Published by Institute of Physics Publishing for SISSA
Received: *November 21, 2006*
Accepted: *December 22, 2006*
Published: *January 10, 2007*

Holography, the Second Law and a $\widetilde{C}$-function in higher curvature gravity

Daniel Cremades
DAMTP, Centre for Mathematical Sciences,
Wilberforce Road, Cambridge, CB3 0WA, U.K.
E-mail: d.cremades@damtp.cam.ac.uk

Ernesto Lozano-Tellechea
Department of Particle Physics, Weizmann Institute of Science
Rehovot 76100, Israel
E-mail: ernesto.lozano@weizmann.ac.il

Abstract: We analyze the Second Law of black hole mechanics and the generalization of the holographic bound for general theories of gravity. We argue that both the possibility of defining a holographic bound and the existence of a Second Law seem to imply each other via the existence of a certain "c-function" (i.e. a never-decreasing function along outgoing null geodesic flow). We are able to define such a "c-function", that we call $\widetilde{C}$, for general theories of gravity. It has the nontrivial property of being well defined on general spacelike surfaces, rather than just on a spatial cross-section of a black hole horizon. We argue that $\widetilde{C}$ is a suitable generalization of the concept of "area" in any extension of the holographic bound for general theories of gravity. Such a function is provided by an algorithm which is similar (although not identical) to that used by Iyer and Wald to define the entropy of a dynamical black hole. In a class of higher curvature gravity theories that we analyze in detail, we are able to prove the monotonicity of $\widetilde{C}$ if several physical requirements are satisfied. Apart from the usual ones, these include the cancellation of ghosts in the spectrum of the gravitational Lagrangian. Finally, we point out that our $\widetilde{C}$-function, when evaluated on a black hole horizon, constitutes by itself an alternative candidate for defining the entropy of a dynamical black hole.

Keywords: Classical Theories of Gravity, Black Holes.

JHEP01(2007)045

Contents

1. Introduction

It is known that, in a general theory of gravity including higher derivative couplings, the entropy of a stationary black hole is no longer given by a quarter of the area of the event horizon. It was shown in [1, 2] that the definition of entropy which obeys the First Law of black hole mechanics is, instead, given by the integral of a particular local quantity on a spatial cross-section Σ of the event horizon. In a theory which does not depend on covariant derivatives of the Riemann tensor, the expression for the entropy is:

$$S_{\rm BH} = -2\pi \int_{\Sigma} \frac{\partial L}{\partial R_{abcd}}\, \epsilon_{ab}\, \epsilon_{cd} \sqrt{h}\, d\Omega\,. \qquad (1.1)$$

where the Lagrangian of the theory is $\mathcal{L} = \sqrt{-g}\, L$ (i.e. L is a scalar), partial derivatives with respect to the Riemann tensor have to be taken as if it was a quantity independent of the metric,[1] ϵ_{ab} denotes the binormal to the horizon cross-section (i.e. the volume element in the 2-space perpendicular to it) normalized so that $\epsilon_{ab}\epsilon^{ab} = -2$, and $\sqrt{h}\, d\Omega$ is the volume element induced on Σ. The overall normalization is for units $G_N = 1$. In General Relativity the above expression reduces to one quarter of the area of Σ but, in more general theories, it differs from it.

This generalization of the expression for BH entropy poses two important questions:

1 Does this expression for $S_{\rm BH}$ obey a Second Law of black hole mechanics?

2 How should we generalize the notion of "area of the boundary" of an arbitrary physical system in establishing any holographic bound in a general theory of gravity?

Question **1** above was considered in [3] for a class of theories where the gravitational Lagrangian is a function of the scalar curvature, i.e. $L = L(R)$. It was shown that, if certain conditions of "positivity of energy" (together with Cosmic Censorship and assuming asymptotic flatness, as in the GR case) are satisfied, then a Second Law holds for $S_{\rm BH}$ in the $L(R)$ theories whenever the equations of motion are obeyed. In general, proving a Second Law is a difficult task, since it is not clear which physical conditions have to be obeyed in a general theory of gravity for the Second Law to hold. Also, imposing the equations of motion is straightforward in principle, but technically difficult in practice.

On the other hand, Question **2** arises because, as it is well known, the upper limit on the degrees of freedom of a system being set by the area of its boundary comes from black hole physics. The original argument of 't Hooft [4] uses the fact that any physical system confined in volume V has to collapse to form a black hole long before its entropy exceeds the area of the boundary of V in Planck units. Adding entropy to the resulting black hole would just increase its size. Therefore, the maximal entropy that can be physically realized in a given volume is that of a black hole of the same size. However, when we go further and use this fact to establish a universal holographic bound, we are concerned with the intrinsic properties of the boundary of an arbitrary spacetime region. i.e. we do not restrict ourselves to the case in which the physical system is a black hole, nor to the case where the boundary of the system is an event horizon. Consequently, when the entropy of a black hole is no longer given by its area, Question **2** arises naturally.

The first natural guess to answer Question **2** would be to use equation (1.1) but, instead of restricting ourselves to the cross-section of the event horizon of a black hole,

[1] It was shown in [2] that, in any covariant theory, the dependence of L on the derivatives of the metric can always be written as a dependence on the Riemann tensor and its covariant derivatives only, i.e. $L = L(g_{ab}, R_{abcd}, \ldots)$. The derivative in (1.1) means the derivative of L when rewritten in such a way.

simply perform the integral over the surface of interest (the boundary of the spacetime region considered). However, this is not satisfactory since (as discussed originally in [1] and further in [5] — see also [2]) the local quantity that one integrates in (1.1) is only well defined when evaluated on an spatial cross-section of a Killing horizon, but it is ambiguous when evaluated on a general spacelike surface.

Below we will show that a generalized concept of "area of the boundary" (which reduces to (1.1) in the case of a black hole horizon) exists and it is well defined for general, spacelike, codimension two surfaces, irrespective of them being cross sections of a Killing horizon or not. Our definition will be very close to the proposal in [2] for the entropy of a dynamical (i.e. non-stationary) black hole, although it will differ slightly from it.

In this paper we would like to emphasize the fact that Questions **1** and **2** are related to a third one, namely:

3 Can we define a "holographic c-function" in general theories of gravity?

Let us stress that we will use the name "c-function" throughout the paper just by analogy, and that by this we do not necessarily mean the true holographic dual of a field theory c-function. By a "c-function" we just want to refer here to a non-decreasing function along "outgoing" null geodesic flow (i.e., for flat/AdS/dS asymptotics, a non-decreasing function of the radial variable). We will see that monotonicity of such a "c-function" is the behaviour required on physical grounds to address the Second Law. In particular, we will show below that, in AdS, this function is a strictly increasing one, and not a constant as one would expect for a true field-theory c-function in a CFT.[2] Nevertheless, we shall keep sometimes the name "c-function" for it, but we prefer to denote it by $\widetilde{C}$. In the case of Einstein gravity coupled to arbitrary matter fields obeying the null energy condition, the existence of such a function in static backgrounds was proved recently in [6]. Earlier attempts to define a ("true") holographic c-function include [7].[3]

Roughly, the basic relation between Questions **1**, **2** and **3** can be put like this: assume that $\widetilde{C}$ exists and that it is well defined on arbitrary spacelike surfaces which need not be the horizon of a black hole. Then, if $\widetilde{C}$ can be shown to equal the black hole entropy when evaluated on a spatial cross-section of a black hole horizon, then it clearly constitutes a natural candidate for the answer of Question **2**. Assume further that $\widetilde{C}$ is non-decreasing along a congruence of outgoing radial null geodesics. Then this means that it cannot decrease along the affine parameter of the null geodesic generators of the horizon, which in turn implies the Second Law. Along the paper we will elaborate more carefully on all these statements.

Our considerations are completely general and, in particular, our definition of the $\widetilde{C}$-function should be applicable in general theories of gravity. However, as an application of our proposal, in this paper we restrict ourselves to a particular class of theories. These are the theories with a Lagrangian $\mathcal{L} = \sqrt{-g}\, L$ given by:

$$L = L_g(R, P, Q) + L_m(g_{ab}, \psi, \partial\psi) \tag{1.2}$$

[2] We will elaborate a bit more on this in section 5.2.3.

[3] Considerations on holography in higher curvature gravity from a different perspective have recently appeared in [14].

JHEP01(2007)045

where R denotes the Ricci scalar, $P = R_{ab} R^{ab}$, $Q = R_{abcd}R^{abcd}$, and ψ denotes any matter fields.[4] The consistency of these theories has been studied recently in [8]. We will show below that, for these theories, if the matter fields obey the null energy condition and, further, if the theory is ghost-free, a $\widetilde{C}$-function exists which satisfies all the above requirements. Throughout the paper we will restrict ourselves to four dimensions, but our results should generalize in a straightforward way to any number of spacetime dimensions.

Finally, let us mention that our definition of $\widetilde{C}$ constitutes, by itself, an alternative candidate to the proposal in [2] for the definition of the entropy of a dynamical black hole. We will comment a bit on this in the Conclusions.

2. Holographic c-functions, Raychaudhuri equation and the Second Law

In order to fix ideas, let us start by reviewing the result of [6] as well as some properties of the Raychaudhuri equation and the Second Law of black hole dynamics in GR.

2.1 A c-function in general relativity

The authors of [6] showed that, in four dimensional Einstein gravity coupled to any matter fields subject to the null energy condition, any spherically symmetric, asymptotically flat spacetime:

$$ds^2 = -a(r)\, dt^2 + a^{-1}(r)\, dr^2 + b(r)\, d\Omega^2\,, \tag{2.1}$$

admits a "c-function", $\widetilde{C}$, which is given by the area $\mathcal{A}$ of a transverse sphere at radius r:

$$\widetilde{C} \equiv \frac{1}{4}\,\mathcal{A}(r) = \pi b(r)\,. \tag{2.2}$$

It was shown in [6] that the equations of motion imply that $\widetilde{C}$ is a never-decreasing function of the radial variable. Note that, in a black hole spacetime, $\widetilde{C}$ at the horizon equals the black hole entropy.

2.2 Relation to Raychaudhuri equation

As already noted in [6], the monotonicity of $\widetilde{C}$ can be understood as a simple consequence of the Raychaudhuri equation. Let us briefly review this.

Consider an arbitrary spacetime (static or otherwise) with metric g_{ab}, and let p be a point of this spacetime. Consider a null geodesic congruence in a vicinity of p and, in particular, the geodesic of the congruence passing through that point. Let $k \equiv k^a\partial_a = d/d\lambda$ be the corresponding tangent vector, λ being the affine parameter of the congruence. Define another null vector on the tangent space of p, $n \equiv n^a\partial_a = d/d\sigma$, satisfying $k^a n_a = -1$ and $k^a\nabla_a n^b = 0$. Consider finally two linearly independent spacelike vectors $\eta^a_{(i)}\partial_a = d/dx^i$, $i = 1, 2$, orthogonal to k and n. The corresponding dual forms define a differential volume form in the space orthogonal to k^a and n^a given by:

$$dV = \sqrt{h}\; dx^1 \wedge dx^2\,, \tag{2.3}$$

[4] We are assuming that L_m that matter fields do not couple explicitly to the curvature, and that further couplings do not arise from terms of the form $\nabla\psi$ (see footnote 1). This is always the case of any theory with arbitrary couplings to scalar, axion and gauge fields.

JHEP01(2007)045

with

$$h_{ij} = \eta^c_{(i)} \eta^d_{(j)} \, g_{cd} \,. \tag{2.4}$$

Raychaudhuri equation tells us about the local behaviour of the expansion ϑ of the congruence. The expansion is given by:

$$\vartheta = \frac{d\sqrt{h}/d\lambda}{\sqrt{h}} = \frac{d \log \mathcal{A}}{d\lambda} \,, \tag{2.5}$$

where $\mathcal{A}$ is the transverse area spanned by the congruence (i.e. the magnitude of an area element of the "wavefront"). From this expression one demonstrates that:

$$\frac{d\vartheta}{d\lambda} = -\frac{1}{2}\vartheta^2 - \sigma^{ab}\sigma_{ab} - R_{ab}k^a k^b, \tag{2.6}$$

where σ^{ab} is the shear of the congruence. This is Raychaudhuri equation. Note that this equation relies only on the geometry of the spacetime, and hence is independent of the dynamics of the theory. From this expression, using the fact that, in General Relativity, $T_{ab}k^a k^b = R_{ab}k^a k^b$, and assuming the null energy condition ($T_{ab}\zeta^a\zeta^b \geq 0$ for every null vector ζ^a), one gets the inequality:

$$\frac{d\vartheta}{d\lambda} \leq -\frac{1}{2}\vartheta^2 \leq 0 \,. \tag{2.7}$$

Assume now asymptotic flatness. This means, in particular, that we have a well defined "radial coordinate" r. If we take the congruence to be "outgoing" (i.e. $dr/d\lambda > 0$) then we have, at infinity, $\mathcal{A} \sim r^2$. Therefore we get:

$$\vartheta \to 0^+ \tag{2.8}$$

asymptotically. Assume now that ϑ cannot diverge. This, together with the fact that $d\vartheta/d\lambda \leq 0$, means that:

$$\vartheta \geq 0 \,, \tag{2.9}$$

for all λ, and therefore $\mathcal{A}$ can never be decreasing along the outgoing null geodesic flow.

Concerning the behaviour of ϑ, we note that it can only diverge at a naked singularity (forbidden if one assumes Cosmic Censorship) or at a caustic. However, it is easy to prove from equation (2.7) that ϑ cannot diverge at any finite affine parameter λ if, initially, we have $\vartheta(\lambda = 0) > 0$. In such a case this implies the existence of a "c-function" $\widetilde{C} \sim \mathcal{A}$. Conversely, if a "c-function" $\widetilde{C}$ exists, then ϑ is never negative along outgoing null geodesic flow. We will next see that, in such a case, this immediately implies the Second Law.

Consider now the particular case of a static spacetime (2.1). In such a case $\mathcal{A}$ in (2.5) is precisely the area of transverse spheres, as in eq. (2.2). In terms of the radial coordinate of the metric (2.1) this just means that the "c-function" $\widetilde{C}$ defined in (2.2) is a non-decreasing function of r [6].[5]

[5]Previous attempts to find holographic c-functions have also been based on the Raychaudhuri equation [7].

2.3 Relation to the Second Law

Consider now a black hole spacetime. Establishing the Second Law of black hole mechanics amounts to prove that:

$$\frac{dS_{\rm BH}}{d\lambda} = \int_\Sigma \vartheta \sqrt{h}\, d\Omega \geq 0\,, \tag{2.10}$$

where, in the expression above, Σ is a cross-section of the event horizon, $\sqrt{h}$ is an area element of it, and ϑ is defined along the outgoing null geodesic congruence orthogonal to Σ (k^a in (2.6) becomes here the null vector field tangent to the horizon generators). Therefore, to prove (2.10) it is enough to prove that $\theta \geq 0$ at every point along the (future directed) generators of the (future) event horizon [12].

The latter requirement ($\vartheta \geq 0$ along outgoing null geodesic flow and, in particular, also along the horizon generators) is precisely what we have just proved in the previous paragraph, eq. (2.9). Moreover, it is well known that $\vartheta < 0$ at any point along the null horizon generators violates Cosmic Censorship [12]. Recall that the requirements to prove that $\vartheta > 0$ this were the EOMs to be satisfied, the null energy condition, asymptotic flatness and Cosmic Censorship.

However, as we have seen, this is equivalent to the statement of the $\widetilde{C}$-function defined in (2.2) being monotonic. Although in [6] the function $\widetilde{C}$ was originally defined just for static spacetimes, this symmetry property was not used at all in the proof of it being monotonic. The only thing that we need is the function $\widetilde{C}$ to be defined along null geodesic flow (the latter being spherically symmetric or otherwise). To relate monotonicity of $\widetilde{C}$ to the Second Law, the only additional requirement is that, if $\widetilde{C}$ is evaluated on a black hole horizon, it should equal its entropy.[6]

3. $L_g(R)$ theories

All these results reviewed above are valid in General Relativity. Next we wish to consider the simplest class of Lagrangians of the form (1.2) in which the gravitational Lagrangian is only a function of the scalar curvature, $L_g = L_g(R)$. The validity of the Second Law for these theories was considered in [3]. Here we want to review their result, discuss the Second Law, and finally obtain a consistency condition that will be used later in the general case of arbitrary Lagrangians of the form (1.2).

3.1 "Generalized" Raychaudhuri equation

Consider an arbitrary spacetime and a null geodesic congruence defined as in section 2.2. We define now the function:

$$\widetilde{C} = -2\pi \frac{\partial L_g}{\partial R_{abcd}} \epsilon_{ab}\epsilon_{cd} \sqrt{h} = 4\pi L_R \sqrt{h}\,, \tag{3.1}$$

[6]Heuristically, note that the fact that $\widetilde{C} = \mathcal{A}/4$ is a non-decreasing function of r fits nicely with the holographic principle ("the total number of degrees of freedom in a region is bounded by the area of its boundary") and the meaning of the c-theorem in field theory ("the number of degrees of freedom cannot decrease along the RG flow").

JHEP01(2007)045

where $\epsilon_{ab} = k_a n_b - k_b n_a$ stands for the binormal of the surface defined by dV in eq. (2.3), and we define L_R as $L_R \equiv \partial L_g / \partial R$. In analogy to the GR case, we now define:

$$\widetilde{\vartheta} \equiv \frac{d \log \widetilde{C}}{d\lambda} = \vartheta + \frac{1}{L_R} k^a \nabla_a L_R \tag{3.2}$$

(where ϑ is the expansion defined as in eq. (2.5)). Hence:

$$\frac{d\widetilde{\vartheta}}{d\lambda} = -\frac{1}{2}\vartheta^2 - \sigma^{ab}\sigma_{ab} - R_{ab}k^a k^b - \frac{1}{L_R^2}\left(k^b \nabla_b L_R\right)^2 + \frac{1}{L_R} k^a k^b \nabla_a \nabla_b L_R, \tag{3.3}$$

where we have used the Raychaudhuri equation. On the other hand, the Einstein equations for a Lagrangian of the form (1.2) with $L_g = L_g(R)$ are given by:

$$L_R \, R_{ab} + \left(\nabla^2 L_R - \frac{1}{2} L_g\right) g_{ab} - \nabla_a \nabla_b L_R = T_{ab}\,, \tag{3.4}$$

where:

$$T_{ab} = \frac{1}{2} L_m \, g_{ab} - \frac{\partial L_m}{\partial g^{ab}}\,. \tag{3.5}$$

Since k^a is null, we get, from (3.4):

$$T_{ab} k^a k^b = L_R \, R_{ab} \, k^a k^b - k^a k^b \nabla_a \nabla_b L_R\,. \tag{3.6}$$

Finally, using the above equation in (3.3), we have:

$$\frac{d\tilde{\vartheta}}{d\lambda} = -\frac{1}{2}\vartheta^2 - \sigma^{ab}\sigma_{ab} - \frac{1}{L_R^2}\left(k^b \nabla_b L_R\right)^2 - \frac{1}{L_R} T_{ab} k^a k^b\,. \tag{3.7}$$

This implies that:

$$\frac{d\widetilde{\vartheta}}{d\lambda} \leq 0 \tag{3.8}$$

whenever the null energy condition holds *if*, additionally:

$$L_R > 0\,. \tag{3.9}$$

Equation (3.8) is the analog to the inequality (2.7) obtained in GR from the Raychaudhuri equation for a congruence of null geodesics, in which the area $\mathcal{A}$ swept by the congruence has been replaced here by $\widetilde{C}$.

The "generalized Raychaudhuri equation" (3.7) implies that $d\widetilde{\vartheta}/d\lambda \leq 0$ if the condition $L_R > 0$ is met.[7] Such a condition is dependent both on the theory and on the specific background considered. However, suppose that $L_R < 0$ for a specific background whose Ricci scalar we denote by R_0. Then, the weak field expansion of the metric around such a background would be the one obtained from the Lagrangian:

$$L_g = L_g(R_0) + \left(\frac{\partial L_g}{\partial R}\right)_{R=R_0} (R - R_0) + \cdots\,, \tag{3.10}$$

[7]Note that this is linked to the matter satisfying the null/"antinull" energy condition.

and therefore we would get a negative gravitational coupling in the spacetime regions where (3.9) does not hold.[8] Note that this condition (positivity of the effective G_N) is a natural generalization of the null energy condition for this class of theories, since the n.e.c. can be rephrased as the requirement of gravity being universally attractive.

These results (with focus on the Second Law, which we now turn to discuss) were derived in [3], and the same consistency condition (3.9) was obtained there. In [3] the condition $L_R > 0$ was derived by going to an auxiliary theory (dynamically equivalent to the $L_g(R)$-theory) described by Einstein gravity coupled to an scalar field.

3.2 A "c-theorem" and the Second Law

Eq. (3.7) means that, when $L_R > 0$:

$$\widetilde{\vartheta} = \vartheta + \frac{dL_R/d\lambda}{L_R} \tag{3.11}$$

is a decreasing function. Consider now an asymptotically flat background (with radial coordinate r) and, as in section 2.2, a congruence of "outgoing" null geodesics (i.e. $dr/d\lambda > 0$). In such a case we have, asymptotically:

$$R \sim 1/r^3\,, \qquad \vartheta \sim 1/r > 0\,, \tag{3.12}$$

since, in asymptotically flat spacetime of dimension D, all the components of the Riemann tensor vanish at least like $\sim 1/r^{D-1}$. (We have that $\vartheta \sim 1/r > 0$ asymptotically by the same argument used in section 2.2.) If the gravitational interaction is of the form $L_g = R + \mathcal{O}(R^2) + \cdots$ we find that:

$$\widetilde{\vartheta} \to 0^+\,, \tag{3.13}$$

asymptotically. Assuming again Cosmic Censorship (and by an argument analogous to that used in section 2.2) this implies, together with the fact of $\widetilde{\vartheta}$ being monotonically decreasing with r, that:

$$\widetilde{\vartheta} \geq 0 \tag{3.14}$$

for all values of r, meaning that the function $\widetilde{C}$ defined in eq. (3.1) is a never decreasing one along the null geodesic congruence. Let us stress again that, conversely, if one proves monotonicity of $\widetilde{C}$, then the condition (3.14) follows immediately.[9]

Note that, by construction, in the case of a black hole spacetime, $\widetilde{C}$ equals the black hole entropy (1.1) when evaluated at the horizon. The analogous of eq. (2.10) for this class of theories is therefore [3]:

$$\frac{dS_{\rm BH}}{d\lambda} = \int_\Sigma \widetilde{\vartheta}\,\sqrt{h}\,d\Omega \geq 0\,. \tag{3.15}$$

The relation the "c-theorem" just proved to the Second Law of black hole mechanics is exactly analogous to the proof sketched in section 2.3 if we replace ϑ by $\widetilde{\vartheta}$. The fact that, also in this case, $\widetilde{\vartheta} < 0$ at the horizon generators also violates Cosmic Censorship was proved in [3].

[8] Remember that the Newton constant G_N is defined by the coefficient of the term $\sim (\partial h)^2$ in a weak field expansion of the metric, $\delta g_{ab} \sim h_{ab}$.

[9] We insist on this since it is (3.14) what implies the Second Law.

4. A $\widetilde{C}$-function in higher curvature theories

In the last sections we have seen that it is possible to define a function $\widetilde{C}$ in GR and in theories with gravitational Lagrangian $L_g(R)$ which has the following properties:

a) It can be evaluated in any arbitrary spacelike surface.

b) When evaluated at the event horizon of a black hole it equals its entropy.

c) If certain physical conditions and certain boundary conditions are satisfied, then $\widetilde{C}$ is a non-decreasing function along outgoing null geodesic flow.

Given a $\widetilde{C}$-function with these properties, the following consequences are implied:

- a) and b) make of $\widetilde{C}$ a natural candidate to generalize the notion of "area" in any generalization of the holographic bound in these higher curvature theories.
- b) and c) imply the Second Law of black hole mechanics in these theories.

These conclusions explicitly answer Questions **1** and **2** in the Introduction. At this point, however, we have to stress the following issue: in more general theories of gravity, a $\widetilde{C}$-function defined as in (3.1), i.e.:

$$\widetilde{C} \sim \frac{\partial L}{\partial R_{abcd}} \epsilon_{ab} \epsilon_{cd} \sqrt{h} \tag{4.1}$$

cannot be well defined in general. The ambiguity arises from the fact that one can always add a total divergence or a topological term[10] to the Lagrangian. This addition leaves the dynamics of the theory unchanged but, on the other hand, if such terms depend on the curvature, they will clearly change the expression of $\widetilde{C}$ if defined as above.[11] Of course, the expression (4.1) *is* unambiguously defined when evaluated on a spacelike cross-section of the event horizon of a stationary black hole (cf. eq. (1.1)), but not on general spacelike surfaces, thus failing to fulfill requirement a) above. Below we briefly review what makes a BH horizon special in this respect.

[10] By this we mean terms like, for instance, a Gauss-Bonnet density in four dimensions. Such term just adds a topological constant to the action (the Euler number of the manifold) and thus it does not affect the equations of motion.

[11] For simplicity, in this paper we restrict ourselves to the case in which L does not depend on derivatives of the Riemann tensor. Therefore, in four dimensions, the only cases we could be concerned with are the ambiguities arising from the addition to the Lagrangian of a Gauss-Bonnet density (first Euler class) or the first Pontrjagin class (however, only the Gauss-Bonnet density will belong to the particular class of theories considered in detail below). Let us point out that, in general theories depending on derivatives of the curvature up to order m, the generalization of eq. (1) is:

$$S_{\rm BH} = -2\pi \int_\Sigma \left(\frac{\partial L}{\partial R_{abcd}} - \nabla_{a_1} \frac{\partial L}{\partial \nabla_{a_1} R_{abcd}} + \cdots + (-1)^m \nabla_{(a_1} \cdots \nabla_{a_m)} \frac{\partial L}{\partial \nabla_{(a_1} \cdots \nabla_{a_m)} R_{abcd}} \right) \epsilon_{ab}\, \epsilon_{cd} \sqrt{h}\, d\Omega \,,$$

and thus all considerations above are qualitatively the same in this case. Therefore we do not foresee any essential difficulty in extending our results in a straightforward manner to more general theories depending on derivatives of the curvature.

JHEP01(2007)045

Our result for the $L_g(R)$-theories is nevertheless valid, since, in four dimensions, one cannot build total derivatives out of R alone. However, any definition of $\widetilde{C}$ in a general theory will necessarily have to face this issue.[12] Hence we see that, in a general theory, we have to refine the "naive" definition of $\widetilde{C}$ so that it has the following property:

d) Its definition has to be absent from any ambiguities. In particular, we require from $\widetilde{C}$ to be insensitive to the addition of a total derivative or a "topological density" to the Lagrangian.

Another nontrivial challenge when considering more general theories of gravity is to find out which are the "physical conditions" to be fulfilled in order to have property c) above. In particular, it is nontrivial to foresee what should be the corresponding conditions of "positivity of energy". In a general theory, one has to find how the generalization of condition (3.9) comes about, if such condition is physically meaningful, and if imposing such condition makes sense on physical grounds.

This section is devoted to the construction of a $\widetilde{C}$-function that satisfies by construction properties a), b) and d) above. The proof of property c) beyond the case $L_g(R)$ is complicated in general. Nevertheless, we will see that, in Lagrangians of the form $L_g = L_g(R, P, Q)$ (see eq. (1.2)), whenever the theory is known to be physically meaningful,[13] property c) above is also obeyed.

4.1 Construction of the $\widetilde{C}$-function

Let us show now how one can define a function $\widetilde{C}$ fulfilling all the requirements a), b) and d) presented at the beginning of this section. In particular, let us start by discussing in detail points a) and d) above when the $\widetilde{C}$-function is taken to be of the form (4.1). That is, we start by analyzing the generalization provided by formula (1.1) when evaluated on a general spacelike surface $\mathcal{S}$ rather than just on a black hole horizon, i.e.:

$$S_{\rm BH}[g, \mathcal{S}] = -2\pi \int_{\mathcal{S}} \frac{\partial L}{\partial R_{abcd}} \epsilon_{ab}\, \epsilon_{bc} \sqrt{h}\, d\Omega\,, \tag{4.2}$$

where, in the l.h.s., g makes explicit reference to the background, and $\mathcal{S}$ indicates the surface on which the functional $S_{\rm BH}$ is to be evaluated. As discussed in the previous section, this expression is, in general, sensitive to the addition to the Lagrangian of a total derivative depending upon the curvature. Consider however the case of a stationary black hole spacetime. In such a case, it is believed that the BH horizon is always a Killing horizon of some Killing vector ξ^a.[14] In [2] it was shown that all additional terms arising in $S_{\rm BH}$ from the addition of an exact form to the Lagrangian are always proportional to ξ^a. Therefore,

[12] In fact, we will give below a general definition of $\widetilde{C}$ which is free of these ambiguities, and we will see that it reduces to the form (4.1) for the particular case of theories with a Lagrangian $L_g = L_g(R)$.

[13] In particular we are talking about the cancellation of ghosts in the spectrum — see below.

[14] As it is well known, this fact was proved in [12] for GR. While this fact has not been proved for an arbitrary higher curvature theory of gravity, no counterexamples are known to the best of our knowledge. In particular, this property automatically holds for static, spherically symmetric black holes, as argued in [1].

if the integration surface is chosen to be the bifurcation surface, $\Sigma_{\rm bif}$, of the Killing horizon (which is, by definition, the surface at which ξ^a vanishes), these contributions cancel.[15] Furthermore, the authors of [5] (see also [2]) showed that:

$$S_{\rm BH}[g, \Sigma_{\rm bif}] = S_{\rm BH}[g, \Sigma]\,, \tag{4.3}$$

where Σ above is any arbitrary spacelike cross-section of the horizon. As a result, $S_{\rm BH}[g, \Sigma]$ is always well defined, but, on the other hand, it will be clearly ambiguous if the surface of integration is taken to be an arbitrary surface $\mathcal{S}$.

The crucial point in this paper is to find a generalization of $S_{\rm BH}[g, \Sigma]$ which is well defined on any arbitrary surface $\mathcal{S}$. Moreover, we necessarily require such a generalization to reduce to $S_{\rm BH}[g, \Sigma]$ when g_{ab} describes a stationary BH spacetime and the surface $\mathcal{S}$ is chosen to be a cross-section Σ of the BH horizon. This problem was essentially solved by Iyer and Wald in [2]. In that reference they provided an algorithm to define the entropy of a dynamical (i.e. non-stationary) black hole. Such a case confronts the same problems we have here since, in general, the event horizon will no longer be a Killing horizon (a non-stationary black hole spacetime may have no Killing vectors at all). Our solution to the problem at hand is clearly inspired by Iyer and Wald's proposal for the entropy of a dynamical black hole; however, it differs from it. For the interested reader, we review their proposal in appendix A (the remaining of the paper is however self-contained and does not require the use of the results reviewed in the appendix).

4.1.1 General idea

In order to provide a generalization of (4.2) which is well defined on arbitrary spacelike surfaces, we proceed as follows. As in [2], instead of modifying the functional form of $S_{\rm BH}[g, \mathcal{S}]$ given in (4.2), we provide an algorithm that deforms the spacetime metric g_{ab} in the vicinity of $\mathcal{S}$. We call $\widetilde{g}_{ab}$ the metric of this deformed spacetime. This deformation is such that $\mathcal{S}$ becomes the bifurcation surface of a bifurcate Killing horizon in the (artificial) spacetime $\widetilde{g}_{ab}$. This will make the expression $S_{\rm BH}[\widetilde{g}, \mathcal{S}]$ to be automatically well defined (for exactly the same reasons that make $S_{\rm BH}[g, \Sigma]$ to be well defined at the horizon of a stationary black hole), hence having fulfilled properties a) and d) above. On the other hand, the particular deformation $g \to \widetilde{g}$ that we choose will be such that properties b) and c) above will also be obeyed.[16]

The basic observation to find such a deformation $\widetilde{g}_{ab}$ is the following. Any spacetime which, in certain coordinates (U, V, x^1, x^2), has a metric of the form:

$$ds^2 = F(UV, x^1, x^2)\, dU dV + g_{ij}(UV, x^1, x^2)\, dx^i dx^j\,, \tag{4.4}$$

[15] Below we will consider explicitly the addition to the Lagrangian of a topological term (in particular, the addition of a Gauss-Bonnet density). We will see that, in such a case, the final expression for the entropy *does* get modified. However, this modification is just the shift of the entropy by a constant.

[16] The main difference between our proposal and that of [2] is that, if we deform the original spacetime in the way proposed in [2], then property c) above does not hold. See the appendix for an explicit comparison.

JHEP01(2007)045

(where F and g_{ij} are arbitrary functions of (x^1, x^2) and the single combination UV) admits the following Killing vector:

$$\xi = U\partial_U - V\partial_V \,, \tag{4.5}$$

and has a bifurcate Killing horizon (the corresponding Killing vector being ξ above) at $UV = 0$, with bifurcation surface (parametrized by (x^1, x^2)) at $U = V = 0$. In the metric above, the null coordinates (U, V) (that we will define below in a precise manner) can be thought as a generalization of the usual Kruskal coordinates of the Schwarzschild black hole. In fact, the simplest spacetime with a metric of the kind (4.4) is Schwarzschild spacetime. The Killing vector (4.5) generates Lorentz boosts in the subspace parametrized by (U, V). Therefore (following the nomenclature used in [2]) we will refer to such metrics as *boost-invariant* metrics.[17]

Given any spacetime metric g_{ab} and an arbitrary spacelike surface $\mathcal{S}$, the basic idea is the following. First, go to an appropriate system of "Kruskal coordinates" (U, V) (to be defined below) in the space orthogonal to $\mathcal{S}$, parametrized in such a way that $\mathcal{S}$ lies at $U = V = 0$. In general (unless $\mathcal{S}$ is already the bifurcation surface of a bifurcate Killing horizon), the resulting metric will not be of the form (4.4). Therefore the next thing to do will be to define a new (unphysical) spacetime $\widetilde{g}_{ab}$ in a vicinity of $U = V = 0$, such that $\widetilde{g}_{ab}$ has the general form (4.4). Of course, the deformation has to be such that $\widetilde{g}_{ab}$ is obtained from g_{ab} in a unique fashion (at least modulo "gauge freedom" — see below). Let us next proceed to explain the construction of $\widetilde{g}_{ab}$.

4.1.2 Boost-invariant projection of g_{ab} at $\mathcal{S}$

Consider an arbitrary spacetime g_{ab} and an arbitrary spacelike surface $\mathcal{S}$. As said, the first thing to do is to find an appropriate system of null coordinates (U, V) perpendicular to $\mathcal{S}$, such that $\mathcal{S}$ is at $U = V = 0$. For a completely generic spacetime, a "universal" coordinate system, which is always possible to define in a vicinity of $\mathcal{S}$, was introduced in [10, 2] (we review this coordinate system in the appendix). It has the right property of automatically bringing the metric to the form (4.4) whenever $\mathcal{S}$ is already the bifurcation surface of a bifurcate Killing horizon.

However, the situation is somewhat simpler in the case of static spacetimes. Since all examples considered below will be of this kind, let us write here the explicit coordinate system that we will be using in such cases. Consider a static spacetime parametrized as in eq. (2.1). In such a case, we define the following system of "Kruskal coordinates" (U, V) given by:

$$\begin{aligned} u &= t + \int \frac{dr}{a(r)}\,, \qquad\qquad & U &= e^{u}\,, \\ v &= t - \int \frac{dr}{a(r)}\,. \qquad\qquad & V &= -e^{-v}\,. \end{aligned} \tag{4.6}$$

[17] Note that defining $U = T + X$, $V = T - X$, the metric (4.4) only depends on the Lorentz-invariant combination $-T^2 + X^2$.

JHEP01(2007)045

In this coordinate system, the metric (2.1) becomes:

$$ds^2 = \frac{a[r(UV)]}{UV}\, dU dV + b[r(UV)]\, d\Omega^2 \,. \tag{4.7}$$

Note that this metric has always the boost-invariant form (4.4). Actually, if we have a regular event horizon at $r = r_H$, we have that $a(r) \sim (r - r_H)$ near the horizon, and therefore:

$$r \to r_H \quad \Leftrightarrow \quad UV \to 0 \,, \tag{4.8}$$

as desired.[18] Note that, in terms of the original coordinates, the Killing vector (4.5) is $\xi = \partial_t$.

Remember that $\widetilde{C}$ is to be defined along outgoing null geodesic flow. Therefore, in the particular case of static metrics (2.1) that we are considering, the surfaces $\mathcal{S}$ we will be interested in will be spheres. An arbitrary sphere $\mathcal{S}$ in the spacetime (4.7) will be located at (U_0, V_0, x^i) with fixed U_0, V_0. Therefore the next step is to shift coordinates:

$$U \to (U + U_0) \,, \qquad V \to (V + V_0) \,, \tag{4.9}$$

so that we get $\mathcal{S}$ at $U = V = 0$. In general, since $\mathcal{S}$ will not be the bifurcation surface of a Killing horizon, the resulting metric will not be boost-invariant there. Therefore we define the *boost-invariant projection of g_{ab} at $\mathcal{S}$*, denoted by $\widetilde{g}_{ab}$, as follows:

$$\widetilde{g}_{ab} \equiv \sum_{n=0}^{N/2} (C_{ab})_n \, (UV)^n \,, \tag{4.10}$$

(where we have omitted the explicit dependence on the x^i), and where N is the order of the higher derivative of the metric appearing in the Lagrangian. The coefficients $(C_{ab})_n$ above are to be chosen as follows:

- First take $(C_{ab})_0$ to be equal to g_{ab} at $U = V = 0$, in order to make both metrics to coincide at $\mathcal{S}$.

- Next, choose the remaining coefficients $(C_{ab})_n$ such that all independent curvature invariants constructed out of $\widetilde{g}_{ab}$ that contribute to $S_{\rm BH}[\widetilde{g}, \mathcal{S}]$ match those constructed out of g_{ab} at $U = V = 0$.

The last point requires some explanation. Notice that, in the case of a boost-invariant metric, the number of curvature invariants that will enter in the final form of $S_{\rm BH}$ is always lower than in a generic (not boost-invariant) case. This is because, for a metric like (4.10), the way in which $\mathcal{S}$ is embedded in $\widetilde{g}_{ab}$ always makes both extrinsic curvatures of $\mathcal{S}$ to vanish. This implies identities between the intrinsic curvature invariants of $\mathcal{S}$ and those of

[18] Since (4.6) always brings a static metric to the form (4.4), it might seem that we will always have a Killing horizon. Of course, this is not the case. If we start with a metric without horizons, the region $UV = 0$ will not be a physical part of the spacetime. The example of AdS considered below will be an explicit example of this.

$\widetilde{g}_{ab}$ at $\mathcal{S}$, which will reduce the number of independent curvature invariants appearing in the formula for the entropy.

In particular, one can derive the following relation (already used in [2], and which we will use below) relating the intrinsic curvature of $\mathcal{S}$ with the curvature invariants of $\widetilde{g}_{ab}$ at $\mathcal{S}$: let k^a and n^a be two null vectors normal to $\mathcal{S}$ along the coordinate U and V, respectively, and normalized so that $k^a n_a = -1$. Then, if ${}^{(2)}R$ is the intrinsic curvature of $\mathcal{S}$, the following identity holds:

$$ {}^{(2)}R = \Big(R - 2t^{ab}R_{ab} + t^{ac}t^{bd}R_{abcd} \Big)_{U=V=0} \tag{4.11} $$

where

$$ t_{ab} \equiv -k_a n_b - n_a k_b \tag{4.12} $$

is the induced metric in the space orthogonal to $\mathcal{S}$. Such relations do not generically hold if the metric is not boost invariant.[19]

Let us finish this section with the following comments:

- Note that the boost-invariant projection of g_{ab} at $\mathcal{S}$ does not always exist. This is because we have at our disposal a maximum number of coefficients to fit and, in principle, it is not ensured that they can be chosen to match all possible invariants contributing to $S_{\rm BH}[\widetilde{g}, \mathcal{S}]$.[20] However, we will see that, in the case of the theories $L_g(R, P, Q)$, the number of curvature invariants entering in the final expression for $S_{\rm BH}[\widetilde{g}, \mathcal{S}]$ precisely coincides with the number of coefficients at our disposal *if* the theory is ghost-free, thus making the construction of $\widetilde{g}_{ab}$ always possible. We therefore expect that, in a general theory of gravity, cancellation of the non-physical degrees of freedom will make the election of a boost-invariant projection of the original metric always possible.

- Note that *any* metric of the form (4.10) is boost-invariant (with $\mathcal{S}$ being the bifurcation surface of a Killing horizon) for whatever choice of the coefficients $(C_{ab})_n$. Iyer and Wald's proposal in [2] is in fact a similar way of deforming the original spacetime, but with a different prescription to choose these coefficients (see appendix). Our particular choice will be justified a posteriori, since we will see that it is precisely this choice what makes property c) at the beginning of this section to be satisfied.

- From the technical point of view, in this section we have focused in the case of static spacetimes. We have done this just for simplicity of the exposition, but our results remain valid (and our prescription for defining $\widetilde{g}_{ab}$ unchanged) for general

[19] The precise statement is that the relation (4.11) holds at a given point if, at that point, the extrinsic curvatures with respect to k^a and n^a ($K_{ab} \equiv h_a{}^i h_b{}^j \nabla_i k_j$ and $N_{ab} \equiv h_a{}^i h_b{}^j \nabla_i n_j$, with $h_{ab} = g_{ab} + k_a n_b + n_a k_b$ the induced metric in $\mathcal{S}$) satisfy $K_a{}^b N_b{}^a = K_c{}^c N_d{}^d$.

[20] Of course, the value of N in the expansion (4.10) can be higher than the order of the highest derivative of the metric appearing in the Lagrangian. However, in such a case, all "extra" coefficients are redundant since they do not contribute to the curvature at $U = V = 0$ (this is what we meant by "gauge freedom" in section 4.1.1).

spacetimes. The only technical difference is that, in a general case, the Kruskal coordinates defined in (4.6) will not be appropriate. However, one can always define the (U, V) coordinates as explained in [10, 2] (see appendix), all the rest being exactly the same.

4.1.3 General definition of $\widetilde{C}$

Finally, let us provide here our definition of the $\widetilde{C}$-function in a general theory of gravity. Given a solution g_{ab} of the equations of motion, and given a spacelike surface $\mathcal{S}$, we define $\widetilde{C}$ as:

$$\widetilde{C}[g, \mathcal{S}] \equiv S_{\rm BH}[\widetilde{g}, \mathcal{S}] = -2\pi \int_{\mathcal{S}} \frac{\partial L(\widetilde{g})}{\partial \widetilde{R}_{abcd}} \, \widetilde{\epsilon}_{ab} \, \widetilde{\epsilon}_{cd} \, \sqrt{h} \, d\Omega \,, \tag{4.13}$$

where $\widetilde{g}_{ab}$ is the boost-invariant projection of g_{ab} at $\mathcal{S}$, as defined in the previous section. Notice that such a $\widetilde{C}$-function automatically satisfies properties a), b) and d) enumerated at the beginning of this section. Properties a) and d) are satisfied by construction. The fact that $\widetilde{C}$ equals the entropy of a stationary black hole when evaluated on a spacelike surface Σ of its event horizon is maybe less evident. However, this fact follows from the definition of $\widetilde{g}_{ab}$: remember that this metric is chosen so that every invariant contributing to the final form of $S_{\rm BH}[\widetilde{g}, \mathcal{S}]$ equals those of the original spacetime at $\mathcal{S}$. At the bifurcation surface $\Sigma_{\rm bif}$ of a regular event horizon, the spacetime metric g_{ab} will be automatically boost-invariant (cf. [10] or last section for the particular case of static spacetimes). Therefore, the same relations like (4.11) will hold both for g_{ab} and $\widetilde{g}_{ab}$ at $\Sigma_{\rm bif}$. This will make appear the same independent curvature invariants in $S_{\rm BH}[\widetilde{g}, \Sigma_{\rm bif}]$ as in $S_{\rm BH}[g, \Sigma_{\rm bif}]$. Since $\widetilde{g}_{ab}$ is chosen such that those are equal, then one necessarily has:

$$\widetilde{C}[g, \Sigma_{\rm bif}] = S_{\rm BH}[g, \Sigma_{\rm bif}] \,. \tag{4.14}$$

This, together with the result (4.3) of [5], implies that:

$$\widetilde{C}[g, \Sigma] = S_{\rm BH}[g, \Sigma] \,. \tag{4.15}$$

Below we will see an explicit example of this in the case of the cosmological horizon of de Sitter space.

Summarizing, we have been able to construct an algorithm that allows to define a $\widetilde{C}$-function in theories of higher curvature gravity, subject to properties a), b) and d) above. However, we have not shown that property c) (namely, that $\widetilde{C}$ is an non-decreasing function along outgoing radial null geodesics) holds. As already emphasized, this is a nontrivial task in a completely generic theory and in an arbitrary background; not only because of the technical difficulty involved in the calculations, but also due to the fact that the physical consistency requirements necessary to prove condition c) are not known in general. Therefore, at this point, we can only leave our construction as a proposal whose validity must in principle be tested in a case by case scenario. In any case, we provide in the following section the explicit construction of the $\widetilde{C}$-function in a particular case: theories whose Lagrangian is given by eq. (1.2) in maximally symmetric backgrounds, where the physical consistency conditions have been recently worked out [8]. We will see that in these cases property c) above is indeed satisfied.

JHEP01(2007)045

5. $\widetilde{C}$ in theories of the form $L_g = L_g(R,P,Q)$

We will devote this section to illustrate the construction of the $\widetilde{C}$-function in a particular case: maximally symmetric backgrounds in theories where the Lagrangian is of the form given by eq. (1.2), i.e. $L_g = L_g(R,P,Q)$ where $P \equiv R_{ab}R^{ab}$ and $Q \equiv R_{abcd}R^{abcd}$. We will first review the consistency conditions for these backgrounds in these kind of theories, and then we will compute $\widetilde{C}$ following the prescription given in section 4. We will show how this function satisfies property c) above whenever the physical consistency conditions are met.

5.1 Consistency of the $L_g(P,Q,R)$ theories

In general, it is known that higher derivative couplings introduce new degrees of freedom in the theory. The case of theories linear in R, R^2, P and Q in four dimensions was studied a long time ago by Stelle [9]. It is found that the spectrum of these theories includes, in addition to the massless graviton, a massive scalar (which can be tachyonic or not) and a massive spin-2 field which is always a ghost.

Gravity theories with a gravitational Lagrangian of the form $L_g(R,P,Q)$ (see eq. (1.2)) have been studied recently in [8]. Their consistency and stability of a given background depends of course of the particular function L_g chosen and on the particular background considered. Unfortunately, the perturbative expansion of these theories about a generic background is not known. However, on maximally symmetric backgrounds, it can be shown that the spectrum and the perturbative expansion of the $L_g(R,P,Q)$-theory are equivalent those of the following Lagrangian:

$$L_g(R,C) = -2\Lambda + \delta R + \frac{1}{6m_0^2}R^2 + \frac{1}{2m_2^2}C_{abcd}C^{abcd}\,, \tag{5.1}$$

about the same maximally symmetric background. C_{abcd} above denotes the Weyl tensor, and the parameters δ, m_0^2 and m_2^2 appearing in $L_g(R,C)$ are calculated from the original Lagrangian $L_g(R,P,Q)$ as follows [8]:

$$\begin{aligned}
\delta &= \Big(L_R - RL_{RR} - R^2(L_{RP} + \tfrac{2}{3}L_{RQ}) - R^3(\tfrac{1}{4}L_{PP} + \tfrac{1}{9}L_{QQ} + \tfrac{1}{3}L_{PQ})\Big)_0\,,\\
m_0^{-2} &= 3\Big(L_{RR} + \tfrac{2}{3}(L_P + L_Q) + R(L_{RP} + \tfrac{2}{3}L_{RQ}) +\\
&\qquad + R^2(\tfrac{1}{4}L_{PP} + \tfrac{1}{9}L_{QQ} + \tfrac{1}{3}L_{PQ})\Big)_0\,,\\
m_2^{-2} &= \Big(L_P + 4L_Q\Big)_0\,.
\end{aligned} \tag{5.2}$$

The subscript "0" in the r.h.s. denotes evaluation in the maximally symmetric background, and L_R, L_{RP}, etc. denote the corresponding partial derivatives of L_g with respect to R, P and Q. The value of the cosmological constant $\Lambda = -\frac{1}{2}(L_g + \cdots)_0$ will not be needed in what follows. The value of Λ is related to the curvature of the maximally symmetric space by the equation of motion:

$$(2L_Q + 3L_P)R^2 + 6L_R R - 12L_g = 0\,. \tag{5.3}$$

The theory (5.1) falls into the class of theories studied in [9],[21] and therefore the spectrum is just as explained above. Actually, the masses of the scalar and the spin-2 field are proportional, respectively, to m_0 and m_2. Thus, while it is true that theories of the kind $L_g(R,P,Q)$ have, in general, a ghost-like spin-2 field in the spectrum, we see that we can decouple the ghost if we have $(L_P + 4L_Q)_0 = 0$. In particular, all theories with a Lagrangian of the form:

$$L_g(R,P,Q) = L_g(R, T = Q - 4P) \tag{5.4}$$

will be automatically ghost-free. In [8] it was shown that this condition to have a ghost-free theory remains valid in the case of a FRW cosmological background.

So the first thing we have to do is to make the theory ghost-free. In order to achieve this, we restrict ourselves to Lagrangians of the form (5.4). Then the ghost decouples from the theory and, further, $L_g(R,C)$ in (5.1) reduces to:

$$L_g(R,C) = -2\Lambda + \delta R + \frac{1}{6m_0^2}R^2\,, \tag{5.5}$$

which is just of the form $L_g = L_g(R)$ studied in section 3.[22] There we showed that a consistency requirement in this kind of theories is (see also [3]):

$$\frac{\partial L_g(R,C)}{\partial R} > 0\,. \tag{5.6}$$

Using (5.2) and the no-ghost condition we see that this translates into:

$$L_R - 2L_T R > 0\,, \tag{5.7}$$

where the functions appearing above are functions only of R and T, and these invariants are those of the maximally symmetric background considered.

5.2 $\widetilde{C}$-function in maximally symmetric backgrounds

Let us see what we get in these theories from our general considerations to find the $\widetilde{C}$-function. First we observe that, for any metric g_{ab}, the functional form of $S_{\rm BH}[g,\mathcal{S}]$ in (4.2) for a generic Lagrangian $L_g(R,P,Q)$ is given by:

$$S_{\rm BH}[g,\mathcal{S}] = 4\pi\int_{\mathcal{S}}(L_R + L_P\,R^{ab}t_{ab} + 2L_Q\,R^{abcd}t_{ac}t_{bd})\,\sqrt{h}\,d\Omega\,, \tag{5.8}$$

[21] This is due to the fact that any theory with L_g linear in R, R^2, P and Q can be put into the form (5.1) by using the identity:

$$C_{abcd}C^{abcd} = Q - 2P + \frac{1}{3}R^2\,,$$

and by using the fact that, in four dimensions, the Gauss-Bonnet density

$$GB = R^2 - 4P + Q = C_{abcd}C^{abcd} - 2P + \frac{2}{3}R^2$$

does not contribute to the equations of motion.

[22] The cosmological constant can be thought as a contribution to the matter Lagrangian satisfying the null energy condition

where we recall that t_{ab} is the induced metric on the spacetime orthogonal to $\mathcal{S}$ (see eq. (4.12)). Now we wish to evaluate $\widetilde{C}[g,\mathcal{S}]$, eq. (4.13). The first thing to do is to find $\widetilde{g}_{ab}$, the boost-invariant projection of g_{ab}. For these theories, $\widetilde{g}_{ab}$ will always be a metric of the form (4.10) with $N = 2$, since the Lagrangian does not include any derivative of the metric of degree greater than two. Maximally symmetric spaces can be always parametrized as in (2.1). Therefore, after we switch to Kruskal coordinates (4.6), we will always get a boost invariant Ansatz of the kind:

$$d\tilde{s}^2 = \Big((C_{UV})_0 + (C_{UV})_1\, UV\Big)\, dUdV + \Big((C_\Omega)_0 + (C_\Omega)_1\, UV\Big)\, d\Omega^2\,, \tag{5.9}$$

where the coefficients $(C_{ab})_n$ will be determined below. Now, as stated in section 4, any metric of the form (4.10), as the metric above, will always satisfy the identity (4.11). Additionally, if we restrict ourselves to the ghost-free theories (5.4), then we will have $L_P = -4L_Q$. Putting these two conditions together, we can see from (5.8) that:

$$S_{\rm BH}[\widetilde{g},\mathcal{S}] = 4\pi \int_{\mathcal{S}} \Big((\widetilde{L_R}) - 2(\widetilde{L_T})\widetilde{R}\Big)\sqrt{h}\, d\Omega + 8\pi \int_{\mathcal{S}} (\widetilde{L_T})\, {}^{(2)}R\, \sqrt{h}\, d\Omega\,, \tag{5.10}$$

where tilded symbols denote the corresponding quantities evaluated in $\widetilde{g}_{ab}$. Now, notice that the final expression for whatever choice of $\widetilde{g}$ depends on just two curvature invariants: $\widetilde{R}$ and $\widetilde{T}$. Following section 4, we now choose the values of the coefficients in (5.9) so that:

- The $n = 0$ coefficients have to be chosen so that $\widetilde{g}_{ab} = g_{ab}$ at $\mathcal{S}$.
- The $n = 1$ coefficients have to be chosen so that $\widetilde{R} = R$ and $\widetilde{T} = T$ at $\mathcal{S}$.

As already emphasized, terms of order $\sim (UV)^2$ or higher will not change the curvature at $U = V = 0$. This means that we have the same number of coefficients to adjust in (5.9) than the number of equations derived from the conditions above. Hence we see that, in general, a boost-invariant projection such that $\widetilde{R} = R$ and $\widetilde{T} = T$ will always exist. Therefore we get a $\widetilde{C}$-function given by:

$$\widetilde{C}[g,\mathcal{S}] = 4\pi \int_{\mathcal{S}} (L_R - 2L_T\, R)\sqrt{h}\, d\Omega + 8\pi \int_{\mathcal{S}} L_T\, {}^{(2)}R\, \sqrt{h}\, d\Omega\,. \tag{5.11}$$

Integrating the above expression on a sphere of radius r_0 (recall that for metrics of the form (2.1) the surface $\mathcal{S}$ is a sphere), we get:

$$\widetilde{C} = 16\pi^2 b^2\, (L_R - 2L_T R)\, r_0^2 + 64\pi^2\, L_T\,, \tag{5.12}$$

where the expressions in the r.h.s. only depend on R and T, the curvature scalars of the maximally symmetric spacetime g_{ab}. Notice that, in general, eq. (5.11) reduces to the $\widetilde{C}$-function of section 3 for theories whose Lagrangian depends only on the Ricci scalar.

5.2.1 Properties of $\widetilde{C}$

By construction, the expression (5.12) has to be well defined. In particular, it has to be insensitive to the addition of a topological term to the Lagrangian.[23] It is straightforward

[23] As mentioned already in footnote 11, in four dimensions we have two possible topological terms built up of local integrals of the curvature: the first Euler class (i.e. the Gauss-Bonnet density) and the first Pontrjagin class. The latter, however, does not belong to the $L_g(R, P, Q)$ theories considered here.

to check that the addition of a Gauss-Bonnet term ($GB = R^2 - 4P + Q$) to the Lagrangian just shifts $\widetilde{C}$ by a constant (proportional to the Euler number of $\mathcal{S}$).[24]

Second, if we were in a black hole spacetime, $\widetilde{C}$ should equal the entropy of the black hole when evaluated at the horizon. However, the result above is for maximally symmetric spacetimes. Below we will consider explicitly the case of de Sitter space in the static patch. In this case, we will verify that, when $\widetilde{C}$ is evaluated at the cosmological horizon it exactly equals its entropy as when computed from eq. (1.1). Moreover, notice also that $\widetilde{C}$ reduces to the entropy formula for Lovelock gravity obtained in [2, 11].

Finally, note that all the curvature scalars are constant in a maximally symmetric space, and hence this function is a non-decreasing function of r_0 if and only if the theory is ghost-free (see eq. (5.7)). Note that this is a natural generalization of the "positivity of energy" condition (3.9) for the $L_g(R)$-theories (which, in turn, was also a reasonable generalization of the null energy condition — see the comment below eq. (3.10) and footnote 8). In fact, condition (3.9) can be interpreted as a no-ghost condition for the graviton, since it means that the graviton kinetic term has to carry the right sign. Here such kind of condition simply extends to the extra degrees of freedom appearing in the theory.

We could have expected to obtain another condition: the surviving scalar in the spectrum being non-tachyonic. The fact that this condition is not needed to have a well-behaved $\widetilde{C}$-function should not come as a surprise. A tachyon in the spectrum means an unstable background, examples of which exist already in GR; nevertheless in GR a $\widetilde{C}$-function exists and the Second Law holds irrespective of the stability of the background.

To end our discussion, let us next consider explicitly the cases of de Sitter and Anti de Sitter space.

5.2.2 de Sitter space

This case is interesting since we have a cosmological horizon. From a technical point of view, a cosmological horizon is the same as an event horizon, so we will be able to check explicitly that our $\widetilde{C}$-function equals its real entropy when evaluated at the cosmological horizon.

The metric of dS-space in the static patch can be written as:

$$ds^2 = b^2\Big(-(1-r^2)\,dt^2 + (1-r^2)^{-1}\,dr^2 + r^2\,d\Omega^2\Big)\,, \tag{5.13}$$

with $0 \le r \le 1$, $-\infty \le t \le \infty$; the cosmological horizon is at $r = 1$. First we have to go to Kruskal coordinates. These are given by:

$$U = e^t\sqrt{\frac{1-r}{1+r}}, \qquad\qquad V = -e^{-t}\sqrt{\frac{1-r}{1+r}}. \tag{5.14}$$

These coordinates only cover the patch $\{U \ge 0\} \cup \{V \le 0\}$ of dS space. We can trivially extend the coordinates to the whole range $-\infty \le U, V \le \infty$ in such a way that r, t are

[24] Hence we see that $\widetilde{C}$ is not exactly invariant against the addition of a Gauss-Bonnet density; however (as in the entropy of a physical system) such a constant is clearly immaterial.

defined in each one of the patches as:

$$r = \frac{1+UV}{1-UV}, \qquad\qquad t = \frac{1}{2}\log\left|\frac{U}{V}\right|. \tag{5.15}$$

The metric is given, for all U, V, by:

$$ds^2 = b^2\left(-\frac{4}{(1-UV)^2}\,dU dV + \left(\frac{1+UV}{1-UV}\right)^2 d\Omega^2\right). \tag{5.16}$$

These coordinates cover all of de Sitter space. The horizon lies at $UV = 0$. Notice also that this metric is a function of the single combination UV; therefore, at the bifurcate horizon $U = V = 0$ (and only there), the metric is automatically boost-invariant. In fact, the vector:

$$\xi = U\partial_U - V\partial_V \tag{5.17}$$

is a Killing vector of this metric, and vanishes at $U = V = 0$.

Next we shift the coordinates in order to have an arbitrary sphere of radius r_0 at $U = V = 0$. We make the shift $U \to (U + U_0)$, $V \to (V + V_0)$, so that $U = V = 0$ becomes a sphere of radius:[25]

$$r_0^2 = \left(\frac{1+U_0V_0}{1-U_0V_0}\right)^2. \tag{5.18}$$

The metric in terms of the shifted coordinates is given by

$$ds^2 = b^2\left(-\frac{4}{\left(1-(U+U_0)(V+V_0)\right)^2}\,dU dV + \left(\frac{1+(U+U_0)(V+V_0)}{1-(U+U_0)(V+V_0)}\right)^2 d\Omega^2\right). \tag{5.19}$$

Notice that this metric is not of the boost-invariant form eq. (4.4). In fact, ξ in (5.17) is no longer a Killing vector. However, it will be a Killing vector of its boost invariant projection at the surface $U = V = 0$ (i.e. $r = r_0$ in some patch), and moreover this surface will become the bifurcation surface of a Killing horizon in the $\widetilde{g}_{ab}$-space.

Let us check explicitly that such a boost invariant projection exists for de Sitter space. The Ansatz for $\widetilde{g}_{ab}$ is given in eq. (5.9). Next have to impose the conditions stated in section 4 to get the correct values for the coefficients $(C_{ab})_n$. In order to match with the actual dS metric (5.19) at $(U, V) = (0, 0)$ we must impose

$$(C_{UV})_0 = -(1 + r_0)^2, \qquad\qquad (C_\Omega)_0 = r_0^2. \tag{5.20}$$

The other two coefficients are adjusted so that $\widetilde{R} = R_{\rm dS} = 12/b^2$ and $\widetilde{T} = T_{\rm dS} = -120/b^4$. We get two possible solutions for each coefficient given by:

$$(C_{UV})_1 = \frac{1}{r_0^2}(1 + r_0)^4\left(\frac{5}{2} - 3r_0^2 \pm \sqrt{6}|1 - r_0^2|\right) \tag{5.21}$$

$$(C_\Omega)_1 = (1 + r_0)^2\left(1 \pm \frac{\sqrt{6}}{2}|1 - r_0^2|\right) \tag{5.22}$$

[25] Note that if the point (U_0, V_0) is located in the initially considered patch, so that $U_0 > 0$, $V_0 < 0$, then r_0 is always smaller than 1.

Finally, using (5.11) and integrating on the sphere of constant radius r_0, we get a $\widetilde{C}$-function given by eq. (5.12), where the r.h.s. is to be evaluated in de Sitter space. It can be explicitly checked that $\widetilde{g}_{ab}$ at $r_0 = 1$ exactly matches the series expansion of the de Sitter metric (5.16) up to second order. Therefore, if $\widetilde{C}$ is evaluated at the cosmological horizon $r_0 = 1$, it will equal its entropy as computed from eq. (1.1).

5.2.3 Anti de Sitter space

The metric of AdS can be written as:

$$ds^2 = b^2\Big(-(1+r^2)\,dt^2 + (1+r^2)^{-1}\,dr^2 + r^2\,d\Omega^2\Big)\,, \tag{5.23}$$

with $0 \le r \le \infty$ and $-\infty \le t \le \infty$. The relevant curvature scalars curvature are given by $R = -12/b^2$ and $T = 120/b^4$. These coordinates already cover the whole of AdS space. Next we define Txuskal coordinates by:

$$\begin{aligned} u &= t + \arctan r\,, & U &= e^u\,, \\ v &= t - \arctan r\,. & V &= -e^{-v}\,. \end{aligned} \tag{5.24}$$

The metric becomes:

$$ds^2 = b^2\left[\frac{1}{UV}\sec^2\left(\frac{1}{2}\log(-UV)\right)dU dV + \tan^2\left(\frac{1}{2}\log(-UV)\right)d\Omega^2\right]. \tag{5.25}$$

This metric is of the boost invariant form eq. (4.4), so it apparently has a bifurcate horizon at $(U,V) = (0,0)$, that does not exist in AdS space. In fact, this "horizon" is just an artifact of the change of coordinates, and the region $UV = 0$ is unphysical and it is not covered by the metric above. Notice that the relation between the original coordinate r and (U,V) is:

$$UV = -\exp(2\arctan r) \tag{5.26}$$

and hence $UV \in (-e^{\pi}, -1]$. It could be argued that other branches of the arctan should be used. Taking a different branch of the arctan changes the relation (5.26) by

$$UV = -\exp(2\arctan r)\exp(2\pi n) \tag{5.27}$$

for some integer n. Now, we can see that this integer must be minus infinity to reach $UV = 0$, therefore we see that the "horizon" generated at $UV = 0$ is definitely an artifact of the change of coordinates and is outside AdS space.

Next we proceed analogously to the dS case. We shift the coordinates as $(U,V) \to (U+U_0, V+V_0)$ (taking into account that $U_0V_0 \in (-e^{\pi}, -1]$, in order to remain in AdS space[26]). The surface $U = V = 0$ becomes a sphere of radius:

$$r_0^2 = \tan^2\left(\frac{1}{2}\log(-U_0V_0)\right)\,. \tag{5.28}$$

[26] Or $e^{\pi n}U_0V_0 \in (-e^{\pi}, -1]$ for some $n \in \mathbf{Z}$ for a different branch of the arctan, considering also branches with negative r.

Next we compute the boost-invariant projection $\widetilde{g}_{ab}$ at $U = V = 0$. Again, we impose that this metric coincides with that of AdS at $U = V = 0$, and that $\widetilde{R} = R_{\text{AdS}}$ and $\widetilde{T} = T_{\text{AdS}}$ at $U = V = 0$. It can be explicitly checked that this metric is given by (5.9) with:

$$\begin{aligned}
(C_{UV})_0 &= -e^{-2\arctan r_0}(1+r_0^2), \\
(C_{UV})_1 &= \frac{e^{-4\arctan r_0}(1+r_0^2)^2}{r_0^2}\left(\frac{5}{2} + 3r_0^2 \pm \sqrt{6}\sqrt{1+2r_0^2+11r_0^4}\right) \\
(C_\Omega)_0 &= r_0^2, \\
(C_\Omega)_1 &= e^{-2\arctan r_0}(1+r_0^2)\left(1 \pm \sqrt{6}\sqrt{1+2r_0^2+11r_0^4}\right)
\end{aligned} \tag{5.29}$$

Finally, the $\widetilde{C}$-function is given by eq. (5.12), where the r.h.s. is now to be evaluated in AdS.

As already mentioned in the Introduction, let us stress here that we are by no means claiming that the above function is the holographic dual of a true field theory c-function (since, for AdS, such a function should to be a constant). In order to address Questions **1** and **2** in the Introduction, the only property that we had to demand from our "c-function" $\widetilde{C}$ was that it has to be non-decreasing along outgoing null geodesic flow. This is exactly what we got.

However, let us mention that, even without using any considerations on the dual CFT, one would maybe expect that any geometric function in AdS should be independent of the radial coordinate, due to the well-known "scale invariance of Anti de Sitter space". Let us recall here that the latter is a property of the Poincaré patch of AdS, but not a property of global AdS (whose metric is the one that we have used here[27]). It is easy to see that had we constructed the $\widetilde{C}$-function from the metric in the Poincaré patch

$$ds^2 = b^2\left(z^2(-d\tau^2 + dx^2 + dy^2) + \frac{dz^2}{z^2}\right) \tag{5.30}$$

with $-\infty < \tau, x, y < \infty$, $0 < z < \infty$, we would have found $\widetilde{C} \sim z^2\, dx \wedge dy$, that is invariant under the symmetry $(\tau, x, y, z) \to (a\tau, ax, ay, a^{-1}z)$ (regardless of the fact of $\widetilde{C}$ being the holographic dual of a field theory c-function or not). Note however that these are not good coordinates to address issues like the entropy contained in a given region of spacetime, since there are no closed, compact spacelike 2-surfaces for constant τ and z in these coordinates.

6. Conclusions

In this paper we have explored to which extent the possibility of establishing a well-defined holographic bound, as well as the Second Law of black hole mechanics, extend to general theories of gravity with higher curvature interactions. As we have shown, these two issues seem to imply each other via the existence of a "c-function", i.e. a never decreasing function along outgoing null geodesic flow.

[27]We wish to thank Ofer Aharony for pointing this out to us.

It would be nice to explore the implications that the general holographic bound that we propose here may have along the lines of the covariant entropy bound of [13]. In order to have well defined holographic screens, it seems crucial to prove that a suitable generalization of the focussing theorem holds for $\widetilde{\vartheta} \equiv d \log \widetilde{C}/d\lambda$. We did not attempt to do this in this paper, but it seems that such a generalization should hold. After all, the focussing theorem plays a crucial role in GR in the proof of the Second Law and in the proof of all singularity theorems. In a sense, a convenient generalization of Raychaudhuri equation and the focussing theorem could be taken as a good starting point to define what a good theory of gravity should be (including its recognition as a holographic theory).

From a more technical point of view, it would be interesting to understand the significance of the boost invariant projection $\widetilde{g}_{ab}$ defined in section 4.1.2. This projection is central in our construction in order to arrive to the final expression for the $\widetilde{C}$-function. However, we do not understand the real significance (if any) of this spacetime. After all, such a spacetime looks in the end like quite a spurious object (having, in addition, a lot of "gauge freedom") which might well be not much more than a mere artifact in order to obtain expressions like (4.11), which seem to play a crucial form in the final form of $\widetilde{C}$. Also, the fact that we have to go to a special coordinate system to find $\widetilde{g}_{ab}$ is somewhat unsatisfactory. It would be very nice to understand the covariant meaning of the boost invariant projection that we define.

Finally, let us mention here that an interesting corollary of our definition of $\widetilde{C}$ is that it constitutes, by itself, a possible candidate to define the entropy of a dynamical black hole, since it satisfies by construction all the properties required in [2]. However, it will in general differ from the proposal of Iyer and Wald (see the appendix). It would be nice to check the implications of this.

Our results are far from general. However, we believe that the ideas presented and developed in this paper should be of quite general applicability. It would be very interesting to understand them in depth and to check our proposal in more general theories and more general backgrounds.

Acknowledgments

We would like to thank Barak Kol, Fernando Quevedo and Aninda Sinha for comments on the manuscript, Sean Hartnoll and Alberto Ramos for useful conversations, Ignacio Navarro for conversations and for helping us much in understanding the consistency of gravity theories with higher curvature couplings, and finally Cdo. Jorge Conde for useful comments on the manuscript and for very useful suggestions on the notation of the paper. We also want to thank the "Bar-Cafetería Mozar" in Madrid for hospitality during the early stages of this project. D.C. also thanks the Weizmann Institute of Science for hospitality during the completion of this work.

The work of D.C. is supported by the University of Cambridge. The work of E. L.-T. has been supported by a Marie Curie Fellowship under contract MEIF-CT-2003-502349, the Spanish Grant BFM2003-01090, The Israel-US Binational Science Foundation, the ISF Centers of Excellence Program and Minerva.

A. Iyer and Wald proposal for the entropy of a dynamical black hole

As mentioned in the body of the paper, our definition of the $\widetilde{C}$-function is based on the proposal of Iyer and Wald [2] to define the entropy of a non-stationary black hole. Such a problem faces the same kind of difficulties that we had to face in order to extend the entropy formula eq. (1.1) to general surfaces which need not be a cross-section of a Killing horizon (since the event horizon of a non-stationary black will not be, in general, a Killing horizon). Also, and as in our definition of $\widetilde{C}$, one would require of any generalization of the entropy formula to reduce to eq. (1.1) when evaluated on the horizon of a stationary black hole. For the interested reader, we review in this appendix the solution to this problem found by Iyer and Wald.

The basic idea is as follows: instead of modifying the functional form of the entropy functional $S_{\rm BH}[g,\Sigma]$ given in eq. (1.1), Iyer and Wald provided a specific algorithm to deform the dynamical metric g_{ab} in a neighbourhood of a spacelike cross-section of the event horizon of a dynamical black hole. Let us denote such a cross-section by $\Sigma_{\rm dyc}$. This deformed spacetime, that here we call $\widehat{g}_{ab}$, is such that $g_{ab}=\widehat{g}_{ab}$ at $\Sigma_{\rm dyc}$. Moreover, it has the property that $\Sigma_{\rm dyc}$ becomes the bifurcation surface of a bifurcate Killing horizon of the metric $\widehat{g}_{ab}$. This will make the quantity $S_{\rm BH}[\widehat{g},\Sigma_{\rm dyc}]$ to be automatically well defined (cf. the discussion at the beginning of section 4.1). Finally, it turns out that $\widehat{g}_{ab}=g_{ab}$ when the "original" spacetime is that of a stationary black hole, and therefore $S_{\rm BH}[\widehat{g},\Sigma_{\rm dyc}] = S_{\rm BH}[g,\Sigma]$. Let us summarize now the prescription of Iyer and Wald to obtain $\widehat{g}_{ab}$ from g_{ab}.

A.1 Boost-invariant part of g_{ab} at $\Sigma_{\rm dyc}$

Their algorithm to find $\widehat{g}_{ab}$ is as follows. First choose, at any point p in $\Sigma_{\rm dyc}$, a couple of independent null vectors (unique up to scale) k^a and n^a orthogonal to $\Sigma_{\rm dyc}$, obeying the (conventional) normalization condition $k^a n_a = -1$. Take now a neighbourhood of $\Sigma_{\rm dyc}$ small enough such that any point x in the vicinity of $\Sigma_{\rm dyc}$ lies on a unique geodesic orthogonal to $\Sigma_{\rm dyc}$. Next define the following coordinate system in the space orthogonal to $\Sigma_{\rm dyc}$:[28] given x, find p in $\Sigma_{\rm dyc}$ and the (unique) geodesic connecting p and x. Parametrize this geodesic such that x is at unit affine parameter from p, and find its tangent vector v^a of such geodesic at p. Finally, assign the coordinates (U,V) to x in the 2-space spanned by k^a and n^a to be the components of v^a along k^a and n^a. Note that, in this coordinate system, $\Sigma_{\rm dyc}$ lies at $U=V=0$.

The next step is to Taylor-expand every component of g_{ab} in U and V around $U = V = 0$ up to some order N (to be fixed below). In this series expansion, remove all terms which do not contain the same number of U's and V's. The resulting metric, $\widehat{g}^N_{ab}$, is called the *boost-invariant part of order* N of g_{ab} at $\Sigma_{\rm dyc}$. Finally, choose the order N of $\widehat{g}^N_{ab}$ to be equal to the higher derivative of the metric appearing in the Lagrangian. Note that the boost-invariant part $\widehat{g}^N_{ab}$ is a function of the single combination UV, and therefore

[28] This coordinate system was used in [10], and their relation to the properties of spacetimes with bifurcate Killing horizons was studied there.

is of the boost-invariant form (4.4).[29] Therefore the surface $U = V = 0$ becomes the bifurcation surface of the Killing horizon $UV = 0$, the corresponding Killing vector being $\xi = U\partial_U - V\partial_V$. All this implies that the quantity:

$$S_{\rm dyc}[g, \Sigma_{\rm dyc}] \equiv S_{\rm BH}[\widehat{g}, \Sigma_{\rm dyc}] = -2\pi \int_{\Sigma_{\rm dyc}} \frac{\partial L(\widehat{g})}{\partial \widehat{R}_{abcd}} \,\widehat{\epsilon}_{ab}\, \widehat{\epsilon}_{cd} \sqrt{h}\, d\Omega\,. \tag{A.1}$$

is automatically well defined and free of all the ambiguities present in $S_{\rm BH}[g, \Sigma_{\rm dyc}]$ [5, 2]. Most importantly, in the case of a stationary black hole (and therefore $\Sigma_{\rm dyc} = \Sigma$), it turns out that:

$$S_{\rm dyc}[g, \Sigma] = S_{\rm BH}[g, \Sigma]\,, \tag{A.2}$$

since, for a stationary black hole, the metric g_{ab} at Σ is automatically boost invariant [10, 2], which implies that g_{ab} equals $\widehat{g}_{ab}$ at Σ. This fact motivated the proposal of [2] to take $S_{\rm BH}[\widehat{g}, \Sigma_{\rm dyc}]$ as a possible candidate for the physical entropy of a dynamical black hole.

Notice that the definition (A.1) for the entropy of a dynamical black hole is in principle evaluated on a cross-section of the event horizon. However, the fact that $\Sigma_{\rm dyc}$ is an event horizon is not required at any point. This is why we used a very similar prescription to extend the definition of the entropy functional (1.1) to arbitrary spacelike surfaces (regardless of them being event horizons or not).

A.2 Comparison between $\widetilde{C}$ and $S_{\rm dyc}$

Notice that the functional $S_{\rm dyc}[g, \mathcal{S}]$ has all the required properties that we discussed in section (4.1) in order for it to be well defined on arbitrary spacelike surfaces. So, in principle, it is a legitimate candidate for a $\widetilde{C}$-function satisfying the properties a)-d) established at the beginning of section 4. Let us therefore compare here our proposal for $\widetilde{C}[g, \mathcal{S}]$ and $S_{\rm dyc}[g, \mathcal{S}]$.

First, note that it is clear that:

$$\widetilde{C}[g, \Sigma] = S_{\rm dyc}[g, \Sigma] = S_{\rm BH}[g, \Sigma] \tag{A.3}$$

when evaluated on a cross-section Σ of the event horizon of a stationary black hole. However, in general:

$$\widetilde{C}[g, \mathcal{S}] \neq S_{\rm dyc}[g, \mathcal{S}] \tag{A.4}$$

on an arbitrary spacelike surface $\mathcal{S}$. This is because, since $\widetilde{g}_{ab}$ and $\widehat{g}_{ab}$ are different metrics, their associated curvature scalars at $\mathcal{S}$ will not coincide in general: only on a black hole horizon it is ensured that (at least to order N in an expansion of the kind of (4.10)) we will have $\widetilde{g}_{ab} = \widehat{g}_{ab} = g_{ab}$. The reason for defining the $\widetilde{C}$-function as we did (as opposed to using Iyer and Wald's boost-invariant part to deform the spacetime metric) is because, at least in the cases that we have been able to check, only if $\widetilde{C}$ is defined as $\widetilde{C}[g, \mathcal{S}] \equiv S_{\rm BH}[\widetilde{g}, \mathcal{S}]$, the conditions for it to be a non-decreasing function along outgoing null geodesic flow match the physical requirements of the theory.

[29] Actually, $\widehat{g}_{ab}$ is a power series of the form (4.10). The difference between $\widehat{g}_{ab}$ and the boost invariant projection $\widetilde{g}_{ab}$ defined in the paper is the prescription to fix the coefficients in (4.10).

JHEP01(2007)045

Let us explicitly verify this in an example. First note that, since $\widehat{g}_{ab}$ is also of the boost-invariant form (4.10), the identity (4.11) will also hold for the curvature invariants associated to $\widehat{g}_{ab}$. This means that a "hatted" analogous of (5.10) also holds for $S_{\rm dyc}[\widehat{g}, \mathcal{S}]$. Considering for definiteness the case of AdS space, using Iyer and Wald's prescription we had obtained a $\widetilde{C}$-function given by:

$$\widetilde{C} \equiv S_{\rm dyc} = 16\pi^2 b^2(\widehat{L_R} - 2\widehat{L_T}\widehat{R}) + 64\pi^2 \widehat{L_T} \tag{A.5}$$

where hats denote evaluation of all expressions in $\widehat{g}_{ab}$ at the surface $\mathcal{S}$. In this case, the final requirement of $\widetilde{C}$ being non-decreasing of r would not have been equivalent to the physical requirements (5.7). This is because the curvature scalars of the boost-invariant part of AdS, when evaluated at a generic surface $r = r_0$, do not coincide with those of AdS itself. In particular, one finds:

$$\begin{aligned} \widehat{R} &= \frac{9r_0^4 - 4r_0^3 + 10r_0^2 + 1}{6r_0^2(1+r_0)^2} R_{\rm AdS}\,, \\ \widehat{T} &= \frac{-12r_0^4 + 36r_0^3 - 45r_0^2 + 8r_0 - 9}{15r_0^2(1+r_0)^2} T_{\rm AdS}\,, \end{aligned} \tag{A.6}$$

Therefore, even if we see that $S_{\rm dyc}[g, \mathcal{S}]$ is well defined, the conditions for it to be non-decreasing do not coincide with the physical requirements (5.7). Notice also that, in general $S_{\rm dyc}[g, \mathcal{S}]$ has the additional unsatisfactory property that the conditions that will have to be satisfied for it to be non-decreasing will depend on the detailed form of the Lagrangian $L_g(R, T)$. However, the ghost-free condition (5.7) is generic for any Lagrangian of the form $L_g(R, P, Q) = L_g(R, T)$.

References

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

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

[3] T. Jacobson, G. Kang and R.C. Myers, *Increase of black hole entropy in higher curvature gravity*, *Phys. Rev.* **D 52** (1995) 3518 [gr-qc/9503020].

[4] G. 't Hooft, *Dimensional reduction in quantum gravity*, gr-qc/9310026.

[5] T. Jacobson, G. Kang and R.C. Myers, *On black hole entropy*, *Phys. Rev.* **D 49** (1994) 6587 [gr-qc/9312023].

[6] K. Goldstein, R.P. Jena, G. Mandal and S.P. Trivedi, *A C-function for non-supersymmetric attractors*, *JHEP* **02** (2006) 053 [hep-th/0512138].

[7] V. Sahakian, *Holography, a covariant C-function and the geometry of the renormalization group*, *Phys. Rev.* **D 62** (2000) 126011 [hep-th/9910099];
E. Álvarez and C. Gómez, *Geometric holography, the renormalization group and the C-theorem*, *Nucl. Phys.* **B 541** (1999) 441 [hep-th/9807226].

[8] I. Navarro and K. Van Acoleyen, *Consistent long distance modification of gravity from inverse powers of the curvature*, *JCAP* **03** (2006) 008 [gr-qc/0511045].

[9] K.S. Stelle, *Renormalization of higher derivative quantum gravity*, *Phys. Rev.* **D 16** (1977) 953.

[10] B.S. Kay and R.M. Wald, *Theorems on the uniqueness and thermal properties of stationary, nonsingular, quasifree states on space-times with a bifurcate Killing horizon*, *Phys. Rept.* **207** (1991) 49.

[11] T. Jacobson and R.C. Myers, *Black hole entropy and higher curvature interactions*, *Phys. Rev. Lett.* **70** (1993) 3684 [hep-th/9305016].

[12] S.W. Hawking, *Black holes in general relativity*, *Commun. Math. Phys.* **25** (1972) 152;
S.W. Hawking and G.R.F. Ellis, *The large scale structure of spacetime*, Cambridge University Press, Cambridge, 1973.

[13] R. Bousso, *Holography in general space-times*, *JHEP* **06** (1999) 028 [hep-th/9906022]; *The holographic principle for general backgrounds*, *Class. and Quant. Grav.* **17** (2000) 997 [hep-th/9911002].

[14] A. Paranjape, S. Sarkar and T. Padmanabhan, *Thermodynamic route to field equations in Lancos-Lovelock gravity*, *Phys. Rev.* **D 74** (2006) 104015 [hep-th/0607240];
A. Mukhopadhyay and T. Padmanabhan, *Holography of gravitational action functionals*, hep-th/0608120;