simple model was proposed in [19, 64] for the evolution of the complexity in terms of an epidemic spreading of infected qubits when the system is evolved in time. The authors were considering the time evolution of a single qubit operator  $W$  given by the precursor  $W(t) = U(t)WU(-t)$  and the suggestion was that the number of infected qubits  $s(t)$  satisfies the following differential equation

$$\ell \frac{ds}{dt} = \frac{K - s}{K - 1} s \quad \longrightarrow \quad s(t) = \frac{K e^{t/\ell}}{K - 1 + e^{t/\ell}} \tag{5.36}$$

where  $K$  is the number of degrees of freedom and  $\ell$  is a characteristic time step in the circuit. The boundary condition for the solution on the right was chosen as  $s(t = 0) = 1$  since originally there was only a single infected qubit. Of course, this solution tends asymptotically to  $K$ . The complexity is then given by integrating over the infected qubits since these count the number of gates in the circuit which do not cancel out, and this leads to

$$\mathcal{C}_{\text{epidemic}}(t) = \frac{1}{\ell} \int_0^t dt s(t) = K \log \left( 1 + e^{(t-t^*)/\ell} \right) \tag{5.37}$$

where  $t^* = \ell \log K$  is the scrambling time. Comparing to the holographic results, one makes the natural identifications that  $K \sim S$  and  $\ell \sim \beta$ .



**Figure 26.** Illustration of the spread of infected qubits in the epidemic model when  $t_L + t_R > 0$  and  $t_w - t_L > 0$ .

It is not hard to generalize the above epidemic model to describe the complexity of the precursor  $U_R(t_R + t_w) \mathcal{O}_R U_R(t_L - t_w)$  appearing in eq. (5.29), with two independent time evolutions from the left and right sides of the perturbation. One interesting feature of the epidemic toy model compared to the previous subsection is that it yields naturally the scrambling time. In addition as we will see, it gives rise to a regime of suppressed exponential growth which is characteristic of chaotic systems and which can be observed in our holographic results. The time evolution can again be pictured as sketched in figure 25. However, cases (a) and (b) will be treated together here as they both have  $t_L - t_w < 0$ , and we begin by describing the behaviour in this regime. Throughout the following, we will assume that  $t_L + t_R > 0$ , and of course, as in the holographic calculations, we assume  $t_R + t_w > 0$ .

The circuit relevant for  $t_L - t_w < 0$  is depicted in figure 26 where the simple operator  $\mathcal{O}_R$  perturbing the circuit is indicated by the red dot and infected qubits are indicated by red stars. The two qubit gates that cancel out on the two sides of the unitary evolution are colored in green. In this case, we have to account for the cancellation of gates and this replaces the upper limit of integration in eq. (5.37) by  $t_w - t_L$ . Of course, we should then add a relevant count of the gates in the part of the circuit  $U_R(t_R + t_w)$  that goes beyond  $t_w - t_L$ , i.e., in the final period of length  $(t_R + t_w) - (t_w - t_L) = t_R + t_L$ . This leads to the following result for the complexity in this simple epidemic model<sup>37</sup>

$$\mathcal{C}_{\text{epidemic}}(t_L, t_R) = K \log \left( 1 + e^{(t_w - t_L - t^*)/\ell} \right) + \frac{K}{2\ell} (t_L + t_R). \quad (5.38)$$

If we further restrict to the symmetric time evolution, we obtain

$$\mathcal{C}_{\text{epidemic}}(t) = K \log \left( 1 + e^{(t_w - \frac{t}{2} - t^*)/\ell} \right) + \frac{K}{2\ell} t. \quad (5.39)$$

Examining this result for a large insertion time  $t_w$ , we see two regimes. At early times  $t \ll 2(t_w - t^*)$ , we see that the complexity behaves as

$$\mathcal{C}_{\text{epidemic}}(t) \approx \frac{K}{\ell} (t_w - t^*) + K e^{-(t_w - t^*)/\ell} e^{\frac{t}{2\ell}}. \quad (5.40)$$

<sup>37</sup>The factor of two in the second term is the same one mentioned in footnote 9 of [19] and is due to the fact that  $K/2$  gates act in a unit time step.

Hence there are two contributions, first the constant and second a term which grows exponentially. However the latter growth is suppressed by the exponentially small prefactor  $e^{-(t_w-t^*)/\ell}$ . Therefore we see that the complexity is approximately constant in this regime and, as we will see below, equal to the complexity of formation. (We will come back to the tiny exponential growth later). This regime (of early times and a simple operator) corresponds to the one in figure 25 (a) and indeed, the leading (constant) behaviour above in eq. (5.40) matches that in eq. (5.30) if we identify  $K/\ell \approx 4M_1$ . Of course, the present epidemic model does not account for the order  $\Delta M$  contribution in eq. (5.30), while the previous simple circuit model does not account for the small exponential growth in eq. (5.40).

At later times  $t \gg 2(t_w - t^*)$ , we obtain

$$\mathcal{C}_{\text{epidemic}}(t) \approx \frac{K}{2\ell}t, \tag{5.41}$$

which restores the linear growth of the unperturbed evolution. This regime corresponds to the sketch in figure 25 (b). This result now matches that in eq. (5.31) with the previous identification, i.e., now the prefactor becomes  $K/(2\ell) \approx 2M_1$ . Again, eq. (5.41) does not describe the order  $\Delta M$  correction found in the circuit model discussion.

Of course, we can also consider the regime  $t_L - t_w > 0$ , which matches the sketch in figure 25 (c), in which case there are no cancellations (i.e., no switchback) and so the count of necessary gates is simply given by

$$\mathcal{C}_{\text{epidemic}}(t_L, t_R) = \frac{K}{2\ell}(t_L - t_w) + \frac{K}{2\ell}(t_w + t_R) = \frac{K}{2\ell}(t_L + t_R). \tag{5.42}$$

Of course, restricting to symmetric time evolution with  $t_L = t_R = t/2$ , yields  $\mathcal{C}_{\text{epidemic}}(t) = \frac{K}{2\ell}t$ , as in eq. (5.41). This case matches the result in eq. (5.32) for the circuit model, up to the order  $\Delta M$  correction.

Now we can also set  $t = 0$  in eq. (5.39) to compare with the complexity of formation, which reads<sup>38</sup>

$$\Delta\mathcal{C}_{\text{epidemic}}(t_w) = K \log \left( 1 + e^{(t_w-t^*)/\ell} \right). \tag{5.43}$$

Of course, while our notation indicates the complexity of formation associated with the precursor, this quantity cannot predict (the part of) the complexity of formation associated with the state  $|TFD\rangle$  appearing in the holographic calculations. In any event, initially with small  $t_w$ , i.e.,  $t_w \ll t^*$ , we obtain  $\Delta\mathcal{C}_{\text{epidemic}}(t_w) \approx Ke^{(t_w-t^*)/\ell}$ , which indicates an exponential growth with  $t_w$  but again this term is suppressed by an exponential factor  $e^{-t^*/\ell}$ . In our discussion of the circuit model, this regime corresponds to the one in figure 25 (b) and the result matches that in eq. (5.31). In the second regime with  $t_w \gg t^*$ , eq. (5.43) yields  $\Delta\mathcal{C}_{\text{epidemic}}(t_w) \approx K(t_w - t^*)/\ell$  which indicates a linear growth of the complexity of formation with respect to  $t_w$  after a delay of duration  $t^*$ . This regime corresponds to the one in figure 25 (a) and our result matches eq. (5.30).

We have seen that the epidemic model is in good agreement with the various different regimes of holographic complexity for light shocks. To obtain a more precise match, it is

---

<sup>38</sup>This result matches eq. (5.37) with the replacement  $t \rightarrow t_w$ . Of course, this is no surprise since with  $t_L = 0 = t_R$ , the precursors in question match after this substitution and equating  $W = \mathcal{O}_R$ .

natural to choose  $\ell = 1/\lambda_L$  where  $\lambda_L = \frac{2\pi}{\beta}$  is the (quantum) Lyapunov exponent of gravitational systems that saturates the bound on chaos [66]. We demonstrate in figure 27 that an exponential growth with this particular exponent is indeed present in our holographic results for BTZ black holes. In addition, our previous identification  $K/\ell \approx 4M_1$  now indicates that  $K$  is proportional to the entropy of the system. The scrambling time  $t^*$  then becomes approximately the fast scrambling time  $t_{\text{scr}}^* \sim \frac{\beta}{2\pi} \log S$  of black holes [49, 55].

Let us add that all of the cases considered above correspond to  $t_L + t_R > 0$ . If instead we take  $t_L + t_R < 0$ , but keep  $t_R + t_w > 0$  as in the holographic model, then the right side of the circuit in figure 26 is longer than the left part and the expression for the complexity becomes

$$\mathcal{C}_{\text{epidemic}}(t_L, t_R) = K \log \left( 1 + e^{(t_R + t_w - t^*)/\ell} \right) + \frac{K}{2\ell} |t_L + t_R|. \quad (5.44)$$

Further, let us mention that as discussed around eq. (5.33) when the energy of the insertion is large the infection will be as fast as a single time step in the system after the insertion which leads to  $t_{\text{del}} \sim 1/T_2$ . In the epidemic model this amounts to changing the boundary conditions to eq. (5.36) and starting with  $K_2 - K_1$  infected qubits as the initial condition for the circuit after the insertion.

In addition, the authors of [15, 65] proposed that a good approximation for the complexity could be derived by looking at the lengths of geodesics stretching across the Einstein-Rosen bridge of a BTZ black hole<sup>39</sup>

$$\mathcal{C}_{\text{simple}}(t_R, t_L) \equiv \tilde{K} \log \left[ \cosh \frac{\pi(t_L + t_R)}{\beta} + c \exp \left[ \frac{\pi}{\beta} (2t_w - 2t_{\text{scr}}^* + t_R - t_L) \right] \right]. \quad (5.45)$$

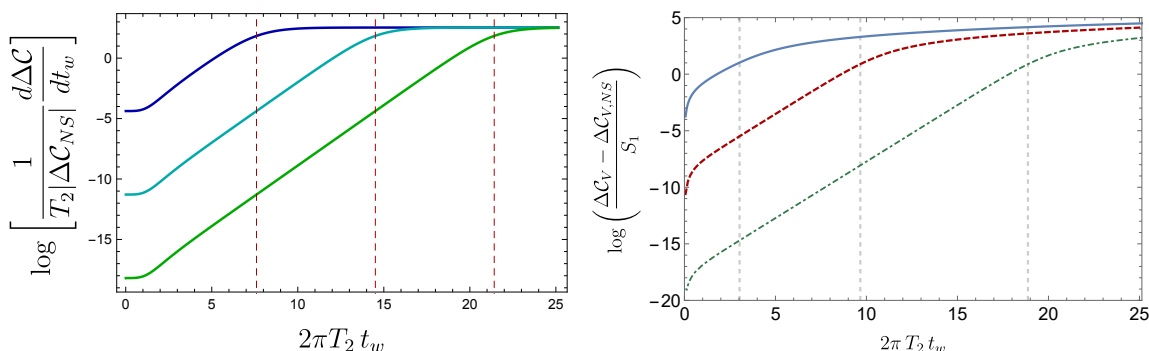
Above, the normalization constant  $\tilde{K}$  should again reflect the number of degrees of freedom while  $c$  is some order one constant. Note that this expression is very similar to the previous one discussed in the context of the epidemic model. When  $t_L + t_R \gg 0$  we may approximate  $\cosh \frac{\pi(t_L + t_R)}{\beta} \approx \frac{1}{2} \exp \left[ \frac{\pi(t_L + t_R)}{\beta} \right]$  which then yields the form (5.38) when identifying  $\tilde{K}$  and  $K$ . On the other hand, with  $t_L + t_R \ll 0$ , we may approximate  $\cosh \frac{\pi(t_L + t_R)}{\beta} \approx \frac{1}{2} \exp \left[ -\frac{\pi(t_L + t_R)}{\beta} \right]$  which then exactly takes the form (5.44). Hence, this expression (5.45) again produces the different behaviours described above.

First, let us consider the symmetric time evolution with  $t_L = t_R = t/2$ , this expression (5.45) reduces to

$$\mathcal{C}_{\text{simple}}(t) = \tilde{K} \log \left[ \cosh(\pi t/\beta) + c \exp \frac{2\pi}{\beta} (t_w - t_{\text{scr}}^*) \right]. \quad (5.46)$$

For early shocks (i.e.,  $t_w \gg t_{\text{scr}}^*$ ), the exponential term dominates the argument of the logarithm at early times. Therefore the holographic complexity is essentially constant until we reach the critical time  $t \sim 2(t_w - t_{\text{scr}}^*)$ , as described in the discussion around eq. (5.6) and in figure 22. After this critical time, there is a transition to a linear growth at late times with  $d\mathcal{C}_{\text{simple}}/dt \simeq \pi\tilde{K}/\beta$ . The latter agrees quantitatively with our holographic results in eq. (5.9) with  $\tilde{K}_A = 2S(d-1)/(d\pi^2)$  for the CA coefficient and  $\tilde{K}_V = 8S/d$  for the CV coefficient, both for planar black holes.

<sup>39</sup>An infinite (but state independent) constant was subtracted off in order to obtain a finite result [15].



**Figure 27.** Left: the derivative of the complexity of formation for BTZ black holes using the CA conjecture with respect to the insertion time  $t_w$ , for light shock waves characterized by  $w = 1 + 10^{-3}$  (dark blue),  $w = 1 + 10^{-6}$  (light blue) and  $w = 1 + 10^{-9}$  (green). We observe a period of exponential growth until times of the order of the scrambling time, which then becomes a linear growth at late times where the plot saturates. Right: early exponential dependence of the complexity of formation on the injection time  $t_w$  from the CV conjecture with  $w = 1 + 10^{-1}$  (blue),  $w = 1 + 10^{-4}$  (red dashed) and  $w = 1 + 10^{-8}$  (green, dot-dashed). We can read from this plot the correct Lyapunov exponent.

Alternatively, we may set  $t_L = t_R = 0$  to examine the contribution of eq. (5.45) to the complexity of formation. In this case, the above expression simplifies to  $\Delta\mathcal{C} \simeq K' \log\left[1 + c \exp\frac{2\pi}{\beta}(t_w - t_{\text{scr}}^*)\right]$ . Here the exponential dominates the argument of the logarithm for  $t_w > t_{\text{scr}}^*$  and in this regime of early shocks, the complexity of formation grows linearly with  $\Delta\mathcal{C} \sim 2\pi\tilde{K}(t_w - t_{\text{scr}}^*)/\beta$ , which once again matches the expectation of the complexity of formation results if  $\tilde{K}$  assumes the values discussed above for CA and CV. We want to emphasize that it is surprising that the simple expression in eq. (5.45), based on geodesics in BTZ, captures so many properties of holographic complexity so well. It would be interesting to better understand this agreement in the future.

While we described the complexity of formation as being constant in the regime  $t_w < t_{\text{scr}}^*$ , the previous discussion indicates an exponential growth with  $t_w$ , as is characteristic of the epidemic model. Of course, we want to stress that this growth is highly suppressed since the prefactor for this exponential carries a factor of  $\exp(-2\pi t_{\text{scr}}^*/\beta)$ . In figure 27, we examine  $\Delta\mathcal{C}$  in this initial regime carefully with a log plot, and we find that there is indeed an exponential growth, even though this is not at all evident in the original plots, e.g., the left panel of figure 19. Further, from the slope of the curves in figure 27, one can infer the correct Lyapunov exponent  $\lambda_L = \frac{2\pi}{\beta}$  (to a good degree of accuracy).

### Firewalls?

The strong sensitivity of the TFD state to small perturbations injected earlier than the scrambling time was already emphasized in [49], where it was pointed out that even a few thermal quanta of energy will be enough to completely distort the finely tuned correlations of the TFD state when sent early enough. In addition, as shown in the late time behaviour of the holographic complexity in figure 22, such deviations become indistinguishable at late times. In holography, this was explained by the fact that the energy of the shock wave is exponentially blueshifted as it falls to the event horizon [49]. Of course, the characteristic

time scale for this to happen is the scrambling time, and is interpreted as the time it takes for these early perturbations to have been scrambled throughout the system. This blueshift also led the authors of [10, 67] to draw connections between such perturbations and firewalls. The point being that the infalling quanta can be viewed as firewall by a (not-too late) infalling observer from the second boundary (i.e., the left boundary in our calculations). This also suggests that the appearance of firewalls depends on the system with which the black hole is entangled (e.g., the measurements made on the radiation exiting the black hole). One intriguing possibility is that the growth of complexity can serve as a diagnostic of firewalls [14, 56, 65], in particular in the context of the shock wave geometries. We have already mentioned that the complexity will actually decrease when only  $t_L$  is pushed forward while holding  $t_R = 0$  fixed<sup>40</sup> as long as  $t_L < t_w - t^*$ . The suggested interpretation of [56, 65] is that the complexity is decreasing as a function of  $t_L$  as long as the shock wave is within a Planck distance from the horizon along the surface and this is precisely a manifestation of a firewall which will be encountered by an observer jumping in from the left side.

### Future directions

In [1] by examining holographic complexity in one-sided Vaidya spacetimes, we found that the null surface counterterm (3.3) was an essential ingredient for the CA proposal (1.2). Our results here have reinforced this point. The most dramatic discrepancy was shown in section 3.3 where without the counterterm, the complexity of formation did not exhibit the switchback effect for  $d = 2$ . In section 3.3 and appendix B, we also found an unusual behaviour for the late time growth rate in the limit of very light shock waves. In particular without the counterterm, the growth rate approached the expected rate found for an eternal black hole, but at some characteristic time (3.96) related to the scrambling time, there was a transition to some more rapid growth, as illustrated in figure 10. The overall lesson here is the importance of testing various proposals for holographic complexity in dynamical spacetimes, such as the Vaidya geometries (2.7).

Some additional topics to explore include generalizing our results to perturbations in charged black hole backgrounds. Another route is to explore localized shocks as in [18], as well as null fluid collapses of finite thickness. In addition, very little is known about higher curvature corrections to properties of complexity [61, 63] in shock wave backgrounds. As we discussed previously in this section, it would be interesting to further investigate the complexity=(spacetime volume) conjecture (5.22), and under which circumstances these proposals diverge from the CA and CV conjectures. Of course, it would also be interesting to better understand the connection between this proposal and the thermodynamic volume [58, 59].

In addition, it would be interesting to investigate to what extent the holographic results can be reproduced by complexity calculations in free field theories. For example, the switchback effect seems to be a very robust feature, which is naturally associated with the geodesic deviation of adjacent trajectories [64] in negatively curved geometries.

---

<sup>40</sup>To be more precise, according to [65], using holographic complexity as a diagnostic of firewalls at the left horizon also requires sending  $t_R = \infty$  to avoid ambiguities.

Therefore, it would be interesting to investigate the complexity of precursors in a field theory context, e.g., using [33, 34] where we have already seen that negatively curved spaces can arise. The Vaidya geometries studied here have an interpretation in terms of a thermal quench, e.g., [68, 69], where some boundary coupling is rapidly varied at  $t_R = -t_w$ . Another interesting direction might be to combine the recent discussions of complexity in the thermofield double state [70–72] and in quantum quenches [73, 74] for Gaussian states to study the case of thermal quenches.

Further, it may be interesting to explore the implications of our results in the context of negative energy shock waves. In this case the negative rate of change of  $M_2 - M_1$  before the scrambling time should be the main effect. Such negative energy shock waves play an important role in the construction of traversable wormholes [75, 76] and it would be interesting to check whether the profile of complexity can serve to diagnose them.

In addition, there has been recent progress on defining complexity for general CFT setups. In particular, it was proposed in [77] that the CV proposal can be related to the quantum information metric. This can be used to find the length of a circuit connecting two ground states whose respective Hamiltonians differ by an insertion of a primary operator. It would be interesting to understand in which cases the quantum information metric could also be used to study the relative complexity of states before and after a global quantum quench. Another proposal [28, 29, 78] ties the complexity to the minimization of a functional given in terms of a generalized Liouville action. It would be interesting to understand how to generalize this proposal to generic time dependent backgrounds and understand if it can be used to study the time dependence of a state after a quantum quench in order to compare with the holographic results here and in [1].

## Acknowledgments

We would like to thank Alice Bernamonti, Adam Brown, Dean Carmi, Lorenzo Di Pietro, Federico Galli, Minyong Guo, Robie Hennigar, Juan Pablo Hernandez, Nick Hunter-Jones, Shan-Ming Ruan, Sotaro Sugishita, Brian Swingle, Beni Yoshida and Ying Zhao for useful comments and discussions. Research at Perimeter Institute is supported by the Government of Canada through the Department of Innovation, Science and Economic Development and by the Province of Ontario through the Ministry of Research, Innovation and Science. HM and RCM thank the Kavli Institute for Theoretical Physics for its hospitality at one stage of this project. At the KITP, this research was supported in part by the National Science Foundation under Grant No. NSF PHY17-48958. SC acknowledges support from an Israeli Women in Science Fellowship from the Israeli Council of Higher Education. RCM is supported by funding from the Natural Sciences and Engineering Research Council of Canada and from the Simons Foundation through the “It from Qubit” collaboration.

## A Counterterm for the null boundaries

As originally discussed in [21], the contributions to the gravitational action from the null boundaries give rise to various ambiguities. In particular,  $I_{\text{WDW}}$  will depend on the

parametrization of the null boundaries and so one must choose a ‘universal’ prescription which allows the comparison of this quantity evaluated for arbitrary boundary time slices in arbitrary bulk backgrounds. However, a simple alternative, which was also suggested in [21], is to add the following counterterm to the action on the null boundaries,

$$I_{\text{ct}} = \frac{1}{8\pi G_N} \int_{\mathcal{B}'} d\lambda d^{d-1}\theta \sqrt{\gamma} \Theta \log(\ell_{\text{ct}}\Theta), \tag{A.1}$$

where  $\ell_{\text{ct}}$  is some length scale, and  $\Theta$  is the expansion scalar of the null boundary generators, i.e.,

$$\Theta = \partial_\lambda \log \sqrt{\gamma}. \tag{A.2}$$

The expansion  $\Theta$  only depends on the intrinsic geometry of the null boundaries and so this surface term (A.1) plays no role in producing a well-defined variational principle for the gravitational action (3.2). However, as shown in [21], this counterterm ensures that the action is independent of the parametrization of the null boundaries. While adding this surface term does not modify the key features of holographic complexity in stationary spacetimes, it was argued in [1] and in the main text here that the counterterm is essential when applying the CA proposal to dynamical spacetimes.

Here we examine in detail the effect of adding the counterterm (A.1) to the gravitational action for the calculations in section 3. In particular, we will confirm that the inclusion of this term removes the dependence on the parametrization of the null boundaries, i.e., on the normalization of the null normal vectors. We also examine the role of the scale  $\ell_{\text{ct}}$  appearing in eq. (A.1) in the UV divergences and in the transient behaviour in the growth of the holographic complexity.

Let us begin by reviewing<sup>41</sup> the computation of the expansion (A.2) on the past null boundary on the right side of the WDW patch, i.e., the past boundary which crosses the shock wave in figure 1. Recall that the null normals are actually tangent vectors along the null boundaries, i.e.,  $k^\mu \partial_\mu = \partial_\lambda$ , and so  $\Theta$  is determined by the normalization of these vectors. For the past boundary extending to the right asymptotic AdS boundary, we can write the null normal (3.19) as

$$k_\mu^p dx^\mu = H(r, v) \left( -dv + \frac{2}{F(r, v)} dr \right) \tag{A.3}$$

where  $F(r, v)$  is the metric function appearing in eq. (2.7) and  $H(r, v)$  takes the form

$$H(r, v) = \alpha \mathcal{H}(r - r_s) + \tilde{\alpha} (1 - \mathcal{H}(r - r_s)), \tag{A.4}$$

Here  $\mathcal{H}$  denotes the Heaviside function and we leave  $\tilde{\alpha}$  unspecified for now. Further, we have  $dr/d\lambda = H(r, v)$  and hence the null expansion (A.2) becomes

$$\Theta = \frac{H(r, v)}{r^{d-1}} \frac{d}{dr} \left( r^{d-1} \right) = \frac{(d-1)H(r, v)}{r}. \tag{A.5}$$

---

<sup>41</sup>See section 2.3 of [1] for further details.



Hence the counterterm contribution (A.1) for  $\mathcal{B}_{\text{past}}$  can be written as

$$I_{\text{ct}}^{(I)} = \frac{\Omega_{k,d-1}(d-1)}{8\pi G_N} \int_{r_{\text{min}}}^{r_{\text{max}}} dr r^{d-2} \log\left(\frac{(d-1)\ell_{\text{ct}}H(r,v)}{r}\right), \quad (\text{A.6})$$

where we replaced  $d\lambda = dr/H(r,v)$ . The upper limit of the radial integral  $r_{\text{max}}$  is the position of the UV regulator surface. We set the lower limit  $r_{\text{min}} = r_m$  where  $r_m$  is the position of the intersection of the two past null boundaries, with the understanding that we set  $r_m = 0$  when these boundaries end on the past singularity. Hence using eq. (A.4), the integral in eq. (A.6) yields

$$I_{\text{ct}}^{(I)} = \frac{\Omega_{k,d-1}}{8\pi G_N} r_{\text{max}}^{d-1} \left[ \log\left(\frac{(d-1)\alpha\ell_{\text{ct}}}{r_{\text{max}}}\right) + \frac{1}{d-1} \right] - \frac{\Omega_{k,d-1}}{8\pi G_N} r_m^{d-1} \left[ \log\left(\frac{(d-1)\tilde{\alpha}\ell_{\text{ct}}}{r_m}\right) + \frac{1}{d-1} \right] + \frac{\Omega_{k,d-1}}{8\pi G_N} r_s^{d-1} \log\left(\frac{\tilde{\alpha}}{\alpha}\right). \quad (\text{A.7})$$

Upon substituting  $\tilde{\alpha}$  as given in eq. (3.22), the above result becomes the expression given in eq. (3.31).

The future null boundary on the left side of the WDW patch also crosses the shock wave at early times, i.e., for  $t_L < t_{L,c2}$ . Hence it is straightforward to carry out the above analysis with the corresponding null normal (3.24) and we find that the counterterm contribution becomes

$$I_{\text{ct}}^{(II)} = \frac{\Omega_{k,d-1}}{8\pi G_N} r_{\text{max}}^{d-1} \left[ \log\left(\frac{(d-1)\alpha\ell_{\text{ct}}}{r_{\text{max}}}\right) + \frac{1}{d-1} \right] + \frac{\Omega_{k,d-1}}{8\pi G_N} r_b^{d-1} \log\left(\frac{\hat{\alpha}}{\alpha}\right). \quad (\text{A.8})$$

Here we have assumed that we are only considering positive times when this boundary ends at  $r = 0$  (i.e., at the future singularity). Further, for  $t_L > t_{L,c2}$ , we would drop the second term above since this boundary no longer crosses the shock wave. Upon substituting  $\hat{\alpha}$  from eq. (3.26), this result matches that given in eq. (3.32).

Of course, it is also straightforward to evaluate the counterterm contributions for the null boundaries which are parallel to the trajectory of the null shell — see figure 1. For the past null boundary extending to the left asymptotic AdS boundary, we find

$$I_{\text{ct}}^{(III)} = \frac{\Omega_{k,d-1}}{8\pi G_N} r_{\text{max}}^{d-1} \left[ \log\left(\frac{(d-1)\alpha\ell_{\text{ct}}}{r_{\text{max}}}\right) + \frac{1}{d-1} \right] - \frac{\Omega_{k,d-1}}{8\pi G_N} r_m^{d-1} \left[ \log\left(\frac{(d-1)\alpha\ell_{\text{ct}}}{r_m}\right) + \frac{1}{d-1} \right], \quad (\text{A.9})$$

where again, we drop the second term for  $t_L < t_{L,c1}$ , i.e., in the regime where this boundary ends on the past singularity. Finally, for the future null boundary on the right side of the WDW patch, we have

$$I_{\text{ct}}^{(IV)} = \frac{\Omega_{k,d-1}}{8\pi G_N} r_{\text{max}}^{d-1} \left[ \log\left(\frac{(d-1)\alpha\ell_{\text{ct}}}{r_{\text{max}}}\right) + \frac{1}{d-1} \right]. \quad (\text{A.10})$$

Here we have assumed that we are in a regime where this boundary surface ends at the future singularity, as with the left future boundary in eq. (A.8).

To confirm that the inclusion of the counterterm (A.1) removes the dependence on the parametrization of the null boundaries, we should combine the four counterterm contributions given above with the joint contributions for the corresponding WDW patch. In the main text, we have already evaluated the joint contributions at  $r = r_s$ ,  $r_b$  and  $r_m$  in eqs. (3.21), (3.25) and (3.29), respectively. Here we are indicating the expressions for these joint contributions before  $\tilde{\alpha}$  and  $\hat{\alpha}$  were fixed.<sup>42</sup> We must also include the joint contributions arising where the null boundaries intersect the UV regulator surface at  $r = r_{\max}$ ,<sup>43</sup> and using the prescription given in [21, 26], we find

$$I_{\text{joint}}^{\text{UV}} = -\frac{\Omega_{k,d-1}}{4\pi G_N} r_{\max}^{d-1} \left[ \log \frac{\alpha}{\sqrt{f_1(r_{\max})}} + \log \frac{\alpha}{\sqrt{f_2(r_{\max})}} \right], \quad (\text{A.11})$$

where the first (second) term corresponds to the contribution from the UV joint at the left (right) asymptotic boundary.

Hence now combining the counterterm and the joint contributions, we find

$$\begin{aligned} I_{\text{ct}}^{\text{tot}} + I_{\text{joint}}^{\text{tot}} = & \frac{\Omega_{k,d-1}}{4\pi G_N} r_{\max}^{d-1} \left[ \log \left( \frac{(d-1)^2 \ell_{\text{ct}}^2 \sqrt{f_1(r_{\max}) f_2(r_{\max})}}{r_{\max}^2} \right) + \frac{2}{d-1} \right] \\ & - \frac{\Omega_{k,d-1}}{4\pi G_N} r_m^{d-1} \left[ \log \left( \frac{(d-1) \ell_{\text{ct}}}{r_m} \sqrt{|f_1(r_m)|} \right) + \frac{1}{d-1} \right] \\ & + \frac{\Omega_{k,d-1}}{8\pi G_N} r_s^{d-1} \log \left( \frac{f_1(r_s)}{f_2(r_s)} \right) - \frac{\Omega_{k,d-1}}{8\pi G_N} r_b^{d-1} \log \left( \frac{f_1(r_b)}{f_2(r_b)} \right). \end{aligned} \quad (\text{A.12})$$

Hence we see that the combined result is completely independent of the normalization constants appearing in the null normals, i.e.,  $\alpha$ ,  $\tilde{\alpha}$  and  $\hat{\alpha}$ . Of course, this is simply an explicit verification that introducing the counterterm (A.1) eliminates the dependence of  $I_{\text{WDW}}$  on the parametrization of the null boundaries [21].

At this point, we reiterate that we have left  $\tilde{\alpha}$  and  $\hat{\alpha}$  arbitrary above, rather than fixing them with the conditions, in eqs. (3.22) and (3.26), that the null boundaries are affinely parametrized across the shock wave. Hence we emphasize that the elimination of these normalization constants in eq. (A.12) was independent of any particular choice we might make for the parametrization of the correspond boundaries. We might also note that if we choose  $\tilde{\alpha} = \alpha$  and  $\hat{\alpha} = \alpha$  (e.g., as was done in [37, 38, 43]), the counterterm contributions at the two crossing points (i.e.,  $r = r_s$  and  $r = r_b$ ) vanish in eqs. (A.7) and (A.8). There remains a contribution at the meeting point  $r = r_m$  coming from eqs. (A.7) and (A.9). However, at late times, these terms make a vanishing contribution to the growth rate and so with this choice (for  $\tilde{\alpha}$  and  $\hat{\alpha}$ ), the joint terms alone capture many of the essential features of the time evolution of the complexity growth, e.g., as can be seen in comparing the results in [38] and [39].<sup>44</sup>

<sup>42</sup>Recall that for general values of the normalization constants,  $\tilde{\alpha}$  and  $\hat{\alpha}$ , the joint contributions in eqs. (3.21) and (3.25) account for the fact that  $\kappa$  is nonvanishing as the corresponding null boundaries cross the shock wave [1].

<sup>43</sup>There is no contribution from the joints where the null boundaries terminate on the future or past singularities because the area of these joints at  $r = 0$  vanishes.

<sup>44</sup>Both references study the complexity growth rate in hyperscaling violating geometries, but [38] does not introduce the counterterm and makes the choice  $\tilde{\alpha} = \alpha = \hat{\alpha}$ , while [39] uses the counterterm.

It is also straightforward to show in the present context of a shock wave propagating into an eternal black hole background, that the key results for the time evolution of the complexity are reproduced with  $\tilde{\alpha} = \alpha = \hat{\alpha}$ . However, the transient early time behaviour again exhibits some differences from the results in the main text where the counterterm is included.

### UV divergences with counterterm

One of the interesting effects of adding the counterterm (3.3) to the action of the WDW patch is that it seems to change the structure of the UV divergences in the corresponding holographic complexity, as first noted in [79]. We would like to review these changes and the leading UV divergences carefully here because it is relevant for the comparison of the holographic complexity with the complexity evaluated in quantum field theories [33, 34]. The latter comparison is considered in more detail in the main text in section 5.

Of course, the action of the WDW patch diverges because this region of spacetime extends all the way to the asymptotic AdS boundaries, e.g., as in figure 1. Hence we regulate our calculations as usual with a UV regulator surface at  $r = r_{\max}$ ,<sup>45</sup> e.g., see [80–82] and also the discussion for holographic complexity in [22, 26]. Of course, this radius can be expressed in terms of the short-distance cutoff  $\delta$  in the boundary theory, e.g., for the present Vaidya geometries (2.7), we have

$$r_{\max} = \frac{L^2}{\delta} \left( 1 - \frac{k}{4} \frac{\delta^2}{L^2} + \dots \right). \tag{A.13}$$

Hence the leading UV divergences in the holographic complexity take the form [26]

$$\tilde{C}_A^{\text{UV}} = \frac{I_{\text{grav}}^{\text{UV}}}{\pi} \simeq \frac{L^{d-1}}{4\pi^2 G_N} \frac{\mathcal{V}(\Sigma)}{\delta^{d-1}} \left[ \log\left(\frac{L}{\alpha \delta}\right) - \frac{1}{d-1} \right] + \dots, \tag{A.14}$$

where  $\mathcal{V}(\Sigma)$  is the total volume of the boundary time slice  $\Sigma$ , e.g.,  $\mathcal{V}(\Sigma) = 2 \Omega_{k,d-1} L^{d-1}$ , including both the left and right boundaries, for the constant time slices used in our calculations. The ellipsis indicates the subleading divergences which will involve integrals of geometric curvatures over the boundary time slice. Note that as indicated in eq. (A.14), we are only considering the contributions from eq. (3.2)<sup>46</sup> and so we have adopted the notation of eq. (3.90) since we are not including the counterterm contribution. An interesting feature of the UV divergences in eq. (A.14) is the appearance of the normalization constant  $\alpha$  in the logarithmic factor. We might add that this factor is essential for the interpretation

---

<sup>45</sup>Let us add that there are number of variations of this regulator procedure that one might consider [26]. For example, one might: a) choose the null boundaries of the WDW patch to be anchored to the desired time slice on the UV regulator surface; b) choose the null boundaries to be anchored to the time slice on the asymptotic AdS boundary (i.e.,  $r \rightarrow \infty$ ) but terminate the WDW patch at  $r = r_{\max}$ , including the GHY surface term on the small (timelike) portion of the regulator surface that becomes part of the boundary; and c) proceed as in (b) and also include the usual AdS boundary counterterms [80], as well as the GHY surface term, on the portion of the boundary at  $r = r_{\max}$ . These different choices will not change the universal features of the holographic complexity but we note that in fact, the UV divergences will agree for procedures (a) and (c). Further, we are implicitly using procedure (a) in the following, as in [26, 79].

<sup>46</sup>In particular, only the bulk integral and the joint at the cutoff surface are contributing to these UV divergences.

of this result as holographic complexity since consistency demands that the sum of the contributions in eq. (A.14) must be positive in order for  $\tilde{\mathcal{C}}_A$  to be positive [26].

However, the counterterm contributions must remove this  $\alpha$  dependence in eq. (A.14). Indeed combining the leading contributions from eqs. (A.7)–(A.10), we find

$$[\mathcal{C}_A^{\text{UV}}]_{\text{ct}} = \frac{L^{d-1}}{4\pi^2 G_N} \frac{\mathcal{V}(\Sigma)}{\delta^{d-1}} \left[ \log\left(\frac{(d-1)\alpha\ell_{\text{ct}}\delta}{L^2}\right) + \frac{1}{d-1} \right] + \dots, \quad (\text{A.15})$$

and then combining these UV contributions with eq. (A.14) yields

$$\mathcal{C}_A^{\text{UV}} = \tilde{\mathcal{C}}_A^{\text{UV}} + [\mathcal{C}_A^{\text{UV}}]_{\text{ct}} = \frac{L^{d-1}}{4\pi^2 G_N} \frac{\mathcal{V}(\Sigma)}{\delta^{d-1}} \log\left(\frac{(d-1)\ell_{\text{ct}}}{L}\right) + \dots. \quad (\text{A.16})$$

We emphasize that the counterterm removes the  $\alpha$  dependence from this leading divergence but also from all of the subleading divergences, as is evident from eq. (A.12). However, the ambiguity which  $\alpha$  introduced in eq. (A.16) has been replaced here by the ambiguity in specifying the counterterm scale  $\ell_{\text{ct}}$ .<sup>47</sup>

Let us add that the AdS scale  $L$  appears in two places in eq. (A.16). The first factor,  $L^{d-1}/G_N$ , yields the central charge  $C_T$  of the boundary theory, e.g., see [83]. However, the factor of  $L$  in the argument of the logarithm must be absorbed by  $\ell_{\text{ct}}$  in order for the final result (which is a quantity in the boundary theory) to be independent of the AdS scale.<sup>48</sup>

## B Complexity = action in higher dimensions

In this appendix, we examine higher dimensional examples of a shock wave in an eternal black hole geometry using the CA proposal. In section 3, we focused on the simple case of  $d = 2$  in detail since much of the analysis could be carried out analytically. Here, we begin by examining the case of  $d = 4$  (i.e., five bulk dimensions) in detail. There will be some interesting differences when comparing the behaviour of the AdS<sub>5</sub> black holes here to the BTZ black holes in section 3, as the sign of  $t_{c2} - t_{c1}$  in eq. (3.42) changes depending on the parameters  $t_w$  (how early the shock wave is inserted) and  $w$  (how heavy the shock wave is). As in section 3, we examine how the CA proposal is affected when we include or do not include the null surface counterterm (3.3) in the WDW action. We conclude the appendix by presenting some results for the complexity of formation (with and without the counterterm) for general  $d$  in the case of planar horizons (i.e.,  $k = 0$ ).

From eqs. (2.9) and (2.14), the parameters describing the  $d = 4$  boundary state dual to the AdS<sub>5</sub> black hole are

$$\begin{aligned} M &= \frac{3\Omega_{k,3}}{16\pi G_N} \omega^2, & \omega^2 &= r_h^2 \left( \frac{r_h^2}{L^2} + k \right), \\ T &= \frac{1}{2\pi r_h} \left( 2 \frac{r_h^2}{L^2} + k \right), & S &= \frac{\Omega_{k,3}}{4G_N} r_h^3. \end{aligned} \quad (\text{B.1})$$

<sup>47</sup>We note that this ambiguity was implicitly fixed in [79] by setting  $\ell_{\text{ct}} = L$ .

<sup>48</sup>It is straightforward to confirm that this factor of  $L$  is the AdS scale, and not the boundary curvature scale, following [23, 84].

From eqs. (2.10) and (2.11), the blackening factor becomes

$$\text{for all } v_L \ \& \ v_R < -t_w : \quad F(r, v) = f_1(r) = \frac{r^2}{L^2} + k - \frac{\omega_1^2}{r^2}, \quad (\text{B.2})$$

$$v_R > -t_w : \quad F(r, v) = f_2(r) = \frac{r^2}{L^2} + k - \frac{\omega_2^2}{r^2}. \quad (\text{B.3})$$

Then following eqs. (2.15) and (2.16), the tortoise coordinates in the different regions of the black hole geometry become: for all  $v_L$  and  $v_R < -t_w$ ,

$$r_1^*(r) = \frac{L^2}{2(2r_{h,1}^2 + kL^2)} \left[ 2\sqrt{r_{h,1}^2 + kL^2} \tan^{-1} \left( \frac{r}{\sqrt{r_{h,1}^2 + kL^2}} \right) - \pi\sqrt{r_{h,1}^2 + kL^2} + r_{h,1} \log \left( \frac{|r - r_{h,1}|}{r + r_{h,1}} \right) \right]; \quad (\text{B.4})$$

and  $v_R > -t_w$ ,

$$r_2^*(r) = \frac{L^2}{2(2r_{h,2}^2 + kL^2)} \left[ 2\sqrt{r_{h,2}^2 + kL^2} \tan^{-1} \left( \frac{r}{\sqrt{r_{h,2}^2 + kL^2}} \right) - \pi\sqrt{r_{h,2}^2 + kL^2} + r_{h,2} \log \left( \frac{|r - r_{h,2}|}{r + r_{h,2}} \right) \right]; \quad (\text{B.5})$$

Recall that the integration constants are chosen here such that  $\lim_{r \rightarrow \infty} r_{1,2}^*(r) = 0$ . Using the dimensionless coordinates in eq. (2.18), eq. (2.19), the ratios of the masses and entropies before and after the shock wave become in  $d = 4$

$$\frac{M_2}{M_1} = w^4 \frac{(1 + kz^2)}{(1 + kz^2w^2)} \quad \text{and} \quad \frac{S_2}{S_1} = w^3. \quad (\text{B.6})$$

### Early time analysis

One interesting difference for higher dimensional AdS black holes, with respect to (three-dimensional) BTZ black holes, is that the spacetime singularities ‘bow’ into the Penrose diagram [54]. As a result, when studying complexity=action for such (unperturbed) black holes (i.e., with  $d \geq 3$ ), there is an initial period during which the WDW patch touches both the future and past singularities and the holographic complexity remains constant, as discussed in [17, 23]. This geometric property also produces some interesting new features in the early time evolution of the holographic complexity when we introduce shock waves in these higher dimensional black holes.

In section 3, we introduced two critical times in the evolution of the CA complexity for  $t \geq 0$ . The critical time  $t_{c1}$  in eq. (3.37) determines the time when the WDW patch lifts off of the past singularity (i.e., when  $r_m = 0$ ). Hence for  $t > t_{c1}$ ,  $r_m$  is a dynamical variable. The critical time  $t_{c2}$  in eq. (3.40) determines the time when the point where the shock wave hits the future singularity just moves inside of the WDW patch (i.e., when  $r_b = 0$ ). That is, for  $t > t_{c2}$ , the (left) future null boundary of the WDW patch does not cross the shock wave and so  $r_b$  is a dynamical variable only for  $t < t_{c2}$ . For the BTZ black

hole discussed in section 3,  $t_{c2}$  was always equal to  $2t_w$  and the difference between the critical times  $t_{c2} - t_{c1} > 0$  was always positive, which meant that there existed a regime with both  $r_m$  and  $r_b$  as dynamical variables. For the higher dimensional black holes, this is not always the case, and we will derive the relevant expressions below in this appendix.

Let us begin by evaluating the time derivative of the holographic complexity at  $t = 0$ . Note that from eq. (3.40), the condition that  $t_{c2} > 0$  implies that  $t_w > -2r_1^*(0)$ . Therefore, if the shock wave is not sent early enough (i.e., the latter inequality is not satisfied), only  $r_s$  is a dynamical variable in the time derivative of complexity, which we will call regime (a). If  $t_w > -2r_1^*(0)$ , then we have to consider both  $r_b$  as well as  $r_s$  as dynamical variables. This was always the case for the BTZ black holes studied in section 3. Here we denote this regime by (b). The rate of change of the holographic complexity in these two regimes reads

$$\frac{d\mathcal{C}_A^{(a)}}{dt} = \frac{M_2}{\pi} - \frac{M_1}{\pi} \frac{f_2(r_s)}{f_1(r_s)} - \frac{M_2}{2\pi} \frac{r_s^{d-2}}{\omega_2^{d-2}} f_2(r_s) \log \frac{f_1(r_s)}{f_2(r_s)}, \quad (\text{B.7})$$

$$\begin{aligned} \frac{d\mathcal{C}_A^{(b)}}{dt} &= \frac{M_2}{\pi} \left( 1 + \frac{f_1(r_b)}{f_2(r_b)} \right) - \frac{M_1}{\pi} \left( 1 + \frac{f_2(r_s)}{f_1(r_s)} \right) + \\ &+ \frac{M_1}{2\pi} \frac{r_b^{d-2}}{\omega_1^{d-2}} f_1(r_b) \log \frac{f_2(r_b)}{f_1(r_b)} - \frac{M_2}{2\pi} \frac{r_s^{d-2}}{\omega_2^{d-2}} f_2(r_s) \log \frac{f_1(r_s)}{f_2(r_s)}. \end{aligned} \quad (\text{B.8})$$

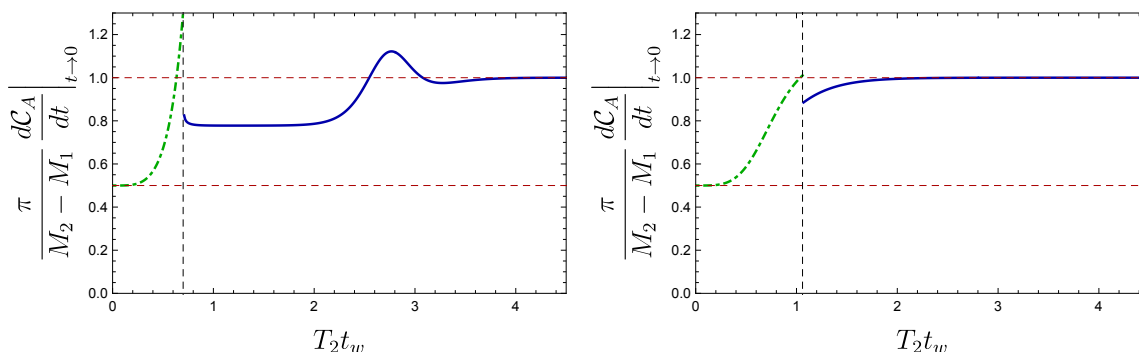
Now we consider the rate of complexity growth at  $t = 0$  in more detail for AdS<sub>5</sub> spherical  $k = 1$  black holes, with  $z$  defined in eq. (2.18) given by  $z = 1/w$ , which means that the smaller black hole with horizon radius  $r_{h,1}$  has the smallest stable horizon radius, right at the Hawking-Page transition. Of course, the overall conclusions are independent of the specific value of  $z$ , but we will focus on this example for concreteness. In the left panel of figure 28, we have a very light shock, with  $w = 10^{-6}$ , and we show the dependence on the perturbation time  $t_w$ . The vertical dashed line represents the time  $t_w = -2r_1^*(0)$  (i.e.,  $t_{c2} = 0$ ), which separates between regimes (a) and (b) in eqs. (B.7) and (B.8). When the shock wave is inserted at very early times (i.e., for large values of  $t_w$ ), the initial rate of change becomes the difference of masses  $\frac{M_2 - M_1}{\pi}$  (represented by the horizontal dashed red line). The right panel in figure 28 shows the analogous results for a heavier shock wave with  $w = 2$ . For the heavier shocks, the critical time  $t_{c2}$  grows in units of  $1/T_2$  as  $w$  grows, and once again the early growth rate approaches  $\frac{M_2 - M_1}{\pi}$  for early enough shock waves.

At  $t_w \rightarrow 0$ ,  $r_s$  approaches the AdS boundary. Therefore, if we expand eq. (B.7) in inverse powers of  $r_s$ , we obtain

$$\left. \frac{d\mathcal{C}_A^{(a)}}{dt} \right|_{t \rightarrow 0} = \frac{M_2 - M_1}{2\pi} + \mathcal{O}\left(\frac{1}{r_s^4}\right). \quad (\text{B.9})$$

For both spherical and planar black holes, the leading order contribution from eq. (B.9) is simply  $(M_2 - M_1)/2$ . In contrast, for large  $t_w$ , it approaches the difference of masses as  $r_s$  and  $r_b$  approach  $r_{h,2}$  and  $r_{h,1}$  respectively,

$$\left. \frac{d\mathcal{C}_A^{(b)}}{dt} \right|_{t \rightarrow 0} = \frac{M_2 - M_1}{\pi} + \mathcal{O}(T_2 t_w e^{-T_2 t_w}). \quad (\text{B.10})$$



**Figure 28.** The growth rate of the holographic complexity at  $t = 0$  for an  $\text{AdS}_5$  spherical black hole as a function of the perturbation time  $t_w$ . We choose  $z = 1/w$  (as defined in eq. (2.18)) so that the smaller black hole is at the HP transition. The left and right panels show the results for  $w = 1 + 10^{-6}$  and  $w = 2$ , respectively. The vertical dashed line indicates the transition from regimes (a) and (b) in eqs. (B.7) and (B.8). The rate of change starts at half of  $(M_2 - M_1)/\pi$  (lower dashed red line), as expected from the expansion in eq. (B.9), and for large  $t_w$  the rate of  $(M_2 - M_1)/\pi$  (indicated by the upper dashed red horizontal line) is reached.

### Time evolution in $\text{AdS}_5$

We now turn our attention to the full time evolution in spherical  $\text{AdS}_5$  black hole geometries with shock waves. As the previous early time analysis suggested, there are some interesting differences between the behaviour of higher dimensional black holes and the BTZ case discussed in section 3. It is still true that for large  $t_w$ , at first the complexity rate of change is approximately given by  $(M_2 - M_1)/\pi$  for a period of time of the order of  $2t_w$ . This is followed by a transient period, and then the final rate of  $(M_2 + M_1)/\pi$  is reached from above, analogously to the unperturbed case in [23]. However, there are two possible transient regimes depending on the sign of  $t_{c2} - t_{c1}$ , which we will analyze next.

As noted above, if the shock wave is sent early enough such that  $t_w > -2r_1^*(0)$ , the dynamics of the growth rate is parametrized by the positions  $r_b$  and  $r_s$ . Now, for light shocks,  $t_{c2} - t_{c1}$  is positive, and therefore there is a regime with  $r_m$ ,  $r_b$  and  $r_s$  contributing to the time derivative, as occurred for the BTZ black hole in section 3. However, for heavier shocks,  $t_{c2} - t_{c1}$  can be negative. That is,  $r_b$  disappears into the future singularity before  $r_m$  becomes dynamical. This leads to a different transition between early and late time behaviours. Of course, the dividing line between these different regimes is determined by solving  $t_{c2} - t_{c1} = 0$  in eq. (3.42), which yields

$$r_1^*(r_s) = 2r_1^*(0). \quad (\text{B.11})$$

Because generally  $r_s$  approaches  $r_{h,2}$  exponentially fast, and we are interested in a regime where  $t_w > -2r_1^*(0)$ , we can approximate the above equation as

$$r_1^*(r_{h,2}) \approx 2r_1^*(0). \quad (\text{B.12})$$

Despite having a simple form, it is still in general a transcendental equation. For AdS<sub>5</sub>, the above expression can be explicitly written as

$$\sqrt{k w_c^2 z_c^2 + 1} \left( \pi + 2 \cot^{-1} \left[ \frac{\sqrt{k w_c^2 z_c^2 + 1}}{w_c} \right] \right) - 2 \coth^{-1} w_c = 0, \quad (\text{B.13})$$

where we denote  $z_c$  and  $w_c$  the parameters at the transition between these regimes. For instance, if we denote  $z = 1/w_c$  as in the previous discussion, then for  $w > w_c$ ,  $t_{c2} - t_{c1} < 0$ . For  $k = 1$  and  $z_c = 1/w_c$ , we find that  $w_c \approx 1.00411$ . In order to probe these two regimes, we solve for a very light shock wave and a heavy shock wave.

If  $t_{c2} - t_{c1} > 0$ , the three regimes to be considered are the same as those discussed in eq. (3.43). If instead  $t_{c2} - t_{c1} < 0$ , the time evolution passes through the three following regimes:

$$\begin{array}{lll} \text{I} : & -t_{c0} < t < t_{c2} & r_b, r_s \text{ exist; } r_m < 0 \\ \text{II} : & t_{c2} < t < t_{c1} & r_s \text{ exists; } r_b, r_m < 0 \\ \text{III} : & t > t_{c1} & r_s, r_m \text{ exist; } r_b < 0. \end{array} \quad (\text{B.14})$$

The rate of change of complexity for regime I is again given by eq. (3.44), and for regime III, by eq. (3.46). For regime II, it is now given by

$$\frac{d\mathcal{C}_A^{(\text{II})}}{dt} = \frac{M_2}{\pi} - \frac{M_1}{\pi} \frac{f_2(r_s)}{f_1(r_s)} - \frac{M_2}{2\pi} \frac{r_s^{d-2}}{\omega_2^{d-2}} f_2(r_s) \log \left[ \frac{f_1(r_s)}{f_2(r_s)} \right]. \quad (\text{B.15})$$

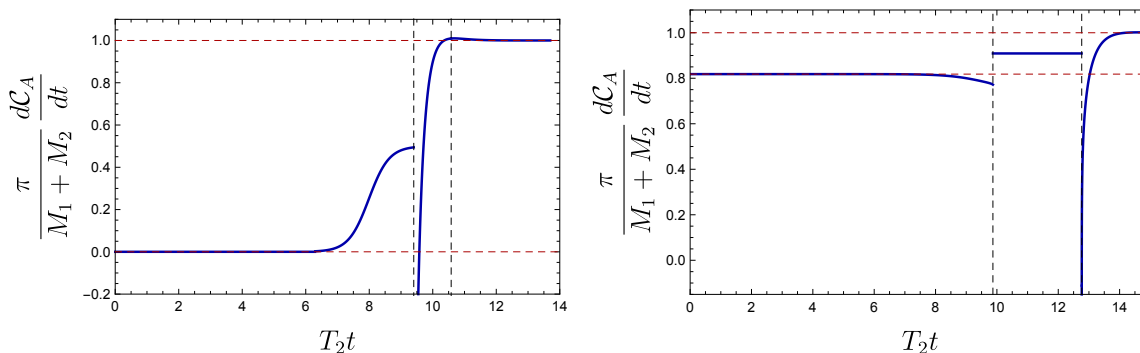
Notice that because generally  $r_s$  approaches  $r_{h,2}$  exponentially fast, the rate in the above expression will be approximately constant and equal to  $\frac{M_2}{\pi}$ .

Figure 29 shows the evolution of the complexity growth rate in an AdS<sub>5</sub> spherical black hole geometry, with  $z = 1/w$  for early shocks sent at  $t_w = 6/T_2$ , with a light energy of  $w = 1 + 10^{-4}$  (left) and a heavy one with  $w = 2$  (right). The lower dashed horizontal line indicates the limit  $(M_2 - M_1)/\pi$ , while the upper dashed line is  $(M_2 + M_1)/\pi$ , in this normalization. In the heavy shock wave example, for a long time, roughly of the order  $2t_w$ , the complexity growth rate is characterized by the difference of masses, then there is a transient regime (i.e.,  $t_{c2} < t < t_{c1}$ ) with a constant growth rate proportional to  $M_2$ , as predicted by eq. (B.15). For the light shock wave we have  $t_{c1} < t_{c2}$ , which means that the transient behaviour is analogous to the one that was studied for the BTZ black hole. At  $t = t_{c1}$  (first vertical line in the left panel and second in the right panel), when  $r_m$  emerges from the past singularity, there is a sharp and negative peak in the rate of change. Finally, the late time limit is approached from above, and the rate is proportional to the sum of the masses.

### Complexity of formation

We now consider the complexity of formation for planar (i.e.,  $k = 0$ ) black holes perturbed by a shock wave in asymptotically AdS<sub>5</sub> geometries. In [22], we studied the complexity of formation for unperturbed eternal black holes and found that for planar black holes,  $\Delta\mathcal{C}$  is proportional to the horizon entropy (i.e., the entanglement entropy between the two copies





**Figure 29.** Complexity growth rate for AdS<sub>5</sub> spherical ( $k = 1$ ) black hole with  $z = 1/w$ ,  $t_w = 6/T_2$  and  $\tilde{\ell}_{ct} = 1$ , light shock wave  $w = 1 + 10^{-4}$  to the left and heavy shock wave  $w = 2$  to the right. For the heavy shock wave,  $r_b$  disappears into the future singularity before  $r_m$  emerges from the past singularity. Around  $t = 0$ , there is a long plateau where the rate is  $(M_2 - M_1)/\pi$  (indicated by the lower dashed horizontal line). For the heavy shock wave, there is a transient regime ( $t_{c2} < t < t_{c1}$ ) where the rate is approximately constant and given by  $M_2/\pi$ . In both examples, when  $r_m$  emerges from the past singularity (right dashed vertical line), there is a sharp negative peak, then the late time limit of  $(M_2 + M_1)/\pi$  is approached from above (upper dashed horizontal line).

of the CFT in the thermofield double state).<sup>49</sup> Using the CA proposal with  $d = 4$  and  $k = 0$ , the unperturbed result reads

$$\Delta\mathcal{C}_A = \frac{S}{2\pi}. \tag{B.16}$$

Since the complexity of formation for stationary planar Schwarzschild-AdS black holes had this simple expression, we will investigate how this quantity behaves in the presence of shock waves.

In order to evaluate the complexity of formation in the Vaidya geometry, we need to examine separately the two regimes, (a) and (b), introduced above. First we consider regime (a), with  $r_b < 0$  at  $t = 0$  (and hence there are no contributions from the counterterm that depends on  $r_b$  to  $\Delta\mathcal{C}$ ). In this case, the contributions from the bulk integration are

$$\begin{aligned} \Delta I_{\text{bulk}}^{(a)} = & \left( \frac{\Omega_{k,d-1}}{16\pi G_N} \right) \left( -\frac{2d}{L^2} \right) \left[ \int_{r_s}^{r_{\text{max}}} dr r^{d-1} (-2r_2^*(r)) + \int_0^{r_{\text{max}}} dr r^{d-1} (-2r_1^*(r)) + \right. \\ & \left. + \int_0^{r_s} dr r^{d-1} (t_w + 2r_1^*(r_s) - 2r_1^*(r)) \right] - 2I_{\text{bulk, Vac}}. \end{aligned} \tag{B.17}$$

The Gibbons-Hawking contributions from the future and past singularities are given in eqs. (3.15) and (3.17), respectively, resulting in

$$\Delta I_{GHY}^{(a)} = \frac{d\Omega_{k,d-1}}{16\pi G_N} \left[ \omega_1^{d-2} (2r_1^*(r_s) - 4r_1^*(0)) - \omega_2^{d-2} (2r_2^*(r_s)) \right]. \tag{B.18}$$

<sup>49</sup>At high temperatures,  $\Delta\mathcal{C}$  could be expressed for the  $k = \pm 1$  cases as the entropy plus curvature corrections, in an expansion in inverse powers of  $LT$ , where here  $L$  stands for the curvature of the sphere in the boundary theory [22].

There are no joint contributions using affine parametrization across the shock wave. With the inclusion of the counterterm to the null boundary that crosses the shock wave at  $r_s$  given by eq. (3.31) with  $r_m = 0$ , we have

$$\Delta I_{\text{ct}}^{(a)} = \frac{\Omega_{k,d-1}}{8\pi G_N} \left[ r_s^{d-1} \log \left( \frac{f_1(r_s)}{f_2(r_s)} \right) \right]. \quad (\text{B.19})$$

Similarly, in the second regime (b), the contribution to the complexity of formation from the counterterm in the future boundary that crosses the shock wave at  $r_b$  is given by eq. (3.32),

$$\Delta I_{\text{ct}}^{(b)} = \frac{\Omega_{k,d-1}}{8\pi G_N} \left[ r_b^{d-1} \log \left( \frac{f_2(r_b)}{f_1(r_b)} \right) \right]. \quad (\text{B.20})$$

The complexity of formation in regime (a) is then the sum of the above contributions in eqs. (B.17), (B.18) and (B.19),

$$\Delta \mathcal{C}_A^{(a)} = \frac{\Delta I_{\text{bulk}}^{(a)} + \Delta I_{\text{GHY}}^{(a)} + \Delta I_{\text{ct}}^{(a)}}{\pi}. \quad (\text{B.21})$$

In the second regime (b), which occurs for larger values of  $t_w$ , the contributions to the complexity of formation are analogous to the expressions arising for the BTZ black hole discussed in section 3. That is, the result here is simply the sum of eqs. (3.71), (3.74), (B.19) and (B.20), now with  $d = 4$ .

In terms of the dimensionless coordinates (3.47), the final result for the complexity of formation for the perturbed planar ( $k = 0$ ) AdS<sub>5</sub> black holes reads

$$\begin{aligned} \Delta \mathcal{C}_A^{(a)} = \frac{S_1}{4\pi^2} \left[ w^3 \log \left( \frac{x_s + 1}{x_s - 1} \right) - 2w^3 \tan^{-1}(x_s) + 2 \tan^{-1}(wx_s) + \pi(w^3 + 1) \right. \\ \left. + 2w^3 x_s^3 \log \left( \frac{w^4 x_s^4 - 1}{w^4 (x_s^4 - 1)} \right) + \log \left( \frac{wx_s + 1}{wx_s - 1} \right) + 2 \log \left( \frac{wx_s - 1}{wx_s + 1} \right) \right], \end{aligned} \quad (\text{B.22})$$

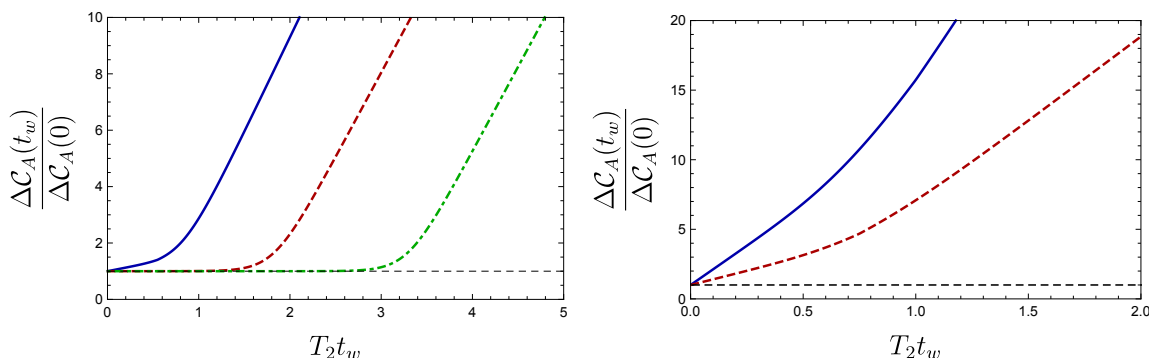
for regime (a), while for regime (b), the expression becomes

$$\begin{aligned} \Delta \mathcal{C}_A^{(b)} = \frac{S_1}{4\pi^2 w} \left[ 4w^4 \tan^{-1} \left( \frac{x_b}{w} \right) + 2w^4 \tan^{-1}(x_s) + 2w \tan^{-1}(wx_s) - 4 \tan^{-1}(x_s) \right. \\ \left. + 2wx_b^3 \log \left( \frac{x_b^4 - w^4}{x_b^4 - 1} \right) + (w^4 - 2) \log \left( \frac{x_s - 1}{x_s + 1} \right) + w \log \left( \frac{wx_s - 1}{wx_s + 1} \right) - 2x_b^4 \tan^{-1} \left( \frac{x_b}{w} \right) \right. \\ \left. + (x_b^4 - 2w^4) \log \left( \frac{x_b + w}{w - x_b} \right) + 2w(x_b^4 - 2w^4) \tan^{-1}(x_b) + (2w^5 - wx_b^4) \log \left( \frac{x_b + 1}{1 - x_b} \right) \right. \\ \left. - \pi \left( (w - 1)x_b^4 - 2w^5 + w^4 + w - 2 \right) - 8w^4 x_s^3 \log(w) + 2w^4 x_s^3 \log \left( \frac{w^4 x_s^4 - 1}{x_s^4 - 1} \right) \right]. \end{aligned} \quad (\text{B.23})$$

Recall that the cross-over between these two regimes occurs at  $t_w = -2r_1^*(0)$  with the tortoise coordinate given in eq. (B.4). Hence this transition occurs when

$$t_w = \frac{\pi L^2 \sqrt{r_{h,1}^2 + kL^2}}{2r_{h,1}^2 + kL^2}, \quad (\text{B.24})$$

which for  $k = 0$ , yields  $t_w = \pi L^2 / (2r_{h,1}) = 1 / (2T_1)$ .



**Figure 30.** The complexity of formation for planar AdS<sub>5</sub> black hole (i.e.,  $k = 0$ ) as a function of the shock wave time  $t_w$ . We normalize by the complexity of formation of the unperturbed black hole in eq. (B.16). In the left panel, we show  $\Delta\mathcal{C}_A(t_w)$  for light shock waves with  $w = 1 + 10^{-1}$  (solid blue),  $w = 1 + 10^{-4}$  (dashed red) and  $w = 1 + 10^{-8}$  (green dot-dashed). For very light shocks, the complexity of formation remains close to the unperturbed value until times of the order of  $t_{scr}^*$ , then increases approximately linearly with  $t_w$ . In the right panel, we show  $\Delta\mathcal{C}_A(t_w)$  for heavier shock waves, with  $w = 2$  (solid blue) and  $w = 1.5$  (dashed red). In these cases, the complexity of formation begins growing immediately as  $t_w$  moves away from zero.

Figure 30 combines these expressions to illustrate the complexity of formation for different temperature ratios  $w$  (which for planar black holes is simply  $w = \frac{T_2}{T_1}$ ). First, we investigate the behaviour of light shock waves in the left panel, with  $w = 1 + 10^{-1}$  (solid blue),  $w = 1 + 10^{-4}$  (dashed red) and  $w = 1 + 10^{-8}$  (green dot-dashed). We see a similar overall behaviour to the BTZ case in figure 7. The complexity does not change significantly from the unperturbed result until times of the order of the scrambling time  $t_{scr}^*$ , after which it grows approximately linearly with  $t_w$ . In the right panel of figure 30, we show  $\Delta\mathcal{C}_A$  for heavier shock waves. In this case, the complexity of formation starts growing immediately from the unperturbed value (B.16), but again we see that there are two distinct regimes.

To better understand the two regimes shown in figure 30, we want then to investigate the rate of growth of  $\Delta\mathcal{C}_A(t_w)$  from eqs. (B.22) and (B.23) for small and large values of  $t_w$ , respectively. One can find such slopes by differentiating eqs. (B.22) and (B.23) with respect to  $x_b$  and  $x_s$ , and finding the derivative of  $x_b$  and  $x_s$  with respect to  $t_w$  in eq. (3.4). As a consequence, we have the simple expressions for the slopes as

$$\left. \frac{d\Delta\mathcal{C}}{dt_w} \right|_{t_w \rightarrow 0} = \frac{M_2 - M_1}{\pi}, \quad (\text{B.25})$$

$$\left. \frac{d\Delta\mathcal{C}}{dt_w} \right|_{t_w \rightarrow \infty} = \frac{2(M_2 + M_1)}{\pi}. \quad (\text{B.26})$$

Hence we see that the initial slope depends on the difference in the masses and hence is essentially zero for the very light shock waves. The expression in eq. (B.25) can be compared to that for the BTZ black holes in eq. (3.85), which contains an additional term proportional to  $\log[M_2/M_1]$ . Similarly, eq. (B.26) can be compared to the slope implied in eq. (3.86) for  $d = 2$ , and in this case, the two slopes for large values of  $t_w$  are identical.

In analogy to eq. (3.86) for  $d = 2$ , we find that the large  $t_w$  behaviour of the complexity of formation follows

$$\Delta\mathcal{C}_A = \frac{S_1}{2\pi} + \frac{2}{\pi} (M_1 + M_2) (t_w - t_{\text{del}}) + \mathcal{O}(T_2 t_w e^{-T_2 t_w}), \quad (\text{B.27})$$

with the delay time given by

$$t_{\text{del}} \equiv \frac{1}{6\pi T_2 (w^4 + 1)} \left[ \pi(w^4 - 1) - 2(2w^4 - 1) \cot^{-1}(w) + 4\pi w - 2w \tan^{-1}(w) \right. \\ \left. + 8w^4 \log(w) - 2w^3(w + 1) \log\left(\frac{w^4 - 1}{8}\right) + (2w^4 + w - 1) \log\left(\frac{w + 1}{w - 1}\right) \right]. \quad (\text{B.28})$$

For light shock waves with  $w = 1 + \epsilon$ , the delay time agrees with the scrambling time at leading order,<sup>50</sup>

$$t_{\text{del}} = \frac{1}{2\pi T_1} \log\left(\frac{2}{\epsilon}\right) + \frac{1}{4T_1} + \mathcal{O}(\epsilon \log \epsilon) = t_{\text{scr}}^* + \frac{1}{4T_1} + \mathcal{O}(\epsilon \log \epsilon), \quad (\text{B.29})$$

which can be compared to eq. (3.88) for the BTZ case. For heavy shock waves, we have that the delay time approaches a constant proportional to  $1/T_2$ ,

$$t_{\text{del}} = \frac{1}{6\pi T_2} \left[ (\pi + 6 \log 2) - \frac{8}{w^3} \log w + \frac{3\pi + \frac{8}{3} + 6 \log 2}{w^3} \right] + \mathcal{O}\left(\frac{1}{T_2 w^4}\right), \quad (\text{B.30})$$

which can be compared to eq. (3.89) for the BTZ case. In figure 31 we show how this characteristic time  $t_{\text{del}}$  in eq. (B.28) generally behaves as a function of  $w$ . Overall, this behaviour is very similar to that for the BTZ black hole in figure 9.

We can also calculate the derivative of the complexity of formation with respect to  $t_w$  in order to show the transition between shock waves sent before and after the delay time defined in eq. (B.28). We show these results in figure 32. For heavy shock waves, we can see that initially the derivative begins as

$$\frac{d\Delta\mathcal{C}_A}{dt_w} = \frac{M_2 - M_1}{\pi} + \mathcal{O}(x_s^{-4}). \quad (\text{B.31})$$

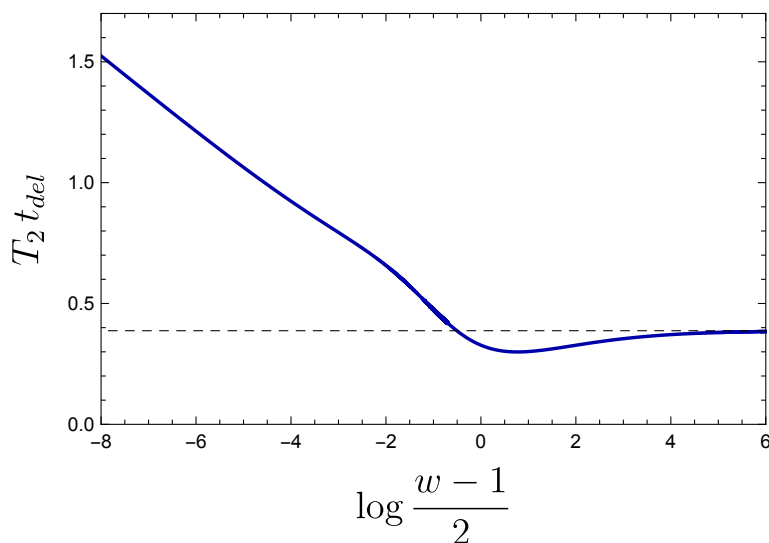
Further, there is a more pronounced regime in which this derivative remains constant for small  $t_w$ , in comparison to the BTZ results in figure 8. However, this plateau is never very long as it ends before  $t_{\text{del}}$ , which we see from eq. (B.30) that becomes  $T_2 t_{\text{del}} \simeq 0.39$  for large  $w$ . Another notable difference for the heavy shock wave is that the derivative has a discontinuity going from regime (a) to (b), given by eqs. (B.22) and (B.23).<sup>51</sup> It is given at leading order by

$$\left( \frac{d\Delta\mathcal{C}_A^{(b)}}{dt_w} - \frac{d\Delta\mathcal{C}_A^{(a)}}{dt_w} \right) \Big|_{t_w \rightarrow \frac{1}{2T_1}} = \frac{4(M_1 + M_2)}{\pi} \frac{\log w}{(1 + w^4)}. \quad (\text{B.32})$$

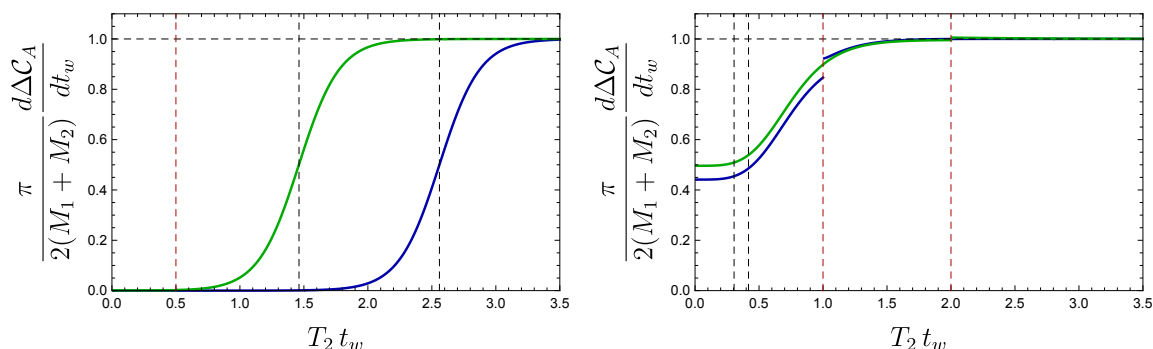
As a consequence, the jump for light shock waves is close to zero, making it imperceptible in figure 32.

<sup>50</sup>Combining the expressions in eq. (B.1), we have  $M \propto T^4$ ,  $S \propto T^3$  and  $M = \frac{3}{4} S T$  for the planar AdS<sub>5</sub> black holes (i.e.,  $k = 0$ ). Then we have  $4\epsilon \simeq E/M_1$  where  $E$  is the energy in the shock wave and so  $2/\epsilon = S_1$  if we choose  $E = 6T_1$ .

<sup>51</sup>However, the complexity of formation is continuous, as can be seen from figure 30.



**Figure 31.** The delay time in eq. (B.28) as a function of  $\log((w - 1)/2)$ . The left part of the plot is linear with a slope minus one, as is characteristic of the scrambling time in eq. (B.29). For heavy shock waves,  $t_{del}$  approaches a constant at large  $w$ , as given in eq. (B.30).



**Figure 32.** The derivative of the complexity of formation with respect to the insertion time  $t_w$ , for planar AdS<sub>5</sub>. In the left panel, we evaluate for a light shock waves with  $w = 1 + 10^{-6}$  (blue) and  $w = 1 + 10^{-3}$  (green), while in the right panel, we have heavier shock waves with  $w = 2$  (blue) and  $w = 4$  (green). The results for light shock waves resemble those for the BTZ black hole in figure 8, with a clear transition between two regimes at the scrambling time, indicated by the vertical dashed black lines. For the heavier shock waves, even though the complexity is a continuous function of  $t_w$  as the regime makes a transition between regimes (a) and (b), as in eqs. (B.22) and (B.23), the derivative has a non-zero jump proportional to  $\log w$ , as shown in eq. (B.32). The vertical dashed black line denotes the transition from regimes (a) to (b), while the dashed red line stands for the delay time in eq. (B.28). Notice that for heavy shock waves, there is a longer period with a constant derivative, in contrast to the BTZ result in figure 8. The profile for heavier shock waves is very similar to the  $w = 4$  example.

### Complexity without the counterterm

We briefly discuss the consequences of not adding eq. (3.3) to the WDW action for  $d = 4$ . The relevant rates of change of complexity in this regime were calculated and discussed in section 3.3. For instance, in the planar case with  $k = 0$  and  $d = 4$ , the rate of change of complexity for very early shock waves in eq. (3.94) become

$$\left. \frac{d\tilde{\mathcal{C}}_A^{(I)}}{dt} \right|_{t_w \rightarrow \infty} = \frac{M_2 - M_1}{3\pi} + \mathcal{O}\left(T_1(2t_w - t)e^{-\pi T_1(2t_w - t)}\right). \quad (\text{B.33})$$

It is proportional to  $M_2 - M_1$  and not simply 0 as for the BTZ black hole. In addition, the late time growth of complexity from eq. (3.97) reads

$$\left. \frac{d\tilde{\mathcal{C}}_A^{(III)}}{dt} \right|_{t \rightarrow \infty} = \frac{1}{\pi} \left( M_1 + \frac{M_2}{3} + \frac{5}{6} M_1^{3/4} M_2^{1/4} \right) + \mathcal{O}\left(T_1 t e^{-\pi T_1(t - 2t_w)}\right), \quad (\text{B.34})$$

which does not reproduce the expected late time behaviour with  $d\mathcal{C}_A/dt \propto M_1 + M_2$ . Further, if the shock wave is very light, with  $M_1 \approx M_2$ , we have

$$\left. \frac{d\tilde{\mathcal{C}}_A^{(III)}}{dt} \right|_{M_1 \approx M_2, t \rightarrow \infty} = \frac{13}{6\pi} M_2, \quad (\text{B.35})$$

which does not recover the expected eternal black hole rate of  $2M_2/\pi$ . The behaviour here is similar to that discussed in the context of the order of limits of light shock wave, without the addition of the counterterm in section 3.3.

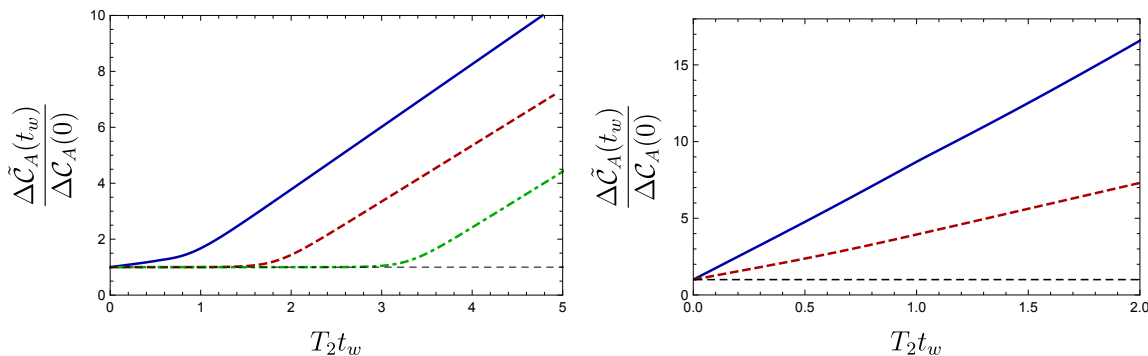
Let us also briefly consider the complexity of formation without the inclusion of the counterterm. The only contributions are from the bulk and Gibbons-Hawking-York boundaries in eqs. (B.17) and (B.18)

$$\begin{aligned} \Delta\tilde{\mathcal{C}}_A^{(a)} = \frac{S_1}{4\pi^2} & \left[ w^3 \log\left(\frac{x_s + 1}{x_s - 1}\right) - 2w^3 \tan^{-1}(x_s) \right. \\ & \left. + \log\left(\frac{wx_s - 1}{wx_s + 1}\right) + 2 \tan^{-1}(wx_s) + \pi(w^3 + 1) \right]. \end{aligned} \quad (\text{B.36})$$

In the second regime, adding eqs. (3.71), (3.74) and (3.75) with  $d = 4$ , the complexity of formation reads,

$$\begin{aligned} \Delta\tilde{\mathcal{C}}_A^{(b)} = \frac{S_1}{4w\pi^2} & \left[ -\pi((w-1)x_b^4 - 2w^5 + w^4 + w - 2) + (w^4 - 2) \log\left(\frac{x_s - 1}{x_s + 1}\right) \right. \\ & + w \log\left(\frac{wx_s - 1}{wx_s + 1}\right) - (2w^4 - x_b^4) \left( \log\left(\frac{w + x_b}{w - x_b}\right) + w \log\left(\frac{1 - x_b}{1 + x_b}\right) \right) + 2w \tan^{-1}(wx_s) \\ & \left. + 2(w^4 - 2) \tan^{-1}(x_s) + 2(2w^4 - x_b^4) \left( \tan^{-1}\left(\frac{x_b}{w}\right) - w \tan^{-1}(x_b) \right) \right]. \end{aligned} \quad (\text{B.37})$$

We show the complexity of formation without the addition of counterterm in figure 33. In contrast to the complexity of formation of BTZ black holes (i.e.,  $d = 2$ ) in figure 12, the overall behaviour of the complexity of formation, up to overall multiplicative constants, is similar to that with the inclusion of the counterterm in figure 30.



**Figure 33.** The complexity of formation for planar AdS<sub>5</sub> black hole as a function of the shock wave time  $t_w$ , without the inclusion of the counterterm. We normalize by the complexity of formation of the unperturbed black hole in eq. (B.16). In the left panel, we show  $\Delta\tilde{\mathcal{C}}_A(t_w)$  for light shock waves with  $w = 1 + 10^{-1}$  (solid blue),  $w = 1 + 10^{-4}$  (dashed red) and  $w = 1 + 10^{-8}$  (green dot-dashed). For very light shocks, the complexity of formation remains close to the unperturbed value until times of the order of  $t_{scr}^*$ , then increases approximately linearly with  $t_w$ . In the right panel, we show  $\Delta\tilde{\mathcal{C}}_A(t_w)$  for heavier shock waves, with  $w = 2$  (solid blue) and  $w = 1.5$  (dashed red). The overall behaviour is similar to the complexity of formation in figure 30, with the inclusion of the counterterm. Therefore, it contrasts with the BTZ black hole results in section 3.3.

### Complexity of formation for general dimensions (and $k = 0$ )

We have shown how the growth rate of complexity behaves in certain regimes in general dimension  $d$  in section 3, with and without the null surface counterterm. It is also possible to derive the dependence of the complexity of formation on  $t_w$  for general dimensions if we consider planar horizons, i.e.,  $k = 0$ . In the case of planar black holes, we can use the tortoise coordinates derived for any dimension  $d$  given in [22], which we rewrite here with the convention that  $r^*$  vanishes at infinity,

$$r_{in,i}^*(r) = \frac{L^2}{r} \left[ {}_2F_1 \left( 1, -\frac{1}{d}; 1 - \frac{1}{d}; \left( \frac{r}{r_{h,i}} \right)^d \right) - 1 \right] - \frac{\pi L^2}{dr_{h,i}} \cot \left( \frac{\pi}{d} \right), \quad (\text{B.38})$$

$$r_{out,i}^*(r) = -\frac{L^2}{(r^d - r_{h,i}^d)^{1/d}} {}_2F_1 \left( \frac{1}{d}, \frac{1}{d}; 1 + \frac{1}{d}; \frac{r_{h,i}^d}{r_{h,i}^d - r^d} \right).$$

In the first regime, we have  $t_w < -2r_1^*(0)$ , which for  $k = 0$  and general  $d$  gives

$$t_w < \frac{1}{2 \cot \left( \frac{\pi}{d} \right) T_1}. \quad (\text{B.39})$$

We solve eqs. (B.17) and (B.18) using the tortoise coordinates given by eq. (B.38), which has a long but analytic expression,

$$\frac{\pi}{S_1} \Delta\tilde{\mathcal{C}}_A^{(a)} = \frac{(d-2) \cot \left( \frac{\pi}{d} \right)}{d} - \frac{d}{\pi(d-1)} \left[ \left( w^d x_s^d - 1 \right)^{\frac{d-1}{d}} {}_2F_1 \left( \frac{1}{d} - 1, \frac{1}{d}; 1 + \frac{1}{d}; \frac{1}{1 - (w x_s)^d} \right) - w^{d-1} \left( x_s^d - 1 \right)^{\frac{d-1}{d}} {}_2F_1 \left( \frac{1}{d} - 1, \frac{1}{d}; 1 + \frac{1}{d}; -\frac{1}{x_s^d - 1} \right) \right] \quad (\text{B.40})$$

$$\begin{aligned}
 & + \frac{w^{d-1}}{2\pi} \frac{2x_s^d + d}{(x_s^d - 1)^{1/d}} {}_2F_1\left(\frac{1}{d}, \frac{1}{d}; 1 + \frac{1}{d}; -\frac{1}{x_s^d - 1}\right) \\
 & + \frac{1}{2\pi} \frac{2(wx_s)^{d-d}}{((wx_s)^d - 1)^{1/d}} {}_2F_1\left(\frac{1}{d}, \frac{1}{d}; \frac{1}{d} + 1; \frac{1}{1 - (wx_s)^d}\right).
 \end{aligned}$$

Here we have begun with the complexity of formation evaluated without the null surface counterterm (3.3). If the shock wave is very light, and  $t_w \ll t_{\text{scr}}^*$ , the above expression simply reduces to

$$\Delta\tilde{\mathcal{C}}_A^{(a)} = \frac{(d-2)\cot\left(\frac{\pi}{d}\right)}{d\pi} S_1, \tag{B.41}$$

which reproduces the complexity of formation for a planar boundary geometry in  $d$  dimensions found in [22].

The second regime, with  $t_w > -2r_1^*(0)$ , consists of solving eqs. (3.71) and (3.74) with the general  $d$  tortoise coordinates in eq. (B.38), which results in

$$\begin{aligned}
 \frac{\pi}{S_1} \Delta\tilde{\mathcal{C}}_A^{(b)} &= \frac{\Gamma\left(\frac{1}{d}-1\right)}{\pi d \Gamma\left(1+\frac{1}{d}\right)} \left[ \left(\frac{1}{(wx_s)^d - 1}\right)^{\frac{1}{d}-1} {}_2F_1\left(\frac{1}{d}-1, \frac{1}{d}; 1+\frac{1}{d}; -\frac{1}{(wx_s)^d - 1}\right) \right. \\
 & \quad \left. - \left(\frac{1}{(wx_s)^d - w^d}\right)^{\frac{1}{d}-1} {}_2F_1\left(\frac{1}{d}-1, \frac{1}{d}; 1+\frac{1}{d}; -\frac{1}{x_s^d - 1}\right) \right] \\
 & + \frac{dw^d - 2x_b}{2\pi x_b} \left[ {}_2F_1\left(1, -\frac{1}{d}; \frac{d-1}{d}; \left(\frac{x_b}{w}\right)^d\right) - {}_2F_1\left(1, -\frac{1}{d}; \frac{d-1}{d}; x_b^d\right) \right] \\
 & + \frac{2(wx_s)^d - d}{2\pi w} \left[ \left(\frac{w^d}{(wx_s)^d - 1}\right)^{1/d} {}_2F_1\left(\frac{1}{d}, \frac{1}{d}; 1+\frac{1}{d}; -\frac{1}{(wx_s)^d - 1}\right) \right. \\
 & \left. - \left(\frac{1}{x_s^d - 1}\right)^{1/d} {}_2F_1\left(\frac{1}{d}, \frac{1}{d}; 1+\frac{1}{d}; -\frac{1}{x_s^d - 1}\right) \right] + \frac{\cot\left(\frac{\pi}{d}\right)}{2dw} \left[ 2x_b^d + (d-4)w - w(2x_b - dw^d) \right].
 \end{aligned} \tag{B.42}$$

Of course, we can also evaluate the complexity of formation including the counterterm for general dimension and  $k=0$ . We would simply add the contributions in eq. (B.19) to eq. (B.40)

$$\Delta\mathcal{C}_A^{(a)} = \Delta\tilde{\mathcal{C}}_A^{(a)} + \frac{S_1}{2\pi^2} \left[ w^{d-1} x_s^{d-1} \log\left(\frac{((wx_s)^d - 1)}{w^d(x_s^d - 1)}\right) \right], \tag{B.43}$$

and eqs. (B.19) and (B.20) to eq. (B.42), which results in

$$\Delta\mathcal{C}_A^{(b)} = \Delta\tilde{\mathcal{C}}_A^{(b)} + \frac{S_1}{2\pi^2} \left[ x_b^{d-1} \log\left(\frac{x_b^d - w^d}{x_b^d - 1}\right) + w^{d-1} x_s^{d-1} \log\left(\frac{((wx_s)^d - 1)}{w^d(x_s^d - 1)}\right) \right]. \tag{B.44}$$

Note that the additional contributions in eq. (B.43) do not modify the result in the limit  $w \rightarrow 1$  and  $t_w \ll t_{\text{scr}}^*$ , i.e., we still recover the expected result for the complexity of formation (B.41) without any shock wave perturbation.<sup>52</sup>

For large  $t_w$ , we can simply expand the expressions for  $x_s = 1 + A \exp^{-2\pi T_2 t_w}$  and  $x_b = 1 - B \exp^{-2\pi T_1 t_w}$ , where  $A$  and  $B$  are constants that depend on the dimension, but

<sup>52</sup>We might note that the original calculations of the complexity of formation in [22] were made without the counterterm (3.3).



that nonetheless are independent of  $t_w$ . In addition, we can evaluate the dependence on  $t_w$  when  $t_w$  is small, which means  $r_s$  is close to the boundary. Evaluating the complexity of formation without counterterm in eqs. (B.40) and (B.42), we find

$$\left. \frac{d\Delta\tilde{\mathcal{C}}_A}{dt_w} \right|_{t_w \rightarrow 0} = \frac{d-2}{d-1} \frac{M_2 - M_1}{\pi}, \quad \left. \frac{d\Delta\tilde{\mathcal{C}}_A}{dt_w} \right|_{t_w \rightarrow \infty} = \frac{d-2}{d-1} \frac{M_1 + M_2}{\pi}. \quad (\text{B.45})$$

Also, we can evaluate the complexity of formation with the addition of the counterterm in eqs. (B.43) and (B.44), which in these regimes give us the expected results

$$\left. \frac{d\Delta\mathcal{C}_A}{dt_w} \right|_{t_w \rightarrow 0} = \frac{M_2 - M_1}{\pi}, \quad \left. \frac{d\Delta\mathcal{C}_A}{dt_w} \right|_{t_w \rightarrow \infty} = \frac{2(M_1 + M_2)}{\pi}, \quad (\text{B.46})$$

e.g., compare to the CA results in eqs. (5.2) and (5.3). Of course, in the limit where  $M_2$  is much larger than  $M_1$ , we should recover the one sided result in [1]. Without the addition of the counterterm, the planar rate of change is just a constant (with the same  $d$  dependence), as can be seen in eq. (3.28) of [1]. When the counterterm is added, the expressions also agree, with the initial rate of  $M_2/\pi$  and the late time rate of  $2M_2/\pi$ , as can be seen in eq. (3.47) of [1]. In addition, the general  $d$  dependence in eq. (B.45) agrees with our analysis of the BTZ black hole in section 3.3, where the complexity of formation without the inclusion of the counterterm saturated to a constant in  $d = 2$ .

### C Details for complexity = volume

In this appendix, we present the detailed derivation behind some of the results presented in section 4 for our maximal volume surfaces in the shockwave geometries. In particular, we will derive the expressions in eqs. (4.14) and (4.15) for the left and right boundary times, and in eqs. (4.16)–(4.18) for the maximal volume and its time derivative. Our derivations are simplified with the following definitions introduced in the main text

$$\tau [P, r] \equiv \frac{1}{f(r)} - \frac{P}{f(r)\sqrt{f(r)r^{2(d-1)} + P^2}}, \quad R [P, r] \equiv \frac{r^{2(d-1)}}{\sqrt{f(r)r^{2(d-1)} + P^2}}, \quad (\text{4.13})$$

as well as defining

$$\tilde{R} [P, r] \equiv \frac{\sqrt{f(r)r^{2(d-1)} + P^2}}{f(r)} - \frac{P}{f(r)}. \quad (\text{C.1})$$

In the following, we will add subscripts  $(\tau_{1,2}, R_{1,2})$  to specify which blackening factor  $f_i$  and conserved momentum  $P_i$  is being used. These functions obey a number of useful identities which we will use repeatedly throughout the derivation

$$R [P, r] = \tilde{R} [P, r] + P\tau [P, r] = \tilde{R} [-P, r] - P\tau [-P, r], \quad (\text{C.2a})$$

$$\partial_P \tilde{R} [P, r] = -\tau [P, r], \quad \partial_P \tilde{R} [-P, r] = \tau [-P, r], \quad (\text{C.2b})$$

$$\tilde{R} [P, r_t] = -\frac{P}{f(r)}, \quad (\text{C.2c})$$

$$\tilde{R} [P, r] = \tilde{R} [-P, r] - \frac{2P}{f(r)} = \dot{v}_+ [P, r] = -\dot{v}_- [-P, r], \quad (\text{C.2d})$$

where  $r_t$  denotes the turning point (a point in which  $\dot{r}$  vanishes) in the relevant black hole which satisfies (cf. eq. (4.12))

$$P^2 + f(r_t)r_t^{2(d-1)} = 0. \quad (\text{C.3})$$

In general the time integrals will be given by integrating the coordinates  $u$  and  $v$  along various parts of the surface according to

$$\begin{aligned} \Delta v_{\pm} &= \int dv_{\pm} = \int \frac{\dot{v}_{\pm}}{\dot{r}_{\pm}} dr = \int \tau[\pm P, r] dr \\ \Delta u_{\pm} &= \int du_{\pm} = \int \frac{\dot{u}_{\pm}}{\dot{r}_{\pm}} dr = - \int \tau[\mp P, r] dr \end{aligned} \quad (\text{C.4})$$

and volume integrals will be given by

$$\mathcal{V} = \Omega_{k,d-1} \int \frac{r^{2(d-1)}}{\dot{r}_{\pm}} dr = \pm \Omega_{k,d-1} \int R[P, r] dr. \quad (\text{C.5})$$

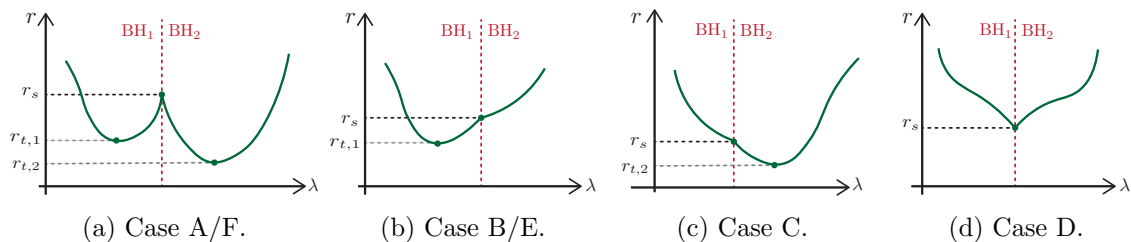
There are six different possibilities to be considered for the shape of our surface:

1. **Case A** — The surface passes behind the future horizon in  $\text{BH}_1$  ( $P_1 > 0$ ), it admits turning points both in  $\text{BH}_1$  and in  $\text{BH}_2$ , see figure 34a.
2. **Case B** — The surface passes behind the future horizon in  $\text{BH}_1$  ( $P_1 > 0$ ), it admits a turning point in  $\text{BH}_1$  but not in  $\text{BH}_2$ , see figure 34b.
3. **Case C** — The surface passes behind the future horizon in  $\text{BH}_1$  ( $P_1 > 0$ ), it admits a turning point in  $\text{BH}_2$  but not in  $\text{BH}_1$ , see figure 34c.
4. **Case D** — The surface passes behind the future horizon in  $\text{BH}_1$  ( $P_1 > 0$ ), it does not admit turning points, see figure 34d.
5. **Case E** — The surface passes behind the past horizon in  $\text{BH}_1$  ( $P_1 < 0$ ), it admits a turning point in  $\text{BH}_1$  but not in  $\text{BH}_2$ , see figure 34b.
6. **Case F** — The surface passes behind the past horizon in  $\text{BH}_1$  ( $P_1 < 0$ ), it admits turning points both in  $\text{BH}_1$  and in  $\text{BH}_2$ , see figure 34a.

We will study these cases one by one and derive in each case expressions for the boundary times and for the maximal volume and its time derivative.

**Case A.** — In this case the surface passes behind the future horizon in  $\text{BH}_1$  ( $P_1 > 0$ ), and admits turning points both in  $\text{BH}_1$  and in  $\text{BH}_2$ , see figure 34a. We will describe this surface using the coordinate  $u_L$  between the left boundary and  $r_{t,1}$  and using the coordinate  $v_R$  everywhere else. To obtain the expression for  $t_L$  we integrate along the surface in  $\text{BH}_1$ , first from the left boundary to  $r_{t,1}$

$$u_{t,1} - u_L = \int_{r_{t,1}}^{\infty} \tau_1[P_1, r] dr, \quad (\text{C.6})$$



**Figure 34.** Different possibilities for the shape of the maximal surface depending on whether it admits a turning point in  $BH_1$  and/or in  $BH_2$ .

then from  $r_{t,1}$  to  $r_s$

$$v_s - v_{t,1} = \int_{r_{t,1}}^{r_s} \tau_1[P_1, r] dr, \quad (C.7)$$

where we have used  $u_{t,1}$  and  $v_{t,1}$  to denote the values of  $u_L$  and  $v_R$  at the turning point of the surface in  $BH_1$ . Summing together and using  $v_{t,1} - u_{t,1} = 2r_1^*(r_{t,1})$  we obtain<sup>53</sup>

$$t_L + v_s - 2r_1^*(r_{t,1}) = \int_{r_{t,1}}^{r_s} \tau_1[P_1, r] dr + \int_{r_{t,1}}^{\infty} \tau_1[P_1, r] dr. \quad (C.8)$$

For the right boundary time, we integrate along the surface first from  $r_s$  to  $r_{t,2}$

$$v_{t,2} - v_s = - \int_{r_{t,2}}^{r_s} \tau_2[-P_2, r] dr, \quad (C.9)$$

then from  $r_{t,2}$  to the right boundary

$$v_R - v_{t,2} = \int_{r_{t,2}}^{\infty} \tau_2[P_2, r] dr, \quad (C.10)$$

and summing up we obtain

$$t_R - v_s = - \int_{r_{t,2}}^{r_s} \tau_2[-P_2, r] dr + \int_{r_{t,2}}^{\infty} \tau_2[P_2, r] dr. \quad (C.11)$$

The volume of the maximal surface is given by integrating along the four segments of the surface

$$\frac{1}{\Omega_{k,d-1}} \mathcal{V} = \int_{r_{t,1}}^{r_s} R_1[P_1, r] dr + \int_{r_{t,1}}^{\infty} R_1[P_1, r] dr + \int_{r_{t,2}}^{r_s} R_2[P_2, r] dr + \int_{r_{t,2}}^{\infty} R_2[P_2, r] dr. \quad (C.12)$$

We can use the identity (C.2a) as well as the relations (C.8) and (C.11) for the boundary times, to rewrite it as follows

$$\begin{aligned} \frac{1}{\Omega_{k,d-1}} \mathcal{V} &= \int_{r_{t,1}}^{r_s} \tilde{R}_1[P_1, r] dr + \int_{r_{t,1}}^{\infty} \tilde{R}_1[P_1, r] dr + \int_{r_{t,2}}^{r_s} \tilde{R}_2[-P_2, r] dr + \int_{r_{t,2}}^{\infty} \tilde{R}_2[P_2, r] dr \\ &\quad + P_1(t_L + v_s - 2r_1^*(r_{t,1})) + P_2(t_R - v_s). \end{aligned} \quad (C.13)$$

<sup>53</sup>To convince oneself that this relation still holds even though the calculations involve both left and right induced Eddington-Finkelstein coordinates, one may simply translate them to Kruskal-Szekeres coordinates which cover the whole spacetime. Further, we note that this is consistent with the synchronization of the clocks described at the beginning of section 3 since in this coordinate system, it is also easy to show that minimal surfaces going through the bifurcation surface are straight lines of constant (and opposite)  $t$  on the two sides.

We note in passing that the fact that  $\dot{v}$  is continuous across the interface implies using (C.2d)

$$\tilde{R}_1[P_1, r_s] = -\tilde{R}_2[-P_2, r_s]. \quad (\text{C.14})$$

For the time derivative of the volume (C.13), we obtain

$$\begin{aligned} \frac{1}{\Omega_{k,d-1}} \frac{d\mathcal{V}}{dt} &= \frac{dr_{t,1}}{dt} \left( -2\tilde{R}_1[P_1, r_{t,1}] - \frac{2P_1}{f_1(r_{t,1})} \right) + \frac{dr_{t,2}}{dt} \left( -\tilde{R}_2[-P_2, r_{t,2}] - \tilde{R}_2[P_2, r_{t,2}] \right) \\ &+ \frac{dP_1}{dt} \left( -\int_{r_{t,1}}^{r_s} \tau_1[P_1, r] dr - \int_{r_{t,1}}^{\infty} \tau_1[P_1, r] dr + (t_L + v_s - 2r_1^*(r_{t,1})) \right) \\ &+ \frac{dP_2}{dt} \left( \int_{r_{t,2}}^{r_s} \tau_2[-P_2, r] dr - \int_{r_{t,2}}^{\infty} \tau_2[P_2, r] dr + (t_R - v_s) \right) \\ &+ \frac{dr_s}{dt} \left( \tilde{R}_1[P_1, r_s] + \tilde{R}_2[-P_2, r_s] \right) + P_1 \frac{dt_L}{dt} + P_2 \frac{dt_R}{dt} \end{aligned} \quad (\text{C.15})$$

where we have defined  $t$  to be a time coordinate specifying the two boundary times via  $t_R(t)$ ,  $t_L(t)$  and used the identity (C.2b) to evaluate some of the derivatives with respect to the momenta. The various contributions inside the brackets can be shown to vanish using the identity (C.2c), the expressions for the left and right boundary times (C.8)–(C.11) and the continuity relation (C.14). So we are left with

$$\frac{1}{\Omega_{k,d-1}} \frac{d\mathcal{V}}{dt} = P_1 \frac{dt_L}{dt} + P_2 \frac{dt_R}{dt}. \quad (\text{C.16})$$

**Case B.** — In this case the surface passes behind the future horizon in  $\text{BH}_1$  ( $P_1 > 0$ ), and admits a turning point in  $\text{BH}_1$  but not in  $\text{BH}_2$ , see figure 34b. We will describe this surface using the coordinate  $u_L$  between the left boundary and  $r_{t,1}$  and using the coordinate  $v_R$  in all the rest of the surface. The expression for  $t_L$  is obtained just like for case A and it reads

$$t_L + v_s - 2r_1^*(r_{t,1}) = \int_{r_{t,1}}^{r_s} \tau_1[P_1, r] dr + \int_{r_{t,1}}^{\infty} \tau_1[P_1, r] dr. \quad (\text{C.17})$$

For the right boundary time, we integrate along the surface from  $r_s$  to the right boundary and obtain

$$t_R - v_s = \int_{r_s}^{\infty} \tau_2[P_2, r] dr. \quad (\text{C.18})$$

The volume of the maximal surface is given by integrating along the three segments of the surface

$$\frac{1}{\Omega_{k,d-1}} \mathcal{V} = \int_{r_{t,1}}^{r_s} R_1[P_1, r] dr + \int_{r_{t,1}}^{\infty} R_1[P_1, r] dr + \int_{r_s}^{\infty} R_2[P_2, r] dr. \quad (\text{C.19})$$

We can use the identity (C.2a) as well as the relations (C.17) and (C.18) for the boundary times, to rewrite it as follows

$$\begin{aligned} \frac{1}{\Omega_{k,d-1}} \mathcal{V} &= \int_{r_{t,1}}^{r_s} \tilde{R}_1[P_1, r] dr + \int_{r_{t,1}}^{\infty} \tilde{R}_1[P_1, r] dr + \int_{r_s}^{\infty} \tilde{R}_2[P_2, r] dr \\ &+ P_1(t_L + v_s - 2r_1^*(r_{t,1})) + P_2(t_R - v_s). \end{aligned} \quad (\text{C.20})$$

We note in passing that the fact that  $\dot{v}$  is continuous across the interface implies using eq. (C.2d)

$$\tilde{R}_1[P_1, r_s] = \tilde{R}_2[P_2, r_s]. \quad (\text{C.21})$$

For the time derivative of the volume (C.20), we obtain

$$\begin{aligned} \frac{1}{\Omega_{k,d-1}} \frac{d\mathcal{V}}{dt} &= \frac{dr_{t,1}}{dt} \left( -2\tilde{R}_1[P_1, r_{t,1}] - \frac{2P_1}{f_1(r_{t,1})} \right) \\ &+ \frac{dP_1}{dt} \left( - \int_{r_{t,1}}^{r_s} \tau_1[P_1, r] dr - \int_{r_{t,1}}^{\infty} \tau_1[P_1, r] dr + (t_L + v_s - 2r_1^*(r_{t,1})) \right) \\ &+ \frac{dP_2}{dt} \left( - \int_{r_s}^{\infty} \tau_2[P_2, r] dr + (t_R - v_s) \right) \\ &+ \frac{dr_s}{dt} \left( \tilde{R}_1[P_1, r_s] - \tilde{R}_2[P_2, r_s] \right) + P_1 \frac{dt_L}{dt} + P_2 \frac{dt_R}{dt} \end{aligned} \quad (\text{C.22})$$

where we have used the identity (C.2b) to evaluate some of the derivatives with respect to the momenta. All the contributions inside the brackets can be shown to vanish using the identity (C.2c), the expressions for the left and right boundary times in eqs. (C.17) and (C.18) and the continuity relation (C.21). So we are left with

$$\frac{1}{\Omega_{k,d-1}} \frac{d\mathcal{V}}{dt} = P_1 \frac{dt_L}{dt} + P_2 \frac{dt_R}{dt}. \quad (\text{C.23})$$

**Case C.** — In this case the surface passes behind the future horizon in  $\text{BH}_1$  ( $P_1 > 0$ ), and admits a turning point in  $\text{BH}_2$  but not in  $\text{BH}_1$ , see figure 34c. We will describe this surface using the coordinate  $u_L$  between the left boundary and  $r_s$  and using the coordinate  $v_R$  in all the rest of the surface. The expression for  $t_R$  is obtained just like for case A and it reads

$$t_R - v_s = - \int_{r_{t,2}}^{r_s} \tau_2[-P_2, r] dr + \int_{r_{t,2}}^{\infty} \tau_2[P_2, r] dr. \quad (\text{C.24})$$

Now  $t_L$  is obtained by integrating  $u_L$  from the left boundary to  $r_s$

$$u_s - u_L = t_L + v_s - 2r_1^*(r_s) = \int_{r_s}^{\infty} \tau_1[P_1, r] dr. \quad (\text{C.25})$$

The volume of the maximal surface is given by integrating along the three segments of the surface

$$\frac{1}{\Omega_{k,d-1}} \mathcal{V} = \int_{r_s}^{\infty} R_1[P_1, r] dr + \int_{r_{t,2}}^{r_s} R_2[P_2, r] dr + \int_{r_{t,2}}^{\infty} R_2[P_2, r] dr. \quad (\text{C.26})$$

We can use the identity (C.2a) as well as the relations (C.24)–(C.25) for the boundary times, to rewrite it as follows

$$\begin{aligned} \frac{1}{\Omega_{k,d-1}} \mathcal{V} &= \int_{r_s}^{\infty} \tilde{R}_1[P_1, r] dr + \int_{r_{t,2}}^{r_s} \tilde{R}_2[-P_2, r] dr + \int_{r_{t,2}}^{\infty} \tilde{R}_2[P_2, r] dr \\ &+ P_1(t_L + v_s - 2r_1^*(r_s)) + P_2(t_R - v_s). \end{aligned} \quad (\text{C.27})$$

We note in passing that the fact that  $\dot{v}$  is continuous across the interface implies using eq. (C.2d)

$$\tilde{R}_1[-P_1, r_s] = \tilde{R}_1[P_1, r_s] + \frac{2P_1}{f_1(r_s)} = \tilde{R}_2[-P_2, r_s]. \quad (\text{C.28})$$

For the time derivative of the volume (C.27), we obtain

$$\begin{aligned} \frac{1}{\Omega_{k,d-1}} \frac{d\mathcal{V}}{dt} &= \frac{dr_{t,2}}{dt} \left( -\tilde{R}_2[-P_2, r_{t,2}] - \tilde{R}_2[P_2, r_{t,2}] \right) \\ &+ \frac{dP_1}{dt} \left( -\int_{r_s}^{\infty} \tau_1[P_1, r] dr + (t_L + v_s - 2r_1^*(r_{t,1})) \right) \\ &+ \frac{dP_2}{dt} \left( \int_{r_{t,2}}^{r_s} \tau_2[-P_2, r] dr - \int_{r_{t,2}}^{\infty} \tau_2[P_2, r] dr + (t_R - v_s) \right) \\ &+ \frac{dr_s}{dt} \left( -\tilde{R}_1[P_1, r_s] + \tilde{R}_2[-P_2, r_s] - \frac{2P_1}{f_1(r_s)} \right) + P_1 \frac{dt_L}{dt} + P_2 \frac{dt_R}{dt} \end{aligned} \quad (\text{C.29})$$

where we have used again the identity (C.2b) to evaluate some of the derivatives with respect to the momenta. All the contributions inside the brackets can be shown to vanish using the identity (C.2c), the expressions for the right and left boundary times (C.24)–(C.25) and the continuity relation (C.28). So we are left with

$$\frac{1}{\Omega_{k,d-1}} \frac{d\mathcal{V}}{dt} = P_1 \frac{dt_L}{dt} + P_2 \frac{dt_R}{dt}. \quad (\text{C.30})$$

**Case D.** — In this case the surface passes behind the future horizon in  $\text{BH}_1$  ( $P_1 > 0$ ), and admits no turning points, see figure 34d. In fact, our numerical analysis revealed that this case is physically irrelevant since eq. (4.10) implies that  $\dot{r}_2 < \dot{r}_1$  for planar and spherical black holes for  $d \geq 3$ , as well as for BTZ black holes, and this is in contradiction with  $\dot{r}_1 < 0$  and  $\dot{r}_2 > 0$  implied by this scenario. This case may become relevant when considering hyperbolic black holes with negative mass, or alternatively in the (unphysical) case of negative energy shocks. We include it here for completeness.

We will describe the surface using the coordinate  $u_L$  between the left boundary and  $r_s$  and using the coordinate  $v_R$  between  $r_s$  and the left boundary. The expression for  $t_R$  is obtained just like for case B and it reads

$$t_R - v_s = \int_{r_s}^{\infty} \tau_2[P_2, r] dr. \quad (\text{C.31})$$

$t_L$  is obtained similarly to case C and reads

$$t_L + v_s - 2r_1^*(r_s) = \int_{r_s}^{\infty} \tau_1[P_1, r] dr. \quad (\text{C.32})$$

The volume of the maximal surface is given by integrating along the two segments of the surface

$$\frac{1}{\Omega_{k,d-1}} \mathcal{V} = \int_{r_s}^{\infty} R_1[P_1, r] dr + \int_{r_s}^{\infty} R_2[P_2, r] dr. \quad (\text{C.33})$$

We can use the identity (C.2a) as well as the relations (C.31) and (C.32) for the boundary times, to rewrite it as follows

$$\frac{1}{\Omega_{k,d-1}} \mathcal{V} = \int_{r_s}^{\infty} \tilde{R}_1[P_1, r] dr + \int_{r_s}^{\infty} \tilde{R}_2[P_2, r] dr + P_1(t_L + v_s - 2r_1^*(r_s)) + P_2(t_R - v_s). \quad (\text{C.34})$$

We note in passing that the fact that  $\dot{v}$  is continuous across the interface implies using (C.2d)

$$-\tilde{R}_1[-P_1, r_s] = -\tilde{R}_1[P_1, r_s] - \frac{2P_1}{f_1(r_s)} = \tilde{R}_2[P_2, r_s]. \quad (\text{C.35})$$

For the time derivative of the volume (C.34) we obtain

$$\begin{aligned} \frac{1}{\Omega_{k,d-1}} \frac{d\mathcal{V}}{dt} &= \frac{dP_1}{dt} \left( - \int_{r_s}^{\infty} \tau_1[P_1, r] dr + (t_L + v_s - 2r_1^*(r_{t,1})) \right) \\ &+ \frac{dP_2}{dt} \left( - \int_{r_s}^{\infty} \tau_2[P_2, r] dr + (t_R - v_s) \right) \\ &+ \frac{dr_s}{dt} \left( -\tilde{R}_1[P_1, r_s] - \tilde{R}_2[P_2, r_s] - \frac{2P_1}{f_1(r_s)} \right) + P_1 \frac{dt_L}{dt} + P_2 \frac{dt_R}{dt} \end{aligned} \quad (\text{C.36})$$

where we have used the identity (C.2b) to evaluate some of the derivatives with respect to the momenta. All the contributions inside the brackets can be shown to vanish using the expressions for the right and left boundary times (C.31)–(C.32) and the continuity relation (C.35). So we are left with

$$\frac{1}{\Omega_{k,d-1}} \frac{d\mathcal{V}}{dt} = P_1 \frac{dt_L}{dt} + P_2 \frac{dt_R}{dt}. \quad (\text{C.37})$$

**Case E.** — In this case the surface passes behind the past horizon in  $\text{BH}_1$  ( $P_1 < 0$ ), and admits a turning point in  $\text{BH}_1$  but not in  $\text{BH}_2$ , see figure 34b. We will describe this surface using the coordinate  $v_L$  between the left boundary and  $r_{t,1}$ , the coordinate  $u_R$  between  $r_{t,1}$  and  $r_s$  and the coordinate  $v_R$  between  $r_s$  and the right boundary. The expression for  $t_R$  is obtained summing the results of the  $v_L$  and  $u_R$  integrations in the first two of these segments

$$\begin{aligned} v_{t,1} - v_L &= v_{t,1} + t_L = - \int_{r_{t,1}}^{\infty} \tau_1[-P_1, r] dr \\ u_s - u_{t,1} &= v_s - 2r_1^*(r_s) - u_{t,1} = - \int_{r_{t,1}}^{r_s} \tau_1[-P_1, r] dr \end{aligned} \quad (\text{C.38})$$

which yields

$$t_L + v_s - 2r_1^*(r_s) + 2r_1^*(r_{t,1}) = - \int_{r_{t,1}}^{r_s} \tau_1[-P_1, r] dr - \int_{r_{t,1}}^{\infty} \tau_1[-P_1, r] dr. \quad (\text{C.39})$$

The right boundary time is obtained just like in case B and reads

$$t_R - v_s = \int_{r_s}^{\infty} \tau_2[P_2, r] dr. \quad (\text{C.40})$$

The volume of the maximal surface is given by integrating along the three segments of the surface

$$\frac{1}{\Omega_{k,d-1}} \mathcal{V} = \int_{r_{t,1}}^{r_s} R_1[P_1, r] dr + \int_{r_{t,1}}^{\infty} R_1[P_1, r] dr + \int_{r_s}^{\infty} R_2[P_2, r] dr. \quad (\text{C.41})$$

We can use the identity (C.2a) as well as the relations (C.39) and (C.40) for the boundary times, to rewrite it as follows

$$\begin{aligned} \frac{1}{\Omega_{k,d-1}} \mathcal{V} &= \int_{r_{t,1}}^{r_s} \tilde{R}_1[-P_1, r] dr + \int_{r_{t,1}}^{\infty} \tilde{R}_1[-P_1, r] dr + \int_{r_s}^{\infty} \tilde{R}_2[P_2, r] dr \\ &+ P_1(t_L + v_s - 2r_1^*(r_s) + 2r_1^*(r_{t,1})) + P_2(t_R - v_s). \end{aligned} \quad (\text{C.42})$$

The fact that  $\dot{v}$  is continuous across the interface implies using (C.2d)

$$\tilde{R}_1[P_1, r_s] = \tilde{R}_1[-P_1, r_s] - \frac{2P_1}{f_1(r_s)} = \tilde{R}_2[P_2, r_s]. \quad (\text{C.43})$$

For the time derivative of the volume (C.42) we obtain

$$\begin{aligned} \frac{1}{\Omega_{k,d-1}} \frac{d\mathcal{V}}{dt} &= \frac{dr_{t,1}}{dt} \left( -2\tilde{R}_1[-P_1, r_{t,1}] + \frac{2P_1}{f_1(r_{t,1})} \right) \\ &\times \frac{dP_1}{dt} \left( \int_{r_{t,1}}^{r_s} \tau_1[-P_1, r] dr + \int_{r_{t,1}}^{\infty} \tau_1[-P_1, r] dr + (t_L + v_s - 2r_1^*(r_s) + 2r_1^*(r_{t,1})) \right) \\ &+ \frac{dP_2}{dt} \left( - \int_{r_s}^{\infty} \tau_2[P_2, r] dr + (t_R - v_s) \right) \\ &+ \frac{dr_s}{dt} \left( \tilde{R}_1[-P_1, r_s] - \tilde{R}_2[P_2, r_s] - \frac{2P_1}{f_1(r_s)} \right) + P_1 \frac{dt_L}{dt} + P_2 \frac{dt_R}{dt} \end{aligned} \quad (\text{C.44})$$

where we have used the identity (C.2b) to evaluate some of the derivatives with respect to the momenta. All the contributions inside the brackets can be shown to vanish using the identity (C.2c), the expressions for the left and right boundary times (C.39)–(C.40) and the continuity relation (C.43). We are finally left with

$$\frac{1}{\Omega_{k,d-1}} \frac{d\mathcal{V}}{dt} = P_1 \frac{dt_L}{dt} + P_2 \frac{dt_R}{dt}. \quad (\text{C.45})$$

**Case F.** — In this case the surface passes behind the past horizon in  $\text{BH}_1$  ( $P_1 < 0$ ), and admits turning points in  $\text{BH}_1$  and  $\text{BH}_2$ , see figure 34a. We will describe this surface using the coordinate  $v_L$  between the left boundary and  $r_{t,1}$ ,  $u_R$  between  $r_{t,1}$  and  $r_s$  and  $v_R$  for the two segments between  $r_s$  and the right boundary. The expression for  $t_R$  is the same as in case E

$$t_L + v_s - 2r_1^*(r_s) + 2r_1^*(r_{t,1}) = - \int_{r_{t,1}}^{r_s} \tau_1[-P_1, r] dr - \int_{r_{t,1}}^{\infty} \tau_1[-P_1, r] dr. \quad (\text{C.46})$$

The right boundary time is obtained just like in case A and reads

$$t_R - v_s = - \int_{r_{t,2}}^{r_s} \tau_2[-P_2, r] dr + \int_{r_{t,2}}^{\infty} \tau_2[P_2, r] dr. \quad (\text{C.47})$$



The volume of the maximal surface is given by integrating along the four segments of the surface

$$\frac{1}{\Omega_{k,d-1}} \mathcal{V} = \int_{r_{t,1}}^{r_s} R_1[P_1, r] dr + \int_{r_{t,1}}^{\infty} R_1[P_1, r] dr + \int_{r_{t,2}}^{r_s} R_2[P_2, r] dr + \int_{r_{t,2}}^{\infty} R_2[P_2, r] dr. \quad (\text{C.48})$$

We can use the identity (C.2a) as well as the relations (C.46) and (C.47) for the boundary times, to rewrite it as follows

$$\begin{aligned} \frac{1}{\Omega_{k,d-1}} \mathcal{V} &= \int_{r_{t,1}}^{r_s} \tilde{R}_1[-P_1, r] dr + \int_{r_{t,1}}^{\infty} \tilde{R}_1[-P_1, r] dr + \int_{r_{t,2}}^{r_s} \tilde{R}_2[-P_2, r] dr + \int_{r_{t,2}}^{\infty} \tilde{R}_2[P_2, r] dr \\ &\quad + P_1(t_L + v_s - 2r_1^*(r_s) + 2r_1^*(r_{t,1})) + P_2(t_R - v_s). \end{aligned} \quad (\text{C.49})$$

The fact that  $\dot{v}$  is continuous across the interface implies using (C.2d)

$$\tilde{R}_1[P_1, r_s] = \tilde{R}_1[-P_1, r_s] - \frac{2P_1}{f_1(r_s)} = -\tilde{R}_2[-P_2, r_s]. \quad (\text{C.50})$$

For the time derivative of the volume (C.27) we obtain

$$\begin{aligned} \frac{1}{\Omega_{k,d-1}} \frac{d\mathcal{V}}{dt} &= \frac{dr_{t,1}}{dt} \left( -2\tilde{R}_1[-P_1, r_{t,1}] + \frac{2P_1}{f(r_{t,1})} \right) \\ &\quad \times \frac{dr_{t,2}}{dt} \left( -\tilde{R}_2[-P_2, r_{t,2}] - \tilde{R}_2[P_2, r_{t,2}] \right) \\ &\quad \times \frac{dP_1}{dt} \left( \int_{r_{t,1}}^{r_s} \tau_1[-P_1, r] dr + \int_{r_{t,1}}^{\infty} \tau_1[-P_1, r] dr + (t_L + v_s - 2r_1^*(r_s) + 2r_1^*(r_{t,1})) \right) \\ &\quad + \frac{dP_2}{dt} \left( \int_{r_{t,2}}^{r_s} \tau_2[-P_2, r] dr - \int_{r_{t,2}}^{\infty} \tau_2[P_2, r] dr + (t_R - v_s) \right) \\ &\quad + \frac{dr_s}{dt} \left( \tilde{R}_1[-P_1, r_s] + \tilde{R}_2[-P_2, r_s] - \frac{2P_1}{f_1(r_s)} \right) + P_1 \frac{dt_L}{dt} + P_2 \frac{dt_R}{dt} \end{aligned} \quad (\text{C.51})$$

where we have used the identity (C.2b) to evaluate some of the derivatives with respect to the momenta. All the contributions inside the brackets can be shown to vanish using the identity (C.2c), the expressions for the left and right boundary times (C.46)–(C.47) and the continuity relation (C.50). So we are again left with

$$\frac{1}{\Omega_{k,d-1}} \frac{d\mathcal{V}}{dt} = P_1 \frac{dt_L}{dt} + P_2 \frac{dt_R}{dt}. \quad (\text{C.52})$$

**Simpler derivation of volume variation formula.** The simple form of the volume variation formula,

$$\frac{1}{\Omega_{k,d-1}} \frac{d\mathcal{V}}{dt} = P_1 \frac{dt_L}{dt} + P_2 \frac{dt_R}{dt}, \quad (\text{C.53})$$

which appears for all of the above cases, is of course not a coincidence. It is a simple consequence of the fact that  $\mathcal{V}$  extremizes an ‘action’, i.e., the volume functional (4.2). A standard result in classical mechanics states that the change in the on-shell action when moving between two nearby solutions of the equations of motion is given by boundary

terms, as we now review. Let  $x^\mu(s, \lambda)$  be a family of solutions to the equations of motion associated with the action

$$S(s) = \int_{\lambda_L}^{\lambda_R} d\lambda \mathcal{L}(x^\mu, \dot{x}^\mu) \tag{C.54}$$

labeled by a parameter  $s$  and where we have denoted  $\dot{x}^\mu = \partial_\lambda x^\mu$ . We also assume that the boundary values of the ‘time’ parameter  $\lambda$  are constant for all the solutions in the family. The variation of the on-shell action with respect to  $s$  is then given by

$$\frac{dS}{ds} = \partial_s x^\mu k_\mu|_{\lambda_L} - \partial_s x^\mu k_\mu|_{\lambda_R} \tag{C.55}$$

where the canonical momentum  $k_\mu$  is given by  $k_\mu \equiv \frac{\partial \mathcal{L}}{\partial \dot{x}^\mu}$ .

In our case the onshell action is given by eq. (4.2) and the canonical momentum  $k_\mu$  reads<sup>54</sup>

$$\frac{1}{\Omega_{k,d-1}} k_\mu dx^\mu = P dv + \frac{r^{d-1} \dot{v}}{\sqrt{-f \dot{v}^2 + 2 \dot{v} \dot{r}}} dr = P du - \frac{r^{d-1} \dot{u}}{\sqrt{-f \dot{u}^2 - 2 \dot{u} \dot{r}}} dr \tag{C.56}$$

where  $P$  is the conserved momentum from eq. (4.4). The family of solutions we are interested in is labeled by a parameter  $s = t$  which is simply our boundary time parameter. The variation of the solutions with respect to  $t$  at the end points of the trajectory is given by

$$\left. \frac{\partial v_R}{\partial t} \right|_{\lambda_R} = \left. \frac{\partial u_R}{\partial t} \right|_{\lambda_R} = \frac{\partial t_R}{\partial t}, \quad \left. \frac{\partial u_L}{\partial t} \right|_{\lambda_L} = \left. \frac{\partial v_L}{\partial t} \right|_{\lambda_L} = -\frac{\partial t_L}{\partial t}, \quad \left. \frac{\partial r}{\partial t} \right|_{\lambda_R} = \left. \frac{\partial r}{\partial t} \right|_{\lambda_L} = 0. \tag{C.57}$$

Substituting into eq. (C.55) we obtain

$$\frac{1}{\Omega_{k,d-1}} \frac{d\mathcal{V}}{dt} = P_1 \frac{dt_L}{dt} + P_2 \frac{dt_R}{dt}. \tag{C.58}$$

Before we conclude this argument let us address two subtle points. First, since the coordinates we are using do not cover the full spacetime, we will need to split our trajectory into different segments using the different coordinate patches. However the new boundary contributions due to this splitting will cancel among themselves, due to the opposite signs of the two contributions in eq. (C.55) and the fact that the canonical momentum varies smoothly along the trajectory. This is true except for at the exact point in which the shell is being crossed which brings us to the second subtle point. At the point in which the shell is being crossed, the canonical momentum (C.56) is not smooth, however the different solutions in the family cross the shell at location who vary only by their values of  $r$ . As a consequence the boundary contributions are proportional there to  $k_r \delta r / \delta t$ . Looking at the equations of motion in the close proximity of the shell  $\partial_\lambda k_r = \partial \mathcal{L} / \partial r$  we conclude by integrating in a small interval around the shell that  $k_r$  is continuous across the shell (in our favourite parametrization (4.3), this takes the form of the continuity of  $\dot{v}$  when crossing the shell). This is enough to assure the cancelation of the extra boundary contributions at the location of the shell.

---

<sup>54</sup>Note that in the following, we are not using the gauge choice in eq. (4.3). Instead, we will choose parametrizations which ensure that the boundary values of the worldvolume parameter  $\lambda$ , i.e.,  $\lambda_L$  and  $\lambda_R$ , are constant for all the solutions in the family  $x^\mu(s, \lambda)$ .

**Crossing between the past and future regimes.** The crossing between the two regimes is associated with the case  $P_1 = 0$ . In this case the minimal surface in  $BH_1$  follows a constant time slice through the bifurcation surface. This constant time slice intersects the shock at

$$v_s + t_L = r_1^*(r_s). \tag{C.59}$$

Collecting all the above results completes the derivation of eqs. (4.14)–(4.18).

**Open Access.** This article is distributed under the terms of the Creative Commons Attribution License ([CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

## References

- [1] S. Chapman, H. Marrochio and R.C. Myers, *Holographic complexity in Vaidya spacetimes. Part I*, *JHEP* **06** (2018) 046 [[arXiv:1804.07410](https://arxiv.org/abs/1804.07410)] [[INSPIRE](#)].
- [2] O. Aharony, S.S. Gubser, J.M. Maldacena, H. Ooguri and Y. Oz, *Large N field theories, string theory and gravity*, *Phys. Rept.* **323** (2000) 183 [[hep-th/9905111](https://arxiv.org/abs/hep-th/9905111)] [[INSPIRE](#)].
- [3] M. Ammon and J. Erdmenger, *Gauge/gravity duality: foundations and applications*, Cambridge University Press, Cambridge, U.K., (2015) [[INSPIRE](#)].
- [4] S. Ryu and T. Takayanagi, *Holographic derivation of entanglement entropy from AdS/CFT*, *Phys. Rev. Lett.* **96** (2006) 181602 [[hep-th/0603001](https://arxiv.org/abs/hep-th/0603001)] [[INSPIRE](#)].
- [5] S. Ryu and T. Takayanagi, *Aspects of holographic entanglement entropy*, *JHEP* **08** (2006) 045 [[hep-th/0605073](https://arxiv.org/abs/hep-th/0605073)] [[INSPIRE](#)].
- [6] V.E. Hubeny, M. Rangamani and T. Takayanagi, *A covariant holographic entanglement entropy proposal*, *JHEP* **07** (2007) 062 [[arXiv:0705.0016](https://arxiv.org/abs/hep-th/0705001)] [[INSPIRE](#)].
- [7] M. Rangamani and T. Takayanagi, *Holographic entanglement entropy*, *Lect. Notes Phys.* **931** (2017) pp.1 [[arXiv:1609.01287](https://arxiv.org/abs/1609.01287)] [[INSPIRE](#)].
- [8] M. Van Raamsdonk, *Comments on quantum gravity and entanglement*, [arXiv:0907.2939](https://arxiv.org/abs/0907.2939) [[INSPIRE](#)].
- [9] M. Van Raamsdonk, *Building up spacetime with quantum entanglement*, *Gen. Rel. Grav.* **42** (2010) 2323 [*Int. J. Mod. Phys. D* **19** (2010) 2429] [[arXiv:1005.3035](https://arxiv.org/abs/1005.3035)] [[INSPIRE](#)].
- [10] J. Maldacena and L. Susskind, *Cool horizons for entangled black holes*, *Fortsch. Phys.* **61** (2013) 781 [[arXiv:1306.0533](https://arxiv.org/abs/1306.0533)] [[INSPIRE](#)].
- [11] L. Susskind, *Entanglement is not enough*, *Fortsch. Phys.* **64** (2016) 49 [[arXiv:1411.0690](https://arxiv.org/abs/1411.0690)] [[INSPIRE](#)].
- [12] J. Watrous, *Quantum computational complexity*, in *Encyclopedia of complexity and systems science*, R.A. Meyers ed., Springer, New York, U.S.A., (2009), pg. 7174 [[arXiv:0804.3401](https://arxiv.org/abs/0804.3401)].
- [13] S. Aaronson, *The complexity of quantum states and transformations: from quantum money to black holes*, [arXiv:1607.05256](https://arxiv.org/abs/1607.05256) [[INSPIRE](#)].
- [14] L. Susskind, *Computational complexity and black hole horizons*, *Fortsch. Phys.* **64** (2016) 24 [*Addendum ibid.* **64** (2016) 44] [[arXiv:1402.5674](https://arxiv.org/abs/1402.5674)] [[arXiv:1403.5695](https://arxiv.org/abs/1403.5695)] [[INSPIRE](#)].

- [15] D. Stanford and L. Susskind, *Complexity and shock wave geometries*, *Phys. Rev. D* **90** (2014) 126007 [[arXiv:1406.2678](#)] [[INSPIRE](#)].
- [16] 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](#)].
- [17] 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](#)].
- [18] D.A. Roberts, D. Stanford and L. Susskind, *Localized shocks*, *JHEP* **03** (2015) 051 [[arXiv:1409.8180](#)] [[INSPIRE](#)].
- [19] L. Susskind and Y. Zhao, *Switchbacks and the bridge to nowhere*, [arXiv:1408.2823](#) [[INSPIRE](#)].
- [20] R.-G. Cai, S.-M. Ruan, S.-J. Wang, R.-Q. Yang and R.-H. Peng, *Action growth for AdS black holes*, *JHEP* **09** (2016) 161 [[arXiv:1606.08307](#)] [[INSPIRE](#)].
- [21] 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](#)].
- [22] S. Chapman, H. Marrochio and R.C. Myers, *Complexity of formation in holography*, *JHEP* **01** (2017) 062 [[arXiv:1610.08063](#)] [[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](#)].
- [24] M. Alishahiha, *Holographic complexity*, *Phys. Rev. D* **92** (2015) 126009 [[arXiv:1509.06614](#)] [[INSPIRE](#)].
- [25] O. Ben-Ami and D. Carmi, *On volumes of subregions in holography and complexity*, *JHEP* **11** (2016) 129 [[arXiv:1609.02514](#)] [[INSPIRE](#)].
- [26] D. Carmi, R.C. Myers and P. Rath, *Comments on holographic complexity*, *JHEP* **03** (2017) 118 [[arXiv:1612.00433](#)] [[INSPIRE](#)].
- [27] A.R. Brown and L. Susskind, *Second law of quantum complexity*, *Phys. Rev. D* **97** (2018) 086015 [[arXiv:1701.01107](#)] [[INSPIRE](#)].
- [28] P. Caputa, N. Kundu, M. Miyaji, T. Takayanagi and K. Watanabe, *Anti-de Sitter space from optimization of path integrals in conformal field theories*, *Phys. Rev. Lett.* **119** (2017) 071602 [[arXiv:1703.00456](#)] [[INSPIRE](#)].
- [29] P. Caputa, N. Kundu, M. Miyaji, T. Takayanagi and K. Watanabe, *Liouville action as path-integral complexity: from continuous tensor networks to AdS/CFT*, *JHEP* **11** (2017) 097 [[arXiv:1706.07056](#)] [[INSPIRE](#)].
- [30] B. Czech, *Einstein equations from varying complexity*, *Phys. Rev. Lett.* **120** (2018) 031601 [[arXiv:1706.00965](#)] [[INSPIRE](#)].
- [31] A. Reynolds and S.F. Ross, *Complexity in de Sitter space*, *Class. Quant. Grav.* **34** (2017) 175013 [[arXiv:1706.03788](#)] [[INSPIRE](#)].
- [32] K. Hashimoto, N. Iizuka and S. Sugishita, *Time evolution of complexity in Abelian gauge theories*, *Phys. Rev. D* **96** (2017) 126001 [[arXiv:1707.03840](#)] [[INSPIRE](#)].
- [33] R. Jefferson and R.C. Myers, *Circuit complexity in quantum field theory*, *JHEP* **10** (2017) 107 [[arXiv:1707.08570](#)] [[INSPIRE](#)].

- [34] 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 [[arXiv:1707.08582](#)] [[INSPIRE](#)].
- [35] L. Hackl and R.C. Myers, *Circuit complexity for free fermions*, [arXiv:1803.10638](#) [[INSPIRE](#)].
- [36] J. Couch, S. Eccles, W. Fischler and M.-L. Xiao, *Holographic complexity and noncommutative gauge theory*, *JHEP* **03** (2018) 108 [[arXiv:1710.07833](#)] [[INSPIRE](#)].
- [37] M. Moosa, *Evolution of complexity following a global quench*, *JHEP* **03** (2018) 031 [[arXiv:1711.02668](#)] [[INSPIRE](#)].
- [38] B. Swingle and Y. Wang, *Holographic complexity of Einstein-Maxwell-dilaton gravity*, [arXiv:1712.09826](#) [[INSPIRE](#)].
- [39] M. Alishahiha, A. Faraji Astaneh, M.R. Mohammadi Mozaffar and A. Mollabashi, *Complexity growth with Lifshitz scaling and hyperscaling violation*, [arXiv:1802.06740](#) [[INSPIRE](#)].
- [40] B. Chen, W.-M. Li, R.-Q. Yang, C.-Y. Zhang and S.-J. Zhang, *Holographic subregion complexity under a thermal quench*, [arXiv:1803.06680](#) [[INSPIRE](#)].
- [41] Y. Zhao, *Uncomplexity and black hole geometry*, [arXiv:1711.03125](#) [[INSPIRE](#)].
- [42] Z. Fu, A. Maloney, D. Marolf, H. Maxfield and Z. Wang, *Holographic complexity is nonlocal*, *JHEP* **02** (2018) 072 [[arXiv:1801.01137](#)] [[INSPIRE](#)].
- [43] C.A. Agón, M. Headrick and B. Swingle, *Subsystem complexity and holography*, [arXiv:1804.01561](#) [[INSPIRE](#)].
- [44] P.C. Vaidya, *The external field of a radiating star in general relativity*, *Curr. Sci.* **12** (1943) 183.
- [45] P.C. Vaidya, *The gravitational field of a radiating star*, *Proc. Indian Acad. Sci. (Math. Sci.)* **33** (1951) 264.
- [46] A. Wang and Y. Wu, *Generalized Vaidya solutions*, *Gen. Rel. Grav.* **31** (1999) 107 [[gr-qc/9803038](#)] [[INSPIRE](#)].
- [47] S.H. Shenker and D. Stanford, *Multiple shocks*, *JHEP* **12** (2014) 046 [[arXiv:1312.3296](#)] [[INSPIRE](#)].
- [48] J.M. Maldacena, *Eternal black holes in anti-de Sitter*, *JHEP* **04** (2003) 021 [[hep-th/0106112](#)] [[INSPIRE](#)].
- [49] S.H. Shenker and D. Stanford, *Black holes and the butterfly effect*, *JHEP* **03** (2014) 067 [[arXiv:1306.0622](#)] [[INSPIRE](#)].
- [50] T. Anous, T. Hartman, A. Rovai and J. Sonner, *Black hole collapse in the  $1/c$  expansion*, *JHEP* **07** (2016) 123 [[arXiv:1603.04856](#)] [[INSPIRE](#)].
- [51] G. Hayward, *Gravitational action for space-times with nonsmooth boundaries*, *Phys. Rev.* **D 47** (1993) 3275 [[INSPIRE](#)].
- [52] D. Brill and G. Hayward, *Is the gravitational action additive?*, *Phys. Rev.* **D 50** (1994) 4914 [[gr-qc/9403018](#)] [[INSPIRE](#)].
- [53] O. Coussaert and M. Henneaux, *Supersymmetry of the  $(2+1)$  black holes*, *Phys. Rev. Lett.* **72** (1994) 183 [[hep-th/9310194](#)] [[INSPIRE](#)].

- [54] L. Fidkowski, V. Hubeny, M. Kleban and S. Shenker, *The black hole singularity in AdS/CFT*, *JHEP* **02** (2004) 014 [[hep-th/0306170](#)] [[INSPIRE](#)].
- [55] Y. Sekino and L. Susskind, *Fast scramblers*, *JHEP* **10** (2008) 065 [[arXiv:0808.2096](#)] [[INSPIRE](#)].
- [56] Y. Zhao, *Complexity, boost symmetry and firewalls*, [arXiv:1702.03957](#) [[INSPIRE](#)].
- [57] R. Khan, C. Krishnan and S. Sharma, *Circuit complexity in fermionic field theory*, [arXiv:1801.07620](#) [[INSPIRE](#)].
- [58] J. Couch, W. Fischler and P.H. Nguyen, *Noether charge, black hole volume and complexity*, *JHEP* **03** (2017) 119 [[arXiv:1610.02038](#)] [[INSPIRE](#)].
- [59] Z.-Y. Fan and M. Guo, *On the Noether charge and the gravity duals of quantum complexity*, [arXiv:1805.03796](#) [[INSPIRE](#)].
- [60] D. Kubiznak, R.B. Mann and M. Teo, *Black hole chemistry: thermodynamics with Lambda*, *Class. Quant. Grav.* **34** (2017) 063001 [[arXiv:1608.06147](#)] [[INSPIRE](#)].
- [61] R.-G. Cai, S.-M. Ruan, S.-J. Wang, R.-Q. Yang and R.-H. Peng, *Action growth for AdS black holes*, *JHEP* **09** (2016) 161 [[arXiv:1606.08307](#)] [[INSPIRE](#)].
- [62] H. Huang, X.-H. Feng and H. Lü, *Holographic complexity and two identities of action growth*, *Phys. Lett. B* **769** (2017) 357 [[arXiv:1611.02321](#)] [[INSPIRE](#)].
- [63] P.A. Cano, R.A. Hennigar and H. Marrochio, *Complexity growth rate in Lovelock gravity*, [arXiv:1803.02795](#) [[INSPIRE](#)].
- [64] A.R. Brown, L. Susskind and Y. Zhao, *Quantum complexity and negative curvature*, *Phys. Rev. D* **95** (2017) 045010 [[arXiv:1608.02612](#)] [[INSPIRE](#)].
- [65] L. Susskind, *The typical-state paradox: diagnosing horizons with complexity*, *Fortsch. Phys.* **64** (2016) 84 [[arXiv:1507.02287](#)] [[INSPIRE](#)].
- [66] J. Maldacena, S.H. Shenker and D. Stanford, *A bound on chaos*, *JHEP* **08** (2016) 106 [[arXiv:1503.01409](#)] [[INSPIRE](#)].
- [67] M. Van Raamsdonk, *Evaporating firewalls*, *JHEP* **11** (2014) 038 [[arXiv:1307.1796](#)] [[INSPIRE](#)].
- [68] A. Buchel, L. Lehner and R.C. Myers, *Thermal quenches in  $N = 2^*$  plasmas*, *JHEP* **08** (2012) 049 [[arXiv:1206.6785](#)] [[INSPIRE](#)].
- [69] A. Buchel, L. Lehner, R.C. Myers and A. van Niekerk, *Quantum quenches of holographic plasmas*, *JHEP* **05** (2013) 067 [[arXiv:1302.2924](#)] [[INSPIRE](#)].
- [70] R.-Q. Yang, *Complexity for quantum field theory states and applications to thermofield double states*, *Phys. Rev. D* **97** (2018) 066004 [[arXiv:1709.00921](#)] [[INSPIRE](#)].
- [71] R.-Q. Yang, C. Niu, C.-Y. Zhang and K.-Y. Kim, *Comparison of holographic and field theoretic complexities for time dependent thermofield double states*, *JHEP* **02** (2018) 082 [[arXiv:1710.00600](#)] [[INSPIRE](#)].
- [72] S. Chapman et al., *Circuit complexity for thermofield double states*, to appear, (2018).
- [73] D.W.F. Alves and G. Camilo, *Evolution of complexity following a quantum quench in free field theory*, [arXiv:1804.00107](#) [[INSPIRE](#)].
- [74] H. Camargo, P. Caputa, D. Das, M.P. Heller and R. Jefferson, *Entanglement is not enough: complexity as a novel probe of quantum quenches*, in preparation.

- [75] P. Gao, D.L. Jafferis and A. Wall, *Traversable wormholes via a double trace deformation*, *JHEP* **12** (2017) 151 [[arXiv:1608.05687](#)] [[INSPIRE](#)].
- [76] J. Maldacena, D. Stanford and Z. Yang, *Diving into traversable wormholes*, *Fortsch. Phys.* **65** (2017) 1700034 [[arXiv:1704.05333](#)] [[INSPIRE](#)].
- [77] M. Miyaji, T. Numasawa, N. Shiba, T. Takayanagi and K. Watanabe, *Distance between quantum states and gauge-gravity duality*, *Phys. Rev. Lett.* **115** (2015) 261602 [[arXiv:1507.07555](#)] [[INSPIRE](#)].
- [78] A. Bhattacharyya, P. Caputa, S.R. Das, N. Kundu, M. Miyaji and T. Takayanagi, *Path-integral complexity for perturbed CFTs*, [arXiv:1804.01999](#) [[INSPIRE](#)].
- [79] A. Reynolds and S.F. Ross, *Divergences in holographic complexity*, *Class. Quant. Grav.* **34** (2017) 105004 [[arXiv:1612.05439](#)] [[INSPIRE](#)].
- [80] R. Emparan, C.V. Johnson and R.C. Myers, *Surface terms as counterterms in the AdS/CFT correspondence*, *Phys. Rev. D* **60** (1999) 104001 [[hep-th/9903238](#)] [[INSPIRE](#)].
- [81] S. de Haro, S.N. Solodukhin and K. Skenderis, *Holographic reconstruction of space-time and renormalization in the AdS/CFT correspondence*, *Commun. Math. Phys.* **217** (2001) 595 [[hep-th/0002230](#)] [[INSPIRE](#)].
- [82] K. Skenderis, *Lecture notes on holographic renormalization*, *Class. Quant. Grav.* **19** (2002) 5849 [[hep-th/0209067](#)] [[INSPIRE](#)].
- [83] A. Buchel, J. Escobedo, R.C. Myers, M.F. Paulos, A. Sinha and M. Smolkin, *Holographic GB gravity in arbitrary dimensions*, *JHEP* **03** (2010) 111 [[arXiv:0911.4257](#)] [[INSPIRE](#)].
- [84] P.A. Cano, *Lovelock action with nonsmooth boundaries*, *Phys. Rev. D* **97** (2018) 104048 [[arXiv:1803.00172](#)] [[INSPIRE](#)].