PHYSICAL REVIEW D VOLUME 52, NUMBER 6 15 SEPTEMBER 1995

Increase of black hole entropy in higher curvature gravity

Ted Jacobson* and Gungwon Kang†
Department of Physics, University of Maryland, College Park, Maryland 20742-4111

Robert C. Myers‡
Department of Physics, McGill University, Montréal, Québec, Canada H3A 2T8

(Received 13 March 1995)

We examine the zeroth law and the second law of black hole thermodynamics within the context of effective gravitational actions including higher curvature interactions. We show that entropy can never decrease for quasistationary processes in which a black hole accretes positive energy matter, independent of the details of the gravitational action. Within a class of higher curvature theories where the Lagrangian consists of a polynomial in the Ricci scalar, we use a conformally equivalent theory to establish that stationary black hole solutions with a Killing horizon satisfy the zeroth law, and that the second law holds in general for any dynamical process. We also introduce a new method for establishing the second law based on a generalization of the area theorem, which may prove useful for a wider class of Lagrangians. Finally, we show how one can infer the form of the black hole entropy, at least for the Ricci polynomial theories, by integrating the changes of mass and angular momentum in a quasistationary accretion process.

PACS number(s): 04.70.Dy, 04.50.+h, 04.70.Bw

I. INTRODUCTION

One of the primary motivations to study black hole thermodynamics is to gain some insight into the nature of quantum gravity. Whatever framework physicists eventually uncover to describe quantum gravity, there should be a low energy effective action which describes the dynamics of a "background metric field" for sufficiently weak curvatures at sufficiently long distances. On general grounds, one expects that this effective gravity action will consist of the classical Einstein action plus a series of covariant, higher-dimension interactions (i.e., higher curvature terms, and also higher derivative terms involving all of the physical fields) induced by quantum effects. While such effective actions are typically pathological when considered as fundamental, they may also be used to define mild perturbations for Einstein gravity coupled to conventional matter fields. It is within this latter context of Einstein gravity "corrected" by higher dimension operators that we wish to consider modifications of black hole thermodynamics.

Naive dimensional analysis suggests that the coefficients of all higher dimension interactions in the effective Lagrangian should be dimensionless numbers of order unity times the appropriate power of the Planck length. Thus one might worry that the effect of the higher dimension terms would be the same order as those of quantum fluctuations, and so there would seem to be little point in studying modifications to *classical* black hole thermodynamics from higher dimension terms. One motivation for studying the classical problem is that it is of course possible that this naive dimensional analysis is incorrect, just as it would be in predicting the observed value of the cosmological constant. So given the lack of any direct experimental evidence, it is possible that the coefficients of some higher dimension terms are larger than expected. Further, we would like to know whether or not consistency with classical black hole thermodynamics places any new restrictions on these coefficients. Moreover, it is interesting to explore black hole thermodynamics in generalized gravity theories in order to see whether the thermodynamic "analogy" is just a peculiar accident of Einstein gravity, is a robust feature of all generally covariant theories of gravity, or is something in between.

In this "analogy," any black hole should behave as a heat bath. Quantum field theory reveals that $\kappa/(2\pi)$ is the black hole temperature [1], independent of the details of the dynamics of the gravity theory [2]. Hence this is in fact a robust feature of black hole physics in general. An important foundation for black hole thermodynamics is then the validity of the zeroth law, namely, that the surface gravity should be constant across a stationary event horizon. If the event horizon is a Killing horizon with a regular bifurcation surface, it is straightforward to show that the zeroth law holds [5]. In Einstein gravity a proof of the zeroth law can also be constructed without the assumption of a regular bifurcation surface, but with the additional assumption that the dominant energy condition [3] holds, i.e., for any future pointing timelike vector u^a the corresponding four momentum density $P^a = -T^{ab}u_b$ is future pointing timelike. Equivalently, for all pairs of future pointing timelike vectors u^a, v^a,

$$T_{ab}\, u^a v^b \geq 0. \tag{1}$$

*Electronic address: jacobson@umdhep.umd.edu
†Electronic address: eunjoo@wam.umd.edu
‡Electronic address: rcm@hep.physics.mcgill.ca

In higher curvature theories, establishing the validity of the zeroth law in general remains an important open question.

Using Euclidean path integral methods [6] it is clear that, if the zeroth law holds, a version of the first law of black hole thermodynamics still applies in these higher curvature theories. Applying these techniques to various specific theories and specific black hole solutions showed that the relation equating the black hole entropy with one-quarter the surface area of the horizon no longer holds in general [7]. However, it is now known that the entropy is given always by a local expression evaluated at the horizon [8–11].

Wald [9] established this result very generally for any diffeomorphism invariant theory via a new Minkowski signature derivation of the first law of black hole mechanics. This law applies for variations (of the dynamical fields) around a *stationary* black hole background to nearby solutions:

$$\frac{\kappa}{2\pi}\delta S = \delta M - \Omega^{(\alpha)}\,\delta J_{(\alpha)} \quad . \tag{2}$$

Here M, $J_{(\alpha)}$, $\Omega^{(\alpha)}$, and κ are the mass, the angular momentum [12], the angular velocity, and the surface gravity of the black hole. Wald found that S can be expressed as a *local* geometric density integrated over a spacelike cross section of the horizon, and that it is associated with the Noether charge of diffeomorphisms under the Killing vector field that generates the horizon. Equation (2) then has the rather remarkable feature that it relates variations in properties of the black hole as measured at asymptotic infinity to a variation of a geometric property of the horizon. Given the identification of the temperature with $\kappa/(2\pi)$, Eq. (2) has a natural interpretation as the first law of thermodynamics where S is the black hole entropy. If this S is truly to play the role of entropy, it should also satisfy the second law of thermodynamics as a black hole evolves—i.e., S should never decrease in any dynamical processes.

For general relativity, one has the celebrated result that the black hole entropy is given by one-quarter the surface area of the horizon, $S = A/(4G)$ [13]. In this context, the second law is established by Hawking's area theorem, which states that in any classical process involving black holes, the total surface area of the event horizon will never decrease [14]. An essential ingredient in the proof of this theorem is the assumption that the null convergence condition $R_{ab}k^a k^b \geq 0$ holds for all null vectors k^a. This condition is implied by the Einstein field equation

$$R_{ab} - \frac{1}{2}g_{ab}R = 8\pi G\, T_{ab} \tag{3}$$

together with the null energy condition

$$T_{ab}\, k^a k^b \geq 0 \qquad \text{for any null vector } k^a \quad . \tag{4}$$

Another essential ingredient is cosmic censorship—i.e., it is assumed that naked singularities do not develop in the processes of interest.

In theories where higher curvature interactions are included along with the Einstein Lagrangian, the equations of motion may still be written in the form of Eq. (3), if the contributions from the higher curvature interactions are included in the stress-energy tensor. Typically, these contributions spoil the energy conditions, and so one cannot establish an area increase theorem in such theories. However, this is not the relevant question for black hole thermodynamics. The relevant question is whether or not the quantity S whose variation appears in the first law (2) satisfies a classical increase theorem. If so, one would have a second law of black hole thermodynamics for these theories, further validating the interpretation of S as the black hole entropy.

In this paper, we investigate the validity of the zeroth law and the classical second law for higher curvature extensions of Einstein gravity. Section II examines entropy increase in quasistationary processes. For such processes, the second law arises directly from the first law for any theory of gravity, as long as the matter stress-energy tensor satisfies the null energy condition. Section III demonstrates that the zeroth law for stationary black holes and the second law for arbitrary dynamical processes hold in a theory where the gravitational Lagrangian is $R + \alpha R^2$. These results are established by relating the higher curvature theory by a conformal field redefinition to a more conventional theory in which Einstein gravity is coupled to an auxiliary scalar field. Our proofs are valid provided that α is positive and that cosmic censorship holds for the conformally related metric. Section IV generalizes the previous results to a larger class of theories where the gravitational action is a polynomial of the Ricci scalar. In Section V, a proof of the second law is constructed directly within the higher curvature theories without making (explicit) use of the conformal field redefinition. This approach follows closely the logic of Hawking's area theorem, and may provide insight for the problem of establishing the second law in more general theories. This analysis makes use of an "extended" Raychaudhuri equation which is also used to establish a new "physical process" derivation of the first law and of the form of the entropy for these higher curvature theories analogous to that of Wald for general relativity [15]. Section VI presents a discussion of our results.

Throughout the paper, we consider only asymptotically flat spaces [16], and we employ the conventions of [4]. We also adopt the standard convention of setting $\hbar = c = 1$. Further, motivated by the fact that many of the recent candidates for a theory of quantum gravity are theories in higher dimensional spacetimes, we will allow spacetime to have an arbitrary dimension, $D \geq 4$.

II. QUASISTATIONARY PROCESSES

Here we demonstrate that in quasistationary processes, the second law is a consequence of the first law for any theory of gravity. We wish to consider a dynamical process in which a small amount of matter enters from a great distance and drops into a vacuum black hole. The initial and final black hole states are (approximately) stationary. In those spacetime regions then, there are Killing vector fields, ξ^a and $\phi^a_{(\alpha)}$, which asymptotically generate time translations and orthogonal rotations [12], respectively. By a *quasistationary* process we mean one in which the background spacetime is only slightly perturbed by the infalling matter. In order to establish a perturbation expansion we introduce a small parameter δ associated with the amplitude of the infalling fields. We assume that the stress energy tensor is order δ^2, as are the resulting perturbations of the metric. The vector fields, ξ^a and $\phi^a_{(\alpha)}$, can then be extended through the intermediate spacetime region where the accretion process occurs such that Killing's equation still applies to leading order [e.g., $\nabla^{(a}\xi^{b)} = O(\delta^2)$]. This extension is further chosen so that the horizon generator is given by $\chi^a = \xi^a + \Omega^{(\alpha)}\phi^a_{(\alpha)}$ throughout the entire evolution, where $\Omega^{(\alpha)}$ is constant to order δ^2.

Two spacelike surfaces Σ_i and Σ_f are introduced at a time before the accretion process begins, and at a time after the process has ended, respectively. The black hole horizon H provides an inner surface, and we introduce a surface O at some large radius in the asymptotically flat region. These four surfaces enclose a spacetime volume V, and we assume that the matter enters from the asymptotically flat region and exits either as it crosses the black hole horizon or after scattering back out across O. One can evaluate the mass and angular momentum carried in with the new matter by flux integrals in the asymptotic region,

$$\Delta M = \int_O T_{ab}\,\xi^a d\Sigma^b \ ,$$
$$\Delta J_{(\alpha)} = -\int_O T_{ab}\,\phi^a_{(\alpha)} d\Sigma^b, \tag{5}$$

where $d\Sigma^b$ is defined with an outward pointing normal. In the framework of the perturbation expansion, these quantities are $O(\delta^2)$. In what follows higher order terms will be ignored. These expressions may be combined to produce a flux integral containing the horizon generating vector field χ^a:

$$\Delta M - \Omega^{(\alpha)}\Delta J_{(\alpha)} = \int_O T_{ab}\,\chi^a d\Sigma^b \ . \tag{6}$$

By the first law (2), this combination of variations is proportional to the variation of the entropy, $\frac{\kappa}{2\pi}\Delta S$. (Here, we apply the first law in what Wald calls the "physical process" form [15], for which one assumes that the accretion process is a mild perturbation, and hence that no singularities develop in V.)

Using stress-energy conservation as well as Killing's equation, one has

$$\nabla^a(T_{ab}\,\chi^b) = \nabla^a T_{ab}\,\chi^b + T_{ab}\,\nabla^a\chi^b = O(\delta^4).$$

Now one can integrate the above expression over V and apply Gauss' law to produce a flux integral. Since the stress-energy vanishes on Σ_i and Σ_f, one finds

$$O(\delta^4) = \int_V \nabla^a(T_{ab}\,\chi^b)\,d\Sigma$$
$$= \int_O T_{ab}\,\chi^a d\Sigma^b + \int_H T_{ab}\,\chi^a d\Sigma^b \ . \tag{7}$$

Here $d\Sigma^b$ is defined with an outward pointing normal which, on the horizon, is antiparallel to the null generator χ^b. Hence the integrand of the second flux integral is everywhere negative or vanishing, as long as the stress tensor satisfies the null energy condition (4). Thus combining Eqs. (6) and (7) with the first law, one has

$$\Delta S = -\frac{2\pi}{\kappa}\int_H T_{ab}\chi^a d\Sigma^b \geq 0 \ . \tag{8}$$

The details of the gravity theory, and also the precise functional form of S, are irrelevant to establish this result for quasistationary processes. The key requirements were the first law relating asymptotic variations to variations of the horizon geometry, and the null energy condition to be satisfied by the matter stress-energy. We emphasize that this stress-energy tensor includes only contributions from the matter fields, and not contributions from any higher curvature interactions, as were considered in the introduction. We have also made an implicit assumption that the black hole solution is stable—i.e., the small perturbation introduced by matter falling in from infinity remains small (and does not lead to any unstable growing or oscillatory excitations) so that the system simply settles down to a new black hole.

We fully expect that the above line of reasoning can be generalized to cover arbitrary quasistationary processes. For example, the result should hold for the case in which (possibly charged) matter and electromagnetic radiation fall into an electrically and/or magnetically charged black hole. Indeed, we have carried out the analysis for the case of charged matter (but no radiation) falling onto a charged black hole. This case is somewhat more complicated than the vacuum case above. If one allows infalling electromagnetic radiation then, due to the nonlinearity of the electromagnetic field stress tensor, there are generically cross terms between the background field and the radiation field, yielding $O(\delta)$ contributions to the stress tensor and to the metric variation. This appears to require a much more involved, or a much more clever, analysis. Also, the above argument does not directly apply in the situation where a packet of gravity waves drops into the black hole although it can probably be adapted to cover positive energy gravitational perturbations.

III. $R + \alpha R^2$ THEORY

In this section we establish the constancy of the surface gravity for stationary solutions and the second law for arbitrary dynamical processes involving black holes in the theory given by a higher curvature action of the form

$$I_o = \int d^D x\sqrt{-g}\left[\frac{1}{16\pi G}(R + \alpha R^2) + L_m(\psi, g)\right] \ , \tag{9}$$

where L_m denotes a conventional Lagrangian for some collection of matter fields, denoted ψ. The matter Lagrangian will also contain the metric, but we assume that it contains no derivatives of the metric.

The gravitational field equations arising from the action (9) are

$$R_{ab}-\frac{1}{2}g_{ab}R=8\pi G\,T^{m}_{ab}(\psi,g)+2\alpha\left(\nabla_a\nabla_b R-g_{ab}\nabla^2 R-R\,R_{ab}+\frac{1}{4}g_{ab}\,R^2\right)\ . \tag{10}$$

We will assume that matter stress-energy tensor $T^{m}_{ab}=-\frac{2}{\sqrt{-g}}\frac{\delta\sqrt{-g}L_m}{\delta g^{ab}}$ does satisfy the dominant energy condition (1). However, if one regards the entire expression on the right-hand side of this equation as the stress-energy tensor in Einstein's equations (3), it is clear that this total stress-energy does not satisfy any energy condition because of the higher curvature contributions (i.e., the terms proportional to α). Thus, as discussed in the introduction, Hawking's proof of the area theorem does not apply here, nor can the proof of the zeroth law for stationary black holes in Einstein gravity be invoked.

For this theory, the black hole entropy appearing in the first law (2) can be written [10, 11, 17]

$$S=\frac{1}{4G}\int_{\mathcal{H}}d^{D-2}x\,\sqrt{h}\,(1+2\alpha\,R)\ , \tag{11}$$

where the integral is taken over a spacelike cross section of the horizon $\mathcal{H}$. The above form of the black hole entropy is not strictly justified unless we know that stationary black holes of the present theory possess a Killing horizon with constant surface gravity. We shall show below that, for $\alpha>0$, the surface gravity of a Killing horizon in this theory is necessarily constant. Note that a number of ambiguities still arise in Wald's construction, and so Eq. (11) is the result after making certain natural choices in the calculation. None of these ambiguities have any effect when the Noether charge is evaluated on a stationary horizon [17], and hence Eq. (11) may be considered in the analysis of Sec. II since there the entropy is only compared between the initial and final stationary black holes. So one knows that the quantity (11) will always increase in a quasistationary process in which a packet of the matter is dropped into a black hole [since by assumption, the matter stress-energy satisfies the dominant energy condition (1), which implies the null energy condition (4)].

We now extend this result to a classical entropy increase theorem for any dynamical process involving black holes in this theory. Our approach will be the following. First, we show that the present higher curvature theory is equivalent to Einstein gravity for a conformally related metric coupled to an auxiliary scalar field, as well as to the original matter fields. Second, we argue that the black hole entropy in the higher curvature theory is identical to that in the conformally related theory. Finally, since Hawking's area theorem holds in the Einstein-plus-scalar theory, we conclude that the entropy never decreases in the original theory (9).

The equivalence of the higher curvature theory (9) to Einstein gravity coupled to an auxiliary scalar field has been discussed previously by many authors [18]. The first step is to introduce a new scalar field ϕ and a new action, which is linear in R:

$$I_1=\int d^Dx\sqrt{-g}\left\{\frac{1}{16\pi G}\left[(1+2\alpha\phi)R-\alpha\phi^2\right]+L_m(\psi,g)\right\}\ . \tag{12}$$

The ϕ equation of motion is simply $\phi=R$, and one recovers the original action upon substituting this equation into Eq. (12)—i.e., $I_1(\phi=R)=I_0$. In the form of Eq. (12), the action contains no terms that are more than quadratic in derivatives. This action contains an unconventional interaction, ϕR, however. Hence in the metric equations of motion,

$$R_{ab}-\frac{1}{2}g_{ab}\,R=8\pi G\,T^{m}_{ab}(\psi,g)+2\alpha\left(\nabla_a\nabla_b\phi-g_{ab}\nabla^2\phi-\phi\,R_{ab}+\frac{1}{2}g_{ab}\,\phi R-\frac{1}{4}g_{ab}\,\phi^2\right)\ ,$$

the total stress-energy tensor appearing on the right-hand side still contains some problematic contributions (e.g., $\nabla_a\nabla_b\phi$), which prevent the dominant or even the null energy conditions from being satisfied.

The ϕR interaction can be removed by performing the conformal transformation

$$g_{ab}=(1+2\alpha\phi)^{-\frac{2}{D-2}}\,\bar{g}_{ab}\ . \tag{13}$$

In terms of $\bar{g}_{ab}$, the action (12) becomes

$$I_2=\int d^Dx\sqrt{-\bar{g}}\left\{\frac{1}{16\pi G}\left[\bar{R}-\frac{D-1}{D-2}\left(\frac{2\alpha}{1+2\alpha\phi}\right)^2\bar{\nabla}_a\phi\bar{\nabla}^a\phi-\alpha(1+2\alpha\phi)^{-\frac{D}{D-2}}\phi^2\right]+(1+2\alpha\phi)^{-\frac{D}{D-2}}L_m[\psi,(1+2\alpha\phi)^{-\frac{2}{D-2}}\bar{g}]\right\}\ , \tag{14}$$

which includes the standard Einstein-Hilbert action for $\bar{g}_{ab}$ and the auxiliary scalar ϕ with less conventional couplings—see below. The $\bar{g}_{ab}$ equations of motion are now

$$\bar{R}_{ab}-\frac{1}{2}\bar{g}_{ab}\,\bar{R}=\frac{8\pi G}{1+2\alpha\phi}T^{m}_{ab}[\psi,(1+2\alpha\phi)^{-\frac{2}{D-2}}\bar{g}]+\frac{D-1}{D-2}\left(\frac{2\alpha}{1+2\alpha\phi}\right)^2\bar{\nabla}_a\phi\bar{\nabla}_b\phi-\frac{1}{2}\bar{g}_{ab}\left[\frac{D-1}{D-2}\left(\frac{2\alpha}{1+2\alpha\phi}\right)^2\bar{\nabla}_c\phi\bar{\nabla}^c\phi+\alpha(1+2\alpha\phi)^{-\frac{D}{D-2}}\phi^2\right]\ . \tag{15}$$

The most important feature of this final theory for our purposes is that, assuming $1 + 2\alpha\phi > 0$, the total stress-energy tensor appearing on the right-hand side above satisfies the dominant (and hence the null) energy condition for positive α if the original matter Lagrangian satisfies this energy condition.

Now suppose we have a stationary black hole solution to (15) whose event horizon is a Killing horizon. Then, since the dominant energy condition holds for $\alpha > 0$, the surface gravity must be constant—i.e., the zeroth law holds [3]. We can therefore identify the entropy via any of the usual methods, e.g., Wald's derivation of the first law.

Given the absence of higher derivative or unconventional gravity couplings, the black hole entropy for I_2 is given by $\bar{S} = \bar{A}/(4G)$, just as for Einstein gravity. Since the equations of motion (15) are Einstein's equations, and the null energy condition is satisfied by the total stress-energy tensor, Hawking's proof of the area theorem is valid for the I_2 theory with the assumptions that cosmic censorship holds for $\bar{g}_{ab}$ and that $1 + 2\alpha\phi > 0$. (The latter assumption will be further discussed below.) Hence there is a classical entropy increase theorem for the theory defined by I_2 in Eq. (14).

Now Eq. (13) along with $\phi = R$ provides a mapping between the solutions for the Einstein-plus-scalar theory defined by I_2, and the original higher curvature theory defined by I_0, in which the metrics are related by a conformal transformation

$$\bar{g}_{ab} = (1 + 2\alpha R)^{\frac{2}{D-2}} g_{ab} \quad . \tag{16}$$

The conformal transformation (16) preserves the causal structure of the solutions and, if g_{ab} is asymptotically flat, then so is $\bar{g}_{ab}$. Thus, if g_{ab} is an asymptotically flat black hole, then so is $\bar{g}_{ab}$, and they have the same horizon and surface gravity [19]. In particular, stationary black hole solutions of the I_0 theory have constant surface gravity, provided they have Killing horizons. (Note that, since a Killing vector remains a Killing vector under *stationary* conformal transformations, the event horizon of a stationary I_0 black hole is a Killing horizon if and only if the same is true for the corresponding I_2 black hole.) The zeroth law therefore holds for I_0.

On the other hand, since the asymptotic forms of g_{ab} and $\bar{g}_{ab}$ agree, the mass and angular momenta of the two spacetimes agree. Also the angular velocities agree, since the null combination of time translation and rotation Killing fields agree on the horizon. In short, we have shown all of the ingredients, other than the entropy, in the first law (2) agree. Thus, for all variations, the changes in the entropies must also agree. Therefore the entropies themselves are equal to within a constant in each connected class of stationary black hole solutions. Since the area increase theorem for the Einstein-plus-scalar theory gives $\delta\bar{S} \geq 0$ in any dynamical process connecting two stationary states, we conclude that $\delta S \geq 0$ for the corresponding process in the higher curvature theory. We have thus established a classical second law in the higher curvature theory defined by the action I_0 in Eq. (9).

It should be emphasized that the first law, which applies to variations away from a stationary black hole background, does not uniquely determine the form of the entropy for nonstationary states [9, 17]. In Einstein gravity, since the entropy is proportional to the horizon area, it has a natural extension to a cross section of an arbitrary nonstationary black hole horizon, and the area theorem shows that this nonstationary entropy never decreases, even *during* a dynamical process. We can obtain a similar result for the higher curvature theory as follows.

Using the conformal relation (16) between the two metrics, the "barred" entropy (area) can be expressed directly in terms of $g_{\mu\nu}$:

$$\bar{S}(\bar{g}) = \frac{1}{4G}\int_{\bar{\mathcal{H}}} d^{D-2}x\,\sqrt{\bar{h}} = \frac{1}{4G}\int_{\bar{\mathcal{H}}} d^{D-2}x\,\sqrt{h}\,(1+2\alpha\, R) \ . \tag{17}$$

For stationary black holes the right-hand side agrees with the entropy in the higher curvature theory (11) as determined directly from the first law in that theory (it was already argued above that $\bar{\mathcal{H}}$ corresponds to a cross section of the event horizon for the metric g_{ab} as well). This agreement is explained by the reasoning given two paragraphs above. In presenting the result (11) for the black hole entropy as determined by the first law, we chose the simplest geometric formula which naturally extends to a dynamical horizon. Here we have shown that, by virtue of the area theorem in the conformally related theory, the entropy given by that particular formula, reproduced in Eq. (17), obeys the second law even *during* a dynamical process.

The relation (16) gives an unambiguous result for the dynamical entropy and so can be used to resolve the ambiguities [17, 11] inherent in the Noether charge construction of [9]. In the present case, the alternate proposal for dynamical entropy of Ref. [11], which used a boost invariant projection, gives a result for the dynamical entropy that differs from (17) for nonstationary black holes. Unless there are two different entropy functionals obeying a local increase law, it appears that the proposal of Ref. [11] is inconsistent with the second law during dynamical processes in the present theories.

There is one considerable assumption in preceding discussion, which we have not yet addressed. For the mapping between the solutions of the two theories (16) to exist and for the total stress-energy tensor in Eq. (15) to satisfy the dominant energy condition (1), it is necessary that the factor $1 + 2\alpha R$ is positive. Thus, given a black hole solution of the higher curvature theory, one must have $R > -\frac{1}{2\alpha}$ (for positive α) everywhere outside of the black hole and on the event horizon.

From the point of view of the Einstein-plus-scalar theory, one requires that $\phi > -\frac{1}{2\alpha}$ everywhere outside of the event horizon for the mapping to a solution of the higher curvature theory to exist. Recall that cosmic censorship was assumed in the proof of the area increase theorem for this theory. This assumption rules out dynamical processes in which a black hole begins with a configuration with $\phi > -\frac{1}{2\alpha}$ everywhere initially, and evolves to one with $\phi \leq -\frac{1}{2\alpha}$ in some region outside of the horizon.

The reason is that, by the equations of motion (15) when $\phi = -\frac{1}{2\alpha}$, the stress-energy tensor is singular and hence the Einstein tensor, $\bar{R}_{ab} - \frac{1}{2}\bar{g}_{ab}\bar{R}$, is singular [20]. Cosmic censorship for $\bar{g}_{ab}$ would only allow such curvature singularities to develop behind the event horizon, and hence rules out any process in which $\phi = -\frac{1}{2\alpha}$ is reached outside of the horizon.

An alternative argument showing that it is consistent to make the restriction $\phi > -\frac{1}{2\alpha}$ can be given by considering the character of the potential term in the Einstein-plus-scalar theory [21]. The nonstandard kinetic term for the scalar field ϕ in Eq. (14) can be replaced with an ordinary one by defining a new scalar field $\varphi := \beta^{-1}\ln(1+2\alpha\phi)$, where $\beta = \sqrt{8\pi G(D-2)/(D-1)}$. In terms of φ the action I_2 becomes

$$I_3 = \int d^Dx\sqrt{-\bar{g}}\left[\frac{1}{16\pi G}\bar{R} - \frac{1}{2}\bar{\nabla}_a\varphi\bar{\nabla}^a\varphi - V(\varphi) + e^{-\frac{D}{D-2}\beta\varphi}L_m(\psi, e^{-\frac{2}{D-2}\beta\varphi}\bar{g})\right], \tag{18}$$

where $V(\varphi) = \frac{1}{64\pi G\alpha}e^{-\frac{D}{D-2}\beta\varphi}(e^{\beta\varphi}-1)^2$. Now the singular point $\phi = -\frac{1}{2\alpha}$ corresponds to $\varphi \to -\infty$. Provided $\alpha > 0$, the potential $V(\varphi)$ rises exponentially as $\varphi \to -\infty$. The term involving the matter Lagrangian has the same exponential for a prefactor, and so one may worry that it may undermine the barrier due to $V(\varphi)$. The kinetic terms for the matter fields will include at least one inverse metric which will bring the rate of exponential growth down by a factor of $\exp[2\beta\varphi/(D-2)]$ for these contributions. We will assume that any matter potential is non-negative, which is implied by the dominant energy condition (1) for T^m_{ab}, so that these terms can only increase the potential barrier as $\varphi \to -\infty$. Thus, as long as the metric and matter fields do not become singular, the dynamics of φ as $\varphi \to -\infty$ will be dominated by the potential barrier so φ will not run off to $-\infty$. Therefore, initial data satisfying the bound $\phi > -\frac{1}{2\alpha}$ will evolve within the bound, as long as the other fields remain nonsingular.

Note that the argument just given breaks down if $\alpha < 0$ since the potential is then *negative* and exponentially *falling* as $\varphi \to -\infty$. Hence, the theory appears unstable for negative α. The previous argument for nondecreasing entropy did not seem to require that α be positive because the null energy condition is satisfied for any α. It did assume cosmic censorship however which, presumably, would be violated in the unstable theory with $\alpha < 0$. Note also that we previously used this condition on α in order to establish the zeroth law.

Given the above arguments that $\phi = -\frac{1}{2\alpha}$ is never reached outside the event horizon for positive α, one may also rule out processes in the higher curvature theory in which a black hole evolves to reach $R = -\frac{1}{2\alpha}$ somewhere outside of the event horizon. From a superficial examination of the higher curvature equations of motion (10), $R = -\frac{1}{2\alpha}$ does not appear to be singular. However, there is no obstruction to mapping the initial part of the evolution to the Einstein-plus-scalar theory, where it becomes a process leading up to a naked singularity at the point where $\phi = -\frac{1}{2\alpha}$, as discussed above. Such processes were ruled out by the assumption of cosmic censorship though, and hence we have also ruled out the corresponding evolution in the higher curvature theory.

IV. ACTIONS POLYNOMIAL IN R

The results of the previous section are easily generalized for higher curvature theories with actions of the form

$$I_o = \int d^Dx\sqrt{-g}\left[\frac{1}{16\pi G}[R + P(R)] + L_m(\psi, g)\right], \tag{19}$$

where P is a polynomial in the Ricci scalar, $P(R) = \sum_{n=2} a_n R^n$. Introducing an auxiliary scalar field ϕ, as in (12) of the preceding section, this theory can be reexpressed using a new action linear in R:

$$I_1 = \int d^Dx\sqrt{-g}\left\{\frac{1}{16\pi G}\left[R + P(\phi) + (R-\phi)P'(\phi)\right] + L_m(\psi, g)\right\}. \tag{20}$$

Here, the primes denote differentiation of P with respect to ϕ—i.e., $P'(\phi) = \sum_{n=2} n\,a_n\phi^{n-1}$. The theory defined by this new action is *not* precisely equivalent to the original theory defined by Eq. (19). Rather the ϕ equation of motion yields two classes of solutions: *i)* $\phi = R$ and *ii)* $\phi = \phi_0$ where ϕ_0 is a constant satisfying $P''(\phi_0) = 0$. Substituting (i) back into the action (20) yields (19). Thus these solutions correspond to solutions of the higher curvature theory, which we wish to study. Substituting (ii) into Eq. (20) yields Einstein gravity with an effective Newton constant, $G_{\rm eff} = G[1 + P'(\phi_0)]$, and an effective cosmological constant $\Lambda_{\rm eff} = [\phi_0 P'(\phi_0) - P(\phi_0)]/2[1 + P'(\phi_0)]$. Thus the latter solutions are spurious for our purposes since they do not correspond to solutions of the original higher curvature theory. In the present analysis we only consider asymptotically flat solutions, which would rule out the second class of solutions because of the presence of an effective cosmological constant. Even in the case of an accidental degeneracy where $\Lambda_{\rm eff} = 0$, one still knows that the asymptotically flat solutions for action (20) include all of those for the original action (19).

As in the preceding section, the action (20) can be transformed to Einstein gravity coupled to a scalar field via the conformal transformation

$$\bar{g}_{ab} = [1 + P'(\phi)]^{\frac{2}{D-2}} g_{ab} \tag{21}$$

yielding the action[18]

$$I_2 = \int d^Dx\sqrt{-\bar{g}}\frac{1}{16\pi G}\left\{\bar{R} - \frac{D-1}{D-2}\left(\frac{P''(\phi)}{1+P'(\phi)}\right)^2\bar{\nabla}_a\phi\bar{\nabla}^a\phi + [1+P'(\phi)]^{-\frac{D}{D-2}}\{P(\phi) - \phi P'(\phi) + 16\pi G\,L_m(\psi, [1+P'(\phi)]^{-\frac{2}{D-2}}\bar{g})\}\right\}. \tag{22}$$

Equivalence of (20) and (22) holds provided the conformal transformation is nonsingular—i.e., $(1 + P') > 0$.

If the matter Lagrangian L_m yields a stress-energy tensor satisfying the dominant energy condition (1), then the action I_2 also yields a stress-energy tensor satisfying the dominant energy condition provided $1 + P'(\phi) > 0$ and $\phi P'(\phi) - P(\phi) > 0$. (These conditions on P will be discussed further below.) In this case, it follows that the surface gravity is constant over the Killing horizon of a stationary black hole for I_2. Further, with the assumption that cosmic censorship applies in this theory, one can prove the area increase theorem, and so then one has shown that $\delta\bar{S} \geq 0$ in any evolution of a black hole.

Assuming we have a solution in class (i), the mapping between the theories takes the form of a conformal transformation that goes to the identity at infinity. As above, this means that a black hole in one theory is mapped to a black hole of the other theory and, further, that the event horizons, surface gravities, and entropies of the two black hole solutions coincide. Thus the constancy of the surface gravity and the area increase theorem of the Einstein-plus-scalar theory translate to the zeroth law and an entropy increase theorem for the original higher curvature theory, respectively.

The mapping between the solutions of the two theories yields a formula for the entropy in the higher curvature theory:

$$\bar{S}(\bar{g}) = \frac{1}{4G}\int_{\bar{\mathcal{H}}} d^{D-2}x\,\sqrt{\bar{h}} = \frac{1}{4G}\int_{\bar{\mathcal{H}}} d^{D-2}x\,\sqrt{h}\,[1+P'(R)]. \tag{23}$$

As expected one recovers the same expression for the black hole entropy in the higher curvature theory that was determined by directly examining the first law in that theory for variations from stationary black holes [17]. Note that the conformal transformation yields an unambiguous definition of the dynamical (nonstationary) black hole entropy for the higher curvature theory, whereas the entropy functional determined from the first law is not unique [9, 17].

Of course, the equivalence of the dynamics defined by the actions I_0 and I_2 required that $\phi = R$, and that $1 + P'(R) > 0$ everywhere for solutions of the higher curvature theory. The latter assumption requires that R lie within some domain including zero, whose precise boundaries will be defined by the couplings a_n appearing in $P(R)$. Alternatively, there are restrictions on allowed values of the auxiliary scalar ϕ in the Einstein-plus-scalar theory. For our purposes it is enough that these restrictions hold outside the horizon and on an open set including the horizon.

One can argue as in the preceding section that cosmic censorship rules out processes in which a black hole evolves from a configuration with $1 + P' > 0$ to one in which this inequality is violated. This is done by rewriting the action yet one more time, as in the preceding section, with the change of variables $\varphi := \beta^{-1}\ln[1 + P'(\phi)]$, where $\beta = \sqrt{8\pi G(D-2)/(D-1)}$. In terms of φ the action I_2 becomes

$$I_3 = \int d^Dx\sqrt{-\bar{g}}\left[\frac{1}{16\pi G}\bar{R} - \frac{1}{2}\bar{\nabla}_a\varphi\bar{\nabla}^a\varphi - V(\varphi) + e^{-\frac{D}{D-2}\beta\varphi}L_m(\psi, e^{-\frac{2}{D-2}\beta\varphi}\bar{g})\right],$$

where now

$$V(\varphi) = \frac{1}{16\pi G}e^{-\frac{D}{D-2}\beta\varphi}(\phi P' - P)\,.$$

Equivalence of I_2 and I_3 requires that one can invert $1 + P'(\phi) = e^{\beta\varphi}$ for $\phi = \phi(\varphi)$. This requires that P'' has a definite sign in the domain of interest for ϕ (which includes $\phi = 0$).

As before, the singular point $1+P'(\phi) = 0$ corresponds to $\varphi \to -\infty$. There is a potential barrier as $\varphi \to -\infty$ provided $\phi P' - P$ is positive. Note that $\phi P' - P$ vanishes at $\phi = 0$, and $(\phi P' - P)' = \phi P''$. If we restrict P'' to be positive, $(\phi P' - P)$ will be positive for all $\phi \neq 0$. Thus we can argue, as in the preceding section, that if $P'' > 0$ (and the matter stress-energy satisfies the dominant energy condition), initial data satisfying the bound $1 + P' > 0$ outside the horizon will evolve within this bound, as long as the other fields remain nonsingular outside the horizon. As a bonus, $P'' > 0$ implies the positivity of $\phi P' - P$, which was required for the dominant energy condition (and therefore the proof of the zeroth law) to hold. Actually, given that $P'' > 0$, $1 + P'$ (which is one at $\phi = 0$) will generically reach zero at some negative value of ϕ, which then defines the boundary for the range of interest. It is enough then to require $P(\phi)'' > 0$ for $\phi \geq \phi_1$, where ϕ_1 denotes the negative value of ϕ nearest the origin for which $1 + P'$ vanishes. If $P'' < 0$ in the range of interest, the theory is probably unstable, as in the $R + \alpha R^2$ theory with negative α.

V. DIRECT PROOF OF THE SECOND LAW

Our method of establishing the second law for certain higher curvature theories used the fact that these theories are conformally related to ordinary Einstein theories in which the area theorem holds. This special feature of these theories is not shared by most higher curvature theories, so it would be interesting to see how the second law could be established *directly* in these theories, without making use of the conformal transformation technique. Such an exercise would be instructive for efforts to establish entropy increase theorems for theories that are not susceptible to the conformal transformation "trick." In order to gain some insight into this question, in the present section we will construct such a direct proof.

Suppose that the black hole entropy of a gravity theory takes the form

$$S = \frac{1}{4G}\int_{\mathcal{H}} d^{D-2}x\,\sqrt{h}\,e^{\rho}, \tag{24}$$

where e^ρ is a scalar function of the local geometry at the horizon. For the class of theories considered in the preceding section one has $e^\rho = 1 + P'(R)$. The method to be used here will rely critically on the fact that e^ρ is necessarily positive, and $\rho = 0$ when the curvature vanishes.

We wish to consider the change of this entropy along the null congruence generating the event horizon under any dynamical evolution. Let k^a be the null tangent vector field of the horizon generators with respect to the affine parameter λ. Then one has

$$\partial_\lambda S = \frac{1}{4G}\int_{\mathcal{H}} d^{D-2}x\sqrt{h}\, e^\rho\, \tilde{\theta}$$

with

$$\tilde{\theta} := \theta + \partial_\lambda \rho, \tag{25}$$

where $\theta = d(\ln\sqrt{h})/d\lambda = \nabla_a k^a$ is the expansion of the horizon generators.

Now the question is whether or not there can exist a point along the null geodesics at which $\tilde{\theta}$ becomes negative. In order to answer this question, we use the Raychaudhuri equation, as in the proof of area theorem, to obtain an expression for $\partial_\lambda\tilde{\theta}$:

$$\begin{aligned}\partial_\lambda\tilde{\theta} &= \partial_\lambda\theta + \partial_\lambda^2\rho \\ &= -\tfrac{1}{D-2}\theta^2 - \sigma^2 - k^a k^b R_{ab} + k^a k^b \nabla_a\nabla_b\rho, \end{aligned}\tag{26}$$

where σ^2 is the square of the shear.

For the $R + P(R)$ theories, it is easy to see that the equations of motion imply that $k^a k^b(R_{ab} - \nabla_a\nabla_b\rho) = (8\pi G\, e^{-\rho} T^m_{ab} + \nabla_a\rho\nabla_b\rho)k^a k^b$, which is non-negative provided the null energy condition holds for the matter fields (and $e^\rho > 0$). Thus in those theories one has

$$\partial_\lambda\tilde{\theta} \le -\tfrac{1}{D-2}\theta^2,$$

or

$$\partial_\lambda[\tilde{\theta}^{-1}] \ge \frac{1}{D-2}(\theta/\tilde{\theta})^2. \tag{27}$$

Now we follow Hawking's proof of the area theorem, with $\tilde{\theta}$ in place of θ. Suppose at some point on the horizon we have $\tilde{\theta} < 0$. Then in a neighborhood of that point one can deform a spacelike slice of the horizon slightly outward to obtain a compact spacelike surface Σ so that $\tilde{\theta} < 0$ everywhere on Σ, $\tilde{\theta}$ being defined along the outgoing null geodesic congruence orthogonal to Σ. If cosmic censorship is assumed, then there is necessarily some null geodesic orthogonal to Σ that remains on the boundary of the future of Σ all the way out to $\mathcal{I}^+$ [14]. Asymptotic flatness (where components of the Riemann tensor in an orthonormal frame all fall off at least as r^{1-D}) implies that $\rho \to 0$ like λ^{1-D} at infinity, whereas θ goes like λ^{-1}, where λ is the affine parameter along an outgoing null geodesic. Therefore $\theta/\tilde{\theta} \to 1 + O(\lambda^{(1-D)})$, so the inequality (27) implies that, as one follows the geodesic outwards from Σ, $\tilde{\theta}$ reaches $-\infty$ at some finite affine parameter. Since $\tilde{\theta} = \theta + \partial_\lambda\rho$, this means that either θ or $\partial_\lambda\rho$ goes to $-\infty$. In the former case we have a contradiction, as in the area theorem, since it implies there is a conjugate point on the geodesic, which cannot happen since the geodesic stays on the boundary of the future of Σ all the way out to $\mathcal{I}^+$. In the latter case we have a naked singularity, since $\partial_\lambda\rho = e^{-\rho}\partial_\lambda e^\rho$, and e^ρ was assumed from the beginning in (24) to be a nonvanishing function of the curvature.

For the $P(R)$ theories we have in particular $\partial_\lambda\rho = k^a\nabla_a\rho = (1+P')^{-1}P'' k^a\nabla_a R$. In the preceding section we argued that $1 + P'$ never goes to zero outside the horizon in the case of a theory with $P'' > 0$, so for these theories the divergence of $\partial_\lambda\rho$ implies divergence of R or $k^a\nabla_a R$, but these divergences also violate cosmic censorship. Therefore we conclude that cosmic censorship and the null energy condition for the matter imply that the black hole entropy (24) can never decrease for the stable theories. Note that in making this argument we have used the condition $1 + P' > 0$ that was established via the conformal transformation trick, so we do not really have a fully "direct proof" of the second law.

The above argument suggests that the "weakest" naked singularity which creates a violation of the second law would be a divergence in $k^a\nabla_a R$. It seems that this could happen even if the curvature itself is nonsingular everywhere. However, if one imposes also the equations of motion of the theory, then a divergence in $k^a\nabla_a R$ would necessarily entail also a divergence in either the curvature tensor or the matter stress tensor.

As a final note, we demonstrate that the "extended" Raychaudhuri equation (26) lends itself to a "physical process" derivation of the first law [15], and also of the form of the entropy, at least for the actions polynomial in the Ricci scalar. A sketch of such a derivation follows: Following the quasistationary process discussion in Sec. II, one concludes that the first law is satisfied if there is an entropy functional S satisfying the equality in (8),

$$\frac{\kappa}{2\pi}\Delta S = \Delta M - \Omega^{(\alpha)}\Delta J_{(\alpha)} = -\int_H T_{ab}\chi^a d\Sigma^b \ .$$

Now on the horizon, where χ^a is null, the equations of motion imply that $8\pi G T_{ab}\chi^a\chi^b = e^\rho\chi^a\chi^b(R_{ab} - \nabla_a\nabla_b\rho - \nabla_a\rho\nabla_b\rho)$ where $e^\rho = 1 + P'(R)$. The horizon generating Killing field $\chi^a\partial_a = \partial_v$ is related to the affinely parametrized null tangent to the horizon $k^a\partial_a = \partial_\lambda$ by $\chi^a\partial_a = \kappa\lambda k^a\partial_a$ where $\lambda = \exp(\kappa v)$ [15]. Further the volume element in the above flux integral may be written $d\Sigma^a = -d^{D-2}x\sqrt{h}\, d\lambda\, k^a$ [15]. Hence using the extended Raychaudhuri equation (26) and the equations of motion, and neglecting terms of higher than linear order in the perturbation, the above equation yields

$$\begin{aligned}\frac{\kappa}{2\pi}\Delta S &= \kappa\int_H d^{D-2}x\sqrt{h}\, d\lambda\, \lambda k^a k^b T_{ab} \\ &= \frac{\kappa}{8\pi G}\int_H d^{D-2}x\sqrt{h}\, d\lambda\, \lambda\, e^\rho\, k^a k^b(R_{ab} - \nabla_a\nabla_b\rho) \\ &= -\frac{\kappa}{8\pi G}\int_H d^{D-2}x\sqrt{h}\, d\lambda\, \lambda\, e^\rho\partial_\lambda\tilde{\theta} \\ &= -\frac{\kappa}{8\pi G}\oint d^{D-2}x\sqrt{h}\lambda e^\rho\tilde{\theta}\Big|_{\Sigma_i}^{\Sigma_f} \\ &\quad + \frac{\kappa}{8\pi G}\int_H d\lambda\, d^{D-2}x\sqrt{h}\, e^\rho\tilde{\theta} \ .\end{aligned}$$

In this final expression, $\tilde{\theta}$ vanishes on the final and initial horizon slices where the horizon is stationary, and so the first contribution is zero. By the definition of $\tilde{\theta}$ in Eq. (25), we see that $\sqrt{h}\, e^\rho\tilde{\theta}$ is the total derivative $\partial_\lambda(\sqrt{h}\, e^\rho)$, so the second contribution is just $\frac{\kappa}{2\pi}\Delta S$

where S is precisely the entropy given in Eq. (23). The form of the entropy functional for these higher curvature theories can thus be inferred directly by consideration of quasistationary accretion processes.

VI. DISCUSSION

In this paper, we have presented two cases where a classical entropy increase theorem applies in higher curvature gravity. These are the following.

(i) For quasistationary processes in which a (vacuum) black hole accretes positive energy matter, i.e., $T^m_{ab}\ell^a\ell^b \geq 0$ for any null vector ℓ^a, the second law is a direct consequence of the first law of black hole mechanics, independent of the details of the gravitational action.

(ii) For higher curvature theories of the form (19) the black hole entropy is given by

$$S(g) = \frac{1}{4G}\int_{\mathcal{H}} d^{D-2}x\,\sqrt{h}\,[1+P'(R)] \quad . \tag{28}$$

This entropy satisfies the second law in any processes involving matter fields that satisfy the null energy condition. Our proof of the second law requires that the coupling constants a_n appearing in $P(R)$ be restricted in such a way that $P''(R)$ is positive for positive R, and also between $R = 0$ and the largest negative value of R where $1 + P'(R)$ vanishes. The latter ensures that $1 + P'(R)$ is positive everywhere outside and on the event horizon of the black hole spacetimes.

The expression $1 + P'(R)$ must be positive in order to implement the conformal transformation between the original higher curvature theory and the Einstein-plus-scalar theory, and also to ensure the null energy condition is satisfied in the latter theory. This positivity is also an essential ingredient for the direct proof in Sec. V. It is interesting that precisely the same expression plays the role of the entropy surface density in Eq. (28). Thus the positivity restriction translates on the horizon to the condition that the local entropy density should be positive everywhere. In particular it requires that the total black hole entropy is always positive. The latter is a minimum requirement that must be satisfied if this entropy is to have a statistical mechanical origin. The fact that we actually require a *local* positivity condition on the entropy density is suggestively consistent with the idea that this density may have a statistical interpretation. In any event, these (and other higher curvature) theories may provide a more refined test of the various proposals to explain the statistical origin of black hole entropy.

The direct proof of the second law (in Sec. V) is essentially a translation, via the conformal transformation, of Hawking's proof of the area theorem applied to the Einstein-plus-scalar theory. Nevertheless, it provides an illustration of how one might hope to prove an entropy increase theorem for other higher curvature theories. Naively, with the assumption of cosmic censorship, this proof can be extended to theories with interactions of the form $R^{2n+1}\nabla^2 R$. A closer examination of the latter theories indicates that they are unstable however, and so the assumption of cosmic censorship appears unlikely to hold. This highlights the problem that, in dealing with the higher curvature theories directly, establishing the stability of asymptotically flat solutions requires an involved analysis. In fact, even for our direct proof in Sec. V, we relied on results about the stability of the theories derived in Secs. III and IV by examining the Einstein-plus-scalar theory.

An obstacle to constructing a direct proof of the second law in general is that the entropy as determined from the first law does not uniquely determine the form of the dynamical entropy. Thus, to begin, one would not know for which entropy density one should be attempting to prove an increase theorem. In the higher curvature theories considered in Secs. III and IV, this ambiguity is resolved by the conformal transformation, which yields precisely Eq. (28) when inserted into $\bar{S} = \bar{A}/(4G)$. In Ref. [17], the present authors introduced an alternative construction for black hole entropy involving field redefinitions, which also appears to avoid any ambiguities. The form of the higher curvature actions for which this approach is applicable is not completely general, but it does extend beyond those theories considered in this paper.

In Einstein gravity, the zeroth law for a Killing horizon can be proved if the dominant energy condition is assumed [3]. Through the conformal transformation technique, this proof was extended to the higher curvature theories introduced in Secs. III and IV, at least with certain restrictions on the coupling constants. In the case of a regular bifurcate Killing horizon, one can show that the surface gravity is constant irrespective of the underlying gravitational dynamics [5]. However, there is no independent proof that Killing horizons in a general theory necessarily possess a regular bifurcation surface, and so for general higher curvature theories the validity of the zeroth law remains an important open question.

It is worth emphasizing that unless the event horizon is a Killing horizon, the above-mentioned proofs of the zeroth law are not applicable. In Einstein gravity, Hawking proved that the event horizon of a stationary black hole must be a Killing horizon [14]. To our knowledge, this proof has not been extended to general higher curvature theories, or even to higher dimensional Einstein gravity. For the higher curvature theories considered in this paper, at least in four dimensions, it seems likely that Hawking's proof can be imported via the conformal transformation relating the theory to Einstein gravity with matter. For stationary black holes in more general higher curvature theories, whether or not stationary event horizons are necessarily Killing horizons is another important open question.

We now come to considering the two shortcomings of the calculations presented in this paper. For all of the cases considered here, the dominant energy condition must hold for the matter fields. Within the present framework, though, it is natural to include higher derivative matter couplings [e.g., $R(\nabla\phi)^2$ or $(\nabla^2\phi)^2$] as well as higher curvature interactions. Generally the former will spoil this positive energy condition. This is perhaps less a criticism of the discussion of actions polynomial in the Ricci scalar since they are already theories with a restricted set of interactions.

The second shortcoming is revealed by the instabilities faced in Secs. III and IV. There our proof of the second law failed for certain values of the coupling constants (e.g., $\alpha < 0$ for the R^2 theory) because we found these theories to be unstable, and hence it appeared that the conformal transformation technique could not be implemented and cosmic censorship was not a valid assumption. Considering the R^2 theory represented by the action (18) with the auxiliary scalar φ, we see that the scalar potential is proportional to $1/\alpha$. Thus the instability is nonperturbative in the higher curvature coupling constant. Further, a perturbative analysis (around flat space) reveals unstable modes with imaginary frequencies of the order of $1/\alpha$. We expect these remarks also to be true of the unstable theories in Sec. IV. The original framework, which we set out for our investigations, though, was Einstein gravity *perturbatively* corrected by higher curvature corrections. The problem with the present analysis is that we are actually taking these theories at their face value, rather than treating the higher curvature terms perturbatively. This suggests that our analysis should be modified to incorporate the ideas of perturbative reduction [22]. One might hope that the problems with instability and cosmic censorship would be avoided in this way. Since the perturbative treatment would extend to all of the higher derivative interactions, including those of the matter fields, such an approach may also be able to circumvent the requirement that the full matter stress-energy tensor satisfy the null energy condition [23].

Establishing the second law for higher curvature theories within a perturbative framework would be a valuable extension of our present results, since we expect nonperturbative instabilities to be a generic feature of these theories when they are considered as fundamental [24]. If one were to find that no second law holds *even perturbatively* for certain interactions or certain values of the coupling constants, one might suspect that those effective actions are unphysical. Perhaps the requirement that the entropy (or even the entropy density) be positive might provide a further restriction on the form of physically relevant effective actions.

In this paper we have only considered an intrinsic or classical second law—i.e., we have only dealt with the increase of the black hole entropy alone. In general relativity, we know that the effective transfer of negative energy from quantum fields to a black hole can lead to a decrease in the horizon entropy (i.e., horizon area), and the same is true for these higher curvature effective theories since black holes still produce Hawking radiation in these theories. Thus it is important to ask whether a generalized second law $[\delta(S_{\rm BH} + S_{\rm outside}) \geq 0]$ holds. In general relativity, there are arguments that the generalized second law applies for quasistationary processes involving positive energy matter [25]. These arguments seem to carry over to stable higher curvature gravity theories as well, since they do not involve the equations of motion but rather lean on the first law and the maximum entropy property of thermal radiation.

Another approach to this question would be to incorporate the effects of the Hawking radiation in the effective action by the introduction of nonlocal contributions [26, 27]. In two dimensions where an explicit nonlocal term can be calculated [27], Wald's techniques have been applied to determine the nonlocal (radiation) contribution to the geometric horizon entropy [28]. Further, in that model, one can show that the generalized entropy, which now includes the contributions of the Hawking radiation, will satisfy a second law, even for evaporating black holes [29]. Some model-independent constructions for the nonlocal action exist in four dimensions [26], and so one could in principle apply Wald's techniques to develop an expression for the black hole entropy in these theories. Perhaps the requirement that this entropy satisfies the second law would impose useful restrictions on the underlying nonlocal action. In any event, addressing the validity of the generalized second law remains an important open problem.

ACKNOWLEDGMENTS

We would like to acknowledge useful discussions with R. Woodard and J. Friedman. R.C.M. was supported by NSERC of Canada, and Fonds FCAR du Québec. T.J. and G.K. were supported in part by NSF Grant No. PHY91-12240. We would also like to thank the ITP, UCSB for their hospitality during the early stages of this work. Research at the ITP, UCSB was supported by NSF Grant No. PHY89-04035. T.J. would like to thank the Institute for Theoretical Physics at the University of Bern for hospitality and Tomalla Foundation Zurich for support while some of this work was done.

[1] See, for example, N.D. Birrell and P.C.W. Davies, *Quantum Fields in Curved Space* (Cambridge University Press, Cambridge, England, 1982).

[2] Alternatively, this identification follows from a semiclassical evaluation of the partition function or the density of states in quantum gravity with any action, as has been shown explicitly in the case of Einstein gravity. See, for example, S.W. Hawking, in *General Relativity, An Einstein Centenary Survey*, edited by S.W. Hawking and W. Israel (Cambridge University Press, Cambridge, England, 1979); J.D. Brown and J.W. York, Phys. Rev. D **47**, 1420 (1993).

[3] J.M. Bardeen, B. Carter, and S.W. Hawking, Commun. Math. Phys. **31**, 161 (1973); see also Ref. [4].

[4] R.M. Wald, *General Relativity* (University of Chicago Press, Chicago, 1984).

[5] I. Rácz and R.M. Wald, Class. Quantum Grav. **9**, 2643 (1992).

[6] G.W. Gibbons and S.W. Hawking, Phys. Rev. D **15**, 2752 (1977); S.W. Hawking, in *General Relativity, An Einstein Centenary Survey*, edited by S.W. Hawking and W. Israel (Cambridge University Press, Cambridge, England, 1979).

[7] C. Callan, R.C. Myers, and M. Perry, Nucl. Phys. **B311**,

673 (1988); R.C. Myers, *ibid.* **B289**, 701 (1987); D.L. Wiltshire, Phys. Lett. **169B**, 36 (1986); Phys. Rev. D **38**, 2445 (1988); R.C. Myers and J.Z. Simon, *ibid.* **38**, 2434 (1988); Gen. Relativ. Gravit. **21**, 761 (1989); B. Whitt, Phys. Rev. D **38**, 3000 (1988); E. Poisson, Class. Quantum Grav. **8**, 639 (1991); M. Lu and M. Wise, Phys. Rev. D **47**, 3095 (1993); M. Visser, *ibid.* **48**, 583 (1993).

[8] T. Jacobson and R.C. Myers, Phys. Rev. Lett. **70**, 3684 (1993).

[9] R.M. Wald, Phys. Rev. D **48**, 3427 (1993).

[10] M. Visser, Phys. Rev. D **48**, 5697 (1993).

[11] V. Iyer and R.M. Wald, Phys. Rev. D **50**, 846 (1994).

[12] In higher dimensions, one has the possibility of commuting rotations in totally orthogonal planes. Hence the black hole is characterized by the set of angular momenta $J_{(\alpha)}$ associated with rotations in the maximal set of mutually orthogonal planes. For example, see R.C. Myers and M.J. Perry, Ann. Phys. (N.Y.) **172**, 304 (1986).

[13] S.W. Hawking, Commun. Math. Phys. **43**, 199 (1975); this result applies for Einstein gravity in any number of spacetime dimensions, but then the surface area A refers to the volume of a spacelike cross section of the horizon, which is a $(D-2)$-dimensional hypersurface in D dimensions.

[14] S.W. Hawking, Commun. Math. Phys. **25**, 152 (1972); S.W. Hawking and G.R.F. Ellis, *The Large Scale Structure of Spacetime* (Cambridge University Press, Cambridge, England, 1973).

[15] R.M. Wald, *Quantum Field Theory in Curved Spacetime and Black Hole Thermodynamics* (University of Chicago Press, Chicago, 1994).

[16] It may be of interest to extend our discussion to spacetimes with asymptotically constant curvature. To start such an investigation, one might begin by adapting Wald's derivation of the first law [9] to spaces with these new asymptotic boundary conditions.

[17] T. Jacobson, G. Kang, and R.C. Myers, Phys. Rev. D **49**, 6587 (1994).

[18] P.W. Higgs, Nuovo Cimento **11**, 816 (1959); B. Whitt, Phys. Lett. **145B**, 176 (1984); G. Magnano, M. Ferraris, and M. Francaviglia, Gen. Relativ. Gravit. **19**, 465 (1987); Class. Quantum Grav. **7**, 557 (1990); K. Maeda, Phys. Rev. D **39**, 3159 (1989); A. Jakubiec and J. Kijowski, *ibid.* **37**, 1406 (1988); J.D. Barrow and S. Cotsakis, Phys. Lett. B **214**, 515 (1988); G. Magnano and L.M. Sokolowski, Phys. Rev. D **50**, 5039 (1994); D. Wands, Class. Quantum Grav. **11**, 269 (1994).

[19] T. Jacobson and G. Kang, Class. Quantum Grav. **10**, L1 (1993).

[20] We assume that the singularities in the individual terms in the stress-energy tensor of Eq. (15) do not cancel to avoid a curvature singularity.

[21] A. Strominger, Phys. Rev. D **30**, 2257 (1984).

[22] J.Z. Simon, Phys. Rev. D **41**, 3720 (1990); **43**, 3308 (1991); L. Parker and J.Z. Simon, *ibid.* **47**, 1339 (1993).

[23] The starting point for such a perturbative analysis is a solution of Einstein's equations. A vacuum solution with $R = 0$ would therefore be unaffected perturbatively for the theories considered in Secs. III and IV. It is only in the presence of matter that $R \neq 0$, and hence that the special class of higher curvature terms considered there would play a role in modifying the solutions and the black hole entropy.

[24] D.A. Eliezer and R.P. Woodard, Nucl. Phys. **B325**, 389 (1989).

[25] W. Zurek, Phys. Rev. Lett. **49**, 1683 (1982); W.H. Zurek and K.S. Thorne, *ibid.* **54**, 2171 (1985); K.S. Thorne, W.H. Zurek, and R.H. Price, in *Black Holes: The Membrane Paradigm*, edited by K.S. Thorne, R.H. Price, and D.A. MacDonald (Yale University Press, New Haven, CT, 1986), p. 280; W.G. Unruh and R.M. Wald, D **25**, 942 (1982); **27**, 2271 (1983); M.J. Radzikowski and W.G. Unruh, *ibid.* **37**, 3059 (1988); R.M. Wald, in *Highlights in Gravitation and Cosmology*, edited by B.R. Iyer *et al.* (Cambridge University Press, Cambridge, England, 1988); V.P. Frolov and D.N. Page, Phys. Rev. Lett. **71**, 3902 (1993).

[26] A.O. Barvinskii, Yu.V. Gusev, G.A. Vilkoviskii, and V.V. Zhytnikov, J. Math. Phys. **35**, 3543 (1994); **35**, 3525 (1994); Nucl. Phys. **B439**, 561 (1995).

[27] C.G. Callan, S.B. Giddings, J.A. Harvey, and A. Strominger, Phys. Rev. D **45**, R1005 (1992); V. Mukhanov, A. Wipf, and A. Zelnikov, Phys. Lett. B **332**, 283 (1994).

[28] R.C. Myers, Phys. Rev. D **50**, 6412 (1994).

[29] A caveat is that the entropy arising from Wald's construction does not appear to provide the full contribution of the long wavelength quantum fields. [See T.M. Fiola, J. Preskill, A. Strominger, and S.P. Trivedi, Phys. Rev. D **50**, 3987 (1994).] It may be that this discrepancy arises from imprecision in specifying the Green's function in the nonlocal action.