

Higher-derivative corrections to entropy and the weak gravity conjecture in Anti-de Sitter space

Sera Cremonini,^a Callum R.T. Jones,^b James T. Liu^b and Brian McPeak^b

^a*Department of Physics, Lehigh University,
Bethlehem, PA, 18018, U.S.A.*

^b*Leinweber Center for Theoretical Physics, Randall Laboratory of Physics,
The University of Michigan, Ann Arbor, MI 48109-1040, U.S.A.*

E-mail: cremonini@lehigh.edu, jonescal@umich.edu, jimliu@umich.edu,
bmcpeak@umich.edu

ABSTRACT: We compute the four-derivative corrections to the geometry, extremality bound, and thermodynamic quantities of AdS-Reissner-Nordström black holes for general dimensions and horizon geometries. We confirm the universal relationship between the extremality shift at fixed charge and the shift of the microcanonical entropy, and discuss the consequences of this relation for the Weak Gravity Conjecture in AdS. The thermodynamic corrections are calculated using two different methods: first by explicitly solving the higher-derivative equations of motion and second, by evaluating the higher-derivative Euclidean on-shell action on the leading-order solution. In both cases we find agreement, up to the addition of a Casimir energy in odd dimensions. We derive the bounds on the four-derivative Wilson coefficients implied by the conjectured positivity of the leading corrections to the microcanonical entropy of thermodynamically stable black holes. These include the requirement that the coefficient of Riemann-squared is positive, meaning that the positivity of the entropy shift is related to the condition that $c - a$ is positive in the dual CFT. We discuss implications for the deviation of η/s from its universal value and a potential lower bound.

KEYWORDS: Black Holes, Effective Field Theories, AdS-CFT Correspondence, Black Holes in String Theory

ARXIV EPRINT: [1912.11161](https://arxiv.org/abs/1912.11161)

Contents

1	Introduction	1
2	Corrections to the geometry	3
2.1	The zeroth order solution	4
2.2	The first order solution	4
2.3	Asymptotic conditions and conserved quantities	6
3	Mass, charge, and entropy from the geometry shift	8
3.1	Mass, charge, and extremality	9
3.1.1	Extremality at leading order	9
3.1.2	Mass shift at fixed charge	10
3.1.3	Charge shift at fixed mass	10
3.1.4	Summary of extremality shifts	11
3.2	Wald entropy	11
3.3	Explicit results for the entropy shifts	13
4	Thermodynamics from the on-shell Euclidean action	14
4.1	Two-derivative thermodynamics	17
4.2	Four-derivative corrections to thermodynamics	19
5	Constraints from positivity of the entropy shift	21
5.1	Thermodynamic stability	23
5.2	Constraints on the EFT coefficients	24
5.3	Flat space limit	26
6	Discussion	27
6.1	Weak gravity conjecture in AdS	29
A	Entropy shifts from the on-shell action	31
A.1	AdS ₅	31
A.2	AdS ₆	32
A.3	AdS ₇	33
B	Another proof of the entropy-extremality relation	34
C	Entropy shifts of very-near extremal black holes	37
C.1	Entropy shift for heterotic black holes	39

1 Introduction

The aim of the swampland program [1] is to understand what subset of the infinite space of possible effective field theories can arise at low energies from theories of quantum gravity. This requires finding simple criteria for classifying these theories; those that admit UV completions including quantum gravity are in the landscape, and those that do not are in the swampland. Such criteria are far more useful if they are detectable in the infrared without any knowledge of the particulars of the UV completion. In practice, one fruitful source of such IR intuition is the physics of black holes, which are believed to conform to general relativity and the laws of black hole thermodynamics in the semiclassical regime.

So far a number of criteria have been proposed (see e.g. [2] for a review of the program). One that has attracted considerable interest is the Weak Gravity Conjecture (WGC), which roughly states that there are states whose mass is smaller than their charge [3] — hence, they are states for which “gravity is the weakest force.” Such states are labelled “superextremal,” and may be provided by fundamental particles or non-perturbative states such as black holes. The original motivation was to provide a mechanism for black holes to decay. These black holes are believed to obey an “extremality bound” on its mass to charge ratio,

$$\frac{M}{Q} > 1, \tag{1.1}$$

which arises from requiring that the solution does not contain a naked singularity. It was believed that if a theory of gravity did not have a superextremal particle in its spectrum, then simple conservation of charge would prevent it from decaying to lighter components without violating cosmic censorship. This, in turn, would be problematic because it would lead to an infinite number of stable states or remnants.

Another mechanism for the decay of nearly extremal black holes in flat space was pointed out in [4]. Generically, the low-energy limit of a theory of quantum gravity should include higher-derivative corrections that encode the UV physics in a highly suppressed way. Such corrections can change the allowed charge to mass ratio of Reissner-Nordström black holes so that they can decay to smaller black holes. In particular, the classical extremality bound would be corrected into a form that schematically may look like

$$\frac{M}{Q} > 1 - \frac{1}{Q^2} (2c_1 + 8c_2 + \dots), \tag{1.2}$$

shifted slightly by the Wilson coefficients c_i of the higher-derivative operators. It then becomes possible for the nearly extremal black holes to decay as long as this combination of coefficients is positive. The statement that black holes can always decay through these higher-derivative corrections is sometimes called the “Black Hole Weak Gravity Conjecture.” This idea has inspired a large amount of work [5–11] on bounding the EFT coefficients, thereby proving the conjecture. While no existing proof is completely general, the work so far covers a large number of possibilities and assumptions.

One intriguing proof of the WGC in flat space relates the extremality shift to the shift in the Wald entropy [7]. The authors first show that, near extremality, the shift

to the extremality bound at fixed charge and temperature is proportional to the shift in entropy at fixed charge and mass. They then present an argument that the higher-derivative corrections should increase the entropy, thereby proving the Black Hole WGC. The argument for the entropy shift positivity is not expected to be fully general; it applies to higher-derivative corrections that arise from integrating out massive particles at tree-level. Nonetheless, it is curious that the entropy shift is proportional to the extremality shift. This fact was given a simple thermodynamic proof in [12], where no assumptions were made about the particulars of the background.

So far, however, these ideas have not fully made their way to Anti-de Sitter space. From the WGC point of view, it is easy to see why: the relationship between mass and charge of an extremal black holes in AdS is *already non-linear at the two-derivative level*.¹ Therefore it is not at all clear what is gained by studying the higher-derivative corrections to the extremal mass-to-charge ratio.² Furthermore, massive particles emitted from a black hole cannot fly off to infinity in AdS as they can in flat space, so if the WGC allows for the instability of black holes in AdS, it must be through a completely different mechanism. (See the Discussion for more commentary on this possibility).

Regardless, the entropy-extremality relationship is expected to hold in AdS as it does in flat space (and indeed, an example in AdS₄ was given in [12]). Therefore, this paper addresses two main issues in Anti-de Sitter space. First, we check the purported relationship between the entropy shift and the extremality shift, and indeed we find that it holds for the AdS-Reissner-Nordström backgrounds. Second, we examine the conjecture that the entropy shift is positive when the leading-order solution is a minimum of the action. By computing the entropy shift explicitly, we see that its positivity for stable black holes implies that the coefficient of Riemann-squared is *universally positive*. This has interesting consequences for the structure of η/s , as we comment on in section 6.

This paper is organized as follows. In section 2 we introduce the theory we will examine, and find the solutions to the equations of motion at first order in the EFT coefficients. In section 3 we use the solution to compute the shift to extremality, considering the result both at fixed charge and fixed mass. We then compute the shift to the Wald entropy, and we find that the conjectured relationship of [7, 12] between the shift to mass and shift to entropy is valid for AdS-Reissner-Nordström black holes. Furthermore, we notice that both the mass shift and entropy shift are also proportional to the charge shift, which allows us to extend the relationship to³

$$(\Delta M)_{Q,T=0} = -T_0 (\Delta S)_{Q,M} = -\Phi_0 (\Delta Q)_{M,T=0} . \tag{1.3}$$

We present a simple thermodynamic derivation of these relationships in appendix B.

In section 4, we reproduce these results from a thermodynamic point of view. It has recently been shown [17] that the first-order corrections to the solutions are not needed

¹By “extremal,” we mean that the temperature is zero. This is not the same as the BPS limit in AdS.

²Other aspects of the WGC have been discussed in AdS. See e.g. [13–16].

³This form of the entropy-extremality relation is valid away from extremality, but a slightly modified form is required for black holes which are very close to extremality. The subtleties involved in this limit are discussed in detail at appendix C.

to compute the first order corrections to thermodynamic quantities. In this section, we verify that this is the case for AdS-Reissner-Nordström backgrounds by computing the four-derivative corrections to the renormalized on-shell action. From this we may compute the free energy and other thermodynamic quantities. We find that the results of this calculation match the results from section 3 in even dimensions, while in odd dimensions the free energy and associated thermodynamic quantities are renormalization scheme dependent, and agree with the geometric calculation in a physically motivated *zero Casimir* scheme.

In section 5, we first review the argument given in [7] for the positivity of the entropy shift, and comment on a potential issue with applying it to AdS. The positivity of the entropy shift requires that the black hole solutions are local minima of the path integral, so we compute the specific heat and electrical permittivity to determine the regions of parameter space where the black holes will be stable. Finally, we determine the constraints placed on the EFT coefficients by assuming that the entropy shift is positive for all stable black holes. We compare this to the results obtained by requiring that the entropy shift is positive for only extremal black holes, which is equivalent to the condition that the extremality shift at fixed charge and temperature is positive. The constraints include the requirement that the coefficient of Riemann-squared is positive. As this coefficient is proportional to the difference $c - a$ between the central charges of the dual CFT, we conclude that the positivity of the entropy shift will be violated in theories where $c - a < 0$.

We summarize our results in the section 6, where we comment on the implications of our results for the behavior of η/s , as well as the nature of the WGC in AdS. We relegate to appendix A the specific form of the entropy shifts and bounds on the EFT coefficients for AdS₅ through AdS₇. In appendix B, we present a general proof of the entropy-extremality relation of [12], and in appendix C we comment on some subtleties involving the extremal limit.

Note added: after completing this work, we noticed that the revised version of [12] obtains the same form of the entropy-extremality relation that we argue for in this paper.

2 Corrections to the geometry

We consider Einstein-Maxwell theory in the presence of a negative cosmological constant in a $(d + 1)$ -dimensional AdS spacetime of size l . The first non-trivial terms in the derivative expansion of the effective action arise at the four-derivative level, and by appropriate field redefinitions we may choose a complete basis of dimension-independent operators:

$$\begin{aligned}
 I = -\frac{1}{16\pi} \int d^{d+1}x \sqrt{-g} & \left[\frac{d(d-1)}{l^2} + R - \frac{1}{4}F^2 \right. \\
 & \left. + l^2 \epsilon \left(c_1 R_{abcd} R^{abcd} + c_2 R_{abcd} F^{ab} F^{cd} + c_3 (F^2)^2 + c_4 F^4 \right) \right]. \tag{2.1}
 \end{aligned}$$

Note that additional CP-odd terms can arise in specific dimensions, but will not contribute to the static, stationary spherically symmetric black holes that we are considering here. This basis parallels that of [18], which used the same set of dimensionless Wilson coefficients, but focused on the $(4 + 1)$ -dimensional case. Depending on the origin of the AdS

length scale l , one may expect these coefficients to be parametrically small, of the form $c_i \sim (\Lambda l)^{-2}$, where Λ denotes the scale at which the EFT breaks down. In particular, this will be the case in order for the action (2.1) to be under perturbative control. We have also introduced the small bookkeeping parameter ϵ , which will allow us to keep track of which terms are first order in the c_i coefficients.

2.1 The zeroth order solution

At the two-derivative level, this action admits a family of AdS-Reissner-Nordström black holes parametrized by uncorrected mass m and charge q ,

$$\begin{aligned}
 ds^2 &= -f(r)dt^2 + g(r)^{-1}dr^2 + r^2 d\Omega_{d-1,k}^2, & f(r) &= g(r) = k - \frac{m}{r^{d-2}} + \frac{q^2}{4r^{2d-4}} + \frac{r^2}{l^2}, \\
 A &= \left(-\frac{1}{c} \frac{q}{r^{d-2}} + \Phi \right) dt, & c &= \sqrt{\frac{2(d-2)}{(d-1)}}, & \Phi &= \frac{1}{c} \frac{q}{r_h^{d-2}}. \quad (2.2)
 \end{aligned}$$

Here r_h is the outer horizon radius, and the parameter $k = 0, \pm 1$ specifies the horizon geometry, with $k = 1$ corresponding to the unit sphere. The constant Φ is chosen so that the A_t component of the gauge field vanishes on the horizon, and represents the potential difference between the asymptotic boundary and the horizon.

Typically, we will consider lower case letters (m, q, \dots) to be parameters in the theory, while upper case letters (M, Q, S, T, \dots) will denote physical quantities that may or may not receive corrections. We will add a subscript zero (e.g. M_0) to denote the uncorrected contribution to quantities that do receive order c_i corrections. The shifts, which are equal to the corrected quantities minus the uncorrected ones, will be denoted by the ϵ derivative. However, we will sometimes use Δ when it is convenient, with subscripts indicating quantities held fixed, for example, we have

$$(\Delta M)_T \equiv \lim_{\epsilon \rightarrow 0} (M(T, \epsilon) - M_0(T)) \equiv \lim_{\epsilon \rightarrow 0} \left(\frac{\partial M}{\partial \epsilon} \right)_T. \quad (2.3)$$

Finally, in sections 4 and 5 we will use dimensionless quantities (ν, ξ) for convenience. These are defined by $\nu = (r_h)_0/l$ and $Q = (1 - \xi)Q_{\text{ext}}$.

2.2 The first order solution

We now turn to the first order solution in terms of the Wilson coefficients c_i . We follow the procedure outlined in ref. [19], but work in an AdS_{d+1} background. While general $(d+1)$ -dimensional results may be worked out analytically, we took a shortcut of working with explicit dimensions four through eight and then fitting the coefficients to extract results for arbitrary dimension. Since the four-derivative terms are built from tensors with eight indices and hence four metric contractions, the resulting expressions will scale at most as d^4 . The coefficients are hence fully determined by working in five different dimensions.

Following [19], we start with the *effective* stress tensor, where corrections come from two sources. The first is from substituting in the corrected Maxwell field to the zeroth order electromagnetic stress tensor, and the second is from the explicit four-derivative corrections

to the stress tensor evaluated on the zeroth order solution. The result of computing both of these contributions to the time-time component of the stress tensor is

$$\begin{aligned}
T_t^t = & -\frac{(d-1)(d-2)q^2}{4r^{2d-2}} + \frac{d(d-1)}{l^2} \\
& + c_1 \left(\frac{(d-2)(8d^3 - 24d^2 + 15d + 3)q^4 l^2}{8r^{4d-4}} - \frac{(d-1)(d-2)(4d^2 - 9d + 3)mq^2 l^2}{r^{3d-2}} \right. \\
& \quad + k \frac{4d(d-1)(d-2)^2 l^2 q^2}{r^{2d}} - \frac{d(d-1)(d-2)(d-3)l^2 m^2}{r^{2d}} \\
& \quad \left. + \frac{(d-2)(2d-3)(2d^2 - 5d + 1)q^2}{r^{2d-2}} + \frac{2d(d-3)}{l^2} \right) \\
& + c_2 \left(\frac{(d-1)^3(d-2)q^4 l^2}{r^{4d-4}} - \frac{(d-1)^2(3d^2 - 8d + 4)q^2 m l^2}{r^{3d-2}} \right. \\
& \quad \left. + k \frac{2d(d-1)^2(d-2)q^2 l^2}{r^{2d}} + \frac{2(d-1)^3(d-2)q^2}{r^{2d-2}} \right) \\
& + (2c_3 + c_4) \left(\frac{(d-1)^2(d-2)^2 q^4 l^2}{2r^{4d-4}} \right). \tag{2.4}
\end{aligned}$$

The shift to the geometry may be obtained from the corrections to the stress tensor [19],

$$\Delta g = \frac{1}{(d-1)r^{d-2}} \int dr r^{d-1} \Delta T_t^t, \tag{2.5}$$

and after integrating the $\mathcal{O}(c_i)$ terms in (2.4), we find

$$\begin{aligned}
\Delta g(r) = & c_1 \left(-\frac{(d-2)(8d^3 - 24d^2 + 15d + 3)q^4 l^2}{8(d-1)(3d-4)r^{4d-6}} + \frac{(d-2)(4d^2 - 9d + 3)mq^2 l^2}{2(d-1)r^{3d-4}} \right. \\
& \quad - k \frac{4(d-2)^2 l^2 q^2}{r^{2d-2}} + \frac{(d-2)(d-3)l^2 m^2}{r^{2d-2}} \\
& \quad \left. - \frac{(2d-3)(2d^2 - 5d + 1)q^2}{(d-1)r^{2d-4}} + \frac{2(d-3)r^2}{(d-1)l^2} \right) \\
& + c_2 \left(-\frac{(d-1)^2(d-2)q^4 l^2}{(3d-4)r^{4d-6}} + \frac{(3d^2 - 8d + 4)q^2 m l^2}{2r^{3d-4}} \right. \\
& \quad \left. - k \frac{2(d-1)(d-2)q^2 l^2}{r^{2d-2}} - \frac{2(d-1)^2 q^2}{r^{2d-4}} \right) \\
& + (2c_3 + c_4) \left(-\frac{(d-1)(d-2)^2 q^4 l^2}{(6d-8)r^{4d-6}} \right). \tag{2.6}
\end{aligned}$$

The time component of the metric can then be obtained using the relation [19]

$$f(r) = (1 + \gamma(r))g(r), \tag{2.7}$$

where $\gamma(r)$ is defined by⁴

$$\gamma(r) = -\frac{1}{(d-2)} \int dr r (T_t^t - T_r^r). \tag{2.8}$$

⁴We note that the definition of γ implies that it is positive provided that the null energy condition holds.

For our particular case we find:

$$\gamma(r) = \left(c_1 \frac{(d-2)(2d^2 - 5d + 1)}{(d-1)} + c_2 d(d-2) \right) \frac{q^2 l^2}{r^{2d-2}}. \quad (2.9)$$

Finally, we have

$$F_{tr} = \sqrt{\frac{(d-2)(d-1)}{2}} \left[(1 - 8c_2) \frac{q}{r^{d-1}} + 4c_2(d-1)(d-2) \frac{qml^2}{r^{2d-1}} \right. \\ \left. + \left(\frac{c_1(2d^2 - 5d + 1)}{2} - \frac{c_2(7d - 12) - 4(2c_3 + c_4)(d-1)}{2} \right) (d-2) \frac{q^3 l^2}{r^{3d-3}} \right], \quad (2.10)$$

which we note is independent of the geometry parameter k , as was the case in [20].

2.3 Asymptotic conditions and conserved quantities

The first order solution can be summarized as

$$ds^2 = -(1 + \gamma(r))g(r)dt^2 + g(r)^{-1}dr^2 + r^2 d\Omega_{d-1,k}^2, \quad (2.11)$$

where

$$g(r) = k - \frac{m}{r^{d-2}} + \frac{q^2}{4r^{2d-4}} + \frac{r^2}{l^2} + \Delta g. \quad (2.12)$$

The corrected metric functions, Δg and $\gamma(r)$, are given in (2.6) and (2.8), respectively. In addition, the full electric field is given in (2.10). For a given zeroth order AdS radius l , this solution is specified by two parameters, m and q , which correspond to the mass and charge of the uncorrected black hole. At the same time, the corrected solution includes a number of integration constants, two of which we have implicitly set to zero in the integral expressions for Δg and $\gamma(r)$. The constant related to Δg can be absorbed by a shift in m , and a third constant from the corrected Maxwell equation can be absorbed by a shift in q . The constant related to $\gamma(r)$ can be absorbed at the linearized level by a rescaling of the time coordinate, and hence can be thought of as a redshift factor.

In order to make the correspondence between the parameters of the solution, m and q , and the physical mass and charge of the black hole more precise, consider the part of Δg that is leading in r . We can see that there is a term that goes like $c_1 \frac{r^2}{l^2}$ that dominates over all other terms in the correction. Therefore, for large values of r , the solution takes the form

$$f(r) \approx g(r) = k - \frac{m}{r^{d-2}} + \left(1 + c_1 \frac{2(d-3)}{d-1} \right) \frac{r^2}{l^2} + \dots, \\ F_{tr} = \sqrt{\frac{(d-2)(d-1)}{2}} (1 - 8c_2) \frac{q}{r^{d-1}} + \dots. \quad (2.13)$$

Our first observation is that the AdS radius gets modified because the Riemann-squared term is non-vanishing on the original uncorrected background. This suggests that we define an effective AdS radius

$$l^2 = \lambda^2 l_{\text{eff}}^2, \quad \lambda^2 = \left(1 + c_1 \frac{2(d-3)}{d-1} \right). \quad (2.14)$$

This shift by λ is unavoidable when turning on the c_1 Wilson coefficient. However, in principle we still have a choice of whether we hold l or l_{eff} fixed when turning on the four-derivative corrections.

In what follows, we always choose to keep l fixed. Then, since the effective AdS radius is shifted, the asymptotic form of the metric is necessarily modified as well. From a holographic point of view, this leads to a modification of the boundary metric

$$ds^2 \sim r^2 \left(\frac{dt^2}{l^2} + d\Omega_{d-1,k}^2 \right) \quad \longrightarrow \quad ds^2 \sim r^2 \left(\frac{dt^2}{l_{\text{eff}}^2} + d\Omega_{d-1,k}^2 \right). \quad (2.15)$$

This is generally undesirable, as we would like to compare thermodynamic quantities in a framework where we hold the boundary metric fixed while turning on the Wilson coefficients. One way to avoid this shift in the boundary metric is to introduce a ‘redshift’ factor

$$t = \bar{t}/\lambda, \quad (2.16)$$

to compensate for the shift in l_{eff} . In terms of the time \bar{t} , the solution now takes the form

$$\begin{aligned} ds^2 &= -\bar{f}(r) d\bar{t}^2 + g(r)^{-1} dr^2 + r^2 d\Omega_{d-1,k}^2, \\ F_{\bar{t}r} &= \lambda^{-1} F_{tr} = \sqrt{\frac{(d-2)(d-1)}{2}} (1 - 8c_2) \frac{q/\lambda}{r^{d-1}} + \dots, \end{aligned} \quad (2.17)$$

where

$$\begin{aligned} \bar{f}(r) &= \lambda^{-2} (1 + \gamma(r)) g(r) = k/\lambda^2 - \frac{m/\lambda^2}{r^{d-2}} + \frac{r^2}{l^2} + \dots, \\ g(r) &= k - \frac{m}{r^{d-2}} + \frac{r^2}{l_{\text{eff}}^2} + \dots. \end{aligned} \quad (2.18)$$

We now turn to the charge and mass of the solution measured with respect to the redshifted \bar{t} time. For the charge Q , we take the conserved Noether charge

$$Q = \frac{1}{16\pi} \int_{\Sigma_{d-1}} *F, \quad (2.19)$$

where F is the effective electric field

$$\mathcal{F}_{\mu\nu} = F_{\mu\nu} + l^2 \left(-4c_2 R_{\mu\nu\rho\sigma} F^{\rho\sigma} - 8c_3 F_{\mu\nu} (F^2) - 8c_4 F_{\nu\rho} F^{\rho\sigma} F_{\sigma\mu} \right). \quad (2.20)$$

The result is

$$Q = \frac{1 + 8c_2}{16\pi} \omega_{d-1} \lambda r^{d-1} F_{\bar{t}r} \Big|_{r \rightarrow \infty} = \sqrt{\frac{(d-2)(d-1)}{2}} \frac{\omega_{d-1}}{16\pi} q, \quad (2.21)$$

where ω_{d-1} is the volume of the unit S^{d-1} . The $1/16\pi$ factor arises from the prefactor in the action (2.1) where we have set Newton’s constant $G = 1$.

Unlike in the asymptotically Minkowski case, some care needs to be taken in obtaining the mass of the black hole. With an eye towards holography, we choose to define the mass

from the boundary stress tensor [21]. The standard approach to holographic renormalization involves the addition of appropriate local boundary counterterms so as to render the action finite. This was performed in [20] for R^2 -corrected bulk actions, and since only the $c_1 R_{abcd} R^{abcd}$ term in (2.1) leads to an additional divergence, we can directly use the result of [20]. The result is

$$M = \frac{\omega_{d-1}}{16\pi} (1 + 4c_1(d-3)) \frac{(d-1)m}{\lambda}, \tag{2.22}$$

where we have taken into account the scaling of the mass by the redshift factor λ . Substituting in λ from (2.14) then gives

$$M = \frac{\omega_{d-1}}{16\pi} (d-1)(1+\rho)m, \tag{2.23}$$

where

$$\rho = c_1 \frac{(d-3)(4d-5)}{d-1}. \tag{2.24}$$

Note that we are taking the mass here to *exclude* the Casimir energy that is normally part of the boundary stress tensor. This will be important when comparing with the thermodynamic quantities extracted from the regulated on-shell action in section 4. Working in the setup of holographic renormalization ensures that the mass M and charge Q defined in (2.23) and (2.21), respectively, yield a consistent framework for black hole thermodynamics.

3 Mass, charge, and entropy from the geometry shift

Given the first-order solution, we now consider shifts to the mass, ΔM , and entropy, ΔS , of the black hole induced by the four-derivative corrections. In these computations it is important to keep in mind what is being held fixed as we turn on the Wilson coefficients c_i . The main parameters we consider here are the mass M and charge Q , which are related to the two parameters, m and q , of the solution by (2.23) and (2.21), respectively. In addition we consider the thermodynamic quantities T (temperature) and S (entropy), although they are not all independent. Note that we always consider the AdS radius l to be fixed, although interesting results have been obtained by mapping it to thermodynamic pressure.

Singly-charged, non-rotating black holes may be described by any two of mass M , charge Q and the horizon radius r_h . Of course, any number of other parameters may be used as well, such as the temperature T or an extremality parameter, such as was used in [7]. If we further impose the extremality condition $T = 0$ on the solution, then only a single parameter is needed. Clearly this is only true for non-rotating black holes with a single gauge field, as more general solutions may have additional charges or angular momenta.

Here we mainly focus on the effect of higher-derivative corrections on extremal or near extremal black holes. In particular, we consider the extremality shift $\Delta(M/Q)$ and the entropy shift ΔS . However, it is important to keep in mind what is being held fixed when we turn on the higher-derivative corrections, as the results will depend on this choice. For example, we will see below that the shift to M/Q depends on whether the mass, charge or horizon radius is held fixed when comparing the corrected with uncorrected quantities.

3.1 Mass, charge, and extremality

Recall that, in our first-order solution, the geometry is essentially given by the radial function

$$g^{rr} = g(r) = k - \frac{m}{r^{d-2}} + \frac{q^2}{4r^{2d-4}} + \frac{r^2}{l^2} + \Delta g, \quad (3.1)$$

where Δg denotes the contributions of the higher-derivative corrections to the geometry, and ϵ is a small parameter we use to keep track of where $\mathcal{O}(c_i)$ corrections come in. Using the fact that both $g(r_h)$ and $g'(r_h)$ vanish at extremality, we may express the extremal mass and charge as a function of the horizon radius,

$$\begin{aligned} M_{\text{ext}} &= 2V(d-1)r_h^{d-2} \left(\left(k + \frac{d-1}{d-2} \frac{r_h^2}{l^2} \right) (1 + \epsilon\rho) + \epsilon \Delta g + \frac{r_h}{2(d-2)} \epsilon \Delta g' \right), \\ Q_{\text{ext}}^2 &= 2V^2(d-1)(d-2)r_h^{2(d-2)} \left(k + \frac{d}{d-2} \frac{r_h^2}{l^2} + \epsilon \Delta g + \frac{r_h}{d-2} \epsilon \Delta g' \right), \end{aligned} \quad (3.2)$$

where M and Q are the asymptotic quantities defined in (2.23) and (2.21), and we have defined $V = \omega_{d-1}/16\pi$. Though we have expressed M and Q as functions of r_h , these expressions are valid regardless of which of the three quantities is being held fixed. For example, if we work at fixed charge, then Q gets no $\mathcal{O}(\epsilon)$ corrections, in which case M and r_h will both receive corrections.

3.1.1 Extremality at leading order

Before discussing the extremality and entropy shifts, we consider the leading order relations between M_0 , Q_0 and $(r_h)_0$ for extremal black holes. We will suppress the 0 subscripts in this subsection, but we mean the uncorrected quantities. Setting $\epsilon = 0$ in (3.2) immediately gives the relations

$$\begin{aligned} M_{\text{ext}} &= 2V(d-1)r_h^{d-2} \left(k + \frac{d-1}{d-2} \frac{r_h^2}{l^2} \right), \\ Q_{\text{ext}}^2 &= 2V^2(d-1)(d-2)r_h^{2(d-2)} \left(k + \frac{d}{d-2} \frac{r_h^2}{l^2} \right). \end{aligned} \quad (3.3)$$

In principle, we can eliminate r_h from these equations to obtain the relation between mass and charge for extremal AdS black holes. However, for general dimension d , there is no simple expression that directly encodes this relation. Nevertheless, we can consider the limit of small and large black holes.

For small black holes ($r_h \ll l$), we take $k = 1$ (i.e. a spherical horizon) and find

$$M_{\text{ext}} \sim Q_{\text{ext}} \sim r_h^{d-2}, \quad (3.4)$$

so one recovers the simple $M \sim Q$ scaling that appears in flat space. (Note that asymptotically Minkowski black holes necessarily have spherical horizons.) For large black holes ($r_h \gg l$), on the other hand, the scaling is very different from that of flat space,

$$M_{\text{ext}} \sim r_h^d, \quad Q_{\text{ext}} \sim r_h^{d-1} \quad \Rightarrow \quad M_{\text{ext}} \sim (Q_{\text{ext}})^{\frac{d}{d-1}}. \quad (3.5)$$

In fact, this is precisely the scaling behavior expected based on the relationship between minimal scaling dimension and charge for boundary operators with large global charges [22].

3.1.2 Mass shift at fixed charge

Now we consider the effect of four-derivative corrections. If we hold the charge fixed, then the shift to extremality is entirely due to the change in the mass. This may be computed from the expression (3.2) for the mass by taking a derivative with respect to ϵ , which parametrizes the higher-derivative corrections, leading to

$$\begin{aligned} \left(\frac{\partial M}{\partial \epsilon}\right)_{Q,T=0} &= V(d-1)r_h^{d-2} \left(2\Delta g + \frac{1}{d-2}r_h\Delta g' + 2\rho \left(k + \frac{d-1}{d-2}\frac{r_h^2}{l^2} \right) \right. \\ &\quad \left. + \frac{2}{(d-2)r_h} \left((d-2)^2k + d(d-1)\frac{r_h^2}{l^2} \right) \left(\frac{\partial r_h}{\partial \epsilon} \right) \right), \end{aligned} \quad (3.6)$$

where we have taken into account the fact that when the charge is fixed, we must allow the horizon radius r_h to vary with ϵ . To compute the shift $\partial r_h/\partial \epsilon$, we use the fact that we are holding Q fixed. Then we use the expression for Q_{ext} in (3.2) and demand that $(\partial Q/\partial \epsilon)_{T=0} = 0$ to obtain an equation for $\partial r_h/\partial \epsilon$. This procedure leads to the rather simple result

$$\left(\frac{\partial M}{\partial \epsilon}\right)_{Q,T=0} = V(d-1)r_h^{d-2} \left(\Delta g + 2\rho \left(k + \frac{d-1}{d-2}\frac{r_h^2}{l^2} \right) \right). \quad (3.7)$$

Note that the dependence on $\Delta g'$ has vanished. From the geometric point of view, this non-trivial cancellation is crucial for the extremality-entropy relation to hold.

3.1.3 Charge shift at fixed mass

If we instead hold the mass fixed, the entire shift in the extremality is due to the shift in charge. Following the same procedure as in the fixed charge case, but this time demanding $\partial M_{\text{ext}}/\partial \epsilon = 0$, we find the relation:

$$\left(\frac{\partial Q^2}{\partial \epsilon}\right)_{M,T=0} = -2V^2(d-1)(d-2)r_h^{2d-4} \left(\Delta g + 2\rho \left(k + \frac{d-1}{d-2}\frac{r_h^2}{l^2} \right) \right). \quad (3.8)$$

Here we also find a cancellation of all $\Delta g'$ terms. Moreover, this shift is proportional to the mass shift at fixed charge

$$\left(\frac{\partial Q^2}{\partial \epsilon}\right)_{M,T=0} = -2V(d-2)r_h^{d-2} \left(\frac{dM}{d\epsilon}\right)_{Q,T=0}. \quad (3.9)$$

This relationship is more clear when we write this as the shift of Q rather than Q^2 . Using $\Delta Q^2 = 2Q\Delta Q$, we find

$$Q \left(\frac{\partial Q}{\partial \epsilon}\right)_{M,T=0} = -V(d-2)r_h^{d-2} \left(\frac{\partial M}{\partial \epsilon}\right)_{Q,T=0}. \quad (3.10)$$

Finally, we use $\Phi = Q/(d-2)Vr^{d-2}$ to write:

$$\left(\frac{\partial M}{\partial \epsilon}\right)_{Q,T=0} = -\Phi \left(\frac{\partial Q}{\partial \epsilon}\right)_{M,T=0}. \quad (3.11)$$

So we see that the mass shift is related to the charge shift times the potential. In appendix B, we derive this statement for a general thermodynamic system and show that it holds for any extensive charge and its conjugate.

One physical consequence of this fact is that the entropy-extremality relationship (with a different proportionality factor) will hold regardless of whether the mass or charge is held fixed. As far as we know, this has not been noticed before in the literature.

3.1.4 Summary of extremality shifts

The shifts to extremality may be obtained from these mass and charge shifts. For completeness, we also present calculation at fixed horizon radius, as this extremality shift has previously been considered in the literature as well [18, 20],

$$\begin{aligned}
 \left(\frac{M}{Q}\right)_{Q,T=0} &= \left(\frac{M}{Q}\right)_0 \left(1 + \rho + \Delta g \frac{1}{2\left(k + \frac{d-1}{d-2} \frac{r_h^2}{l^2}\right)}\right), \\
 \left(\frac{M}{Q}\right)_{M,T=0} &= \left(\frac{M}{Q}\right)_0 \left(1 + \rho \frac{k + \frac{d-1}{d-2} \frac{r_h^2}{l^2}}{k + \frac{d}{d-2} \frac{r_h^2}{l^2}} + \Delta g \frac{1}{2\left(k + \frac{d}{d-2} \frac{r_h^2}{l^2}\right)}\right), \\
 \left(\frac{M}{Q}\right)_{r_h,T=0} &= \left(\frac{M}{Q}\right)_0 \left(1 + \rho + \frac{\Delta g \left(k + \frac{d+1}{d-2} \frac{r_h^2}{l^2}\right) + r_h \Delta g' \frac{1}{(d-2)^2} \frac{r_h^2}{l^2}}{2\left(k + \frac{d-1}{d-2} \frac{r_h^2}{l^2}\right) \left(k + \frac{d}{d-2} \frac{r_h^2}{l^2}\right)}\right),
 \end{aligned} \tag{3.12}$$

where the corrections are encoded in ρ and Δg given in (2.24) and (2.6), respectively (and $\Delta g'$ as well for the fixed r_h case). For these final results, we have set $\epsilon = 1$. However, the expressions are only valid to first order in the Wilson coefficients c_i . Here the uncorrected charge to mass ratio may be obtained from (3.3), and takes the form

$$\left(\frac{M}{Q}\right)_0 = \sqrt{\frac{2(d-1)}{d-2}} \frac{k + \frac{d-1}{d-2} \frac{r_h^2}{l^2}}{\sqrt{k + \frac{d}{d-2} \frac{r_h^2}{l^2}}}. \tag{3.13}$$

Note that, in (3.12), the horizon radius r_h may be taken to be the uncorrected radius, and can be obtained from either M or Q using the leading order expressions (3.3). In (3.13), the leading order expression for r_h should be used. Finally, note that Δg depends on the parameters m and q as well as the radius r . The m and q parameters are directly obtained from M and Q using (2.23) and (2.21), and again the leading order horizon radius can be used in Δg .

3.2 Wald entropy

We now compare the shift in mass at fixed charge and temperature to the shift in entropy at fixed mass and charge. The entropy for black holes in higher-derivative theories is given by the Wald entropy [23]:

$$S = -2\pi \int_{\Sigma} \frac{\delta \mathcal{L}}{\delta R_{\mu\nu\rho\sigma}} \epsilon_{\mu\nu} \epsilon_{\rho\sigma}. \tag{3.14}$$

For spherically symmetric backgrounds, the integral over the horizon Σ gives a factor of the area A . The two-derivative contribution to the entropy is simply $S^{(2)} = A/4$, while the four-derivative terms yield

$$S^{(4)} = -2\pi A \frac{\delta \Delta \mathcal{L}}{\delta R_{\mu\nu\rho\sigma}} \epsilon_{\mu\nu} \epsilon_{\rho\sigma} \Big|_{\partial^4} = -\frac{A}{4} l^2 (4c_1 R_{trtr} + 2c_2 F_{tr} F_{tr}). \quad (3.15)$$

The total entropy is the sum of these terms,

$$S = \frac{A}{4} \left(1 - \epsilon (4c_1 l^2 R_{trtr} + 2c_2 l^2 F_{tr} F_{tr}) \right) \Big|_{r_h}, \quad (3.16)$$

where we once again introduced ϵ to parametrize the expansion. Here the horizon area is given by $A = \omega_{d-1} r_h^{d-1}$, where r_h is the corrected horizon radius. On the other hand, the R_{trtr} and $F_{tr} F_{tr}$ terms need only be computed on the zeroth-order background,

$$\begin{aligned} R_{trtr} &= \frac{1}{l^2} + \frac{(2d-3)(Q/V)^2}{2(d-1)r^{2d-2}} - \frac{(d-2)M/V}{2r^d}, \\ F_{tr} F_{tr} &= \frac{(Q/V)^2}{r^{2d-2}}. \end{aligned} \quad (3.17)$$

It does not matter whether we use the corrected or uncorrected quantities here because they already show up in a term that is order ϵ . Note also that, while the expression for the Wald entropy (3.16) is given in terms of M , Q and r_h of the fully corrected solution, only two of these quantities are independent.

We now examine the entropy shift for a given solution at fixed mass M and charge Q . For the moment, we work at arbitrary M and Q , and not necessarily at extremality. The general expression for the entropy shift is then

$$\left(\frac{\partial S}{\partial \epsilon} \right)_{Q,M} = \frac{A}{4} \left((d-1) \left(\frac{1}{r_h} \frac{\partial r_h}{\partial \epsilon} \right)_{Q,M} - (4c_1 l^2 R_{trtr} + 2c_2 l^2 F_{tr} F_{tr}) \right), \quad (3.18)$$

where the first term was obtained by

$$\frac{1}{A} \frac{\partial A}{\partial \epsilon} = (d-1) \frac{1}{r_h} \frac{\partial r_h}{\partial \epsilon}. \quad (3.19)$$

Here, it is important to note that the horizon radius r_h receives a correction when working at fixed M and Q . If, on the other hand, we were to keep the horizon radius fixed (as is done in [18]), we would find only the second (interaction) term in (3.18), and the entropy shift would be independent of c_3 and c_4 .

To compute $\partial r_h / \partial \epsilon$, we start with the horizon condition $g(r_h) = 0$ where $g(r)$ is given by (3.1) with m and q rewritten in terms of M and Q . Taking a derivative and solving for $\partial r_h / \partial \epsilon$ then gives

$$\frac{1}{r_h} \frac{\partial r_h}{\partial \epsilon} = -\frac{\rho M + V(d-1)r_h^{d-2} \Delta g}{(d-2)(M - (M_{\text{ext}})_0)}. \quad (3.20)$$

where $(M_{\text{ext}})_0$ is the leading order extremal mass given in (3.3). As we can see, this expression diverges if the leading order solution is extremal. This is in fact not a surprise, as

leading order extremality implies a double root at the horizon. The higher order corrections will lift this double root and hence cannot be parametrized as a linear shift in ϵ . The correct shift will be proportional to $\sqrt{\epsilon}$; however, this will not affect the bounds implied on the EFT coefficients. For a discussion of the entropy shift of very-near extremal black holes, see appendix C.

In order to avoid the divergence, we can instead consider a leading order solution taken slightly away from extremality. As long as we are sufficiently close to extremality, the first term in (3.18) will dominate the entropy shift. Noting further that, at extremality, the numerator of (3.20) becomes proportional to the mass shift (3.7) at fixed charge, we can rewrite (3.18) as

$$\left(\frac{\partial S}{\partial \epsilon}\right)_{Q,M} = -\frac{A}{4} \left(\frac{d-1}{(d-2)(M-(M_{\text{ext}})_0)} \left(\frac{\partial M}{\partial \epsilon}\right)_{Q,T=0} + \frac{d-1}{d-2} \rho + 4c_1 l^2 R_{trtr} + 2c_2 l^2 F_{tr} F_{tr} \right). \quad (3.21)$$

The deviation away from extremality can be written in terms of the leading order temperature,

$$4\pi T_0 = |g'((r_h)_0)|_{\epsilon=0} = \frac{(d-2)(M-(M_{\text{ext}})_0)}{V(d-1)(r_h)_0^{d-1}}. \quad (3.22)$$

The total shift to the entropy is then given by

$$\left(\frac{\partial S}{\partial \epsilon}\right)_{Q,M} = -\frac{1}{T_0} \left(\frac{\partial M}{\partial \epsilon}\right)_{Q,T=0} - \frac{A}{4} \left(\frac{d-1}{d-2} \rho + 4c_1 l^2 R_{trtr} + 2c_2 l^2 F_{tr} F_{tr} \right). \quad (3.23)$$

Finally, as $T_0 \rightarrow 0$ we reproduce the relation [7, 12]

$$\left(\frac{\partial M}{\partial \epsilon}\right)_{Q,T=0} = -T_0 \left(\frac{\partial S}{\partial \epsilon}\right)_{Q,M}. \quad (3.24)$$

Note that this relation was obtained using only the general feature that the corrected geometry may be written in terms of a shift Δg to the radial function $g(r)$. In particular, we never had to use the explicit form of Δg given in (2.6).

3.3 Explicit results for the entropy shifts

In order to compare with the next section, we include some explicit results for the mass shifts. In section 5, we will see what constraints may be placed on the EFT coefficients by imposing that entropy shift is positive. We'll use the mass shift here, to remove the factor of T_0 . The entropy shift is positive when the mass shift at constant charge is negative. It is easy to see that the shifts here are positive when all the coefficients are positive.

For AdS_4 , we find:

$$T_0 \Delta S = \frac{1}{5r_h l^2} \left(4c_1 (l^2 + 3r_h^2)^2 + 2c_2 (l^2 + 3r_h^2)(l^2 + 18r_h^2) + 8(2c_3 + c_4)(l^2 + 3r_h^2)^2 \right). \quad (3.25)$$

For AdS₅, we get:

$$T_0\Delta S = \frac{\pi}{16l^2} \left(c_1(31l^4 + 128l^2r_h^2 + 138r_h^4) + c_224(l^2 + 2r_h^2)(l^2 + 6r_h^2) + (2c_3 + c_4)72(l^2 + 2r_h^2)^2 \right). \quad (3.26)$$

AdS₆:

$$T_0\Delta S = \frac{2\pi}{99l^2} \left(c_1r_h(369l^4 + 1263l^2r_h^2 + 1124r_h^4) + c_24r_h(3l^2 + 5r_h^2)(27l^2 + 100r_h^2) + (2c_3 + c_4)96r_h(3l^2 + 5r_h^2)^2 \right). \quad (3.27)$$

AdS₇:

$$T_0\Delta S = \frac{\pi^2}{224l^2} \left(c_1(1384l^4r_h^2 + 4236l^2r_h^4 + 3345r_h^6) + c_240(2l^2 + 3r_h^2)(16l^2 + 45r_h^2) + (2c_3 + c_4)800(2l^2 + 3r_h^2)^2 \right). \quad (3.28)$$

4 Thermodynamics from the on-shell Euclidean action

The ultimate goal of this paper is to determine the leading higher-derivative corrections to relations between certain global properties of black hole solutions. These relations are of a thermodynamic nature, and arise by taking various derivatives of the free-energy corresponding to the appropriate ensemble. As is well-known [24], the classical free-energy of a black hole can be calculated using the saddle-point approximation of the Euclidean path integral with appropriate boundary conditions. In the *Gibbs* or *grand canonical ensemble*, the appropriate quantity is the Gibbs free-energy, which may be calculated from the on-shell Euclidean action

$$\beta G(T, \Phi) = I_E[g_{\mu\nu}^E(T, \Phi), A_\mu^E(T, \Phi)], \quad (4.1)$$

where $\beta = T^{-1}$, and $g_{\mu\nu}^E(T, \Phi)$ and $A_\mu^E(T, \Phi)$ are Euclideanized solutions to the classical equations of motion with temperature T and potential Φ . Similarly in the *canonical ensemble* the corresponding quantity is the Helmholtz free-energy, given by

$$\beta F(T, Q) = I_E[g_{\mu\nu}^E(T, Q), A_\mu^E(T, Q)], \quad (4.2)$$

where $g_{\mu\nu}^E(T, Q)$ and $A_\mu^E(T, Q)$ are Euclideanized solutions with temperature T and electric charge Q . In both expressions, I_E is the *renormalized* Euclidean on-shell action.

The Euclidean action with cosmological constant is IR divergent when evaluated on a solution. However, it may be given a satisfactory finite definition by first regularizing the integral with a radial cutoff R . To render the variation principle well-defined on a spacetime with boundary we must add an appropriate set of Gibbons-Hawking-York (GHY) [25, 26]

(in the case of the canonical ensemble, also Hawking-Ross [27]) terms *in addition to* a set of boundary counterterms. The complete on-shell action then consists of three contributions

$$I_E = I_{\text{bulk}} + I_{\text{GHY}} + I_{\text{CT}}. \tag{4.3}$$

If the counterterms are chosen correctly, they will cancel the divergence of the bulk and Gibbons-Hawking-York terms, rendering the results finite as $R \rightarrow \infty$. In AdS there is a systematic approach to generating such counterterms via the method of *holographic renormalization* [21, 28, 29]; since the logic of this approach is well-described in detail elsewhere (see e.g. [30]) we will not review it further, but simply make use of known results. Explicit expressions for the needed GHY and counterterms (including the four-derivative corrections used in this paper) valid in AdS_d , $d = 4, 5, 6$ can be found in [20, 31].

Once the free-energy is calculated, the remaining thermodynamic quantities can be determined straightforwardly by using the definitions of the free-energies and the first-law of black hole thermodynamics

$$F = E - TS, \quad G = E - TS - \Phi Q, \quad dE = TdS + \Phi dQ. \tag{4.4}$$

The expressions calculated using these Euclidean methods should agree with the Lorentzian or geometric calculations in the previous section. Note, however, that there is a bit of a subtlety with the notion of black hole mass here, as the thermodynamic relations are for the energy E of the system. In holographic renormalization, there is always an ambiguity in the addition of finite counterterms that shift the value of the on-shell action. The standard approach is to fix the ambiguity by demanding that even-dimensional global AdS has zero vacuum energy while odd-dimensional global AdS has non-zero vacuum energy that is interpreted as a Casimir energy in the dual field theory. In this case the thermodynamic energy is the sum of the black hole mass and the Casimir energy

$$E = M + E_c, \tag{4.5}$$

and the mass M of the black hole is only obtained after subtracting out the Casimir energy contribution, as we did in section 2.

The purpose of introducing this alternative approach is not just to give a cross-check on the results of the previous section, but also to verify a recent general claim by Reall and Santos [17]. In this paper, the $\mathcal{O}(\epsilon)$ corrections we are considering can be calculated by first evaluating the free-energy or on-shell action at the same order. Naively, this would require evaluating three contributions

$$I_E[g_{\mu\nu}^E, A_\mu^E] = I_E^{(2)}[g_{\mu\nu}^{(2)E}, A_\mu^{(2)E}] + \epsilon \left(\frac{\partial}{\partial \epsilon} I_E^{(2)}[g_{\mu\nu}^{(2)E} + \epsilon g_{\mu\nu}^{(4)E}, A_\mu^{(2)E} + \epsilon A_\mu^{(4)E}] \right) \Big|_{\epsilon=0} + \epsilon I_E^{(4)}[g_{\mu\nu}^{(2)E}, A_\mu^{(2)E}] + \mathcal{O}(\epsilon^2), \tag{4.6}$$

where (2) and (4) denote two and four derivative terms in the action and their corresponding perturbative contributions to the solution. The central claim in [17] is that the first term at $\mathcal{O}(\epsilon)$ is actually *zero*, and that therefore we do not need to explicitly calculate the $\mathcal{O}(\epsilon)$

corrections to the equations of motion. For black hole solutions of the type considered in this paper, we can evaluate the leading corrections without much difficulty, but for more general situations with less symmetry this may not be possible. In such a case the Euclidean method is more powerful, as has recently been demonstrated with calculation of corrections involving angular momentum [32] or dilaton couplings [33].

Although the result of [17] was demonstrated in the grand canonical ensemble, it is straightforward to see that it implies an identical claim about the leading corrections in the canonical ensemble. While the quantities of interest can be extracted from either, the explicit expressions encountered in the latter are usually far simpler and therefore more convenient. Recall that we can change ensemble by a Legendre transform of the free-energy

$$F(T, Q) = G(T, \Phi(Q)) + \Phi(Q)Q, \quad Q = - \left(\frac{\partial G}{\partial \Phi} \right)_T, \quad (4.7)$$

where the right-hand-side is defined in terms of the implicit inverse function $\Phi(Q)$. At fixed T and Q , the potential Φ receives corrections from the higher-derivative interactions, and so, expanding the right-hand-side to $\mathcal{O}(\epsilon)$, we have

$$F(T, Q) = G^{(2)}(T, \Phi^{(2)}(Q)) + \epsilon \left(\frac{\partial}{\partial \epsilon} G^{(2)}(T, \Phi^{(2)}(Q) + \epsilon \Phi^{(4)}(Q)) \right) \Big|_{\epsilon=0} + \epsilon G^{(4)}(T, \Phi^{(2)}(Q)) + \Phi^{(2)}(Q)Q + \epsilon \Phi^{(4)}(Q)Q + \mathcal{O}(\epsilon^2). \quad (4.8)$$

Recognizing that

$$\left(\frac{\partial}{\partial \epsilon} G^{(2)}(T, \Phi^{(2)}(Q) + \epsilon \Phi^{(4)}(Q)) \right) \Big|_{\epsilon=0} = \Phi^{(4)}(Q) \left(\frac{\partial G^{(2)}}{\partial \Phi} \right) \Big|_{\Phi=\Phi^{(2)}(Q)} = -\Phi^{(4)}(Q)Q, \quad (4.9)$$

we see that the leading correction to the Helmholtz free energy is simply given by

$$F(T, Q) = F^{(2)}(T, Q) + \epsilon G^{(4)}(T, \Phi^{(2)}(Q)) + \mathcal{O}(\epsilon^2). \quad (4.10)$$

In terms of the on-shell Euclidean action, using the result of Reall and Santos, this is then equivalent to

$$F^{(4)}(T, Q) = \frac{1}{\beta} I_E^{(4)} \left(g_{\mu\nu}^{(2)E}(T, Q), A_\mu^{(2)E}(T, Q) \right), \quad (4.11)$$

where here $I_E^{(4)}$ denotes the contribution of the four-derivative terms to the renormalized on-shell action. Note that this includes potential four-derivative Gibbons-Hawking-York terms, but as this argument makes clear, will not include any additional Hawking-Ross terms. This expression is the analogue of the Reall-Santos result, but in the canonical ensemble. It says that the leading correction to the Helmholtz free-energy is given by evaluating the four-derivative part of the renormalized on-shell action on a solution to the two-derivative equations of motion with temperature T and charge Q .

Below we will give a brief review of the well-known thermodynamic relations at two-derivative order, and then using the above result we will calculate the leading corrections and verify explicitly that they agree with the results of the previous section.

4.1 Two-derivative thermodynamics

As described above, the regularized on-shell action has a bulk as well as various boundary contributions. At two-derivative order and in d -dimensions these have the explicit form

$$\begin{aligned}
 I_{\text{bulk}}^{(2)} &= -\frac{1}{16\pi} \int d^{d+1}x \sqrt{g} \left(\frac{d(d-1)}{l^2} + R - \frac{1}{4}F^2 \right), \\
 I_{\text{GHY}}^{(2)} &= -\frac{1}{8\pi} \int d^d x \sqrt{h} K, \\
 I_{\text{CT}}^{(2)} &= \frac{1}{8\pi} \int d^d x \sqrt{h} \left(\frac{d-1}{l} + \frac{l}{2(d-2)} \mathcal{R} \right),
 \end{aligned}
 \tag{4.12}$$

where h_{ab} and \mathcal{R}_{ab} are the metric and Ricci tensor of the induced geometry on the boundary at $r = R$. Note that in $I_{\text{CT}}^{(2)}$ we have included the minimal set of counterterms necessary to cancel the IR divergence in $d = 3$ and $d = 4$. For $d > 4$, additional counterterms beginning at quadratic order in the boundary Riemann tensor are necessary to cancel further divergences.

The regularized bulk action has a well-defined variational principle provided that $\delta A_a = 0$ at $r = R$. This amounts to holding Φ fixed, and thus it corresponds to boundary conditions compatible with the grand canonical ensemble. For many applications, we will want to hold the charge fixed. From a thermodynamic point of view, we want to use the extensive quantity Q instead of the intensive Φ , so we must compute the Helmholtz free energy instead of the Gibbs free energy. Holding Q fixed requires different boundary conditions, and in particular the further addition of a Hawking-Ross boundary term [27]

$$I_{\text{HR}}^{(2)} = \frac{1}{16\pi} \int d^d x \sqrt{h} n_\mu F^{\mu b} A_b,
 \tag{4.13}$$

where n_μ is the normal vector on the boundary and A_a is the pull-back of the gauge potential. To summarize, the total two-derivative on-shell action

$$I_E^{(2)} = I_{\text{bulk}}^{(2)} + I_{\text{GHY}}^{(2)} + I_{\text{HR}}^{(2)} + I_{\text{CT}}^{(2)},
 \tag{4.14}$$

evaluated on the Euclideanized solution to the two-derivative equations of motion

$$\begin{aligned}
 ds_E^2 &= f(r) d\tau^2 + g(r)^{-1} dr^2 + r^2 d\Omega_{d-1}^2, \quad f(r) = g(r) = 1 - \frac{m}{r^{d-2}} + \frac{q^2}{4r^{2d-4}} + \frac{r^2}{l^2}, \\
 A_E &= i \left(-\frac{1}{c} \frac{q}{r^{d-2}} + \Phi \right) d\tau, \quad c = \sqrt{\frac{2(d-2)}{(d-1)}}, \quad \Phi = \frac{1}{c} \frac{q}{l^{d-2} \nu^{d-2}},
 \end{aligned}
 \tag{4.15}$$

is equal to $\beta F^{(2)}(T, Q)$, where $F^{(2)}$ is the two-derivative contribution to the Helmholtz free-energy. In the above we have introduced the dimensionless variable $\nu \equiv (r_h)_0/l$, where $(r_h)_0$ is the location of the outer-horizon of the two-derivative solution with temperature T and charge Q . Note also that here, and for the remainder of this section, we will consider only spherical $k = 1$ black holes. Since ν satisfies $f(\nu) = 0$, we can solve for the parameter m as

$$m = \nu^{d-2} + \frac{q^2}{4\nu^{d-2}} + \frac{\nu^d}{l^2}.
 \tag{4.16}$$

In the Euclidean approach to calculating the leading corrections to the thermodynamics, it will prove natural to continue to use ν and q to parametrize the space of black hole solutions, even when the four-derivative corrections are included. This means that it is also natural to write all thermodynamic quantities in these variables, which requires the use of standard thermodynamic derivative identities to rewrite derivatives. Recall that the parameter q and the physical charge Q are not the same, but are related by an overall constant given in (2.21). Therefore holding Q fixed is the same as holding q fixed. Explicitly, the two-derivative free-energy calculated in this way in AdS₄ is given by

$$F_{d=3}^{(2)}(q, \nu) = -\frac{l\nu^3}{4} + \frac{l\nu}{4} + \frac{3q^2}{16l\nu}, \tag{4.17}$$

and in AdS₅ by

$$F_{d=4}^{(2)}(q, \nu) = -\frac{1}{8}\pi l^2 \nu^4 + \frac{1}{8}\pi l^2 \nu^2 + \frac{5\pi q^2}{32l^2 \nu^2} + \frac{3\pi l^2}{32}. \tag{4.18}$$

Once the free-energy is calculated, the entropy and energy are given by

$$S = -\left(\frac{\partial F}{\partial T}\right)_Q, \quad E = F + TS. \tag{4.19}$$

In terms of our natural variables, we can re-express the entropy as

$$S(q, \nu) = \left(\frac{\partial F}{\partial \nu}\right)_q \left[\left(\frac{\partial T}{\partial \nu}\right)_q\right]^{-1}, \tag{4.20}$$

where the temperature is given by

$$T(q, \nu) = \frac{(d-2)q^2 l^{1-d} \nu^{1-d}}{4\pi} + \frac{(d-1)\nu^2 + d-2}{4\pi l}. \tag{4.21}$$

Note that this expression is exact, meaning it does not receive corrections when we include the four-derivative interactions. It is therefore useful to introduce the function

$$q_{\text{ext}}^2(\nu) = -\frac{2(d\nu^2 + d - \nu^2 - 2)(l\nu)^{d-2}}{(d-2)}, \tag{4.22}$$

such that taking the limit $q^2 \rightarrow q_{\text{ext}}^2(\nu)$ is equivalent to taking the extremal limit $T \rightarrow 0$.

If we extract the energy $E = F + TS$ from the expressions (4.17) and (4.18), we find that it agrees with the mass, (2.23), for AdS₄ but not AdS₅. This is not surprising as the thermodynamic energy E and mass M of the black hole in AdS₅ differ by a Casimir energy contribution that is independent of q and ν . We can, of course remove the Casimir energy by the addition of *finite boundary counterterms*, or equivalently by a change in holographic renormalization scheme. The expression (4.18) is calculated in a *minimal subtraction* scheme, in which the possible finite counterterms are zero and the Casimir energy is present.

Physically, it is useful work in a scheme in which the energy E coincides with the mass M of the black hole, without a Casimir contribution. In such a *zero Casimir* scheme, the energy of pure AdS₅ is defined to be zero. Calculating the free-energy from the on-shell

action of pure AdS₅ with generically parametrized four-derivative counterterms we find that this scheme requires the following modification from the minimal subtraction counterterms

$$I_{\text{CT}}^{(2)} \longrightarrow I_{\text{CT}}^{(2)} + \frac{1}{8\pi} \int d^4x \sqrt{h} \left(-\frac{l^3}{96} \right) \mathcal{R}^2. \quad (4.23)$$

The free energy calculated with this modified on-shell action agrees exactly with the expectation using (2.23). Note that the entropy, since it is given by a derivative of the free-energy, is independent of the choice of scheme. The zero Casimir scheme is a physically motivated choice, but certainly not unique.

4.2 Four-derivative corrections to thermodynamics

To evaluate the four-derivative corrections we make use of the result (4.11). As in the two-derivative contribution, the on-shell action is properly defined by a regularization and renormalization procedure. For the operators in (2.1) with Wilson coefficients c_2 , c_3 and c_4 the required $I_{\text{bulk}}^{(4)}$ contribution is actually finite, while for the term in (2.1) proportional to c_1 , we must again regularize and renormalize by adding infinite boundary counterterms. The required explicit expressions, as well as the complete set of four-derivative GHY terms, can be found in [20, 31]. The calculation is otherwise identical to the two-derivative contribution described above, and in AdS₄ we find

$$F_{d=3}^{(4)}(q, \nu) = c_1 \left(-\frac{(20l^4\nu^4 - 5l^2\nu^2q^2 + q^4)}{20l^5\nu^5} - \frac{3\nu}{l} \right) + \frac{c_2q^2(l^2(20l^2\nu^2 - 7q^2) - 60l^4\nu^4)}{80l^7\nu^5} - \frac{c_3q^4}{5l^5\nu^5} - \frac{c_4q^4}{10l^5\nu^5}. \quad (4.24)$$

The complete free-energy, up to $\mathcal{O}(\epsilon^2)$ contributions, is then given by

$$F_{d=3}(q, \nu) = F_{d=3}^{(2)}(q, \nu) + \epsilon F_{d=3}^{(4)}(q, \nu) + \mathcal{O}(\epsilon^2). \quad (4.25)$$

From this explicit expression we can then calculate the entropy

$$S_{d=3} = \pi l^2 \nu^2 - \frac{4\pi c_1 \epsilon (4l^4\nu^4 (1 - 3\nu^2) - 3l^2\nu^2q^2 + q^4)}{4l^2(3\nu^2 - 1)\nu^4 + 3\nu^2q^2} - \frac{\pi c_2 q^2 \epsilon (12l^2\nu^2(\nu^2 - 1) + 7q^2)}{4l^2(3\nu^2 - 1)\nu^4 + 3\nu^2q^2} - \frac{16\pi c_3 q^4 \epsilon}{4l^2(3\nu^2 - 1)\nu^4 + 3\nu^2q^2} - \frac{8\pi c_4 q^4 \epsilon}{4l^2(3\nu^2 - 1)\nu^4 + 3\nu^2q^2} + \mathcal{O}(\epsilon^2), \quad (4.26)$$

and mass (which coincides with the thermal energy)

$$M_{d=3} = \frac{1}{2}l(\nu^3 + \nu) + \frac{q^2}{8l\nu} + \frac{c_1 q^4 \epsilon (q^2 - 4l^2\nu^2(9\nu^2 + 2))}{40l^5(3\nu^2 - 1)\nu^7 + 30l^3\nu^5q^2} + \frac{c_2 q^2 \epsilon (80l^4\nu^4(-9\nu^4 + 6\nu^2 + 1) - 8l^2\nu^2(39\nu^2 + 7)q^2 + 7q^4)}{40l^3\nu^5(4l^2\nu^2(3\nu^2 - 1) + 3q^2)} + \frac{2c_3 q^4 \epsilon (q^2 - 4l^2\nu^2(9\nu^2 + 2))}{5l^3\nu^5(4l^2\nu^2(3\nu^2 - 1) + 3q^2)} + \frac{c_4 q^4 \epsilon (q^2 - 4l^2\nu^2(9\nu^2 + 2))}{5l^3\nu^5(4l^2\nu^2(3\nu^2 - 1) + 3q^2)} + \mathcal{O}(\epsilon^2). \quad (4.27)$$

Taking the extremal limit we find the following expression for the mass shift

$$(\Delta M_{d=3})_{Q,T=0} = -\frac{4c_1 l (3\nu^2 + 1)^2}{5\nu} - \frac{2c_2 l (3\nu^2 + 1)(18\nu^2 + 1)}{5\nu} - \frac{16c_3 l (3\nu^2 + 1)^2}{5\nu} - \frac{8c_4 l (3\nu^2 + 1)^2}{5\nu}, \quad (4.28)$$

which agrees exactly with (3.25). Strictly, the two expressions are parameterized in terms of different variables (ν the uncorrected horizon vs. r_h the corrected horizon), but these differ by $\mathcal{O}(\epsilon)$, and so when we take $\epsilon \rightarrow 0$ the two functions are the same.

Similarly we can calculate the shift in the microcanonical entropy, which will be important in the subsequent section for analyzing conjectured bounds on the Wilson coefficients. The actual expression is given in (5.10), and can be calculated straightforwardly using standard thermodynamic derivative identities

$$(\Delta S)_{Q,E} = \lim_{\epsilon \rightarrow 0} \left[\left(\frac{\partial S}{\partial \epsilon} \right)_{q,\nu} - \left(\frac{\partial E}{\partial \epsilon} \right)_{q,\nu} \frac{\left(\frac{\partial S}{\partial \nu} \right)_q}{\left(\frac{\partial E}{\partial \nu} \right)_q} \right]. \quad (4.29)$$

The calculation for AdS₅ is similar, but in this case we have to be cautious about the Casimir energy. We calculate the free-energy in the physically motivated zero Casimir scheme. To do so, we again fix the finite counterterms by evaluating the four-derivative on-shell action on pure AdS₅. Requiring the Casimir energy to vanish requires the following modification from the minimal subtraction counterterm action

$$I_{\text{CT}}^{(4)} \rightarrow I_{\text{CT}}^{(4)} + \frac{1}{8\pi} \int d^4x \sqrt{h} \left(-\frac{5c_1 l^3}{48} \right) \mathcal{R}^2. \quad (4.30)$$

Using this we calculate the four-derivative contribution to the renormalized free-energy

$$F_{d=4}^{(4)} = \frac{1}{256} \pi c_1 \left(-\frac{43q^4}{l^8 \nu^8} + \frac{24(5\nu^2 + 8)q^2}{l^4 \nu^4} - 32(13\nu^4 + 41\nu^2 + 18) \right) + \frac{3\pi c_2 (8l^4 \nu^4 q^2 - 3q^4)}{32l^8 \nu^8} - \frac{9\pi c_3 q^4}{16l^8 \nu^8} - \frac{9\pi c_4 q^4}{32l^8 \nu^8}. \quad (4.31)$$

We also obtain the entropy

$$S_{d=4} = \frac{1}{2} \pi^2 l^3 \nu^3 + \frac{\pi^2 c_1 \epsilon (8l^8 (26\nu^2 + 41) \nu^{10} + 6l^4 (5\nu^2 + 16) \nu^4 q^2 - 43q^4)}{4l^3 \nu^3 (4l^4 (2\nu^2 - 1) \nu^4 + 5q^2)} + \frac{6\pi^2 c_2 \epsilon (4l^4 \nu^4 q^2 - 3q^4)}{l^7 (8\nu^9 - 4\nu^7) + 5l^3 \nu^3 q^2} - \frac{36\pi^2 c_3 q^4 \epsilon}{l^7 (8\nu^9 - 4\nu^7) + 5l^3 \nu^3 q^2} - \frac{18\pi^2 c_4 q^4 \epsilon}{l^7 (8\nu^9 - 4\nu^7) + 5l^3 \nu^3 q^2} + \mathcal{O}(\epsilon^2), \quad (4.32)$$

and mass

$$M_{d=4} = \frac{3\pi (4l^4 (\nu^2 + 1) \nu^4 + q^2)}{32l^2 \nu^2} + c_1 \left[\frac{\pi \epsilon (384l^{12} (\nu^2 + 1) (26\nu^4 + 23\nu^2 + 6) \nu^{12} - 32l^8 (27\nu^4 + 32\nu^2 + 18) \nu^8 q^2)}{256l^8 \nu^8 (4l^4 (2\nu^2 - 1) \nu^4 + 5q^2)} \right]$$

$$\begin{aligned}
& + \frac{\pi\epsilon(-4l^4(684\nu^2+253)\nu^4q^4+129q^6)}{256l^8\nu^8(4l^4(2\nu^2-1)\nu^4+5q^2)} \Big] \\
& + \frac{3\pi c_2 q^2 \epsilon (32l^8(10\nu^2+3)\nu^8 - 4l^4(54\nu^2+19)\nu^4q^2 + 9q^4)}{32l^8\nu^8(4l^4(2\nu^2-1)\nu^4+5q^2)} \\
& + \frac{9\pi c_3 q^4 \epsilon (3q^2 - 4l^4\nu^4(18\nu^2+7))}{16l^8\nu^8(4l^4(2\nu^2-1)\nu^4+5q^2)} \\
& + \frac{9\pi c_4 q^4 \epsilon (3q^2 - 4l^4\nu^4(18\nu^2+7))}{32l^8\nu^8(4l^4(2\nu^2-1)\nu^4+5q^2)} + \mathcal{O}(\epsilon^2). \tag{4.33}
\end{aligned}$$

The extremal mass shift is given by

$$\begin{aligned}
(\Delta M_{d=4})_{Q,T=0} = & -\frac{1}{16}\pi c_1(138\nu^4+128\nu^2+31) - \frac{3}{2}\pi c_2(2\nu^2+1)(6\nu^2+1) \\
& - 9\pi c_3(2\nu^2+1)^2 - \frac{9}{2}\pi c_4(2\nu^2+1)^2, \tag{4.34}
\end{aligned}$$

which agrees exactly with the result (3.26). Likewise we can calculate the correction to the microcanonical entropy using (4.29), the explicit expression is given in (A.2).

5 Constraints from positivity of the entropy shift

Having derived the general entropy shift at fixed mass, we may now determine what constraints on the EFT coefficients are implied by the assumption that it is positive. Recall that the argument of [7] for the positivity of the entropy shift assumes the existence of a number of quantum fields ϕ with mass m_ϕ , heavy enough so that they can be safely integrated out. In particular, such fields are assumed to couple to the graviton and photon in such a way that, after being integrated out, they generate *at tree-level* the higher-dimension operators we are considering (with the corresponding operator coefficients scaling as $c_i \sim 1/m_\phi$). This assumption is essential to the proof; it may be that the entropy shift is universally positive (see [32] for a number of examples), but proving such a statement for non-tree-level completions would require a different argument from the one laid out here.

We revisit the logic of [7] in the context of flat space, before discussing how it may be extended to AdS asymptotics, and denote the Euclidean on-shell action of the theory that includes the heavy scalars ϕ by $I_{UV}[g, A, \phi]$. First, note that when the scalars are set to zero and are non-dynamical, the action reduces to that of the pure Einstein-Maxwell theory,

$$I_{UV}[g, A, 0] = I^{(2)}[g, A]. \tag{5.1}$$

This is a statement relating the value of the functionals I_{UV} and $I^{(2)}$ (the two-derivative action) when we pick particular configurations for the fields. These fields may or may not be solutions to the equations of motion. Next, consider the corrected action, $I_C = I^{(2)} + I^{(4)}$, and note that it obeys

$$I_C[g + \Delta g, A + \Delta A] \simeq I_{UV}[g, A, \phi]. \tag{5.2}$$

Here we have in mind that the fields are valid solutions of the respective theories, i.e. $[g, A, \phi]$ is a solution of the UV theory and $[g + \Delta g, A + \Delta A]$ is a solution to the four-derivative

corrected theory. The UV theory and that with an infinite series of higher-derivative corrections should have exactly the same partition function; therefore, this expression is an equality up to quantum corrections and corrections that are $\mathcal{O}(\epsilon^2)$. Finally, let us choose $[g, A, \phi]$ to be solutions of the UV theory with charge Q and temperature T , and $[g_0, A_0]$ to be field configurations in the pure Einstein-Maxwell theory *with the same charge and temperature* as those of the UV theory. One then finds the following inequality,

$$I_C[g + \Delta g, A + \Delta A]_{T,Q} \simeq I_{UV}[g, A, \phi]_{T,Q} < I_{UV}[g_0, A_0, 0]_{T,Q} = I^{(2)}[g_0, A_0]_{T,Q}. \quad (5.3)$$

Since $[g, A, \phi]$ is a solution of the UV theory, it extremizes the action. To ensure the inequality that appears in (5.3), one must further require that this solution is a *minimum* of the action. The inequality then follows because $[g_0, A_0, 0]$ is *not* a solution to the equations of motion, for the same charge and temperature. Finally, as long as one works in the same ensemble, the boundary terms will be the same for both actions and thus don't affect the argument.

In general, different theories will have different relationships between mass, charge, and temperature. We are interested in the entropy shift at fixed mass and charge. Therefore we must compare the two action functionals at different temperatures. For simplicity, we use T_4/T_2 for the temperature that corresponds to mass M and charge Q for the theory with/without higher-derivative corrections, respectively. Then we have:

$$\begin{aligned} F_C(Q, T_4) &< F_2(Q, T_4), \\ F_C(Q, T_4) &< F_2(Q, T_2) + (T_4 - T_2)\partial_T F_2(Q, T_2), \\ F_C(Q, T_4) &< F_2(Q, T_2) - (T_4 - T_2)S_2, \\ M - S_4 T_4 &< M - S_2 T_2 - (T_4 - T_2)S_2, \\ -S_4 T_4 &< -T_4 S_2, \\ \Delta S &> 0, \end{aligned} \quad (5.4)$$

at fixed M and Q (and in the zero Casimir energy scheme).

Now that we have outlined the argument in flat space, we can ask whether it can be immediately extended to AdS. One subtle point in the derivation outlined above is that the free-energy is only finite after the subtraction of the free-energy of a reference background. In the flat space context, the contributions of such terms to the two actions are identical because the asymptotic charges are the same. Thus, this issue does not affect the validity of the argument.

In AdS, the story is a little different— the free-energy is computed using holographic renormalization. Different counterterms are required to render the two-derivative action $I^{(2)}$ and the corrected action I_C finite. Moreover, I_{UV} may also require a different set of counterterms involving contributions from the scalar, and unlike the bulk contribution, there is no reason to expect that their on-shell values are less than their off-shell values. This is a potential hole in the positivity argument in AdS. Apart from this issue, the rest of the argument can be immediately applied to AdS.

5.1 Thermodynamic stability

As we've seen, the above proof requires that the uncorrected backgrounds are minima of the action. Thermodynamically, this amounts to the condition that the black holes are stable under thermal and electrical fluctuations. This translates to the following requirements on the free-energies,

$$\left(\frac{\partial^2 F}{\partial T^2}\right)_Q \leq 0, \quad \left(\frac{\partial^2 G}{\partial T^2}\right)_\Phi \leq 0, \quad \epsilon_T = \left(\frac{\partial^2 F}{\partial Q^2}\right)_T \geq 0. \quad (5.5)$$

These conditions may be rewritten in terms of the specific heat and permittivity of the black hole, which can be used to determine, respectively, the thermal stability and electrical stability of the black hole [34, 35]. We'll ignore the specific heat at constant Φ now, as we are interested in the stability in the canonical ensemble, and consider

$$C_Q = T \left(\frac{\partial S}{\partial T}\right)_Q \geq 0, \quad \epsilon_T = \left(\frac{\partial Q}{\partial \Phi}\right)_T \geq 0. \quad (5.6)$$

Positivity of the specific heat is equivalent to the statement that larger black holes should heat up and radiate more, while smaller ones should become colder and radiate less. When the quantity ϵ_T is negative the black hole is unstable to electrical fluctuations, meaning that when more charge is placed into it, its chemical potential *decreases*. We expect that it should instead increase, to make it more difficult to move a charge from outside to inside the black hole — thus making it harder to move away from equilibrium [35]. We may compute these quantities using the results of the previous section. For AdS₄, we find

$$C_Q = \frac{2\pi l^2 \nu^2 (1 + 3\nu^2)(2 - \xi)\xi}{2 - 6\xi + 3\xi^2 + 3\nu^2(4 - 6\xi + 3\xi^2)}, \quad \epsilon_T = \frac{(\xi - 2)\xi + 3\nu^2(2 - 2\xi + \xi^2)}{\nu l (2 - 6\xi + 3\xi^2 + 3\nu^2(4 - 6\xi + 3\xi^2))}, \quad (5.7)$$

where we recall that $\nu = r_h/l$ and $Q = (1 - \xi)Q_{\text{ext}}$. These results have been obtained previously e.g. in [36]. We find that both of these coefficients are positive when either

$$\nu < \nu^* = \frac{1}{\sqrt{3}}, \quad \xi < \xi^* = 1 - \sqrt{\frac{1 - 3\nu^2}{1 + 3\nu^2}}, \quad (5.8)$$

holds, or when

$$\nu > \nu^* = \frac{1}{\sqrt{3}}, \quad 0 < \xi < 1, \quad (5.9)$$

is satisfied.

Thus, for small black holes stability requires that the extremality parameter be less than some function of the radius, $\xi < \xi^*$. In particular, extremal black holes, for which $\xi \rightarrow 0$, are stable while neutral black holes, which correspond to $\xi \rightarrow 1$, are not. The implication of (5.9) is that above a certain radius ($r_h > l/\sqrt{3}$) all black holes are thermodynamically stable. This behavior is visible from figure 1, where we have plotted the allowed parameter space based on the C_Q and ϵ_T conditions separately. This raises an interesting point in making contact with the flat space limit: if we require both parameters to be positive, there

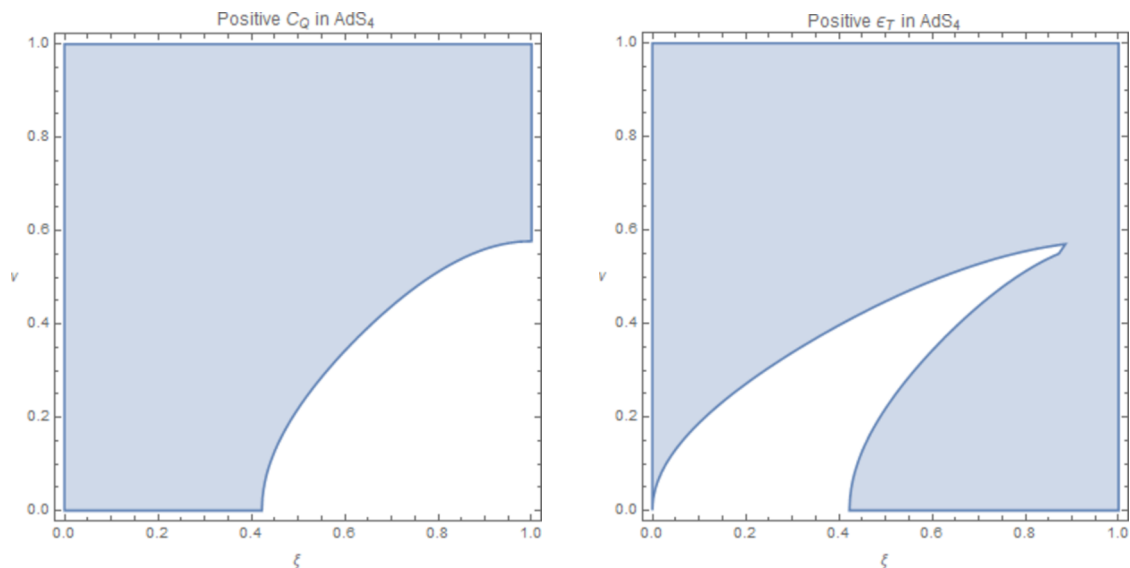


Figure 1. Blue represents the regions of parameter space where each quantity is positive.

are no stable black holes at $\nu = 0$. Note that in [7] only C_Q was considered. However, in applications involving AdS/CFT, we believe that both the specific heat and electrical permittivity should be taken into account.

Here we have only considered the leading-order stability. The higher-derivative corrections will shift the point where the specific heat crosses from positive to negative. However, in proving the extremality-entropy relation, we are only interested in the extremal limit, which is not affected by this consideration. In principal we could compute the order ϵ shifts to the stability conditions to obtain small corrections to the entropy bounds.

5.2 Constraints on the EFT coefficients

The entropy shift in AdS₄ for a black hole with an *arbitrary size and charge* takes the following form,

$$\begin{aligned} \left(\frac{\partial S}{\partial \epsilon}\right)_{Q,M} &= \frac{l(1+3\nu^2)}{5\nu T} \left(c_1 (4 - 6\xi + 19\xi^2 - 16\xi^3 + 4\xi^4 + 12\nu^2(\xi - 1)^4) \right. \\ &\quad + c_2(\xi - 1)^2 (2 - 14\xi + 7\xi^2 + 3\nu^2(12 - 14\xi + 7\xi^2)) \\ &\quad \left. + 8(2c_3 + c_4)(1 + 3\nu^2)(\xi - 1)^4 \right), \end{aligned} \quad (5.10)$$

where the temperature is given by the expression

$$T(r_h, \xi) = -\frac{(1+3\nu^2)(\xi-2)\xi}{4\pi\nu l}.$$

We can see from the ξ dependence of the latter that in the $\xi \rightarrow 0$ limit the shift to the entropy blows up. If we examine the leading part in $1/\xi$, we find that it is proportional to

the mass shifts we have computed above. Thus, in the extremal limit we have

$$\left(\frac{\partial S}{\partial \epsilon}\right)_{\xi \rightarrow 0} = \frac{l^2}{5r_h T} \left(4c_1(1+3\nu^2)^2 + 2c_2(1+3\nu^2)(1+18\nu^2) + 8(2c_3+c_4)(1+3\nu^2)^2\right). \quad (5.11)$$

It is also interesting to note that in the chargeless limit $\xi \rightarrow 1$ the dependence of (5.10) on c_2, c_3 and c_4 drops out entirely, and we are left with an entropy shift of the simple form

$$\left(\frac{\partial S}{\partial \epsilon}\right)_{\xi \rightarrow 1} = \frac{l}{\nu T} c_1 (1+3\nu^2). \quad (5.12)$$

Our results above show that large black holes are stable in the chargeless limit. Therefore, under the assumption that the four-derivative corrections yield a positive entropy shift for all possible stable black hole backgrounds, we find

$$c_1 \geq 0. \quad (5.13)$$

In figure 2, we have graphed the constraints on the coefficients that arise from demanding that the entropy shift is positive. We have included both the constraints from the extremal entropy shift and from considering the shift of all stable black holes. Considering only extremal black holes may be interesting because it is equivalent to the condition that the extremality shift, $\Delta(M/Q)$, is negative. Thus we may look at the constraints implied by positive entropy shift and by negative extremality shift independently. Note that we have divided by c_1 , which we have already proven to be positive. We may write out the all the constraints obtained:

$$\begin{aligned} c_1 &\geq 0, \\ c_2 &\geq 0, \\ c_3 &\geq -\frac{1}{8}c_1(2+c_2). \end{aligned} \quad (5.14)$$

We have computed the corresponding bounds for AdS₅ through AdS₇. The results may be found in appendix A. We would, however, like to comment on AdS₅, where the positivity of the coefficient of the Riemann-squared term is of particular interest. The stability analysis yields results that are qualitatively similar to (5.8) and (5.9), but with the following definitions

$$\xi^* = 1 - \sqrt{\frac{1-2\nu^2}{1+2\nu^2}}, \quad \nu^* = \frac{1}{\sqrt{2}}. \quad (5.15)$$

Once again, we see that large black holes are stable for all values of the charge.

When we examine the entropy shift in the neutral limit, we find

$$\frac{\pi l^2}{32T} c_1 (87 + 164\nu^2 + 52\nu^4), \quad (5.16)$$

whose overall sign is completely determined by that of c_1 . This means that there are stable black holes where the sign of the entropy shift is the same as the sign of the coefficient of

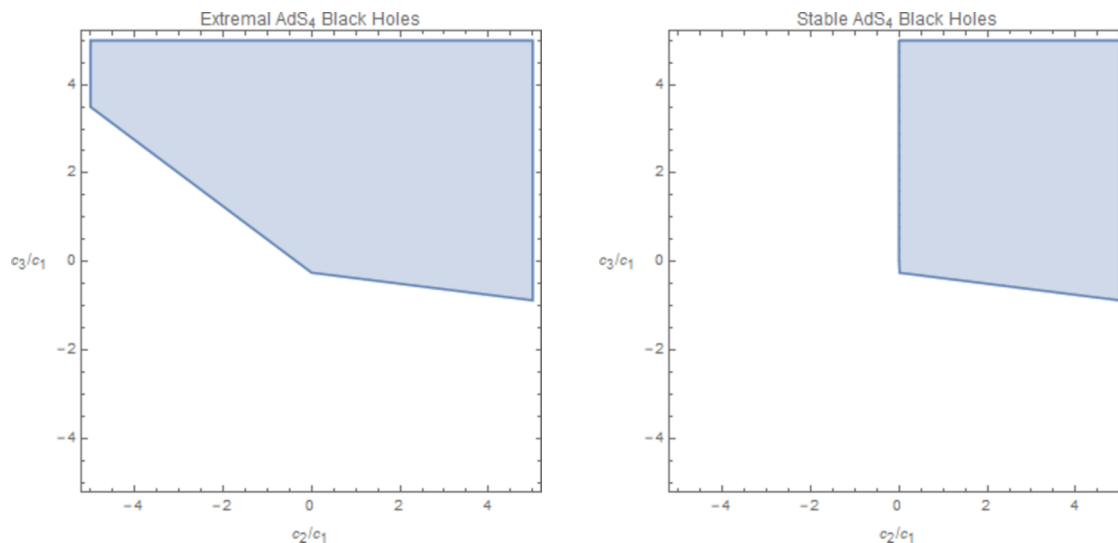


Figure 2. Blue regions are allowed after imposing that the entropy shift is positive. (Left): allowed region after imposing that extremal black holes have positive entropy shift (Right): allowed region after imposing that all stable black holes have positive entropy shift,

R_{abcd}^2 . Thus, a positive entropy shift for stable black holes implies that c_1 is positive. In fact, a positive value of c_1 was the necessary ingredient in [4] for obtaining the violation of the KSS bound.⁵ It is also interesting to note that in $d > 3$, this sign constraint was shown to follow from the assumption of a unitary tree-level UV completion [37]. The entropy constraints given in this paper are then strictly stronger since they also apply in $d = 3$.

In closing, we stress that we are not claiming that the entropy shift should be universally positive; the proof outlined above only applies when the higher-derivative corrections are generated by integrating out massive fields at tree-level (and relies on assuming that the corresponding solutions minimize the effective action). However, it is interesting that the conjecture that the entropy shift is universally positive appears to suggest that violations of the KSS bound are required to occur. Our results extend and make more precise the earlier claim by some of us [20] of a link between the WGC and the violation of the KSS bound. We will come back to this point in section 6.

5.3 Flat space limit

As we have pointed out above, we can not compare the results we have given above to the flat space limit. This is because if we impose both $C_Q > 0$ and $\epsilon_T > 0$, we find that there are no stable black holes in the flat space limit $\nu \rightarrow 0$ (as suggested by figure 1). In AdS/CFT, we expect that both conditions are necessary to ensure thermodynamic stability; nonetheless, we may remove the condition $\epsilon_T > 0$ in order to compare with the flat space limit. In this case, we find that stability requires

$$\xi^* = 1 - \frac{1}{\sqrt{3}} \sqrt{\frac{1 - 3\nu^2}{1 + 3\nu^2}}, \quad \nu^* = \frac{1}{\sqrt{3}}, \quad (5.17)$$

⁵We have checked the calculation with a different basis, choosing to use Gauss-Bonnet instead of Riemann squared. As expected, we find that the coefficient of the Gauss-Bonnet term is positive.

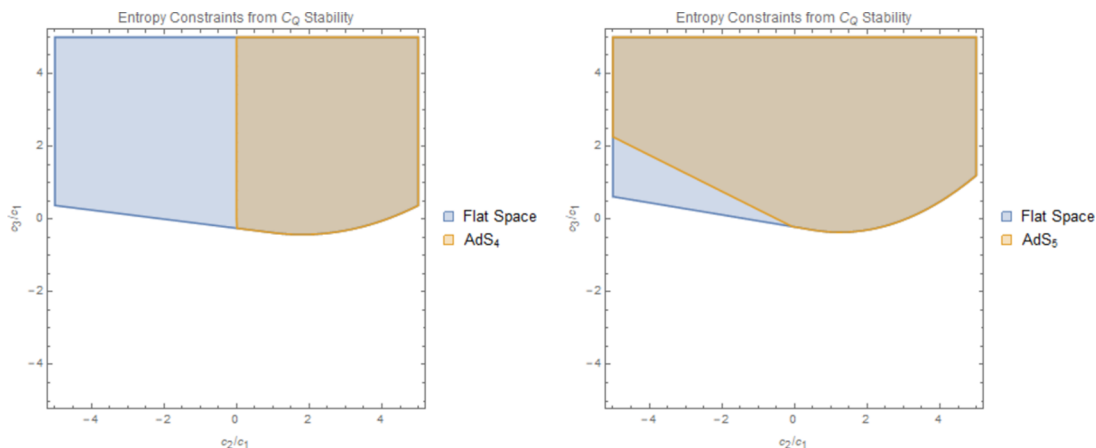


Figure 3. The blue regions are allowed in flat space and the orange in AdS— note that the AdS regions are a subset of those from flat space.

for the AdS₄ black holes, and

$$\xi^* = 1 - \frac{1}{\sqrt{2}} \sqrt{\frac{1 - 2\nu^2}{1 + 2\nu^2}}, \quad \nu^* = \frac{1}{\sqrt{2}}, \quad (5.18)$$

for the AdS₅ black holes. This allows for a more direct comparison between the two cases. In figure 3, we contrast the bounds obtained in AdS and flat space. The bounds in AdS are stronger, as they should be given that there is an extra parameter’s worth of stable black holes. Note also that $c_1 > 0$ is implied by positivity in AdS, but not in flat space, because in flat space there are no stable neutral black holes.

6 Discussion

In this paper, we have examined the relationship between the higher-derivative corrections to entropy and extremality in Anti-de Sitter space. As we have seen, extremality is considerably more complicated in AdS because the relationship between mass, charge, and horizon radius at extremality is non-linear. Nonetheless, we have verified the relation [7, 12] between the entropy shift at fixed charge and mass and the extremality shift at fixed charge and temperature. There is a sharp dependence on which quantities are held fixed in AdS. This is in contrast to flat space, where the linear relationship between mass, charge, and horizon radius removes this issue. We have also provided a more general proof of this relation in appendix B, and extended the result to show that there is a third proportional quantity, which is the extremality shift at fixed mass and temperature.

When viewed geometrically, these statements seem almost accidental. In section 4, we performed the same calculation from a thermodynamic point of view by computing the free energy from the renormalized on-shell action. From this point of view, issues concerning “which quantity is held fixed” translate to “which ensemble is used.” In addition to providing an additional check on the results from section 3, this provides a non-trivial confirmation of the results of [17], which states that the shifted geometry is not needed to compute the thermodynamic quantities.

Assuming that the entropy shift is positive places constraints on the Wilson coefficients. However, a crucial difference appears in AdS when compared to flat space. The stability criterion depends on the horizon radius over the AdS length, and goes to zero at large horizon radius. This means that there are stable neutral black holes that are asymptotically AdS. For neutral black holes, the entropy shift is dominated by c_1 , which is the coefficient of the Riemann squared term, so the positivity of the entropy shift implies the positivity of this coefficient. In AdS₅, this coefficient may be related to the central charges of the dual field theory [28, 38, 39] by

$$c_1 = \frac{1}{8} \frac{c - a}{c}. \tag{6.1}$$

Thus, the positivity of the entropy shift appears to be violated in theories where $c - a < 0$. In [40], a number of superconformal field theories were examined, and all were found to satisfy $c - a > 0$. It is worth noting there are non-interacting theories where $c - a < 0$; for example, $\frac{a}{c} = \frac{31}{18}$ for a free theory of only vector fields [41]. However, such theories do not have weakly curved gravity duals.⁶

The question of whether holographic theories with gravity and gauge fields necessarily correspond to $c - a$ non-negative is interesting for a number of reasons — both from a fundamental point of view and for phenomenological applications.

In particular, recall that the range of the Wilson coefficients and the sign of $c - a$ played an important role in the physics of the shear viscosity to entropy ratio η/s and how it deviates from its universal $1/4\pi$ result [42, 43], as discussed extensively in the literature (see [44] for a review of the status of the shear viscosity to entropy bound). Indeed, it is interesting to compare our results to the higher-derivative corrections to η/s , which (for the AdS₅ case of interest to us here) were shown [18] to be given by

$$\frac{\eta}{s} = \frac{1}{4\pi} \left(1 - 8c_1 + 4(c_1 + c_2) \frac{q^2}{r_0^6} \right), \tag{6.2}$$

where r_0 is a parameter of the solution defined in [18]; the factor q^2/r_0^6 goes from 0 (for neutral black holes) to 2 (at extremality). Our bounds on c_1 imply that neutral black holes will necessarily have a negative viscosity shift, violating the KSS bound. Models where this is realized are known to exist — the first UV complete counter-example to the KSS bound was given in [19]. For extremal black holes, the dependence on c_1 drops out and only the sign of c_2 matters, $\eta/s = \frac{1}{4\pi} (1 + 8c_2)$. For AdS₅, the c_2 coefficient may have both positive and negative values. However, imposing the null energy condition implies an additional constraint on the range of c_2 , which in AdS₅ takes the form

$$\frac{13}{12} c_1 + c_2 > 0. \tag{6.3}$$

⁶In fact, there are holographic theories with $c < a$. These are the \mathcal{T}_N theories, which arise as M_5 branes wrapping punctured Riemann surfaces. As these theories enjoy $\mathcal{N} = 2$ supersymmetry on the boundary, their bulk duals necessarily include massless scalars in the graviton multiplet. Therefore our analysis does not include this case, and it may be interesting to try to extend our work to include scalars in the massless spectrum. We thank Eric Perlmutter for making us aware of this interesting example.

This may be seen by first noticing that the definition of the parameter γ in equation (2.8) implies $\gamma > 0$ as long as the null energy condition holds. Then the bound in (6.3) may be derived from the specific form of γ given in (2.9). This alone is sufficient to bound c_2 from below, when c_1 is non-negative. Thus, one can see that utilizing such constraints it is at least in principle possible to bound η/s *from below*, in specific cases. To what extent this can be done generically is still an open question.

It might be interesting to try to relate the extremality bounds to the transport coefficients of the boundary theory in a more concrete way. As the corrections to η/s depend only on c_1 and c_2 in five dimensions, it is clear that the shift to extremality is not captured by the physics that controls η/s alone. One might wonder, however, if some other linear combination of transport coefficients, such as the conductivity or susceptibility,⁷ might be related to the extremality shift. From a purely CFT point of view, this is certainly not that strange; the philosophy of conformal hydrodynamics is that scaling symmetry ties together ultraviolet quantities (a, c) that characterize the CFT to the transport coefficients, which characterize the IR, long-wavelength behavior of the theory. If we believe that EFT coefficients in the bulk are related to these UV quantities (as is known in the case of c_1), then a correspondence between higher-derivatives and hydrodynamics is very natural. The question is to what extent this can be used to efficiently constrain IR quantities. Finally, we should note that extending our analysis to holographic theories that couple gravity to scalars would be useful to make contact with the efforts to generate non-trivial temperature dependence for η/s (see e.g. the discussion in [46, 47]), which is expected to play a key role in understanding the dynamics of the strongly coupled quark gluon plasma.

Our results also have potential to make contact with the work on CFTs at large global charge [48]. As we've seen above, the extremality curve for AdS-Reissner-Nordström black holes is non-linear even at the two-derivative level. In an analysis of the minimum scaling dimension for highly charged 3D CFTs states of a given charge, it was found [22] that $\Delta \sim q^{3/2}$. This is in striking agreement with the extremality relationship $m \sim q^{3/2}$ that holds for large black holes. The large charge OPE may be powerful because it offers an expansion parameter, $1/q$, which may be used even for CFTs which are strongly coupled. In principle, it should be possible to match our higher-derivative corrections to the extremality bound with corrections to the minimum scaling dimension that are subleading in $1/q$. This might allow one to use the large charge OPE to compute the EFT coefficients of the bulk dual of specific theories where the minimum scaling dimensions are known.

6.1 Weak gravity conjecture in AdS

One of the motivations for this work is to address to question of to what extent the WGC is constraining in Anti-de Sitter space. It is not obvious that it should be. In flat space, one looks for higher-derivative corrections to shift the extremality bound $m(q)$ to have a slope that is greater than one. In that case, a single nearly extremal black holes is (kinematically) allowed to decay to two smaller black holes, which can fly apart off to infinity and decay further if they wish.

⁷These have been considered in [45], which already in 2008 had an interesting comment about a possible relation to the WGC.

In AdS, the extremality bound $m(q)$ has a slope that is greater than one at the two-derivative level. Therefore one might expect that large black holes are already able to decay without any new particles or higher-derivative corrections. This picture may be too naive, however; the AdS radius introduces a long range potential that is proportional to $\frac{r^2}{l^2}$. This causes all massive states emitted from the black hole to fall back in, contrary to the situation in flat space.

A different decay path is provided by the dynamical instability [49–52], whereby charged black branes are unstable to formation of a scalar condensate. This occurs only if the theory also has a scalar with charge q and dimension Δ that satisfies

$$(m_\phi l)^2 \leq \frac{1}{2}(q_\phi g M_{Pl} l)^2 - \frac{3}{2}. \tag{6.4}$$

Note that, even in the limit of large AdS-radius l , this does not approach the bound we have for small black holes, which is $m \leq q$. Numerical work in [50] suggests that the endpoint of the instability is a state where all the charge is carried by the scalar condensate. Similar requirements appear for the superradiant instability of small black holes [53, 54]. For a more thorough review, see [13]. In either case, it is curious that in AdS, a condition similar to the flat space WGC allows for black holes to decay through an entirely different mechanism.

Another remarkable hint of the WGC comes from its connection to cosmic censorship. In [55, 56], it is shown that a class of solutions of Einstein-Maxwell theory in AdS_4 that appear to violate cosmic censorship [57] are removed if the theory is modified to include a scalar whose charge is great enough to satisfy the weak gravity bound.⁸

It may be possible to study these solutions in the presence of higher-derivative corrections. One might ask whether there is a choice of higher-derivative terms such that the singular solutions are removed. It would be interesting to check if this occurs when the higher-derivative terms are those that are obtained by integrating out a scalar of sufficient charge. It would also be interesting to compare constraints obtained by requiring cosmic censorship with constraints due to positivity of the entropy shift.

A more general proof of the WGC in AdS was given in [16]. In that paper, it was shown that, under mild assumptions, entanglement entropy for the boundary dual of an extremal black brane should go like the surface area of the entangling subregion, which is in tension with the volume law scaling predicted by the Ryu-Takayanagi formula. The contradiction is removed when one introduces a WGC-satisfying state. This violates one of the assumptions that imply the area law for the entropy— that is, the assumption that correlations decay exponentially with distance.

This form of the WGC is particularly interesting to us because it makes no reference to whether or not the WGC-satisfying state is a particle, or a non-perturbative object like a black hole. Therefore, the contradiction pointed out in that paper may be lifted if the higher-derivative corrections allow for black holes with charge greater than mass. Heavy black holes in AdS have masses far greater than their charge— therefore we expect that

⁸The bound they consider is the bound for superradiance of small black holes, which requires $\Delta \leq ql$.

the WGC-satisfying states might be provided by small black holes whose higher-derivative corrections shift the extremality bound to allow slightly more charge.

Acknowledgments

It is a pleasure to thank Anthony Charles, Cliff Cheung, Simeon Hellerman, Finn Larsen, Grant Remmen, and Gary Shiu for useful conversations and input on this project. This work was supported in part by the U.S. Department of Energy under grant DE-SC0007859. S.C. is supported in part by the National Science Foundation Grant PHY-1915038. CRTJ was supported in part by a Leinweber Student Fellowship and in part by a Rackham Predoctoral Fellowship from the University of Michigan.

A Entropy shifts from the on-shell action

In section 5, we computed the constraints on the coefficients in AdS₄. Here we will present the results of this calculation for AdS₅ through AdS₇ using the entropies computed in section 3, which corresponds to working in the zero Casimir energy scheme. For completeness, we also present the Casimir energies for AdS₅ and AdS₇ that show up in the thermodynamic energy of section 4 when using a minimal set of counterterms.

A.1 AdS₅

In AdS₅ we find that the stability condition obtained by demanding positive specific heat and permittivity is given by $\xi < \xi^*$ for $\nu < \nu^*$, with

$$\xi^* = 1 - \sqrt{\frac{1 - 2\nu^2}{1 + 2\nu^2}}, \quad \nu^* = \frac{1}{\sqrt{2}}, \quad (\text{A.1})$$

and that all black holes with $\nu > \nu^*$ are stable for all values of the charge. The full entropy shift is simpler to express as a function of charge q than extremality parameter ξ . We find

$$\begin{aligned} \left(\frac{\partial S}{\partial \epsilon}\right)_{M,Q} &= \frac{\pi}{256l^6\nu^8 T} \left(c_1 (43q^4 - 24l^4 q^2 \nu^4 (8 + 5\nu^2) + 32l^8 \nu^8 (18 + 41\nu^2 + 13\nu^4)) \right. \\ &\quad \left. + 24c_2 q^2 (3q^2 - 8l^4 \nu^4) + 72(2c_3 + c_4)q^4 \right). \end{aligned} \quad (\text{A.2})$$

Note that holographic renormalization in AdS₅ with a Riemann-squared correction yields a Casimir energy

$$E_c = \frac{\omega_3}{16\pi} \left(\frac{3}{4}l^2 - \frac{15}{4}c_1 l^2 \right), \quad (\text{A.3})$$

where $\omega_3 = 2\pi^2$. This Casimir energy must be removed from the thermodynamic energy in order to obtain the mass M of the black hole. Alternatively, it can be cancelled right from the beginning by adding an appropriate finite counterterm to the action, in which case the thermodynamic energy would then correspond directly to the mass. If the Casimir energy is not removed, then the thermodynamic energy shift becomes a combination of mass shift and Casimir energy shift since E_c depends explicitly on the c_1 Wilson coefficient.

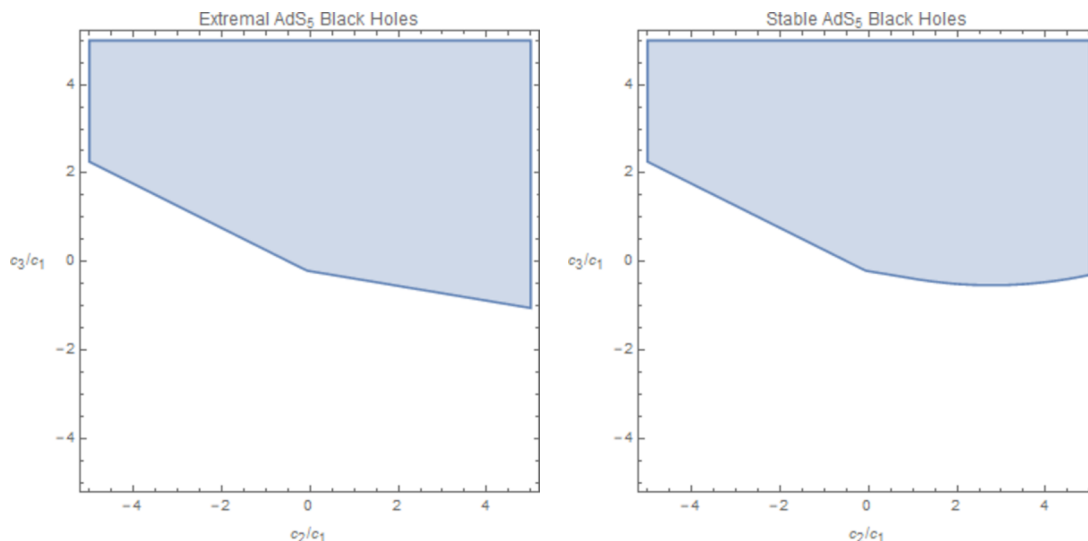


Figure 4. Allowed regions for AdS₅ EFT coefficients.

We find the following expression for the extremal limit,

$$\left(\frac{\partial S}{\partial \epsilon}\right)_{M,Q} = \frac{\pi l^2}{16T} \left(c_1(31 + 128\nu^2 + 138\nu^4) + 24c_2(1 + 2\nu^2)(1 + 6\nu^2) + 72(2c_3 + c_4)(1 + 2\nu^2)^2 \right), \quad (\text{A.4})$$

while in the neutral limit we have

$$\left(\frac{\partial S}{\partial \epsilon}\right)_{M,Q} = \frac{\pi l^2}{16T} c_1 (18 + 41\nu^2 + 13\nu^4). \quad (\text{A.5})$$

Once again, the entropy shift is proportional to c_1 in this limit. It is interesting that we do not find a positivity constraint on c_2 , as we did in AdS₄. There is a lower bound on c_3/c_1 of about -0.5339. The general constraints obtained by the Reduce function of Mathematica are extremely complicated and probably of little interest.

A.2 AdS₆

In AdS₆ the stability condition obtained by demanding positive specific heat and permissivity is of the same general structure as in AdS₅, but with the following identifications:

$$\xi^* = 1 - \sqrt{\frac{3 - 5\nu^2}{3 + 5\nu^2}}, \quad \nu^* = \sqrt{\frac{3}{5}}. \quad (\text{A.6})$$

The entropy shift is given by:

$$\left(\frac{\partial S}{\partial \epsilon}\right)_{M,Q} = \frac{\pi}{264l^9\nu^{11}T} \left(c_1(189q^4 - 22l^6q^2\nu^6(36 + 29\nu^2) + 264l^{12}\nu^{12}(8 + 17\nu^2 + 7\nu^4)) + 2c_2q^2(153q^2 - 44l^6\nu^6(9 + 5\nu^2)) + 288(2c_3 + c_4)q^4 \right), \quad (\text{A.7})$$

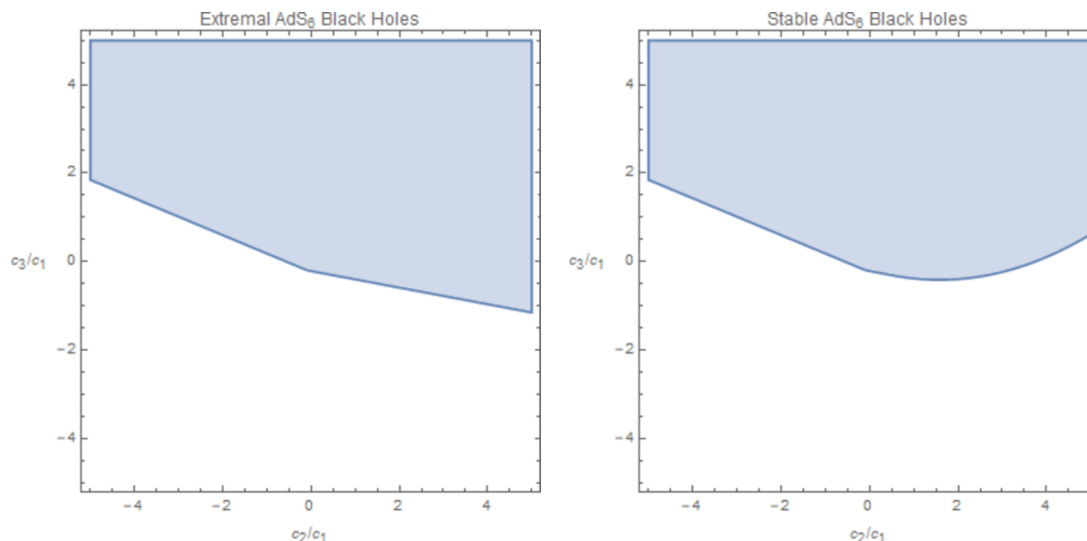


Figure 5. Allowed regions for AdS₆ EFT coefficients.

and in the extremal limit takes the form:

$$\left(\frac{\partial S}{\partial \epsilon}\right)_{M,Q} = \frac{2\nu\pi l^3}{99T} \left(c_1(369 + 1263\nu^2 + 1124\nu^4) + 4c_2(3 + 5\nu^2)(27 + 100\nu^2) + 96(2c_3 + c_4)(3 + 5\nu^2)^2 \right). \quad (\text{A.8})$$

Finally, in the neutral limit we find

$$\left(\frac{\partial S}{\partial \epsilon}\right)_{M,Q} = \frac{\nu\pi l^3}{T} c_1 (8 + 17\nu^2 + 7\nu^4). \quad (\text{A.9})$$

Note that no Casimir energy subtraction is needed in AdS₆. We again find that c_1 is positive. The other bounds are displayed in figure 5. In AdS₆ and AdS₇, the Reduce function of Mathematica was not able to find the general constraints over all stable values of ξ and ν . However, we believe that the strongest constraints will come from the boundaries of the region of stable black holes. Specifically, we imposed positivity at the neutral $\xi \rightarrow 1$ limit, the extremal $\xi \rightarrow 0$ limit, the planar limit $\nu \rightarrow \infty$ limit, and at $\xi = \xi^*$. We believe this method should give the same answer, and we have checked explicitly that it does in the case for AdS₄ and AdS₅.

A.3 AdS₇

In AdS₇ the stability window is determined by

$$\xi^* = 1 - \sqrt{\frac{2 - 3\nu^2}{2 + 3\nu^2}}, \quad \nu^* = \sqrt{\frac{2}{3}}, \quad (\text{A.10})$$

and the entropy shift is:

$$\left(\frac{\partial S}{\partial \epsilon}\right)_{M,Q} = \frac{\pi^2}{896l^{12}\nu^{14}T} \left(c_1(556q^4 - 14q^2l^8\nu^8(160 + 141\nu^2) + 56l^{16}\nu^{16}(100 + 207\nu^2 + 8\nu^4)) + 80c_2q^2(11q^2 - 7l^8\nu^8(4 + 3\nu^2)) + 800(2c_3 + c_4)q^4 \right). \quad (\text{A.11})$$

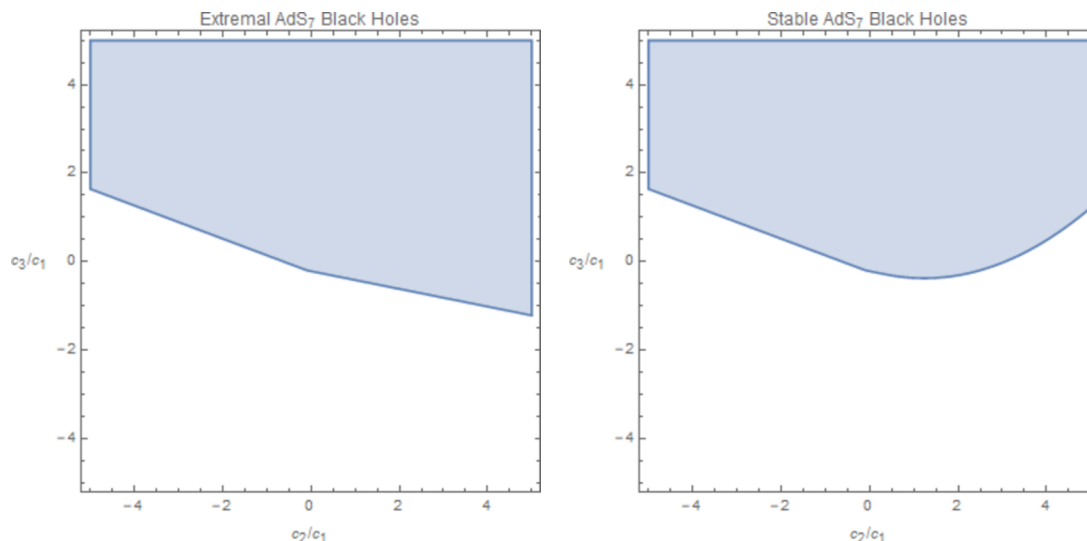


Figure 6. Allowed regions for AdS₇ EFT coefficients.

The Casimir energy that must be removed from the thermodynamic energy in AdS₇ is

$$E_c = \frac{\omega_5}{16\pi} \left(-\frac{5}{8}l^4 + \frac{35}{8}c_1l^4 \right), \quad (\text{A.12})$$

where $\omega_5 = \pi^3$.

We find the following expression for the extremal limit,

$$\begin{aligned} \left(\frac{\partial S}{\partial \epsilon} \right)_{M,Q} &= \frac{\pi^2 \nu^2 l^4}{224T} \left(c_1 (1384 + 4236\nu^2 + 3345\nu^4) \right. \\ &\quad \left. + 40c_2(2 + 3\nu^2)(16 + 45\nu^2) + 800(2c_3 + c_4)(2 + 3\nu^2)^2 \right), \end{aligned} \quad (\text{A.13})$$

while in the neutral limit we find

$$\left(\frac{\partial S}{\partial \epsilon} \right)_{M,Q} = \frac{\pi^2 l^2 \nu^2}{16T} c_1 (100 + 207\nu^2 + 93\nu^4). \quad (\text{A.14})$$

Once again, c_1 is positive. The other bounds are displayed in figure 6. Again, we used the method of extremizing over the boundaries of the space of stable black holes.

B Another proof of the entropy-extremality relation

Recent work [7, 12] suggests a remarkable universal relationship between the corrections to extremality and corrections to entropy. Here we will present a simple derivation of this relation using standard thermodynamic identities, including a slight generalization of the relation away from extremality. The statement itself is not specific to black holes, and is in fact a relatively universal statement about infinitesimal deformations of thermodynamic systems. As we explain in detail both here and in appendix C, the relation we obtain is *formally* correct, but has a subtle physical interpretation.

Consider a thermodynamic system, let E be the total thermal energy, T the temperature, S the entropy and X collectively label a set of extensive thermodynamic variables (for black holes this could be the charge Q and spin J). Now consider a small deformation of this system parametrized by a continuous parameter ϵ . The key assumption we will make about this deformation is that it preserves the third law of thermodynamics in the form

$$\lim_{T \rightarrow 0} TS(T, X, \epsilon) = 0, \tag{B.1}$$

for all ϵ on an open neighbourhood of $\epsilon = 0$. We begin with the first law of thermodynamics in the form

$$1 = T \left(\frac{\partial S}{\partial E} \right)_{X, \epsilon}. \tag{B.2}$$

Making use of the triple product identity

$$\left(\frac{\partial S}{\partial E} \right)_{X, \epsilon} \left(\frac{\partial E}{\partial \epsilon} \right)_{X, S} \left(\frac{\partial \epsilon}{\partial S} \right)_{X, E} = -1, \tag{B.3}$$

we have

$$\left(\frac{\partial E}{\partial \epsilon} \right)_{X, S} = -T \left(\frac{\partial S}{\partial \epsilon} \right)_{X, E}. \tag{B.4}$$

Formally inverting $S(T, X, \epsilon)$ gives $T(S, X, \epsilon)$. We can use this to write

$$\left(\frac{\partial E}{\partial \epsilon} \right)_{X, S} = \left(\frac{\partial E}{\partial \epsilon} \right)_{X, T} + \left(\frac{\partial E}{\partial T} \right)_{X, \epsilon} \left(\frac{\partial T}{\partial \epsilon} \right)_{X, S}. \tag{B.5}$$

Combining these

$$\left(\frac{\partial E}{\partial \epsilon} \right)_{X, T} = -T \left(\frac{\partial S}{\partial \epsilon} \right)_{X, E} - \left(\frac{\partial E}{\partial T} \right)_{X, \epsilon} \left(\frac{\partial T}{\partial \epsilon} \right)_{X, S}. \tag{B.6}$$

Next, we use (B.2) again

$$\begin{aligned} \left(\frac{\partial E}{\partial \epsilon} \right)_{X, T} &= -T \left(\frac{\partial S}{\partial \epsilon} \right)_{X, E} - \left(\frac{\partial E}{\partial T} \right)_{X, \epsilon} \left(\frac{\partial T}{\partial \epsilon} \right)_{X, S} \\ &= -T \left(\frac{\partial S}{\partial \epsilon} \right)_{X, E} - T \left(\frac{\partial S}{\partial E} \right)_{X, \epsilon} \left(\frac{\partial E}{\partial T} \right)_{X, \epsilon} \left(\frac{\partial T}{\partial \epsilon} \right)_{X, S} \\ &= -T \left(\frac{\partial S}{\partial \epsilon} \right)_{X, E} - T \left(\frac{\partial S}{\partial T} \right)_{X, \epsilon} \left(\frac{\partial T}{\partial \epsilon} \right)_{X, S}, \end{aligned} \tag{B.7}$$

one final application of the triple product identity gives a *generalized entropy-extremality relation*

$$\left(\frac{\partial E}{\partial \epsilon} \right)_{X, T} + T \left(\frac{\partial S}{\partial \epsilon} \right)_{X, E} = T \left(\frac{\partial S}{\partial \epsilon} \right)_{X, T}. \tag{B.8}$$

Next we make use of the assumption that the deformation does not violate the third law of thermodynamics. Taylor expanding (B.2) we have

$$\lim_{T \rightarrow 0} \left[TS(T, X, \epsilon = 0) + \epsilon T \left(\frac{\partial S}{\partial \epsilon} \right)_{T, X} \Big|_{\epsilon=0} + \mathcal{O}(\epsilon^2) \right] = 0. \tag{B.9}$$

By assumption this is true on an open neighbourhood of $\epsilon = 0$ and so it must be true order-by-order in the expansion. This gives

$$\lim_{T \rightarrow 0} \lim_{\epsilon \rightarrow 0} T \left(\frac{\partial S}{\partial \epsilon} \right)_{T,X} = 0. \tag{B.10}$$

Using this together with (B.8) gives the entropy-extremality relation [12]

$$\lim_{T \rightarrow 0} \left[\left(\frac{\partial E}{\partial \epsilon} \right)_{X,T} \Big|_{\epsilon=0} + T \left(\frac{\partial S}{\partial \epsilon} \right)_{X,E} \Big|_{\epsilon=0} \right] = 0. \tag{B.11}$$

For the specific application to black hole thermodynamics we identify E with the mass M of the black hole, X with the black hole parameters measured at infinity such as charge Q or angular momentum J , and ϵ with a Wilson coefficient of a four-derivative effective operator.

In section 3, we have pointed out that shift in charge at fixed mass is also proportional to the entropy shift and mass shift. This statement can be derived similarly. By the triple product identity,

$$\left(\frac{\partial E}{\partial \epsilon} \right)_{X_i,T} = - \left(\frac{\partial X_i}{\partial \epsilon} \right)_{E,T} \left(\frac{\partial E}{\partial X_i} \right)_{\epsilon,T}. \tag{B.12}$$

This holds for any extensive quantity. Now we choose $X_i = Q$, and we may identify

$$\left(\frac{\partial E}{\partial X_i} \right)_{\epsilon,T} = \Phi. \tag{B.13}$$

So we find

$$\left(\frac{\partial E}{\partial \epsilon} \right)_{Q,T} = -\Phi \left(\frac{\partial Q}{\partial \epsilon} \right)_{E,T}. \tag{B.14}$$

For black holes, this means that the shift in charge is related to the shift in mass. Neither of them is related to the entropy except at extremality. The result of this is that the entropy shift at extremality may be related to the extremality shift at constant charge or at constant mass,

$$\lim_{T \rightarrow 0} \left(\frac{\partial E}{\partial \epsilon} \right)_{Q,T} = - \lim_{T \rightarrow 0} \Phi \left(\frac{\partial Q}{\partial \epsilon} \right)_{E,T} = - \lim_{T \rightarrow 0} T \left(\frac{\partial S}{\partial \epsilon} \right)_{Q,E}. \tag{B.15}$$

While the derivation in this appendix is formally valid, the consequences of the expansion in (B.9) are subtle and require some commentary. Throughout we have made the implicit assumption that the various thermodynamic quantities are differentiable functions of ϵ , and moreover that $TS(T, X, \epsilon)$ is analytic on an open neighbourhood of $\epsilon = 0$, permitting the use of the Taylor series. In general this is a valid assumption for $T \neq 0$, but may *fail* to be valid at $T = 0$. As a consequence, the limits in (B.10) fail to commute

$$\lim_{T \rightarrow 0} \lim_{\epsilon \rightarrow 0} T \left(\frac{\partial S}{\partial \epsilon} \right)_{T,X} \neq \lim_{\epsilon \rightarrow 0} \lim_{T \rightarrow 0} T \left(\frac{\partial S}{\partial \epsilon} \right)_{T,X}, \tag{B.16}$$

making the physical interpretation of the relation problematic. The formally correct entropy-extremality relation (B.11) and its corollaries (B.15), treats ϵ as a free parameter in the model that is taken to zero at the end of the calculation. In a physical application of the low-energy EFT (2.1), which is derived by matching onto a UV completion, the parameter ϵ takes a fixed finite value, and thermodynamic quantities at extremality are correctly calculated by taking $T \rightarrow 0$ *before* making a small ϵ expansion. In such a situation, the entropy-extremality relation is only approximately valid in the *near-extremal* regime $T \sim \epsilon$. A correct treatment at extremality for finite ϵ is described in detail in appendix C.

C Entropy shifts of very-near extremal black holes

The entropy shift in (3.24), which we rewrite as

$$(\Delta M)_{Q,T=0} = -T_0 (\Delta S)_{Q,M} \quad (T_0 > 0), \tag{C.1}$$

was derived at *finite* temperature, and as described in appendix B, is a valid approximation in the near-extremal regime $T_0 \sim \sqrt{\epsilon}$. For black holes with $T_0 \lesssim \sqrt{\epsilon}$, which we will call very-near extremal black holes, the shift in the horizon is not analytic in ϵ and (C.1) is no longer valid. The presence of a double root at extremality implies that the shift to the horizon radius, and therefore to the entropy, is proportional to $\sqrt{\epsilon}$. Nonetheless, a proportionality between the entropy shift and extremality shift still holds. This has been stressed in [9, 33], where it was shown that, at extremality, the same combination of EFT coefficients that appear in the extremality shift is required to be positive in order to ensure that the entropy shift is real. In fact, near extremality the form of the entropy-extremality relationship is nearly identical to the one of [12], up to an extra factor of 1/2. The correct relation takes the form

$$(\Delta M)_{Q,T=0} = -\frac{1}{2}T (\Delta S)_{Q,M} \quad (T_0 = 0), \tag{C.2}$$

where now T is the temperature of the corrected black hole. In particular, while the leading order temperature vanishes, the corrected solution is evaluated at fixed mass M and charge Q , and hence is no longer extremal.

In order to see the transition between (C.1) and (C.2), we may revisit the derivation of the mass and entropy shifts of section 3. Since the entropy shift is calculated at fixed mass and charge, it is convenient to rewrite the geometry shift in terms of M and Q . In this case, the radial function takes the form

$$g(r) = g_0(r) + \Delta g = k - \frac{\hat{M}}{r^{d-2}} + \frac{\hat{Q}^2}{4r^{2(d-2)}} + \frac{r^2}{l^2} + \epsilon \Delta g, \tag{C.3}$$

where here Δg takes into account the higher-derivative corrections to the geometry along with the contribution of the ρ factor to the physical mass M in (2.23). Note that, to avoid extra volume factors, we have defined

$$M = \frac{\omega_{d-1}(d-1)}{16\pi} \hat{M}, \quad Q = \frac{\omega_{d-1}}{16\pi} \sqrt{\frac{(d-2)(d-1)}{2}} \hat{Q}. \tag{C.4}$$

The entropy shift at fixed (M, Q) can then be obtained from the horizon shift according to (3.18)

$$(\Delta S)_{Q,M} = \frac{\omega_{d-1}(d-1)}{4} r_0^{d-2} \Delta r_h + \dots, \quad (\text{C.5})$$

where r_0 is the leading order horizon location, and the ellipses denote the additional higher order corrections to the Wald entropy. These additional terms will be unimportant in either of the near extremal or very near extremal limits, where the shift to the horizon radius dominates.

Our goal now is to rewrite the horizon shift Δr_h in terms of the mass shift and temperature. For the mass shift, we always consider the shift of the extremal mass at fixed charge, which can be taken from (3.7)

$$(\Delta M)_{Q,T=0} = \frac{\omega_{d-1}(d-1)}{16\pi} r_0^{d-2} \epsilon \Delta g. \quad (\text{C.6})$$

Note that the term in (3.7) proportional to ρ is absorbed into the shift Δg in this expression. In any case, we see that in order to connect the entropy shift (C.5) to the mass shift (C.6), we would like to have a relation between Δr_h and Δg . The first step here is to use the fact that the radial function $g(r)$ vanishes at the horizon

$$0 = g(r_h) = g_0(r_0 + \Delta r_h) + \epsilon \Delta g = g_0(r_0) + \Delta r_h g'_0(r_0) + \frac{1}{2} (\Delta r_h)^2 g''_0(r_0) + \dots + \epsilon \Delta g. \quad (\text{C.7})$$

Note that we have not expanded $\Delta g(r_h) = \Delta g(r_0) + \dots$ since we already consider it to be a small quantity. Since r_0 is the uncorrected horizon, $g_0(r_0)$ always vanishes. However, we need to keep the next two terms in the expansion since $g'_0(r_0)$ will vanish for a leading order extremal black hole [9]. The horizon shift can then be obtained by solving

$$\Delta r_h g'_0(r_0) + \frac{1}{2} (\Delta r_h)^2 g''_0(r_0) = -\epsilon \Delta g, \quad (\text{C.8})$$

or equivalently

$$\Delta r_h g'_0(r_0) + \frac{1}{2} (\Delta r_h)^2 g''_0(r_0) = -\frac{16\pi}{\omega_{d-1}(d-1)} r_0^{-d+2} (\Delta M)_{Q,T=0}. \quad (\text{C.9})$$

In principle, we can now insert the horizon shift from (C.9) into the expression (C.5) to obtain a general relation between the entropy shift and the mass shift. However, as it stands, this result would depend on the derivatives $g'_0(r_0)$ and $g''_0(r_0)$ of the radial function. To replace this with more physical quantities, we note that the temperature of the corrected black hole can be obtained from $g(r)$ according to

$$T = \frac{1}{4\pi} g'(r_h) = \frac{1}{4\pi} g'(r_0 + \Delta r_h) = \frac{1}{4\pi} (g'_0(r_0) + \Delta r_h g''_0(r_0) + \dots + \epsilon \Delta g'). \quad (\text{C.10})$$

Although this depends on $\Delta g'$, this term can be ignored as long as we are in the near (or very near) extremal limit since either $g'_0(r_0)$ or $\Delta r_h g''_0(r_0)$ will dominate. More precisely, $g'_0(r_0) = \mathcal{O}(\epsilon^0)$ unless it vanishes, and when it does, (C.8) along with $g''_0(r_0) = \mathcal{O}(\epsilon^0)$ demonstrates that $\Delta r_h g''_0(r_0) = \mathcal{O}(\sqrt{\epsilon})$. As a result, we can take

$$T = \frac{1}{4\pi} (g'_0(r_0) + \Delta r_h g''_0(r_0)). \quad (\text{C.11})$$

In the general case, this provides only a single relation between the derivatives $g'(r_0)$ and $\Delta r_h g''(r_0)$ and the temperature T . However, we can also introduce the temperature of the uncorrected black hole, $T_0 = g'_0(r_0)/4\pi$, which can be allowed to vanish. We now solve for the derivatives

$$g'_0(r_0) = 4\pi T_0, \quad \Delta r_h g''_0(r_0) = 4\pi(T - T_0). \quad (\text{C.12})$$

Inserting this into (C.9) gives us the desired relation between the horizon shift and mass shift

$$(\Delta M)_{Q,T=0} = -\frac{\omega_{d-1}(d-1)}{4} r_0^{d-2} \frac{T_0 + T}{2} \Delta r_h, \quad (\text{C.13})$$

which yields

$$(\Delta M)_{Q,T=0} = -\frac{T_0 + T}{2} (\Delta S)_{Q,M}, \quad (\text{C.14})$$

upon substitution of (C.5). Note that this relation is expected to break down away from extremality where the various approximations taken above will no longer be valid.

The relation (C.14) encompasses both the near and very near extremal cases, (C.1) and (C.2). When working somewhat away from extremality, we start with $T_0 > 0$. In this case, we have the scaling

$$\Delta r_h \sim \epsilon, \quad \Delta S \sim \epsilon, \quad \Delta M \sim \epsilon, \quad (\text{C.15})$$

along with $T = T_0 + \mathcal{O}(\epsilon)$. Taking this into account then directly yields (C.1). On the other hand, if we start from extremality, $T_0 = 0$, we have instead

$$\Delta r_h \sim \sqrt{\epsilon}, \quad \Delta S \sim \sqrt{\epsilon}, \quad \Delta M \sim \epsilon, \quad (\text{C.16})$$

and $T = \mathcal{O}(\sqrt{\epsilon})$, which leads to the extremal case (C.2). The transition between these two cases occurs when $\Delta T \sim T$, which more precisely defines the very near extremal limit.

C.1 Entropy shift for heterotic black holes

The above discussion can be applied to an interesting example. Recent work has considered the leading α' -corrections to dyonic Reissner-Nordström black holes embedded in heterotic string theory [58, 59]. Though the four-dimensional backgrounds considered in these papers are asymptotically flat, we would like to briefly comment on them in connection with the universal entropy-extremality relationship.

To connect with the above, we would like to consider the corrections to the geometry. These are given [59] by the radial function

$$g(r) = 1 - \frac{2M}{r} + \frac{p^2}{2r^2} + \Delta g(r), \quad \Delta g(r) = -\alpha' \frac{p^2}{4r^4} \left[1 - \frac{3M}{2r} + \frac{11p^2}{40r^2} \right], \quad (\text{C.17})$$

where the constant p denotes the charge in the notation of [59]. With these variables, the leading-order solution is extremal when $M = p/\sqrt{2}$. Using this expression and (C.6), we can compute the shift in the extremal mass at a given charge p ,

$$(\Delta M)_{p,T=0} = \frac{1}{2} r_0 \Delta g(r_0) = -\alpha' \frac{1}{4(p/\sqrt{2})} \left(1 - \frac{3}{2} + \frac{22}{40} \right) = -\frac{1}{80} \frac{\alpha'}{p/\sqrt{2}}, \quad (\text{C.18})$$

where we have computed Δg at extremality in the last equality. This agrees with [59].

Next, we consider the entropy. In [59], this is shown to be

$$S = \pi r_h^2 \left\{ 1 + \alpha' \left[\frac{M}{r_h^3} - \frac{3p^2}{8r_h^4} \right] \right\}. \quad (\text{C.19})$$

Just as we have found in the body of the paper, this entropy includes corrections from the higher-derivative terms in the Lagrangian; these are the explicit α' terms that appear in the expression above. It also includes corrections from the horizon radius, which are implicit in the r_h^2 that appears in the first term. Thus, we may rewrite the entropy more explicitly as

$$S = \pi r_0^2 \left\{ 1 + 2 \frac{\Delta r_h}{r_0} + \alpha' \left[\frac{M}{r_0^3} - \frac{3p^2}{8r_0^4} \right] \right\}. \quad (\text{C.20})$$

The corrections are then given by

$$(\Delta S)_{M,p} = 2\pi r_0 \Delta r_h + \alpha' \pi \left[\frac{M}{r_0} - \frac{3p^2}{8r_0^2} \right]. \quad (\text{C.21})$$

We need the temperature to complete the relation. Recall that

$$T = \frac{1}{4\pi} (g'_0(r_0) + \Delta r g''_0(r_0) + \dots). \quad (\text{C.22})$$

The order of both of these corrections in α' will depend on the order of Δr_h . Thus, to determine the entropy-extremality relationship we need to compute the correction to the horizon radius, which may be obtained from the geometry shift via equation (C.8). This procedure gives different results away from extremality and near extremality. Away from extremality, we have

$$T \sim T_0 = \frac{1}{4\pi} g'_0(r_0), \quad (\text{C.23})$$

and the horizon shift is⁹

$$\Delta r = -\frac{\Delta g}{4\pi T_0}. \quad (\text{C.24})$$

The corresponding entropy shift becomes

$$(\Delta S)_{M,p} = \frac{1}{T_0} \alpha' \frac{p^2}{8r_0^3} \left[1 - \frac{3M}{2r_0} + \frac{11p^2}{40r_0^2} \right] + \alpha' \pi \left[\frac{M}{r_0} - \frac{3p^2}{8r_0^2} \right]. \quad (\text{C.25})$$

If we then take this expression to extremality, we take the limit where $T_0 \rightarrow 0$. Since the first term dominates in this limit, we may write the entropy shift as

$$T_0 (\Delta S)_{M,p} = \frac{1}{80} \frac{\alpha'}{p/\sqrt{2}}, \quad (\text{C.26})$$

⁹it is not easy to compare this to (4.12) of [59] because they have numerically inverted the expression to ensure that Δr is only a function of M and p . However, both results for Δr lead to the same entropy shift at extremality.

where we have made use of the extremal value of the mass. This gives the correct relation away from extremality,

$$(\Delta M)_{p,T=0} = -T_0(\Delta S)_{M,p}. \tag{C.27}$$

Now let us perform the same analysis in the very-near extremal limit. This is the limit where $T_0 = 4\pi g'_0 \rightarrow 0$, so that for the temperature we have

$$T = \frac{1}{4\pi} \Delta r_h g''_0(r_0) \tag{C.28}$$

to leading order in α' . In this limit, we now extract a different result from (C.8),

$$\Delta r_h = \sqrt{-2 \frac{\Delta g}{g''_0(r_0)}}. \tag{C.29}$$

In this case, the first term of (C.21) still dominates the entropy corrections, but here this is because it is $\mathcal{O}(\sqrt{\alpha'})$, rather than because the denominator diverges. Thus, very-near extremality the entropy shift is

$$(\Delta S)_{M,p} = 2\pi r_0 \Delta r_h, \tag{C.30}$$

and therefore in this limit we find

$$T(\Delta S)_{M,p} = \frac{1}{2} r_0 \Delta r_h^2 g''_0(r_0) = -r_0 \Delta g = \frac{1}{40} \frac{\alpha'}{p/\sqrt{2}}. \tag{C.31}$$

This gives the very-near extremal relation:

$$(\Delta M)_{p,T=0} = -\frac{1}{2} T(\Delta S)_{M,p}. \tag{C.32}$$

This agrees with the discussion of [59] and confirms their claim that this is the correct form of the relationship in the extremal limit. It also confirms the scaling relations in (C.15) and (C.16). In both cases, ΔM is order α' . However, in the very-near extremal limit we see the non-analytic behavior discussed above, namely that $T \sim \sqrt{\alpha'}$ and $\Delta S \sim \sqrt{\alpha'}$. It is also important to note that in both cases, we took the extremal limit. However, the relative size of T_0 and α' dictates the correct order of limits to take. The derivation “away from extremality” remains valid for $T_0 \geq \sqrt{\alpha'}$. For very small temperatures — that is, for $T_0 \lesssim \sqrt{\alpha'}$ — the additional subtleties discussed in this section arise, and we need to use the “very-near extremal” derivation. Nonetheless, in both cases there is a proportionality between the mass shift and the temperature times the entropy shift.

Open Access. This article is distributed under the terms of the Creative Commons Attribution License ([CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

References

- [1] C. Vafa, *The string landscape and the swampland*, [hep-th/0509212](#) [INSPIRE].
- [2] E. Palti, *The swampland: introduction and review*, *Fortsch. Phys.* **67** (2019) 1900037 [[arXiv:1903.06239](#)] [INSPIRE].
- [3] N. Arkani-Hamed, L. Motl, A. Nicolis and C. Vafa, *The string landscape, black holes and gravity as the weakest force*, *JHEP* **06** (2007) 060 [[hep-th/0601001](#)] [INSPIRE].
- [4] Y. Kats and P. Petrov, *Effect of curvature squared corrections in AdS on the viscosity of the dual gauge theory*, *JHEP* **01** (2009) 044 [[arXiv:0712.0743](#)] [INSPIRE].
- [5] C. Cheung and G.N. Remmen, *Naturalness and the weak gravity conjecture*, *Phys. Rev. Lett.* **113** (2014) 051601 [[arXiv:1402.2287](#)] [INSPIRE].
- [6] G. Shiu, P. Soler and W. Cottrell, *Weak gravity conjecture and extremal black holes*, *Sci. China Phys. Mech. Astron.* **62** (2019) 110412 [[arXiv:1611.06270](#)] [INSPIRE].
- [7] C. Cheung, J. Liu and G.N. Remmen, *Proof of the weak gravity conjecture from black hole entropy*, *JHEP* **10** (2018) 004 [[arXiv:1801.08546](#)] [INSPIRE].
- [8] S. Andriolo, D. Junghans, T. Noumi and G. Shiu, *A tower weak gravity conjecture from infrared consistency*, *Fortsch. Phys.* **66** (2018) 1800020 [[arXiv:1802.04287](#)] [INSPIRE].
- [9] Y. Hamada, T. Noumi and G. Shiu, *Weak gravity conjecture from unitarity and causality*, *Phys. Rev. Lett.* **123** (2019) 051601 [[arXiv:1810.03637](#)] [INSPIRE].
- [10] A.M. Charles, *The weak gravity conjecture, RG flows, and supersymmetry*, [arXiv:1906.07734](#) [INSPIRE].
- [11] C.R.T. Jones and B. McPeak, *The black hole weak gravity conjecture with multiple charges*, *JHEP* **06** (2020) 140 [[arXiv:1908.10452](#)] [INSPIRE].
- [12] G. Goon and R. Penco, *Universal relation between corrections to entropy and extremality*, *Phys. Rev. Lett.* **124** (2020) 101103 [[arXiv:1909.05254](#)] [INSPIRE].
- [13] Y. Nakayama and Y. Nomura, *Weak gravity conjecture in the AdS/CFT correspondence*, *Phys. Rev. D* **92** (2015) 126006 [[arXiv:1509.01647](#)] [INSPIRE].
- [14] D. Harlow, *Wormholes, emergent gauge fields, and the weak gravity conjecture*, *JHEP* **01** (2016) 122 [[arXiv:1510.07911](#)] [INSPIRE].
- [15] M. Montero, G. Shiu and P. Soler, *The weak gravity conjecture in three dimensions*, *JHEP* **10** (2016) 159 [[arXiv:1606.08438](#)] [INSPIRE].
- [16] M. Montero, *A holographic derivation of the weak gravity conjecture*, *JHEP* **03** (2019) 157 [[arXiv:1812.03978](#)] [INSPIRE].
- [17] H.S. Reall and J.E. Santos, *Higher derivative corrections to Kerr black hole thermodynamics*, *JHEP* **04** (2019) 021 [[arXiv:1901.11535](#)] [INSPIRE].
- [18] R.C. Myers, M.F. Paulos and A. Sinha, *Holographic hydrodynamics with a chemical potential*, *JHEP* **06** (2009) 006 [[arXiv:0903.2834](#)] [INSPIRE].
- [19] Y. Kats, L. Motl and M. Padi, *Higher-order corrections to mass-charge relation of extremal black holes*, *JHEP* **12** (2007) 068 [[hep-th/0606100](#)] [INSPIRE].
- [20] S. Cremonini, J.T. Liu and P. Szepietowski, *Higher derivative corrections to R-charged black holes: boundary counterterms and the mass-charge relation*, *JHEP* **03** (2010) 042 [[arXiv:0910.5159](#)] [INSPIRE].

- [21] V. Balasubramanian and P. Kraus, *A stress tensor for Anti-de Sitter gravity*, *Commun. Math. Phys.* **208** (1999) 413 [[hep-th/9902121](#)] [[INSPIRE](#)].
- [22] O. Loukas, D. Orlando, S. Reffert and D. Sarkar, *An AdS/EFT correspondence at large charge*, *Nucl. Phys. B* **934** (2018) 437 [[arXiv:1804.04151](#)] [[INSPIRE](#)].
- [23] R.M. Wald, *Black hole entropy is the Noether charge*, *Phys. Rev. D* **48** (1993) 3427 [[gr-qc/9307038](#)] [[INSPIRE](#)].
- [24] S.W. Hawking and W. Israel, *General relativity*. Cambridge University Press, Cambridge U.K. (1979).
- [25] J.W. York, *Role of conformal three geometry in the dynamics of gravitation*, *Phys. Rev. Lett.* **28** (1972) 1082 [[INSPIRE](#)].
- [26] G.W. Gibbons and S.W. Hawking, *Action integrals and partition functions in quantum gravity*, *Phys. Rev. D* **15** (1977) 2752 [[INSPIRE](#)].
- [27] S.W. Hawking and S.F. Ross, *Duality between electric and magnetic black holes*, *Phys. Rev. D* **52** (1995) 5865 [[hep-th/9504019](#)] [[INSPIRE](#)].
- [28] M. Henningson and K. Skenderis, *The holographic Weyl anomaly*, *JHEP* **07** (1998) 023 [[hep-th/9806087](#)] [[INSPIRE](#)].
- [29] 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](#)].
- [30] K. Skenderis, *Lecture notes on holographic renormalization*, *Class. Quant. Grav.* **19** (2002) 5849 [[hep-th/0209067](#)] [[INSPIRE](#)].
- [31] J.T. Liu and W.A. Sabra, *Hamilton-Jacobi counterterms for Einstein-Gauss-Bonnet gravity*, *Class. Quant. Grav.* **27** (2010) 175014 [[arXiv:0807.1256](#)] [[INSPIRE](#)].
- [32] C. Cheung, J. Liu and G.N. Remmen, *Entropy bounds on effective field theory from rotating Dyonic black holes*, *Phys. Rev. D* **100** (2019) 046003 [[arXiv:1903.09156](#)] [[INSPIRE](#)].
- [33] G.J. Loges, T. Noumi and G. Shiu, *Thermodynamics of 4D dilatonic black holes and the weak gravity conjecture*, *Phys. Rev. D* **102** (2020) 046010 [[arXiv:1909.01352](#)] [[INSPIRE](#)].
- [34] A. Chamblin, R. Emparan, C.V. Johnson and R.C. Myers, *Charged AdS black holes and catastrophic holography*, *Phys. Rev. D* **60** (1999) 064018 [[hep-th/9902170](#)] [[INSPIRE](#)].
- [35] A. Chamblin, R. Emparan, C.V. Johnson and R.C. Myers, *Holography, thermodynamics and fluctuations of charged AdS black holes*, *Phys. Rev. D* **60** (1999) 104026 [[hep-th/9904197](#)] [[INSPIRE](#)].
- [36] J.-y. Shen, R.-G. Cai, B. Wang and R.-K. Su, *Thermodynamic geometry and critical behavior of black holes*, *Int. J. Mod. Phys. A* **22** (2007) 11 [[gr-qc/0512035](#)] [[INSPIRE](#)].
- [37] C. Cheung and G.N. Remmen, *Positivity of curvature-squared corrections in gravity*, *Phys. Rev. Lett.* **118** (2017) 051601 [[arXiv:1608.02942](#)] [[INSPIRE](#)].
- [38] S. Nojiri and S.D. Odintsov, *On the conformal anomaly from higher derivative gravity in AdS/CFT correspondence*, *Int. J. Mod. Phys. A* **15** (2000) 413 [[hep-th/9903033](#)] [[INSPIRE](#)].
- [39] M. Blau, K.S. Narain and E. Gava, *On subleading contributions to the AdS/CFT trace anomaly*, *JHEP* **09** (1999) 018 [[hep-th/9904179](#)] [[INSPIRE](#)].
- [40] A. Buchel, R.C. Myers and A. Sinha, *Beyond $\eta_s = 1/4\pi$* , *JHEP* **03** (2009) 084 [[arXiv:0812.2521](#)] [[INSPIRE](#)].

- [41] D.M. Hofman and J. Maldacena, *Conformal collider physics: Energy and charge correlations*, *JHEP* **05** (2008) 012 [[arXiv:0803.1467](#)] [[INSPIRE](#)].
- [42] G. Policastro, D.T. Son and A.O. Starinets, *The Shear viscosity of strongly coupled $N = 4$ supersymmetric Yang-Mills plasma*, *Phys. Rev. Lett.* **87** (2001) 081601 [[hep-th/0104066](#)] [[INSPIRE](#)].
- [43] A. Buchel and J.T. Liu, *Universality of the shear viscosity in supergravity*, *Phys. Rev. Lett.* **93** (2004) 090602 [[hep-th/0311175](#)] [[INSPIRE](#)].
- [44] S. Cremonini, *The shear viscosity to entropy ratio: a status report*, *Mod. Phys. Lett. B* **25** (2011) 1867 [[arXiv:1108.0677](#)] [[INSPIRE](#)].
- [45] P. Kovtun and A. Ritz, *Universal conductivity and central charges*, *Phys. Rev. D* **78** (2008) 066009 [[arXiv:0806.0110](#)] [[INSPIRE](#)].
- [46] S. Cremonini and P. Szepietowski, *Generating temperature flow for η_s with higher derivatives: from Lifshitz to AdS*, *JHEP* **02** (2012) 038 [[arXiv:1111.5623](#)] [[INSPIRE](#)].
- [47] S. Cremonini, U. Gürsoy and P. Szepietowski, *On the temperature dependence of the shear viscosity and holography*, *JHEP* **08** (2012) 167 [[arXiv:1206.3581](#)] [[INSPIRE](#)].
- [48] S. Hellerman, D. Orlando, S. Reffert and M. Watanabe, *On the CFT operator spectrum at large global charge*, *JHEP* **12** (2015) 071 [[arXiv:1505.01537](#)] [[INSPIRE](#)].
- [49] S.S. Gubser, *Breaking an abelian gauge symmetry near a black hole horizon*, *Phys. Rev. D* **78** (2008) 065034 [[arXiv:0801.2977](#)] [[INSPIRE](#)].
- [50] S.A. Hartnoll, C.P. Herzog and G.T. Horowitz, *Holographic superconductors*, *JHEP* **12** (2008) 015 [[arXiv:0810.1563](#)] [[INSPIRE](#)].
- [51] S.A. Hartnoll, C.P. Herzog and G.T. Horowitz, *Building a holographic superconductor*, *Phys. Rev. Lett.* **101** (2008) 031601 [[arXiv:0803.3295](#)] [[INSPIRE](#)].
- [52] F. Denef and S.A. Hartnoll, *Landscape of superconducting membranes*, *Phys. Rev. D* **79** (2009) 126008 [[arXiv:0901.1160](#)] [[INSPIRE](#)].
- [53] S. Bhattacharyya, S. Minwalla and K. Papadodimas, *Small hairy black holes in $AdS_5 \times S^5$* , *JHEP* **11** (2011) 035 [[arXiv:1005.1287](#)] [[INSPIRE](#)].
- [54] O.J.C. Dias, R. Monteiro, H.S. Reall and J.E. Santos, *A scalar field condensation instability of rotating anti-de Sitter black holes*, *JHEP* **11** (2010) 036 [[arXiv:1007.3745](#)] [[INSPIRE](#)].
- [55] T. Crisford, G.T. Horowitz and J.E. Santos, *Testing the weak gravity — Cosmic censorship connection*, *Phys. Rev. D* **97** (2018) 066005 [[arXiv:1709.07880](#)] [[INSPIRE](#)].
- [56] G.T. Horowitz and J.E. Santos, *Further evidence for the weak gravity — Cosmic censorship connection*, *JHEP* **06** (2019) 122 [[arXiv:1901.11096](#)] [[INSPIRE](#)].
- [57] G.T. Horowitz, J.E. Santos and B. Way, *Evidence for an electrifying violation of cosmic censorship*, *Class. Quant. Grav.* **33** (2016) 195007 [[arXiv:1604.06465](#)] [[INSPIRE](#)].
- [58] P.A. Cano, T. Ortín and P.F. Ramirez, *On the extremality bound of stringy black holes*, *JHEP* **02** (2020) 175 [[arXiv:1909.08530](#)] [[INSPIRE](#)].
- [59] P.A. Cano, S. Chimento, R. Linares, T. Ortín and P.F. Ramírez, *α' corrections of Reissner-Nordström black holes*, *JHEP* **02** (2020) 031 [[arXiv:1910.14324](#)] [[INSPIRE](#)].