Entropy of null surfaces and dynamics of spacetime

T. Padmanabhan*
IUCAA, Post Bag 4, Ganeshkhind, Pune - 411 007, India

Aseem Paranjape†
Tata Institute of Fundamental Research, Homi Bhabha Road, Colaba, Mumbai - 400 005, India
(Received 17 January 2007; published 2 March 2007)

The null surfaces of a spacetime act as oneway membranes and can block information for a corresponding family of observers (timelike curves). Since lack of information can be related to entropy, this suggests the possibility of assigning an entropy to the null surfaces of a spacetime. We motivate and introduce such an entropy functional for any vector field in terms of a fourth-rank divergence-free tensor P_{ab}^{cd} with the symmetries of the curvature tensor. Extremizing this entropy for all the null surfaces then leads to equations for the *background metric* of the spacetime. When P_{ab}^{cd} is constructed from the metric alone, these equations are identical to Einstein's equations with an undetermined cosmological constant (which arises as an integration constant). More generally, if P_{ab}^{cd} is allowed to depend on both metric and curvature in a polynomial form, one recovers the Lanczos-Lovelock gravity. In all these cases: (a) We only need to extremize the entropy associated with the null surfaces; the metric is *not* a dynamical variable in this approach. (b) The extremal value of the entropy agrees with standard results, when evaluated on shell for a solution admitting a horizon. The role of the full quantum theory of gravity will be to provide the specific form of P_{ab}^{cd} which should be used in the entropy functional. With such an interpretation, it seems reasonable to interpret the Lanczos-Lovelock type terms as quantum corrections to classical gravity.

DOI: 10.1103/PhysRevD.75.064004 PACS numbers: 04.70.Dy, 04.20.Fy, 04.50.+h

I. INTRODUCTION

The strong mathematical resemblance between the dynamics of spacetime horizons and thermodynamics has led several authors [1] to argue that a gravitational theory built upon the principle of equivalence must be thought of as the *macroscopic* limit of some underlying microscopic theory. In particular, this paradigm envisages gravity as analogous to the theory of elasticity of a deformable solid. The unknown, microscopic degrees of freedom of spacetime (which should be analogous to the atoms in the case of solids), will play a role only when spacetime is probed at Planck scales (which would be analogous to the lattice spacing of a solid [2]). Candidate models for quantum gravity, like e.g., string theory, do suggest the existence of such microscopic degrees of freedom for gravity. The usual picture of treating the metric as incorporating the dynamical degrees of freedom of the theory is therefore not fundamental and the metric must be thought of as a coarse grained description of the spacetime at macroscopic scales (somewhat like the density of a solid which has no meaning at atomic scales).

In such a picture, we expect the microscopic structure of spacetime to manifest itself only at Planck scales or near singularities of the classical theory. However, in a manner which is not fully understood, the horizons—which block information from certain classes of observers—link [3] certain aspects of microscopic physics with the bulk dynamics, just as thermodynamics can provide a link between statistical mechanics and (zero temperature) dynamics of a solid. (The reason is probably related to the fact that horizons lead to infinite redshift, which probes *virtual* high energy processes; it is, however, difficult to establish this claim in mathematical terms). It has been known for several decades that one can define the thermodynamic quantities entropy S and temperature T for a spacetime horizon [4]. If the above picture is correct, then one should be able to link the equations describing bulk spacetime dynamics with horizon thermodynamics in a well-defined manner.

There have been several recent approaches which have attempted to quantify this idea with different levels of success [1,5,6]. An explicit example was [7] the case of spherically symmetric horizons in four dimensions. In this case, Einstein's equations can be interpreted as a thermodynamic relation $TdS = dE + PdV$ arising out of virtual radial displacements of the horizon. More recently, it has been shown [8] that this interpretation is not restricted to Einstein's general relativity (GR) alone, but is in fact true for the case of the generalized, higher derivative Lanczos-Lovelock gravitational theory in D dimensions as well. Explicit demonstration of this result has also been given for the case of Friedmann models in the Lanczos-Lovelock theory [9] as well as for rotating and time dependent horizons in Einstein's theory [10]. In a related development, there have been attempts to interpret other gravitational Lagrangians (like $f(R)$ models) in terms of nonequilibrium thermodynamics [11].

In standard thermodynamics, extremization of the functional form of the entropy (treated as a function of the

*Electronic address: paddy@iucaa.ernet.in
†Electronic address: aseem@tifr.res.in

relevant dynamical variables) leads to the equations governing the equilibrium state of the system. This suggests that in the context of gravity as well, one should be able to define an *entropy functional*, which—on extremization—will lead to the equations describing the macroscopic, long-wavelength behavior of the system (which in this case is the spacetime.) That is, if our analogy is to be taken seriously, we should be able to define an ''entropy functional for spacetime,'' the extremization of which should lead to the gravitational field equations for the metric of the spacetime. At the lowest order, this should give Einstein's equations but the formalism will continue to be valid even in the quantum regime. It is important to recall that—even in the case of ordinary matter—there is no such thing as ''quantum thermodynamics''; only quantum statistical mechanics. What quantum theory does is to modify the form of the entropy functional (or some other convenient thermodynamic potential, like free energy); the extremization now leads to equations of motion which incorporate quantum corrections. Similarly, we expect our entropy functional to pick up corrections to the lowest order term, thereby leading to corrections to Einstein's equations. In that sense, this approach is very general and this is what we will develop in this paper.

A crucial point to note is that the concept of entropy, both in the standard thermodynamics as well as in the gravitational context, stems from the fact that certain degrees of freedom are *not observable* for certain classes of observers. (Throughout the paper, we will use the term observers to mean family of timelike curves, without any extra connotations.) In the context of a metric theory of gravity this is inevitably linked with the existence of one-way membranes which are provided by *null surfaces* in the spacetime. The classical black hole event horizons, like e.g., the one in the Schwarzschild spacetime, are special cases of such oneway membranes. This also leads to the conclusion—suggested by several authors e.g., [5,12]—that the concept of entropy of spacetime horizons is intrinsically observer dependent, since a null surface (one-way membrane) may act as a horizon for a certain class of observers but not for some other class of observers. In flat Minkowski spacetime, the light cone at some event can act as a horizon for the appropriate class of uniformly accelerated Rindler observers, but not so for inertial observers. Similarly, even the black hole horizon (which can be given a ''geometrical'' definition, say, in terms of a Penrose diagram) will be viewed differently by an observer falling into the black hole compared to another who is orbiting at a radius $r > 2M$. The fact that the dynamics of the spacetime should nevertheless be described in an observer independent manner leads to very interesting consequences, as we shall see.

The above discussion also points to the possibility that the null surfaces of the spacetime should play a key role in the extremum principle we want to develop. It is also important, within this conceptual framework, that the metric is not a dynamical variable but an emergent, long-wavelength concept [13]. In this paper, we will construct such an entropy functional in a metric theory of gravity and derive the equations resulting from its extremization. We will show that not only Einstein's GR, but also the higher derivative Lanczos-Lovelock theory can be naturally incorporated in this framework. The formalism also allows us to write down higher order quantum corrections to Einstein's theory in a systematic, algorithmic procedure.

The paper is organized as follows: In Sec. II we motivate a definition for the entropy functional S related to a vector field ξ^a in the spacetime in the context of the Einstein and Lanczos-Lovelock theories, and elaborate upon the variational principle we employ to determine the spacetime dynamics from this functional. In Sec. III we compute the extremized value of $S[\xi]$ and show that under appropriate circumstances it is identical to the expression for the horizon entropy as derived by other authors in the context of Lanczos-Lovelock theories, thus justifying (at least partially) the name ''entropy functional.'' Section IV rephrases our results in the language of forms, to give its geometrical meaning. We conclude in Sec. V by discussing some implications of our results.

II. AN ENTROPY FUNCTIONAL FOR GRAVITY

Our key task is to define a suitable entropy functional for the spacetime. Since this is similar to introducing the action functional for the theory, it is obvious that we will not be able to *derive* its form without knowing the microscopic theory. So we shall do the next best thing of motivating its choice. (If the reader is unhappy with the motivating arguments, (s)he may take the final form of the entropy functional in Eq. (1) below as the basic postulate of our approach.)

The first clue comes from the theory of elasticity. We know that, in the theory of elasticity [14], the key quantity is a *displacement vector field* $\xi^a(x)$ which describes the elastic displacement of the solid through the equation $x^a \rightarrow x^a + \xi^a(x)$. (Of course, in elasticity, one usually deals with three-vectors while we need to work in D-dimensions.) All thermodynamic potentials, including the entropy of a deformable solid can be written as an integral over a quadratic functional of the displacement vector field, which can capture the relevant dynamics in the long-wavelength limit. In the context of gravity, the ''solid'' in question is spacetime itself [15]. The crucial difference from the theory of elasticity is the following: In elasticity, extremizing the entropy function will lead to an equation *for* the displacement field and determine ξ^a. In the case of spacetime, the equations should determine the *background metric*. This is a nontrivial constraint on the structure of the theory and we will show how this can be achieved.

In the case of elasticity as well as gravity, we will expect the entropy functional to be an integral over a local entropy density, so that extensivity on the volume is ensured. In the case of an elastic solid, we expect the entropy density to be translationally invariant and hence depend only upon the derivatives of ξ^a quadratically to the lowest order. We would expect this to be true for *pure* gravity as well, and hence the entropy density should have a form $P_{ab}{}^{cd}\nabla_c\xi^a\nabla_d\xi^b$, where the fourth-rank (tensorial) object $P_{ab}{}^{cd}$ is built out of metric and other geometrical quantities like curvature tensor of the background spacetime. But in the presence of nongravitational matter distribution in spacetime (which, alas, has no geometric interpretation), one cannot demand translational invariance. Hence, the entropy density can have quadratic terms in both the derivatives $\nabla_a\xi^b$ as well as ξ^a itself. We will denote the latter contribution as $T_{ab}\xi^a\xi^b$ where the second rank tensor T_{ab} (which is taken to be symmetric, since only the symmetric part is relevant to this expression) is determined by matter distribution and will vanish in the absence of matter. (We will later see that T_{ab} is just the energy-momentum tensor of matter; the notation anticipates this but does not demand it at this stage.) So our entropy functional can be written as:

$$S[\xi] = \int_{\mathcal{V}} d^Dx\sqrt{-g}(4P_{ab}{}^{cd}\nabla_c\xi^a\nabla_d\xi^b - T_{ab}\xi^a\xi^b), \quad (1)$$

where $\mathcal{V}$ is a D-dimensional region in the spacetime with boundary $\partial\mathcal{V}$, and we have introduced some additional factors and signs in the expression for later convenience. We will now impose two additional conditions on $P_{ab}{}^{cd}$ and T_{ab}. (i) For the case of the elastic solid, the coefficients of the quadratic terms are constants (related to the bulk modulus, the modulus of rigidity, and so on). We take the analogues of these constant coefficients to be quantities with vanishing covariant divergences. That is, we postulate the "constancy" conditions:

$$\nabla_b P_a{}^{bcd} = 0 = \nabla_a T^{ab}. \quad (2)$$

(ii) The second requirement we impose is that the tensor P_{abcd} should have the algebraic symmetries similar to the Riemann tensor R_{abcd} of the D-dimensional spacetime; viz., P_{abcd} is antisymmetric in ab and cd and symmetric under pair exchange. Equation (2) then implies that the P^{abcd} will be divergence free in *all* its indices. Because of these symmetries, the notation P^{ab}_{cd} with two upper and two lower indices is unambiguous.

In summary, we associate with every vector field ξ^a in the spacetime an entropy functional in Eq. (1), with the conditions: (i) The tensor $P_a{}^{bcd}$ is built from background geometrical variables, like the metric, curvature tensor, etc. and has the algebraic index symmetries of the curvature tensor. It is also divergence free. (ii) The tensor T_{ab} is related to the matter variables and vanishes in the absence of matter. It also has zero divergence. One key feature of the functional in Eq. (1) is that the entropy associated with *null* vector fields is invariant under the shift $T_{ab} \rightarrow T_{ab} + \rho g_{ab}$ where ρ is a scalar. This fact will play an interesting role later on.

A. Explicit form of P^{abcd}

Obviously, the structure of the gravitational sector is encoded in the form of P^{abcd} and we need to consider the possible choices for P^{abcd} which determine the form of the entropy functional. In a complete theory, the form of P^{abcd} will be determined by the long-wavelength limit of the microscopic theory just as the elastic constants can—in principle—be determined from the microscopic theory of the lattice. However, our situation in gravity is similar to that of the physicists of the 18th century with respect to solids and—just like them—we need to determine the "elastic constants" of spacetime by general considerations. Taking a cue from the standard approaches in renormalization group, we expect P^{abcd} to have a derivative expansion in powers of number of derivatives of the metric:

$$P^{abcd}(g_{ij}, R_{ijkl}) = c_1 \overset{(1)}{P}{}^{abcd}(g_{ij}) + c_2 \overset{(2)}{P}{}^{abcd}(g_{ij}, R_{ijkl}) + \cdots, \quad (3)$$

where $c_1, c_2, \cdots$ are coupling constants. The lowest order term must clearly depend only on the metric with no derivatives. The next term depends on the metric and curvature tensor. Note that since P^{abcd} is a tensor, its expansion in derivatives of the metric necessarily involves the curvature tensor as a "package" comprising of products of first derivatives of the metric (the $\Gamma\Gamma$ terms) and terms linear in the second derivatives ($\partial\Gamma$), where Γ symbolically denotes the Christoffel connection. Higher order terms can involve both higher powers of the curvature tensor, as well as its covariant derivatives.

These terms can, in fact, be listed from the required symmetries of P_{abcd}. For example, let us consider the possible fourth-rank tensors P^{abcd} which (i) have the symmetries of curvature tensor; (ii) are divergence free; (iii) are made from g^{ab} and $R^a{}_{bcd}$ but not derivatives of $R^a{}_{bcd}$. If we do not use the curvature tensor, then we have just one choice made from the metric:

$$\overset{(1)}{P}{}^{ab}_{cd} = \frac{1}{32\pi}(\delta^a_c\delta^b_d - \delta^a_d\delta^b_c). \quad (4)$$

We have fixed an arbitrary constant in the above expression for later convenience. Next, if we allow for P^{abcd} to depend linearly on curvature, then we have the following additional choice of tensor with the required symmetries:

$$\overset{(2)}{P}{}^{ab}_{cd} = \frac{1}{8\pi}(R^{ab}_{cd} - G^a_c\delta^b_d + G^b_c\delta^a_d + R^a_d\delta^b_c - R^b_d\delta^a_c). \quad (5)$$

We have again chosen an arbitrary constant for convenience, but in this case the constant can always be specified in the factor c_2 of Eq. (4).

The expressions in Eqs. (4) and (5) can be expressed in a more illuminating form. Note that, the expression in

Eq. (4) is just

$$\overset{(1)}{P}{}^{a_1a_2}_{b_1b_2} = \frac{1}{16\pi}\frac{1}{2}\delta^{a_1a_2}_{b_1b_2}, \tag{6}$$

where we have the introduced the alternating or "determinant" tensor $\delta^{a_1a_2}_{b_1b_2}$. Similarly, the expression in Eq. (5) above can be rewritten in the following form:

$$\overset{(2)}{P}{}^{a_1a_2}_{b_1b_2} = \frac{1}{16\pi}\frac{1}{2}\delta^{a_1a_2a_3a_4}_{b_1b_2b_3b_4}R^{b_3b_4}_{a_3a_4}, \tag{7}$$

where we have again introduced the alternating tensor $\delta^{a_1a_2a_3a_4}_{b_1b_2b_3b_4}$

$$\delta^{a_1a_2a_3a_4}_{b_1b_2b_3b_4} = \frac{-1}{(D-4)!}\epsilon^{c_1\cdots c_{D-4}a_1a_2a_3a_4}\epsilon_{c_1\cdots c_{D-4}b_1b_2b_3b_4}. \tag{8}$$

The alternating tensors are totally antisymmetric in both sets of indices and take values +1, −1, and 0. They can be written in any dimension as an appropriate contraction of the Levi-Civita tensor density with itself [16]. (In 4 dimensions the expression in Eq. (5) is essentially the double-dual of R_{abcd}.) We see a clear pattern emerging from Eqs. (6) and (7) with the mth order contribution being a term involving $(m-1)$ factors of the curvature tensor. Following this pattern it is easy to construct the mth order term which satisfies our constraints. This is unique and is given by

$$\overset{(m)}{P}{}_{ab}{}^{cd} \propto \delta^{cda_3\ldots a_{2m}}_{abb_3\ldots b_{2m}}R^{b_3b_4}_{a_3a_3}\cdots R^{b_{2m-1}b_{2m}}_{a_{2m-1}a_{2m}}. \tag{9}$$

These terms have a close relationship with the Lagrangian for Lanczos-Lovelock theory, which is a generalized higher derivative theory of gravity. Before proceeding further, we shall briefly recall the properties of Lanczos-Lovelock theory and describe this connection.

The Lanczos-Lovelock Lagrangian is a specific example from a general class of Lagrangians which describe a (possibly semiclassical) theory of gravity and are given by

$$\mathcal{L} = Q_a{}^{bcd}R^a{}_{bcd}, \tag{10}$$

where $Q_a{}^{bcd}$ is the most general fourth-rank tensor sharing the algebraic symmetries of the Riemann tensor $R^a{}_{bcd}$ and further satisfying the criterion $\nabla_b Q_a{}^{bcd} = 0$ (several general properties of this class of Lagrangians are discussed in Ref. [17]). The D-dimensional Lanczos-Lovelock Lagrangian is given by [18] a polynomial in the curvature tensor:

$$\begin{aligned}\mathcal{L}^{(D)} &= \sum_{m=1}^{K} c_m \mathcal{L}_m^{(D)};\\ \mathcal{L}_m^{(D)} &= \frac{1}{16\pi}2^{-m}\delta^{a_1a_2\ldots a_{2m}}_{b_1b_2\ldots b_{2m}}R^{b_1b_2}_{a_1a_2}R^{b_{2m-1}b_{2m}}_{a_{2m-1}a_{2m}},\end{aligned} \tag{11}$$

where the c_m are arbitrary constants and $\mathcal{L}_m^{(D)}$ is the mth order Lanczos-Lovelock term. Here the generalized alternating tensor $\delta^{\cdots}_{\cdots}$ is the natural extension of the one defined in Eq. (8) for $2m$ indices, and we assume $D \geq 2K+1$. The mth order Lanczos-Lovelock term $\mathcal{L}_m^{(D)}$ given in Eq. (11) is a homogeneous function of the Riemann tensor of degree m. For each such term, the tensor $Q_a{}^{bcd}$ defined in Eq. (10) carries a label m and becomes

$$^{(m)}Q_{ab}{}^{cd} = \frac{1}{16\pi}2^{-m}\delta^{cda_3\ldots a_{2m}}_{abb_3\ldots b_{2m}}R^{b_3b_4}_{a_3a_3}\cdots R^{b_{2m-1}b_{2m}}_{a_{2m-1}a_{2m}}. \tag{12}$$

The full tensor $Q_{ab}{}^{cd}$ is a linear combination of the $^{(m)}Q_{ab}{}^{cd}$ with the coefficients c_m. Einstein's GR is a special case of Lanczos-Lovelock gravity in which only the coefficient c_1 is nonzero. Since the tensors $^{(m)}Q_{ab}{}^{cd}$ appear linearly in the Lanczos-Lovelock Lagrangian and consequently in all other tensors constructed from it, it is sufficient to concentrate on the case where a single coefficient c_m is nonzero. All the results that follow can be easily extended to the case where more than one of the c_m are nonzero, by taking suitable linear combinations of the tensors involved. Hence, to avoid displaying cumbersome notation and summations, we will now restrict our attention to a single mth order Lanczos-Lovelock term $\mathcal{L}_m^{(D)}$, and will also drop the superscript (m) on the various quantities. Comparing with our expression in Eq. (9) it is clear that $P_a{}^{ijk}$ can be taken to be proportional to $Q_a{}^{ijk}$ which can be conveniently expressed as a derivative of the Lanczos-Lovelock Lagrangian with respect to the curvature tensor. To be concrete, we shall take the mth order term in Eq. (9) to be:

$$P_a{}^{ijk} = mQ_a{}^{ijk} = M_a{}^{ijk} \equiv \frac{\partial\mathcal{L}_m^{(D)}}{\partial R^a{}_{ijk}}. \tag{13}$$

This equation defines the divergence-free tensor $M_a{}^{ijk}$, where the partial derivatives are taken treating g^{ab}, $\Gamma^a{}_{bc}$, and $R^a{}_{bcd}$ as independent quantities. The numerical coefficients are chosen for later convenience and can be, of course, absorbed into the definitions of the c_m. With this choice, we have completely defined the geometrical structure of the entropy functional, except for the coupling constants c_m which appear at each order [19].

Just to see explicitly and in gory detail what we have, let us write down the entropy functional in the absence of matter ($T_{ab}=0$), correct up to first order in the curvature tensor in P^{abcd}. To this order, our entropy functional (1) takes the form $S = S_1 + S_2$ where

$$\begin{aligned}S_1[\xi] &= \int_{\mathcal{V}}\frac{d^Dx}{8\pi}((\nabla_c\xi^c)^2 - \nabla_a\xi^b\nabla_b\xi^a),\\ S_2[\xi] &= c_2\int_{\mathcal{V}}d^Dx(R^{cd}_{ab}\nabla_c\xi^a\nabla_d\xi^b - (G^c_a + R^c_a)\\ &\quad\times(\nabla_c\xi^a\nabla_b\xi^b - \nabla_c\xi^b\nabla_b\xi^a)),\end{aligned} \tag{14}$$

where c_2 is a coupling constant and we have used Eqs. (4) and (5). We will later see that the entropy given by S_1 leads to Einstein's equations in general relativity while S_2 and higher order terms can be interpreted as corrections to this. If we choose ξ^a to be the normal vector field of a sequence of hypersurfaces foliating the spacetime, then the integrand in S_1 has the familiar structure, $(\mathrm{Tr}K)^2 - \mathrm{Tr}(K^2)$ where K_{ab} is the extrinsic curvature. (This could offer an alternative interpretation of Arnowitt-Deser-Misner (ADM) action; however, we will not discuss this aspect here except to cast our results in the familiar language when appropriate.) The expression in Eq. (14) can be further simplified by integrating it by parts where appropriate and writing the right-hand side as a sum of a contribution from the bulk and a surface term. The general expression after such a splitting is given later in Eq. (29) and can also be found in Sec. IV where it arises transparently in the language of forms (see Eq. (53)).

B. Field equations from extremizing the entropy

Having made these general observations regarding the choice of P^{ab}_{cd} let us now return to the entropy functional in Eq. (1). This expression is well defined for *any* displacement vector field ξ^a. We can, therefore, associate an entropy functional with any hypersurface in the spacetime, by choosing the normal to the hypersurface as ξ^a. Among all such hypersurfaces, the null hypersurfaces will play a key role since they act as oneway membranes which block information for a specific class of observers. Given this motivation, we will now extremize this S with respect to variations of the null vector field ξ^a and demand that the resulting condition holds for *all null vector fields*. That is, the equilibrium configurations of the "spacetime solid" are the ones in which the entropy associated with *every* null vector is extremized. Varying the null vector field ξ^a after adding a Lagrange multiplier λ for imposing the null condition $\xi_a \delta\xi^a = 0$, we find:

$$\begin{aligned}\delta S &= 2\int_{\mathcal{V}} d^D x\sqrt{-g}(4P_{ab}{}^{cd}\nabla_c\xi^a(\nabla_d\delta\xi^b) - T_{ab}\xi^a\delta\xi^b \\ &\quad - \lambda g_{ab}\xi^a\delta\xi^b) \\ &\equiv 2\int_{\mathcal{V}} d^D x\sqrt{-g}(4P_{ab}{}^{cd}\nabla_c\xi^a(\nabla_d\delta\xi^b) - \bar{T}_{ab}\xi^a\delta\xi^b),\end{aligned} \tag{15}$$

where we have used the symmetries of $P_{ab}{}^{cd}$ and T_{ab} and set $\bar{T}_{ab} = T_{ab} + \lambda g_{ab}$. (As we said before such a shift leaves entropy associated with null vectors unchanged so the Lagrange multiplier will turn out to be irrelevant; nevertheless, we will use $\bar{T}_{ab}$ for the moment.) An integration by parts and the condition $\nabla_d P_{ab}{}^{cd} = 0$, leads to

$$\begin{aligned}\delta S &= 2\int_{\mathcal{V}} d^D x\sqrt{-g}(-4P_{ab}{}^{cd}(\nabla_d\nabla_c\xi^a) - \bar{T}_{ab}\xi^a)\delta\xi^b \\ &\quad + 8\int_{\partial\mathcal{V}} d^{D-1}x\sqrt{h}(n_d P_{ab}{}^{cd}(\nabla_c\xi^a))\delta\xi^b,\end{aligned} \tag{16}$$

where n^a is the D-vector field normal to the boundary $\partial\mathcal{V}$ and h is the determinant of the intrinsic metric on $\partial\mathcal{V}$. As usual, in order for the variational principle to be well defined, we require that the variation $\delta\xi^a$ of the null vector field should vanish on the boundary. The second term in Eq. (16) therefore vanishes, and the condition that $S[\xi]$ be an extremum for arbitrary variations of ξ^a then becomes

$$2P_{ab}{}^{cd}(\nabla_c\nabla_d - \nabla_d\nabla_c)\xi^a - \bar{T}_{ab}\xi^a = 0, \tag{17}$$

where we used the antisymmetry of $P_{ab}{}^{cd}$ in its upper two indices to write the first term. The definition of the Riemann tensor in terms of the commutator of covariant derivatives reduces the above expression to

$$(2P_b{}^{ijk}R^a{}_{ijk} - \bar{T}^a_b)\xi_a = 0, \tag{18}$$

and we see that the equations of motion *do not contain* derivatives with respect to ξ. This peculiar feature arose because of the symmetry requirements we imposed on the tensor $P_{ab}{}^{cd}$. We further require that the condition in Eq. (18) hold for *arbitrary* null vector fields ξ^a. A simple argument based on local Lorentz invariance then implies that

$$2P_b{}^{ijk}R^a{}_{ijk} - T^a_b = F(g)\delta^a_b, \tag{19}$$

where $F(g)$ is some scalar functional of the metric and we have absorbed the $\lambda\delta^a_b$ in $\bar{T}^a_b = T^a_b + \lambda\delta^a_b$ into the definition of F. The validity of the result in Eq. (19) is obvious if we take a dot product of Eq. (18) with ξ^b. (A formal proof can be found in Appendix A 1.) The scalar $F(g)$ is arbitrary so far and we will now show how it can be determined in the physically interesting cases.

1. Lowest order theory: Einstein's equations

To do this, let us substitute the derivative expansion for P^{abcd} in Eq. (3) into Eq. (19). To the lowest order we find that the equation reduces to:

$$\frac{1}{8\pi}R^a_b - T^a_b = F(g)\delta^a_b, \tag{20}$$

where F is an arbitrary function of the metric. Writing this equation as $(G^a_b - 8\pi T^a_b) = Q(g)\delta^a_b$ with $Q = 8\pi F - (1/2)R$ and using $\nabla_a G^a_b = 0$, $\nabla_a T^a_b = 0$ we get $\partial_b Q = \partial_b[8\pi F - (1/2)R] = 0$; so that Q is an undetermined integration constant, say Λ, and F must have the form $8\pi F = (1/2)R + \Lambda$. The resulting equation is

$$R^a_b - (1/2)R\delta^a_b = 8\pi T^a_b + \Lambda\delta^a_b \tag{21}$$

which leads to Einstein's theory if we identify T_{ab} as the matter energy-momentum tensor *with a cosmological constant appearing as an integration constant*. (For the importance of the latter with respect to the cosmological constant problem, see Ref. [20]; we will not discuss this issue here.)

The same procedure works with the first order term in Eq. (3) as well and we reproduce the Gauss-Bonnet gravity with a cosmological constant. In this sense, we can interpret the first term in the entropy functional in Eq. (14) as the entropy in Einstein's general relativity and the term proportional to c_2 as a Gauss-Bonnet correction term. Instead of carrying out this analysis explicitly order by order, we shall now describe the most general structure in the family of theories starting with Einstein's GR, Gauss-Bonnet gravity, etc.—and will show that we reproduce the Lanczos-Lovelock theory by our approach.

2. Higher order corrections: Lanczos-Lovelock gravity

To see this result, let us briefly recall some aspects of Lanczos-Lovelock theory. It can be shown that (see e.g., [17]) the equations of motion for a general theory of gravity derived from the Lagrangian in Eq. (10) using the standard variational principle with g^{ab} as the dynamical variables, are given by

$$E_{ab} = \frac{1}{2}T_{ab}; \qquad E_{ab} \equiv \frac{1}{\sqrt{-g}}\frac{\partial}{\partial g^{ab}}(\sqrt{-g}\mathcal{L}) - 2\nabla^m\nabla^n M_{amnb}. \tag{22}$$

Here T_{ab} is the energy-momentum tensor for the matter fields. The tensor M_{abcd} defined through $M_a{}^{bcd} \equiv (\partial\mathcal{L}/\partial R^a{}_{bcd})$ is a generalization of the one defined for the Lanczos-Lovelock case in Eq. (13). The partial derivatives are as before taken treating g^{ab}, $\Gamma^a{}_{bc}$, and $R^a{}_{bcd}$ as independent quantities. For the mth order Lanczos-Lovelock Lagrangian $\mathcal{L}_m^{(D)}$, since M^{abcd} is divergence free, the expression for the tensor E_{ab} in Eq. (22) becomes

$$E_{ab} = \frac{\partial\mathcal{L}_m^{(D)}}{\partial g^{ab}} - \frac{1}{2}\mathcal{L}_m^{(D)} g_{ab}, \tag{23}$$

where we have used the relation $\partial(\sqrt{-g})/\partial g^{ab} = -(1/2)\sqrt{-g}g_{ab}$. The first term in the expression for E_{ab} in Eq. (23) can be simplified to give

$$\frac{\partial\mathcal{L}_m^{(D)}}{\partial g^{ab}} = mQ_a{}^{ijk}R_{bijk} = M_a{}^{ijk}R_{bijk}, \tag{24}$$

where the expressions in Eq. (24) can be verified by direct computation, or by noting that $\mathcal{L}_m^{(D)}$ is a homogeneous function of the Riemann tensor $R^a{}_{bcd}$ of degree m. To summarize, the Lanczos-Lovelock field equations are given by

$$E_{ab} = \tfrac{1}{2}T_{ab}; \qquad E_{ab} = mQ_a{}^{ijk}R_{bijk} - \tfrac{1}{2}\mathcal{L}_m^{(D)} g_{ab}. \tag{25}$$

Further, diffeomorphism invariance implies that the tensor E_{ab} defined in Eq. (22) is divergence free, $\nabla_a E^a_b = 0$. The equations of motion for the matter imply that the energy-momentum tensor T_{ab} is also divergence free (as required by Eq. (2)). Using these conditions in Eq. (19) together with the choice in Eq. (13) for $P_a{}^{ijk}$ leads to

$$\partial_a F = \partial_a \mathcal{L}_m^{(D)}, \tag{26}$$

which fixes $F(g)$ as $F = \mathcal{L}_m^{(D)} + \Lambda/8\pi$ where Λ is a constant with the normalization chosen so as to conform with the usual definition of the cosmological constant. The resulting field equations for Lanczos-Lovelock gravity will be:

$$16\pi[P_b{}^{ijk}R^a{}_{ijk} - \tfrac{1}{2}\delta^a_b\mathcal{L}_m^{(D)}] = 8\pi T^a_b, \tag{27}$$

where we have included a possible cosmological constant, that arises as an undetermined integration constant in the definition of T^a_b. Taking the trace of this equation, we find that $\mathcal{L}_m^{(D)} = (2m - D)^{-1}T$. In other words, the on shell value of the Lagrangian is proportional to the trace of the stress tensor in all Lanczos-Lovelock theories, just like in GR. In the absence of the source term, this implies that $\mathcal{L}_m^{(D)} = 0$ and the equation of motion reduces to $P_b{}^{ijk}R^a{}_{ijk} = 0$. The case $m = 1$ with $\mathcal{L}_{m=1}^{(D)} = (1/16\pi)R$ is easily seen to reduce to that of Einstein's gravity.

To summarize, if we take the derivative expansion in Eq. (3) to correspond to a polynomial form in the curvature tensor, then it has the form given by Eq. (13). In this case, extremizing the entropy leads to the Lanczos-Lovelock theory. We stress that the resulting field equations have the form of Einstein's equations with higher order corrections. In our picture, we consider this as emerging from the form of the entropy functional which has an expansion in powers of the curvature.

Before concluding this section, we want to comment on an interesting property of the entropy functional. The derivation of Eq. (19) was based upon a variational principle which closely resembles the usual variational principle used in other areas of physics in which some quantity is varied within an integral arbitrarily, except for it being fixed at the boundary. Instead of such an arbitrary variation, let us consider a *subset* of all possible variations of the null vector field ξ^a, given by $\xi^a(x) \to (1 + \epsilon(x))\xi^a(x)$; namely, infinitesimal *rescalings* of ξ^a. We assume that the scalar $\epsilon(x)$ is infinitesimal and also that it vanishes on the boundary $\partial\mathcal{V}$. In this case it is easy to see that the variation of $S[\xi]$ in Eq. (16) becomes

$$\delta S|_{\text{rescale}} = 2\int_{\mathcal{V}} d^Dx\sqrt{-g}(2P_b{}^{ijk}R^a{}_{ijk} - T^a_b)\xi_a\xi^b\epsilon(x). \tag{28}$$

Clearly, requiring that the functional $S[\xi]$ be *invariant* under rescaling transformations of ξ^a leads to the same requirement as before, namely, that Eq. (19) be satisfied. We can understand the physical motivation behind imposing such a symmetry condition on $S[\xi]$ as follows. Let us begin by noting the fact that the causal structure of a spacetime, which can be thought of as the totality of all possible families of null hypersurfaces in the spacetime, is left invariant under rescalings of the generators of these null hypersurfaces. To see this symmetry, note that *any* curve in the spacetime, with tangent vector field $t^a = dx^a/d\lambda$, say, is invariant under the rescaling $t^a \rightarrow f(\lambda)t^a$ of the tangent field, where f is some scalar. This is so because a rescaling of the tangent vector field is equivalent to a reparametrization of the curve. A null hypersurface $\mathcal{H}$ can be thought of as being "filled" by null geodesics contained in it, and is hence invariant under rescalings of its generator field ξ^a. The result for the full causal structure then follows. The functional $S[\xi]$ depends only on the generator ξ^a of some null hypersurface (apart from the metric and matter fields which we consider as given quantities). It is therefore natural to demand that this symmetry of the causal structure also be a symmetry of $S[\xi]$. We do not, however, use this feature in this paper. (For a completely different, purely classical, approach to general relativity based on null surfaces, see [21]).

III. EVALUATING THE ENTROPY FUNCTIONAL ON SHELL

The results in the previous section show that our approach provides—at the least—an alternative variational principle to obtain not only Einstein's theory but also Lanczos-Lovelock theory. In the conventional approaches to these theories, one can obtain the field equations by varying the action functional. But once the field equations are obtained, the extremum *value* of the action functional is not of much concern. (The only exceptions are in the semiclassical limit in which it appears as the phase of the wave function or in some specific Euclidean extension of the solution.) In our approach, it is worthwhile to proceed further and ask what this extremal ["on shell"] value means in specific contexts. We will be able to provide an interpretation under some specific contexts, justifying the term entropy functional but it should be stressed that the results in this section are logically independent of the derivation of field equations in the previous section. In particular, we need to consider non-null vectors to provide a natural interpretation of the extremum value of the functional.

The term "on shell" refers to satisfying the relevant equations of motion, which in this case are given by Eq. (18). Manipulating the covariant derivatives in Eq. (1), we can write

$$\begin{aligned} S[\xi] &= \int_{\mathcal{V}} d^D x\sqrt{-g}[4\nabla_d(P_{ab}{}^{cd}(\nabla_c\xi^a)\xi^b) \\ &\quad - 4P_{ab}{}^{cd}(\nabla_d\nabla_c\xi^a)\xi^b - T_{ab}\xi^a\xi^b] \\ &= 4\int_{\partial\mathcal{V}} d^{D-1}x\sqrt{h}n_d(P_{ab}{}^{cd}\xi^b\nabla_c\xi^a) \\ &\quad + \int_{\mathcal{V}} d^D x\sqrt{-g}(2P_{mb}{}^{cd}R^m{}_{acd} - T_{ab})\xi^a\xi^b. \end{aligned} \tag{29}$$

In writing the first equality, we have used the condition $\nabla_d P_a{}^{bcd} = 0$. As before, in the first term of second equality, n^a is the vector field normal to the boundary $\partial\mathcal{V}$ and h is the determinant of the intrinsic metric on $\partial\mathcal{V}$. (In general, the boundary is $(D-1)$ dimensional. We will soon see that the really interesting case occurs, in fact, when part of the boundary $\partial\mathcal{V}$ is null and hence intrinsically $(D-2)$ dimensional. This case needs to be handled by a limiting procedure and in what follows we will elaborate on the procedure we use.) The second term of the second equality in Eq. (29) vanishes in the absence of matter because, when $T_{ab} = 0$, the equation of motion reduces to $Q_a{}^{ijk}R_{bijk} = 0$, thereby allowing us to interpret the first term as the on shell value of entropy from the gravity sector. Even in the presence of matter, the second term can be expressed in terms of matter variables as an integral over the trace of the stress tensor (see the discussion around Eq. (27)) and, of course, is not a surface term. We will, therefore, concentrate on the surface term arising from the gravitational sector which reduces to

$$\begin{aligned} S|_{\text{on shell}} &= 4\int_{\partial\mathcal{V}} d^{D-1}x\sqrt{h}n_a(P^{abcd}\xi_c\nabla_b\xi_d) \\ &\rightarrow \frac{1}{8\pi}\int_{\partial\mathcal{V}} d^{D-1}x\sqrt{h}n_a(\xi^a\nabla_b\xi^b - \xi^b\nabla_b\xi^a), \end{aligned} \tag{30}$$

where we have manipulated a few indices using the symmetries of P^{abcd}. The second expression after the arrow is the result for general relativity; we give this explicitly to show the form of the expression in a familiar setting. Note that, when ξ^a is chosen as the normal to a set of surfaces foliating the spacetime, the integrand has the familiar structure of $n_i(\xi^i K + a^i)$ where $a^i = \xi^b\nabla_b\xi^i$ is the acceleration associated with the vector field ξ^a and $K \equiv -\nabla_b\xi^b$ is the trace of extrinsic curvature in the standard context. This is the standard surface term which arises in the ADM formulation and the cognoscenti will immediately see its connection with entropy of horizons in GR.

At this stage, we have not put any restriction of the boundary $\partial\mathcal{V}$ or on the choice of the vector field ξ^a. The expression in Eq. (30) is valid for *any* vector field ξ^a—not necessarily null. (Our entropy functional is defined for any vector field; to obtain the equation of motion we consider only the variation of null vector fields but having done that, we can study the on shell entropy for any vector field.) The only restriction is that the expression in Eq. (30) should be evaluated on a solution to the field equations. It is clear that one cannot say much about the value of this expression in such a general context, keeping the boundary and ξ^a totally arbitrary. Further, even in the case of a null vector field ξ^a, the integrand $\mathcal{I}$ in Eq. (30) changes by $\mathcal{I} \rightarrow f^2(x)\mathcal{I}$, under a rescaling $\xi^a \rightarrow f(x)\xi^a$ which keeps the null vector as null. Since the value of the integral can be changed even by such a rescaling, it is clear that a choice has to be made for the overall scaling of the null vector field before we can evaluate $S[\xi]$ on shell [22].

The fact that Eq. (30) has no clear interpretation in general should not be surprising since we expect to obtain a nontrivial value for the entropy only in specific cases in which the solution has a definite thermodynamic interpretation and the surface and the vector field is chosen appropriately. Obviously, making this connection will require choosing a particular solution to the field equations, a particular domain of integration for the entropy functional, etc. and making other specific assumptions. We shall now calculate the extremum value in specific situations and demonstrate that it gives the standard result for the gravitational entropy when the latter is well defined and understood. We will also discuss several features of this issue and, in particular, will demonstrate that in the standard cases with horizons, the extremal value of the entropy correctly reproduces the known results, *not only in GR but even for Lanczos-Lovelock theory.*

The most important case corresponds to solutions with a *stationary* horizon which can be locally approximated as Rindler spacetime. Many of the results which had motivated us to develop the current formalism were proved in this specific context and hence this will act as a natural testing ground. In this case, the relevant part of the boundary will be a null surface and we will choose ξ^a to be a spacelike vector which can be interpreted as describing the displacement of the horizon normal to itself. To define this properly, we will use a limiting procedure and provide the physical motivation for the choice of ξ^a (based on certain locally accelerated observers) thereby leading to a meaningful interpretation of the on shell value of $S[\xi]$.

To set the stage for calculations that follow, we will begin by briefly recalling the notion of *Rindler* observers in flat (Minkowski) spacetime. In Minkowski spacetime with inertial coordinates $x^i_{\rm M} = (t_{\rm M}, x^\alpha_{\rm M})$, $\alpha = 1, 2, \ldots, D-1$, observers undergoing constant acceleration along the $x^1_{\rm M}$ direction (the Rindler observers) follow hyperbolic trajectories [5,16] described by $(x^1_{\rm M})^2 - (t_{\rm M})^2 = \text{constant}$. A natural set of coordinates for these observers is given by $x^a_{\rm R} = (t_{\rm R}, N, x^A_\perp)$, $A = 2, 3, \ldots, D-1$, where the transformation between $x^a_{\rm R}$ and $x^a_{\rm M}$ is given by

$$t_{\rm M} = \frac{N}{\kappa}\sinh(\kappa t_{\rm R}); \qquad x^1_{\rm M} = \frac{N}{\kappa}\cosh(\kappa t_{\rm R}),$$
$$N = \kappa((x^1_{\rm M})^2 - (t_{\rm M})^2)^{1/2}; \qquad t_{\rm R} = \frac{1}{\kappa}\tanh^{-1}\left(\frac{t_{\rm M}}{x^1_{\rm M}}\right), \tag{31}$$
$$x^A_{\rm M} = x^A_\perp, \qquad A = 2, 3, \ldots, D-1,$$

with constant κ. The metric in the Rindler coordinates becomes

$$ds^2 = -N^2 dt^2_{\rm R} + dN^2/\kappa^2 + dL^2_\perp, \tag{32}$$

where $dL^2_\perp$ is the (flat) metric in the transverse spatial directions. It is easy to see that the surface described by $(x^1_{\rm M})^2 - (t_{\rm M})^2 = 0$ (or $N = 0$ in the Rindler coordinates) is simply the null light cone in the $t_{\rm M}$-$x^1_{\rm M}$ plane at the origin, and that it acts as a horizon for the observers maintaining $N = \text{constant} \neq 0$.

In a general curved spacetime, one can introduce a notion of local Rindler frames along similar lines. We first go to the local inertial frame (LIF, hereafter) around any event $\mathcal{P}$ and introduce the LIF coordinates $x^i_{\rm M} = (t_{\rm M}, x^\alpha_{\rm M})$, $\alpha = 1, 2, \ldots, D-1$. We then use the transformations in Eq. (31) to define a local Rindler frame (LRF, hereafter). The choice of $x^1_{\rm M}$ axis is of course arbitrary and one could have chosen any direction in the LIF as the $x^1_{\rm M}$ axis by a simple rotation. In particular, a general null surface $\mathcal{H}$ in the original spacetime passing through $\mathcal{P}$ can be locally mapped to the null cone in LIF which—in turn—can be locally identified with the $N = 0$ surface for the *local* Rindler frame. This local patch $\mathcal{H}_{\rm LIF} \subset \mathcal{H}$ of the original null surface acts as a horizon for these observers. We will make good use of this observation below. (The local nature of the construction is more transparent in the Euclidean description. If we choose a LIF around any event and then transform to a LRF, then the null surface in the Minkowski coordinates gets mapped to the origin of the Euclidean coordinates. Our constructions in a local region around the origin in the Euclidean sector capture the physics near the Rindler horizon in the Minkowski frame.) The crucial fact to notice is that locally, $\mathcal{H}$ is the *Killing* horizon for a suitable class of Rindler observers. To see this, choose some point $\mathcal{P} \in \mathcal{H}$ and erect a D-ad (the D-dimensional generalization of a tetrad) in the LIF at $\mathcal{P}$, endowed with Minkowski coordinates. Let $\mathcal{H}_{\rm LIF} \subset \mathcal{H}$ denote the part of $\mathcal{H}$ contained in the LIF. Choose the D-ad in such a way that the only nonvanishing components of the generator χ^a of $\mathcal{H}_{\rm LIF}$ are χ^0 and χ^1. In other words, with respect to this D-ad, $\mathcal{H}_{\rm LIF}$ is defined by $(x^1_{\rm M})^2 - (x^0_{\rm M})^2 = 0$, where $x^i_{\rm M}$ are the Minkowski coordinates in the LIF. Now transform to the local Rindler frame using the transformation in Eq. (31) and consider the vector $v^a = (1, \vec{0})$ in the Rindler frame. Clearly v^a is the Killing vector associated with time translations in the Rindler frame, with norm $v^a v_a = -N^2$, and hence $\mathcal{H}_{\rm LIF}$ (given by $N = 0$) is a Killing horizon for the Rindler observers, generated by v^a. (It can be shown that the original generator χ^a of $\mathcal{H}_{\rm LIF}$ when transformed to the Rindler frame, is proportional to the Rindler Killing vector v^a *on* the horizon $\mathcal{H}_{\rm LIF}$.)

We will now give a prescription for the evaluation of $S|_{\text{on shell}}$ in a specific LIF (i.e. on $\mathcal{H}_{\rm LIF} \subset \mathcal{H}$), which extends to the entire surface $\mathcal{H}$ in an obvious way. For notational convenience therefore, we will drop the subscript "LIF" on $\mathcal{H}$. Instead of the surface $\mathcal{H}$, consider the surfaces in the local Rindler frame at $\mathcal{P}$ given by $N = \epsilon = \text{constant}$. Take $\xi^a = n^a$ as the unit *spacelike normal* to these surfaces, so that

$$n^a = \xi^a = (0, \kappa, 0, 0, \ldots); \qquad \sqrt{h} = \epsilon\sqrt{\sigma}, \tag{33}$$

where σ is the metric determinant on the $t_{\rm R} = \text{constant}$, $N = \text{constant}$ surfaces. We will evaluate the surface integral for $S|_{\text{on shell}}$ on a surface with $N = \epsilon = \text{constant}$, and take the limit $\epsilon \to 0$ at the end of the calculation. The vector displacement field ξ^a then has the natural interpretation of moving the surface $N = \epsilon =$ constant normal to itself and previous work has shown that these displacements play a crucial role in the thermodynamic interpretation [7,8]. In our limiting procedure, we use the normal vector to the surface $N = \epsilon =$ constant, fix its norm when the surface is not null (i.e., $\epsilon \neq 0$), and take the $\epsilon \to 0$ limit right at the end so that it remains properly normalized. In this process, we are considering the null surface as a limit of a sequence of timelike surfaces. This is clearly only an ansatz and—as we have said before—the final result will be different for a different ansatz. In this sense, it is the end result that we obtain which provides further justification for this choice. This prescription can be understood in a natural way in the Euclidean continuation ($t_R \to it_R$) of the Rindler frame, where the $N = \epsilon$ surfaces are circles of radius ϵ/κ in the t-x plane of the Euclideanized Minkowski coordinates.

Computing the entropy functional using this vector field, and taking the $\epsilon \to 0$ limit at the very end, we find that

$$S|_{\mathcal{H}} = \sum_{m=1}^{K} 4\pi m c_m \int_{\mathcal{H}} d^{D-2}x_\perp \sqrt{\sigma} \mathcal{L}_{(m-1)}^{(D-2)}, \tag{34}$$

where $x_\perp$ denotes the transverse coordinates on $\mathcal{H}$, σ is the determinant of the intrinsic metric on $\mathcal{H}$, and we have restored a summation over m thereby giving the result for the most general Lanczos-Lovelock case. The proof of Eq. (34) can be found in Appendix A 2. The expression in Eq. (34) *is precisely the entropy of a general Killing horizon in Lanczos-Lovelock gravity* based on the general prescription given by Wald and others [23] and computed by several authors [24]. This result justifies the choice of vector field ξ^a used to compute the entropy functional (as well as the nomenclature "entropy functional" itself). For a wide class of Killing horizons ($\mathcal{H}_K$) it is possible to take the *Rindler limit* of the near-horizon geometry, and write $ds^2 = -N^2dt^2 + dN^2/\kappa^2 + dL_\perp^2$ near the horizon, where $dL_\perp^2$ denotes the line element on the transverse surfaces (and, in particular, on the $N = 0$ surface which is the horizon; around any point $\mathcal{P} \in \mathcal{H}_K$ the transverse directions will be locally flat in the Rindler limit). In this case, κ is the surface gravity of the horizon (being constant over the entire horizon) and we choose ξ^a by the above limiting prescription. Our construction is then valid over the entire surface $\mathcal{H}_K$ and the resulting on shell value of the entropy functional is precisely the standard entropy of the horizon.

To summarize, we get meaningful results in two cases of importance. First, whenever we have a solution to the field equations which possesses a stationary horizon with Rindler limit, we have a natural choice for ξ^a through a limiting procedure such that the extremum value of the entropy functional on shell matches with the standard result for the entropy of the horizon. Second, in any spacetime, if we take a local Rindler frame around any event we will obtain an entropy for the locally defined Rindler horizon. In the case of GR, this entropy per unit transverse area is just 1/4 as expected. This requires working in a local patch and accepting the notions of local Rindler observers and local Rindler horizons about which there is still no universal agreement. (Not everyone is comfortable with the de Sitter universe having an observer dependent entropy let alone Rindler horizon, but we believe this is the correct paradigm.) Finally, in the situation in which the boundary $\partial\mathcal{V}$ is nowhere null and ξ^a is an arbitrary vector field—as we pointed out before—it is hard to say anything about the value of S. We hope to investigate this case fully in a later work.

IV. THE RESULTS IN THE LANGUAGE OF FORMS

The purpose of this section is to recast our formalism in the language of forms for the sake of those who find such things attractive. This will, hopefully, help in further work because of two reasons. First, the expression for Wald entropy [23] can be expressed in the language of forms nicely. Second, the action for Lanczos-Lovelock gravity can be expressed in terms of the wedge product of curvature forms. We shall begin with brief pedagogy to set the stage and notation and then will derive the key results.

Since the tensor $P^{ab}{}_{cd}$ has the same algebraic structure as curvature tensor, one can express it in terms of a 2-form analogous to the curvature 2-form. We define a 2-form $\mathbf{P}^{ab}$ related to our tensor $P^{ab}{}_{cd}$ by:

$$\mathbf{P}^{ab} \equiv \frac{1}{2!} P^{ab}{}_{cd} \boldsymbol{\omega}^c \wedge \boldsymbol{\omega}^d. \tag{35}$$

Throughout this section we will assume a coordinate 1-form basis $\boldsymbol{\omega}^i = \mathbf{d}x^i$. If $\mathbf{v} = \mathbf{e}_a v^a$ is a vector, then the vector-valued 1-form $\mathbf{dv}$ is given by

$$\mathbf{dv} = \mathbf{e}_a(\mathbf{d}v^a + \boldsymbol{\omega}^a{}_b v^b) = \mathbf{e}_a(\nabla_b v^a)\boldsymbol{\omega}^b; \qquad \boldsymbol{\omega}^a{}_b = \Gamma^a_{bc}\boldsymbol{\omega}^c. \tag{36}$$

We will work in the case of pure gravity (that is, $T_{ab} = 0$) since this is more geometrical and since it does not affect the value of the final on shell entropy functional. Our entropy functional in Eq. (1) has the integrand (D-form):

$$\mathcal{I} = (4P_{ab}{}^{cd}\nabla_c\xi^a\nabla_d\xi^b)\boldsymbol{\epsilon}, \tag{37}$$

where $\boldsymbol{\epsilon} = (1/D!)\epsilon_{a_1\cdots a_D}\boldsymbol{\omega}^{a_1} \wedge \cdots \wedge \boldsymbol{\omega}^{a_D}$ is the natural D-form on the integration domain $\mathcal{V}$. The first point to note is that the D-form in Eq. (37) is the same as the following:

$$\mathcal{I} = 4(*\mathbf{P}_{ab} \wedge (\mathbf{d}\boldsymbol{\xi})^a \wedge (\mathbf{d}\boldsymbol{\xi})^b), \tag{38}$$

where $(\mathbf{d}\boldsymbol{\xi})^a = (\nabla_c\xi^a)\boldsymbol{\omega}^c$, etc. To see this, we use

$$*\,\mathbf{P}_{ab} = \frac{1}{(D-2)!}\frac{1}{2!}P_{ab}{}^{cd}\epsilon_{cda_1\cdots a_{D-2}}\boldsymbol{\omega}^{a_1} \wedge \cdots \wedge \boldsymbol{\omega}^{a_{D-2}}, \tag{39}$$

which allows us to expand the D-form in Eq. (38) as

$$\mathcal{I} = \frac{2}{(D-2)!} P_{ab}{}^{cd} \epsilon_{cda_1\cdots a_{D-2}} (\nabla_{a_{D-1}} \xi^a) \times (\nabla_{a_D} \xi^b) \boldsymbol{\omega}^{a_1} \wedge \cdots \wedge \boldsymbol{\omega}^{a_D}. \quad (40)$$

Since this is a D-form in a D-dimensional space, the right-hand side of Eq. (40) should be expressible as $f\boldsymbol{\epsilon}$ where f is a scalar. This implies

$$\frac{2}{(D-2)!} P_{ab}{}^{cd} \epsilon_{cda_1\cdots a_{D-2}} (\nabla_{a_{D-1}} \xi^a)(\nabla_{a_D} \xi^b) = \frac{1}{D!} f \epsilon_{a_1\cdots a_D}. \quad (41)$$

We now use the identity

$$\epsilon^{b_1\cdots b_{D-j} a_1 \cdots a_j} \epsilon_{b_1\cdots b_{D-j} c_1 \cdots c_j} = (-1)^s (D-j)! \delta^{a_1 a_2 \ldots a_j}_{c_1 c_2 \ldots c_j}, \quad (42)$$

where s is the number of minus signs in the metric and $\delta^{\cdots}_{\cdots}$ is the alternating tensor with our normalization. This also implies

$$\epsilon^{a_1\cdots a_D} \epsilon_{a_1\cdots a_D} = (-1)^s D!. \quad (43)$$

Contracting both sides of Eq. (41) with $\epsilon^{a_1\cdots a_D}$ and using the symmetries of $\epsilon_{a_1\cdots a_D}$ leads to

$$f = 2P_{ab}{}^{cd} (\nabla_i \xi^a)(\nabla_j \xi^b) \delta^{ij}_{cd} = 4P_{ab}{}^{cd} (\nabla_c \xi^a)(\nabla_d \xi^b), \quad (44)$$

which is exactly what is required in Eq. (37). Therefore the entropy functional for any vector field can be written, somewhat more geometrically, as

$$S[\xi] = \int_{\mathcal{V}} 4(*\mathbf{P}_{ab} \wedge (\mathbf{d}\boldsymbol{\xi})^a \wedge (\mathbf{d}\boldsymbol{\xi})^b). \quad (45)$$

We will now do two things. The first is to derive the equations of motion and hence the on shell value of $S[\xi]$, and the second is to relate this value to the Wald entropy. To derive the equations of motion, it is convenient to introduce some new notation which makes the derivation more compact. (Essentially, we will suppress all explicit index occurrences in, say Eq. (45), etc.). The relation to Wald entropy, however, is more easily seen *with* the indices explicitly in place, so we will revert to the explicit notation at that stage.

To get rid of tagging along the indices, we will first introduce the convention that in a tensor valued p-form $\mathbf{T}$, the tensor indices will always be thought of as being superscripts. For example if $\mathbf{T}$ is a 2-tensor valued p-form, we have $\mathbf{T} = \mathbf{e}_a \mathbf{e}_b \mathbf{T}^{ab}$ where $\mathbf{T}^{ab}$ is a p-form and we have suppressed the direct product sign for the basis vectors. We come across p-forms like $\mathbf{P}^{ab}$ which are antisymmetric in a and b. These can be denoted by the "bivector valued" p-form $\mathbf{P} = (1/2!)\mathbf{e}_a \wedge \mathbf{e}_b \mathbf{P}^{ab}$. This will allow us to consistently introduce a dot product for the tensor valued forms. Let us now introduce the notation "wedge-dot" $\dot{\wedge}$ as follows in terms of a couple of examples: (1) If $\mathbf{A}$ is a 2-tensor valued p-form and $\mathbf{B}$ is a vector valued q-form then

$$\mathbf{A} \dot{\wedge} \mathbf{B} \equiv \mathbf{e}_a (\mathbf{e}_b \cdot \mathbf{e}_c) \mathbf{A}^{ab} \wedge \mathbf{B}^c = \mathbf{e}_a \mathbf{A}^a{}_b \wedge \mathbf{B}^b. \quad (46)$$

That is, the wedge in $\dot{\wedge}$ acts in the usual fashion on the p-forms and the dot acts on the two nearest basis vectors it finds. This last condition also implies that: (2) If $\mathbf{A}$ is a bivector valued p-form ($\mathbf{A}^{ab} = -\mathbf{A}^{ba}$) and $\mathbf{B}$ is a vector valued q-form, then

$$\mathbf{A} \dot{\wedge} \mathbf{B} = (-1)^{pq+1} \mathbf{B} \dot{\wedge} \mathbf{A}, \quad (47)$$

where the $(-1)^{pq}$ is the usual factor on exchange of the wedge product, and the additional factor of (-1) arises due to the dot in $\dot{\wedge}$ shifting from one index of $\mathbf{A}$ to the other.

In this notation, we have

$$*\mathbf{P}_{ab} \wedge (\mathbf{d}\boldsymbol{\xi})^a \wedge (\mathbf{d}\boldsymbol{\xi})^b = (-1)^{D-2} (\mathbf{d}\boldsymbol{\xi})^a \wedge *\mathbf{P}_{ab} \wedge (\mathbf{d}\boldsymbol{\xi})^b = (-1)^{D-2} \mathbf{d}\boldsymbol{\xi} \dot{\wedge} * \mathbf{P} \dot{\wedge} \mathbf{d}\boldsymbol{\xi}. \quad (48)$$

Further, in this notation one can show that for the bivector valued 2-form $\mathbf{P}$, the following relation holds

$$* \mathbf{d} * \mathbf{P} = \frac{1}{2!} \mathbf{e}_a \wedge \mathbf{e}_b (\nabla_c P^{abc}{}_d) \boldsymbol{\omega}^d, \quad (49)$$

and the condition $\nabla_c P^{abcd} = 0$ is equivalent to $\mathbf{d} * \mathbf{P} = 0$. This also means $\mathbf{d}(*\mathbf{P} \cdot \boldsymbol{\xi}) = (-1)^{D-2} * \mathbf{P} \dot{\wedge} \mathbf{d}\boldsymbol{\xi}$ (where the ordinary dot is defined in the obvious way on the nearest basis vectors), and the entropy functional becomes

$$S[\xi] = \int_{\mathcal{V}} (-1)^{D-2} 4[\mathbf{d}\boldsymbol{\xi} \dot{\wedge} * \mathbf{P} \dot{\wedge} \mathbf{d}\boldsymbol{\xi}] = \int_{\mathcal{V}} 4[\mathbf{d}\boldsymbol{\xi} \dot{\wedge} \mathbf{d}(*\mathbf{P} \cdot \boldsymbol{\xi})]. \quad (50)$$

Using the identity

$$\mathbf{d}\,(\mathbf{d}\boldsymbol{\xi} \dot{\wedge} * \mathbf{P} \cdot \boldsymbol{\xi}) = \mathbf{d}^2 \boldsymbol{\xi} \dot{\wedge} * \mathbf{P} \cdot \boldsymbol{\xi} - \mathbf{d}\boldsymbol{\xi} \dot{\wedge} \mathbf{d}(*\mathbf{P} \cdot \boldsymbol{\xi}), \quad (51)$$

and the definition of the (bivector valued) Riemann curvature 2-form via

$$\mathbf{d}^2 \boldsymbol{\xi} = \mathbf{R} \cdot \boldsymbol{\xi} = -\boldsymbol{\xi} \cdot \mathbf{R};$$
$$\mathbf{R} = \frac{1}{2!} \mathbf{e}_a \wedge \mathbf{e}_b \mathbf{R}^{ab}; \qquad \mathbf{R}^{ab} = \frac{1}{2!} R^{ab}{}_{cd} \boldsymbol{\omega}^c \wedge \boldsymbol{\omega}^d, \quad (52)$$

the entropy functional can be rewritten as

$$S[\xi] = -\int_{\mathcal{V}} 4[\boldsymbol{\xi} \cdot \mathbf{R} \dot{\wedge} * \mathbf{P} \cdot \boldsymbol{\xi}] - \int_{\partial\mathcal{V}} 4(\mathbf{d}\boldsymbol{\xi} \dot{\wedge} * \mathbf{P} \cdot \boldsymbol{\xi}), \quad (53)$$

where $\partial\mathcal{V}$ is the $(D-1)$-dimensional boundary of the volume $\mathcal{V}$.

We will now specialize to a subset of vector fields which are null and vary the entropy functional with respect to them. The first term in Eq. (53) is manifestly symmetric in $\boldsymbol{\xi}$ and will give a contribution with a factor of 2. (This can also be verified using the rules of exchanging the $\dot{\wedge}$, etc., or by explicitly working with the indices in place.) The second term of Eq. (53) will not contribute since the variation $\delta\boldsymbol{\xi}$ vanishes on the boundary. The condition $\boldsymbol{\xi}\cdot\boldsymbol{\xi}=0$ is preserved by adding a Lagrange multiplier term $\lambda\boldsymbol{\xi}\cdot\boldsymbol{\xi}$ to the entropy functional. The resulting equations of motion are therefore

$$-4\boldsymbol{\xi}\cdot\mathbf{R}\dot{\wedge}*\mathbf{P}+\lambda\boldsymbol{\xi}=0, \tag{54}$$

which is a set of vector equations. Taking a dot product with $\boldsymbol{\xi}$ we get

$$-4\boldsymbol{\xi}\cdot\mathbf{R}\dot{\wedge}*\mathbf{P}\cdot\boldsymbol{\xi}=0, \tag{55}$$

which must hold for all null vector fields $\boldsymbol{\xi}$. Equation (55) can clearly also be written (since $\mathbf{d}*\mathbf{P}=0$) as

$$4[\mathbf{d}(\mathbf{d}\boldsymbol{\xi}\dot{\wedge}*\mathbf{P})]\cdot\boldsymbol{\xi}=0, \tag{56}$$

which can also be derived directly by the variation of Eq. (50). It is easy to show that Eq. (55) is equivalent (after bringing back all indices) to the vacuum version of Eq. (18). To see this, consider

$$\begin{aligned}-4\boldsymbol{\xi}\cdot\mathbf{R}\dot{\wedge}*\mathbf{P}\cdot\boldsymbol{\xi}&=-4\xi^a\mathbf{R}_{ab}\wedge*\mathbf{P}^{bc}\xi_c\\&=-\frac{1}{(D-2)!}\xi^aR_{abk_1k_2}P^{bcmn}\\&\quad\times\xi_c\epsilon_{mnk_3\cdots k_D}\boldsymbol{\omega}^{k_1}\wedge\cdots\wedge\boldsymbol{\omega}^{k_D}.\end{aligned} \tag{57}$$

On equating the right-hand side of Eq. (57) to $f\boldsymbol{\epsilon}$ where f is a scalar and using arguments similar to the ones following Eq. (40), we find

$$f=2P_{mb}{}^{ij}R^m{}_{aij}\xi^a\xi^b=2P_b{}^{ijk}R^a{}_{ijk}\xi_a\xi^b, \tag{58}$$

which is what appears in the first term on the left-hand side of Eq. (18). The rest of the derivation of field equations follows as before.

Having derived the equations of motion by extremizing the entropy functional with respect to the null vector fields, we go back to the entropy functional for an arbitrary vector field and evaluate it on shell. Using Eqs. (53) and (55), the on shell value of the entropy functional for an arbitrary vector field is

$$S|_{\text{on shell}}=-4\int_{\partial\mathcal{V}}(\mathbf{d}\boldsymbol{\xi}\dot{\wedge}*\mathbf{P}\cdot\boldsymbol{\xi}). \tag{59}$$

To show that the on shell value agrees with Wald entropy for our specific choice described earlier, we take $\partial\mathcal{V}$ to be a null surface and denote an arbitrary spacelike cross section of $\partial\mathcal{V}$ by $\mathcal{H}$. Reverting to (semi)index notation, we have

$$\begin{aligned}S|_{\text{on shell}}&=-4\int_{\partial\mathcal{V}}(\mathbf{d}\boldsymbol{\xi})_a\wedge*\mathbf{P}^{ab}\xi_b\\&=-4\int_{\partial\mathcal{V}}(\nabla_k\xi_a)\Big(\frac{1}{(D-2)!}\frac{1}{2!}P^{abcd}\xi_b\epsilon_{cda_1\cdots a_{D-2}}\Big)\\&\quad\times\boldsymbol{\omega}^k\wedge\boldsymbol{\omega}^{a_1}\wedge\cdots\wedge\boldsymbol{\omega}^{a_{D-2}}.\end{aligned} \tag{60}$$

To show that this expression is the same as the Wald entropy, we will again employ the limiting procedure used in Sec. III; namely, we will consider the Rindler limit of the geometry near the null surface $\partial\mathcal{V}$. We consider the vector field $\boldsymbol{\xi}$ to be normal to surfaces of $N=$ constant for the Rindler metric Eq. (32) and take the $N\to 0$ limit at the end. The spacelike cross section $\mathcal{H}$ of $\partial\mathcal{V}$ corresponds to the transverse directions of the Rindler metric and has coordinates labeled with uppercase indices $x_\perp^A$, $A=2,3,\cdots,D-1$.

For the vector $\boldsymbol{\xi}=\xi^a\mathbf{e}_a=(0,\kappa,0,\ldots)$, we have $\nabla_k\xi_a=-(1/\kappa)\Gamma^N_{ka}$ of which only the component $\Gamma^N_{00}=\kappa^2N\neq 0$. The integrand $\mathcal{I}$ of Eq. (60) becomes

$$\begin{aligned}\mathcal{I}&=4N\Big(\frac{1}{(D-2)!}\frac{1}{2!}\Big)\\&\quad\times P^{0Ncd}\epsilon_{cda_1\cdots a_{D-2}}\boldsymbol{\omega}^0\wedge\boldsymbol{\omega}^{a_1}\wedge\cdots\wedge\boldsymbol{\omega}^{a_{D-2}}.\end{aligned} \tag{61}$$

Since the integrand is a $(D-1)$-form restricted to a surface of constant N, and since the basis 1-form $\boldsymbol{\omega}^0$ appears explicitly, it follows that the remaining basis 1-forms must be intrinsic to the spacelike cross section $\mathcal{H}$. The indices $a_1\cdots a_{D-2}$ in Eq. (61) can then be replaced by $A_1\cdots A_{D-2}$, and the integrand reduces to

$$\begin{aligned}\mathcal{I}&=4N(g^{00}g^{NN}P_{0N}{}^{0N})\sqrt{-g_{00}g_{NN}}\frac{1}{(D-2)!}\\&\quad\times\sqrt{\sigma}e_{A_1\cdots A_{D-2}}\boldsymbol{\omega}^0\wedge\boldsymbol{\omega}^{A_1}\wedge\cdots\wedge\boldsymbol{\omega}^{A_{D-2}},\end{aligned} \tag{62}$$

where $e_{A_1\cdots A_{D-2}}$ is the alternating symbol which takes values $+1(-1)$ when the indices are an even (odd) permutation of $2,3,\cdots,D-1$, and is zero otherwise. The metric coefficients refer to the Rindler metric in Eq. (32), and cancel the N dependence of the integrand. Recognizing the natural $(D-2)$-form on $\mathcal{H}$ given by

$$\tilde{\boldsymbol{\epsilon}}=\frac{1}{(D-2)!}\sqrt{\sigma}e_{A_1\cdots A_{D-2}}\boldsymbol{\omega}^{A_1}\wedge\cdots\wedge\boldsymbol{\omega}^{A_{D-2}}, \tag{63}$$

the integrand becomes

$$\mathcal{I}=-4\kappa P_{0N}{}^{0N}\boldsymbol{\omega}^0\wedge\tilde{\boldsymbol{\epsilon}}. \tag{64}$$

In a standard derivation of the expression for the Wald entropy—e.g. Ref. [25]—one would be dealing with the binormal to the cross section $\mathcal{H}$, defined in terms of two null vectors normal to $\mathcal{H}$. Since we are dealing with a limiting procedure in which the horizon (null surface $\partial\mathcal{V}$)

is approached in the limit $N \to 0$, we define a 2-form $\mathbf{n} = (1/2!)n_{ij}\boldsymbol{\omega}^i \wedge \boldsymbol{\omega}^j$ which reduces to the standard binormal to $\mathcal{H}$ on the horizon. To do this we use the fact that in the standard case, the natural $(D-2)$-form $\tilde{\boldsymbol{\epsilon}}$ on $\mathcal{H}$ is simply the dual of the 2-form binormal $\mathbf{n}$. We therefore define our "binormal" via the relation

$$n^{ij} = \frac{(-1)^s}{(D-2)!}\sqrt{\sigma}e_{A_1\cdots A_{D-2}}\epsilon^{ijA_1\cdots A_{D-2}}. \tag{65}$$

It is now straightforward to check, using $\sqrt{-g} = \sqrt{-g_{00}g_{NN}}\sqrt{\sigma}$, that $P^{abcd}n_{ab}n_{cd} = -4P_{0N}{}^{0N}$. In Appendix A 2 we have explicitly shown that this quantity is independent of N in the Lanczos-Lovelock case, and hence the $N \to 0$ limit is trivial. The on shell value of the entropy functional therefore becomes

$$S|_{\text{on shell}} = \kappa \int_{\partial\mathcal{V}} P^{abcd}n_{ab}n_{cd}\boldsymbol{\omega}^0 \wedge \tilde{\boldsymbol{\epsilon}}. \tag{66}$$

Since $\boldsymbol{\omega}^0 = \mathbf{d}t_{\mathrm{R}}$, when the on shell entropy is evaluated in a solution which is stationary, the above integral splits into a time integration and an integral over the arbitrary cross section $\mathcal{H}$ of $\partial\mathcal{V}$. Then restricting the time integration range to from 0 to $2\pi/\kappa$, we get

$$S|_{\text{on shell}} = 2\pi \oint_{\mathcal{H}} P^{abcd}n_{ab}n_{cd}\tilde{\boldsymbol{\epsilon}}, \tag{67}$$

which is precisely the expression for Wald's entropy when we recall that $P^{abcd} = (\partial\mathcal{L}/\partial R_{abcd})$ for the Lanczos-Lovelock type theories (up to a sign which depends on the sign convention used for $\tilde{\boldsymbol{\epsilon}}$). Using the definition of n_{ij} and following the algebra in Appendix A 2, one can also easily recover the expression in Eq. (34) thereby proving the equivalence of the two approaches.

V. DISCUSSION

Since we have described the ideas fairly extensively in the earlier sections as well as the introduction, we will be brief in this section and concentrate on the broader picture.

We take the point of view that the gravitational interaction described through the metric of a smooth spacetime is an emergent, long-wavelength phenomenon. Einstein's equations provide the lowest order description of the dynamics and one would expect higher order corrections to these equations as we probe the smaller scales. It was shown in Ref. [7] that the Einstein equation $G^0_0 = 8\pi T^0_0$ for spherically symmetric spacetimes with horizons can be rewritten in terms of thermodynamic variables and is in fact identical to the first law of thermodynamics $TdS = dE + PdV$. In Ref. [8] this result was extended to the $E^0_0 = (1/2)T^0_0$ equation for spherically symmetric spacetimes in Lanczos-Lovelock gravity where E^a_b was defined in Eq. (22). In fact, the Lorentz invariance of the theory, taken together with the equation $E^0_0 = (1/2)T^0_0$ then leads to the full set of equations $E^a_b = (1/2)T^a_b$ governing the dynamics of the metric g_{ab}. The invariance under Lorentz boosts in a local inertial frame maps to translation along the Rindler time coordinate in the local Rindler frame. Hence the validity of local thermodynamic description for *all* Rindler observers allows one to obtain the full set of equations from the time-time component of the equations. Given the key role played by horizons in all these, it seems natural that we should have an alternative formulation of the theory in terms of the entropy associated with the horizons. This is precisely what has been attempted in this paper.

The first key result of this paper is an alternative variational principle to obtain not only the Einstein's theory but also the more general Lanczos-Lovelock theory. We have shown that there is a natural procedure for obtaining the dynamics of the metric (which is now interpreted as a macroscopic variable like the density of a solid) using the functional $S[\xi]$ given in Eq. (1). This functional can be defined for any vector field ξ^a. When we restrict attention to null vector fields, and demand that the entropy associated with all the null vectors should be an extremum, we obtain a condition on the background geometry that is equivalent to the dynamical equations of the theory. This provides an alternative route without the usual problems which arise in the handling of surface terms, etc. when the metric is varied. Interestingly, our approach selects out the Lanczos-Lovelock type of theories, which are known to have nice properties regarding the integrability of the field equations, etc.

As an aside, we want to make a remark on the usual derivations of the field equations in Lanczos-Lovelock theory. To illustrate the point consider the familiar $D = 4$ case. We know that in 4D, the second order Lanczos-Lovelock term $\mathcal{L}^{(4)}_2$ (the Gauss-Bonnet term) is a total divergence, and the higher terms identically vanish. Usually in the literature, one will ignore the Gauss-Bonnet term, since it is a total divergence, and claim that the equations of motion are identical to Einstein's equations. One must note, however, that in a situation wherein the Lagrangian contains a total derivative term, the conventional action principle is well defined only when all surviving surface terms are held fixed. It is well known that in Einstein's GR this can be achieved by adding the Gibbons-Hawking term to cancel certain surface contributions. The case of the Gauss-Bonnet term in 4 dimensions proves to be trickier in terms of defining consistent boundary conditions. (See Ref. [26] for an attempt to address this issue, and for further references.) So the Lanczos-Lovelock theory, in the conventional formalism based on varying an action principle, faces certain difficulties associated with the boundary term. Our formalism, of course, reproduces the standard result and the correct entropy and—as a bonus—the above difficulties do not arise in our approach since we do not vary the metric to get the equations of motion.

The second result in the paper is related to providing an interpretation of the extremum value of this functional. Here we find that one needs to make some ansatz which—though physically well motivated and natural—is logically independent of the derivation of field equations. The ideas we have used are closely connected with previous attempts to interpret the radial displacements of horizons as the key to obtaining a thermodynamic interpretation of gravitational theories [7,8]. When we evaluate the on shell value of the entropy functional associated with a normal to a sequence of timelike surfaces and take the limit when these timelike surfaces approach the horizon, we obtain the standard entropy of the horizon. In the case of GR, the result is intuitively obvious since the integrand of the entropy functional has the structure $n_j(K\xi^j + a^j)$ which is essentially the surface term in Einstein-Hilbert action from which we know that the correct entropy can be obtained. Remarkably enough the same prescription works even in the case Lanczos-Lovelock gravity for which there is no simple intuitive interpretation.

Thus, at the least, we have provided an alternative variational principle to obtain not only Einstein gravity but also its closely related extensions, without varying the metric in the functional. This, by itself, is worth further study from three points of view. First, it is important to understand why it works. In conventional approaches one interprets the extremum value of action in terms of the path integral prescription in which alternative histories are explored by a quantum system; here this should correspond to fluctuations of the light cone structure in some sense. It is not clear how to make this notion more precise and useful. Second, the matter sector is—as usual—quite ugly and nongeometrical and one could even claim that it was added by hand. It is not clear whether the entropy functional, including the matter term has a geometric interpretation [19]. Finally, the work clearly endows a special status to the Lanczos-Lovelock theory as a natural extension of GR within the thermodynamic paradigm. Several previous results, especially Ref. [17], have already pointed in this direction. Given the rich geometrical structure of the Lanczos-Lovelock theory (compared to, for example, theories based on $f(R)$ Lagrangians), it is worth investigating this issue further.

These results certainly indicate a deep connection between gravity and thermodynamics *which goes well beyond Einstein's theory.* The general Lanczos-Lovelock theory, which is expected to partially account for an effective action for gravity in the semiclassical regime, satisfies the same relations between the dynamics of horizons and thermodynamics, as Einstein's GR. This suggests that these results have possible consequences concerning a quantum theory of gravity as well. In a previous work [17] it was shown that this class of theories exhibits a type of classical "holography" which assumes special significance in the backdrop of current results.

ACKNOWLEDGMENTS

We thank R. Bousso and R. Nityananda for useful comments. A. P. thanks all the members of IUCAA for their warm hospitality during his stay there, during which a part of this work was completed.

APPENDIX A

In this appendix we provide proofs for some of the results quoted in the text.

1. Proof of Eq. (19)

It is sufficient to prove that if a symmetric tensor S_{ab} satisfies

$$S_{ab}\xi^a\xi^b = 0, \tag{A1}$$

for an arbitrary null vector field ξ^a, then it must satisfy $S_{ab} = F(x)g_{ab}$ where F is some scalar function. Consider a point $\mathcal{P}$ in the spacetime and construct the local inertial frame LIF at $\mathcal{P}$, endowed with Minkowski coordinates. The vector field ξ^a appearing in Eq. (A1) is arbitrary, and the choice of the D-ad erected in the LIF is also arbitrary up to a local Lorentz transformation (LLT) (which has $D(D-1)/2$ degrees of freedom). We will utilize the $(D-1)(D-2)/2$ (spatial) rotational degrees of freedom available in the LLT to choose the D-ad axes such that the purely spatial components of S_{ab} with respect to these axes, vanish; namely $S_{\alpha\beta} = 0$, $\alpha, \beta = 1, 2, \ldots, D-1$. Similarly, we utilize the $(D-1)$ boost degrees of freedom in the LLT to choose the D-ad axes such that $S_{0\alpha} = 0$, $\alpha = 1, 2, \ldots, D-1$. Further, since the null vector field ξ is arbitrary, we will in turn consider the vector fields given by

$$\xi^a_{(\alpha)} = \delta^a_0 + \delta^a_\alpha, \qquad \alpha = 1, 2, \ldots, D-1, \tag{A2}$$

where the components are understood to be with respect to the local Minkowski coordinates. Substituting for this choice of ξ^a in Eq. (A1), we find

$$S_{00} + 2S_{0\alpha} + S_{\alpha\alpha} = 0. \tag{A3}$$

The middle term drops out because of our choice of D-ad axes. Repeated application of Eq. (A3) for all allowed values of α then gives us $S_{\alpha\alpha} = -S_{00}$ for all α, and combined with $S_{0\alpha} = 0 = S_{\alpha\beta}$, we obtain

$$S_{ab}(\mathcal{P}) = F(\mathcal{P})\eta_{ab}, \tag{A4}$$

where F is a scalar depending on the choice of $\mathcal{P}$. Since Eq. (A4) is a tensor equation in the LIF, it immediately generalizes as required, to

$$S_{ab}(x) = F(x)g_{ab}(x), \tag{A5}$$

2. Proof of Eq. (34)

We now consider the algebraic details in the evaluation of the on shell value of entropy functional. As explained in the text, we expect to obtain a meaningful result only for certain solutions, when a specific choice is made for the boundary and the vector field. To do this we define it through a limiting process involving a sequence of timelike surfaces and their normals with the limit taken at the end of the calculation. According to the prescription laid down in the text, we take $\xi^a = n^a$ to be the unit spacelike normal to the $N = \epsilon$ surfaces in the Rindler frame, and taking the $\epsilon \rightarrow 0$ limit at the very end of the calculation. (This limiting procedure is physically well motivated; in the case of standard GR and a Schwarzschild black hole, for example, it will correspond to approaching the $r = 2M$ surface as the $\epsilon \rightarrow 0$ limit of $r = 2M + \epsilon$ sequence of surfaces.) In the Rindler frame, with the metric $ds^2 = -N^2 dt_{\mathrm{R}}^2 + dN^2/\kappa^2 + dL_\perp^2$,

$$\begin{aligned} n_a &= \xi_a = (0, 1/\kappa, 0, 0, \ldots); \\ n^a &= \xi^a = (0, \kappa, 0, 0, \ldots); \qquad \sqrt{h} = \epsilon\sqrt{\sigma}, \end{aligned} \tag{A6}$$

where h is the metric determinant for the $N = \epsilon$ surfaces and σ the metric determinant for the $N =$ constant, $t_{\mathrm{R}} =$ constant surfaces. (When $\epsilon \neq 0$, the surface and its normal are not null but $\epsilon = 0$ is the null Rindler horizon). The entropy functional is

$$S[\xi] = \int_{\partial\mathcal{V},\epsilon} d^{D-1}x\sqrt{h}n_a(4P^{abcd}\xi_c\nabla_b\xi_d). \tag{A7}$$

Here, $d^{D-1}x = dt_{\mathrm{R}} d^{D-2}x_\perp$, and $\nabla_b\xi_d = -\Gamma^a{}_{bd}\xi_a = -(1/\kappa)\Gamma^N{}_{bd}$, of which only $\Gamma^N{}_{00} = \kappa^2\epsilon \neq 0$. The ϵ on the integration symbol reminds us that we are not actually on the given boundary $\partial\mathcal{V}$, but will approach it as $\epsilon \rightarrow 0$. As mentioned in the text, our choice of the tensor P^{abcd} is a single tensor ${}^{(m)}P^{abcd} = m{}^{(m)}Q^{abcd}$, and appears linearly in the expression for $S|_{\text{on shell}}$. It is sufficient to analyze this expression using the single ${}^{(m)}P^{abcd}$ and to take the appropriate linear combination for the general Lanczos-Lovelock case at the end. This will not interfere with the process of taking the limit $\epsilon \rightarrow 0$. As in the text, we will drop the superscript (m) for notational convenience. The integrand for a single m can be evaluated as follows

$$\begin{aligned} \sqrt{h}n_a(4P^{abcd}\xi_c\nabla_b\xi_d) &= \frac{\epsilon\sqrt{\sigma}}{\kappa^2}(4P^{NbNd}\nabla_b\xi_d) \\ &= \frac{\epsilon\sqrt{\sigma}}{\kappa^2}\left(-4P^{N0N0}\frac{1}{\kappa}\Gamma^N{}_{00}\right) \\ &= \frac{\epsilon^2\sqrt{\sigma}}{\kappa}(-4P^{N0N0}) \\ &= \frac{\epsilon^2\sqrt{\sigma}}{\kappa}(-4mg^{00}g^{NN}Q_{N0}{}^{N0}) \\ &= \kappa\sqrt{\sigma}(4mQ_{N0}{}^{N0}). \end{aligned} \tag{A8}$$

Consider the quantity $Q_{N0}{}^{N0}$ which—for the mth order Lanczos-Lovelock action—is given by

$$Q_{N0}{}^{N0} = \frac{1}{16\pi}\frac{1}{2^m}\delta^{N0a_3\ldots a_{2m}}_{N0b_3\ldots b_{2m}}(R^{b_3b_4}_{a_3a_4}\ldots R^{b_{2m-1}b_{2m}}_{a_{2m-1}a_{2m}})_{N=\epsilon}. \tag{A9}$$

The presence of both 0 and N in each row of the alternating tensor forces all other indices to take the values $2, 3, \ldots, D-1$. In fact, we have $\delta^{N0a_3\ldots a_{2m}}_{N0b_3\ldots b_{2m}} = \delta^{A_3A_4\ldots A_{2m}}_{B_3B_4\ldots B_{2m}}$ with $A_i, B_i = 2, 3, \ldots, D-1$ (the remaining combinations of Kronecker deltas on expanding out the alternating tensor are all zero since $\delta^0_A = 0 = \delta^N_A$ and so on). Hence $Q_{N0}{}^{N0}$ reduces to

$$Q_{N0}{}^{N0} = \frac{1}{2}\left(\frac{1}{16\pi}\frac{1}{2^{m-1}}\right)\delta^{A_3A_4\ldots A_{2m}}_{B_3B_4\ldots B_{2m}}(R^{B_3B_4}_{A_3A_4}\ldots R^{B_{2m-1}B_{2m}}_{A_{2m-1}A_{2m}})_{N=\epsilon}. \tag{A10}$$

In the $\epsilon \rightarrow 0$ limit therefore, recalling that $R^{AB}_{CD}|_{\mathcal{H}} = {}^{(D-2)}R^{AB}_{CD}|_{\mathcal{H}}$, we find that $Q_{N0}{}^{N0}$ is essentially the Lanczos-Lovelock Lagrangian of order $(m-1)$:

$$Q_{N0}{}^{N0} = \tfrac{1}{2}\mathcal{L}^{(D-2)}_{(m-1)}|_{\mathcal{H}}, \tag{A11}$$

and the entropy functional becomes

$$S|_{\mathcal{H}} = 2m\kappa\int_{\mathcal{H}} dt_{\mathrm{R}} d^{D-2}x_\perp\sqrt{\sigma}\mathcal{L}^{(D-2)}_{(m-1)}. \tag{A12}$$

Restricting the t_{R} integral to the range $(0, 2\pi/\kappa)$ as usual and using stationarity (which cancels the κ dependence of the result), we get

$$S^{(m)}|_{\mathcal{H}} = 4\pi m\int_{\mathcal{H}} d^{D-2}x_\perp\sqrt{\sigma}\mathcal{L}^{(D-2)}_{(m-1)}, \tag{A13}$$

where we have restored the superscript (m) in the last expression. Finally, taking the appropriate linear combination we find

$$\begin{aligned} S|_{\mathcal{H}} &= \sum_{m=1}^{K} c_m S^{(m)}|_{\mathcal{H}} \\ &= \sum_{m=1}^{K} 4\pi m c_m\int_{\mathcal{H}} d^{D-2}x_\perp\sqrt{\sigma}\mathcal{L}^{(D-2)}_{(m-1)}. \end{aligned} \tag{A14}$$

This is precisely the entropy in the Lanczos-Lovelock theory. We have also verified that this expression is explicitly recovered when working with a spherically symmetric metric of the form $ds^2 = -f(r)dt^2 + dr^2/f(r) + r^2 d\Omega^2_{(D-2)}$ which admits a horizon at some value of $r = r_{\mathcal{H}}$. The calculation in this case can be done using this metric (instead of the Rindler metric) and considering the limit $r \rightarrow r_{\mathcal{H}}$.

Note that, instead of the above limiting procedure, if we had just foliated the spacetime with null surfaces, chosen ξ_a to be the null normal vector field to the foliating surfaces, and taken the boundary to be one of the foliating surfaces (so that the normal vector n_a coincides with the null vector on the boundary) then the surface term $S|_{\text{on shell}}$ will give zero on $\mathcal{H}$. This is transparent in Einstein gravity, where the integrand of $S[\xi]$ becomes $\sim n_a(-\xi^a(\nabla_b\xi^b) + \xi^b\nabla_b\xi^a)$, which vanishes on $\mathcal{H}$ where n^a and ξ^a coincide and are null. It can be shown that the same result holds more generally and is only dependent on the algebraic symmetries of P^{abcd}. Similarly, if we choose $\xi^a = v^a = (1, \vec{0})$ in the local Rindler frame—which is the Rindler time translation Killing vector that becomes null on the horizon—and use the same limiting procedure, we get a vanishing entropy. On the other hand, the *normalized* vector in the timelike direction $N^{-1}v^a$ gives the same (correct) result as our choice. Clearly the final result depends on our choice and no general statement can be made.

Finally, to make contact with results in a more familiar setting, we point out that some of these features are not unique to the above context. In fact, a similar caveat also applies to the well-known Gibbons-Hawking term which is a similar surface term arising in the Einstein-Hilbert Lagrangian. Apart from some constant proportionality factors which are irrelevant to this discussion, this term actually arises as the integral of a total derivative (in 4 dimensions) as

$$\begin{aligned} A^{\text{GH}} &\sim \int_{\mathcal{V}} d^4x\sqrt{-g}\nabla_a(n^a\nabla_b n^b) \\ &= -\int_{\mathcal{V}} d^4x\sqrt{-g}\nabla_a(n^a K) \sim \int_{\partial\mathcal{V}} d^3x\sqrt{h}(v_a n^a)K, \end{aligned} \tag{A15}$$

where n^a is the normal to the foliating surfaces, v_a is the normal to the boundary $\partial\mathcal{V}$ of the 4-dimensional region $\mathcal{V}$, and $K = -\nabla_b n^b$ is the trace of the extrinsic curvature. We now consider the case in which the boundary surfaces are chosen to be members of the set of foliating surfaces (e.g., we can foliate the spacetime by $t =$ constant surfaces and choose part of $\partial\mathcal{V}$ to be given by $t = t_1$ and $t = t_2$) so that $v_a = n_a$. Then we have

$$A^{\text{GH}} \sim \int_{\partial\mathcal{V}} d^3x\sqrt{h}(n_a n^a)K \sim \int_{\partial\mathcal{V}} d^3x\sqrt{h}K, \tag{A16}$$

provided $\partial\mathcal{V}$ is *not* null, so that the normal vector can be assumed to have unit norm $n_a n^a = \pm 1$. This is the familiar expression often quoted in the literature. Clearly, the above naive argument breaks down (but the result still holds) when the spacetime is foliated by a series of null surfaces ($n_a n^a = 0$) and the boundary is one of these surfaces. But this case also can be handled by a limiting procedure similar to the one we used for computing our surface integral. In fact, our prescription essentially foliates the Rindler limit of the horizon by a series of timelike surfaces (like the $r = 2M + \epsilon =$ constant surfaces in the Schwarzschild) approaching the null horizon in a particular limit (like $\epsilon \to 0$). In the case of GR, this is equivalent to the standard calculation of integrating the extrinsic curvature (defined by this foliation) over the surface and—of course—we get the standard result of entropy density being a quarter of transverse area. What is more interesting and nontrivial is that the same prescription works in a much wider context and reproduces the Lanczos-Lovelock entropy.

[1] A. D. Sakharov, Sov. Phys. Dokl. **12**, 1040 (1968); T. Jacobson, Phys. Rev. Lett. **75**, 1260 (1995); T. Padmanabhan, Mod. Phys. Lett. A **17**, 1147 (2002); **18**, 2903 (2003); G. E. Volovik, Phys. Rep. **351**, 195 (2001); *The Universe in a Helium Droplet* (Oxford University, New York, 2003); B. L. Hu, Int. J. Theor. Phys. **44**, 1785 (2005), and references therein.

[2] H. S. Snyder, Phys. Rev. **71**, 38 (1947); B. S. DeWitt, Phys. Rev. Lett. **13**, 114 (1964); T. Yoneya, Prog. Theor. Phys. **56**, 1310 (1976); T. Padmanabhan, Ann. Phys. (N.Y.) **165**, 38 (1985); Classical Quantum Gravity **4**, L107 (1987); A. Ashtekar *et al.*, Phys. Rev. Lett. **69**, 237 (1992); T. Padmanabhan, Phys. Rev. Lett. **78**, 1854 (1997); Phys. Rev. D **57**, 6206 (1998); For a review, see L. J. Garay, Int. J. Mod. Phys. A **10**, 145 (1995).

[3] See e.g., T. Padmanabhan, Phys. Rev. Lett. **81**, 4297 (1998); Phys. Rev. D **59**, 124012 (1999), and references therein.

[4] S. A. Fulling, Phys. Rev. D **7**, 2850 (1973); S. Hawking, Commun. Math. Phys. **43**, 199 (1975); W. G. Unruh, Phys. Rev. D **14**, 870 (1976); G. W. Gibbons and S. W. Hawking, Phys. Rev. D **15**, 2752 (1977); K. Srinivasan *et al.*, Phys. Rev. D **60**, 024007 (1999); L. Sriramkumar *et al.*, Int. J. Mod. Phys. D **11**, 1 (2002).

[5] For a recent review, see: T. Padmanabhan, Phys. Rep. **406**, 49 (2005); Mod. Phys. Lett. A **17**, 923 (2002).

[6] T. Padmanabhan, Braz. J. Phys. **35**, 362 (2005); Gen. Relativ. Gravit. **34**, 2029 (2002); **35**, 2097 (2003).

[7] T. Padmanabhan, Classical Quantum Gravity **19**, 5387 (2002).

[8] A. Paranjape, S. Sarkar, and T. Padmanabhan, Phys. Rev. D **74**, 104015 (2006).

[9] R.-G. Cai and L.-M. Cao, gr-qc/0611071; M. Akbar and R.-G. Cai, hep-th/0609128.

[10] Dawood Kothawala, Sudipta Sarkar, and T. Padmanabhan, gr-qc/0701002.

[11] M. Akbar and Rong-Gen Cai, gr-qc/0612089; C. Eling *et al.*, Phys. Rev. Lett. **96**, 121301 (2006).

[12] See e.g., T. Jacobson and R. Parentani, Found. Phys. **33**, 323 (2003); T. Padmanabhan, Classical Quantum Gravity

21, 4485 (2004).

[13] These are the two reasons why we do not want to think of the standard action functional itself as some kind of an entropy, but are looking for alternatives. At a somewhat ill-defined level, one can think of Euclidean action as analogous to entropy, but this has no relationship with null surfaces and blocking of information. Also, in using the action functional, one considers the metric as dynamical and the extremization is done with respect to variations in the metric. We will in contrast have a completely different and more rigorous choice for the entropy functional.

[14] L. D. Landau and E. M. Lifshitz, *Theory of Elasticity* (Butterworth-Heinemann, Oxford, England, 1986), 3rd ed.

[15] T. Padmanabhan, Int. J. Mod. Phys. D **13**, 2293 (2004); gr-qc/0609012. This idea was earlier developed for general relativity in these papers and more detailed conceptual comparison with elasticity can be found in these papers.

[16] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973).

[17] A. Mukhopadhyay and T. Padmanabhan, Phys. Rev. D **74**, 124023 (2006).

[18] C. Lanczos, Z. Phys. **73**, 147 (1932); Ann. Math. **39**, 842 (1938); D. Lovelock, J. Math. Phys. (N.Y.) **12**, 498 (1971).

[19] The following curious fact is worth mentioning. If we treat R^{ab}_{cd} and g^{ab} as independent variables, then the gravitational Lagrangian *density* in Lanczos-Lovelock theory is formally independent of g^{ab} and is constructed from R^{ab}_{cd} and the alternating tensor. On the other hand, the matter Lagrangian density is dependent on g^{ab} but is usually independent of R^{ab}_{cd}. If $\mathcal{L} = \mathcal{L}_{\rm grav} + \mathcal{L}_{\rm matter}$ is the total Lagrangian density, then our integrand in the entropy functional is essentially the sum $4(\partial \mathcal{L}/\partial R^{ab}_{cd})\nabla_c \xi^a \nabla_d \xi^b + (1/2)(\partial \mathcal{L}/\partial g^{ab})\xi^a \xi^b$ for any *null* vector field ξ^a, because the term proportional to g_{ab} in the definition stress tensor does not contribute. We do not use this result in this paper.

[20] T. Padmanabhan, Gen. Relativ. Gravit. **38**, 1547 (2006); Int. J. Mod. Phys. D **15**, 1659 (2006); Classical Quantum Gravity **22**, L107 (2005); Dark Energy: Mystery of the Millennium, section 5; AIP Conf. Proc. **861**, 179 (2006).

[21] For a formulation of GR using null surfaces, see S. Frittelli *et al.*, J. Math. Phys. (N.Y.) **36**, 4984 (1995); **36**, 4975 (1995); **36**, 5005 (1995); for a different perspective on null surfaces, related to holography, see D. Cremades *et al.*, J. High Energy Phys. 01 (2007) 045.

[22] To avoid a possible confusion we make the following comment: We mentioned earlier that Eq. (18) can be obtained by requiring that $S[\xi]$ be invariant under scaling transformations of ξ^a while Eq. (30) on the other hand shows that the on shell value of $S[\xi]$ is *not* invariant under such a rescaling. These results are not contradictory because the earlier result required that the infinitesimal rescaling function $\epsilon(x)$ vanish on $\partial \mathcal{V}$ while here we are concerned with this *surface term* evaluated on $\partial \mathcal{V}$.

[23] R. M. Wald, Phys. Rev. D **48**, R3427 (1993); V. Iyer and R. M. Wald, Phys. Rev. D **52**, 4430 (1995).

[24] There is extensive literature on the topic of entropy in the context of higher derivative theories of gravity. For a sample, see T. Jacobson and R. C. Myers, Phys. Rev. Lett. **70**, 3684 (1993); R. C. Myers and J. Z. Simon, Phys. Rev. D **38**, 2434 (1988); R.-G. Cai, Phys. Rev. D **65**, 084014 (2002); R.-G. Cai, Phys. Lett. B **582**, 237 (2004); S. Nojiri, S. D. Odintsov, and S. Ogushi, Phys. Rev. D **65**, 023521 (2002); S. Nojiri and S. D. Odintsov, Phys. Lett. B **521**, 87 (2001); M. Cvetic, S. Nojiri, and S. D. Odintsov, Nucl. Phys. **B628**, 295 (2002); T. Clunan, S. F. Ross, and D. J. Smith, Classical Quantum Gravity **21**, 3447 (2004); I. P. Neupane, Phys. Rev. D **67**, 061501 (2003); Y. M. Cho and I. P. Neupane, Phys. Rev. D **66**, 024044 (2002); N. Deruelle, J. Katz, and S. Ogushi, Classical Quantum Gravity **21**, 1971 (2004); G. Kofinas and R. Olea, Phys. Rev. D **74**, 084035 (2006). For related work see R.-G. Cai and S. P. Kim, J. High Energy Phys. 02 (2005) 050; M. Akbar and R.-G. Cai, Phys. Lett. B **635**, 7 (2006).

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

[26] R. Aros *et al.*, Phys. Rev. Lett. **84**, 1647 (2000); Phys. Rev. D **62**, 044002 (2000); R. Olea, J. High Energy Phys. 06 (2005) 023.