

Generalized entropy in higher curvature gravity and entropy of algebra of observables

Mohd Ali^{*} and Vardarajan Suneeta[†]

The Indian Institute of Science Education and Research (IISER), Pune, India—411008

 (Received 1 August 2023; accepted 13 September 2023; published 29 September 2023)

Recently, Chandrasekaran, Penington, and Witten (CPW) have shown that the generalized entropy of the Schwarzschild black hole at the bifurcation surface equals the entropy of an extended von Neumann algebra of quantum observables in the black hole exterior, in semiclassical Einstein gravity. They also derive a version of the generalized second law. We generalize these results to a static black hole in an arbitrary diffeomorphism invariant theory of gravity. Thus, a version of the generalized second law for an arbitrary diffeomorphism invariant theory of gravity follows.

DOI: [10.1103/PhysRevD.108.066017](https://doi.org/10.1103/PhysRevD.108.066017)

I. INTRODUCTION

Generalized entropy in Einstein gravity was introduced by Bekenstein in order that the second law of thermodynamics be valid near black holes [1,2]. He suggested the generalized second law (GSL) holds, namely that the generalized entropy increases under future evolution along the black hole horizon. The generalized entropy for a quantum black hole coupled to matter, in the semiclassical $G \rightarrow 0$ limit was defined to be

$$S_{\text{gen}} = \left\langle \frac{A}{4\hbar G} \right\rangle + S_{\text{QFT}}, \quad (1.1)$$

where A is the black hole horizon area and S_{QFT} is the entanglement entropy of the quantum fields in the black hole exterior. As has been pointed out, both the terms in (1.1) are individually ultraviolet (UV) divergent (the first term due to loop effects which renormalize G and the second term, entanglement entropy, which is UV divergent), but the sum is UV finite [3–10]. The GSL for Einstein gravity was proved by Wall [11]. In an arbitrary diffeomorphism invariant theory of gravity, we can analogously define the generalized entropy for a black hole with quantum fields to be the sum of its horizon entropy S in that theory and the entanglement entropy of quantum fields in the black hole exterior,

$$S_{\text{gen}} = S + S_{\text{QFT}}. \quad (1.2)$$

A candidate for the horizon entropy for a stationary black hole is Wald entropy [12,13]. The Wald entropy is ambiguous for a nonstationary black hole—these ambiguities were first discussed in [14]. A linearized GSL (ignoring gravitons) was proved for Lovelock gravity in [15].

In an interesting recent development, the generalized entropy in Einstein gravity appeared as the entropy of a von Neumann algebra of observables in the black hole exterior where the black hole is the Schwarzschild solution in Einstein gravity [16]. Leutheusser and Liu [17] (see also [18]) studied the holographic boundary operator algebra of the CFT dual to gravity in the asymptotically anti-de Sitter (AdS) black hole spacetime. They found an emergent type III₁ von Neumann algebra for single trace operators in the large N limit of the boundary CFT (see also [19]). Later, Witten [20] showed that by enlarging the set of operators to include the boundary Hamiltonian and enlarging the Hilbert space to include a degree of freedom corresponding to a boundary time shift, the algebra becomes a type II_∞ von Neumann algebra. In [21], Chandrasekaran, Longo, Penington, and Witten discuss how this construction can be generalized to asymptotically flat black holes. Starting with quantum fields in the exterior of a Schwarzschild black hole, by including the ADM Hamiltonian in the set of operators and the time shift degree of freedom in the Hilbert space, the algebra of these operators in the bulk spacetime is type II_∞. A type II von Neumann algebra (unlike a type III algebra) has a notion of a (renormalized) trace, density matrix and corresponding von Neumann entropy associated with the density matrix (for a review of von Neumann algebras and their classification, see [22]). In [16], Chandrasekaran, Penington, and Witten (CPW) showed that the entropy of semiclassical states in the boundary algebra is equal to the generalized entropy of the

^{*} mohd.ali@students.iiserpune.ac.in

[†] suneeta@iiserpune.ac.in

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.

black hole at the bifurcation surface. This has a bulk interpretation—in semiclassical gravity in the $G \rightarrow 0$ limit, the generalized entropy is the entropy of the algebra of operators in the black hole exterior. Finally, CPW use the monotonicity of the entropy of the algebra under trace preserving inclusions to prove a version of the generalized second law (GSL) showing monotonicity of generalized entropy between early and late times.

In this paper, we study the generalization of these results to an arbitrary diffeomorphism invariant theory of gravity by also including gravitons. We first write the black hole entropy in such a theory at an arbitrary horizon cut, which is the Wald entropy [12,13] with an extra term representing an ambiguity in the Wald entropy for a nonstationary black hole [14]. We work in semiclassical gravity. We consider a static (therefore stationary) black hole that is slightly perturbed due to infalling quantum matter and gravitons. In the limit when the cut $v \rightarrow \infty$, the perturbed black hole approaches a stationary black hole (v is the affine parameter along the null generator of the horizon). We compute the entropy at $v \rightarrow \infty$ minus the entropy at the bifurcation surface up to quadratic order in the perturbation, and we take into account the contribution due to gravitons to the stress-energy tensor. It is possible to fix the ambiguity in the Wald entropy in such a way that this difference of entropies takes a simplified form proportional to the vv component of the stress-energy tensor. Generalizing the computations of CPW [16], we find the difference of *generalized* entropies at $v \rightarrow \infty$ and at the bifurcation surface to be proportional to the relative entropy, which is non-negative. By computing the entropy of the extended von Neumann algebra of the black hole exterior [20], we show that the entropy of the algebra is the generalized entropy at the bifurcation surface just as in [16] for Einstein gravity. All the above constructions can be done for asymptotically flat static black holes [21]. Finally, we discuss the monotonicity result of CPW [16] who show that monotonicity of relative entropy under trace preserving inclusions can be used to argue that the generalized entropy at late times is more than that at early times. For this, we need asymptotically AdS black holes with a holographic dual, but modulo this change, the monotonicity result of CPW goes through for the generalized entropy of a black hole in an arbitrary diffeomorphism invariant theory of gravity.

In Sec. II, we discuss the difference of entropies at $v \rightarrow \infty$ and at the bifurcation surface for a slightly perturbed black hole. We use boost arguments which we summarize in Sec. II A to simplify this difference of entropies. By expanding the Raychaudhuri equation order by order in the perturbation parameter, we compute the change in entropies to quadratic order in Sec. II D both without graviton contributions to the stress-energy tensor, and with the graviton contribution included. In Sec. III, we discuss the entropy of the algebra of operators in the black hole exterior. We first summarize salient results from earlier

papers of Witten [20], Chandrasekaran, Longo, Penington and Witten [21] and CPW [16] who discussed how the entropy of the algebra was related to the generalized entropy in Einstein gravity. We then generalize these results to an arbitrary diffeomorphism invariant theory of gravity. We conclude in Sec. IV with a discussion.

II. ENTROPY CHANGE IN HIGHER CURVATURE THEORY

In what follows and the rest of the paper, we work in units where $G = 1$. Consider an entropy function for a black hole horizon in an arbitrary diffeomorphism invariant theory of gravity with matter,

$$S = \frac{1}{4} \int \rho \sqrt{h} d^{D-2}x \quad (2.1)$$

where $\rho = 1 + \rho_w + \Omega$, where h is the induced metric on the $D - 2$ dimensional transverse cut on the horizon and $1 + \rho_w$ is the Wald local entropy density [12,13]. As is well known, the Wald entropy is unambiguously defined for a stationary black hole, but suffers from ambiguities when evaluated on a non-stationary black hole. These ambiguities were pointed out by Jacobson, Kang and Myers (JKM) [14] and by Iyer and Wald [13]. Ω is a correction to Wald entropy density representing this JKM ambiguity, such that it vanishes for a stationary solution. We are interested in a black hole spacetime with a regular bifurcation surface \mathcal{B} , which is slightly perturbed from stationarity by throwing some quantum matter. Let v be an affine parameter along the null generator of the future horizon, such that $v = 0$ is the bifurcation surface as shown in Fig. 1. Then, the entropy at an arbitrary horizon cut (given by $v = \text{constant}$) is

$$S[v] = \frac{1}{4} \int_v \rho \sqrt{h} d^{D-2}x \quad (2.2)$$

where the subscript v in the integral indicates that the integral is over the transverse space at fixed v on the horizon. We can compute the change in the entropy along the horizon,

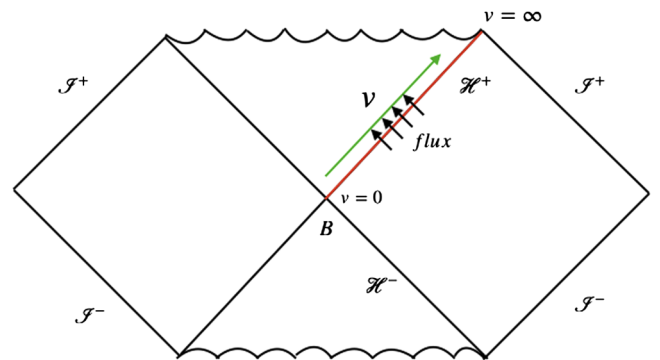


FIG. 1. Accretion of matter across the horizon in an asymptotically flat black hole.

$$\frac{dS}{dv} = \frac{1}{4} \int_v \sqrt{h} d^{D-2}x \left(\frac{d\rho}{dv} + \theta\rho \right) \quad (2.3)$$

where expansion $\theta \equiv \frac{1}{\sqrt{h}} \frac{d\sqrt{h}}{dv}$. To compute change in the entropy from $v = 0$ to $v \rightarrow \infty$, we can integrate both the sides with respect to v . This yields

$$\Delta S = \frac{1}{4} \int_0^\infty dv \int_v \sqrt{h} d^{D-2}x \left(\frac{d\rho}{dv} + \theta\rho \right). \quad (2.4)$$

Here, $\Delta S = S(\infty) - S(0)$. Using integration by parts,

$$\begin{aligned} \Delta S &= \frac{1}{4} \int_v \left\{ v\sqrt{h} \left(\frac{d\rho}{dv} + \theta\rho \right) \right\} \Big|_{v=0}^{v \rightarrow \infty} d^{D-2}x \\ &\quad - \frac{1}{4} \int_0^\infty dv \int \sqrt{h} d^{D-2}x v \\ &\quad \times \left\{ \frac{d^2\rho}{dv^2} + \frac{d\theta}{dv}\rho + 2\frac{d\rho}{dv}\theta + \theta^2\rho \right\}. \end{aligned} \quad (2.5)$$

We will assume that $\sqrt{h} \left(\frac{d\rho}{dv} + \theta\rho \right)$ goes to zero faster than $\frac{1}{v}$, therefore the first term in the above equation is identically zero and we are left with

$$\begin{aligned} \Delta S &= -\frac{1}{4} \int_0^\infty dv \int \sqrt{h} d^{D-2}x v \\ &\quad \times \left\{ \frac{d^2\rho}{dv^2} + \frac{d\theta}{dv}\rho + 2\frac{d\rho}{dv}\theta + \theta^2\rho \right\}. \end{aligned} \quad (2.6)$$

To compute ΔS order by order, we will now consider the metric perturbation sourced by a stress-energy tensor of order ϵ , i.e., $\langle T_{vv} \rangle \sim O(\epsilon)$. We will also assume that the perturbation is about a stationary black hole background and at late times, the black hole will again settle down into a stationary state. The perturbation expansion we are interested in is

$$g_{\mu\nu} = g_{\mu\nu}^{(0)} + \epsilon^{\frac{1}{2}} g_{\mu\nu}^{(\frac{1}{2})} + \epsilon g_{\mu\nu}^{(1)} + O(\epsilon^{\frac{3}{2}}), \quad (2.7)$$

where the zeroth order term corresponds to the stationary black hole solution with regular bifurcation surface, the $\sqrt{\epsilon}$ term is due to quantized graviton fluctuations, and the ϵ term is due to the gravitational field of matter or gravitons. We can think of ϵ as \hbar . We want to emphasize that Ω vanishes at order $\sqrt{\epsilon}$ at the bifurcation surface [23], a fact which will be useful later in the calculations.

A. Boost argument

We now use boost arguments first used in [24] and later in [25,26]. The metric near any null hypersurface and therefore near the event horizon can be given in Gaussian null coordinates as

$$\begin{aligned} ds^2 &= 2dvdu - u^2 X(u, v, x^k) dv^2 \\ &\quad + 2u\omega_i(u, v, x^k) dv dx^i + h_{ij}(u, v, x^k) dx^i dx^j \end{aligned} \quad (2.8)$$

where v is an affine parameter along the null generator of the horizon, x^i corresponds to coordinates on the $D-2$ transverse surface (cut) and u is chosen in a way that $\partial_v \cdot \partial_u = 1$ and $\partial_u \cdot \partial_i = 0$. In this coordinate system, $u = 0$ is the future horizon and $u = 0, v = 0$ corresponds to the bifurcation surface \mathcal{B} . These coordinates may not cover the entire spacetime, but the near-horizon region of any dynamical black hole spacetime can always be written in this form. Now, the black hole spacetime we consider in this paper is a static (therefore stationary) black hole spacetime which is a solution in a diffeomorphism invariant theory of gravity. Then, the black hole event horizon is a Killing horizon [12]. First consider the case where this horizon is a Killing horizon with respect to the boost field $\xi = v\partial_v - u\partial_u$ (see [25]). This is true for any stationary black hole spacetime which, near the horizon, looks like a Rindler spacetime, hence the terminology ‘‘boost field.’’ The near-horizon metric of this stationary black hole will then be boost invariant, i.e., the Lie derivative of the metric $\mathcal{L}_\xi g_{\mu\nu} = 0$. Then, the near-horizon metric (2.8) is of the form

$$ds_{st}^2 = 2dvdu - u^2 X(uv, x^k) dv^2 + h_{ij}(uv, x^k) dx^i dx^j \quad (2.9)$$

Here, $\omega_i = 0$ since the spacetime is static. This is the most general form of a static spacetime with a Killing horizon near the horizon. It can easily be seen from (2.8) that along the horizon, any nonzero tensor A , which is constructed out of metric components can always be written as $A = \partial_v^n \partial_u^m B$, where m, n are integers and B is constructed out of metric components X, ω_i, h_{ij} and their derivatives with respect to ∇_i . Then, we can associate a boost weight with these tensors as boost weight = $\#v$ index $- \#u$ index. Furthermore, we can write the schematic form for the vv component of any 2-tensor A_{vv} constructed from metric components as

$$A_{vv} = \tilde{X} \partial_v^2 Y + C \partial_v A \partial_v B. \quad (2.10)$$

Here, \tilde{X}, Y, C, A, B have boost weight zero and are constructed out of metric components. Now for the stationary black hole spacetime, the above equation reduces down to

$$A_{vv}|_{st} = u^2 \tilde{X} \partial_{uv}^2 Y + u^2 C \partial_{uv} A \partial_{uv} B. \quad (2.11)$$

This is because the stationary black hole has a Killing symmetry which reduces on the horizon to a scaling symmetry under $u \rightarrow pu$ and $v \rightarrow v/p$. Thus, the metric components in the stationary case only depend on uv at the horizon. This implies that the vv component of any 2-tensor A_{vv} constructed from metric components in a

stationary black hole spacetime vanishes at the future horizon, i.e., at $u = 0$.

$$A_{vv}|_{st}^{u=0} = 0. \quad (2.12)$$

Now, the vv component of the equation of motion for any higher curvature theory takes the following form,

$$R_{vv} + H_{vv} = 8\pi T_{vv} \quad (2.13)$$

where H_{vv} corresponds to a higher curvature contribution to the equation of motion. Using (2.11) for the stationary black hole $R_{vv} = 0$ and $H_{vv} = 0$, this implies that

$$T_{vv}|^{u=0} = 0 \quad (2.14)$$

for any *classical* matter stress-energy tensor. Furthermore, whenever a v derivative acts on the stationary background metric component, it gives a factor of u as well, since the metric component depends on v only through uv . Hence such a term will vanish at the future horizon $u = 0$. Therefore, from (2.10), the vv component of any 2-tensor A_{vv} linear in the metric perturbation at the future horizon $u = 0$ can always be written in the following form,

$$A_{vv}|^{u=0} = \partial_v^2 \zeta, \quad (2.15)$$

where ζ has boost weight zero and is constructed from the background metric and the linear perturbation over stationarity.

B. Semiclassical gravity equations

Following Chandrasekaran, Penington, and Witten [16] and Wall [11], we will work in semiclassical gravity where the expectation value of the matter stress energy tensor is a source term in the gravity equations. Now let us look at the semiclassical equations of motion. The vv component is

$$R_{vv} + H_{vv} = 8\pi \langle T_{vv} \rangle \quad (2.16)$$

Writing this order by order in ϵ , we get:

$$\epsilon^0: R_{vv} + H_{vv} = 0 \quad (2.17)$$

$$\epsilon^{\frac{1}{2}}: R_{vv}^{(\frac{1}{2})} + H_{vv}^{(\frac{1}{2})} = 0 \quad (2.18)$$

$$\epsilon^1: R_{vv}^{(1)} + H_{vv}^{(1)} = 8\pi \langle T_{vv}^Q \rangle. \quad (2.19)$$

where $T_{vv}^Q = T_{vv}^M + t_{vv}$, $\langle T_{vv}^M \rangle \sim O(\epsilon)$ is a matter stress energy tensor and $\langle t_{vv} \rangle \sim O(\epsilon)$, the pseudo-stress-energy tensor of the graviton. Further, $R_{vv}^{(\frac{1}{2})}$, $H_{vv}^{(\frac{1}{2})}$ are linear in $g_{\mu\nu}^{(\frac{1}{2})}$ perturbation, and $R_{vv}^{(1)}$, $H_{vv}^{(1)}$ are linear in $g_{\mu\nu}^{(1)}$ perturbation.

C. Raychaudhuri equation order by order

As is well known, the Raychaudhuri equation (2.18) plays a key role in the proof of the second law and we will use it later in our computation. The Raychaudhuri equation is given by,

$$\frac{d\theta}{dv} = - \left\{ \frac{\theta^2}{D-2} + \sigma^{\alpha\beta} \sigma_{\alpha\beta} + R_{vv} \right\} \quad (2.20)$$

Now, if we write it order by order in ϵ , we get,

$$\epsilon^0: \frac{d\theta^{(0)}}{dv} = 0 \quad (2.21)$$

$$\epsilon^{\frac{1}{2}}: \frac{d\theta^{(\frac{1}{2})}}{dv} = -R_{vv}^{(\frac{1}{2})} \quad (2.22)$$

$$\epsilon^1: \frac{d\theta^{(1)}}{dv} = - \left\{ \frac{\theta^{(\frac{1}{2})} \theta^{(\frac{1}{2})}}{D-2} + \sigma_{(\frac{1}{2})}^{\alpha\beta} \sigma_{\alpha\beta}^{(\frac{1}{2})} + R_{vv}^{(1)} + R_{vv}^{(\frac{1}{2})} \right\} \quad (2.23)$$

Further using (2.19), (2.23) can be written as

$$\frac{d\theta^{(1)}}{dv} = - \left\{ \frac{\theta^{(\frac{1}{2})} \theta^{(\frac{1}{2})}}{D-2} + \sigma_{(\frac{1}{2})}^{\alpha\beta} \sigma_{\alpha\beta}^{(\frac{1}{2})} + 8\pi \langle T_{vv}^Q \rangle - H_{vv}^{(1)} + R_{vv}^{(\frac{1}{2})} \right\}. \quad (2.24)$$

Furthermore, if we compute $R_{vv}^{(\frac{1}{2})}$ in TT (transverse traceless) gauge at the horizon,¹ we will get,

$$R_{vv}^{(\frac{1}{2})}{}_{TT} = -\frac{1}{4} \frac{dg_{(\frac{1}{2})}^{ij}}{dv} \frac{dg_{ij}^{(\frac{1}{2})}}{dv} + \frac{1}{2} \frac{d}{dv} \left(g_{(\frac{1}{2})}^{ij} \frac{dg_{ij}^{(\frac{1}{2})}}{dv} \right). \quad (2.25)$$

Now using the fact that $\frac{1}{2} \frac{dg_{ij}^{(\frac{1}{2})}}{dv} = \sigma_{ij}^{(\frac{1}{2})} + \frac{1}{D-2} g_{ij}^{(0)} \theta^{(\frac{1}{2})}$ [11] (i, j are transverse coordinate indices), we can write Eq. (2.25) as

$$R_{vv}^{(\frac{1}{2})}{}_{TT} = - \left(\frac{\theta^{(\frac{1}{2})} \theta^{(\frac{1}{2})}}{D-2} + \sigma_{(\frac{1}{2})}^{\alpha\beta} \sigma_{\alpha\beta}^{(\frac{1}{2})} \right) + \frac{1}{4} \frac{d^2}{dv^2} (g_{(\frac{1}{2})}^{ij} g_{ij}^{(\frac{1}{2})}) \quad (2.26)$$

which yields

$$\frac{d\theta^{(1)}}{dv} = - \left\{ 8\pi \langle T_{vv}^Q \rangle - H_{vv}^{(1)} + \frac{1}{4} \frac{d^2}{dv^2} (g_{(\frac{1}{2})}^{ij} g_{ij}^{(\frac{1}{2})}) \right\}. \quad (2.27)$$

(2.21) follows from the fact that the background solution is stationary. The other equations are obtained by expanding the Raychaudhuri equation order by order.

¹The perturbative expansion of the Ricci tensor to quadratic order can be found in [27].

D. Entropy change due to accretion of quantum matter across the horizon

In this subsection, we will compute the order-by-order change in the entropy due to the accretion of quantum matter and gravitons across the horizon. We will assume that the background black hole solution is stationary, as well as the final state of the black hole at late times. Also, we assume the perturbation will fall off fast enough, so that all boundary terms at late times vanish.² To compute the entropy change order by order, first we will write the perturbative expansion of entropy density as

$$\rho = \rho^{(0)} + \epsilon^{\frac{1}{2}}\rho^{(\frac{1}{2})} + \epsilon\rho^{(1)} + O(\epsilon^{\frac{3}{2}}) \quad (2.28)$$

Now, we have all the tools to compute the change in entropy. First we will do the change in entropy computation in the absence of the graviton. Then we will do the computation in which we will include gravitons.

1. Entropy change without graviton contribution

When there is no graviton, all the terms with $g_{\mu\nu}^{(\frac{1}{2})}$ perturbation will go away in all of the above equations. Also, the stress-energy tensor will have only the matter contribution, i.e., $\langle T_{vv}^Q \rangle = \langle T_{vv}^M \rangle$, which we will take to be $O(\epsilon)$. The background solution is a stationary black hole, with a Killing horizon and regular bifurcation sphere. For the stationary black hole, the expansion coefficient of the horizon is zero and the entropy density will be independent of the chosen horizon cut. Therefore

$$\epsilon^0: \Delta S^{(0)} = 0 \quad (2.29)$$

Now, we will compute the change in the entropy due to accretion of matter by the stationary black hole, which takes it away from stationarity. As we have already mentioned, the black hole will settle down into a new stationary state at late times. Now using (2.6) and (2.27) with the fact that there is no $\rho^{(\frac{1}{2})}$ and $\Theta^{\frac{1}{2}}$ we will get

$$\epsilon: \Delta\delta S^{(1)} = -\frac{1}{4} \int_0^\infty dv \int d^{D-2}x \sqrt{h} v \left\{ \frac{d^2\rho^{(1)}}{dv^2} + \frac{d\theta^{(1)}}{dv} \rho^{(0)} \right\}. \quad (2.30)$$

Here, δ is the perturbation away from stationarity. Since $\rho^{(0)} = 1 + \rho_w^{(0)}$,

²This implies that at late times all the perturbations would have either crossed the horizon or gone to asymptotic infinity. For AdS black hole spacetime with reflecting boundary conditions, all perturbations will cross the horizon. Also, we dynamically impose the gauge (2.8) at all times.

$$\frac{d\theta^{(1)}}{dv} \rho^{(0)} = \frac{d\theta^{(1)}}{dv} + \frac{d\theta^{(1)}}{dv} \rho_w^{(0)} = -8\pi \langle T_{vv}^Q \rangle + H_{vv}^{(1)} - R_{vv}^{(1)} \rho_w^{(0)}. \quad (2.31)$$

Equation (2.31) is obtained using (2.23), (2.27) and the fact that there is no $g^{(\frac{1}{2})}$ perturbation, which will make $\frac{d\theta^{(1)}}{dv} = -R_{vv}^{(1)}$. Further, we used the equation of motion to rewrite $R_{vv}^{(1)}$ in terms of the stress energy tensor. Putting the Eq. (2.31) in the Eq. (2.30), we get

$$\begin{aligned} \epsilon: \Delta\delta S^{(1)} = & -\frac{1}{4} \int_0^\infty dv \int d^{D-2}x \sqrt{h} v \\ & \times \left\{ \left(\frac{d^2\rho^{(1)}}{dv^2} - R_{vv}^{(1)} \rho_w^{(0)} + H_{vv}^{(1)} \right) - 8\pi \langle T_{vv}^Q \rangle \right\}. \end{aligned} \quad (2.32)$$

We note that $(\frac{d^2\rho^{(1)}}{dv^2} - R_{vv}^{(1)} \rho_w^{(0)} + H_{vv}^{(1)})$ is constructed out of background metric components and the perturbation and is linear in the perturbation.³ Therefore, using boost arguments we can write $(\frac{d^2\rho^{(1)}}{dv^2} - R_{vv}^{(1)} \rho_w^{(0)} + H_{vv}^{(1)}) = \partial_v^2 \zeta_{(1)}$, which will yield

$$\Delta\delta S^{(1)} = -\frac{1}{4} \int_0^\infty dv \int d^{D-2}x \sqrt{h} v \{ \partial_v^2 \zeta_{(1)} - 8\pi \langle T_{vv}^Q \rangle \}. \quad (2.33)$$

We can simplify the above term using integration by parts,

$$\begin{aligned} \Delta\delta S^{(1)} = & 2\pi \int_0^\infty dv \int d^{D-2}x \sqrt{h} v \langle T_{vv}^Q \rangle \\ & + \frac{1}{4} \int d^{D-2}x \sqrt{h} \int_0^\infty dv \partial_v \zeta_{(1)} \\ & - \frac{1}{4} \int d^{D-2}x \sqrt{h} (v \partial_v \zeta_{(1)})|_{v=0}^\infty. \end{aligned} \quad (2.34)$$

Now we assume fall-off conditions at late times i.e., all perturbations and their derivatives should fall fast enough such that this boundary term goes to zero at late times. The contribution from the last term in (2.34) also vanishes at $v = 0$. Let us recall that we are in the gauge (2.8) in which the horizon is always at $u = 0$. Then we will get

$$\begin{aligned} \Delta\delta S^{(1)} = & 2\pi \int_0^\infty dv \int d^{D-2}x \sqrt{h} v \langle T_{vv}^Q \rangle \\ & - \frac{1}{4} \int d^{D-2}x \sqrt{h} \zeta_{(1)}|_{v=0}. \end{aligned} \quad (2.35)$$

We can get rid of the second term by assuming that we can fix the JKM ambiguity in such a way that

³We can replace the ordinary derivatives with respect to v in the first term with covariant derivatives in the gauge we are in.

$$\frac{d^2\rho^{(1)}}{dv^2} - R_{vv}^{(1)}\rho_w^{(0)} + H_{vv}^{(1)} = \partial_v^2\zeta_{(1)} = 0 \quad (2.36)$$

everywhere on the horizon. This will get rid of the term which is giving rise to the second term in (2.35). This is because then $\zeta_{(1)}(v) = av + b$, where a and b are only functions of transverse coordinate. $\zeta_{(1)}$ is constructed out of the background metric and the linear perturbation in the gauge (2.8). We have to further impose the fall-off conditions on the perturbation, i.e., the perturbation and its derivatives with respect to v must go to zero at late times. Thus $\zeta_{(1)}(v) = 0$. There is no contradiction with the fact that Ω vanishes at the bifurcation surface in linear order. It is shown in [28] that the second term in (2.35) is zero for $F(R)$ gravity and arbitrary order Lovelock gravity. It is also argued there that this will be true for an arbitrary diffeomorphic theory at linear order. Therefore, assuming this,

$$\Delta\delta S^{(1)} = 2\pi \int_0^\infty dv \int d^{D-2}x\sqrt{h}v \langle T_{vv}^Q \rangle \quad (2.37)$$

The above derivation is of course true even when the accreting matter is classical. For classical matter, imposing the null energy condition i.e., $T_{vv}^Q \geq 0$ will give the second law.

2. Entropy change with the graviton contribution included

In this section, we include the graviton contribution, and therefore we will work with the full perturbation expansion. We will again do an order by order expansion. Using (2.6) and the fact that background solution is stationary,

$$\epsilon^0: \Delta S^{(0)} = 0 \quad (2.38)$$

Now, let us compute change in entropy at $\epsilon^{\frac{1}{2}}$ order. Writing (2.6) at $\epsilon^{\frac{1}{2}}$ will give

$$\epsilon^{\frac{1}{2}}: \Delta\delta S^{(\frac{1}{2})} = -\frac{1}{4} \int_0^\infty dv \int d^{D-2}x\sqrt{h}v \left\{ \frac{d^2\rho^{(\frac{1}{2})}}{dv^2} + \frac{d\theta^{(\frac{1}{2})}}{dv} \rho^{(0)} \right\}. \quad (2.39)$$

Here, δ corresponds to entropy change due to a perturbation that takes the solution away from stationarity. Using (2.22), we can write (2.39) as

$$\Delta\delta S^{(\frac{1}{2})} = -\frac{1}{4} \int_0^\infty dv \int d^{D-2}x\sqrt{h}v \left\{ \frac{d^2\rho^{(\frac{1}{2})}}{dv^2} - R_{vv}^{(\frac{1}{2})}\rho^{(0)} \right\} \quad (2.40)$$

Furthermore using the boost argument in (2.15), we can write $R_{vv}^{(\frac{1}{2})} = \partial_v^2\zeta_{(\frac{1}{2})}$, where $\zeta_{(\frac{1}{2})}$ is constructed out of the background metric and linear perturbation in $g_{\mu\nu}^{(\frac{1}{2})}$ where we

work in the gauge (2.8). Further, $\rho^{(0)}$ is independent of v . Therefore, the Eq. (2.40) becomes

$$\Delta\delta S^{(\frac{1}{2})} = -\frac{1}{4} \int_0^\infty dv \int d^{D-2}x\sqrt{h}v \frac{d^2}{dv^2} \left\{ \rho^{(\frac{1}{2})} - \rho^{(0)}\zeta_{(\frac{1}{2})} \right\}. \quad (2.41)$$

Using integration by parts,

$$\begin{aligned} \Delta\delta S^{(\frac{1}{2})} &= -\frac{1}{4} \int d^{D-2}x\sqrt{h}v \frac{d}{dv} \left\{ \rho^{(\frac{1}{2})} - \rho^{(0)}\zeta_{(\frac{1}{2})} \right\} \Big|_{v=0}^{v\rightarrow\infty} \\ &\quad + \frac{1}{4} \int_0^\infty dv \int d^{D-2}x\sqrt{h} \frac{d}{dv} \left\{ \rho^{(\frac{1}{2})} - \rho^{(0)}\zeta_{(\frac{1}{2})} \right\}. \end{aligned} \quad (2.42)$$

Using the fall-off condition on the perturbation at late times, the first term vanishes. Therefore, we get

$$\Delta\delta S^{(\frac{1}{2})} = \frac{1}{4} \int_0^\infty dv \int d^{D-2}x\sqrt{h} \frac{d}{dv} \left\{ \rho^{(\frac{1}{2})} - \rho^{(0)}\zeta_{(\frac{1}{2})} \right\}. \quad (2.43)$$

Integrating and using fall-off condition gives

$$\Delta\delta S^{(\frac{1}{2})} = -\frac{1}{4} \int d^{D-2}x\sqrt{h} \left\{ \rho^{(\frac{1}{2})} - \rho^{(0)}\zeta_{(\frac{1}{2})} \right\} \Big|_{v=0}. \quad (2.44)$$

Using the fact that $\theta = \frac{d}{dv} \log(\sqrt{h})$ and (2.22), we can write $\delta \log(\sqrt{h}) = -\zeta_{(\frac{1}{2})}$. Therefore, we can write (2.44) as

$$\Delta\delta S^{(\frac{1}{2})} = -\frac{1}{4} \int_{\mathcal{B}} d^{D-2}x\sqrt{h} \left\{ \rho^{(\frac{1}{2})} + \delta \log(\sqrt{h})\rho^{(0)} \right\}. \quad (2.45)$$

Using (2.1), it is straightforward to verify that

$$\Delta\delta S^{(\frac{1}{2})} = -\delta S^{(\frac{1}{2})} \Big|_{\mathcal{B}} \quad (2.46)$$

\mathcal{B} is the bifurcation surface. Now, we use Theorem 6.1 in Iyer and Wald (IW) [13], i.e. $(\delta S = \delta\mathcal{E} - \Omega_H\delta\mathcal{J}) \Big|_{\mathcal{B}}$, where \mathcal{E} is the canonical energy and \mathcal{J} is the canonical angular momentum of the black hole in the covariant phase space formalism.⁴ This was proved by IW at the bifurcation surface for any non-stationary perturbation satisfying the linearized equation of motion.⁵ Now in our case, there is no stress-energy tensor at $\epsilon^{(\frac{1}{2})}$ order. This implies

$$\Delta\delta S^{(\frac{1}{2})} = -\delta S^{(\frac{1}{2})} \Big|_{\mathcal{B}} = 0 \quad (2.47)$$

⁴Since we have a static black hole, angular velocity at the horizon is zero.

⁵IW's first law at the bifurcation surface \mathcal{B} is unaffected by the JKM ambiguity.

Hence, $\delta S^{(\frac{1}{2})}(0) = \delta S^{(\frac{1}{2})}(\infty) = 0$. Therefore, from (2.47), if $\delta S^{(\frac{1}{2})}$ is nonzero at any cut, then for some range of v entropy will definitely decrease. This violates the second law. The only way for the second law to be true is to assume that $\delta S^{(\frac{1}{2})}$ will vanish at arbitrary cut. It was shown explicitly by the authors in [23,28] that

$$\mathcal{R}_{vv} \equiv \frac{d^2\rho}{dv^2} - R_{vv}\rho_w^{(0)} + H_{vv} \quad (2.48)$$

vanishes for $F(R)$ theory and Lovelock theory of arbitrary order, at the linear order in perturbation theory about the stationary black hole (perturbation can be nonstationary). Using (2.18), it can be checked that the term in curly brackets in (2.39) is the same as $\mathcal{R}_{vv}^{(\frac{1}{2})}$. The authors in [23,28] also argued that this relation may be true for an arbitrary theory of gravity with an appropriate definition of local entropy density. Vanishing of $\delta S^{(\frac{1}{2})}$ in general will yield,

$$\frac{d^2\rho^{(\frac{1}{2})}}{dv^2} = -\rho^{(0)}\partial_v\theta^{(\frac{1}{2})} \quad (2.49)$$

which after integration and using the boundary condition that the perturbation vanishes at late times will give $\rho^{(\frac{1}{2})} = \rho^{(0)}\zeta^{(\frac{1}{2})}$.

Now, let us compute the ϵ order change in entropy, writing (2.6) to the ϵ order. We get

$$\begin{aligned} \Delta\delta S^{(1)} = & -\frac{1}{4}\int_0^\infty dv \int v \left\{ \frac{d^2\rho^{(1)}}{dv^2} + \rho^{(0)}\frac{d\theta^{(1)}}{dv} + \frac{d\theta^{(\frac{1}{2})}}{dv}\rho^{(\frac{1}{2})} \right. \\ & \left. + 2\theta^{(\frac{1}{2})}\frac{d\rho^{(\frac{1}{2})}}{dv} + \rho^{(0)}\theta^{(\frac{1}{2})}\theta^{(\frac{1}{2})} \right\} \sqrt{h}d^{D-2}x. \end{aligned} \quad (2.50)$$

Using (2.19), (2.23), and (2.48), the first two terms in the above expression can be written as

$$\frac{d^2\rho^{(1)}}{dv^2} + \rho^{(0)}\frac{d\theta^{(1)}}{dv} = -8\pi\langle T_{vv}^Q \rangle + \mathcal{R}_{vv}^{(1)} - \frac{1}{4d}\frac{d^2}{dv^2}(g_{(\frac{1}{2})}^{ij}g_{ij}^{(\frac{1}{2})}\rho^{(0)}). \quad (2.51)$$

Using (2.48), (2.15) and the fact that $H_{vv}^{(1)}$ and $R_{vv}^{(1)}$ are constructed out of background metric components and the perturbation and are linear in $g_{\mu\nu}^{(1)}$ perturbation in the gauge (2.8), $\mathcal{R}_{vv}^{(1)}$ can be written as $\mathcal{R}_{vv}^{(1)} = \partial_v^2\zeta^{(1)}$. This yields

$$\frac{d^2\rho^{(1)}}{dv^2} + \rho^{(0)}\frac{d\theta^{(1)}}{dv} = -8\pi\langle T_{vv}^Q \rangle + \partial_v^2\zeta^{(1)}. \quad (2.52)$$

where, $\zeta^{(1)} = \zeta_{(1)} - \frac{1}{4}(g_{(\frac{1}{2})}^{ij}g_{ij}^{(\frac{1}{2})}\rho^{(0)})$. Putting the above equation in (2.50) we get,

$$\begin{aligned} & -\frac{1}{4}\int_0^\infty dv \int v\sqrt{h}d^{D-2}x \left(\frac{d^2\rho^{(1)}}{dv^2} + \rho^{(0)}\frac{d\theta^{(1)}}{dv} \right) \\ & = 2\pi \int_0^\infty dv \int v\sqrt{h}d^{D-2}x \langle T_{vv}^Q \rangle \\ & \quad - \frac{1}{4}\int \sqrt{h}\zeta'_{(1)}d^{D-2}x|_{v=0}. \end{aligned} \quad (2.53)$$

We get the above equation using integration by parts in the $\partial_v^2\zeta'_{(1)}$ integral and the fact that the term at $v \rightarrow \infty$ will vanish due to the fall-off condition. Let us consider the rest of the terms in (2.50), we will call it $A(\frac{1}{2}, \frac{1}{2})$,

$$\begin{aligned} A\left(\frac{1}{2}, \frac{1}{2}\right) = & -\frac{1}{4}\int_0^\infty dv \int v \left\{ \frac{d\theta^{(\frac{1}{2})}}{dv}\rho^{(\frac{1}{2})} \right. \\ & \left. + 2\theta^{(\frac{1}{2})}\frac{d\rho^{(\frac{1}{2})}}{dv} + \rho^{(0)}\theta^{(\frac{1}{2})}\theta^{(\frac{1}{2})} \right\} \sqrt{h}d^{D-2}x. \end{aligned} \quad (2.54)$$

After integrating (2.49) once, we get $\frac{d\rho^{(\frac{1}{2})}}{dv} = -\rho^{(0)}\theta^{(\frac{1}{2})}$. Putting this in (2.54) will yield

$$A\left(\frac{1}{2}, \frac{1}{2}\right) = -\frac{1}{4}\int_0^\infty dv \int v \left\{ \frac{d\theta^{(\frac{1}{2})}}{dv}\rho^{(\frac{1}{2})} + \theta^{(\frac{1}{2})}\frac{d\rho^{(\frac{1}{2})}}{dv} \right\} \sqrt{h}d^{D-2}x \quad (2.55)$$

which can be further simplified using integration by parts and using fall-off conditions as $v \rightarrow \infty$,

$$A\left(\frac{1}{2}, \frac{1}{2}\right) = \frac{1}{4}\int_0^\infty dv \int \theta^{(\frac{1}{2})}\rho^{(\frac{1}{2})}\sqrt{h}d^{D-2}x. \quad (2.56)$$

Using $\frac{d\rho^{(\frac{1}{2})}}{dv} = -\rho^{(0)}\theta^{(\frac{1}{2})}$ in (2.56), we get

$$A\left(\frac{1}{2}, \frac{1}{2}\right) = -\frac{1}{8}\int_0^\infty dv \int \frac{d}{dv} \left(\frac{(\rho^{(\frac{1}{2})})^2}{\rho^{(0)}} \right) \sqrt{h}d^{D-2}x. \quad (2.57)$$

After integrating and using fall-off at late times, we will get

$$A\left(\frac{1}{2}, \frac{1}{2}\right) = \frac{1}{8}\int \left(\frac{(\rho^{(\frac{1}{2})})^2}{\rho^{(0)}} \right) \sqrt{h}|_{v=0}d^{D-2}x \quad (2.58)$$

This quantity is thus manifestly positive. From (2.53) and (2.58), we get

$$\begin{aligned} \Delta\delta S^{(1)} = & 2\pi \int_0^\infty dv \int v\sqrt{h}d^{D-2}x \langle T_{vv}^Q \rangle \\ & - \frac{1}{4}\int \sqrt{h}\zeta'_{(1)}d^{D-2}x|_{v=0} + A\left(\frac{1}{2}, \frac{1}{2}\right). \end{aligned} \quad (2.59)$$

We know that the Wald entropy has JKM ambiguities when the metric is not stationary. That was a motivation for

putting Ω as the correction to the Wald entropy in the definition of local entropy density. We now fix Ω such that the last two terms in (2.59) vanish. These terms are anyway zero for a stationary black hole, as can be seen from the expression for $A(\frac{1}{2}, \frac{1}{2})$ in (2.58) and $\mathcal{R}_{vv}^{(1)}$ is zero for a stationary black hole from results in the subsection on the boost argument.⁶ So Ω will be nonzero only when the metric is not stationary. For the cases when it is possible to set these two terms to zero by a choice of Ω , we will get

$$\Delta\delta S^{(1)} = 2\pi \int_0^\infty dv \int v\sqrt{h}d^{D-2}x \langle T_{vv}^Q \rangle \quad (2.60)$$

Some of the ambiguities in Wald's entropy were fixed for some class of theories at linear order in perturbation theory [24]. Moreover, this entropy was shown to be equal to holographic entanglement entropy computed by Dong [29]. It is also pointed out in [24] that considering the second law at linear order does not fix the ambiguities at higher order. Therefore these results are not in contradiction with our computation. One can also get rid of the second and third term in (2.59), by restricting the perturbation to a special class of perturbations which vanishes at $v = 0$. One physical case in which such perturbation can be realized is when matter falls after $v = 0$.

III. THE ENTROPY OF ALGEBRA AND GENERALIZED ENTROPY IN HIGHER CURVATURE THEORY

In this section, we utilize the algebraic approach to quantum field theory, specifically the constructions of Chandrasekaran, Longo, Penington, and Witten (CLPW) in [21], and Chandrasekaran, Penington, and Witten (CPW) in [16], to study black holes in higher curvature theories. We have employed their construction to prove a version of the GSL (generalized second law) for an arbitrary diffeomorphism invariant theory of gravity, with certain appropriate assumptions. The setup that we are interested in involves both the asymptotically flat and asymptotically AdS stationary black hole solutions in an arbitrary diffeomorphism invariant theory. Throughout this section we follow the notation of CPW. In this section, we will first review the work of CPW and then generalize it to our case.

A. Brief review of the recent papers

The series of recent papers by CLPW, Witten and CPW [16,20,21] have helped to understand the generalized entropy introduced by Bekenstein [1] better. They have addressed the question of why the generalized entropy is well-defined, whereas the gravity contribution and the

⁶For a stationary black hole, v derivatives of ρ are zero and R_{vv} and H_{vv} are zero for a stationary metric as discussed using boost arguments.

quantum field contribution in the generalized entropy are not well-defined separately. CPW showed for an eternal black hole that is either asymptotically flat and asymptotically AdS, that there is an entropy associated with the von Neumann algebra of quantum fields (including the graviton) in the black hole left or right exterior, together with either the left or right ADM/boundary Hamiltonian. Further, they showed that this algebra entropy is the generalized entropy at the bifurcation surface of the black hole. They have also discussed the monotonicity of the generalized entropy of asymptotically AdS black holes by using techniques from von Neumann algebras.

We first provide a brief review of the salient features of these constructions and then generalize them to our case of black holes in arbitrary diffeomorphism invariant theories of gravity.

Let M be the asymptotically flat, maximally extended Schwarzschild black hole in Einstein's theory or the maximally extended AdS-Schwarzschild black hole. We consider quantum fields in this spacetime, including gravitons. The left and right exterior regions of M will be denoted by ℓ and r respectively, while L and R will be used to denote left and right spatial infinity. Let \mathcal{H}_0 be the Hilbert space of this theory that we get by quantizing the fields and the local algebra of observables of the left and right exterior region be $\mathcal{A}_{\ell,0}$ and $\mathcal{A}_{r,0}$ respectively as shown in Fig. 2. It is well known that algebras $\mathcal{A}_{\ell,0}$ and $\mathcal{A}_{r,0}$ are type III₁ factors (their centers are trivial) [22,30,31].⁷ Moreover, $\mathcal{A}_{\ell,0}$ and $\mathcal{A}_{r,0}$ are each other's commutants (i.e., all the operators of $\mathcal{A}_{\ell,0}$ commute with all the operators in $\mathcal{A}_{r,0}$).

The spacetime is stationary and equipped with a time translation Killing field V . V is future directed in the right exterior region and past-directed in the left exterior region. Due to background diffeomorphism invariance, one can define a conserved quantity \hat{h} associated with the time translation vector field V . Let S be the bulk Cauchy surface going from the spatial infinity of the right exterior region to the spatial infinity of the left exterior region, through the bifurcation surface as shown in Fig. 2. Then \hat{h} can be defined as

$$\hat{h} = \int_S d\Sigma^\mu V^\nu T_{\mu\nu}. \quad (3.1)$$

Here, $T_{\mu\nu}$ is the stress-energy tensor of the bulk fields.⁸ In Tomita-Takesaki theory of the quantum fields in the black hole exterior, $\beta\hat{h}$ is the modular Hamiltonian associated with the Hartle-Hawking state $|\Psi_{HH}\rangle$ of the black

⁷The algebra of operators in quantum field theory in a causal wedge is always a von Neumann algebra of type III [17]. When the center is trivial it is a type III₁ algebra.

⁸ $T_{\mu\nu}$ includes the contribution from the pseudo-stress tensor of gravitons.

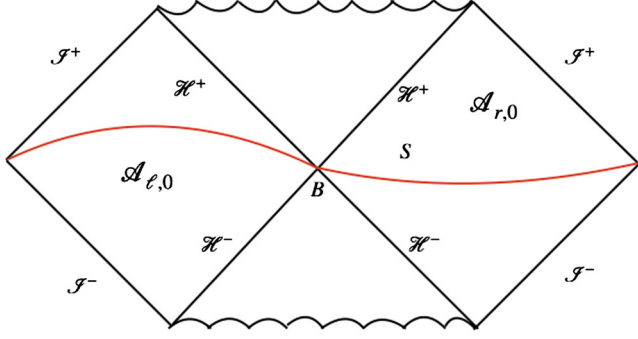


FIG. 2. Maximally extended stationary black hole with Cauchy surface S . The bulk algebra of the left and the right exterior region is $\mathcal{A}_{\ell,0}$ and $\mathcal{A}_{r,0}$ respectively.

hole and β is the Hawking temperature⁹ [32,33]. It is well known that \hat{h} in Einstein's gravity is the difference between the right ADM Hamiltonian H_R and the left ADM Hamiltonian H_L i.e., $\hat{h} = H_R - H_L$.

CLPW and CPW now extend the type III₁ algebra $\mathcal{A}_{\ell,0}$ and $\mathcal{A}_{r,0}$ by including one more operator h_L with $\mathcal{A}_{\ell,0}$ and h_R similarly for the right algebra. This extended (crossed product) algebra acts on an extended Hilbert space $\mathcal{H} = \mathcal{H}_0 \otimes L^2(\mathbb{R})$ where the extra degree of freedom that has been introduced is the time-shift (the sum of the times in the left and the right exteriors). The extended crossed product right algebra is denoted $\mathcal{A}_r = \mathcal{A}_{r,0} \rtimes \mathbb{R}_h$ and similarly for the left algebra. Here,

$$h_L = H_L - M_0 \quad h_R = H_R - M_0 \quad (3.2)$$

M_0 is the ADM mass of some reference black hole. CPW work in a microcanonical ensemble, i.e., an energy eigenstate centered around some energy M_0 (mass of the reference black hole).¹⁰ The algebra of observables for the right exterior region is studied in a semiclassical limit, i.e., $G \rightarrow 0$. In this limit, the ADM masses H_R and H_L diverge because the black hole mass M_0 (Schwarzschild radius divided by $2G$) diverges. So, CPW work with the nondivergent subtracted Hamiltonians h_L and h_R .

We can also write $\hat{h} = h_r - h_\ell$ where

$$h_r = \int_{S_1} d\Sigma^\mu V^\nu T_{\mu\nu} \quad (3.3)$$

$$h_\ell = - \int_{S_2} d\Sigma^\mu V^\nu T_{\mu\nu} \quad (3.4)$$

where S_1 and S_2 are the right exterior and the left exterior part of the Cauchy surface S as shown in Fig. 3. As pointed

⁹The modular operator is defined as $\Delta = \exp(-H_{\text{mod}})$ and for $|\Psi_{\text{HH}}\rangle$, $H_{\text{mod}} = \beta\hat{h}$.

¹⁰CPW explicitly do the microcanonical ensemble construction in the boundary CFT using a thermofield double state.

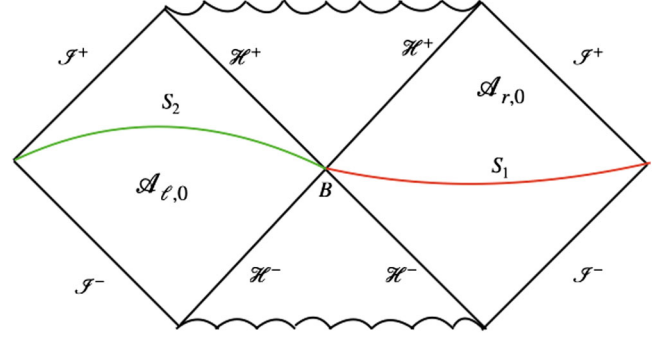


FIG. 3. This figure depicts the split of Cauchy surface S into union of the red Cauchy surface S_1 in the right exterior and the green Cauchy surface S_2 in the left exterior.

out by CPLW, h_r and h_ℓ have divergent fluctuations. Thus, such a splitting is not true, strictly speaking [20], but in the extended algebra of Witten, the modular operator has a factorization into a product of operators in the left and right algebra.

$$\hat{h} = h_r - h_\ell = h_R - h_L \quad (3.5)$$

Further,

$$h_R = \hat{h} + h_L \equiv \frac{h_\Psi}{\beta} + x \quad (3.6)$$

where $h_\Psi = H_{\text{mod}}$ is the modular Hamiltonian for the Hartle-Hawking state $|\Psi\rangle = |\Psi_{\text{HH}}\rangle$. $\mathcal{A}_r = \mathcal{A}_{r,0} \rtimes \mathbb{R}_h$ is the crossed product algebra of the algebra $\mathcal{A}_{r,0}$ by the modular group for the cyclic separating vector $|\Psi\rangle$. As discussed in [20], the algebra \mathcal{A}_r is then a von Neumann algebra of type II _{∞} . As explained in [20], unlike a type III algebra, a type II algebra has a notion of trace, density matrix, and entropy.

An operator in \mathcal{A}_r which we will denote as \hat{a} has the form $\hat{a} = a e^{(ish_\Psi/\beta)} \otimes e^{(isx)}$, where $a \in \mathcal{A}_{r,0}$. The states on which this operator acts has the form $|\hat{\Psi}\rangle = |\Psi\rangle \otimes g(x) \in \mathcal{H}$ where $|\Psi\rangle \in \mathcal{H}_0$ and $g(x) \in L^2(\mathbb{R})$. The most generic operator in \mathcal{A}_r can be written as

$$\hat{a} = \int_{-\infty}^{\infty} a(s) e^{is(x+\hat{h})} ds \quad (3.7)$$

where $a(s) \in \mathcal{A}_{r,0}$. Similarly, we can write the most general state as

$$|\hat{\Psi}\rangle = \int dx f(x) |\Psi\rangle |x\rangle \quad (3.8)$$

As we have already mentioned, a type II algebra has a trace, which is a positive linear functional on operators in the algebra satisfying $\text{tr}[\hat{a}\hat{b}] = \text{tr}[\hat{b}\hat{a}]$. The trace of an operator $\hat{a} \in \mathcal{A}_r$ can be denoted as [16],

$$\text{tr}[\hat{a}] = \int_{-\infty}^{\infty} dx e^{\beta x} \langle \hat{\Psi} | \hat{a} | \hat{\Psi} \rangle \quad (3.9)$$

where the expectation value $\langle \hat{\Psi} | \hat{a} | \hat{\Psi} \rangle$ is in general some nontrivial function of x . This trace is not the same as a standard trace on Hilbert space \mathcal{H} but one can think of it more as a renormalized trace. Also, tr is defined up to a scaling, which follows from the fact that the algebra \mathcal{A}_r has a one-parameter family of outer automorphisms $x \rightarrow x + c$, which just rescales tr by e^c . This rescaling has the physical interpretation of shifting the renormalization constant M_0 to $M_0 - c$. Now, using the trace, we can define the density matrix for the algebra \mathcal{A}_r . For any given state $|\hat{\Phi}\rangle \in \mathcal{H}$, the density matrix $\rho_{\hat{\Phi}} \in \mathcal{A}_r$ is defined by

$$\text{tr}[\rho_{\hat{\Phi}} \hat{a}] = \langle \hat{\Phi} | \hat{a} | \hat{\Phi} \rangle \quad \forall \hat{a} \in \mathcal{A}_r \quad (3.10)$$

Moreover, $\rho_{\hat{\Phi}}$ exists and is unique, which follows from the fact that the trace is nondegenerate. If we are able to define a density matrix, then we can also define the entropy of the algebra, by

$$S(\hat{\Phi})_{\mathcal{A}_r} = -\text{tr}[\rho_{\hat{\Phi}} \log \rho_{\hat{\Phi}}] = -\langle \hat{\Phi} | \log \rho_{\hat{\Phi}} | \hat{\Phi} \rangle. \quad (3.11)$$

As emphasized by CPW, $S(\hat{\Phi})$ should not be thought of as the entanglement entropy of \mathcal{A}_r but is a renormalized entropy. Also, because of the ambiguity in the definition of trace (trace is defined up to a scaling), it is only entropy differences that are unambiguous, not the entropy itself. CPW now work with a semiclassical state, i.e., the state with fluctuation in timeshift p , $\Delta p \sim O(\varepsilon)$, where ε is some parameter much smaller than unity. Then, with $x = h_L$, $\Delta x \sim O(\frac{1}{\varepsilon})$. CPW consider the AdS-Schwarzschild black hole and write down the semiclassical state in the *boundary* CFT. However, it could have equally been defined in the bulk and we will assume that the formulas below correspond to the equivalent bulk statement. The general form of such a state is,

$$|\hat{\Phi}\rangle = \int_{-\infty}^{\infty} \varepsilon^{\frac{1}{2}} g(\varepsilon x) |\Phi\rangle |x\rangle \quad \text{where } |\Phi\rangle \in \mathcal{H}, \quad g(x) \in L^2(\mathbb{R}) \quad (3.12)$$

It is shown by CPW that the density matrix $\rho_{\hat{\Phi}}$ for the state $|\hat{\Phi}\rangle$ is approximately¹¹ given by

$$\rho_{\hat{\Phi}} \approx \varepsilon \bar{g}(\varepsilon h_R) e^{-\beta x} \Delta_{\Phi|\Psi} g(\varepsilon h_R) \quad (3.13)$$

¹¹The expression is only valid up to corrections suppressed by $O(\varepsilon)$ terms.

where $\Delta_{\Phi|\Psi} = e^{-h_{\Psi|\Phi}}$ is a relative modular operator and $h_{\Psi|\Phi}$ is the relative modular Hamiltonian.¹² The relative modular Hamiltonian $h_{\Phi|\Psi}$ is defined such that $h_{\Psi|\Psi} = h_{\Psi}$. As we have already mentioned, the type II algebra modular operator factorizes, i.e., $\Delta_{\hat{\Phi}|\hat{\Psi}} = \rho_{\hat{\Psi}} \rho'_{\hat{\Phi}}^{-1}$ (where prime denotes the element of the commutant of the algebra \mathcal{A}_r) [21]. Putting (3.13) in (3.11) yields

$$S(\hat{\Phi})_{\mathcal{A}_r} = \langle \hat{\Phi} | \beta h_R | \hat{\Phi} \rangle - \langle \hat{\Phi} | h_{\Psi|\Phi} | \hat{\Phi} \rangle - \langle \hat{\Phi} | \log(\varepsilon |g(\varepsilon h_R)|^2) | \hat{\Phi} \rangle + O(\varepsilon) \quad (3.14)$$

By definition, the second term in the above equation is the relative entropy,

$$S_{\text{rel}}(\Phi|\Psi) = -\langle \hat{\Phi} | h_{\Psi|\Phi} | \hat{\Phi} \rangle \quad (3.15)$$

B. Generalization to higher curvature gravity

We now generalize the construction of the previous subsection to an arbitrary diffeomorphism invariant theory of gravity. We note that some of the constructions in the previous sub-section such as the semiclassical state were done by CPW for the AdS-Schwarzschild black hole in the *boundary* CFT. But we can analogously define such a semiclassical state in the bulk using the same construction. In fact, only in the sections on monotonicity of generalized entropy in CPW, are results in the boundary theory crucially used. Therefore, except while discussing monotonicity of the generalized entropy, we can confine our analysis to the bulk, and we can even work with an asymptotically flat black hole, as discussed by CPLW. So, in our case, M is the $(3+1)$ dimensional, asymptotically flat, maximally extended static (therefore stationary) black hole solution in an arbitrary diffeomorphism invariant theory of gravity. Therefore, its horizon is a Killing horizon. Now, the expectation value of the energy-momentum tensor of the quantum fields $\langle \Phi | T_{\mu\nu} | \Phi \rangle$ is covariantly conserved as a consequence of invariance of the action under background diffeomorphisms [see appendix]. Equation (3.1) will thus define a conserved quantity even in the arbitrary theory of gravity if V^μ is a Killing vector. Let us define the 1-form $J = T_{\mu\nu} V^\nu dx^\mu$ where V is the timelike Killing vector of M . Then, divergence of J_μ is zero, i.e., J^μ is covariantly conserved. This implies $d * J = 0$, where $*$ is the Hodge dual. This further implies $*J = dQ$ for some 2 form Q [34]. Also, notice that the integral of $*J$ over the 3 dimensional Cauchy surface is \hat{h} . Since $*J = dQ$, this reduces to an integral over the codimension 2 surface which is the boundary of the Cauchy slice. Therefore we can write $\hat{h} = H_R - H_L$ where H_R and H_L are codimension 2 integrals at right and left

¹² Ψ is the Hartle Hawking state and Φ is any arbitrary state of quantum fields in the black hole spacetime.

spatial infinity respectively. We note that the canonical energy \mathcal{E} in the covariant phase space formalism is given by (3.1) apart from a surface term ambiguity (see appendix of IW [13]).

$$\mathcal{E} = \int_{\Sigma} J + \text{Surface term} = \int_{\Sigma} d\Sigma^{\mu} T_{\mu\nu} V^{\nu} + \text{Surface term} \quad (3.16)$$

So we can think of H_R and H_L as being the right and left canonical energy, respectively, apart from ambiguities in the canonical energy.¹³ We will call them right and left Hamiltonian. Although the results of IW are only for classical gravity and matter, and we are interested in quantum fields, the IW equations can be understood as expectation values in semiclassical gravity as discussed in appendix. Now, \hat{h} is the modular Hamiltonian corresponding to the Hartle-Hawking state as before. This follows from the analysis of Sewell [33] for any metric of the following form:

$$ds^2 = A(t^2 - w^2, y)(-dt^2 + dw^2) + B(t^2 - w^2, y)d\sigma^2(y). \quad (3.17)$$

The Schwarzschild spacetime in Kruskal coordinates is of this form. We will assume that our static black hole solution also has this form (i.e., we assume the existence of Kruskal-like coordinates).

Now we can proceed by defining h_L and h_R as defined in (3.2). Following the argument in the previous section that including gravity changes the algebra to type II, we can split \hat{h} as done in (3.3) and (3.4). Further, we can straightforwardly obtain the equation in (3.5) and (3.6). The only difference is now h_R in (3.6) is the renormalized Hamiltonian in the higher curvature theory which generates the time translation on the boundary of the right exterior region. Afterward, the construction of the crossed-product algebra (extended algebra) and other constructions like defining the trace and entropy will analogously go through as done in the previous section. We will work with a semiclassical state as defined in (3.12). Therefore we can define the entropy of the algebra in the right exterior region by the same formula (3.14), i.e.,

$$S(\hat{\Phi})_{\mathcal{A}_r} = \langle \hat{\Phi} | \beta h_R | \hat{\Phi} \rangle - \langle \hat{\Phi} | h_{\Psi|\Phi} | \hat{\Phi} \rangle - \langle \hat{\Phi} | \log(\varepsilon |g(\varepsilon h_R)|^2) | \hat{\Phi} \rangle + O(\varepsilon) \quad (3.18)$$

where $S_{\text{rel}}(\Phi|\Psi) = -\langle \hat{\Phi} | h_{\Psi|\Phi} | \hat{\Phi} \rangle$ is relative entropy as defined earlier. So the form of the Eq. (3.14) remains

¹³As discussed in the Appendix of IW [13], the ambiguity in the canonical energy is the sum of two surface terms, one of which vanishes for common matter theories in a background spacetime. There is an ambiguity due to the second surface term which is a function of the background metric, Killing vector, matter fields and their derivatives. In what follows, we ignore this ambiguity.

intact—the only change is that h_R is the renormalized canonical energy in higher curvature theory and $h_{\Phi|\Psi}$ is the relative modular Hamiltonian in that particular theory.

In our case of interest, the black hole settles down to a stationary state at very late times. This is plausible since at late times all the flux of matter would either have crossed the horizon or would have escaped through future null infinity. So, at late times we will not be able to distinguish between $|\Psi\rangle$ and $|\Phi\rangle$. We get

$$S_{\text{bulk}}(\infty)_{\Phi} = S_{\text{bulk}}(\infty)_{\Psi} = S_{\text{bulk}}(b)_{\Psi} \quad (3.19)$$

Now let us analyze h_r , using (3.3) and the fact that the deformation of Cauchy surfaces S does not affect the conserved quantity \hat{h} . We deform S_1 such that $S'_1 = \mathcal{H}^+ \cup \mathcal{I}^+$, where \mathcal{H}^+ is the future horizon and \mathcal{I}^+ is future null infinity.¹⁴ Therefore

$$\beta h_r^{\mathcal{H}^+} = \beta(h_r - h_r^{\mathcal{I}^+}) = \int_0^{\infty} dv \int_{\mathcal{H}^+} d^{D-2}x \sqrt{h} v T_{vv} \quad (3.20)$$

where $\beta h_r^{\mathcal{I}^+}$ is the time translation generator at future null infinity and $\beta h_r^{\mathcal{H}^+}$ is the boost generator on the horizon. The second equality in the above equation can be obtained using (3.3) and the fact that $h_r^{\mathcal{I}^+}$ is just the integral of the stress tensor supported at future null infinity. Let us define a one-sided modular operator (boost operator) at arbitrary cut $v = v_*$ (which is the $D - 2$ dimensional transverse surface) at the horizon. It is well known that the density matrix $(\rho_r)_{\text{HH}}$ of the Hartle-Hawking state in the region r is thermal with respect to [11,32]

$$K_r(v_*) = \int_{v_*}^{\infty} dv \int_{\mathcal{H}^+} d^{D-2}x \sqrt{h}(v - v_*) T_{vv} + K_r^{\mathcal{I}^+} \quad (3.21)$$

where, $K_r^{\mathcal{I}^+} = \beta h_r^{\mathcal{I}^+}$ is the modular energy at \mathcal{I}^+ , which accounts for energy which goes to \mathcal{I}^+ without crossing the horizon. Also notice that $K_r^{\mathcal{I}^+}$ is independent of v_* . When $v_* = 0$, then the first term in the above equation will become $\beta h_r^{\mathcal{H}^+}$ as defined in (3.20), and therefore

$$K_r(b) = \beta h_r = \beta h_r^{\mathcal{H}^+} + \beta h_r^{\mathcal{I}^+} \quad (3.22)$$

where b is the bifurcation surface $v_* = 0$. We can also define

$$K_r(\infty) = \lim_{v_* \rightarrow \infty} K_r(v_*) \quad (3.23)$$

Using the result from the previous section and Eq. (3.20), we get

¹⁴In the case of an asymptotically AdS black hole, the deformed Cauchy surface is just \mathcal{H}^+ .

$$\langle \Phi | \beta h_r^{\mathcal{H}^+} | \Phi \rangle = \Delta \delta S \quad (3.24)$$

If a density matrix were to exist for the algebra $\mathcal{A}_{r,0}$, then using the definition of modular Hamiltonian for state $|\Psi\rangle$ and the fact that $\Delta_\Psi = \rho_\Psi \rho_\Psi'^{-1}$, we will be able to write

$$\log \rho_\Psi = -K_r(b) + C \quad (3.25)$$

where C is some constant. The density matrix of the Hartle-Hawking state in region r can be written as $\Psi(\mathcal{H}^+ \cup \mathcal{I}^+) = (\rho_r)_{\text{HH}} \otimes \sigma$, which corresponds to the ground state¹⁵ ρ_{HH} at \mathcal{H}^+ , product taken with some arbitrary density matrix defining a faithful state at \mathcal{I}^+ [11]. Therefore,

$$\langle \Psi | \log \rho_\Psi | \Psi \rangle = -\langle \Psi | K_r^{\mathcal{I}^+} | \Psi \rangle + C \quad (3.26)$$

Here we use the fact that $\langle \Psi | h_r^{\mathcal{H}^+} | \Psi \rangle = 0$ since it is the Hartle-Hawking state. Further, since $S_{\text{bulk}}(b)_\Psi = -\langle \Psi | \log \rho_\Psi | \Psi \rangle$, we get $S_{\text{bulk}}(b)_\Psi = \langle \Psi | K_r^{\mathcal{I}^+} | \Psi \rangle - C$.

As mentioned before, it is not strictly true of the algebra $\mathcal{A}_{r,0}$ that the modular operator factorizes, but by extending the algebra to \mathcal{A}_r , it is true that the modular operator factorizes as $\hat{\Delta}_\Psi = \rho_\Psi \rho_\Psi'^{-1}$ (in the notation of [20]). We will ignore this detail just for illustrative purposes following [16].

$$S_{\text{gen}}(\infty) - S_{\text{gen}}(b) = S(\infty) - S(b) + S_{\text{bulk}}(b)_\Psi - S_{\text{bulk}}(b)_\Phi. \quad (3.27)$$

It can be written in terms of the one-sided modular operator,

$$S_{\text{gen}}(\infty) - S_{\text{gen}}(b) = -\langle \Phi | (K_r(\infty) - K_r(b)) | \Phi \rangle + S_{\text{bulk}}(b)_\Psi - S_{\text{bulk}}(b)_\Phi. \quad (3.28)$$

Putting $S_{\text{bulk}}(b)_\Psi$ and $\langle \Phi | K_r(b) | \Phi \rangle$ using the Eq. (3.25), we get

$$S_{\text{gen}}(\infty) - S_{\text{gen}}(b) = -\langle \Phi | \log \rho_\Psi | \Phi \rangle + C - \langle \Phi | K_r(\infty) | \Phi \rangle + \langle \Psi | K_r^{\mathcal{I}^+} | \Psi \rangle - C - S_{\text{bulk}}(b)_\Phi \quad (3.29)$$

Now, we use the fact that at late times, every state is indistinguishable from $|\Psi\rangle$ and $K_r(\infty) = K_r^{\mathcal{I}^+}$. Further, $K_r^{\mathcal{I}^+}$ is independent of the cut. The expectation value $K_r^{\mathcal{I}^+}$ in state $|\Phi\rangle$ will be equal to its expectation value in state $|\Psi\rangle$. Therefore, we get

$$S_{\text{gen}}(\infty) - S_{\text{gen}}(b) = -\langle \Phi | \log \rho_\Psi | \Phi \rangle - S_{\text{bulk}}(b)_\Phi \quad (3.30)$$

¹⁵Hartle-Hawking state is a ground state with respect to the time v .

therefore we got,

$$S_{\text{gen}}(\infty) - S_{\text{gen}}(b) = S_{\text{rel}}(\Phi || \Psi) \quad (3.31)$$

As we see, the difference of generalized entropies in (3.31) is manifestly finite and non-negative. For Einstein gravity, the above expression has been already obtained by Wall in [11]. The result (3.31) is in an arbitrary theory of gravity—the difference between generalized entropy at late times and generalized entropy at the bifurcation surface is relative entropy of the state of the black hole with respect to the Hartle-Hawking state. We now need to show, as in [16], that the generalized entropy at the bifurcation surface is the entropy of the algebra. We thus need to show

$$S_{\text{gen}}(\infty) = \langle \hat{\Phi} | \beta h_R | \hat{\Phi} \rangle - \langle \hat{\Phi} | \log(\epsilon | g(\epsilon h_R) |^2) | \hat{\Phi} \rangle + \text{Const} \quad (3.32)$$

Since both terms in the above equation are only functions of h_R , and since we have interpreted h_R as the renormalized canonical energy, the above terms are some distributions of energy in the semiclassical state $|\hat{\Phi}\rangle$. Also, these terms are independent of the state $|\Phi\rangle$. To see that, choose $a(s)$ such that

$$a(s) = \int e^{-ih'_R s} f(h'_R) dh'_R \quad (3.33)$$

where $f(h_R)$ is some chosen function. Putting the Eq. (3.33) in the equation (3.7) and using the fact that $h_R = \hat{h}_\Psi + x$, will yield $\hat{a} = f(h_R)$. Now let us compute the expectation value $\langle \hat{\Phi} | \beta \hat{a} | \hat{\Phi} \rangle$ for (3.33) with the semiclassical state $|\hat{\Phi}\rangle$ defined in (3.12). Using the results (3.25) and (3.26) in CPW [16],

$$\langle \hat{\Phi} | \hat{a} | \hat{\Phi} \rangle = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} ds | \epsilon g(\epsilon x) |^2 e^{isx} \langle \Psi | \Delta_{\Phi|\Psi} a(s) | \Psi \rangle \quad (3.34)$$

Now put (3.33) in (3.34). Using the fact that $h_R = \hat{h}_\Psi + x$ and $\hat{h}_\Psi |\Psi\rangle = 0$, we get

$$\langle \hat{\Phi} | f(h_R) | \hat{\Phi} \rangle = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} ds \int_{-\infty}^{\infty} dy | \epsilon g(\epsilon x) |^2 \times e^{is(x-y)} f(y) \langle \Psi | \Delta_{\Phi|\Psi} | \Psi \rangle \quad (3.35)$$

By definition, $\langle \Psi | \Delta_{\Phi|\Psi} | \Psi \rangle = 1$. Therefore the above equation is independent of $|\Phi\rangle$, it will only depend on $f(h_R)$ and $g(\epsilon x)$. Therefore, both the terms on the right-hand side of (3.32) will give the same result either when we compute them in the state $|\Phi\rangle$ or in the Hartle-Hawking state $|\Psi\rangle$ at late times. Since both terms can be determined from the

late-time behavior of the black holes, the relation (3.32) is plausible. This is because at late times, all the fields have either fallen across the horizon or to infinity.

Using Eq. (3.2), we can write $\langle \hat{\Psi} | \beta h_R | \hat{\Psi} \rangle = \beta(\Delta\mathcal{E})$, where $\Delta\mathcal{E}$ is the energy difference between the black hole we are studying and the reference black hole. Since both the black holes are taken in an equilibrium state, we can apply the first law of black hole mechanics for two equilibrium configurations in phase space which yields

$$\langle \hat{\Psi} | \beta h_R | \hat{\Psi} \rangle = \Delta S \quad (3.36)$$

where ΔS is the difference of entropy of the equilibrium black hole state we get at late times to the reference black hole. Therefore, the first term in (3.32) describes the change in $S_{\text{gen}}(\infty)$ due to a change in black hole entropy. At very late times, all the matter would have either crossed the horizon or would have escaped to null infinity. The second term should be thought of as the contribution of entropy of fluctuations in black hole entropy [16]. Finally, we add and subtract the entanglement entropy of the quantum fields in the Hartle-Hawking state and lump one of the pieces in the constant in (3.32) using (3.26). This is because at late times, all the fields have either fallen across the horizon or to infinity.

Combining everything, we get

$$S(\hat{\Phi})_{\mathcal{A}_r} = S_{\text{gen}}(b) + \text{Const.} \quad (3.37)$$

The equation (3.37) tells us that $S_{\text{gen}}(b)$ for black holes at the bifurcation surface in the arbitrary theory of gravity can be thought of as the entropy of the algebra \mathcal{A}_r modulo a constant. Notice, we have shown that the generalized entropy at the bifurcation surface is equal to the entropy of the algebra up to a constant, but we are not making any statement about entropy at an arbitrary cut of the horizon. In algebraic QFT, relative entropy is positive. This implies $S_{\text{gen}}(\infty) - S_{\text{gen}}(b) \geq 0$. Can we go beyond this result and prove that the generalized entropy is monotonic? The entropy of the algebra is monotonic under trace-preserving inclusions [35]. To obtain a GSL (monotonicity of the generalized entropy), CPW consider an AdS Schwarzschild black hole with a holographic dual CFT. Then, they have the following clever argument: In the dual CFT, they consider operator algebras at two different times (early and late times), $\mathcal{A}_{R,0}$ and $\mathcal{B}_{R,0}$ respectively, separated by a timescale much larger than the thermal timescale β . The correlation functions of operators at these different times factorize into a product of correlators of early and late times. Thus, the algebra generated by both early and late time operators is $\mathcal{C}_{R,0} = \mathcal{A}_{R,0} \otimes \mathcal{B}_{R,0}$. The Hilbert space factorizes similarly. These algebras are extended similar to what was done before, to obtain a type II algebra which has an associated entropy. Now, consider three different situations: first, the quantum fields at both early and late times

are in an arbitrary state, a second situation where the fields at early times have fallen into the horizon, so that the state of these fields is the vacuum times any state of the late time fields, and finally, a situation where both sets of fields have fallen into the horizon and the state of the fields in the exterior is the vacuum. CPW argue that the generalized entropies at these three horizon cuts is the generalized entropy for the extended algebra \mathcal{C}_R for these three different states at a hypothetical *bifurcation surface* in the limit of large time gap. From the property of the monotonicity of the entropy of the algebra under trace preserving inclusions, it follows that the generalized entropy is increasing in going from the first to the third situation above. This argument is then a version of the GSL. We can use these results to prove this version of the GSL in an arbitrary diffeomorphism invariant theory of gravity if we start with an asymptotically AdS black hole which has a holographic dual. We can repeat all the steps in this section for such a black hole. The only thing we need is for the first law as in the paper of Iyer and Wald [13] to be true in this situation. Although the statement of the first law is only for asymptotically flat black holes, the same will be true for an asymptotically AdS black hole provided the integrals involved in the presymplectic form and the canonical energy are finite after assuming appropriate fall-offs for the fields. In this case, the computation of CPW generalizes to these black holes in a higher curvature theory of gravity and a version of increase of generalized entropy (comparing the entropy at early and late times) is true. The ambiguity in the Wald entropy which we fixed in the bifurcation surface in a previous section was by terms quadratic in the half-order perturbation. This is not affected by the different states of the quantum fields in the argument of CPW, so their argument goes over to our case. Can we show a stronger monotonicity result $\frac{dS_{\text{gen}}}{dv} \geq 0$? This is what Wall [11] has done for Einstein gravity, using an expression for the entropy at any cut of the horizon. Due to JKM ambiguities in the Wald entropy, this expression will probably need one to specify the particular theory of gravity.

IV. DISCUSSION

In the context of Einstein gravity, it was shown by CPW [16] that for the system of quantum fields in a perturbed Schwarzschild black hole spacetime in the $G \rightarrow 0$ limit with infalling quantum matter across the horizon, the generalized entropy at the bifurcation surface was equal to the entropy of the von Neumann algebra of operators in the black hole exterior. This was achieved by enlarging the operator algebra by including the ADM Hamiltonian, and by enlarging the Hilbert space. This had the effect of changing the algebra of operators to a type II_∞ von Neumann algebra, to which we can associate a notion of trace and entropy. Furthermore, CPW showed that

the difference of the generalized entropy of an arbitrary cut of the horizon, in the limit when the cut $v \rightarrow \infty$ and the generalized entropy at the bifurcation surface was equal to the relative entropy, and therefore nonnegative. For this, CPW worked in semiclassical gravity and up to quadratic order in perturbations. They then obtained a monotonicity result (GSL) for the generalized entropy from the monotonicity of relative entropy under trace preserving inclusions. In this paper, we consider quantum fields in a slightly perturbed static black hole in an arbitrary diffeomorphism invariant theory of gravity in the $G \rightarrow 0$ limit. In this setup, generalized entropy is the sum of Wald entropy and entanglement entropy of quantum fields in the black hole exterior. We consider the difference in Wald entropy at infinity and at the bifurcation surface up to quadratic order in the perturbations and obtain (2.59). Wald entropy has ambiguities on nonstationary geometries. We fix the ambiguity in order to get (2.60) which matches the result for Einstein gravity in the paper of CPW and we obtain a simplified result for the difference of entropies which enables us to employ the CPW construction. We then consider the difference in *generalized* entropies and show that this difference equals the relative entropy of the state of the quantum fields and the Hartle-Hawking state—it is thus non-negative. We next consider the von Neumann algebra of the quantum fields in the black hole exterior, extended to include the Hamiltonian, and an enlarged Hilbert space as in CPW. Evaluated on the semiclassical states defined by CPW, we show that the entropy of the von Neumann algebra equals the generalized entropy of the bifurcation surface. Finally, we see that the derivation of the increase of generalized entropy by CPW in Einstein gravity goes through for black holes in an arbitrary gravity theory, provided the black hole is asymptotically AdS, which has a holographic dual.

Going forward, it would be interesting to not merely restrict to semiclassical states and see what the entropy of the algebra corresponding to more general states is. Recently, there has also been an interesting proposal to study the algebra of operators in general subregions in Einstein gravity with matter in the $G \rightarrow 0$ limit, and it has been shown that the entropy of this algebra equals the generalized entropy of this subregion up to a constant [36] (see also the latest papers [37,38]). It would be interesting to see how these results generalize to an arbitrary diffeomorphism invariant theory of gravity.

ACKNOWLEDGMENTS

We would like to thank Sanved Kolekar for useful comments. M. A. also acknowledges the Council of Scientific and Industrial Research (CSIR), Government of India for financial assistance.

APPENDIX

1. Quantum canonical energy in covariant phase space formalism

Consider quantum fields in the stationary black hole background, where the quantum fields include gravitons. Using diffeomorphism invariance of the Lagrangian (which is an n -form) with respect to the background metric, we can write

$$\frac{\delta L}{\delta \phi} \mathcal{L}_\xi \phi = -d\tilde{J} - \frac{\delta L}{\delta g_{ab}} \mathcal{L}_\xi g_{ab}. \quad (\text{A1})$$

ϕ here corresponds to all quantum fields and g_{ab} is the background metric. ξ is a vector field generating the diffeomorphism and \mathcal{L}_ξ denotes the Lie derivative with respect to ξ . Here, \tilde{J} is the Noether current ($n-1$) form in the covariant phase space formalism given by Eq. (49) in IW [13]. This is defined in terms of the symplectic potential form $\Theta(\delta\phi, \phi)$ as

$$\tilde{J} = \Theta(\phi, \mathcal{L}_\xi \phi) - \xi \cdot L. \quad (\text{A2})$$

As we have a Lagrangian for quantum fields to begin with, we look at the expectation value of the relation (A1) in an arbitrary state. Since the terms in the equations are products of operators at the same point, the expectation values need to be regularized using some procedure. Using a prescription involving point-splitting and background subtraction, we regularize these expectation values.

We want to compute the expectation value of (A1) in some state $|\Phi\rangle$. Consider the point-split expression

$$\begin{aligned} & \lim_{y \rightarrow x} \langle \Phi | \frac{\delta L}{\delta \phi}(x) \mathcal{L}_\xi \phi(y) | \Phi \rangle \\ &= \lim_{y \rightarrow x} \langle \Phi | \left(-d\tilde{J}(x, y) - \frac{\delta L}{\delta g_{ab}}(x) \mathcal{L}_\xi g_{ab}(y) \right) | \Phi \rangle \end{aligned} \quad (\text{A3})$$

From the Schwinger-Dyson equation for expectation value in an arbitrary state, we can compute the left hand side, and in the limit of coincident points, it is just a state independent divergent term. Therefore if we consider the difference of the quantity $(-d\tilde{J}(x, y) - \frac{\delta L}{\delta g_{ab}}(x) \mathcal{L}_\xi g_{ab}(y))$ in any two states $|\Phi\rangle$ and $|\Psi\rangle$, then the state independent divergent term will cancel out. Now we can take the coincidence limit $y \rightarrow x$.

$$\begin{aligned} & \langle \Phi | \left(-d\tilde{J}(x) - \frac{\delta L}{\delta g_{ab}}(x) \mathcal{L}_\xi g_{ab}(x) \right) | \Phi \rangle \\ & - \langle \Psi | \left(-d\tilde{J}(x) - \frac{\delta L}{\delta g_{ab}}(x) \mathcal{L}_\xi g_{ab}(x) \right) | \Psi \rangle = 0 \end{aligned} \quad (\text{A4})$$

Now, (A4) can be written as

$$\begin{aligned} d\langle\Phi|(\tilde{J} + k.\epsilon)|\Phi\rangle - d\langle\Psi|(\tilde{J} + k.\epsilon)|\Psi\rangle \\ = \nabla_\mu(\langle\Phi|T^{\mu\nu}|\Phi\rangle)\xi_\nu - \nabla_\mu(\langle\Psi|T^{\mu\nu}|\Psi\rangle)\xi_\nu. \end{aligned} \quad (\text{A5})$$

Here, we use the fact that $\frac{\delta L}{\delta g_{\mu\nu}} = T^{\mu\nu}\epsilon$, where ϵ is the volume form and $k^\mu = T^{\mu\nu}\xi_\nu$. Notice that the left-hand side in the above equation is a total derivative for any ξ while the right-hand side is not. The only way that can happen is if $\nabla_\mu(\langle\Phi|T^{\mu\nu}|\Phi\rangle)\xi_\nu = \nabla_\mu(\langle\Psi|T^{\mu\nu}|\Psi\rangle)\xi_\nu$. Since this has to be true for any two states and any vector field ξ^μ , it can only be if $\nabla_\mu(\langle\Phi|T^{\mu\nu}|\Phi\rangle)$ vanishes for any state $|\Phi\rangle$ up to a local term independent of state. We may modify our prescription and get rid of this extra state independent term by doing background subtraction with respect to some standard state. We do the background subtraction (A4) to make sense of

expectation values of Eq. (A1). Thus the expectation value of the stress energy tensor in any state is conserved. It has already been extensively discussed in Wald [39] how to define the expectation value of the stress energy tensor in such a way that it is conserved, so alternatively, we could use that result. By the same argument, $d\langle\Phi|(\tilde{J} + k.\epsilon)|\Phi\rangle = 0$. The rest of the analysis of the Appendix of IW [13] follows. Now we can choose ξ^μ to be the Killing field of the static background spacetime. Following IW [13], we will get

$$\langle\mathcal{E}\rangle_\Phi = \int_\Sigma d\Sigma^\mu \langle T_{\mu\nu}\rangle_\Phi \xi^\nu + \text{Surface term} \quad (\text{A6})$$

where \mathcal{E} is the canonical energy in the covariant phase space formalism and it is independent of the choice of Cauchy slice.

-
- [1] J. D. Bekenstein, Black holes and entropy, *Phys. Rev. D* **7**, 2333 (1973).
- [2] J. D. Bekenstein, Black holes and the second law, *Lett. Nuovo Cimento* **4**, 737 (1972).
- [3] L. Susskind and J. Uglum, Black hole entropy in canonical quantum gravity and superstring theory, *Phys. Rev. D* **50**, 2700 (1994).
- [4] T. Jacobson, Black hole entropy and induced gravity, [arXiv:gr-qc/9404039](https://arxiv.org/abs/gr-qc/9404039).
- [5] F. Larsen and F. Wilczek, Renormalization of black hole entropy and of the gravitational coupling constant, *Nucl. Phys.* **B458**, 249 (1996).
- [6] V. P. Frolov, D. Fursaev, and A. Zelnikov, Statistical origin of black hole entropy in induced gravity, *Nucl. Phys.* **B486**, 339 (1997).
- [7] S. N. Solodukhin, Entanglement entropy of black holes, *Living Rev. Relativity* **14**, 8 (2011).
- [8] J. H. Cooperman and M. A. Luty, Renormalization of entanglement entropy and the gravitational effective action, *J. High Energy Phys.* **12** (2014) 045.
- [9] R. Bousso, Z. Fisher, S. Leichenauer, and A. C. Wall, Quantum focusing conjecture, *Phys. Rev. D* **93**, 064044 (2016).
- [10] E. Gesteau, Large N von Neumann algebras and the renormalization of Newton's constant, [arXiv:2302.01938](https://arxiv.org/abs/2302.01938).
- [11] Aron Wall, A proof of the generalized second law for rapidly changing fields and arbitrary horizon slices, *Phys. Rev. D* **85**, 104049 (2012).
- [12] R. M. Wald, Black hole entropy is Noether charge, *Phys. Rev. D* **48**, R3427 (1993).
- [13] V. Iyer and R. M. Wald, Some properties of Noether charge and a proposal for dynamical black hole entropy, *Phys. Rev. D* **50**, 846 (1994).
- [14] T. Jacobson, G. Kang, and R. C. Myers, On black hole entropy, *Phys. Rev. D* **49**, 6587 (1994).
- [15] S. Sarkar and A. C. Wall, Generalized second law at linear order for actions that are functions of Lovelock densities, *Phys. Rev. D* **88**, 044017 (2013).
- [16] Venkatesa Chandrasekaran, Geoff Penington, and Edward Witten, Large N algebras and generalized entropy, *J. High Energy Phys.* **04** (2023) 009.
- [17] Samuel Leutheusser and Hong Liu, Emergent times in holographic duality, [arXiv:2112.12156](https://arxiv.org/abs/2112.12156).
- [18] Samuel Leutheusser and Hong Liu, Subalgebra-subregion duality: Emergence of space and time in holography, [arXiv:2212.13266](https://arxiv.org/abs/2212.13266).
- [19] Keiichiro Furuya, Nima Lashkari, Mudassir Moosa, and Shoy Ouseph, Information loss, mixing and emergent type III₁ factors, *J. High Energy Phys.* **08** (2023) 111.
- [20] Edward Witten, Gravity and the crossed product, *J. High Energy Phys.* **10** (2022) 008.
- [21] Venkatesa Chandrasekaran, Roberto Longo, Geoff Penington, and Edward Witten, An algebra of observables for de Sitter space, *J. High Energy Phys.* **02** (2023) 082.
- [22] Edward Witten, Notes on some entanglement properties of quantum field theory, *Rev. Mod. Phys.* **90**, 045003 (2018).
- [23] A. Chatterjee and S. Sarkar, Physical Process First Law and Increase of Horizon Entropy for Black Holes in Einstein-Gauss-Bonnet Gravity, *Phys. Rev. Lett.* **108**, 091301 (2012).
- [24] A. Wall, A second law for higher curvature gravity, *Int. J. Mod. Phys. D* **24**, 1544014 (2015).
- [25] S. Bhattacharyya, P. Dhivakar, A. Dinda, N. Kundu, M. Patra, and S. Roy, An entropy current and the second law in higher derivative theories of gravity, *J. High Energy Phys.* **09** (2021) 169.
- [26] J. Bhattacharya, S. Bhattacharyya, A. Dinda, and N. Kundu, An entropy current for dynamical black holes in four-

- derivative theories of gravity, *J. High Energy Phys.* **06** (2020) 017.
- [27] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Princeton University Press, 1973).
- [28] A. Mishra, S. Chakraborty, A. Ghosh, and S. Sarkar, On the physical process first law for dynamical black holes, *J. High Energy Phys.* **09** (2018) 034.
- [29] X. Dong, Holographic entanglement entropy for general higher derivative gravity, *J. High Energy Phys.* **01** (2014) 044.
- [30] Huzihiro Araki, Von Neumann algebras of local observables for free scalar field, *J. Math. Phys. (N.Y.)* **5**, 1 (1964).
- [31] Edward Witten, Why does quantum field theory in curved spacetime make sense? and what happens to the algebra of observables in the thermodynamic limit?, [arXiv:2112.11614](https://arxiv.org/abs/2112.11614).
- [32] J. J. Bisognano and E. H. Wichmann, On the duality condition for quantum fields, *J. Math. Phys. (N.Y.)* **17**, 303 (1976).
- [33] Geoffrey L Sewell, Quantum fields on manifolds: PCT and gravitationally induced thermal states, *Ann. Phys. (N.Y.)* **141**, 201 (1982).
- [34] G. Compere and A. Fiorucci, Advanced lectures on general relativity, [arXiv:1801.07064](https://arxiv.org/abs/1801.07064).
- [35] Roberto Longo and Edward Witten, A note on continuous entropy, [arXiv:2202.03357](https://arxiv.org/abs/2202.03357).
- [36] Kristan Jensen, Jonathan Sorce, and Antony Speranza, Generalized entropy for general subregions in quantum gravity, [arXiv:2306.01837](https://arxiv.org/abs/2306.01837).
- [37] Shadi Ali Ahmad and Ro Jefferson, Crossed product algebras and generalized entropy for subregions, [arXiv:2306.07323](https://arxiv.org/abs/2306.07323).
- [38] Marc S. Klinger and Robert G. Leigh, Crossed products, extended phase spaces and the resolution of entanglement singularities, [arXiv:2306.09314](https://arxiv.org/abs/2306.09314).
- [39] R. Wald, *Quantum Field theory in Curved Spacetime and Black Hole Thermodynamics*, Chicago Lectures in Physics (University of Chicago Press, 1994).