Novel aspects of the extended first law of entanglement

Felipe Rosso[a] **and Andrew Svesko**[b]

[a]*Department of Physics and Astronomy, University of Southern California,*
Los Angeles, CA 90089-0484, U.S.A.

[b]*Department of Physics, Arizona State University,*
Tempe, Arizona 85287, U.S.A.

E-mail: felipero@usc.edu, asvesko@asu.edu

ABSTRACT: Using AdS/CFT an extended first law of entanglement has been previously derived for the vacuum reduced to a ball in Minkowski. The statement not only includes perturbations of the state but also of the conformal field theory (CFT), via variations of the generalized central charge. We clarify some subtleties previously overlooked and use simple arguments to generalize prior derivations to arbitrary gravity theories in the bulk as well as new regions in the boundary CFT. Our construction also applies to two-dimensional bulk theories and admits an interesting extension for a three-dimensional bulk, providing a curious result regarding the thermodynamic volume in extended black hole thermodynamics. We discuss future prospects regarding the extended first law of entanglement.

JHEP08(2020)008

Contents

1 Introduction

The first law of entanglement is a natural generalization of the first law of thermodynamics that applies to non-equilibrium states. As first shown in refs. [1, 2], it is a consequence of positivity of relative entropy, and determines the first order variation of entanglement entropy under state perturbations. Its most interesting application is arguably given in refs. [3, 4], where it plays a crucial role in deriving the bulk linearized Einstein's equations about a perturbed AdS background from boundary entanglement correlations of the CFT.

Motivated by extended black hole thermodynamics [5–7], where the cosmological constant Λ is interpreted as a thermodynamic pressure $p \equiv -\Lambda/8\pi G$, an extension of the first law of entanglement was proposed in ref. [8], which includes not only variations of the state but also of the CFT itself. It can be written as

$$\delta S_{EE} = \delta\langle K_B \rangle + \frac{S_{EE}}{a_d^*}\delta a_d^*, \tag{1.1}$$

where S_{EE} is the vacuum entanglement entropy associated to a ball in Minkowski and K_B its modular hamiltonian. The constant a_d^* is defined for an arbitrary CFT as

$$a_d^* = \begin{cases} A_d\,, & \text{for d even} \\ (-1)^{\frac{d-1}{2}}\ln[Z(S^d)]/2\pi\,, & \text{for d odd}\,. \end{cases} \tag{1.2}$$

Here A_d is the coefficient in the trace anomaly proportional to Euler's density, while for odd dimensions a_d^* is determined by the partition function of the CFT placed on a unit

sphere S^d (see ref. [9] for some examples in free theories). Since a_d^* has a monotonous behavior under renormalization group flows [10], we can interpret it as counting the number of degrees of freedom in the CFT. The generalized central charge a_d^* has appeared in a number of holographic c-theorems in arbitrary dimensions and higher curvature theories of gravity [11].

The first term in (1.1) is the ordinary contribution to the first law obtained by perturbing the state, while the second gives the behavior of the entanglement entropy when varying the CFT. We must emphasize that this second contribution is *not* equivalent to a renormalization group flow, since the variation continuously interpolates between CFTs. It simply gives the dependence of the entanglement entropy on the CFT data.

The extended first law (1.1) was initially derived in ref. [8] for a holographic CFT dual to Einstein gravity, and later generalized to specific higher curvature gravity theories in refs. [12–14]. These derivations start by considering a particular Killing horizon in pure AdS and deriving an extended bulk first law which considers variations of the cosmological constant, using either Hamiltonian perturbation theory [12] or the Iyer-Wald formalism [13]. The horizon entropy associated to this Killing horizon is then identified as the entanglement entropy of the boundary CFT, while the variation of the cosmological constant maps to changing the generalized central charge a_d^*.

Given the importance and wide range of applications of the first law of entanglement, we should take any reasonable generalization seriously, as it has the potential of providing new insights into the structure of space-time and entanglement in QFTs. In this work we explore the extended first law of entanglement (1.1) by generalizing previous derivations to include arbitrary theories of gravity, clarifying some of its subtle features and studying its low dimensional limit.

The outline of this article is as follows. We start in section 2 by showing that a remarkably simple argument allows us to derive the bulk analog of (1.1) for perturbations of any Killing horizon in pure AdS. Contrary to previous derivations, our computation is novel in its simplicity and the fact that it holds for arbitrary bulk gravity theories and Killing horizons in pure AdS, finding no need to resort to technical calculations as in refs. [8, 12–14]. We discuss how each of the bulk quantities is mapped to the boundary CFT, carefully analyzing some subtleties previously overlooked. Applying our construction to certain bulk Killing horizons, we derive the extended first law (1.1) for the vacuum state of a CFT reduced to the following regions: a ball and the half-space in Minkowski, a spherical cap in the Lorentzian cylinder $\mathbb{R} \times S^{d-1}$ and de Sitter, and a ball in AdS_d. The method used to find the appropriate bulk Killing horizons crucially relies on the freedom to choose conformal frames at the AdS boundary.

We continue in section 3, where we revisit the calculations from section 2 but carefully analyzing the case in which the bulk theory is two-dimensional. While for a class of two-dimensional gravity theories we find no obstructions when deriving the extended first law for Killing horizons in pure AdS_2, there are certain Einstein-dilaton theories where the end result takes a different form. We illustrate this in subsection 3.2 for Jackiw-Teitelboim gravity [15, 16], where we show the extended first law for Killing horizons is distinct.

In section 4 we show that in three dimensional gravity an extended first law can be derived for Killing horizons in space-times that are locally but not globally AdS. This

allows us to obtain an extended first law for the boundary CFT_2 that is analogous to (1.1) but involving thermal instead of entanglement entropy. From the bulk perspective we find some interesting results for extended black hole thermodynamics, where we obtain a curious formula for the thermodynamic volume (see eq. (4.8)), the conjugate variable to the pressure p.

We conclude in section 5 by expanding some discussions on the calculations in the main text. We clarify some aspects regarding the structure of divergences in the extended first law of entanglement (1.1) and critically analyze the extent to which it can hold for arbitrary regions and CFTs. We briefly comment on the bulk constraints implied by assuming both the RT holographic entropy formula [17] and the extended first law of entanglement hold for arbitrary setups in the boundary CFT. Finally, we discuss some interesting aspects of the thermodynamic volume in three dimensional gravity and its connection to the microscopic interpretation of black hole super-entropicity [18].

2 Killing horizons in pure AdS and extended first law

In this section we present a derivation of the extended first law of entanglement for holographic CFTs described by arbitrary covariant theories of gravity in the bulk

$$I[\lambda_i, g_{\mu\nu}] = \int d^{d+1}x \, \sqrt{-g} \, \mathcal{L} \left(g_{\mu\nu}, \mathcal{R}_{\mu\nu\rho\sigma}, \nabla_\lambda \mathcal{R}_{\mu\nu\rho\sigma}, \dots \right), \qquad (2.1)$$

where $\mathcal{R}_{\mu\nu\rho\sigma}$ is the Riemann tensor. Each theory is characterized by a family of coupling constants $\{\lambda_i\}$ that are chosen such that the action admits a pure AdS vacuum solution of radius L. This length scale is a non-trivial function of the coupling constants of the theory $L = L(\lambda_i)$, and the pure AdS metric only depends on $\{\lambda_i\}$ through L. A concrete and simple example of a higher curvature theory is Einstein Gauss-Bonnet gravity. Although we could also add some matter to the action, for the most part we consider pure gravity and set matter fields to zero.

Consider a Killing vector ξ^μ of the pure AdS metric $g_{\mu\nu}^{\text{AdS}}(L)$ which is time-like over some region

$$\xi^2 \equiv g_{\mu\nu}^{\text{AdS}} \xi^\mu \xi^\nu \leq 0 \qquad \Longleftrightarrow \qquad \text{Some region of AdS} . \qquad (2.2)$$

The surface in which the vector vanishes defines a Killing horizon. One of the central quantities characterizing this horizon is its entropy, that for an arbitrary theory is computed from Wald's functional according to [19, 20]

$$S_\xi \left[g_{\mu\nu}^{\text{AdS}}(L), \lambda_i \right] = -2\pi \int dV \left[\frac{\delta \mathcal{L}}{\delta \mathcal{R}^{\mu\nu}_{\ \ \rho\sigma}} n^{\mu\nu} n_{\rho\sigma} \right], \qquad (2.3)$$

where the integral is over the bifurcation Killing surface with induced volume element dV. The anti-symmetric tensor $n^{\mu\nu}$ is the binormal to the horizon normalized so that $n^{\mu\nu} n_{\mu\nu} = -2$. Our aim is to study the behavior of this entropy functional under general perturbations and to determine its consequences for the boundary CFT.

Let us start by considering the behavior of the entropy under metric perturbations $g_{\mu\nu}^{\text{AdS}}(L) \to g_{\mu\nu}^{\text{AdS}}(L) + \delta g_{\mu\nu}$.[1] Since we are working with a Killing horizon we can apply the same methods used to study black hole thermodynamics. As computed in ref. [20] in the context of black holes, the first order variation of (2.3) is determined by the Noether charge Q_ξ. This can be related to the energy E_ξ associated with the Killing vector ξ measured at asymptotic spatial infinity according to [20][2]

$$\delta S_\xi = \frac{2\pi}{\kappa} \delta E_\xi\,, \qquad \kappa^2 = -\frac{1}{2}\left(\nabla^\mu \xi^\nu\right)\left(\nabla_\mu \xi_\nu\right)\,, \tag{2.4}$$

where κ is the surface gravity.

We now consider another type of perturbation obtained by changing the gravitational theory itself, i.e. $\mathcal{L} \to \mathcal{L} + \delta\mathcal{L}$, implemented by slightly changing the coupling constants of the theory $\lambda_i \to \lambda_i + \delta\lambda_i$. Since the pure AdS metric $g_{\mu\nu}^{\text{AdS}}(L)$ is a function of λ_i through $L = L(\lambda_i)$, the perturbation induces a variation of the metric. If we did not take this metric variation into account, the perturbed metric would not be a solution of the perturbed Lagrangian. Hence, the first order variation of Wald's functional is explicitly given by

$$\delta S_\xi = S_\xi\left[g_{\mu\nu}^{\text{AdS}}(\lambda_i + \delta\lambda_i), \lambda_i + \delta\lambda_i\right] - S_\xi\left[g_{\mu\nu}^{\text{AdS}}(\lambda_i), \lambda_i\right]\,. \tag{2.5}$$

From the definition of Wald's entropy in (2.3) we can compute this in full generality, the key feature being that both terms are evaluated in the pure AdS metric of each theory. Since AdS is maximally symmetric, the integrand in (2.3) can be evaluated explicitly [11] and written as[3]

$$\left.\frac{\delta\mathcal{L}}{\delta\mathcal{R}^{\mu\nu}{}_{\rho\sigma}}\right|_{\text{AdS}} = -\frac{L^2}{4d}\left(\delta_\mu^\rho \delta_\nu^\sigma - \delta_\mu^\sigma \delta_\nu^\rho\right)\mathcal{L}|_{\text{AdS}}\,, \tag{2.6}$$

where $\mathcal{L}\big|_{\text{AdS}}$ is the Lagrangian density (2.1) evaluated in the pure AdS solution. Using this, we can evaluate Wald's functional and write it as

$$S_\xi\left[g_{\mu\nu}^{\text{AdS}}(\lambda_i), \lambda_i\right] = \frac{4\pi a_d^*(\lambda_i)}{\text{Vol}(S^{d-1})}\widetilde{\mathcal{A}}_{\text{horizon}}\,, \tag{2.7}$$

where $\widetilde{\mathcal{A}}_{\text{horizon}}$ is the horizon area $\mathcal{A}_{\text{horizon}}$ divided by the AdS radius L^{d-1}. We have identified a_d^* according to [11, 21]

$$a_d^*(\lambda_i) = -\frac{1}{2d}\text{Vol}(S^{d-1})L^{d+1}\mathcal{L}\big|_{\text{AdS}}\,, \tag{2.8}$$

---

[1]The perturbation $\delta g_{\mu\nu}$ can be any metric which satisfies the equations of motion obtained from (2.1) linearized around pure AdS.

[2]The Iyer and Wald formalism derives the ordinary first law from a $(d-1)$-form $\chi = \delta Q_\xi - \xi \cdot \Theta$ that is closed $d\chi = 0$ on shell, where Θ is the symplectic potential and δQ_ξ is variation of the Noether charge associated to ξ. The integral of χ vanishes on-shell so that one obtains a relation between the two boundary contributions, at the bifurcate Killing horizon and the asymptotic boundary. At the horizon we have $\xi = 0$ so that $\chi = \delta Q_\xi$, and $\int_{\text{horizon}} \chi = \frac{\kappa}{2\pi}\delta S_\xi$, while at the asymptotic boundary $\int_\infty \chi = \delta E_\xi$ (see also ref. [4]).

[3]To obtain this general expression the only thing that is required is that the metric is locally AdS, see section 5.2 of ref. [11] for details. This becomes very useful in section 4, where it allows us to extend some of our results beyond pure AdS in three dimensional gravity.

where $\mathrm{Vol}(S^{d-1}) = 2\pi^{d/2}/\Gamma(d/2)$. The coefficient $a_d^* = a_d^*(\lambda_i)$ is in general a complicated function of the coupling constants of the theory. Using (2.7) we can easily evaluate the variation in (2.5) and find

$$\delta_{\lambda_i} S_\xi = \frac{S_\xi}{a_d^*} \delta_{\lambda_i} a_d^*, \qquad \delta_{\lambda_i} a_d^*(\lambda_i) = \sum_i \left(\frac{\partial a_d^*}{\partial \lambda_i} \right) \delta \lambda_i . \qquad (2.9)$$

This expression relies on the fact that the pure AdS metric $g_{\mu\nu}^{\mathrm{AdS}}(L)$ is only a function of the length scale $L = L(\lambda_i)$, which means the dimensionless horizon area $\widetilde{\mathcal{A}}_{\mathrm{horizon}} = \mathcal{A}_{\mathrm{horizon}}/L^{d-1}$ is independent of λ_i. In section 4 we revisit this when considering more general metrics in three dimensional gravity.

Since we are considering linear perturbations we can put together the results in eqs. (2.4) and (2.9) and obtain the following bulk extended first law

$$\delta S_\xi = \frac{2\pi}{\kappa} \delta E_\xi' + \frac{S_\xi}{a_d^*} \delta a_d^* . \qquad (2.10)$$

We can already see the similarities of this bulk relation with the extended first law of entanglement (1.1). For a particular Killing vector ξ in AdS, this result was first obtained in ref. [8] for Einstein gravity and later in refs. [12–14] for specific higher curvature gravity theories.[4] Our derivation generalizes to arbitrary covariant theories of gravity as well as any Killing horizon in pure AdS. The method is quite simple and follows almost immediately upon evaluating Wald's functional in (2.7).[5]

Before analyzing the holographic consequences of this relation, let us comment on the prime we have added on the charge E_ξ' in (2.10). From the derivation of (2.9) it is clear that when the variation is only given by $\lambda_i \to \lambda_i + \delta\lambda_i$, the first term in (2.10) vanishes, $\delta_{\lambda_i} E_\xi' = 0$, i.e.

$$E_\xi \left[g_{\mu\nu}^{\mathrm{AdS}}(\lambda_i + \delta\lambda_i), \lambda_i + \delta\lambda_i \right] - E_\xi \left[g_{\mu\nu}^{\mathrm{AdS}}(\lambda_i), \lambda_i \right] = 0 . \qquad (2.11)$$

Given that there is no reason for these terms to cancel each other for arbitrary values of λ_i, both must vanish separately. This is achieved by defining the normalized quantity E_ξ' as

$$E_\xi' [g_{\mu\nu}, \lambda_i] = E_\xi [g_{\mu\nu}, \lambda_i] - E_\xi \left[g_{\mu\nu}^{\mathrm{AdS}}(\lambda_i), \lambda_i \right] . \qquad (2.12)$$

While this normalization plays no role in (2.4) when considering metric perturbations, it gives the appropriate behavior under more general variations. This prescription is equivalent to subtracting the Casimir energy contribution in pure AdS, that is present for certain foliations of the space-time (see ref. [23] for some examples). The procedure is common in extended black hole thermodynamics, where the Casimir energy is not included in the first law [5].

---

[4]In some of these papers this relation is not written in terms of the coefficient a_d^*, but in terms of the coupling constants $\{\lambda_i\}$ of particular theories.

[5]Comparing with the methods in [13, 14] we find we do not have to explicitly deal with additional divergences that arise (and ultimately cancel) from evaluating the Iyer-Wald form at the asymptotic boundary when implementing the extended Iyer-Wald formalism (see also [22]).

2.1 Mapping to boundary CFT

We are mainly interested in the first law in (2.10) from the perspective of a holographic CFT. Taking a bulk coordinate z so that the AdS boundary is located at $z \to 0$, the d-dimensional space-time in which the CFT is defined is given by

$$\lim_{z \to 0} ds^2_{\text{bulk}} = w^2(x^\mu) ds^2_{\text{CFT}} + \dots . \tag{2.13}$$

Applying a bulk diffeomorphism or changing the definition of $w^2(x^\mu)$ results in a different boundary space-time. A particular way of taking this limit corresponds to choosing a conformal frame. We will shortly take advantage of this freedom, which from the CFT perspective is equivalent to a conformal transformation.

What about the quantum state of the boundary CFT? Although the bulk space-time is pure AdS, the CFT is technically not in the vacuum state since there is a horizon and therefore an associated temperature, given by the surface gravity in (2.4) according to $\beta = 2\pi/\kappa$. This means the boundary state is thermal with respect to the Killing flow evaluated at the boundary, i.e.

$$\rho = \frac{1}{Z} \exp\left(-\beta K_\xi\right) , \tag{2.14}$$

where the operator K_ξ generates the flow of ξ^μ as we approach the boundary. It can be written explicitly in terms of the boundary coordinates x^a and the pullback of the Killing vector ξ^a as

$$K_\xi = \int_{\Sigma_\xi} \xi^a T_{ab} dS^b , \tag{2.15}$$

where T_{ab} is the stress tensor of the CFT and the integral is over a boundary codimension one space-like surface Σ_ξ where the vector ξ^a is time-like. The directed surface element dS^a is given by $dS^a = dS n^a$, with n^a a unit vector normal to Σ_ξ.

The variation of the conserved quantity E'_ξ appearing in the gravitational first law (2.10) is given by the variation of the expectation value of K_ξ in the state (2.14). The normalization condition for E'_ξ in (2.12) translates into the following normalization of the stress tensor T_{ab}

$$T_{ab} \quad \to \quad T'_{ab} = T_{ab} - \langle T_{ab} \rangle_\rho , \tag{2.16}$$

with ρ in (2.14). Since a bulk Killing vector gives a *conformal* Killing vector at the boundary, the operator K_ξ does not correspond to the Hamiltonian in general. We shall shortly consider some examples which illustrate this.

Putting everything together, the gravitational first law (2.10) maps to the boundary CFT according to

$$\delta S = \beta \, \delta \langle K_\xi \rangle_\rho + \frac{S}{a_d^*} \delta a_d^* , \tag{2.17}$$

where we identified the horizon entropy S_ξ with the Von Neumann entropy $S(\rho) = -\text{Tr}(\rho \ln(\rho))$ of ρ in (2.14). From the field theory perspective it might not be entirely clear what each of these terms corresponds to, so let us write them more explicitly.

For perturbations in which we keep the CFT fixed it is clear that $\delta a_d^* = 0$ while the state is deformed according to $\rho + \delta\rho$. In this case, the relation (2.17) is similar to the first law of thermodynamics. When $\delta a_d^* \neq 0$ we must be more careful since in this case the CFT is changing, which in particular implies that the Hilbert space shifts $\mathcal{H} \to \bar{\mathcal{H}}$. The state ρ cannot remain fixed, meaning that $\delta a_d^* \neq 0$ induces a variation of ρ given by

$$\rho \qquad \longrightarrow \qquad \bar{\rho} = \frac{1}{Z} \exp\left(-\beta \bar{K}_\xi\right), \qquad (2.18)$$

where $\bar{\rho}$ and $\bar{K}_\xi$ are the same operators but acting on the Hilbert space $\bar{\mathcal{H}}$ instead. In this case the extended first law (2.17) can be written explicitly as

$$S(\bar{\rho}) - S(\rho) = \beta \left[\langle \bar{K}_\xi \rangle_{\bar{\rho}} - \langle K_\xi \rangle_\rho\right] + \frac{S(\rho)}{a_d^*} \delta a_d^*. \qquad (2.19)$$

Notice that the first terms on the right-hand side involve operators on different Hilbert spaces. Moreover, the normalization of K_ξ given in (2.16) (and an analogous expression for $\bar{K}_\xi$) implies that both terms between square brackets vanish independently. This is equivalent to the gravitational case, where we obtained (2.9).

Putting everything together, the most general perturbation of the Von Neumann entropy of ρ is given by

$$S(\bar{\rho} + \delta\bar{\rho}) - S(\rho) = \beta \operatorname{Tr}\left(\bar{K}_\xi \, \delta\bar{\rho}\right) + \frac{S(\rho)}{a_d^*} \delta a_d^*, \qquad (2.20)$$

where we have used $\langle K_\xi \rangle_\rho = \langle \bar{K}_\xi \rangle_{\bar{\rho}} = 0$. This expression considers the simultaneous variations $a_d^* \to a_d^* + \delta a_d^*$ and $\rho \to \bar{\rho} + \delta\bar{\rho}$, and clarifies the precise meaning of (2.17), which without any explanation is rather obscure.

2.2 Extended first law of entanglement

So far we have shown that (2.17) follows from AdS/CFT when studying Killing horizons in pure AdS. We now consider particular horizons that will allow us to identify this relation as the extended first law of entanglement. Let us start with the simplest example of a Killing horizon in AdS, obtained by writing pure AdS in a hyperbolic slicing

$$ds^2 = -\left(\frac{\rho^2 - L^2}{R^2}\right) d\tau^2 + \left(\frac{L^2}{\rho^2 - L^2}\right) d\rho^2 + \rho^2 dH_{d-1}^2, \qquad (2.21)$$

where R is an arbitrary positive constant and dH_{d-1} is the line element of a unit hyperbolic plane. This space-time is often referred as Rindler-AdS since it describes a section of anti-de Sitter. The vector $\xi = \partial_\tau$ trivially satisfies Killing's equation and is time-like over the whole patch $\rho \geq L$, generating a horizon at $\rho = L$. It therefore satisfies all the conditions leading to the first law in (2.10) and (2.20).

A simple computation shows that the surface gravity is $\kappa = 1/R$, while the boundary metric is given by $ds_{\text{CFT}}^2 = -d\tau^2 + R^2 dH_{d-1}^2 \equiv \mathbb{R} \times \mathbb{H}^{d-1}$. From this we see that $\xi = \partial_\tau$ is also a Killing vector of ds_{CFT}^2, so that K_ξ in (2.15) is equal to the Hamiltonian and can be written as

$$K_\xi = \int_{\tau=0} T'_{\tau\tau} dS^\tau \equiv H_\tau. \qquad (2.22)$$

This means the boundary state is an ordinary thermal state $\rho_\beta \propto \exp(-\beta H_\tau)$, where the inverse temperature is fixed by the surface gravity to $\beta = 2\pi R$. The extended first law (2.17) then becomes

$$\delta S(\rho_\beta) = \beta\, \delta\langle H_\tau\rangle + \frac{S(\rho_\beta)}{a_d^*}\delta a_d^* \, . \tag{2.23}$$

While the first term is nothing more than the first law of thermodynamics, the second contribution is unique to the case of inverse temperature $\beta = 2\pi R$. This is clear from the holographic perspective, since moving away from this temperature is equivalent to leaving pure AdS, where the analysis of the previous section is no longer valid. In section 4 we show that for $d = 2$ this expression remains valid for arbitrary values of β. Although (2.23) is not the extended first law of entanglement (since it involves a thermal state in $\mathbb{R}\times\mathbb{H}^{d-1}$), this simple example will be very useful in what follows.

2.2.1 Shifting conformal frames

Building on the canonical example we just described, we can obtain the more complicated setups we are actually interested in. To obtain the extended first law of entanglement we take advantage of the freedom present when taking the boundary limit in (2.13). Different ways of taking this limit correspond to distinct conformal frames and result in different setups for the boundary CFT. We still consider the bulk Killing vector $\xi = \partial_\tau$, but written in a different set of coordinates corresponding to distinct conformal frames.

Ball in Minkowski. Let us first show how we can recover the extended first law of entanglement for the Minkowski vacuum reduced to a ball. We first apply a change of coordinates on the Rindler-AdS metric (2.21), which is given in eq. (4.7) of ref. [24], so that the metric becomes

$$ds^2 = \left(\frac{L}{\hat{r}\sin(\psi)}\right)^2\left(-dt^2 + d\hat{r}^2 + \hat{r}^2\left(d\psi^2 + \cos^2(\psi)d\Omega_{d-2}^2\right)\right), \tag{2.24}$$

where $\hat{r} \geq 0$, $\psi \in [0, \pi/2]$ and $d\Omega_{d-2}$ is the line element of a unit sphere S^{d-2}. This is nothing more than the AdS Poincaré patch, as can be seen by defining the usual coordinates $(z, r) = \hat{r}(\sin(\psi), \cos(\psi))$. At the boundary $\psi \to 0$ we recover d-dimensional Minkowski space-time with $\hat{r} = r$ the spatial radial coordinate. We use the convention in which the boundary coordinate r refers to the bulk coordinate $\hat{r}$ when $\psi \to 0$. This same notation is used in the following examples.

It is straightforward to write the Killing vector $\xi = \partial_\tau$ in these new coordinates and find

$$\xi = \left(\frac{R^2 - \hat{r}_+^2}{2R^2}\right)\partial_{\hat{r}_+} - \left(\frac{R^2 - \hat{r}_-^2}{2R^2}\right)\partial_{\hat{r}_-}\, , \tag{2.25}$$

where $\hat{r}_\pm = \hat{r} \pm t$. The important difference with respect to the hyperbolic example is that this Killing vector is time-like only in a section of the metric (2.24), given by $|\hat{r}_\pm| \leq R$. For the Minkowski boundary this corresponds to the causal domain of a ball of radius R. The operator generating the flow of ξ inside the ball can be written from (2.15) as

$$K_\xi = \int_{r\leq R}\left(\frac{R^2 - r^2}{2R^2}\right)T_{tt}'\, dS^t \, . \tag{2.26}$$

While this is clearly not the Hamiltonian generating t translations in Minkowski, it is proportional to the modular hamiltonian characterizing the Minkowski vacuum reduced to the ball [21]. The proportionality constant missing to make the identification is given by $K_{\text{Ball}} = 2\pi R K_\xi$, that is precisely the inverse temperature $\beta = 2\pi R$ obtained from the surface gravity of the bulk Killing vector (2.25). Altogether, the quantum state ρ in (2.14) is exactly given by the Minkowski vacuum reduced to the ball. The Von Neumann entropy is equivalent to the entanglement entropy, so that (2.17) becomes the extended first law of entanglement (1.1).

Half-space in Minkowski. Another interesting case is obtained by applying the change of coordinates given in eq. (4.4) of ref. [24] (see also ref. [25]) to the Rindler-AdS space-time, so that the bulk metric (2.21) becomes

$$ds^2 = (L/z)^2 \left(dz^2 - dt^2 + dx^2 + d\vec{y}.d\vec{y}\right) , \qquad (2.27)$$

where $(x, \vec{y}) \in \mathbb{R} \times \mathbb{R}^{d-2}$. Once again we recognize the Poincaré patch of AdS, so that we recover a d-dimensional Minkowski boundary when $z \to 0$. The Killing vector $\xi = \partial_\tau$ in these coordinates is given by

$$\xi = (x_+/R)\partial_{x_+} - (x_-/R)\partial_{x_-} , \qquad (2.28)$$

where $x_\pm = x \pm t$. This vector is time-like when $x_\pm \geq 0$, which from the boundary perspective corresponds to the Rindler region, i.e. the causal domain of the half space $x \geq 0$. Using (2.15) to compute the operator generating the Killing flow at the boundary we find

$$K_\xi = \int_{x>0} (x/R)T'_{tt}\, dS^t . \qquad (2.29)$$

Since the surface gravity of (2.28) is still given by $\kappa = 1/R$, the inverse temperature is $\beta = 2\pi R$ and we recognize $\rho \propto \exp(-\beta K_\xi)$ as the Minkowski vacuum reduced to Rindler [26, 27]. Similarly to the previous case, (2.17) becomes the extended first law of entanglement (1.1) but in this case, for the Minkowski vacuum reduced to the half-space.

Spherical cap in Lorentzian cylinder. Let us now show how we can obtain the extended first law of entanglement for holographic CFTs defined on curved backgrounds. Consider the following change of coordinates on the AdS metric (2.24)

$$\hat{r}_\pm(\hat{\theta}_\pm) = R\frac{\tan(\hat{\theta}_\pm/2)}{\tan(\theta_0/2)} , \qquad (2.30)$$

where $\hat{\theta}_\pm = \hat{\theta} \pm \sigma/R$ and $\theta_0 \in [0,\pi]$ is a fixed parameter. The metric (2.24) becomes

$$ds^2 = \left[\frac{L/R}{\sin(\psi)\sin(\hat{\theta})}\right]^2 \left(-d\sigma^2 + R^2 d\hat{\theta}^2 + R^2\sin^2(\hat{\theta})\left(d\psi^2 + \cos^2(\psi)d\Omega_{d-2}^2\right)\right), \qquad (2.31)$$

where $\sigma \in \mathbb{R}$ is the time coordinate and $\hat{\theta}$ is restricted to $\hat{\theta} \in [0,\pi]$. As we take the boundary limit $\psi \to 0$ and remove the conformal factor between square brackets we find that the CFT
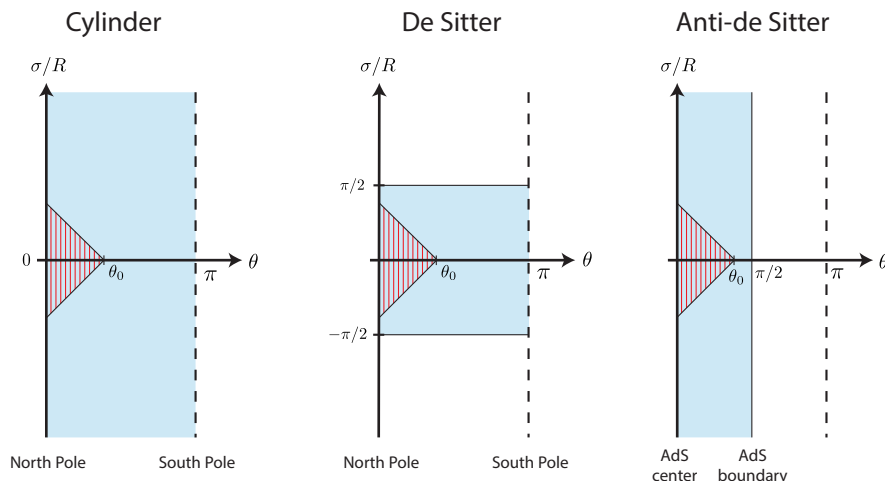
Figure 1. Boundary space-times represented in the $(\sigma/R, \theta)$ plane. The blue region corresponds to the section of the $(\sigma/R, \theta)$ plane covered by the boundary metrics (2.31) (in the limit $\psi \to 0$ and without the conformal factor), (2.34) and (2.36). In red we see the region in which the boundary vector ξ^a is time-like and therefore the extended first law of entanglement applies.

is defined in the Lorentzian cylinder $\mathbb{R} \times S^{d-1}$ with metric $ds^2_{\mathrm{CFT}} = -d\sigma^2 + R^2 d\Omega^2_{d-1}$. The bulk coordinate $\hat{\theta}$ becomes the polar angle $\hat{\theta} = \theta$ on the spatial sphere S^{d-1}, with $\theta = 0, \pi$ corresponding to the North and South poles respectively.

The Killing vector ξ in (2.25) can be written in these coordinates as

$$\xi = \left(\frac{\cos(\hat{\theta}_+) - \cos(\theta_0)}{R\sin(\theta_0)} \right) \partial_{\hat{\theta}_+} - \left(\frac{\cos(\hat{\theta}_-) - \cos(\theta_0)}{R\sin(\theta_0)} \right) \partial_{\hat{\theta}_-} . \tag{2.32}$$

Computing its magnitude we see that the bulk region in which this vector is time-like is given by $|\hat{\theta}_\pm| < \theta_0$. For the boundary CFT in the Lorentzian cylinder, this corresponds to the causal domain of a spherical cap on the spatial S^{d-1} given by $\theta \in [0, \theta_0]$ at $\sigma = 0$. Plotting this region in the $(\sigma/R, \theta)$ plane we obtain the left diagram in figure 1. The whole infinite strip in blue corresponds to the Lorentzian cylinder $\mathbb{R} \times S^{d-1}$, with the North and South pole located at $\theta = 0, \pi$.

The operator generating the Killing flow at the boundary is computed from (2.15) as

$$K_\xi = \int_{\theta \leq \theta_0} \left(\frac{\cos(\theta) - \cos(\theta_0)}{R\sin(\theta_0)} \right) T'_{\sigma\sigma} \, dS^\sigma . \tag{2.33}$$

In a similar way to the previous case, we recognize the state $\rho \propto \exp\left(-\beta K_\xi\right)$ with $\beta = 2\pi R$ as the vacuum state of the cylinder reduced to the spherical cap [21]. This gives the extended first law of entanglement for a CFT in the Lorentzian cylinder (1.1).

Spherical cap in de Sitter. Using the same coordinates as in (2.31) we can obtain a CFT defined on a de Sitter background by taking the limit $\psi \to 0$ and choosing the conformal factor so that the boundary metric is given by

$$ds^2_{\mathrm{CFT}} = \frac{-d\sigma^2 + R^2 d\Omega^2_{d-1}}{\cos^2(\sigma/R)} . \tag{2.34}$$

– 10 –

This is d-dimensional global de Sitter space-time, as can be seen by changing the time coordinate to $\cosh(t_s/R) = 1/\cos(\sigma/R)$, so that we get

$$ds^2_{\text{CFT}} = -dt_s^2 + R^2 \cosh^2(t_s/R) d\Omega^2_{d-1} \ . \tag{2.35}$$

It is convenient to work in the time coordinate σ, since the Killing vector ξ has the simple form given in (2.32) and is time-like when $|\theta_{\pm}| \le \theta_0$. Plotting this region in the $(\sigma/R, \theta)$ plane for the boundary metric (2.34), we obtain the center diagram in figure 1. The main difference with respect to the case of the Lorentzian cylinder is that the full de Sitter space-time (blue region) is constrained to $|\sigma/R| \le \pi/2$ due to the denominator in (2.34). Since the topology of dS is the same as the cylinder $\mathbb{R} \times S^{d-1}$, the region in which ξ^a is time-like also corresponds to the causal domain of a spherical cap $\theta \in [0, \theta_0]$, but with θ_0 restricted to $\theta_0 \le \pi/2$.

The operator generating the flow of the Killing vector at the boundary is still given by (2.33),[6] which is equivalent to the modular hamiltonian of the dS vacuum after multiplying by $\beta = 2\pi R$. Altogether, this results in the extended first law of entanglement (1.1) for the de Sitter vacuum reduced to a spherical cap.

Ball in anti-de Sitter. Finally, we can obtain a CFT defined in an AdS_d space-time by taking the limit $\psi \to 0$ in (2.31) and choosing the conformal factor so that we get

$$ds^2_{\text{CFT}} = \frac{-d\sigma^2 + R^2(d\theta^2 + \sin^2(\theta)d\Omega^2_{d-2})}{\cos^2(\theta)} \ . \tag{2.36}$$

Changing coordinates to $\varrho = R\tan(\theta) \ge 0$ we recognize global AdS_d, with ϱ the usual radial coordinate. Similar to the dS case, it is convenient to describe the AdS_d boundary in terms of the (σ, θ) coordinates, where the Killing vector ξ and operator K_ξ are still given by (2.32) and (2.33). The main difference is that the region in which ξ is time-like $|\theta_{\pm}| \le \theta_0$, now corresponds to the causal domain of a ball in AdS_d of radius $\varrho_{\max} = R\tan(\theta_0)$. We plot this in the right diagram of figure 1, where $\theta = 0, \pi/2$ in (2.36) now correspond to the AdS center and boundary. The entanglement entropy associated to the vacuum state reduced on this ball satisfies the extended first law of entanglement in (1.1).

3 Killing horizons in pure AdS_2

Our calculations so far have been in the context of the $\text{AdS}_{d+1}/\text{CFT}_d$ correspondence for $d \ge 2$, where the duality is well understood. In this section we revisit the construction for the case in which $d = 1$, where the gravity theory is highly constrained and there is no clear holographic picture.

Let us start by briefly reviewing some basic notions of two dimensional gravity (see ref. [28] for a comprehensive review). In two space-time dimensions the most general scalar curvature invariant is built from the Ricci scalar $\mathcal{R}$ and contractions of its covariant

---

[6]The only difference with respect to the case of the cylinder is given by the induced surface element dS^σ, which is now computed from (2.34).

derivatives, e.g. $(\nabla \mathcal{R})^2 = (\nabla_\mu \mathcal{R})(\nabla^\mu \mathcal{R})$. Both the Riemann and Ricci tensor are fixed by $\mathcal{R}$ and $g_{\mu\nu}$ according to

$$\mathcal{R}_{\mu\nu\rho\sigma} = \frac{\mathcal{R}}{2} \left(g_{\mu\rho} g_{\nu\sigma} - g_{\mu\sigma} g_{\nu\rho} \right), \qquad \mathcal{R}_{\mu\nu} = \frac{\mathcal{R}}{2} g_{\mu\nu} . \tag{3.1}$$

This means there is a single gravitational degree of freedom, determined by $\mathcal{R}$. Similarly to the general d case in (2.1), the most general two dimensional gravity theory is given by

$$I[g_{\mu\nu}, \lambda_i] = \int d^2 x \sqrt{-g} \, \mathcal{L}(\mathcal{R}, \nabla_\mu \mathcal{R}, \dots), \tag{3.2}$$

where the coefficients λ_i are the coupling constants of the theory. The only constraint we impose is that there is a pure AdS solution with some radius $L = L(\lambda_i)$. Notice that the relations in (3.1) imply the Einstein tensor $G_{\mu\nu} = \mathcal{R}_{\mu\nu} - g_{\mu\nu} \mathcal{R}/2$ vanishes for every two dimensional metric, so that $\mathcal{L} = \mathcal{R}$ gives a trivial theory.

Just as in the higher dimensional case, let us consider a Killing vector ξ^μ of pure AdS$_2$ which is time-like over some region and generates a horizon (2.2). The associated entropy is computed from Wald's functional (2.3), that in the two dimensional case is given by

$$S_\xi[g_{\mu\nu}(L), \lambda_i] = -2\pi \left[\frac{\delta \mathcal{L}}{\delta R^{\mu\nu}{}_{\rho\sigma}} n^{\mu\nu} n_{\rho\sigma} \right]_{\text{Horizon}}, \tag{3.3}$$

where there is no integral since the bifurcate horizon is a single point. Evaluating in pure AdS we can use (2.6) to write this as

$$S_\xi[g_{\mu\nu}^{\text{AdS}}(L), \lambda_i] = 2\pi a_1^*(\lambda_i), \qquad \text{where} \qquad a_1^*(\lambda_i) = -L^2 \mathcal{L}\big|_{\text{AdS}} . \tag{3.4}$$

An important difference with respect to the higher dimensional case, is that in two dimensions this expression is always finite and only depends on the global features of the theory, i.e., it is insensitive to the details of the Killing vector ξ^μ. The entropy in (3.4) only depends on the pure AdS$_2$ radius and the Lagrangian density evaluated on AdS$_2$. Altogether, there is no obstruction in applying the same reasoning as in higher dimensions and write the extended first law for Killing horizons in pure AdS exactly as in (2.10)

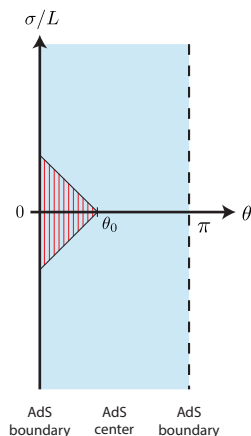$$\delta S_\xi = \frac{2\pi}{\kappa} \delta E'_\xi + \frac{S_\xi}{a_1^*} \delta a_1^* . \tag{3.5}$$

Let us construct a concrete example by first writing pure AdS$_2$ in global coordinates

$$ds^2 = \frac{-d\sigma^2 + L^2 d\theta^2}{\sin^2(\theta)}, \tag{3.6}$$

where $\sigma \in \mathbb{R}$ and $\theta \in (0, \pi)$. Notice this notation is different from the previous section, since θ is now a bulk coordinate and the boundary is just described by σ. Two-dimensional AdS is distinct from higher dimensions, since there are two disjoint boundaries at $\theta = 0, \pi$. A sketch of its Penrose diagram is given in figure 2.

We can easily check that the following is a Killing vector

$$\xi = \left(\frac{\cos(\theta_+) - \cos(\theta_0)}{L \sin(\theta_0)} \right) \partial_{\theta_+} - \left(\frac{\cos(\theta_-) - \cos(\theta_0)}{L \sin(\theta_0)} \right) \partial_{\theta_-}, \tag{3.7}$$

Figure 2. The blue region corresponds to AdS$_2$ space-time represented in the $(\sigma/L, \theta)$ plane, with the two boundaries at $\theta = 0, \pi$. In red we see the region in which the bulk Killing vector ξ^μ (3.7) is time-like and therefore the extended first law in (3.5) applies.

with surface gravity $\kappa = 1/L$. From its norm we see that it is time-like in the domain of dependence of the bulk surface $(\sigma = 0, \theta)$ with $\theta \in (0, \theta_0)$, meaning the boundary time coordinate is restricted to $|\sigma/L| \leq \theta_0$. This corresponds to the red region in figure 2.

As an example, let us compute the horizon entropy explicitly for a particular gravity theory, that we take as

$$\mathcal{L} = f(\mathcal{R}) = \lambda_0 + \lambda_2 \mathcal{R}^2 . \tag{3.8}$$

The AdS radius L is determined by solving the equations of motion evaluated at $\mathcal{R} = -2/L^2$, which can be written as

$$\nabla_\mu \nabla_\nu f'(\mathcal{R}) + \frac{1}{2} g_{\mu\nu} \left(\mathcal{R} f'(\mathcal{R}) - f(\mathcal{R}) \right) = 0 \qquad \Longrightarrow \qquad L^4 = \frac{4\lambda_2}{\lambda_0} . \tag{3.9}$$

Using this we can evaluate Wald's entropy in (3.4) as

$$S_\xi \left[g_{\mu\nu}^{\text{AdS}}(L), \lambda_i \right] = 2\pi \left(-8\lambda_2/L^2 \right) , \tag{3.10}$$

where between parenthesis we identify the factor a_1^*, which is positive if and only if $\lambda_2 < 0$.

This raises the question regarding the holographic interpretation of the extended first law as written in (3.5), since a_1^* is supposed to capture the number of degrees of freedom of the boundary theory. The usual AdS/CFT correspondence for a two dimensional bulk does not yield a clear picture as in the higher dimensional case. Although there has been very interesting work on the subject (see refs. [29–34]), there continues to be debate about what is meant by the dual "CFT$_1$", whether it is conformal quantum mechanics or the chiral sector of a two-dimensional CFT. Moreover in the context of Jackiw-Teitelboim (JT) gravity [15, 16] it is understood that the boundary is not a single theory but an ensemble average [35]. For these reasons, we refrain from giving a boundary interpretation of the extended first law and leave this aspect to future investigations.

3.1 Einstein-dilaton theories

So far we have considered two dimensional theories of gravity in which the only field is given by the metric $g_{\mu\nu}$. We now discuss the extended first law for Einstein-dilaton theories, which are widely studied in the context of two dimensional gravity.

One disadvantage of the pure gravity action considered in (3.2) is that since non-trivial theories must have $\mathcal{L} \sim \mathcal{O}(\mathcal{R}^2)$, the equations of motion for the metric are at least fourth order differential equations. This issue can be avoided by the introduction of an auxiliary dilaton field $\phi(x^\mu)$ coupled to ordinary Einstein gravity

$$I_\phi[g_{\mu\nu}, \lambda_i] = \int d^2x \sqrt{-g}\left[\phi\mathcal{R} - V(\phi)\right] . \tag{3.11}$$

The equations of motion obtained from this action are second order. In particular, varying with respect to the dilaton field we get the algebraic constraint $\mathcal{R} = V'(\phi)$. If the potential has non-vanishing second derivative, one can invert this relation and substitute back into the action (3.11) to obtain a purely gravitational theory of the type $\mathcal{L} = f(\mathcal{R})$. As an example, if we take $V(\phi) = \phi^2/4\lambda_2 - \lambda_0$, the equation of motion for ϕ sets $\phi_0 = 2\lambda_2\mathcal{R}$ and we get

$$I_{\phi=\phi_0}[g_{\mu\nu}, \lambda_i] = \int d^2x \sqrt{-g}\left[\lambda_0 + \lambda_2\mathcal{R}^2\right] , \tag{3.12}$$

that is the gravity theory previously considered in (3.8). This allows us to study two dimensional gravity from the simpler action (3.11). We should interpret the dilaton field as a gravitational degree of freedom, which gets non-trivial dynamics from varying (3.11) with respect to the metric

$$\nabla_\mu\nabla_\nu\phi = \frac{1}{2}g_{\mu\nu}V(\phi) . \tag{3.13}$$

Since the Einstein-dilaton theories in (3.11) (with $V''(\phi) \neq 0$) are equivalent to the purely gravitational action previously considered in (3.2), the results obtained for the extended first law also hold in this setup. We should mention that while JT gravity is given by (3.11) with $V(\phi) \propto \phi$, it cannot be written as a purely gravitational theory since $V''(\phi) = 0$ and the dilaton equation simply fixes the curvature to a constant $\mathcal{R} = \text{const}$. We analyze the case of JT gravity separately in the next subsection.

There are more general Einstein-dilaton actions than (3.11) that yield interesting two dimensional theories. For instance, there is a particular way of taking the two-dimensional limit of higher dimensional Einstein gravity which results in the following action [36]

$$I_\phi[g_{\mu\nu}, \Lambda_2] = \int d^2x \sqrt{-g}\left[\phi\mathcal{R} + \frac{1}{2}(\nabla\phi)^2 - 2\Lambda_2\right] , \tag{3.14}$$

where Λ_2 is a coupling constant. This theory was studied in ref. [37] from the perspective of extended black hole thermodynamics. Although this action is clearly different from (3.11), if we redefine the metric according to $\tilde{g}_{\mu\nu} = e^{\phi/2}g_{\mu\nu}$ it can be written as

$$I_\phi[\tilde{g}_{\mu\nu}, \Lambda_2] = \int d^2x \sqrt{-\tilde{g}}\left[\phi\tilde{\mathcal{R}} - V(\phi)\right], \qquad \text{where} \qquad V(\phi) = 2\Lambda_2 e^{-\phi/2} . \tag{3.15}$$

Once we have the action in this form, we can solve the dilaton field equation and substitute back into the action to get a purely gravitational theory for the metric $\tilde{g}_{\mu\nu}$

$$I_{\phi=\phi_0}\left[\tilde{g}_{\mu\nu}, \Lambda_2\right] = \int d^2x \sqrt{-\tilde{g}} f(\tilde{\mathcal{R}}), \qquad \text{where} \qquad f(x) = 2x\left(1 - \ln(-x/\Lambda_2)\right). \quad (3.16)$$

This raises the question of which is the "physical" gravitational metric, either $g_{\mu\nu}$ (usually called the Jordan frame) or $\tilde{g}_{\mu\nu}$ (the Einstein frame).[7] The distinction between the frames is important as the solutions obtained in either case are very different. For instance, if we consider a constant curvature solution for $\tilde{g}_{\mu\nu}$, the equation of motion from (3.16) is given by

$$\tilde{\mathcal{R}}f'(\tilde{\mathcal{R}}) - f(\tilde{\mathcal{R}}) = 0 \qquad \Longrightarrow \qquad \tilde{\mathcal{R}} = 0. \quad (3.17)$$

From (3.1), this implies the metric $\tilde{g}_{\mu\nu}$ vanishes, so that the theory does not admit a pure AdS$_2$ solution and we cannot consider the extended first law in (3.5).

On the other hand, working in the Jordan frame with the metric $g_{\mu\nu}$ the action (3.14) allows a pure AdS$_2$ solution [37]. This means it is sensible to consider the extended first law for the metric $g_{\mu\nu}$, although the derivation leading to (3.5) does not apply. An extended first law of black hole thermodynamics (which studies the behavior of the black hole entropy under variations of the cosmological constant) was derived in ref. [37] for the Einstein-dilaton theory in (3.14). In order to obtain a sensible result, the authors of ref. [37] use an unconventional approach that involves rescaling Newton's constant according to $G_{d+1} = \frac{(1-d)}{2}G_2$. Starting from the results in ref. [8], this procedure can also be applied to derive an extended first law for perturbations of Killing horizons in the AdS$_2$ metric $g_{\mu\nu}$.

3.2 Jackiw-Teitelboim gravity

In this subsection we consider the extended first law in the context of Jackiw-Teitelboim gravity [15, 16], that correspond to an Einstein-dilaton theory that cannot be written as a purely gravitational theory of the type $\mathcal{L} = f(\mathcal{R})$. The action defining the theory can be written as

$$I_{JT} = I_\phi[g_{\mu\nu}; \phi_0, L] = \int d^2x\sqrt{-g}\left[\phi_0\mathcal{R} + \phi(x)(\mathcal{R} + 2/L^2)\right]. \quad (3.18)$$

The dilaton field $\phi(x)$ is dimensionless and there are two coupling constants that define the theory $\lambda_i = (\phi_0, L)$. As usual, the action must be supplemented with appropriate boundary terms to yield a well defined variational problem. The equations of motion can be easily computed and written as

$$\mathcal{R} + 2/L^2 = 0$$
$$\left[\nabla_\mu\nabla_\nu - \frac{g_{\mu\nu}}{L^2}\right]\phi(x) = 0. \quad (3.19)$$

The first equation fixes the Ricci scalar to a negative constant value and since the theory is two dimensional, it completely determines the Riemann tensor (3.1). This means the

_____

[7]See refs. [38, 39] for a discussion around this issue.

only metric solution in JT gravity is pure AdS$_2$. The analysis of the extended first law in JT gravity is extremely simple given that all we have to do is analyze the thermodynamic behavior of Killing horizons in pure AdS$_2$. The theory does not admit any real black hole solution.[8]

Writing the metric in global coordinates (σ, θ) as in (3.6) the only Killing horizon is generated by the vector in (3.7), which is time-like in the region $\theta_{\pm} < \theta_0 \in (0, \pi)$, sketched in figure 2. The equation of motion of the dilaton $\phi(x)$ can be easily solved in global coordinates and written as

$$\phi(\sigma, \theta) = \phi_h \frac{\cos(\sigma/L)\sin(\theta_0)}{\sin(\theta)}\,, \tag{3.20}$$

where $\phi_h > 0$ is an integration constant that gives the value of the dilaton at the horizon. The full solution is parametrized by the value of the single constant ϕ_h.[9]

To compute the horizon entropy we use Wald's functional (3.3) together with the fact that the Riemann tensor is fixed by the Ricci scalar $\mathcal{R}$ in (3.1)

$$S_\xi = 4\pi \left.\frac{\delta\mathcal{L}}{\delta R}\right|_{\text{Horizon}} = 4\pi\phi_0 + 4\pi\phi(x)\big|_{\theta_\pm=\theta_0} = 4\pi(\phi_0 + \phi_h)\,. \tag{3.21}$$

This agrees with the result obtained from the semi-classical computation of the Euclidean path integral [41]. The extended first law involves computing the entropy variation with respect to the coupling constants of the theory $\lambda_i = (\phi_0, L)$ and checking whether it can be written as

$$\delta_{\lambda_i} S_\xi = \frac{S_\xi}{a_1^*}\delta_{\lambda_i} a_1^*\,, \tag{3.22}$$

where a_1^* is some function of the coupling constants $a_1^* = a_1^*(\phi_0, L)$. In this setup we have no natural definition of a_1^* in terms of the on-shell Lagrangian (3.4), so in principle we can allow any function that depends exclusively on the coupling constants (ϕ_0, L). However, since a_1^* and ϕ_0 are dimensionless quantities and L has dimensions of length we have it can only depend on ϕ_0.[10] From the simple expression of the entropy given in (3.21) we can compute the entropy variation explicitly and find it is not compatible with the extended first law as written in (3.22) for any definition of $a_1^*(\phi_0)$

$$\delta_{\lambda_i} S_\xi = 4\pi\delta\phi_0 \neq \frac{S_\xi}{a_1^*}\delta_{\lambda_i} a_1^*\,. \tag{3.23}$$

This means the form of the extended first law for JT gravity is not the same as in the previous cases we studied so far. The difference is that the solution in JT gravity depends on the additional parameter ϕ_h, that appears in the horizon entropy and is not related to the AdS radius L. In the previous derivations in section 2 we used the fact that the pure AdS solution only depends on the radius L.

---

[8]While the classical theory is almost trivial, interesting dynamics arise by introducing a fluctuating boundary. These boundary effects give one loop contributions to the Euclidean partition function [40, 41] and therefore lie beyond the semi-classical analysis captured by horizon thermodynamics.

[9]While it seems the solution also depends on $\theta_0 \in (0, \pi)$, we can use the isometries of AdS$_2$ to fix $\theta_0 = \pi/2$.

[10]Note that if we naively apply the definition of a_1^* in (3.4), we get $a_1^* = 2\phi_0$.

We expect a similar situation for other Einstein-dilaton theories that cannot be written as pure gravity theories. For any particular theory one can still compute the variation of the horizon entropy on pure AdS_2 as in (3.21), but there is no guarantee there exists a function $a_1^* = a_1^*(\lambda_i)$ such that it can be written as in the extended first law (3.22).

4 Beyond pure AdS in three dimensional gravity

Given that all our calculations so far have been for Killing horizons in pure AdS, a natural question is whether these results can be extended to horizons in more general space-times. In this section we investigate this in the context of three dimensional gravity, making contact with some concepts in extended black hole thermodynamics [7].

Consider a general three dimensional metric $g_{\mu\nu}$ which solves the equations of motion obtained from (2.1) and admits a time-like Killing horizon generated by the vector ξ^μ. The horizon entropy is obtained from Wald's functional (2.3) evaluated on $g_{\mu\nu}$, which for a general metric we cannot evaluate explicitly. However, three dimensional gravity theories admit interesting black hole solutions which are locally but not globally AdS, i.e., which satisfy

$$\mathcal{R}_{\mu\nu\rho\sigma} = -\frac{1}{L^2}\left(g_{\mu\rho}g_{\nu\sigma} - g_{\mu\sigma}g_{\nu\rho}\right) . \tag{4.1}$$

For this class of black holes we can evaluate the integrand in Wald's functional using (2.6) and find

$$S_\xi\left[g_{\mu\nu}, \lambda_i\right] = 2a_2^*(\lambda_i)\widetilde{\mathcal{A}}, \tag{4.2}$$

where $\widetilde{\mathcal{A}} = \mathcal{A}_{\mathrm{horizon}}/L^{d-1}$ and a_2^* in (2.8) is proportional to the Virasoro central charge c of the dual CFT_2. This expression is equivalent to the pure AdS relation (2.7) evaluated at $d = 2$.

Let us now consider the behavior of the entropy under deformations of the theory, i.e., $\lambda_i \to \lambda_i + \delta\lambda_i$ in (2.1). In this case, apart from the obvious contribution given by the coefficient $a_2^*(\lambda_i)$ in (4.2), we must take into account the variation of the dimension-less horizon area $\widetilde{\mathcal{A}}$. For the pure AdS metric, $\widetilde{\mathcal{A}}$ is independent of λ_i since the metric $g_{\mu\nu}^{\mathrm{AdS}}(L)$ only depends on the dimensionful parameter L, so that dimensional analysis implies $\mathcal{A}_{\mathrm{horizon}} \propto L^{d-1}$. This is no longer true for more general metrics which satisfy (4.1) but are not globally pure AdS, as the metric can also depend on some integration constants $\{c_j\}$ (e.g. mass, angular momentum, charge, etc.) so that the horizon area $\mathcal{A}_{\mathrm{horizon}}$ is no longer proportional to L^{d-1}. Altogether, the variation of (4.2) is now given by

$$\delta S_\xi = S_\xi\, \delta\left[\ln(a_2^*) + \ln(\widetilde{\mathcal{A}})\right] . \tag{4.3}$$

As we will shortly see in a simple example, computing this extra variation for a particular solution is straightforward. However, while the first term involving a_2^* has a clear meaning in the boundary CFT (given in (1.2)), this is not the case for $\widetilde{\mathcal{A}}$. Only by restricting ourselves to black holes in which $\delta\widetilde{\mathcal{A}} = 0$, the boundary CFT satisfies the extended first law given by

$$\delta\widetilde{\mathcal{A}} = 0 \qquad \Longrightarrow \qquad \delta S(\rho_\beta) = \beta\,\delta\langle H\rangle + \frac{S(\rho_\beta)}{a_2^*}\delta a_2^*, \tag{4.4}$$

where ρ_β is a thermal state and we have included the usual energy term $(2\pi/\kappa)\delta E'_\xi$ in (4.3) which maps to H, the hamiltonian of the CFT. Additional conserved quantities such as angular momentum or charges, can be added to this relation in the usual way. The first law in (4.4) is similar to the one obtained for the thermal state at temperature $\beta = 2\pi R$ in the background $\mathbb{R} \times \mathbb{H}^{d-1}$ (2.23), with the crucial difference that β in this case is unconstrained.

Let us illustrate how everything works by considering a simple example in Einstein gravity

$$I[g_{\mu\nu}; G, L] = \frac{1}{16\pi G} \int d^3x \sqrt{-g} \left(\mathcal{R} + \frac{2}{L^2} \right) . \tag{4.5}$$

The coupling constants of the theory are $\{\lambda_i\} = \{G, L\}$, where L is also the radius of the pure AdS solution. The rotating BTZ black hole solution satisfies (4.1) and is given by [42]

$$ds^2 = -f(r)dt^2 + \frac{dr^2}{f(r)} + r^2\Big(d\theta - \frac{GJ}{2r^2}dt\Big)^2 , \tag{4.6}$$

where $f(r) = -8GM + (r/L)^2 + (JG/2r)^2$. Different black holes are labeled by the integration constants $\{c_j\} = \{M, J\}$, which also give the global charges associated to the Killing vectors ∂_t and ∂_θ respectively.

The outer horizon radius r_+ is obtained from $f(r_+) = 0$ and is a non-trivial function of (G, L, M, J). We can easily write the dimensionless horizon area $\widetilde{\mathcal{A}}$ in terms of r_+

$$\widetilde{\mathcal{A}} = \frac{2\pi r_+}{L} = 4\pi\sqrt{MG} \left[1 + \sqrt{1 - \left(\frac{J}{8ML}\right)^2} \right]^{1/2} . \tag{4.7}$$

This expression depends explicitly on both G and L, meaning that the second term in (4.3) gives a non-trivial contribution, which we can easily write explicitly. However, if we consider the static black hole $J = 0$ we get $\widetilde{\mathcal{A}} = 4\pi\sqrt{2MG}$, which is independent of L. Therefore, if we restrict to variations of L (while keeping G fixed), we obtain the extended first law given in (4.4).

4.1 Extended thermodynamics and volume

Let us now restrict to a particular type of theory deformation, in which we take the radius of the pure AdS solution L as one of the coupling constants defining the theory and consider $\delta(\lambda_i, L) = (0, \delta L)$. This corresponds to the variations studied in the extended black hole thermodynamics [7], in which the thermodynamic pressure is identified with L according to $p \equiv d(d-1)/(16\pi G L^2)$. Its conjugate variable is referred as the volume V and can be defined from the entropy as

$$V \equiv -T\frac{\partial S_\xi}{\partial p} = -TS_\xi \frac{\partial}{\partial p} \left[\ln(a_2^*) + \ln(\widetilde{\mathcal{A}}) \right] . \tag{4.8}$$

where the second equality is obtained from (4.3). The p derivative is computed while keeping all the remaining parameters fixed.

This volume formula holds for locally AdS black holes in any three dimensional theory of gravity. Similar to (4.3), there are two distinct contributions to the volume. While the

variation of a_2^* has a natural boundary interpretation in terms of the number of degrees of freedom, the dimensionless area $\widetilde{\mathcal{A}}$ does not. For cases in which $\widetilde{\mathcal{A}}$ is independent of L, the thermodynamic volume takes the following simple form

$$\frac{\partial \widetilde{\mathcal{A}}}{\partial L} = 0 \quad \Longrightarrow \quad V = -\left(\frac{TS_\xi}{a_2^*}\right)\frac{\partial a_2^*}{\partial p} \; . \tag{4.9}$$

This gives a class of three dimensional black holes whose thermodynamic volume is directly related to changing the central charge of the boundary CFT. Since the meaning of V for the boundary theory is not completely understood (see refs. [8, 13, 43–47]), this formula might help give further insights. Let us use it in some concrete examples to compute the volume of some black hole solutions.

Thermodynamic volume in Einstein gravity. Consider the simple setup of a BTZ black hole (4.6) in Einstein gravity (4.5). As previously noted, for the static black hole $J = 0$ the dimensionless horizon area $\widetilde{\mathcal{A}}$ in (4.7) is independent of L, meaning that we can directly use the volume formula in (4.9). Simple calculations give $a_2^* = L/8G$ and $T = r_+/2\pi L^2$, so that we can compute the volume as

$$V_{J=0} = -\left(\frac{TS_\xi}{a_2^*}\right)\frac{\partial a_2^*}{\partial p} = \pi r_+^2 \; . \tag{4.10}$$

which agrees with the result obtained from a more standard approach in extended thermodynamics [37].

For the rotating BTZ solution with $J \neq 0$ the dimensionless horizon area $\widetilde{\mathcal{A}}$ in (4.7) is a non-trivial function of L, meaning that we must use the more general volume formula in (4.8). Although the calculation in this case is slightly more involved, the final result is again very simple and given by

$$V_{J\neq 0} = -TS_\xi\frac{\partial}{\partial p}\left[\ln(a_2^*) + \ln(\widetilde{\mathcal{A}})\right] = \pi r_+^2 \, , \tag{4.11}$$

in agreement with the previously known relation [37]. It is interesting to see that the extra variation with respect to $\widetilde{\mathcal{A}}$ is exactly what is needed in order to obtain this simple final answer. An interesting microscopic analysis of this expression was recently given in ref. [47].[11]

Thermodynamic volume in higher curvature theories. Since the volume formula (4.8) is particularly powerful in the context of higher curvature gravity theories, let us apply it in an example by considering the following generalization of new massive gravity [49–51]

$$I[g_{\mu\nu}] = \frac{1}{16\pi G}\int d^3x\sqrt{-g}\left(\mathcal{R} + \frac{2}{\ell^2} + \ell^2\mathcal{R}_2 + \ell^4\mathcal{R}_3\right) \, , \tag{4.12}$$

---

[11]We should mention that while the charged BTZ black hole in Einstein-Maxwell theory [48] is not locally AdS (4.1), if we naively apply the volume formula in (4.9) we obtain $V = \pi r_+^2 - \pi(QL/2)^2$, which agrees with the previously known result [37]. The reason it works is due to the fact that in Einstein gravity Wald's entropy functional always reduces to the Bekenstein-Hawking area expression, i.e. $S_\xi = \mathcal{A}/4G$. For higher curvature theories we do not expect the volume formula (4.8) to reproduce the correct result for the charged black hole.

where

$$\mathcal{R}_2 = 4(\lambda_1 \mathcal{R}_{\mu\nu}\mathcal{R}^{\mu\nu} + \lambda_2 \mathcal{R}^2) \,,$$

$$\mathcal{R}_3 = \frac{17}{12}(\mu_1 \mathcal{R}^\nu_{\ \mu}\mathcal{R}^\rho_{\ \nu}\mathcal{R}^\mu_{\ \rho} + \mu_2 \mathcal{R}_{\mu\nu}\mathcal{R}^{\mu\nu}\mathcal{R} + \mu_3 \mathcal{R}^3) \,. \tag{4.13}$$

The coupling constants of the theory are given by $\{G, \ell, \lambda_1, \lambda_2, \mu_i\}$ with $i = 1, 2, 3$, where new massive gravity [49, 50] is obtained by setting $\mu_i = 0$ and $\lambda_2 = -3\lambda_1/8$.

To apply the volume formula in (4.8) we must first compute the a_2^* factor, which depends on the pure AdS solution of the theory. We can find such solution by varying the action (4.12) with respect to the metric, which gives the following equations of motion [51]

$$\mathcal{R}_{\mu\nu} - \frac{1}{2}\mathcal{R}g_{\mu\nu} - \frac{1}{\ell^2}g_{\mu\nu} - H_{\mu\nu} = 0 \,, \tag{4.14}$$

where

$$\begin{aligned}
H_{\mu\nu} = 4\ell^2 &\left[\lambda_1 \left(-2\mathcal{R}^\rho_{\ \mu}\mathcal{R}_{\rho\nu} + \frac{1}{2}g_{\mu\nu}\mathcal{R}_{\rho\sigma}\mathcal{R}^{\rho\sigma} \right) + \lambda_2 \left(-2\mathcal{R}\mathcal{R}_{\mu\nu} + \frac{1}{2}g_{\mu\nu}\mathcal{R}^2 \right) \right] \\
&+ \frac{17}{12}\ell^4 \left[\mu_1 \left(-3\mathcal{R}_{\mu\rho}\mathcal{R}^\rho_{\ \sigma}\mathcal{R}^\sigma_{\ \nu} + \frac{1}{2}g_{\mu\nu}\mathcal{R}^\rho_{\ \sigma}\mathcal{R}^\alpha_{\ \rho}\mathcal{R}^\sigma_{\ \alpha} \right) + \mu_3 \left(-3\mathcal{R}^2\mathcal{R}_{\mu\nu} + \frac{1}{2}g_{\mu\nu}\mathcal{R}^3 \right) \right. \\
&\left. + \mu_2 \left(-\mathcal{R}^\rho_{\ \sigma}\mathcal{R}^\sigma_{\ \rho}\mathcal{R}_{\mu\nu} - 2\mathcal{R}\mathcal{R}_{\mu\rho}\mathcal{R}^\rho_{\ \nu} + \frac{1}{2}g_{\mu\nu}\mathcal{R}\mathcal{R}_{\rho\sigma}\mathcal{R}^{\rho\sigma} \right) \right] + \mathcal{O}(\nabla^2 \mathcal{R}, \nabla^2 \mathcal{R}^2, \dots) \,,
\end{aligned} \tag{4.15}$$

and we are omitting derivative terms that do not contribute to the pure AdS solution.

We can evaluate these complicated terms in a pure AdS metric $g_{\mu\nu}^{\mathrm{AdS}}(L)$ of some radius L using that it is a maximally symmetric space-time (4.1). Taking the trace of (4.14) and writing the AdS radius as $L = \ell/\sqrt{f_\infty}$ we obtain the following algebraic constraint for the factor f_∞

$$L = \ell/\sqrt{f_\infty} \qquad \Longrightarrow \qquad 1 - f_\infty - 8f_\infty^2(\lambda_1 + 3\lambda_2) + 17f_\infty^3(\mu_1 + 3\mu_2 + 9\mu_3) = 0 \,. \tag{4.16}$$

The solution f_∞ of this algebraic equation that is smoothly connected to Einstein gravity determines the pure AdS radius L. We can now write a_2^* from (2.8) by evaluating the Lagrangian density (4.12) in AdS, so that we find

$$a_2^* = -\frac{1}{2}\pi L^3 \mathcal{L}\big|_{\mathrm{AdS}} = \frac{L}{8G}\left[1 - 16f_\infty(\lambda_1 + 3\lambda_2) + 17f_\infty^2(\mu_1 + 3\mu_2 + 9\mu_3) \right] \,, \tag{4.17}$$

where we have used

$$\mathcal{R}_2 = \frac{48}{L^4}(\lambda_1 + 3\lambda_2)\,, \qquad\qquad \mathcal{R}_3 = -\frac{34}{L^6}(\mu_1 + 3\mu_2 + 9\mu_3) \,. \tag{4.18}$$

The expression for a_2^* and the constraint in (4.16) defining f_∞ reduce to the ones given in ref. [51] when setting $\lambda_2 = -3\lambda_1/8$ and $(\mu_1, \mu_2) = \mu_3(64, -72)/17$. Moreover, if we take $\lambda_1 = \lambda_2 = \mu_i = 0$ we get $f_\infty = 1$ and $a_2^* = L/8G$, in agreement with the Einstein gravity results. Notice that the dependence of a_2^* with the AdS radius L is linear, as in the Einstein case.

We can now consider a black hole solution for this theory. Given that the BTZ black hole in (4.6) is locally AdS, it solves the equations of motion in (4.14) as long as we take L according to (4.16). The horizon entropy is obtained from (4.2) with a_2^* and $\widetilde{\mathcal{A}}$ as given in (4.17) and (4.7). For the rotating solution with $J \neq 0$ we can use the volume formula in (4.8) and find

$$V_{J \neq 0} = \pi r_+^2 \left[1 - 16 f_\infty (\lambda_1 + 3\lambda_2) + f_\infty^2 (\mu_1 + 3\mu_2 + 9\mu_3) \right] . \tag{4.19}$$

To our knowledge, higher curvature contributions to the BTZ thermodynamic volume have not been computed before.

5 Discussion

The extended first law of entanglement has been previosuly derived for the Minkowski vacuum reduced to a ball by considering particular gravity theories in the bulk [8, 12–14]. In this work, we have shown a novel and simple procedure that generalizes the proof to arbitrary gravity theories in the bulk and new setups in the boundary CFT. From the bulk perspective we have found no obstructions in working in two dimensional gravity and also obtain some intriguing results concerning extended black hole thermodynamics in three dimensions. Let us discuss some additional aspects regarding the calculations in the main text.

Divergent terms in the extended first law of entanglement. One important feature of the ordinary first law of entanglement $\delta S_{EE} = \delta \langle K_B \rangle$ is that although the entanglement entropy always diverges, the left-hand side is well defined since the difference between entropies associated to different states is finite.[12] For the extended first law of entanglement this is no longer the case. Let us consider a variation of the theory without perturbing the state, so that the first term on the right-hand side of (2.20) drops out and we are left with

$$S_{EE}(\bar\rho) - S_{EE}(\rho) = \frac{S_{EE}(\rho)}{a_d^*} \delta a_d^* . \tag{5.1}$$

Both sides of this equality diverge, the left-hand side due to the fact that the divergences of the entanglement entropies corresponding to different theories do not cancel each other. This raises the question regarding how we should interpret (5.1), which seems to depend on the regularization procedure.

Let us illustrate the issue by considering the simple case of the Minkowski vacuum reduced to a ball of radius R in $d = 3$, where the entanglement entropy is [21]

$$S_{EE}(\rho_B) = \mu_1 \frac{R}{\epsilon} - 2\pi a_3^*, \tag{5.2}$$

with μ_1 a dimensionless and non-universal constant and a_3^* given by (1.2). The short distance cut-off ϵ can be properly defined using mutual information, see ref. [53]. If we

---

[12]As shown in ref. [52] this is not entirely true, since there are cases in which the entanglement entropy acquires state dependent divergences, so that δS_{EE} diverges. However, the relative entropy remains finite.

consider the same setup but for a CFT in which $\bar{a}_3^* = a_3^* - \delta a_3^*$, the entanglement entropy is given by

$$S_{EE}(\bar{\rho}_B) = \bar{\mu}_1 \frac{R}{\bar{\epsilon}} - 2\pi \bar{a}_3^* \,, \tag{5.3}$$

where the cut-off $\bar{\epsilon}$ and the constant $\bar{\mu}_1$ are not necessarily related to the ones appearing in (5.2).

How should we understand (5.1) in this context? A practical approach is to simply ignore the non-universal contributions to the entanglement entropy and regard (5.1) as a relation between the universal terms, where it is clear that the extended first law is satisfied. A different procedure is instead given by relating the cut-offs of each theory in a particular way such that the extended first law is satisfied to every order. Assuming there is a relation $\epsilon = \epsilon(\bar{\epsilon})$ which can be expanded around the origin as

$$\epsilon(\bar{\epsilon}) = \bar{\epsilon} \left(b_0 + b_2 (\bar{\epsilon}/R)^2 + b_4 (\bar{\epsilon}/R)^4 + \dots \right) \,, \tag{5.4}$$

we can fix the coefficients b_{2n} such that (5.1) is satisfied to every order. For the case of a ball in three dimensional Minkowski we find

$$\epsilon(\bar{\epsilon}) = \bar{\epsilon} \frac{\mu_1}{\bar{\mu}_1} \left(1 - \delta \ln(a_3^*) \right) + \dots \,, \tag{5.5}$$

where higher order terms are unconstrained. An analogous construction can be considered for the higher dimensional case and other setups in the CFT. This subtle aspect regarding the extended first law of entanglement has not been previously discussed in the literature.

Extended first law of entanglement for general setups. Given that we have shown that the extended first law of entanglement holds in a wide variety of setups, a natural question is whether it is valid for arbitrary CFTs, regions and states. While the ordinary first law follows from positivity of relative entropy [1] and therefore holds in full generality, the extended version can only be formulated for CFTs since the coefficient a_d^* in even dimensions is only defined for conformal theories (1.2). Although trying to directly prove the extended first law for arbitrary CFTs seems a complicated task, we can check whether the results for the entanglement entropy present in the literature are consistent with (1.1), which essentially implies $S_{EE} \propto a_d^*$ to first order in a_d^*.

Let us consider two dimensional CFTs, where a_2^* is proportional to the Virasoro central charge c. For the vacuum entropy associated to any number of disjoint intervals of a holographic CFT in Minkowski space, refs. [54–56] showed that $S_{EE} \propto a_2^*$. The same is true for a thermal state reduced to an interval [57] and analogous setups in curved backgrounds [58]. For more general situations, the entanglement entropy is only known for particular CFTs, mostly free theories. In each of these cases the entropy depends on the details of the theory in a complicated way. However, we are not aware of any result where the entanglement entropy in two dimensions is not proportional to the central charge and, consequently, in contradiction with (1.1).

For space-time dimensions larger than two, it becomes evident that the extended first law of entanglement as written in (1.1) cannot hold in full generality. The simplest example is to consider the Minkowski vacuum in $d = 4$ reduced to a cylinder. Here the entanglement

entropy is independent of a_d^* and is instead proportional to the coefficient appearing in the square of the Weyl tensor in the trace anomaly [59]. For more complicated regions the entropy is a combination of these coefficients. While this shows the extended first law as written in (1.1) cannot hold in general for $d = 4$, it suggests the following generalization might still be true[13]

$$\delta S_{EE} = \delta\langle K_B \rangle + \sum_i \frac{S_{EE}}{a_i} \delta a_i \,, \qquad (5.6)$$

where B is a region in four-dimensional Minkowski and a_i are the coefficients of the terms appearing in the trace anomaly (see for example [11]). This generalization has a better chance of applying to more general regions.

It would be interesting to understand how holography is able to capture the extended first law of entanglement in these more general cases where it is expected to hold. The $d = 2$ case stands out as the simplest one in which concrete progress might be possible, maybe using similar techniques as the ones developed in [56]. This deserves further study, in order to determine whether a general derivation of the extended first law of entanglement in this context is possible.

Bulk constraints from extended first law of entanglement. Assuming the RT holographic formula for entanglement entropy together with the ordinary first law of entanglement in the boundary, implies Einstein's bulk equations about a perturbed AdS background. What are the consequences of assuming the extended first law of entanglement instead?[14]

Let us address this question in the simplest setup of AdS$_3$/CFT$_2$, where the bulk theory is described by Einstein gravity, so that the coupling constants are $\lambda_i = (G, L)$. Let us assume (the non-trivial statement that) the extended first law of entanglement holds in the boundary CFT for arbitrary states ρ and regions B, together with the RT formula

$$\delta S_{EE} = \delta\langle K_B \rangle + \frac{S_{EE}}{c}\delta c \,, \qquad S_{EE} = \frac{A(\gamma_{\text{ext}})}{4G} \,, \qquad (5.7)$$

where γ_{ext} is an extremal bulk curve homologous to the region B at the boundary. Using that in Einstein gravity the central charge c is given by $c = 3L/2G$, the "extended" contribution of the first law of entanglement on the bulk becomes

$$\delta_{\lambda_i}\left(\frac{A(\gamma_{\text{ext}})}{4G}\right) = \frac{A(\gamma_{\text{ext}})}{4G}\delta_{\lambda_i}\ln(L/G) \qquad \Longrightarrow \qquad A(\gamma_{\text{ext}}) \propto L \,. \qquad (5.8)$$

The extended first law of entanglement translates into the statement that the length of the extremal curve on the bulk is proportional to the AdS radius L.

If the boundary state is the vacuum $|0\rangle$ the bulk metric is pure AdS$_3$, which only depends on L, and $A(\gamma_{\text{ext}}) \propto L$ immediately follows from dimensional analysis. The constraint becomes more interesting when considering excited states at the boundary, such as a thermal state $\rho(\beta)$ with inverse temperature β. In this case we can easily compute

---

[13]We thank Manus Visser for suggesting this generalization.

[14]We thank Manus Visser for suggesting this question.

$A(\gamma_{\text{ext}})$ and find the non-trivial statement $A(\gamma_{\text{ext}}) \propto L$ is indeed true [17]. For more general setups this gives a bulk constraint coming from the boundary extended first law of entanglement.

It is also interesting to consider the inverse logic. We can directly compute $A(\gamma_{\text{ext}})$ for complicated holographic setups and check whether the end result is proportional to L. This could help understand in which situations the extended first law of entanglement holds for the boundary theory.

Extended first law in a single dimension. In this work we have explored the extended first law for two-dimensional gravitational theories. While we have shown interesting results can be obtained from the gravitational perspective we have not analyzed the boundary interpretation of our calculations. In future work it would be interesting to study this further, maybe in the setup of JT gravity that it has been recently understood as a holographic description of an ensemble average of SYK models [35].

JT gravity also offers an arena to study the relation between quantum bulk effects and the extended first law of entanglement at the boundary. While in general it is very difficult to take these contributions into account, this simple setup allows for very explicit calculations in the bulk [60]. Hence, it might be possible to write down an extended first law that incorporates bulk quantum corrections. On a more speculative note, it would be interesting to investigate the extended first law in dynamical space-times, in the hope it sheds a new perspective regarding recent progress on the black hole information paradox [61, 62].

Three dimensional gravity and thermodynamic volume. For three dimensional bulk duals we have derived a modification of the extended first law (4.4) that holds for space-times that are not necessarily (globally) pure AdS, such as the BTZ black hole. In the context of extended black hole thermodynamics, we obtain a curious formula for the thermodynamic volume (4.8), which we verified gives the correct expressions found using standard means. In particular, we obtain a result for the thermodynamic volume of the BTZ black hole in a higher curvature theory of gravity (4.19).

It would be interesting to see whether the formula for the thermodynamic volume in (4.8) provides anything new to the field of extended thermodynamics. Particularly, it would be beneficial to see if it gives another microscopic viewpoint of V, along the lines of [47]. In ref. [47] it was shown that the thermodynamic volume sometimes constrains the number of available CFT states dual to AdS$_3$ gravity, revealing that the Bekenstein-Hawking entropy (given by the Cardy formula) overcounts the number of CFT degrees of freedom. This chain of reasoning provides a microscopic explanation for black hole super-entropicity, a designation for black holes whose entropy exceeds that of Schwarzschild-AdS, and violate the reverse isoperimetric inequality [18]. In three space-time dimensions, the reverse isoperimetric inequality takes the form

$$\pi V \geq 4S^2 G^2 \ . \tag{5.9}$$

When we input our expression for the volume in (4.8), the reverse isoperimetric inequality

imposes a lower bound on the L derivative of $\log(a_2^*)$,

$$\frac{\partial}{\partial L}\left[\log(a_2^*) + \log(\tilde{\mathcal{A}})\right] \geq \frac{SG}{\pi^2 L^3 T} \geq 0 \; . \tag{5.10}$$

Black holes which satisfy this inequality, e.g., rotating BTZ, are said to be sub-entropic. Super-entropic black holes, such as the charged BTZ, violate the inequality (5.9) and impose the following upper bound

$$\frac{\partial}{\partial L}\left[\log(a_2^*) + \log(\tilde{\mathcal{A}})\right] \leq \frac{SG}{\pi^2 L^3 T} \; . \tag{5.11}$$

Since a_2^* relates to the number of degrees of freedom of the dual CFT$_2$, these bounds are expected to tell us something about the availability of CFT microstates to be counted by the Cardy formula. It would be interesting to study these bounds in further detail, where $\tilde{\mathcal{A}}$ might acquire a boundary interpretation.

Acknowledgments

References

[1] D.D. Blanco, H. Casini, L.-Y. Hung and R.C. Myers, *Relative entropy and holography*, *JHEP* **08** (2013) 060 [arXiv:1305.3182] [INSPIRE].

[2] G. Wong, I. Klich, L.A. Pando Zayas and D. Vaman, *Entanglement temperature and entanglement entropy of excited states*, *JHEP* **12** (2013) 020 [arXiv:1305.3291] [INSPIRE].

[3] N. Lashkari, M.B. McDermott and M. Van Raamsdonk, *Gravitational dynamics from entanglement 'thermodynamics'*, *JHEP* **04** (2014) 195 [arXiv:1308.3716] [INSPIRE].

[4] T. Faulkner, M. Guica, T. Hartman, R.C. Myers and M. Van Raamsdonk, *Gravitation from entanglement in holographic CFTs*, *JHEP* **03** (2014) 051 [arXiv:1312.7856] [INSPIRE].

[5] D. Kastor, S. Ray and J. Traschen, *Enthalpy and the mechanics of AdS black holes*, *Class. Quant. Grav.* **26** (2009) 195011 [arXiv:0904.2765] [INSPIRE].

[6] B.P. Dolan, *The cosmological constant and the black hole equation of state*, *Class. Quant. Grav.* **28** (2011) 125020 [arXiv:1008.5023] [INSPIRE].

[7] D. Kubiznak, R.B. Mann and M. Teo, *Black hole chemistry: thermodynamics with Lambda*, *Class. Quant. Grav.* **34** (2017) 063001 [arXiv:1608.06147].

[8] D. Kastor, S. Ray and J. Traschen, *Chemical potential in the first law for holographic entanglement entropy*, *JHEP* **11** (2014) 120 [arXiv:1409.3521] [INSPIRE].

[9] S.S. Pufu, *The F-theorem and F-maximization*, *J. Phys. A* **50** (2017) 443008 [arXiv:1608.02960] [inSPIRE]

[10] H. Casini, E. Testé, and G. Torroba, *Markov property of the conformal field theory vacuum and the a theorem*, *Phys. Rev. Lett.* **118** (2017) 261602 [arXiv:1704.0187] [inSPIRE]

[11] R.C. Myers and A. Sinha, *Holographic c-theorems in arbitrary dimensions*, *JHEP* **01** (2011) 125 [arXiv:1011.5819] [inSPIRE].

[12] D. Kastor, S. Ray and J. Traschen, *Extended first law for entanglement entropy in Lovelock gravity*, *Entropy* **18** (2016) 212 [arXiv:1604.04468] [inSPIRE]

[13] E. Caceres, P.H. Nguyen and J.F. Pedraza, *Holographic entanglement chemistry*, *Phys. Rev. D* **95** (2017) 106015 [arXiv:1605.00595] [inSPIRE]

[14] S.-Q. Lan, G.-Q. Li, J.-X. Mo and X.-B. Xu, *On the first law of entanglement for quasi-topological gravity*, *Gen. Rel. Grav.* **50** (2018) 106 [arXiv:1710.01553] [inSPIRE]

[15] C. Teitelboim, *Gravitation and Hamiltonian structure in two space-time dimensions*, *Phys. Lett. B* **126** (1983) 41 [inSPIRE].

[16] R. Jackiw, *Lower dimensional gravity*, *Nucl. Phys.* **B 252** (1985) 343 [inSPIRE].

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

[18] M. Cvetič, G.W. Gibbons, D. Kubiznak and C.N. Pope, *Black hole enthalpy and an entropy inequality for the thermodynamic volume*, *Phys. Rev. D* **84** (2011) 024037 [arXiv:1012.2888] [inSPIRE].

[19] R.M. Wald, *Black hole entropy is the Noether charge*, *Phys. Rev. D* **48** (1993) 3427 [gr-qc/9307038] [inSPIRE].

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

[21] H. Casini, M. Huerta and R.C. Myers, *Towards a derivation of holographic entanglement entropy*, *JHEP* **05** (2011) 036 [arXiv:1102.0440] [inSPIRE].

[22] T. Jacobson and M. Visser, *Gravitational thermodynamics of causal diamonds in (A)dS*, arXiv:1812.01596.

[23] 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].

[24] F. Rosso, *Localized thermal states and negative energy*, *JHEP* **12** (2019) 246 [arXiv:1907.07699] [inSPIRE]

[25] R. Emparan, *AdS/CFT duals of topological black holes and the entropy of zero energy states*, *JHEP* **06** (1999) 036 [hep-th/9906040] [inSPIRE].

[26] J. Bisognano and E.H. Wichmann, *On the duality condition for quantum fields*, *J. Math. Phys.* **17** (1976) 303 [inSPIRE].

[27] W.G. Unruh, *Notes on black hole evaporation*, *Phys. Rev. D* **14** (1976) 870 [inSPIRE].

[28] T. Strobl, *Gravity in two space-time dimensions*, Ph.D. thesis, Technische Hochschule, Aachen, Germany (1999) [hep-th/0011240] [inSPIRE].

[29] A. Strominger, *AdS$_2$ quantum gravity and string theory*, *JHEP* **01** (1999) 007 [hep-th/9809027] [inSPIRE].

[30] M. Cadoni and S. Mignemi, *Asymptotic symmetries of AdS_2 and conformal group in $d = 1$*, *Nucl. Phys. B* **557** (1999) 165 [`hep-th/9902040`] [ɪɴSPIRE].

[31] T. Hartman and A. Strominger, *Central charge for AdS_2 quantum gravity*, *JHEP* **04** (2009) 026 [`arXiv:0803.3621`] [ɪɴSPIRE].

[32] A. Castro, D. Grumiller, F. Larsen and R. McNees, *Holographic description of AdS_2 black holes*, *JHEP* **11** (2008) 052 [`arXiv:0809.4264`] [ɪɴSPIRE].

[33] M. Alishahiha and F. Ardalan, *Central charge for 2D gravity on AdS_2 and AdS_2/CFT_1 correspondence*, *JHEP* **08** (2008) 079 [`arXiv:0805.1861`] [ɪɴSPIRE].

[34] M. Cvetič and I. Papadimitriou, *AdS_2 holographic dictionary*, *JHEP* **12** (2016) 008 [*Erratum ibid.* **01** (2017) 120] [`arXiv:1608.07018`] [ɪɴSPIRE]

[35] P. Saad, S.H. Shenker and D. Stanford, *JT gravity as a matrix integral*, `arXiv:1903.11115`.

[36] R.B. Mann and S.F. Ross, *The $D \to 2$ limit of general relativity*, *Class. Quant. Grav.* **10** (1993) 1405 [`gr-qc/9208004`] [ɪɴSPIRE].

[37] A.M. Frassino, R.B. Mann and J.R. Mureika, *Lower-dimensional black hole chemistry*, *Phys. Rev. D* **92** (2015) 124069 [`arXiv:1509.05481`] [ɪɴSPIRE]

[38] V. Faraoni and E. Gunzig, *Einstein frame or Jordan frame?*, *Int. J. Theor. Phys.* **38** (1999) 217 [`astro-ph/9910176`] [ɪɴSPIRE].

[39] M. Postma and M. Volponi, *Equivalence of the Einstein and Jordan frames*, *Phys. Rev. D* **90** (2014) 103516 [`arXiv:1407.6874`] [ɪɴSPIRE].

[40] J. Maldacena, D. Stanford and Z. Yang, *Conformal symmetry and its breaking in two dimensional nearly Anti-de-Sitter space*, *PTEP* **2016** (2016) 12C104 [`arXiv:1606.01857`] [ɪɴSPIRE]

[41] D. Harlow and D. Jafferis, *The factorization problem in Jackiw-Teitelboim gravity*, *JHEP* **02** (2020) 177 [`arXiv:1804.01081`] [ɪɴSPIRE]

[42] M. Bañados, C. Teitelboim and J. Zanelli, *The black hole in three-dimensional space-time*, *Phys. Rev. Lett.* **69** (1992) 1849 [`hep-th/9204099`] [ɪɴSPIRE].

[43] B.P. Dolan, *The compressibility of rotating black holes in D-dimensions*, *Class. Quant. Grav.* **31** (2014) 035022 [`arXiv:1308.5403`] [ɪɴSPIRE].

[44] C.V. Johnson, *Holographic heat engines*, *Class. Quant. Grav.* **31** (2014) 205002 [`arXiv:1404.5982`] [ɪɴSPIRE].

[45] B.P. Dolan, *Bose condensation and branes*, *JHEP* **10** (2014) 179 [`arXiv:1406.7267`] [ɪɴSPIRE].

[46] J. Couch, W. Fischler and P.H. Nguyen, *Noether charge, black hole volume, and complexity*, *JHEP* **03** (2017) 110 [`arXiv:1610.02038`] [ɪɴSPIRE]

[47] C. V. Johnson, V. L. Martin, and A. Svesko, *A microscopic description of thermodynamic volume in extended black hole thermodynamics*, `arXiv:1911.05286`.

[48] C. Martinez, C. Teitelboim and J. Zanelli, *Charged rotating black hole in three space-time dimensions*, *Phys. Rev. D* **61** (2000) 104013 [`hep-th/9912259`] [ɪɴSPIRE].

[49] E.A. Bergshoeff, O. Hohm and P.K. Townsend, *Massive gravity in three dimensions*, *Phys. Rev. Lett.* **102** (2009) 201301 [`arXiv:0901.1766`] [ɪɴSPIRE].

[50] E.A. Bergshoeff, O. Hohm and P.K. Townsend, *More on massive* 3*D gravity*, *Phys. Rev. D* **79** (2009) 124042 [`arXiv:0905.1259`] [inSPIRE].

[51] A. Sinha, *On the new massive gravity and AdS/CFT*, *JHEP* **06** (2010) 061 [`arXiv:1003.0683`] [inSPIRE].

[52] D. Marolf and A.C. Wall, *State-dependent divergences in the entanglement entropy*, *JHEP* **10** (2016) 109 [`arXiv:1607.01246`] [inSPIRE]

[53] H. Casini, M. Huerta, R.C. Myers and A. Yale, *Mutual information and the F-theorem*, *JHEP* **10** (2015) 003 [`arXiv:1506.06195`] [inSPIRE]

[54] S. Ryu and T. Takayanagi, *Aspects of holographic entanglement entropy*, *JHEP* **08** (2006) 045 [`hep-th/0605073`] [inSPIRE].

[55] T. Hartman, *Entanglement entropy at large central charge*, `arXiv:1303.6955` [inSPIRE].

[56] T. Faulkner, *The entanglement Renyi entropies of disjoint intervals in AdS/CFT*, `arXiv:1303.7221` [inSPIRE].

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

[58] J. Cardy and E. Tonni, *Entanglement Hamiltonians in two-dimensional conformal field theory*, *J. Stat. Mech.* **1612** (2016) P123103 [`arXiv:1608.01283`] [inSPIRE]

[59] S.N. Solodukhin, *Entanglement entropy, conformal invariance and extrinsic geometry*, *Phys. Lett. B* **665** (2008) 305 [`arXiv:0802.3117`] [inSPIRE].

[60] D.L. Jafferis and D.K. Kolchmeyer, *Entanglement entropy in Jackiw-Teitelboim gravity*, `arXiv:1911.10663`.

[61] G. Penington, *Entanglement wedge reconstruction and the information paradox*, `arXiv:1905.08255`.

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