Published for SISSA by Springer

Received: June 17, 2022
Accepted: June 26, 2022
Published: August 25, 2022

The second law of black hole mechanics in effective field theory

Stefan Hollands,[a,b] **Áron D. Kovács**[c,d] **and Harvey S. Reall**[e]

[a] *Institute for Theoretical Physics, Leipzig University, Brüderstrasse 16, Leipzig 04103, Germany*
[b] *MPI-MiS, Inselstrasse 22, Leipzig 04103, Germany*
[c] *SISSA and INFN Sezione di Trieste, Via Bonomea 265, Trieste 34136, Italy*
[d] *IFPU — Institute for Fundamental Physics of the Universe, Via Beirut 2, Trieste 34014, Italy*
[e] *Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, U.K.*

E-mail: stefan.hollands@uni-leipzig.de, arkovac@sissa.it, hsr1000@cam.ac.uk

Abstract: We investigate the second law of black hole mechanics in gravitational theories with higher derivative terms in the action. Wall has described a method for defining an entropy that satisfies the second law to linear order in perturbations around a stationary black hole. We show that this can be extended to define an entropy that satisfies the second law to quadratic order in perturbations, provided that one treats the higher derivative terms in the sense of effective field theory. We also address some outstanding issues with Wall's method, in particular its gauge invariance and its relation to the Iyer-Wald entropy.

Keywords: Black Holes, Classical Theories of Gravity, Effective Field Theories, Gauge Symmetry

ArXiv ePrint: 2205.15341

Open Access, © The Authors.
Article funded by SCOAP[3].

https://doi.org/10.1007/JHEP08(2022)258

JHEP08(2022)258

Contents

1 Introduction and overview **1**
- 1.1 Introduction 1
- 1.2 Summary of conventions 4
- 1.3 Gaussian null coordinates 4
- 1.4 Second law 5
- 1.5 Stationary black holes 6
- 1.6 Wall's procedure 7
- 1.7 Second law in EFT 9
- 1.8 Example: vacuum gravity 12
- 1.9 Example: scalar-tensor EFT 13
- 1.10 Example: Ricci squared gravity 14
- 1.11 Discussion 15

2 Preliminaries **17**
- 2.1 More on Gaussian Null Coordinates (GNCs) 17
- 2.2 Local covariant tensors and GNCs 19
- 2.3 Covariance of GNC tensors and BRST-formalism 29

3 Entropy current associated with a null surface **35**
- 3.1 Covariant phase space formalism 35
- 3.2 Covariant phase space derivation of IWW entropy current 37
- 3.3 Transformation of IWW entropy under a change of GNCs 43

4 Improved entropy current **48**
- 4.1 Vacuum gravity 48
- 4.2 Scalar-tensor effective field theories 54

A Curvature components in Gaussian null coordinates **57**
- A.1 Expressions on $\mathcal{N}$ 59

1 Introduction and overview

1.1 Introduction

The laws of black hole mechanics are a set of classical laws governing the behaviour of black holes. When combined with Hawking's discovery of black hole radiation, these laws provide compelling evidence for the interpretation of black holes as thermodynamic objects. These laws were first proved for Einstein gravity minimally coupled to matter. It is natural to ask whether they are also valid in extensions of Einstein gravity, for example in the presence

JHEP08(2022)258

of higher-derivative terms in the action that are expected from an effective field theory (EFT) perspective.

The first law of black hole mechanics concerns linear perturbations of a stationary (i.e. time-independent) black hole. Wald has shown that a version of the first law holds for *any* diffeomorphism invariant theory of gravity coupled to matter [1]. In particular, this leads to a definition of the entropy — the Wald entropy — of a stationary black hole in such a theory. It provides a fully satisfactory definition of the entropy of an equilibrium black hole in this very large class of theories.

The Wald entropy is defined unambiguously only for a stationary black hole. Iyer and Wald have made a proposal for the entropy of a dynamical (i.e. non-stationary) black hole solution of a general diffeomorphism-invariant theory [2]. The procedure is based on classifying possible terms according to their "boost weight", which determines how they transform under a constant rescaling of affine parameter of the horizon generators. The Iyer-Wald entropy is built from quantities of zero boost weight. Iyer and Wald showed that their definition is independent of any choice of coordinates or basis. However, they left open the question of whether or not it satisfies a second law.

Jacobson, Kang and Myers (JKM) investigated black hole entropy for the class of theories for which the gravitational Lagrangian is a function of the Ricci scalar, so-called $f(R)$ theories [3]. Such theories can be transformed into a conventional scalar-tensor theory using a field redefinition. Using this, JKM were able to define an entropy that satisfies a second law. Their entropy is proportional to the integral of $f'(R)$ over a horizon cross-section. For a stationary black hole, the JKM entropy coincides with the Wald entropy. However, for a dynamical black hole, the JKM entropy differs from the Iyer-Wald entropy.

In general, one expects the Lagrangian of a gravitational EFT to include scalars built from contractions of the Riemann tensor and its derivatives. For such a theory, Wall has sketched a procedure for constructing an entropy that satisfies the second law to *linear order* in perturbations around a stationary black hole [4]. As we shall explain, the Wall procedure supplements the Iyer-Wald entropy with terms that are linear in quantities with positive boost weight. The simplest example of such a term is the integral over a horizon cross-section of $K\bar{K}$ where K and $\bar{K}$ are the expansions of outgoing and ingoing null geodesic congruences orthogonal to the cross-section. We shall refer to the result of Wall's construction as the Iyer-Wald-Wall (IWW) entropy. For a stationary black hole it reduces to the Wald entropy.

To linear order, the second law does not imply an entropy increase, but only that the entropy does not change in time. To see why, assume that there exists a linear perturbation that leads to an entropy increase over time. Now multiply this perturbation by minus one. The result is a linear perturbation that decreases the entropy over time, in violation of the second law. Thus, to linear order, Wall's result implies that the IWW entropy can neither decrease nor increase but must remain constant in time.[1]

[1] Wall allowed for the presence of matter, which was described by an energy-momentum tensor satisfying the null energy condition. This acts as a source for the linearized gravitational field. In this case, the entropy does increase but the increase arises entirely from the matter, not from gravitational waves. We will not follow this approach for reasons explained in section 1.6.

JHEP08(2022)258

For $f(R)$ theories, the JKM and IWW entropies agree if $f(R)$ is quadratic in R but not if cubic or higher order terms are present. This suggests that, except for special theories (e.g. 2-derivative Einstein gravity), the IWW entropy is unlikely to satisfy a second law beyond the linearized approximation. For $f(R)$ theories, the JKM entropy involves adding terms to the IWW entropy that are quadratic, or higher, order in quantities with positive boost weight. So, except for special theories, it seems likely that such higher order terms will need to be included if the second law is to be extended beyond the linearized approximation.

In this paper we will establish a second law beyond the linearized approximation for a large class of theories. Our approach will be to treat these theories as EFTs, ordering terms in the equations of motion according to how many derivatives they contain. The lowest order terms are the 2-derivative terms familiar from conventional Einstein theory. These are then supplemented by terms with 4 derivatives, then 6 derivatives and so on. We assume that the coupling constants multiplying these terms are all powers of some fundamental (UV) length scale ℓ. We restrict attention to solutions lying within the regime of validity of EFT. Roughly this means that if a solution varies over a length (or time) scale L then $\ell/L \ll 1$. One then expects terms with many derivatives to be less important than terms with few derivatives.

Consider truncating such a theory by retaining only terms with N, or fewer, derivatives. Rather than attempting to prove a second law that holds *exactly* in the truncated theory, our approach will be to prove a second law that holds *to the same expected accuracy as the theory itself*, i.e., up to neglect of terms with more than N derivatives. We will show that, for any N, one can define an entropy that satisfies the second law to *quadratic order* in perturbations around a stationary black hole, in the sense that any violation of the second law will be of the same order (in ℓ/L) as the terms with more than N derivatives that have been neglected in truncating the theory. By increasing N one improves the result, so the better one knows the EFT, the better the accuracy to which the second law is satisfied.

Our entropy is defined by adding new terms to the IWW entropy. The new terms are of quadratic (or higher) order in quantities with positive boost weight. The simplest example is the integral over a horizon cross-section of $(K\bar{K})^2$. By counting derivatives, one can see that the new terms are required only once $N \geq 6$. In particular, if $N = 4$ then our result implies that the IWW entropy satisfies the second law to quadratic order in perturbations, in the EFT sense described above.

An important question concerning the entropy is its gauge-invariance. The Iyer-Wald entropy of a horizon cross-section is independent of any choice of coordinates — it depends only on the local geometry of the cross-section. Wall's procedure, which our approach extends, is based on a fixed Gaussian null coordinate system. Such a coordinate system has gauge freedom corresponding to a rescaling of the affine parameter along each horizon generator. The terms generated by Wall's procedure are not manifestly invariant if this rescaling differs from generator to generator. Nevertheless, we will prove below that the IWW entropy is invariant to linear order in perturbations around a stationary black hole. Thus the IWW entropy is gauge invariant to the same extent that it satisfies the second law, namely to linear order in perturbations. In fact, with quite a lot more work, we shall prove that the freedom to adjust higher order terms in the IWW entropy can be used to

JHEP08(2022)258

bring it to a form that is gauge invariant in a fully nonlinear sense. It is natural to guess that our extension of the IWW entropy is gauge invariant at least to quadratic order in perturbations, in the sense of EFT. We have not yet been able to prove this and we will discuss this issue further below. This is not an issue for $N = 4$ since our entropy reduces to the IWW entropy in this case.

In the rest of this section we shall explain our main results in more precise terms but omitting the proofs. Then section 2 will prove various technical results used in later sections. Section 3 will present proofs of our results concerning the IWW entropy. Section 4 presents our construction of the improved entropy that satisfies the second law to quadratic order.

1.2 Summary of conventions

The metric signature is $(- + + \cdots +)$, and the conventions for the curvature tensors are as in Wald's text [5]. Bold face letters denote differential forms, with the conventions for $\wedge$ and d as in [5]. The spacetime dimension is $n > 2$. Lower case Greek letters $\mu, \nu, \ldots$ are spacetime coordinate indices, upper case Roman letters $A, B, \ldots$ are coordinate indices on an $(n-2)$ dimensional spacelike cut C associated with the Gaussian null coordinates defined in section 1.3. $(\mu_1 \ldots \mu_s)$ resp. $[\mu_1 \ldots \mu_a]$ denotes a total symmetrization resp. anti-symmetrization with combinatorial factors included to make these operations projections onto the space symmetric resp. anti-symmetric tensors.

The notation $L[\Psi]$ (square brackets) indicates that L is a local functional of some fields Ψ, i.e. at each point, x, $L|_x$ depends on finitely many derivatives $\partial_{\mu_1} \ldots \partial_{\mu_d} \Psi|_x$. We generally set $16\pi G = 1$ and ℓ denotes a UV length scale.

1.3 Gaussian null coordinates

Consider a spacetime of dimension n containing a *smooth* null hypersurface $\mathcal{N}$ that is ruled with affinely parameterized null geodesics with future-directed tangent vector l^μ. We assume that the generators are future-complete, i.e., they extend to infinite affine parameter to the future. This hypersurface might be the event horizon of a black hole but some of our analysis applies more generally. Note that the smoothness assumption excludes important physical situations such as black hole mergers. However, it includes spacetimes describing a black hole "settling down to equillibrium", which is the physical situation we have in mind.

Assume that every null geodesic generator of $\mathcal{N}$ intersects a spacelike cross section C precisely once. We can introduce Gaussian Null Coordinates (GNCs) in a neighbourhood of $\mathcal{N}$ as follows. On $\mathcal{N}$ we let v be an affine parameter along each null generator such that $v = 0$ on C and such that $l^\mu \partial_\mu v = 1$. Then we transport C by affine time v into cross sections $C(v)$ thereby obtaining a null foliation of $\mathcal{N}$. On each $C(v)$, a (past-directed) null vector field $n = n^\mu \partial_\mu$ is next defined uniquely by demanding that it is orthogonal to $C(v)$ and $g_{\mu\nu} l^\mu n^\nu = 1$. We consider the affinely parameterized null geodesics tangent to n and call the affine parameter r with $r = 0$ on $\mathcal{N}$. Finally, we choose on C a coordinate chart $x^A, A = 1, \ldots, n-2$. Then we transport this first along the geodesics tangent to l along $\mathcal{N}$ and then along n off $\mathcal{N}$ at each fixed value of v. It can be shown that the metric and

JHEP08(2022)258

vector fields n, l take the form

$$g = 2\mathrm{d}v(\mathrm{d}r - \frac{1}{2}r^2\alpha\mathrm{d}v - r\beta_A\mathrm{d}x^A) + \mu_{AB}\mathrm{d}x^A\mathrm{d}x^B, \quad l = \partial_v, \quad n = \partial_r \tag{1.1}$$

in the GNCs (v, r, x^A), where α, β_A and μ_{AB} are, at least for small r, smooth functions of the coordinates. $\mathcal{N}$ is the surface $r = 0$ and C is the surface $v = r = 0$. We denote the inverse of μ_{AB} as μ^{AB} and $A, B, \ldots$ indices will be raised and lowered with μ^{AB} and μ_{AB}. The covariant derivative on $C(v)$ defined by μ_{AB} is denoted D_A.

We also define

$$K_{AB} \equiv \frac{1}{2}\partial_v\mu_{AB}, \quad \bar{K}_{AB} \equiv \frac{1}{2}\partial_r\mu_{AB} \qquad K \equiv K^A_A \qquad \bar{K} \equiv \bar{K}^A_A. \tag{1.2}$$

On $\mathcal{N}$, K is the expansions of the null geodesics tangent to l. Similarly $\bar{K}$ is the expansion of the (past-directed) null geodesics tangent to n. The tracefree parts of K_{AB} and $\bar{K}_{AB}$ define the shear of these families of geodesics.

The definition of GNCs is not unique. If one fixes the initial cut C then one still has the freedom to rescale the affine parameter along each generator of $\mathcal{N}$: if $a(x^A) > 0$ then $v' \equiv v/a$ is also an affine parameter along the generators. The corresponding tangent vector to these generators is $l' = al$. This change of affine parameter leads to a new set of GNCs $(v', r', x^{A'})$.

If a is *constant* then we have simply

$$v' = \frac{v}{a} \qquad r' = ar \qquad x^{A'} = x^A \qquad \text{constant } a. \tag{1.3}$$

If a quantity X transforms homogeneously under such a change of coordinates then we define its *boost weight* b by

$$X'(v', r', x^{A'}) = a^b X(v, r, x^A) \tag{1.4}$$

(see section 2.1 for a precise definition); for example α, β_A and μ_{AB} have boost weight zero, K_{AB} has boost weight $+1$ and $\bar{K}_{AB}$ has boost weight -1. A quantity $D_{A_1} \ldots D_{A_m}\partial_v^p\partial_r^q\psi$ with $\psi \in \{\alpha, \beta_A, \mu_{AB}\}$ has boost weight $p - q$. The boost weight of a tensor component $T^{\mu\nu\cdots}{}_{\rho\sigma\ldots}$ is given by the sum of $+1$ for each subscript v or superscript r index, and -1 for each superscript v or subscript r index e.g. R_{vv} has boost weight $+2$.

If a is not constant then the transformation between the GNCs is highly non-trivial away from $\mathcal{N}$. This implies that many quantities (e.g. α, β_A, $\partial_v\partial_r\mu_{AB}$) transform inhomogeneously, even on $\mathcal{N}$, with terms involving the derivative of a.

1.4 Second law

We now review a simple proof of the second law in conventional GR. The proof is simple because it makes various strong assumptions, specifically that the horizon is smooth, the horizon generators are future-complete, and that the black hole "settles down to equilibrium" at late time. One can of course prove the second law in conventional GR under much weaker assumptions than these [6].

JHEP08(2022)258

Let $\mathcal{H}$ be the future event horizon of a black hole and take $\mathcal{N} = \mathcal{H}$ to define GNCs. The area of $C(v)$ is

$$A(v) = \int_{C(v)} \mathrm{d}^{n-2}x\sqrt{\mu} \tag{1.5}$$

with rate of change

$$\dot{A}(v) = \int_{C(v)} \mathrm{d}^{n-2}x\sqrt{\mu}K\,. \tag{1.6}$$

Recall Raychaudhuri's equation:

$$\partial_v K = -K^{AB}K_{AB} - R_{vv} \qquad \text{on } \mathcal{N}\,. \tag{1.7}$$

We now *assume* that the expansion of the generators of $\mathcal{H}$ vanishes at late time, i.e., $K \to 0$ as $v \to \infty$ on $\mathcal{H}$. This would be the case if the black hole is "settling down to equilibrium". We can now write

$$\begin{aligned}\dot{A}(v) &= \int_{C(v)} \mathrm{d}^{n-2}x\sqrt{\mu}\int_v^\infty dv'(-\partial_v K)(v',x)\\ &= \int_{C(v)} \mathrm{d}^{n-2}x\sqrt{\mu}\int_v^\infty dv'\left(K^{AB}K_{AB} + R_{vv}\right)(v',x)\end{aligned} \tag{1.8}$$

where we used Raychaudhuri's equation in the second step. If the spacetime satisfies the null convergence condition ($R_{\mu\nu}V^\mu V^\nu \geq 0$ for any null V^μ) then the r.h.s. is manifestly non-negative (as $R_{vv} = R_{\mu\nu}l^\mu l^\nu \geq 0$) and so we have $\dot{A}(v) \geq 0$, i.e, $A(v)$ is an increasing function.

If we consider a theory consisting of conventional 2-derivative GR coupled to matter satisfying the null energy condition then spacetimes satisfying the Einstein equation will obey the null convergence condition and so the second law holds in such a theory. However, we will be interested in more general theories in which higher derivatives are present in the Lagrangian. In this case there is no reason to expect the null convergence condition to be satisfied by solutions of the equations of motion and so the above argument no longer applies.

1.5 Stationary black holes

Much of this paper will concern perturbations of stationary black holes. We will now describe the class of stationary black holes to be considered. In conventional GR coupled to various types of matter fields it is known that the event horizon of a stationary black hole must be a Killing horizon. This result has not been extended to theories of the type that we will be considering. We will simply *assume* that the theory admits a family $\mathcal{F}$ of stationary black hole solutions for which the event horizon is a Killing horizon. Moreover, we will assume that this is a *bifurcate* Killing horizon. This implies that the zeroth law of black hole mechanics is satisfied,[2] i.e., that the surface gravity is constant on the horizon (conversely, the zeroth law implies that the spacetime can be extended to contain a bifurcation surface [8]).

[2]Very recently, ref. [7] has proved a zeroth law for a large class of theories without assuming a bifurcate Killing horizon, under the assumption that higher derivative terms in the action are multiplied by powers of a dimensionful constant (analogous to our ℓ below) and that the metric depends analytically on this constant.

JHEP08(2022)258

These assumptions ensure that if $\mathcal{H}$ is the future event horizon of a black hole in $\mathcal{F}$ then all positive boost-weight quantities vanish on $\mathcal{H}$. Furthermore, all non-zero boost-weight quantities vanish on the bifurcation surface.

1.6 Wall's procedure

Wall [4] has sketched an approach that, for any diffeomorphism invariant theory of gravity, produces an entropy S which satisfies the second law to linear order, i.e., $\delta\dot{S} = 0$ for perturbations of a black hole in the family $\mathcal{F}$.[3] Wall's approach has been discussed in more detail in [11, 12], where it has been reformulated in terms of an *entropy current*. To explain this, let $\sqrt{-g}E^{\mu\nu} = \delta I/\delta g_{\mu\nu}$, where I is the action, so the Einstein equation of our theory is $E_{\mu\nu} = 0$. Then, for a general null hypersurface $\mathcal{N}$ the entropy current is a vector field $s^v\partial_v + s^A\partial_A$ tangent to $\mathcal{N}$, where s^v and s^A are functions of the GNC components of the metric, and their derivatives, with boost-weights 0 and 1 respectively. It satisfies the identity

$$E_{vv} = \partial_v \left[\frac{1}{\sqrt{\mu}} \partial_v \left(\sqrt{\mu} s^v \right) + D_A s^A \right] + \dots \tag{1.9}$$

where the ellipsis denotes terms that are of quadratic or higher order in quantities of positive boost weight. This is an off-shell identity, i.e., it holds independently of the equations of motion. To obtain the linearized second law, we consider this equation linearized around a member of $\mathcal{F}$, taking $\mathcal{N}$ to be the event horizon $\mathcal{H}$. Since positive boost weight quantities vanish on the event horizon of the unperturbed spacetime we obtain

$$\delta E_{vv} = \partial_v \delta \left[\frac{1}{\sqrt{\mu}} \partial_v \left(\sqrt{\mu} s^v \right) + D_A s^A \right] . \tag{1.10}$$

On-shell this becomes

$$\partial_v \delta \left[\frac{1}{\sqrt{\mu}} \partial_v \left(\sqrt{\mu} s^v \right) + D_A s^A \right] = 0 . \tag{1.11}$$

The quantity inside square brackets vanishes for the background solution and can be expected to decay as $v \to \infty$ if the perturbed spacetime settles down to a stationary black hole belonging to $\mathcal{F}$. Hence integrating the above equation w.r.t. v gives

$$\frac{1}{\sqrt{\mu}} \delta \partial_v \left(\sqrt{\mu} s^v \right) + D_A \delta s^A = 0 . \tag{1.12}$$

Equation (1.11) shows that if this condition holds on C then it holds on all of $\mathcal{H}$. This equation can be regarded as a gauge condition on the metric perturbation. It arises because we have chosen our coordinates such that the perturbation does not change the location of $\mathcal{H}$, i.e., the horizon of the perturbed black hole remains at $r = 0$. For conventional GR (where (1.12) is simply $\delta K = 0$) this was explained in [13].

[3] See [9, 10] for earlier work establishing a linearized second law in particular theories.

JHEP08(2022)258

Integrating (1.12) over $C(v)$ (with measure $\sqrt{\mu}$) the divergence drops out and we obtain $\delta\dot{S}_{\rm IWW} = 0$ where the Iyer-Wald-Wall entropy is[4]

$$S_{\rm IWW}(v) = 4\pi \int_{C(v)} {\rm d}^{n-2}x\sqrt{\mu}s^v \,. \tag{1.13}$$

Thus this definition of entropy satisfies the second law to linear order in perturbations.

Wall considered coupling the original theory to a matter source so that the Einstein equation becomes $E_{\mu\nu} = (1/2)T_{\mu\nu}$ (units: $16\pi G = 1$) where $T_{\mu\nu}$ is the energy-momentum tensor of the matter. If one treats $T_{\mu\nu}$ as a term of linear order then equation (1.11) has $(1/2)T_{\mu\nu}$ on the r.h.s. and if the matter obeys the null energy condition $T_{vv} \geq 0$ then $\delta\dot{S}_{\rm IWW} \geq 0$ so one can have a genuine increase in entropy driven by the matter source. However, we shall not include such a matter source below for several reasons: (i) if the matter fields are treated using the linearized approximation, as with the gravitational field, then $T_{\mu\nu}$ is of quadratic, not linear, order; (ii) in EFT one does not expect a clear division of the Lagrangian into a "gravitational part" and a "matter part"; instead there will be higher derivative terms mixing matter and gravitational fields; (iii) higher derivative matter terms will not satisfy the null energy condition. We take the view that one should treat the gravitational and matter fields on an equal footing.[5]

Section 3 of this paper will address some outstanding issues concerning Wall's approach.

First, one needs to make sure that this definition of entropy also satisfies the *first* law of black hole mechanics, which relates δS evaluated on the bifurcation surface of the black hole to the perturbations in mass and angular momentum. Wall argues that this must be true as follows. His result for s^v can be divided into terms built entirely from quantities with vanishing boost weight and a part that is at least quadratic in quantities with non-vanishing boost weight. Wall claims that the integral of the former part is the same as the entropy $S_{\rm IW}$ defined by Iyer and Wald. From this claim it follows that $\delta S_{\rm IWW} = \delta S_{\rm IW}$ on the bifurcation surface of a member of $\mathcal{F}$ and since $S_{\rm IW}$ is known to satisfy the first law, so must $S_{\rm IWW}$. However, a proof of this claim is lacking. We shall present one in section 3.2.

Second, Wall's procedure is carried out using a particular set of GNCs and produces an expression for s^v that depends on the metric components, and their derivatives, w.r.t. those GNCs. These quantities transform in a complicated way under transformations between GNCs with non-constant $a(x^A)$. Therefore it is very unclear whether the result of Wall's procedure must be gauge-invariant. (This has also been noted in [14].)

The issue of gauge invariance is important if we wish to compare the entropy of two cuts C, C' of $\mathcal{H}$ with C' strictly to the future of $\mathcal{H}$. The linearized second law described above lets us do this for the special case where $C' = C(v)$. So to compare the IWW entropy of C' and C we must choose our GNCs such that C' is a constant v cut of $\mathcal{H}$. This can always be achieved by rescaling the affine parameters of the generators of $\mathcal{H}$, i.e., by a choice of the function $a(x^A)$ defined in section 1.3. But we do not want the definition of the entropy

[4]Recall our choice of units $16\pi G = 1$ so $1/(4G) = 4\pi$. This factor is included for convenience, so that we have $s^v = 1$ for standard GR.

[5]Wall's analysis does allow for a scalar field which is treated on an equal footing with the metric, i.e., this field is not included in $T_{\mu\nu}$.

JHEP08(2022)258

of C to depend on our choice of C' (or vice versa). Hence we must demonstrate that this definition is gauge-invariant under changes of GNCs involving non-constant $a(x^A)$.

In section 3.3, we shall give a short proof that the IWW entropy is gauge-invariant at the linearized level. However, in the class of examples studied by Wall [4] (a Lagrangian that is a function of the Riemann tensor) one can see that the result is actually gauge-invariant at the fully nonlinear level and one can ask whether this is true generally. In Wall's analysis, he implicitly makes use of the fact that his procedure only determines terms in $S_{\rm IWW}$ that are of up to linear order in quantities with positive boost weight since terms of quadratic (e.g. $K^2\bar{K}^2$) or higher order vanish when linearized around a member of $\mathcal{F}$. Wall makes a specific choice of these higher order terms in his explicit examples. The question is whether it is always possible to choose these higher order terms to render $S_{\rm IWW}$ fully gauge-invariant. We shall show in section 3.3 that the answer is yes.

1.7 Second law in EFT

We can now explain our main result, derived in section 4.1. For simplicity, we shall consider pure gravity, without matter fields. Consider an EFT for a diffeomorphism invariant theory of gravity. We assume that the Lagrangian is a formal sum of terms with increasing numbers of derivatives of the fields, multiplied by suitable powers of some UV length scale ℓ. A term with $k+2$ derivatives will be multiplied by ℓ^k. For pure gravity (assuming parity symmetry if n is odd) only terms with even numbers of derivatives can occur. The terms with 2 or fewer derivatives are assumed to take the standard Einstein-Hilbert form, with a possible cosmological constant.

Validity of EFT requires that terms with increasing numbers of derivatives are increasingly less important. This is the case if we restrict ourselves to spacetimes varying over some length/time scale L satisfying $\ell/L \ll 1$. This L should be a lower bound for the size of the final black hole equilibrium state and any length/time scales associated with the perturbation away from equilibrium. More precisely, we assume that there exists $L > 0$ with $\ell/L \ll 1$ such that, in a neighbourhood of the event horizon, w.r.t. some set of GNCs, any quantity X_k involving k derivatives of the metric obeys a uniform bound of the form $|X_k| \leq C_k/L^k$ for some dimensionless constants C_k depending on the initial data. We assume that the cosmological constant satisfies $|\Lambda|L^2 \leq 1$.

In practice, an EFT Lagrangian will not be known to all orders. We assume that only the terms with N or fewer derivatives are known explicitly. The EFT equation of motion is then

$$E_{\mu\nu} = \mathcal{O}(\ell^N) \tag{1.14}$$

where $E_{\mu\nu} = -\Lambda g_{\mu\nu} - G_{\mu\nu} + \ldots$ denotes the terms with up to N derivatives and the r.h.s. represents the effects of the terms with $N+2$ or more derivatives. We should really write this as $\mathcal{O}(\ell^N/L^{N+2})$ but we shall mostly suppress the L-dependence below.

Consider first the case $N = 2$, i.e., conventional GR viewed as an EFT. Without matter we have $R_{vv} = -E_{vv} = \mathcal{O}(\ell^2)$. Hence equation (1.8) is

$$\dot{A}(v) = \int_{C(v)} \mathrm{d}^{n-2}x\sqrt{\mu}\int_v^\infty \mathrm{d}v' \left(K^{AB}K_{AB}(v',x) + \mathcal{O}(\ell^2)\right) . \tag{1.15}$$

JHEP08(2022)258

The r.h.s. may become negative but only by a small $\mathcal{O}(\ell^2)$ amount. From an EFT perspective the second law no longer holds exactly, but only to the same $\mathcal{O}(\ell^2)$ accuracy as the theory itself. If $\dot{A}(v_0) < 0$ then the above equation implies (reinstating L) $||K_{AB}||_{L^2(v_0)} = \mathcal{O}(\ell/L^2)$ where $||\cdot||_{L^2(v_0)}$ denotes the norm defined by the above integral on the portion $v \geq v_0$ of $\mathcal{H}$. Hence if $\dot{A}(v_0) < 0$ then the expansion and shear of the generators of $\mathcal{H}$ must be small for all times $v \geq v_0$. This suggests that the black hole is close to equilibrium. Thus the process of relaxation to equilibrium seems the most likely situation in which the higher derivative EFT terms could cause a (small) decrease in horizon area.

Our main result is to show that, for general N, the following EFT generalization of (1.9) (evaluated on-shell) holds on a null hypersurface $\mathcal{N}$:[6]

$$\partial_v \left[\frac{1}{\sqrt{\mu}} \partial_v \left(\sqrt{\mu} S^v \right) + D_A S^A \right] = - \left(K^{AB} + X^{AB} \right) \left(K_{AB} + X_{AB} \right) - D_A Y^A + \mathcal{O}(\ell^N) \,. \tag{1.16}$$

Here $S^A = s^A$, and S^v is defined by adding new terms to the expression for s^v arising from $E_{\mu\nu}$. These new terms are quadratic (or higher order) in terms with positive boost weight so they do not affect any of the results described in the previous section, i.e., such terms are not fixed by Wall's procedure. On the r.h.s., X_{AB} (which is symmetric) and Y^A are both $\mathcal{O}(\ell^2)$ and have boost weight $1, 2$ respectively, and Y^A is a sum of terms that each contain at least two factors of positive boost weight. In contrast with (1.9), the above equation holds on-shell, i.e., its derivation makes use of most of the components of the Einstein equation (1.14). The $\mathcal{O}(\ell^N)$ term arises not only from the $\mathcal{O}(\ell^N)$ corrections on the r.h.s. of (1.14) but also from various steps in the derivation. In other words, even if we worked with the exact equation $E_{\mu\nu} = 0$ we would still generate a $\mathcal{O}(\ell^N)$ term. (In this regard the $N = 2$ case is exceptional.)

To obtain a second law we take $\mathcal{N}$ to be the horizon $\mathcal{H}$ of a black hole and use S^v to define an entropy S as in (1.13):

$$S(v) = 4\pi \int_{C(v)} \mathrm{d}^{n-2}x \sqrt{\mu} S^v \,. \tag{1.17}$$

Equation (1.16) implies

$$\dot{S}(v) = \frac{1}{4} \int_{C(v)} \mathrm{d}^{n-2}x \sqrt{\mu} \int_v^\infty \mathrm{d}v' \left[\left(K^{AB} + X^{AB} \right) \left(K_{AB} + X_{AB} \right) + D_A Y^A + \mathcal{O}(\ell^N) \right] (v', x) \tag{1.18}$$

where again we assume that our black hole settles down to a member of $\mathcal{F}$ at late time, which implies that all quantities with positive boost weight vanish at late time, hence the (boost-weight 1) quantity in square brackets on the l.h.s. of (1.16) vanishes as $v \to \infty$. In constrast with the $N = 2$ case, we do not know anything about the sign of $D_A Y^A$ on the r.h.s. of (1.18) in a fully nonlinear situation. However, we can make progress by resorting to second order perturbation theory about a member of $\mathcal{F}$.

[6] A similar equation was derived for Lovelock theories in [15].

JHEP08(2022)258

Positive boost weight quantities vanish on the horizon of a member of $\mathcal{F}$, which implies that S^v and s^v agree at zeroth and first order in perturbation theory. Since Y^A is at least quadratic in positive boost weight quantities we have

$$\delta^2 \int_{C(v)} \mathrm{d}^{n-2}x\sqrt{\mu} \int_v^\infty \mathrm{d}v'(D_A Y^A)(v',x) = \int_{C(v)} \mathrm{d}^{n-2}x\sqrt{\mu} \int_v^\infty \mathrm{d}v'(D_A \delta^2 Y^A)(v',x)\,. \tag{1.19}$$

We are working to quadratic order so on the r.h.s. D_A and $\sqrt{\mu}$ are evaluated in the background spacetime, where they are independent of v. Hence we can interchange the order of integration and use the divergence theorem to see that the r.h.s. vanishes. This leaves

$$\delta^2 \dot{S}(v) = \frac{1}{4} \int_{C(v)} \mathrm{d}^{n-2}x\sqrt{\mu} \int_v^\infty \mathrm{d}v' \left[\mu^{AC}\mu^{BD}\delta\left(K_{AB}+X_{AB}\right)\delta\left(K_{CD}+X_{CD}\right)+\mathcal{O}(\ell^N)\right](v',x) \tag{1.20}$$

where we used the fact that K_{AB} and X_{AB} have boost weight one and hence vanish on $\mathcal{N}$ in the background spacetime. Since the first term on the r.h.s. is positive definite, this shows that, to quadratic order in perturbations, $\dot{S}(v)$ can become negative only by an amount $\mathcal{O}(\ell^N)$, i.e., of the same size as the unknown effects caused by our ignorance of higher order EFT terms. In particular, the better we know the EFT, the larger N is, and the smaller the amount by which $\dot{S}(v)$ can become negative.[7]

Since our equations of motion involve N or fewer derivatives, it follows that S^v and S^A involve $N-2$ or fewer derivatives. These quantities are sums of monomials where each monomial is at most quadratic in terms with positive boost weight. It is natural to group such monomials into those with zero, one or two (or more) factors with positive boost weight. The former terms generate the Iyer-Wald entropy, including the second set of terms gives the Iyer-Wald-Wall entropy and including the final set of terms gives our generalization of this. A term with non-zero boost weight b costs at least b derivatives so a term quadratic in positive boost weight quantities, with overall boost weight zero, must contain at least 4 derivatives and therefore occurs only for $N \geq 6$. Hence for $N = 4$ we have $S^v = s^v$ and $S^A = s^A$ so in this case our result implies that, without further modification, the Iyer-Wald-Wall entropy satisfies the second law to quadratic order in perturbations around a member of $\mathcal{F}$.

Equation (1.16) holds for a purely gravitational EFT. We shall make a few remarks about the generalization to include matter fields. With matter fields there is a greater variety of higher derivative terms that can occur on the r.h.s. of (1.16). However, there are also further terms on the r.h.s. of (1.16) arising from the energy-momentum tensor of the 2-derivative part of the matter Lagrangian. If this 2-derivative Lagrangian satisfies the null energy condition then these terms have a "good sign" (i.e. they are negative definite). The hope is that this can be used to help control the behaviour of higher derivative terms

[7] There is another way of looking at the second law to quadratic order. If S satisfies a second law then any horizon cross-section satisfies $\dot{S} \geq 0$. A horizon cross-section of a stationary black hole satisfies $\dot{S} = 0$ and therefore minimizes $\dot{S}$ within the "space of cross-sections of black hole horizons". Hence if the second law holds then variations around a stationary black hole must satisfy the conditions for a local minimum, i.e., $\delta\dot{S} = 0$ and $\delta^2\dot{S} \geq 0$.

JHEP08(2022)258

involving the matter fields, by completing the square in the same way that we did with the $K_{AB}K^{AB}$ term in vacuum gravity. For example, a minimally coupled 2-derivative scalar field contributes $-(1/2)(\partial_v\Phi)^2$ to the r.h.s. of (1.16). So one can try to use this to control higher derivative terms involving $\partial_v\Phi$ by completing the square. This will lead to an extra term of the form $-(1/2)(\partial_v\Phi + P)^2$ on the r.h.s. of (1.16). (We will study an example of this in section 1.9.) This will give an extra term $[\delta(\partial_v\Phi + P)]^2$ inside the integral of (1.20). Since this term has a good sign, the above argument that the second law holds to quadratic order, in the sense of EFT, is still valid. In section 4.2 we shall sketch a proof that this can be done for any scalar-tensor EFT.

1.8 Example: vacuum gravity

Field redefinitions can be used to simplify the Lagrangian of an EFT. For example, in vacuum gravity, a field redefinition can be used to bring the terms with up to 4 derivatives to the "Einstein-Gauss-Bonnet" form:[8]

$$L = -2\Lambda + R + \frac{1}{4}k\ell^2 L_{\rm GB} \tag{1.21}$$

where k is a dimensionless constant and $L_{\rm GB}$ is the Euler density associated with the Gauss-Bonnet invariant:

$$L_{\rm GB} = \frac{1}{4}\delta^{\mu_1\mu_2\mu_3\mu_4}_{\nu_1\nu_2\nu_3\nu_4} R_{\mu_1\mu_2}{}^{\nu_1\nu_2} R_{\mu_3\mu_4}{}^{\nu_3\nu_4}\,. \tag{1.22}$$

This term is topological in $n = 4$ dimensions but non-trivial in higher dimensions. This theory has second order equations of motion and admits a well-posed initial value problem [17, 18] as long as it remains within the regime of validity of EFT.

Since we have $N = 4$ here, we have $S^v = s^v$ and $S^A = s^A$ as explained above. Using previous results for s^v and s^A [4, 11, 12] gives

$$S^v = 1 + \frac{k}{2}\ell^2 R[\mu] \qquad S^A = -k\ell^2\left(D_B K^{AB} - D^A K\right) \tag{1.23}$$

where $R[\mu]$ is the Ricci scalar of μ_{AB}. In this case S^v involves only quantities of zero boost weight, so our entropy is simply the Iyer-Wald entropy $S_{\rm IW}$. Hence, for Einstein-Gauss-Bonnet theory, our result implies that the Iyer-Wald entropy satisfies the second law to quadratic order, in the EFT sense explained above, i.e., $\delta^2\dot{S}_{\rm IW}$ is positive up to $\mathcal{O}(\ell^4)$ terms.

We can also demonstrate how to obtain (1.16) for this theory. A calculation gives the off-shell identity

$$\begin{aligned} -E_{vv} = &-\partial_v\left[\frac{1}{\sqrt{\mu}}\partial_v\sqrt{\mu}\right] - K^{AB}K_{AB} \\ &+\partial_v\left[\frac{k\ell^2}{\sqrt{\mu}}\partial_v\left(-\frac{1}{2}\sqrt{\mu}R[\mu]\right) + k\ell^2 D^A\left(D^B K_{AB} - D_A K\right)\right] - K_{AB}W^{AB} - \mathcal{D}_A Y^A \end{aligned} \tag{1.24}$$

[8] See [16] for previous work on the second law in this theory under the assumption of spherical symmetry.

JHEP08(2022)258

where the terms on the first line are the usual terms arising from the Einstein-Hilbert Lagrangian, and the terms on the second line arise from L_{GB}. These involve

$$\begin{aligned} W^{AB} = -k\ell^2 \Big[& \frac{1}{2}\mu^{AB}\partial_v R[\mu] - \partial_v R[\mu]^{AB} + R_{vv}\left(\bar{K}^{AB} - \bar{K}\mu^{AB}\right) \\ & + 2R_v{}^{(A}{}_{vC}\bar{K}^{B)C} - R_{vCvD}\bar{K}^{CD}\mu^{AB} - R_v{}^A{}_v{}^B\bar{K} \\ & + \mathcal{D}^{(A}R_{vC}{}^{B)C} - \mathcal{D}_C R_{vD}{}^{CD}\mu^{AB} + \mathcal{D}_C R_v{}^{(A|C|B)} \Big] \end{aligned} \tag{1.25}$$

and

$$Y^A = -k\ell^2 \left[K R_{vC}{}^{AC} - K^{AC} R_{vBC}{}^B - K_{BC} R_v{}^{CAB} \right]. \tag{1.26}$$

where $R[\mu]_{AB}$ is the Ricci tensor of μ_{AB}. These expressions involve a new derivative operator $\mathcal{D}_A$. Acting on a quantity of boost weight b this is defined as $\mathcal{D}_A = D_A - b\beta_A/2$ ($b = 1$ in the above expressions). As we shall explain below, this a connection on the normal bundle of the horizon cross-section. Note that each term in Y^A contains a factor of K_{BC} so we can write $K^{AB}K_{AB} + K_{AB}W^{AB} + \mathcal{D}_A Y^A = (K_{AB} + X_{AB})(K^{AB} + X^{AB}) + D_A Y^A + \mathcal{O}(\ell^4)$ for some $\mathcal{O}(\ell^2)$ quantity X_{AB}, with the $\mathcal{O}(\ell^4)$ "error" term arising from completing the square. So in this case, we have the off-shell result that $-E_{vv}$ can be written as the difference between the l.h.s. and r.h.s. of (1.16) up to $\mathcal{O}(\ell^4)$ terms. On-shell, the equation of motion gives $E_{vv} = \mathcal{O}(\ell^4)$ so (1.16) holds as claimed.[9] We have used only the vv-component of the Einstein equation. This is atypical: in general, the derivation of (1.16) makes use of multiple components of the Einstein equation.

In this example, the entropy is simply the IW entropy, which is manifestly gauge-invariant, in agreement with our general result. However note that S^A is not gauge-invariant. This is because it is written in terms of D_A rather than $\mathcal{D}_A$. So the entropy current is not gauge-invariant in this example (see also [14] for discussion of this point).

1.9 Example: scalar-tensor EFT

As a second example, we will consider the EFT of gravity coupled to a scalar field in $n = 4$ dimensions. As above, field redefinitions can be used to simplify the Lagrangian [19]. In particular, assuming a parity symmetry, terms with up to 4 derivatives can be written

$$L = -V(\Phi) + R + X + \frac{1}{2}\ell^2\alpha(\Phi)X^2 + \frac{1}{4}\ell^2\beta(\Phi)L_{\rm GB} \tag{1.27}$$

where $X \equiv -\frac{1}{2}g^{\mu\nu}\partial_\mu\Phi\partial_\nu\Phi$, and V, α, β are arbitrary functions. (The coupling functions α, β should not be confused with the GNC metric components α, β_A.) The coupling functions α, β are dimensionless. This theory has second order equations of motion and admits a well-posed initial value problem [17, 18] as long as it remains within the regime of validity of EFT.

[9] In this example (and the one of the next subsection) we could also include extra matter fields satisfying the null energy condition, as discussed in section 1.6, which would contribute negatively to the r.h.s. of (1.16) and so the second law would still hold to quadratic order in the gravitational perturbation.

JHEP08(2022)258

The Einstein equation is $E_{\mu\nu} = \mathcal{O}(\ell^4)$ where the r.h.s. arises from the unkown EFT terms with 6 or more derivatives and the l.h.s. is

$$-E_{\mu\nu} = G_{\mu\nu} - \left(\frac{1}{2} + \alpha\ell^2 X\right)\partial_\mu\Phi\partial_\nu\Phi - \frac{1}{2}g_{\mu\nu}\left(X - V + \frac{1}{2}\ell^2\alpha X^2\right) + \\ -\frac{1}{4}\ell^2\epsilon_\mu{}^{\alpha_1\alpha_2\alpha_3}\epsilon_\nu{}^{\beta_1\beta_2\beta_3}R_{\alpha_1\alpha_2\beta_1\beta_2}\nabla_{\alpha_3}\nabla_{\beta_3}\beta\,. \tag{1.28}$$

A calculation gives the off-shell result

$$-E_{vv} = -\partial_v\left[\frac{1}{\sqrt{\mu}}\partial_v\left(\sqrt{\mu}S^v\right) + D_A S^A\right] - K^{AB}K_{AB} - \left(\frac{1}{2} + \alpha\ell^2 X\right)(\partial_v\Phi)^2 + \\ -D_A Y^A - K_{AB}W^{AB} - \beta'\partial_v\Phi W \tag{1.29}$$

where

$$S^v = 1 + \frac{1}{2}\ell^2\beta(\Phi)R[\mu] \\ S^A = \ell^2\left(\beta(D_B K^{AB} - D^B K) - D_B\beta(K^{AB} - K\mu^{AB})\right). \tag{1.30}$$

and Y^A, W^{AB}, W are $\mathcal{O}(\ell^2)$ quantities that we shall not write out explicitly. We can complete the square on K_{AB} as in our previous example, generating an $\mathcal{O}(\ell^4)$ error term. We can also complete the square on $\partial_v\Phi$, again generating an $\mathcal{O}(\ell^4)$ error term. So on-shell we obtain an equation of the form (1.16) with an extra term of the form $-(1/2)(\partial_v\Phi + P)^2$ on the r.h.s. for some $\mathcal{O}(\ell^2)$ quantity P. This term has a good sign and so the second law holds to quadratic order in perturbations, in the sense of EFT, as explained above.

As in our previous example, only the vv component of the Einstein equation is used above, and S^v involves only quantities of zero boost weight, so for this theory our entropy is simply the Iyer-Wald entropy $S_{\rm IW}$, which is manifestly gauge-invariant.

1.10 Example: Ricci squared gravity

The above examples were atypical because the IWW entropy coincides with the IW entropy, and because equations (1.24) and (1.29) hold off-shell. As a more typical example consider a theory of vacuum gravity with

$$L = -2\Lambda + R + k_1\ell^2 R_{\mu\nu}R^{\mu\nu} + k_2\ell^2 R^2\,. \tag{1.31}$$

For $n = 4$ dimensions this includes the most general (non-topological) terms with $N = 4$. These terms could be eliminated via a field redefinition but we choose not to do so here. Since we have $N = 4$ we must have $S^v = s^v$ and $S^A = s^A$ as in our previous examples. Using previous results for s^v, s^A [4, 11, 12] gives

$$S^v = 1 + k_1\ell^2(2R_{rv} - K\bar{K}) + 2k_2\ell^2 R \\ S^A = k_1\ell^2\left(2D_B K^{AB} - D^A K + \mu^{AB}\partial_v\beta_B\right)\,. \tag{1.32}$$

The k_2 term in S^v agrees with the JKM entropy [3]. In this example, the IWW entropy differs from the IW entropy because of the $K\bar{K}$ term and similar terms contained in R_{rv}

JHEP08(2022)258

and R. On C, $R_{rv} \equiv R_{\mu\nu}\ell^\mu n^\nu$ and $K\bar{K}$ are both invariant under a change of GNCs. (This is explained in more detail in section 2.1.) Hence the entropy is gauge invariant, in agreement with our general result. Once again, the entropy current is not gauge-invariant if $k_1 \neq 0$.

We shall not write out equation (1.16) explicitly. However, we emphasise that, for this example, it is necessary to use several components of the Einstein equation to obtain (1.16). In particular the vA component is used to eliminate a term on the r.h.s. quadratic in $\partial_v \beta_A$.

1.11 Discussion

We shall conclude this overview with a discussion of some important open issues, namely the gauge invariance and uniqueness of the entropy.

Section 3 of this paper establishes that the IWW entropy can be defined in a gauge-invariant manner. In more detail, what this means is that if we pick a cross-section C of the horizon and define GNCs based on C then the resulting definition of the IWW entropy of C is the same for *any* choice of these GNCs. Clearly it is important to determine whether our improved entropy (based on S^v) is also gauge-invariant in this sense. In the examples discussed above, i.e. EFTs with up to 4 derivatives, our improved entropy is the same as the IWW entropy and therefore the entropy is gauge-invariant in these examples. But it is unclear whether this remains true if we include terms with 6 or more derivatives.

To understand why this is important, let C, C' be two cross-sections of $\mathcal{H}$ with C' lying entirely to the future of C. We can choose GNCs based on C and normalize the affine parameter along the horizon generators such that C' is given by $v = v_0$ for some $v_0 > 0$. The entropy defined using our approach will then satisfy a second law: $S[C'] \geq S[C]$ (to quadratic order, in the sense of EFT). However, if the entropy is not gauge-invariant then the definition of $S[C]$ depends on the choice of GNCs, and hence on C', which is clearly unsatisfactory.

Our entropy satisfies the second law only to quadratic order in perturbations, in the sense of EFT (i.e. modulo higher derivative terms), so maybe a proof of gauge invariance would also only hold to quadratic order, in the sense of EFT. We leave the construction of such a proof to future work.

Now we turn to uniqueness of the entropy current. The definition of entropy in non-equilibrium thermodynamics is known to suffer from ambiguities. An example is a relativistic viscous fluid. The fluid equations of motion can be viewed as an expansion in increasing numbers of derivatives of the fields and one can define an entropy current also in terms of an expansion in increasing numbers of derivatives of the fields. The aim is to find a definition of the entropy current that satisfies a second law on-shell. It has been shown that such an entropy current is not unique: there are multi-parameter families of entropy currents that satisfy the second law [20–22].

Something similar can be seen for black hole entropy in the perturbative context we are considering. For example, consider conventional vacuum GR. The IWW entropy is simply given by the horizon area: $s^v = 1$. Now consider $\tilde{s}^v = 1 + c\ell^2 K\bar{K}$. Recall that $K = 0$ on $\mathcal{H}$ in the background, and that, on-shell, linear perturbations satisfy (1.12), which here reduces to $\delta K = 0$. Hence we have $\delta(\sqrt{\mu}\tilde{s}^v) = \delta(\sqrt{\mu}s^v)$ on-shell. So, to linear order, s^v

JHEP08(2022)258

and $\tilde{s}^v$ define the same entropy on shell. However, off-shell they differ at linear order. If we restrict to linear perturbations then the only sense in which s^v is preferred over $\tilde{s}^v$ appears to be that the former satisfies the off-shell equation (1.9).

Going beyond linear order we must adopt the EFT perspective. For $n = 4$ vacuum gravity, field redefinitions can be used to eliminate 4-derivative terms from the equations of motion, i.e., the equation of motion is $G_{\mu\nu} + \Lambda g_{\mu\nu} = \mathcal{O}(\ell^4)$. Our "improved" entropy S^v differs from s^v by terms with at least 4 derivatives, which appear at higher order in EFT than the order to which we are working. Hence we have $S^v = s^v = 1$ here. Now we could use $\tilde{s}^v$ as the starting point in our algorithm for improving the entropy and we would then obtain $\tilde{S}^v = \tilde{s}^v = 1 + c\ell^2 K\bar{K}$. On-shell S^v and $\tilde{S}^v$ agree at linear order but they differ at quadratic order. Since c is arbitrary, we therefore have a non-unique entropy current.

Generalising this example, for a general theory we denote the term in square brackets on the r.h.s. of (1.9) as $\partial \cdot s$. Equation (1.12) says that on-shell we have $\delta(\partial \cdot s) = 0$. Now consider

$$\tilde{s}^v = s^v + \mathcal{L}(\partial \cdot s) \qquad \tilde{s}^A = s^A + \mathcal{L}^A(\partial \cdot s) \tag{1.33}$$

where $\mathcal{L}, \mathcal{L}^A$ are arbitrary linear operators built from ∂_v, D_A and the metric components. On-shell we have $\delta(\sqrt{\mu}\tilde{s}^v) = \delta(\sqrt{\mu}s^v)$ as above, so the entropies agree to linear order. But, as we saw above, $\tilde{S}^v$ and S^v will differ at quadratic order. Clearly there is a lot of freedom in the choice of $\mathcal{L}$ and $\mathcal{L}^A$ so there seems to be a lot of freedom in defining an entropy that satisfies a second law in the perturbative sense we have discussed.[10]

This non-uniqueness in the entropy can arise from field redefinitions.[11] In general, a redefinition of the metric would change the location of $\mathcal{H}$ but we can restrict to field redefinitions that are trivial on-shell to avoid having to deal with this. An example is $n = 4$ vacuum gravity viewed as an EFT: as mentioned above we can eliminate 4-derivative terms from the equations of motion to obtain the equations arising from the Lagrangian $L = R + \mathcal{O}(\ell^4)$ (ignoring Λ for simplicity). Starting from this Lagrangian we could perform a field redefinition of the form $g_{\mu\nu} \to g_{\mu\nu} + a_1\ell^2 R_{\mu\nu} + a_2\ell^2 R g_{\mu\nu}$ to obtain a new Lagrangian L' of the form (1.31). This field redefinition is trivial on-shell (as the original equation of motion is $R_{\mu\nu} = \mathcal{O}(\ell^4)$). The IWW entropy resulting from L' is given by (1.32), which differs off-shell from that of L, although they agree on-shell to linear order. Going beyond linear order, the expressions for S^v for the two Lagrangians differ on-shell by a multiple of $K\bar{K}$, so at quadratic order the entropy is different before and after the field redefinition.[12]

[10] We should also note that we can freely adjust terms in S^v, S^A that are of cubic or higher order in positive boost weight quantities whilst continuing to satisfy the second law to quadratic order. This is because such terms vanish to quadratic order.

[11] In a "note added" to the published version of [2], Iyer and Wald criticized their own definition of black hole entropy because it is not invariant under (non-derivative) field redefinitions in an arbitrary theory of gravity and matter.

[12] Note that the procedure just described differs from substituting the field redefinition into the original expression for the entropy [23], which does not change the on-shell entropy because the field redefinition is trivial on-shell.

JHEP08(2022)258

2 Preliminaries

2.1 More on Gaussian Null Coordinates (GNCs)

In a spacetime $(\mathcal{M}, g)$, consider a smooth co-dimension 1 null surface $\mathcal{N}$ that is ruled by affinely parameterized null geodesics with tangent $l = l^\mu \partial_\mu$ such that every null geodesic intersects a spacelike cross section C precisely once. Associated with this structure one can construct GNCs (1.1) as described in section 1.3.

The tensors $\beta_A \mathrm{d}x^A$ and $\mu_{AB}\mathrm{d}x^A\mathrm{d}x^B$ appearing in the Gaussian null form of g (1.1) have an invariant geometric meaning on the cross section C (or more generally on each fixed leaf $C(v)$ of the foliation) of $\mathcal{N}$. The meaning of μ_{AB} is obvious: it is the induced metric on C, with Levi-Civita connection D. To understand the role of β_A, note that the foliation defines a split

$$T_C\mathcal{M} = T_C C \oplus (T_C C)^\perp, \tag{2.1}$$

where the normal bundle $(T_C C)^\perp$ is spanned by the null vectors $n = \partial_r, l = \partial_v$. Now $\boldsymbol{\beta} = \beta_A \mathrm{d}x^A$ can be regarded as a connection 1-form on the 2-dimensional Lorentzian vector bundle $(T_C C)^\perp$. Said differently, on C we can reduce the $\mathrm{SO}(n-1,1)$ principal fibre bundle $F^g\mathcal{M}|_C$ of orthonormal frames defined from the metric g to the product of the $\mathrm{SO}(n-2)$ principal fibre bundle $F^\mu C$ associated with orthonormal frames of $\mu_{AB}\mathrm{d}x^A\mathrm{d}x^B$ and the (trivial) $\mathrm{SO}(1,1)$ principal fibre P bundle of pairs of null directions (rather than vectors) in $(T_C C)^\perp$. The group $\mathrm{SO}(1,1)$ may be identfied with $\mathbb{R}_+$ and it acts in this parameterization on a pair of null vectors n, l by the local rescaling $al, a^{-1}n$, where $a > 0$ is a smooth function on C identified with a local gauge transformation. The representation of $\mathbb{R}_+$ on a line $\mathbb{R}$ given by $\pi_b : a \mapsto a^b$ for a given boost weight b gives rise to a line bundle $P \ltimes_{\pi_b} \mathbb{R}$ over C via the associated vector bundle construction, with corresponding covariant derivative operator $\mathcal{D} = D - \frac{1}{2}b\boldsymbol{\beta}$. If a given tensor field on $\mathcal{M}$ is decomposed into its tangential components along C, its components along l and along n (corresponding to the coordinate indices A, v, r in GNCs), then b corresponds exactly to the boost weight of such a component as described briefly in section 1.3 and more precisely in section 2.2 below.

We shall consider theories including matter fields, assumed to be either scalar fields or abelian p-form fields. In the latter case, we shall fix the gauge as follows. A p-form field $\boldsymbol{A} = A_{\mu_1\ldots\mu_p}\mathrm{d}x^{\mu_1} \wedge \cdots \wedge \mathrm{d}x^{\mu_p}$ transforms under a gauge transformation by a $(p-1)$-form $\boldsymbol{\Lambda}$ as in $\boldsymbol{A} \to \boldsymbol{A} + \mathrm{d}\boldsymbol{\Lambda}$. By a suitable choice of $\boldsymbol{\Lambda}$, we can always achieve that $n \cdot \boldsymbol{A} = 0$ in some neighborhood of $\mathcal{N}$ and $l \cdot \boldsymbol{A} = 0$ on $\mathcal{N}$. For a 1-form field this means that in such a gauge,

$$\boldsymbol{A} = \phi r \mathrm{d}v + A_A \mathrm{d}x^A, \tag{2.2}$$

so the differential $\mathrm{d}r$ does not appear and the other differentials occur in a boost invariant combination.

As mentioned in section 1.3, there is freedom in the choice of GNCs on $\mathcal{N}$. Understanding this freedom will be important in our discussions of gauge-invariance below.[13] The freedom present in the choice of GNCs is: (1) we may start from a different cross

[13] Some of this discussion overlaps with the recent paper [14] which appeared while the current work was in preparation.

JHEP08(2022)258

section, C'; (2) we may choose a different set of coordinates x'^A on C'; (3) we may choose a different affine parameterization. Making such a change implies that the relationship between the GNCs is

$$x^A = x^A(x'^C), \quad v = a(x'^C)v' + b(x'^C) \quad \text{on } \mathcal{N}, \tag{2.3}$$

where a is positive and gives the change of affine parameter and where b corresponds to a change of the reference cross section C' defined by $v' = 0$.

The relationship between $(x^\mu) = (v, r, x^C)$ and $(x'^\mu) = (v', r', x'^C)$ away from $\mathcal{N}$ is in general complicated. Let us, for later purposes, consider only a change of the affine parameter leaving C as it is and keep the coordinates on C as they are (invariance under a change of coordinates on C will be manifest below). Thus, we take $b = 0$ and $x'^C = x^C$ on the cut C for simplicity, so that on $\mathcal{N}$, the change of GNCs is

$$x^C = x'^C, \quad v = a(x'^C)v' \quad \text{on } \mathcal{N}. \tag{2.4}$$

We will determine the relationship between $(x^\mu) = (v, r, x^C)$ and $(x'^\mu) = (v', r', x'^C)$ away from $\mathcal{N}$ order by order in r' in the following way. First note that

$$l' = \frac{\partial}{\partial v'} = a\frac{\partial}{\partial v}, \quad \frac{\partial}{\partial x'^B} = \frac{\partial}{\partial x^B} + v'\frac{\partial a}{\partial x'^B}\frac{\partial}{\partial v} \quad \text{on } \mathcal{N}, \tag{2.5}$$

and then imposing the defining relations for n' on $\mathcal{N}$, $g(n', n') = g(n', \partial/\partial x'^C) = 0, g(l', n') = 1$ gives that

$$n' = a^{-1}\frac{\partial}{\partial r} - v'\mu^{AB}\frac{\partial \log a}{\partial x'^B}\frac{\partial}{\partial x^A} - \frac{av'^2}{2}\mu^{AB}\frac{\partial \log a}{\partial x'^A}\frac{\partial \log a}{\partial x'^B}\frac{\partial}{\partial v} \quad \text{on } \mathcal{N}. \tag{2.6}$$

In other words, since $n' = \partial/\partial r'$ in the primed GNCs,

$$\left(\frac{\partial r}{\partial r'}\right)_{\mathcal{N}} = a^{-1}, \quad \left(\frac{\partial v}{\partial r'}\right)_{\mathcal{N}} = -\frac{av'^2}{2}\mu^{AB}\frac{\partial \log a}{\partial x'^A}\frac{\partial \log a}{\partial x'^B}, \quad \left(\frac{\partial x^A}{\partial r'}\right)_{\mathcal{N}} = -v'\mu^{AB}\frac{\partial \log a}{\partial x'^B}. \tag{2.7}$$

Using (2.4), we then have, in a neighborhood of $\mathcal{N}$,

$$\begin{aligned} r &= \frac{r'}{a(x'^C)} + O(r'^2), \quad v = av' - \frac{ar'v'^2}{2}\mu^{AB}\frac{\partial \log a}{\partial x'^A}\frac{\partial \log a}{\partial x'^B} + O(r'^2) \\ x^C &= x'^C - r'v'\mu^{AB}\frac{\partial \log a}{\partial x'^B} + O(r'^2). \end{aligned} \tag{2.8}$$

Eqs. (2.7), (2.8) are the initial conditions for the geodesic equation for n', which since $n'^\mu = \partial x^\mu/\partial r'$, is

$$0 = \frac{\partial^2 x^\mu}{\partial r'^2} + \Gamma^\mu_{\sigma\rho}(x^\alpha)\frac{\partial x^\sigma}{\partial r'}\frac{\partial x^\rho}{\partial r'}. \tag{2.9}$$

Here, $\Gamma^\mu_{\sigma\rho}$ is the Christoffel symbol of the metric in the coordinates $(x^\alpha) = (r, v, x^C)$ (see appendix A). By integrating the geodesic equations we can determine $x^\mu = x^\mu(r', v', x'^C)$ order by order in r'. Note that the integration is trivial for $v = v' = 0$ and yields

$$r = a(x'^C)^{-1}r', \quad v = 0, \quad x^C = x'^C \quad \text{for } v' = 0. \tag{2.10}$$

JHEP08(2022)258

In general the GNC components of the metric in (r', v', x'^C) coordinates are rather complicated functions of the original metric components. However, some expressions simplify on $\mathcal{N}$ or C. For example using (2.8) one obtains, on $\mathcal{N}$ (i.e. for $r = r' = 0$)

$$\begin{aligned} \mu'_{AB} &= \mu_{AB} \\ \partial_{v'}^N K'_{AB} &= a^{N+1} \partial_v^N K_{AB} \quad \text{on } \mathcal{N}, \end{aligned} \tag{2.11}$$

subject to the identification $(v, x^A) = (a(x'^C)v', x'^A)$ where $K_{AB}, \bar{K}_{AB}$ are the extrinsic curvatures (1.2). The transformation law of β_A, though not of its r-derivatives, is also relatively simple on $\mathcal{N}$,

$$\beta'_A = \beta_A + 2D_A \log a - 2v K_A{}^B D_B \log a \quad \text{on } \mathcal{N}, \tag{2.12}$$

subject to the same identification. The formulae for the transformation of $\bar{K}_{AB}$ and its r-derivatives, or the Riemann tensor $R_{ABCD}[\mu]$ of μ_{AB}, are complicated on $\mathcal{N}$. However, they simplify on C, i.e., for $v = v' = r = r' = 0$:

$$\begin{aligned} D'_{(B_1} \cdots D'_{B_r)} R_{ABCD}[\mu'] &= D_{(B_1} \cdots D_{B_r)} R_{ABCD}[\mu] \\ \partial_{r'}^N \bar{K}'_{AB} &= a^{-N-1} \partial_r^N \bar{K}_{AB} \quad \text{on } C. \end{aligned} \tag{2.13}$$

A quick way to determine the transformation law of a given tensor in GNCs of the form $D_{(B_1} \cdots D_{B_r)} \partial_v^p \partial_r^q \{\mu_{AB}, \beta_A, \alpha\}$ is to transform to one of the covariant bases of monomials described in lemmas 2.2, 2.3 below.

We can think of the above change of GNCs as a diffeomorphism f of some open neighborhood of $\mathcal{N}$ by assigning to a point p with a given set (x^μ) of GNCs a new point $p' = f(p)$ with the same values (x'^μ) in the unique GNC system related by $(v, x^A) = (a(x'^C)v', x'^A)$ on $\mathcal{N}$. Then we can say that $g' = f^* g$, where g' is the metric (1.1) defined by the transformed $\mu'_{AB}, \beta'_A, \alpha'$. It is important to note that, although the restriction of f to $\mathcal{N}$ only refers to the function $a > 0$ on C but not to a particular metric, the diffeomorphism f depends on both a and the metric g off of $\mathcal{N}$, because its construction involves solving for the transverse geodesics relative to the metric g. Thus, we should write $f[a, g]$ to indicate properly the dependencies on a and g. Then we have the cocycle condition

$$f[g, a] \circ f[g', a'] = f[g, aa'], \tag{2.14}$$

where $\circ$ indicates the composition of diffeomorphisms and $g' = f[g, a]^* g$. This relationship can be proven by noting that both sides trivially have the same action on $\mathcal{N}$ which together with the preservation of the Gaussian null form uniquely determines the diffeomorphisms off of $\mathcal{N}$ on both sides.

2.2 Local covariant tensors and GNCs

Consider a tensor field such as $t^{\alpha_1 \ldots \alpha_r}{}_{\beta_1 \ldots \beta_s}[g]$ that is constructed locally and covariantly out of the metric. By the Thomas replacement lemma [2] it is built from contractions of $g_{\mu\nu}, g^{\mu\nu}, \nabla_{(\alpha_1} \cdots \nabla_{\alpha_k)} R_{\mu\nu\alpha\beta}$ (and possibly the volume form). We may evaluate such a tensor in GNCs, laboriously computing the curvature tensors and covariant derivatives in these coordinates. This will lead to, in general complicated, expressions involving $\mu_{AB}, \beta_A, \alpha, r$ and their derivatives, see appendix A for the Riemann and

JHEP08(2022)258

Ricci components, for example. Here we would like to make certain general statements about the resulting expressions. These general features arise from the functional property $t^{\alpha_1\cdots\alpha_r}{}_{\beta_1\ldots\beta_s}[f^*g] = f^*t^{\alpha_1\cdots\alpha_r}{}_{\beta_1\ldots\beta_s}[g]$, taking f to be a diffeomorphism preseriving the Gaussian null form, as described in the previous section.

Recall that a general change of coordinates preserving the Gaussian null form is given on $\mathcal{N}$ by (2.3). We first specialize this equation to $a = 1, b = 0$. Then it is easy to see that, also away from $\mathcal{N}$, the diffeomorphism f is $v = v', r = r', x^A = f^A(x'^C)$. If we now take a GNC component of $t^{\alpha_1\cdots\alpha_r}{}_{\beta_1\ldots\beta_s}$, and view this as a tensor relative to the indices chosen as $\alpha_i = A_i, \beta_j = B_j$ (i.e. those not chosen as r, v), then the resulting tensor is on C a local covariant functional of $\mu_{AB}, \beta_A, \alpha$ and r. In other words, it is a functional which is built out of contractions (with the inverse metric μ^{AB}) of $D_{(A_1}\cdots D_{A_k)}\partial_v^p\partial_r^q\{\mu_{AB}, \alpha, \beta_A\}$ or $D_{(A_1}\cdots D_{A_k)}R_{ABCD}[\mu]$, multiplied by non-negative powers of r. The latter will be absent on $\mathcal{N}$ (as $r = 0$). It is convenient to introduce the following terminology (here $\epsilon[\mu]$ denotes the volume form induced by μ_{AB} on $C(v)$):

Definition 2.1. *A "primitive factor" is one of the following:* μ^{AB}, $\epsilon[\mu]_{A_1\ldots A_{n-2}}$, $D_{(A_1}\cdots D_{A_k)}\partial_v^p\partial_r^q\psi$ *with* $\psi \in \{\mu_{AB}, \alpha, \beta_A\}$, *or* $D_{(A_1}\cdots D_{A_k)}R_{ABCD}[\mu]$. *A "primitive monomial" is a (possibly contracted) product of primitive factors.*

On $\mathcal{N}$, any component of the Riemann tensor (or its derivatives) is a sum of primitive monomials.

If we specialize the change of GNCs defined by (2.4) to the case when $x^C = x'^C$, $b = 0$ and $a > 0$ is *constant* then globally we have $v = av', r = a^{-1}r', x^C = x'^C$. Recall that we say that a quantity X has *boost weight* b if it transforms as in (1.4) under such a change of GNCs. Recall also the simple rule described in section 1.3 for computing the boost weight of a tensor component:

$$\begin{aligned}&\text{boost weight of GNC component of } t^{\alpha_1\cdots\alpha_r}{}_{\beta_1\ldots\beta_s}\\ &\quad = +\#(\text{up } r) - \#(\text{down } r) - \#(\text{up } v) + \#(\text{down } v),\end{aligned} \tag{2.15}$$

A given GNC component of $t^{\alpha_1\cdots\alpha_r}{}_{\beta_1\ldots\beta_s}$ is a sum of monomials built from μ^{AB}-contractions of $D_{(A_1}\cdots D_{A_k)}\partial_v^p\partial_r^q\{\mu_{AB}, \alpha, \beta_A\}$ or $D_{(A_1}\cdots D_{A_k)}R_{ABCD}[\mu]$, multiplied by non-negative powers of r. The boost weight of the various terms appearing in such expressions is given as follows:

Definition 2.2. *The boost weight of* r *is* $+1$, *the boost weight of* v *is* -1; *the boost weight of* $D_{(A_1}\cdots D_{A_k)}\partial_v^p\partial_r^q\psi$ *with* $\psi \in \{\mu_{AB}, \alpha, \beta_A\}$ *is* $p - q$, *the boost weight of* μ^{AB}, $\epsilon[\mu]_{A_1\ldots A_{n-2}}$, $D_{(A_1}\cdots D_{A_k)}R_{ABCD}[\mu]$ *is zero. The boost weight is additive under products.*

The boost weight of each monomial in the expression for a given tensor component is equal to the boost weight of that tensor component. For example, the boost weight of each monomial in the Riemann component R_{rAvB} is zero, as can be seen explicitly by the expressions in appendix A. Likewise, the boost weight of each monomial in the Ricci component R_{rr} is -2, etc.

When matter fields with a tensorial character such as 1-forms $\boldsymbol{A}_J$ are present, then we should first decompose them into GNC components and pick a suitable gauge as e.g.

JHEP08(2022)258

in (2.2) for a 1-form field. In such a gauge, A_{JA}, ϕ_J count as having boost weight zero and are treated on the same footing as the $\mu_{AB}, \beta_A, \alpha$ or scalar fields Φ_I, so in this case tensor components are expressed in terms of derivatives of $\psi \in \{\Phi_I, A_{JA}, \phi_J, \mu_{AB}, \alpha, \beta_A\}$ and non-negative powers of r, and similarly for higher form fields. D_A should be replaced by an appropriate charged covariant derivative using $A_{BJ}\mathrm{d}x^B$ on any charged scalar fields. Quantities descending from a diffeomorphism and gauge covariant quantity remain gauge covariant under the restricted gauge transformations preserving the gauge (2.2), that is Λ_J that do not depend on r, v. We will usually suppress discussion of matter fields below and only comment on then where they lead to important differences.

The notion of boost weight depends on the chosen cut C and it refers to a particular choice of GNCs. If we perform a non-trivial change of GNCs, corresponding to non-constant a, then the definition of boost weight w.r.t. the new GNCs will not agree with the definition w.r.t. the old GNCs. However, (for fixed C) the definitions will agree on $\mathcal{N}$. More precisely, let $\psi'(x'^\mu) \in \{\mu'_{A'B'}(x'^\mu), \beta'_{A'}(x'^\mu), \alpha'(x'^\mu)\}$ be the quantities defined by the components of the metric g w.r.t. a second set of GNCs x'^μ. Then we have the following lemma.

Lemma 2.1. *On $\mathcal{N}$, the definition of boost weight is independent of the choice of GNCs. In other words, the expressions for $D'_{(A_1}\cdots D'_{A_j)}\partial_{v'}^{p'}\partial_{r'}^{q'}\psi'(x'^\mu)$ in terms of $D_{(A_1}\cdots D_{A_k)}a(x^C)$, $D_{(A_1}\cdots D_{A_j)}\partial_v^p\partial_r^q\psi(x^\mu)$ and powers of v' will only contain terms of boost weight $p'-q'$, with v' counting as boost weight -1 and a counting as boost weight 0.*

Proof. On $\mathcal{N}$, the quantities $\alpha, \beta_A, \mu_{AB}$ and their derivatives can all be written as $\partial_{\mu_1}\ldots\partial_{\mu_N}g_{\nu\rho}$ for some choice of indices (e.g. $\beta_A = -\partial_r g_{vA}$ on $\mathcal{N}$). If we consider the corresponding quantity $\partial_{\mu'_1}\ldots\partial_{\mu'_N}g_{\nu'\rho'}$ in another set of GNCs then this will be given by some contraction of $\partial_{\mu_1}\ldots\partial_{\mu_N}g_{\sigma\tau}$ with expressions of the form $J^\mu_{\nu'_1\ldots\nu'_p} \equiv \partial^p x^\mu/\partial x'^{\nu_1}\ldots\partial x'^{\nu_p}$ for $p = 1, \ldots, N$. We claim that, on $\mathcal{N}$, each component of such an expression has a definite boost weight determined by the same rule described above for tensor components, i.e., by counting the number of up and down indices of each type. (For example $J^r_{v'A'}|_{\mathcal{N}}$ will have boost weight $1+1+0=2$.) Hence, when contracted with $\partial_{\mu_1}\ldots\partial_{\mu_N}g_{\nu\rho}$, the additivity of boost weight under products will ensure that the result will have a definite boost weight. An example of this can be seen in (2.12) where both the l.h.s. and r.h.s. have boost weight zero (since we assign a boost weight zero).

To see why each component of $J^\mu_{\nu'_1\ldots\nu'_p}|_{\mathcal{N}}$ has definite boost weight, we proceed inductively. Consider first the case where $\nu'_i \neq r'$ for all i. The result is then immediate from (2.4) (and $r'=0$ on $\mathcal{N}$). For example $(\partial^2 v/\partial v'\partial x'^A)_{\mathcal{N}} = \partial'_A a$, which has boost weight zero, in agreement with the above rules applied to the l.h.s. Next assume exactly one of the ν'_i is r'. For $p=1$ the result is immediate from (2.7), where each term on the r.h.s. has the correct boost weight to match the above rules applied to the l.h.s. Taking derivatives w.r.t. v' and $x^{A'}$ respects these rules and so this result extends to any p. If exactly two of the ν'_i are r' then, for $p=2$, the result follows by evaluating (2.9) on $\mathcal{N}$ and using the fact that the components of the Christoffel symbols have definite boost weight. The result for $p>2$ follows by taking v' and $x^{A'}$ derivatives of (2.9). If exactly three of the ν'_i are r' then one takes another r' derivative of (2.9) and evaluates on $\mathcal{N}$ and so on. □

JHEP08(2022)258

Since we will be dealing with equations involving higher derivatives, we will sometimes need to keep track of the number of derivatives associated with a given quantity. We are mostly interested in doing this on $\mathcal{N}$. In the above proof we saw that, on $\mathcal{N}$, the quantities of interest can all be expressed in terms of partial derivatives of the metric tensor: α involves 2 r-derivatives and β_A involves 1 r-derivative. We make the following definition to count derivatives:

Definition 2.3. *The "dimension" of α, β_A and μ_{AB} are $2, 1, 0$ respectively. Taking a derivative w.r.t. v, r or x^A increases the dimension by 1. Dimension is additive under products.*

For example $D_{(A_1} \cdots D_{A_k)} \partial_v^p \partial_r^q \alpha$ has dimension $k + p + q + 2$. A quantity with boost weight b involves at least $|b|$ derivatives and so its dimension is bounded below by $|b|$. The dimension of a component of the kth covariant derivative of the Riemann tensor is $k + 2$.

The primitive factors of definition 2.1 are employed in Wall's method. However, in various places we will find it more convenient to express quantities on $\mathcal{N}$ in terms of a different set of quantities. These are described in lemmas 2.2 and 2.3.

Lemma 2.2. *On $\mathcal{N}$, any of the expressions $D_{(A_1} \cdots D_{A_k)} \partial_v^p \partial_r^q \psi$ with $\psi \in \{\mu_{AB}, \alpha, \beta_A\}$ can be expressed uniquely as a sum of monomials where each monomial is a product of factors of the following form (and possible factors of μ^{AB})*

- $D_{(A_1} \cdots D_{A_j)} \partial_v^N K_{AB}$, $D_{(A_1} \cdots D_{A_k)} \partial_r^{\bar{N}} \bar{K}_{AB}$,
- $D_{(A_1} \cdots D_{A_j)} \beta_B$,
- $D_{(A_1} \cdots D_{A_j)} R[\mu]_{BCDE}$. *Note that, by the Bianchi identities, these components are not all independent. When $j \geq 1$, an independent set is obtained e.g. by choosing $D_{(A_1} \cdots D_{A_j} R[\mu]^B{}_{A_{j+1} A_{j+2})C}$.*[14]
- *A GNC component of $\nabla_{(\alpha_1} \cdots \nabla_{\alpha_j)} R_{\mu\nu}$, except for: $(\mu\nu) = (rr)$ and $\alpha_j \in \{r, A\}$ or $(\mu\nu) = (vv)$ and $\alpha_j \in \{v, A\}$, and except for GNC components that are linearly dependent on a suitably chosen minimal set via the Bianchi identities.*

Proof. First consider the case of $\psi = \alpha$. To express $D_{(A_1} \cdots D_{A_k)} \partial_v^p \partial_r^q \alpha$ in the desired form, we start by writing the Ricci component R_{vr} in terms of GNC components (see appendix A)

$$R_{vr} = -\frac{1}{2} \partial_r^2 \left(r^2 \alpha \right) - \frac{1}{2} \bar{K} \partial_r \left(r^2 \alpha \right) + \ldots \tag{2.16}$$

where the ellipsis denotes terms that do not depend on α. Evaluating this equation at $r = 0$ uniquely determines $\alpha|_{r=0}$ in terms of β_A, $D_A \beta_B$ and quantities of the form $\partial_v^{p'} \partial_r^{q'} \mu_{AB}$ with $p', q' \leq 1$. Next, we act on the above equation with ∂_r and use[15]

$$\partial_r R_{rv} = \nabla_r R_{rv} + \Gamma^A_{rv} R_{rA} + \ldots \tag{2.17}$$

[14] In fact, [26] shows that in Riemann normal coordinates any number of partial derivatives of the metric can be expressed using this set of quantities.

[15] To clarify, in this equation $\nabla_r R_{rv}$ means the rrv component of the tensor obtained by taking the covariant derivative of the Ricci tensor. Of course, the component $\nabla_r R_{rv}$ depends on multiple Ricci components $R_{\mu\nu}$, not just on R_{rv}. We continue to use similar conventions hereafter. Note, however, that covariant derivatives mix up components only at the subprincipal level.

JHEP08(2022)258

where the ellipsis denotes terms that vanish at $r = 0$. By virtue of the identities of appendix A, this can be used to determine $\partial_r \alpha|_{r=0}$ in terms of quantities of the following form: $D_A \partial_r^{q'} \beta_B$ with $q' \leq 2$; $\partial_v^{p''} \partial_r^{q''} \mu_{AB}$ with $p'' \leq 1$, $q'' \leq 2$; and the GNC component $\nabla_r R_{rv}$. Proceeding inductively, we assume that we can express $\partial_r^{q-1} \alpha|_{r=0}$ in terms of the following quantities:

- $D_A \partial_r^{q'} \beta_B$ with $q' \leq q-1$, and
- $\partial_v^{p''} \partial_r^{q''} \mu_{AB}$ with $p'' \leq 1$ $q'' \leq q$, and
- a GNC component $\nabla_{\alpha_1} \cdots \nabla_{\alpha_j} R_{\mu\nu}$ with $j \leq q-1$.

We have already shown that this is true for $q = 1$ and $q = 2$. Now we act on (2.16) with ∂_r^q and evaluate at $r = 0$. Then we determine $\partial_r^q \alpha|_{r=0}$ in terms of primitive factors of the form $D_A \partial_r^{q'} \beta_B$ with $q' \leq q$, $\partial_v^{p''} \partial_r^{q''} \mu_{AB}$ with $p'' \leq 1$, $q'' \leq q+1$, and $\partial_r^q R_{rv}$. One can then rewrite $\partial_r^q R_{rv}$ in terms of GNC components $\nabla_{\alpha_1} \cdots \nabla_{\alpha_j} R_{\mu\nu}$ with $j \leq q$ and terms of the form $\partial_r^{\bar{q}} \Gamma^{\mu}_{r\nu}$ with $\bar{q} < q$. It follows from the formulae of appendix A that the terms involving Christoffel symbols can be written in terms of primitive factors of the form $\partial_r^{q'} \beta_A$ with $q' \leq q$, $\partial_r^{q''} \mu_{AB}$ with $q'' \leq q+1$ and $\partial_r^{q'''} \alpha$ with $q''' \leq q-1$. However, the terms of the form $\partial_r^{q'''} \alpha$ with $q''' \leq q-1$ can be eliminated in favour of derivatives of β_A, μ_{AB} and covariant derivatives of $R_{\mu\nu}$, as specified above. This shows that if the induction hypothesis holds then $\partial_r^q \alpha|_{r=0}$ can be expressed with the quantities

- $D_A \partial_r^{q'} \beta_B$ with $q' \leq q$, and
- $\partial_v^{p''} \partial_r^{q''} \mu_{AB}$ with $p'' \leq 1$ $q'' \leq q+1$, and
- a GNC component $\nabla_{\alpha_1} \cdots \nabla_{\alpha_j} R_{\mu\nu}$ with $j \leq q$.

Since this statement is true for $q = 1$, it follows that it is true for general q.

By taking v or x^A derivatives of this result, we determine $D_{A_1} D_{A_2} \ldots D_{A_m} \partial_v^p \partial_r^q \alpha|_{r=0}$ as a sum of products of factors depending only on

- $D_{A_1} D_{A_2} \ldots D_{A_{m'}} \partial_v^{p'} \partial_r^{q'} \beta_A$ with $q' \leq q$, and
- $D_{A_1} D_{A_2} \ldots D_{A_{m''}} \partial_v^{p''} \partial_r^{q''} \mu_{AB}$ with $q'' \leq q+1$, and
- a GNC component $\nabla_{\alpha_1} \cdots \nabla_{\alpha_j} R_{\mu\nu}$ with $j \leq m+p+q$.

Each monomial in this sum must have boost weight $p-q$. In the rest of this argument it is understood that whenever $D_{A_1} D_{A_2} \ldots D_{A_m} \partial_v^p \partial_r^q \alpha|_{r=0}$ appears, we will rewrite it in terms of the above quantities.

The next step is to show that we can use similar methods to eliminate certain derivatives of β_A. First we consider the following expression for R_{rA} (see appendix A)

$$R_{rA} = -\frac{1}{2}\partial_r^2 (r\beta_A) - \frac{1}{2}\bar{K}\partial_r (r\beta_A) + \bar{K}_A{}^B \partial_r (r\beta_B) + r\beta^B \left(\partial_r \bar{K}_{AB} + \ldots\right) + \ldots \tag{2.18}$$

JHEP08(2022)258

where the first ellipsis on the r.h.s. denotes terms quadratic in $\bar{K}_{AB}$ and the second ellipsis denotes terms linear in $D_A\bar{K}_{BC}$ (contracted in some way with μ^{AB} in both cases). Evaluating this equation at $r=0$ determines $\partial_r\beta_A|_{r=0}$ in terms of β_A, μ_{AB}, $\partial_r\mu_{AB}$, $D_C\partial_r\mu_{AB}$ and R_{rA}. Acting with ∂_r^{q-1} for $q\geq 1$ and evaluating at $r=0$ lets us determine, inductively (similarly to the $\psi=\alpha$ case above), $\partial_r^q\beta_A|_{r=0}$ in terms of β_A, $D_C\partial_r^{q'}\mu_{AB}$ with $q'\leq q$ and GNC components $\nabla_{\alpha_1}\dots\nabla_{\alpha_j}R_{\mu\nu}$ with $j\leq q$.

Next we consider R_{vA} evaluated at $r=0$:

$$R_{vA}|_{r=0}=\frac{1}{2}\partial_v\beta_A+\frac{1}{2}K\beta_A-D_AK+D_BK_A{}^B\,. \tag{2.19}$$

This determines $\partial_v\beta_A|_{r=0}$ in terms of β_A, μ_{AB}, $\partial_v\mu_{AB}$, $D_C\partial_v\mu_{AB}$ and R_{vA}. Acting with ∂_v^{p-1} for $p\geq 1$ lets us determine, inductively, $\partial_v^p\beta_A|_{r=0}$ in terms of β_A, terms of the form $D_C\partial_v^{p'}\mu_{AB}$ with $p'\leq p$ and $\partial_v^{p-1}R_{vA}$. The last one of these can then be written in terms of GNC components $\nabla_{\alpha_1}\dots\nabla_{\alpha_j}R_{\mu\nu}$ with $j\leq p-1$ and $\partial_v^{\bar{p}}\Gamma^{\mu}_{v\nu}\big|_{r=0}$ with $\bar{p}<p$. Employing the expressions of appendix A relating Christoffel symbols to primitive factors then gives a formula for $\partial_v^p\beta_A|_{r=0}$ in terms of β_A, terms of the form $D_C\partial_v^{p'}\mu_{AB}$ with $p'\leq p$ and GNC components $\nabla_{\alpha_1}\dots\nabla_{\alpha_j}R_{\mu\nu}$ with $j\leq p-1$.

Now we return to (2.18): taking a v-derivative of this equation and using our result for $\partial_v\beta_A$, we can write $\partial_v\partial_r\beta_A|_{r=0}$ in terms of R_{vA}, ∇_vR_{rA}, β_A, μ_{AB}, and terms of the form $\partial_v^{p'}\partial_r^{q'}\mu_{AB}$ and $D_C\partial_v^{p'}\partial_r^{q'}\mu_{AB}$ with $p'\leq 1$ and $q'\leq 1$. Similarly, taking multiple v-derivatives of (2.18) and employing an induction on the number of v-derivatives gives us an expression for $\partial_v^p\partial_r\beta_A|_{r=0}$ containing the following quantities: β_A, primitive factors with $\psi=\mu_{AB}$ and components of the form $\nabla_{\alpha_1}\dots\nabla_{\alpha_j}R_{\mu\nu}$ with $j\leq p$. Furthermore, taking v-derivatives of our previous expressions for $\partial_r^q\beta|_{r=0}$ and proceeding inductively as before we obtain an expression for $\partial_v^p\partial_r^q\beta_A|_{r=0}$ $(p,q\geq 1)$ in terms of β_A, μ_{AB}, primitive factors of the form $D_{A_1}\dots D_{A_m}\partial_v^{p'}\partial_r^{q'}\mu_{AB}$ with $p'\leq p$, $q'\leq q$, and components $\nabla_{\alpha_1}\dots\nabla_{\alpha_j}R_{\mu\nu}$. Taking derivatives w.r.t. x^A of this result lets us write $D_{A_1}\dots D_{A_m}\partial_v^p\partial_r^q\beta_B|_{r=0}$ in terms of

- $D_{A_1}\dots D_{A_n}\beta_B$ with $0\leq n\leq m$,
- $D_{A_1}\dots D_{A_{n'}}\partial_v^{p'}\partial_r^{q'}\mu_{AB}$ with $p'\leq p$, $q'\leq q$
- and covariant components $\nabla_{\alpha_1}\dots\nabla_{\alpha_j}R_{\mu\nu}$ with $j\leq m+p+q$.

From now on, it is assumed that at every occurrence of $D_{A_1}\dots D_{A_m}\partial_v^p\partial_r^q\beta_B|_{r=0}$ it is expressed with the above quantities.

Consider finally a factor of the form $D_{A_1}\dots D_{A_k}\partial_v^p\partial_r^q\mu_{CD}$. We wish to argue that this factor can be written as a sum of monomials whose factors are of the form specified in the statement of the lemma. To this end, we consider the expression for the Ricci component R_{AB} in terms of GNCs (given in more detail in appendix A)

$$\begin{aligned}R_{AB}=&-2\partial_v\bar{K}_{AB}-K\bar{K}_{AB}+4K_{(A}{}^C\bar{K}_{B)C}-K_{AB}\bar{K}\\&+R[\mu]_{AB}-D_{(B}\beta_{A)}-\frac{1}{2}\beta_A\beta_B+\dots\end{aligned} \tag{2.20}$$

JHEP08(2022)258

where the ellipsis now stands for terms that vanish at $r = 0$. This equation allows us to express $\partial_v \bar{K}_{AB}|_{r=0}$ in terms of R_{AB}, $R[\mu]_{AB}$, K_{AB}, $\bar{K}_{AB}$, β_A and $D_A\beta_B$. Next, we act with ∂_r on equation (2.20) and use

$$\partial_r R_{AB} = \nabla_r R_{AB} + 2\Gamma^{\mu}_{r(A} R_{B)\mu} = -2\partial_v \partial_r \bar{K}_{AB} - \beta_{(A}\partial_r \beta_{B)} - \partial_r D_{(A}\beta_{B)} - 2\alpha \bar{K}_{AB} + \dots . \tag{2.21}$$

Here the ellipsis stands for terms that are of the required form and terms involving $\partial_v \bar{K}_{AB}$ that can be written in the required form as described in the previous paragraph. As for the term involving α, we have shown it previously that it can be expressed (at $r = 0$) in terms of R_{rv}, μ_{AB}, K_{AB}, $\bar{K}_{AB}$, β_A, $D_A\beta_B$ and $\partial_v \bar{K}_{AB}$. Combining this with our result on $\partial_v \bar{K}_{AB}$ yields a formula for α in terms of R_{AB}, R_{rv}, μ_{AB}, $R[\mu]_{AB}$, K_{AB}, $\bar{K}_{AB}$, β_A and $D_A\beta_B$. Regarding the terms involving $\partial_r\beta_A$ and $\partial_r D_A\beta_B$ in (2.21), we have a prescription to write these (at $r = 0$) in terms of β_A, $D_A\beta_B$, μ_{AB}, $\bar{K}_{AB}$, $D_{A_1}\bar{K}_{AB}$, $D_{A_1}D_{A_2}\bar{K}_{AB}$ and covariant components involving the Ricci tensor. Therefore, equation (2.21) provides an expression for $\partial_v\partial_r \bar{K}_{AB}|_{r=0}$ as a sum of monomials that are products of factors of the following quantities: μ_{AB}, β_A, $D_A\beta_B$, K_{AB}, $\bar{K}_{AB}$, $D_{A_1}\bar{K}_{AB}$, $D_{A_1}D_{A_2}\bar{K}_{AB}$, $\partial_r\bar{K}_{AB}$, $R[\mu]_{AB}$ and covariant quantities of the form $R_{\mu\nu}$, $\nabla_\alpha R_{\mu\nu}$.

Proceeding inductively, one can fix $\partial_v\partial_r^q \bar{K}_{AB}|_{r=0}$ (by taking multiple derivatives of (2.20) w.r.t. r and using previous results) in terms of the quantities

- μ_{AB}, $R[\mu]_{AB}$,
- K_{AB}, $\partial_r^{q'}\bar{K}_{AB}$, $D_{A_1}\partial_r^{q'}\bar{K}_{AB}$, $D_{A_1}D_{A_2}\partial_r^{q'}\bar{K}_{AB}$ with $q' \leq q$
- factors of β_A and $D_A\beta_B$
- covariant components of the form $\nabla_{\alpha_1}\dots\nabla_{\alpha_j}R_{\mu\nu}$ with $j \leq q$

Similarly, an inductive argument establishes that taking multiple v-derivatives of our expressions for $\partial_v\partial_r^q\bar{K}_{AB}|_{r=0}$ allows us to write $\partial_v^p\partial_r^q\mu_{AB}|_{r=0}$ using only the factors specified below:

- μ_{AB}, $R[\mu]_{AB}$,
- $\partial_v^{p'}K_{AB}$, $D_{A_1}\partial_v^{p'}K_{AB}$, $D_{A_1}D_{A_2}\partial_r^{p'}K_{AB}$ with $p' \leq p-1$
- $\partial_r^{q'}\bar{K}_{AB}$, $D_{A_1}\partial_r^{q'}\bar{K}_{AB}$, $D_{A_1}D_{A_2}\partial_r^{q'}\bar{K}_{AB}$ with $q' \leq q-1$
- factors of β_A and $D_A\beta_B$
- covariant components of the form $\nabla_{\alpha_1}\dots\nabla_{\alpha_j}R_{\mu\nu}$ with $j \leq p+q-2$

Taking D_A derivatives of this result then lets us write $D_{(A_1}\dots D_{A_k)}\partial_v^p\partial_r^q\mu_{AB}$ in terms of the factors

- μ_{AB}, $D_{A_1}\dots D_{A_j}R[\mu]_{AB}$ with $0 \leq j \leq k$,
- $D_{A_1}\dots D_{A_j}\partial_v^{p'}K_{AB}$ with $0 \leq j \leq k+2$ and $p' \leq p-1$

JHEP08(2022)258

- $D_{A_1} \dots D_{A_j} \partial_v^{q'} \bar{K}_{AB}$ with $0 \leq j \leq k+2$ and $p' \leq p-1$
- $D_{A_1} \dots D_{A_j} \beta_A$ with $0 \leq j \leq k+1$
- covariant components of the form $\nabla_{\alpha_1} \dots \nabla_{\alpha_j} R_{\mu\nu}$ with $j \leq p+q+m-2$

Combining this with our previous results on primitive factors of α and β, we can determine $D_{(A_1} \dots D_{A_k)} \partial_v^p \partial_r^q \psi|_{r=0}$ with $\psi \in \{\mu_{AB}, \alpha, \beta_A\}$ as a sum of monomials with each monomial being a product of factors

(i) μ_{AB}, $D_{A_1} \dots D_{A_j} R[\mu]_{AB}$

(ii) $D_{A_1} \dots D_{A_j} \partial_v^{p'} K_{AB}$ and $D_{A_1} \dots D_{A_j} \partial_v^{q'} \bar{K}_{AB}$

(iii) $D_{A_1} \dots D_{A_j} \beta_A$

(iv) covariant components of the form $\nabla_{\alpha_1} \dots \nabla_{\alpha_j} R_{\mu\nu}$

To obtain the same dependencies as in the statement of the lemma, as well as uniqueness, we first note that (on $\mathcal{N}$) the covariant quantities $\nabla_{\alpha_1} \dots \nabla_{\alpha_j} R_{vv}$ with $\alpha_j \in \{v, A\}$ and $\nabla_{\alpha_1} \dots \nabla_{\alpha_j} R_{rr}$ with $\alpha_j \in \{r, A\}$ can be determined by all the other factors (i)-(iv) listed above. To see this, we consider the identity (see appendix A)

$$R_{rr} = -\partial_r \bar{K} - \bar{K}_{AB} \bar{K}^{AB}. \tag{2.22}$$

Hence, R_{rr} is clearly expressible in terms of $\bar{K}_{AB}$ and $\partial_r \bar{K}_{AB}$. Taking derivatives of this result w.r.t. r and x^A, one can inductively express $\nabla_{\alpha_1} \dots \nabla_{\alpha_j} R_{rr}$ with $\alpha_j \in \{r, A\}$ in terms of factors of the form $D_{A_1} \dots D_{A_j} \partial_v^{p'} K_{AB}$ and other factors listed in (i)-(iv). A very similar argument establishes that $\nabla_{\alpha_1} \dots \nabla_{\alpha_j} R_{vv}$ with $\alpha_j \in \{v, A\}$ is also redundant in the list of factors above.

Furthermore, we may express all the quantities listed in (i)-(iv) with totally symmetrized derivatives such as $D_{(A_1} \dots D_{A_j)}$. The reason for this is that an anti-symmetrization of a quantity $D_{A_1} \dots D_{A_j} \Psi$ over a subset of the indices $A_1, \dots A_j$ is expressible via the Riemann tensor $R[\mu]_{ABCD}$ and fewer than j derivatives of the quantity Ψ (using the Bianchi identities). Similarly, the quantities $\nabla_{\alpha_1} \dots \nabla_{\alpha_j} R_{\mu\nu}$ can be expressed with the totally symmetrized derivatives $\nabla_{(\alpha_1} \dots \nabla_{\alpha_m)} R_{\mu\nu}$ and (covariant derivatives of) the Riemann tensor associated with ∇. To eliminate the dependency on $R_{\mu\nu\rho\sigma}$, one can use the formulae of appendix A and the procedure described above to write the components of the Riemann tensor (at $r = 0$) in terms of μ_{AB}, β_A, $D_A \beta_B$, K_{AB}, $\bar{K}_{AB}$, $\partial_v K_{AB}$, $\partial_r \bar{K}_{AB}$, $D_C K_{AB}$, $D_C \bar{K}_{AB}$, $R[\mu]_{AB}$ and Ricci components $R_{\mu\nu}$. Then one can argue that any derivative of the Riemann tensor can be expressed in terms of the quantities listed in the statement of the lemma by using induction on the number of derivatives and by making use of identities of the form

$$\nabla_{\alpha_1} \nabla_{(\alpha_2} \dots \nabla_{\alpha_j)} R_{\mu\nu} = \nabla_{(\alpha_1} \nabla_{\alpha_2} \dots \nabla_{\alpha_j)} R_{\mu\nu} + \dots \tag{2.23}$$

where the ellipsis on the r.h.s. stands for terms involving fewer than j derivatives of the curvature tensor.

JHEP08(2022)258

Finally, the Bianchi identities give us a freedom when expressing a primitive factor $D_{(A_1}\dots D_{A_j)}\partial_v^p\partial_r^q\psi$ in terms of the factors in the statement of the lemma. It can be seen that this is all the remaining freedom that one has using induction, because the leading derivative terms of the Bianchi identities cancel and these are the only possible linear dependencies among the leading derivative terms used in our induction argument. □

Lemma 2.3. *Any of the expressions $D_{(A_1}\cdots D_{A_k)}\partial_v^p\partial_r^q\psi$ with $\psi\in\{\mu_{AB},\alpha,\beta_A\}$ can be expressed (on $\mathcal{N}$) as a sum of monomials where each monomial is a product of factors of the following form (and possible factors of μ^{AB})*

- $D_{(A_1}\cdots D_{A_j}K_{A)B}$, $D_{(A_1}\cdots D_{A_j}\bar{K}_{A)B}$,
- $D_{(A_1}\cdots D_{A_j}\beta_{B)}$,
- *A GNC component of $\nabla_{(\alpha_1}\cdots\nabla_{\alpha_j)}R_{\mu\nu\sigma\rho}$. By the Bianchi-identities, not all these components are independent. An independent set for $j\geq 1$ can be obtained by choosing e.g. $\nabla_{(\alpha_1}\cdots\nabla_{\alpha_j}R^{\mu}{}_{\alpha_{j+1}\alpha_{j+2})\nu}$.*

Proof. By lemma 2.2, we only need to eliminate $D_{(A_1}\cdots D_{A_j)}\partial_v^N K_{AB}$, $D_{(A_1}\cdots D_{A_k)}\partial_r^{\bar{N}}\bar{K}_{AB}$ in favor of $D_{(A_1}\cdots D_{A_j}K_{A)B}$, $D_{(A_1}\cdots D_{A_j}\bar{K}_{A)B}$, we need to eliminate $D_{(A_1}\cdots D_{A_j)}\beta_B$ in favor of $D_{(A_1}\cdots D_{A_j}\beta_{B)}$, and we need to eliminate $D_{(A_1}\cdots D_{A_j)}R_{ABCD}[\mu]$, modulo GNC components of $\nabla_{(\alpha_1}\cdots\nabla_{\alpha_j)}R_{\mu\nu\sigma\rho}$.

First we consider the evolution equation for the expansion and shear in the r-direction, which is

$$R_{rArB}=\bar{K}_A{}^C\bar{K}_{BC}-\partial_r\bar{K}_{AB}, \tag{2.24}$$

which lets us eliminate $\partial_r\bar{K}_{AB}$ in terms of $\bar{K}_{AB}$, plus a GNC component of the Riemann tensor. Next, we consider, on $\mathcal{N}$, the Riemman component

$$R_{rvAB}=2K_{[A}{}^C\bar{K}_{B]C}+D_{[A}\beta_{B]}, \tag{2.25}$$

which lets us write $D_A\beta_B$ in terms of $D_{(A}\beta_{B)}$ plus terms containing K_{AB}, $\bar{K}_{AB}$ or a GNC component of the Riemann tensor. The evolution equation for the expansion and shear in the v-direction is on $\mathcal{N}$ given by

$$R_{vAvB}=K_A{}^CK_{BC}-\partial_vK_{AB}, \tag{2.26}$$

which lets us eliminate ∂_vK_{AB} in terns of K_{AB}, plus a GNC component of the Riemann tensor. Then we consider the Gauss-Codacci equation on $\mathcal{N}$,

$$R_{ABCD}=R[\mu]_{ABCD}+2K_{B[C}\bar{K}_{D]A}+2K_{A[D}\bar{K}_{C]B} \tag{2.27}$$

which lets us eliminate $R[\mu]_{ABCD}$ in favor of terms containing K_{AB}, $\bar{K}_{AB}$ or a GNC component of the Riemann tensor. Finally, we have

$$\begin{aligned}R_{ABrC}&=-\left(D_A+\frac{1}{2}\beta_A\right)\bar{K}_{BC}+\left(D_B+\frac{1}{2}\beta_B\right)\bar{K}_{AC}\\R_{ABvC}&=-\left(D_A-\frac{1}{2}\beta_A\right)K_{BC}+\left(D_B-\frac{1}{2}\beta_B\right)K_{AC}.\end{aligned} \tag{2.28}$$

JHEP08(2022)258

This lets us eliminate $D_{[A}K_{B]C}, D_{[A}\bar{K}_{B]C}$ in favor of terms containing K_{AB}, $\bar{K}_{AB}, \beta_A$ or a GNC component of the Riemann tensor

We now continue this process by taking suitable derivatives of the above equations, using the non-zero Christoffel symbols (on $\mathcal{N}$) given in appendix A.

First we show inductively that $\partial_r^{\bar{N}}\bar{K}_{AB}$ can be expressed with only $\bar{K}_{AB}$ and covariant components of the form $\nabla_{(\alpha_1}\cdots\nabla_{\alpha_j)}R_{\mu\nu\sigma\rho}$ with $j \leq \bar{N}-1$. We have already shown above that this is true for $\bar{N}=1$. Now assume that $\partial_r^{\bar{q}}\bar{K}_{AB}$ with $q \leq \bar{N}-1$ can be expressed using $\bar{K}_{AB}$ and GNC components $\nabla_{(\alpha_1}\cdots\nabla_{\alpha_j)}R_{\mu\nu\sigma\rho}$ with $j \leq \bar{q}-1$. Now act $\bar{N}$ times with ∂_r on (2.24), noting that (2.24) holds off $\mathcal{N}$. We can convert any partial derivatives of Riemann tensor components to components of covariant derivatives of the Riemann tensor. This lets us eliminate $\partial_r^{\bar{N}}\bar{K}_{AB}$ in terms of $\partial_r^{q}\bar{K}_{AB}$ with $0 \leq q \leq \bar{N}-1$, $\partial_r^{q'}\Gamma^A_{rB}$ with $0 \leq q' \leq \bar{N}-1$, plus GNC components of the Riemann tensor. Note that other Christoffel symbols cannot appear in this expression since Γ^μ_{rr} and $\Gamma^v_{r\mu}$ vanish (on and off $\mathcal{N}$). Furthermore, $\Gamma^A_{rB} = \bar{K}^A{}_B$ on and off $\mathcal{N}$. By the induction hypothesis we can eliminate the terms $\partial_r^q\bar{K}_{AB}$ in favor of $\bar{K}_{AB}$ and covariant derivatives of the Riemann tensor, closing the induction loop.

A very similar inductive argument lets us write $\partial_v^N K_{AB}$ in terms of K_{AB} and GNC components $\nabla_{(\alpha_1}\cdots\nabla_{\alpha_j)}R_{\mu\nu\sigma\rho}$ with $j \leq N-1$. This argument is based on acting with ∂_v multiple times on (2.26) and noting that $\Gamma^r_{v\mu} = \Gamma^\mu_{vv} = 0$ and $\Gamma^A_{vB} = K^A{}_B$ on $\mathcal{N}$.

Next, we express the factors $D_{(A_1}\cdots D_{A_j)}K_{AB}$, $D_{(A_1}\cdots D_{A_j)}\bar{K}_{AB}$, $D_{(A_1}\cdots D_{A_j)}\beta_B$ and $D_{(A_1}\cdots D_{A_{j-1})}R_{ABCD}[\mu]$ in terms of the quantities listed in the statement of the lemma by using induction on j. These expressions have been obtained for $j=1$ above. Now suppose we have the corresponding expressions for $D_{(A_1}\cdots D_{A_{j-1})}K_{AB}$, $D_{(A_1}\cdots D_{A_{j-1})}\bar{K}_{AB}$, $D_{(A_1}\cdots D_{A_{j-1})}\beta_B$ and $D_{(A_1}\cdots D_{A_{j-2})}R_{ABCD}[\mu]$. We next use an identity of the form

$$\begin{aligned} D_{(A_1}\ldots D_{A_j)}K_{BC} = &-\frac{2j}{j+1}\left(D_{(A_1}\ldots D_{A_{j-1})}D_{[A_j}K_{B]C} + \ldots + D_{(A_2}\ldots D_{A_j)}D_{[A_1}K_{B]C}\right) \\ &+D_{(A_1}\ldots D_{A_j}K_{B)C} + \ldots \end{aligned} \tag{2.29}$$

where the ellipsis in the second line stands for a sum of monomials containing factors of the form $D_{(A_1}\ldots D_{A_m)}R[\mu]_{ABCD}$ with $m \leq j-2$ and a factor $D_{(A_1}\ldots D_{A_k)}K_{BC}$ or a factor $D_{(A_1}\ldots D_{A_{k-1})}D_{[A_k}K_{B]C}$ with $k \leq j-2$. Then we consider $\nabla_{A_1}\cdots\nabla_{A_{j-1}}R_{ABvC}$ and write it in terms of primitive factors. This lets us eliminate $D_{(A_1}\cdots D_{A_{j-1})}D_{[A}K_{B]C}$ in favor of $D_{(A_1}\cdots D_{A_k)}K_{BC}$, $D_{(A_1}\cdots D_{A_l)}\beta_B$, $D_{(A_1}\cdots D_{A_{m-1})}R[\mu]_{ABCD}$ with $k,l,m \leq j-1$ and covariant derivatives of the Riemann tensor. Making use of the induction hypothesis yields an expression for $D_{(A_1}\cdots D_{A_j)}K_{BC}$ in terms of the desired quantities. A similar argument establishes the corresponding result for $D_{(A_1}\ldots D_{A_j)}\bar{K}_{BC}$.

Writing $\nabla_{A_1}\cdots\nabla_{A_j-1}R_{ABCD}$ in terms of primitive factors yields an expression relating $D_{(A_1}\cdots D_{A_{j-1})}R_{ABCD}[\mu]$ to $D_{(A_1}\cdots D_{A_k)}\bar{K}_{BC}$, $D_{(A_1}\cdots D_{A_l)}\bar{K}_{BC}$, $D_{(A_1}\cdots D_{A_{m-1})}R_{ABCD}[\mu]$ with $k,l,m \leq j-1$ and covariant components. By employing the induction hypothesis, we get an expression for $D_{(A_1}\cdots D_{A_{j-1})}R_{ABCD}[\mu]$ in terms of the quantities listed in the statement of the lemma.

JHEP08(2022)258

Regarding $D_{(A_1}\cdots D_{A_j)}\beta_B$, we first use an identity similar to (2.29):

$$D_{(A_1}\dots D_{A_j)}\beta_B = -\frac{2j}{j+1}\left(D_{(A_1}\dots D_{A_{j-1})}D_{[A_j}\beta_{B]} + \dots + D_{(A_2}\dots D_{A_j)}D_{[A_1}\beta_{B]}\right) + D_{(A_1}\dots D_{A_j}\beta_{B)} + \dots \tag{2.30}$$

where again the ellipsis in the second line stands for a sum of monomials containing factors of $D_{(A_1}\dots D_{A_m)}R[\mu]_{ABCD}$ with $m \leq j-2$ and a factor $D_{(A_1}\dots D_{A_k)}\beta_B$ or a factor $D_{(A_1}\dots D_{A_{k-1})}D_{[A_k}\beta_{B]}$ with $k \leq j-2$. The terms in the first line of the r.h.s. of this equation can be eliminated by writing $\nabla_{(A_1}\cdots\nabla_{A_{l-1})}R_{rvAB}$ in terms of primitive factors. The terms in the ellipsis can be dealt with by invoking the induction assumption. This yields an expression for $D_{(A_1}\cdots D_{A_j)}\beta_B$ in terms of the desired quantities, closing the induction loop.

Finally, consider the case of $D_{(A_1}\cdots D_{A_j)}\partial_r^{\bar{N}}\bar{K}_{AB}$ with $\bar{N} \geq 1$. Let us take derivatives w.r.t. x^A on our previous expression relating $\partial_r^{\bar{N}}\bar{K}_{AB}$ to $\bar{K}_{AB}$ and covariant components involving the Riemann tensor. This gives expressions relating $D_{(A_1}\cdots D_{A_j)}\partial_r^{\bar{N}}\bar{K}_{AB}$ to $D_{(A_1}\cdots D_{A_k)}\bar{K}_{AB}$, $D_{(A_1}\cdots D_{A_l)}\beta_B$, $D_{(A_1}\cdots D_{A_k)}R[\mu]_{ABCD}$ and covariant components. Using previous substitutions yields the desired expression for $D_{(A_1}\cdots D_{A_j)}\partial_r^{\bar{N}}\bar{K}_{AB}$. A similar argument establishes that $D_{(A_1}\cdots D_{A_j)}\partial_r^{N}K_{AB}$ is also expressible in the desired form. □

2.3 Covariance of GNC tensors and BRST-formalism

We now consider how quantities on our cut C transform under a non-trivial change of GNCs, i.e., a change with non-constant $a(x^A) > 0$. It follows from the results of section 2.1 that under such a change, on C we have[16]

$$\begin{gathered}\mu_{AB} \to \mu_{AB} \qquad \partial_v^N K_{AB} \to a^{+(N+1)}\partial_v^N K_{AB} \quad \partial_r^{\bar{N}}\bar{K}_{AB} \to a^{-(\bar{N}+1)}\partial_r^{\bar{N}}\bar{K}_{AB} \\ \nabla_{(\alpha_1}\cdots\nabla_{\alpha_j)}R_{\mu\nu\sigma\rho} \to a^b\nabla_{(\alpha_1}\cdots\nabla_{\alpha_j)}R_{\mu\nu\sigma\rho} \qquad \beta_A \to \beta_A + 2\partial_A \log a\end{gathered} \tag{2.31}$$

where b is the boost weight of the GNC component of the covariant tensor $\nabla_{(\alpha_1}\cdots\nabla_{\alpha_j)}R_{\mu\nu\sigma\rho}$, see eq. (2.15). A useful way to think about the replacements (2.31) on the cut C is that $\boldsymbol{\beta} = \beta_A \mathrm{d}x^A$ is a gauge potential and that $K_{AB}, \bar{K}_{AB}$ and the GNC components of $\nabla_{(\alpha_1}\cdots\nabla_{\alpha_j)}R_{\mu\nu\sigma\rho}$ are charged under the gauge group $\mathbb{R}_+$, with charges $+1, -1$ and b respectively. The function a corresponds to a finite gauge transformation associated with the gauge group $\mathbb{R}_+$. It is therefore useful to define the gauge covariant derivative

$$\mathcal{D}_A := D_A - \frac{1}{2}b\beta_A, \tag{2.32}$$

with b the boost weight (= charge) of the quantity that it is acting on. See the start of section 2.1 for the geometrical interpretation of $\mathcal{D}$ as a connection on the normal bundle of C.

Our eventual aim is to write an expression for an entropy as an integral of a boost-weight zero $n-2$ form over the cut C. For this entropy to be invariant under a change of GNCs, the $n-2$ form should be gauge-invariant up to addition of an exact form. The following lemma characterizes the most general structure of such a form:

[16] The first two transformations hold on all of $\mathcal{N}$, not just on C.

JHEP08(2022)258

Lemma 2.4. *Let $\boldsymbol{w}_m$ be a local m-form on C of boost weight 0 built from contractions of $D_{(A_1}\cdots D_{A_k)}\partial_v^p\partial_r^q\psi$ with $\psi\in\{\mu_{AB},\beta_A,\alpha\}$, or μ^{AB}, or $\epsilon_{A_1\ldots A_{n-2}}[\mu]$, or $D_{(A_1}\cdots D_{A_k)}R_{ABCD}[\mu]$, such that $\boldsymbol{w}_m\to\boldsymbol{w}_m+\mathrm{d}\boldsymbol{w}_{m-1}$ on C under a change of GNCs. Then*

$$\boldsymbol{w}_m=\boldsymbol{\beta}\wedge\boldsymbol{I}_{m-1}+\boldsymbol{I}_m-\mathrm{d}\boldsymbol{b}_{m-1} \tag{2.33}$$

where $\boldsymbol{\beta}=\beta_A\mathrm{d}x^A$; $\boldsymbol{I}_m$ and $\boldsymbol{I}_{m-1}$ are gauge-invariant local forms such that $\boldsymbol{I}_{m-1}$ is identically closed ($\mathrm{d}\boldsymbol{I}_{m-1}=0$), and $\boldsymbol{b}_{m-1}$ is a local form on C. The forms $\boldsymbol{I}_m$, $\boldsymbol{I}_{m-1}$ can be expressed as boost-weight zero contractions of the following quantities:

- *μ_{AB}, μ^{AB}, $\epsilon_{A_1\ldots A_{n-2}}[\mu]$,*
- *$\mathcal{D}_{(A_1}\cdots\mathcal{D}_{A_j}K_{A)B}$, $\mathcal{D}_{(A_1}\cdots\mathcal{D}_{A_j}\bar{K}_{A)B}$, and*
- *GNC components of the covariant tensors $\nabla_{(\alpha_1}\cdots\nabla_{\alpha_j)}R[g]_{\mu\nu\sigma\rho}$. (By the Bianchi identities, not all these Riemann components are independent, see lemma 2.3.)*

To prove lemma 2.4 and other similar lemmas which we shall need later on, we now reformulate the gauge transformations (2.31) using the BRST method (see e.g. [24]). First, we write $a=e^{t\Lambda/2}$, and then obtain the transformation, on $\mathcal{N}$, of any quantity under an infinitesimal gauge transformation by differentiating its transformation law w.r.t. t and evaluating at $t=0$. This gives — in general very complicated — formulas for the infinitesimal change of any monomial $D_{(A_1}\cdots D_{A_k)}\partial_v^p\partial_r^q\psi$ under an infinitesimal change of GNCs on $\mathcal{N}$.

On C, these formulas simplify if we pass from this basis of primitive monomials to one of the bases provided by lemmas 2.2 or 2.3. Following the usual BRST approach, we now declare Λ to be an anti-commuting field (of boost weight 0) and the infinitesimal version of the transformations (2.31) becomes a "BRST transformation", which we denote as γ:

$$\begin{aligned}
\gamma\mu_{AB}&=0\\
\gamma\partial_v^N K_{AB}&=+\frac{1}{2}(N+1)\Lambda\partial_v^N K_{AB}\\
\gamma\partial_r^{\bar{N}}\bar{K}_{AB}&=-\frac{1}{2}(\bar{N}+1)\Lambda\partial_r^{\bar{N}}\bar{K}_{AB}\\
\gamma\nabla_{(\alpha_1}\cdots\nabla_{\alpha_j)}R_{\mu\nu\sigma\rho}&=\frac{1}{2}b\Lambda\nabla_{(\alpha_1}\cdots\nabla_{\alpha_j)}R_{\mu\nu\sigma\rho}\\
\gamma\beta_A&=\partial_A\Lambda\\
\gamma\Lambda&=0
\end{aligned} \tag{2.34}$$

where the last transformation is imposed as usual to ensure that $\gamma^2=0$ and where b is the boost weight of the corresponding GNC component. The action of γ is extended to a product via the graded Leibniz rule. The degree in Λ of a monomial in the fields and their derivatives is referred to as the ghost number and Λ as the ghost field. Consistent with the last two equations equation in (2.34), we declare that the boost weight of Λ is zero, so that γ does not change the boost weight of any quantity. We furthermore follow the custom to

JHEP08(2022)258

declare the action of the exterior differential d on a ghost number g monomial to be $(-1)^g$ times the usual definition, and with this convention we have $\mathrm{d}\gamma = -\gamma\mathrm{d}$, for example.

For *constant* Λ, the action of γ on a quantity of boost weight b is to multiply that quantity by $b\Lambda/2$. However, in general Λ is not constant. The transformation law assumed in lemma 2.4 may be stated as saying that $\gamma\boldsymbol{w}_m = \mathrm{d}\boldsymbol{w}_{m-1}$ and $\boldsymbol{w}_m$ has ghost number 0 whereas $\boldsymbol{w}_{m-1}$ has ghost number 1. A form such that $\gamma\boldsymbol{w}_m = 0$ is called BRST-closed (γ-closed) and a form of the type $\gamma\boldsymbol{w}_m$ is called BRST-exact.

Lemma 2.5. *Let $\boldsymbol{w}_m$ be a local m form on C of boost weight 0 and ghost number $g > 0$ such that $\gamma\boldsymbol{w}_m = \mathrm{d}\boldsymbol{w}_{m-1}$. Then $\boldsymbol{w}_m$ is a sum of the following:*

- *A local γ-exact local form of boost weight 0.*
- *A local* d*-exact local form of boost weight 0.*
- *Λ times a boost weight 0 contraction of $\mathcal{D}_{(A_1}\cdots\mathcal{D}_{A_j}K_{A)B}$, $\mathcal{D}_{(A_1}\cdots\mathcal{D}_{A_j}\bar{K}_{A)B}$, of GNC component of covariant tensors $\nabla_{(\alpha_1}\cdots\nabla_{\alpha_j)}R[g]_{\mu\nu\sigma\rho}$ [not all of these are independent by the Bianchi identities, see lemma 2.3], of μ^{AB} or of $\epsilon[\mu]_{A_1\ldots A_{n-2}}$. This case only occurs for $g = 1$.*

Poof of lemmas 2.4 and 2.5. The proofs of lemmas 2.4 and 2.5 are quite similar and interrelated so we will give them together. We shall start with the special case that $\gamma\boldsymbol{w}_m = 0$ where $\boldsymbol{w}_m$ is a local form of ghost number $g = 0$. In $\boldsymbol{w}_m$ we make the replacements as in lemma 2.3. Then we further replace $D_{(A_1}\cdots D_{A_j}K_{A)B}$, $D_{(A_1}\cdots D_{A_j}\bar{K}_{A)B}$ with the gauge covariantized forms $\mathcal{D}_{(A_1}\cdots\mathcal{D}_{A_j}K_{A)B}$, $\mathcal{D}_{(A_1}\cdots\mathcal{D}_{A_j}\bar{K}_{A)B}$ at the expense of additional terms of the form $D_{(A_1}\cdots D_{A_j}\beta_{B)}$. Now we impose $\gamma\boldsymbol{w}_m = 0$. Since the factors $D_{(A_1}\cdots D_{A_j}\beta_{B)}$ are the only factors which transform with derivatives of Λ and since the terms with undifferentiated Λ drop out using that $\boldsymbol{w}_m$ has boost weight 0, it follows

$$0 = \gamma\boldsymbol{w}_m = \sum_{j\geq 1}\frac{\partial\boldsymbol{w}_m}{\partial D_{(A_1}\cdots D_{A_{j-1}}\beta_{A_j)}}D_{(A_1}\cdots D_{A_j)}\Lambda. \tag{2.35}$$

At any point in C the terms $D_{(A_1}\cdots D_{A_j)}\Lambda$ can be chosen as linearly independent from each other so we learn that $\partial\boldsymbol{w}_m/\partial D_{(A_1}\cdots D_{A_{j-1}}\beta_{A_j)} = 0$ for all j. Thus there is no dependence on β_A in the chosen basis of monomials and we have demonstrated lemma 2.4 in our special case $\gamma\boldsymbol{w}_m = 0$.

We remain in the same special case but now we go to ghost number $g > 0$. In the basis of monomials given by $\mathcal{D}_{(A_1}\cdots\mathcal{D}_{A_j}K_{A)B}$, $\mathcal{D}_{(A_1}\cdots\mathcal{D}_{A_j}\bar{K}_{A)B}$, by GNC components of $\nabla_{(\alpha_1}\cdots\nabla_{\alpha_j)}R_{\mu\nu\sigma\rho}$, by $D_{(A_1}\cdots D_{A_{j-1}}\beta_{A_j)}$ and by $D_{(A_1}\cdots D_{A_j)}\Lambda$, we realize that the BRST-transformations (2.34) have $D_{(A_1}\cdots D_{A_{j-1}}\beta_{A_j)}$ and $D_{(A_1}\cdots D_{A_j)}\Lambda$ as "contractible pairs" in the terminology of [24], meaning that $\gamma D_{(A_1}\cdots D_{A_{j-1}}\beta_{A_j)} = D_{(A_1}\cdots D_{A_j)}\Lambda$, $\gamma D_{(A_1}\cdots D_{A_j)}\Lambda = 0$ and the rest of the BRST transformations are independent of these variables. By a standard result [24], appendix 2.B, we know that $\boldsymbol{w}_m$ is up to a γ-exact local form equal to a local form that has no explicit dependence on β_A in the chosen basis of monomials and only depends on Λ in undifferentiated form. The proof of this works as

JHEP08(2022)258

follows: first define

$$N = \sum_{j\geq 1}\left(D_{(A_1}\cdots D_{A_{j-1}}\beta_{A_j)}\frac{\partial}{\partial D_{(A_1}\cdots D_{A_{j-1}}\beta_{A_j)}} + D_{(A_1}\cdots D_{A_j)}\Lambda\frac{\partial}{\partial(D_{(A_1}\cdots D_{A_j)}\Lambda)}\right) \tag{2.36}$$

to be the number operator for the contractible pair, and let

$$\rho = \sum_{j\geq 1} D_{(A_1}\cdots D_{A_{j-1}}\beta_{A_j)}\frac{\partial}{\partial D_{(A_1}\cdots D_{A_j)}\Lambda}. \tag{2.37}$$

One has $[N,\gamma]=0$ and $N=\gamma\rho+\rho\gamma$. We now decompose $\boldsymbol{w}_m=\sum_{k\geq 0}\boldsymbol{w}_m^k$ where $\boldsymbol{w}_m^k$ are eigenvectors of N: $N\boldsymbol{w}_m^k = k\boldsymbol{w}_m^k$. Since N and γ commute, $\gamma\boldsymbol{w}_m=0$ implies $\gamma\boldsymbol{w}_m^k=0$. For $k>0$ we have $\boldsymbol{w}_m^k=(1/k)N\boldsymbol{w}_m^k=(1/k)(\gamma\rho+\rho\gamma)\boldsymbol{w}_m^k=\gamma[(1/k)\rho\boldsymbol{w}_m^k]$. Therefore $\boldsymbol{w}_m=\boldsymbol{w}_m^0$ modulo γ exact. By construction, $\boldsymbol{w}_m^0$ only depends on Λ but not on β_A. Therefore, by lemma 2.3,

$$\boldsymbol{w}_m = \Lambda^g\boldsymbol{I}_m + \gamma\text{-exact}, \tag{2.38}$$

where γ-exact means an m-form in the image of the local m-forms and where $\boldsymbol{I}_m$ has boost weight zero and is gauge invariant, i.e. it is a local m-form constructed from $\mathcal{D}_{(A_1}\cdots\mathcal{D}_{A_j}K_{A)B}$, $\mathcal{D}_{(A_1}\cdots\mathcal{D}_{A_j}\bar{K}_{A)B}$ and GNC components of $\nabla_{(\alpha_1}\cdots\nabla_{\alpha_j)}R_{\mu\nu\sigma\rho}$. Since $\Lambda^g=0$ for $g>1$, it follows that when $g>1$, $\boldsymbol{w}_m$ is γ-exact. When $g=1$, we have $\boldsymbol{w}_m=\Lambda\boldsymbol{I}_m$ plus γ-exact. Thus, we have demonstrated lemma 2.5 in our special case $\gamma\boldsymbol{w}_m=0$.

We now turn to the general case of lemmas 2.4 and 2.5 where we only assume that $\gamma\boldsymbol{w}_m=\mathrm{d}\boldsymbol{w}_{m-1}$ for some local $m-1$ form $\boldsymbol{w}_{m-1}$. This case will be analyzed using the standard technique of descent equations. To construct the descent equations, we need to appeal to the algebraic Poincare lemma [25]:

Lemma 2.6 (algebraic Poincare lemma). *Let $\boldsymbol{\xi}([\Psi,\Phi],x)$ be an m form on a manifold C that is locally constructed out of smooth fields Ψ and "background" fields Φ such that $\mathrm{d}\boldsymbol{\xi}([\Psi,\Phi],x)=0$ for all Ψ in a neighborhood of a special configuration Ψ_0. Assume that all Ψ in that neighborhood can be joined to Ψ_0 by a differentiable path $\Psi_\lambda, \lambda\in[0,1]$ of smooth configurations such that $\Psi_\lambda|_x$ depends on finitely many derivatives $\partial^k\Psi|_x$ for any $x\in C$ and such that $\Psi_{\lambda=1}=\Psi$. Then $\boldsymbol{\xi}([\Psi,\Phi],x)-\boldsymbol{\xi}([\Psi_0,\Phi],x)=\mathrm{d}\boldsymbol{\eta}([\Psi,\Psi_0,\Phi],x)$ for some local $m-1$ form $\boldsymbol{\eta}([\Psi,\Psi_0,\Phi],x)$.*

Proof. [25] We sketch the proof here since we later want to see how $\boldsymbol{\eta}$ preserves certain structures of $\boldsymbol{\xi}$ below. We will suppress the dependence on the background fields Φ in this proof. We fix an auxiliary covariant derivative D_A on C (coordinate indices $A,B,\dots$). Consider the linearization of $\boldsymbol{\xi}$,

$$\delta\boldsymbol{\xi}([\Psi],x)_{B_1\dots B_m} = \sum_{j=0}^{k}\frac{\partial\boldsymbol{\xi}_{B_1\dots B_m}}{\partial(D_{(A_1}\dots D_{A_j)}\Psi)}([\Psi],x)D_{(A_1}\dots D_{A_j)}\delta\Psi(x), \tag{2.39}$$

where k is the maximum number of derivatives of Ψ that $\boldsymbol{\xi}$ depends on. Then we let

$$\boldsymbol{\gamma}_k([\Psi,\delta\Psi],x)_{C_1\dots C_{m-1}} := \frac{mk}{(m+1)(k+1)}\frac{\partial\boldsymbol{\xi}_{C_1\dots C_{m-1}B}}{\partial(D_{(A_1}\dots D_{A_{k-1}}D_{B)}\Psi)}([\Psi],x)D_{(A_1}\dots D_{A_{k-1})}\delta\Psi(x). \tag{2.40}$$

JHEP08(2022)258

The arguments given in [25] show that $\delta\boldsymbol{\xi}([\Psi],x) - \mathrm{d}\boldsymbol{\gamma}_k([\Psi,\delta\Psi],x)$ is a closed local functional depending on up to $(k-1)$ derivatives of $\delta\Psi$ and so verifies the assumptions of the theorem because $\delta\Psi$ is an arbitrary variation. The argument is then iterated to determine $\boldsymbol{\gamma}_{k-1}([\Psi,\delta\Psi],x)$ and so on and then set $\boldsymbol{\gamma}([\Psi,\delta\Psi],x) = \sum_{j\leq k}\boldsymbol{\gamma}_j([\Psi,\delta\Psi],x)$. Finally, we set (using a dot to denote a derivative w.r.t. λ)

$$\boldsymbol{\eta}([\Psi_0,\Psi],x) = \int_0^1 \mathrm{d}\lambda\boldsymbol{\gamma}([\Psi_\lambda,\dot{\Psi}_\lambda],x), \tag{2.41}$$

which is local and satisfies the desired equation. □

Remark 1. In the applications below, the space of Ψ will consist of certain monomials as in lemmas 2.2, 2.3, and will have the structure of vector space. Thus, the interpolating family can be chosen to be simply $\lambda\Psi = \Psi_\lambda$ and $\Psi_0 = 0$. Furthermore, the forms $\boldsymbol{\xi}$ in question will be at least quadratic in positive boost weight monomials. It is clear from the above proof that also $\boldsymbol{\eta}$ will then be at least quadratic in positive boost weight quantities.

We now return to the proof of lemmas 2.4 and 2.5 in the general case. Taking γ of $\gamma\boldsymbol{w}_m = \mathrm{d}\boldsymbol{w}_{m-1}$, using $\gamma^2 = \mathrm{d}\gamma + \gamma\mathrm{d} = 0$, we learn that $\mathrm{d}(\gamma\boldsymbol{w}_{m-1}) = 0$. $\gamma\boldsymbol{w}_{m-1}$ depends at least linearly on Λ, so we can apply the algebraic Poincare lemma for $\Psi = \Lambda, \Psi_0 = 0$ (viewing the other fields as background fields), to learn that $\gamma\boldsymbol{w}_{m-1} = \mathrm{d}\boldsymbol{w}_{m-2}$, where $\boldsymbol{w}_{m-2}$ is again local. Iterating this process we get the descent equations

$$\begin{aligned} \gamma\boldsymbol{w}_m &= \mathrm{d}\boldsymbol{w}_{m-1} \\ \gamma\boldsymbol{w}_{m-1} &= \mathrm{d}\boldsymbol{w}_{m-2} \\ &\dots \\ \gamma\boldsymbol{w}_k &= 0. \end{aligned} \tag{2.42}$$

where the ghost number increases by 1 at each step. The last equation does not involve a boundary term. Such an equation must exist for some k — in the worst case, i.e. the longest descent, we have $k = 0$. The important point is now that (as $\boldsymbol{w}_k$ has positive ghost number) we can apply our previous results to $\gamma\boldsymbol{w}_k = 0$. We learn $\boldsymbol{w}_k = \gamma\boldsymbol{b}_k + \Lambda\boldsymbol{I}_k$, where the last term $\Lambda\boldsymbol{I}_k$ is present only if $\boldsymbol{w}_k$ has ghost number 1. Thus, if $\boldsymbol{w}_m$ had had ghost number 0 (lemma 2.4), the last term is present when $k = m-1$, whereas if $\boldsymbol{w}_m$ had had ghost number $g > 0$ (lemma 2.5), the last term is present only when $g = 1$ and $k = m$. So long as the last term $\Lambda\boldsymbol{I}_k$ is not present, we can consider $\tilde{\boldsymbol{w}}_{k+1} := \boldsymbol{w}_{k+1} + \mathrm{d}\boldsymbol{b}_k, \tilde{\boldsymbol{w}}_j := \boldsymbol{w}_j$ for $j > k+1$, and then the new ladder $\{\tilde{\boldsymbol{w}}_m, \dots, \tilde{\boldsymbol{w}}_{k+1}\}$ will satisfy by construction the descent

$$\begin{aligned} \gamma\tilde{\boldsymbol{w}}_m &= \mathrm{d}\tilde{\boldsymbol{w}}_{m-1} \\ \gamma\tilde{\boldsymbol{w}}_{m-1} &= \mathrm{d}\tilde{\boldsymbol{w}}_{m-2} \\ &\dots \\ \gamma\tilde{\boldsymbol{w}}_{k+1} &= 0, \end{aligned} \tag{2.43}$$

i.e. we have shortened the descent. We continue with this until we get a nontrivial term $\Lambda\boldsymbol{I}_k$.

JHEP08(2022)258

Lemma 2.5. Since $g \geq 1$, the ghost number of $\tilde{\boldsymbol{w}}_j$ is greater than 1 for $j < m$ so we can shorten the descent until we reach $\gamma\tilde{\boldsymbol{w}}_m = 0$. Our previous result gives $\tilde{\boldsymbol{w}}_m = \gamma\boldsymbol{b}_m + \Lambda^g\boldsymbol{I}_m$ and hence $\boldsymbol{w}_m = \gamma\boldsymbol{b}_m - \mathrm{d}\boldsymbol{b}_{m-1} + \Lambda^g\boldsymbol{I}_m$, establishing the result (as $\Lambda^g = 0$ for $g > 1$).

Lemma 2.4. In this case we have $g = 0$ and we can shorten the descent until we reach $\gamma\tilde{\boldsymbol{w}}_{m-1} = 0$, which implies $\tilde{\boldsymbol{w}}_{m-1} = \gamma\boldsymbol{b}_{m-1} + \Lambda\boldsymbol{I}_{m-1}$, so $\boldsymbol{w}_{m-1} = \gamma\boldsymbol{b}_{m-1} + \Lambda\boldsymbol{I}_{m-1} - \mathrm{d}\boldsymbol{b}_{m-2}$. This gives

$$\begin{aligned}\gamma\boldsymbol{w}_m = \mathrm{d}(\gamma\boldsymbol{b}_{m-1} + \Lambda\boldsymbol{I}_{m-1}) &\Longrightarrow \gamma(\boldsymbol{w}_m + \mathrm{d}\boldsymbol{b}_{m-1}) = \mathrm{d}(\Lambda\boldsymbol{I}_{m-1}) \\ &\Longrightarrow \gamma(\boldsymbol{w}_m + \mathrm{d}\boldsymbol{b}_{m-1} - \boldsymbol{\beta}\wedge\boldsymbol{I}_{m-1}) = -\Lambda\mathrm{d}\boldsymbol{I}_{m-1}.\end{aligned} \tag{2.44}$$

The left side depends only on Λ in differentiated form by the definition of γ and because all forms have boost weight 0, whereas the right side depends on Λ only in undifferentiated form. Therefore, since the equation must hold for all Λ, and since $D_{(A_1}\dots D_{A_j)}\Lambda$ are linearly independent at each point, both sides must vanish. Vanishing of the l.h.s. implies by our previous results (for ghost number 0) that $\boldsymbol{w}_m + \mathrm{d}\boldsymbol{b}_{m-1} - \boldsymbol{\beta}\wedge\boldsymbol{I}_{m-1} = \boldsymbol{I}_m$ and vanishing of the r.h.s. gives $\mathrm{d}\boldsymbol{I}_{m-1} = 0$ which completes the proof. □

Although it is not essential for our proof of gauge invariance of the IWW entropy, it is interesting to note that we can be more explicit about the form of $\boldsymbol{I}_{m-1}$ in lemma 2.4. In the following lemma note that $\mathrm{d}\boldsymbol{\beta}$ (the curvature of the connection on the normal bundle of C) can be expressed in terms of the quantities listed in lemma 2.4 via equation (2.25).

Lemma 2.7. *The gauge-invariant forms $\boldsymbol{I}_{m-1}$, $\boldsymbol{I}_m$ of lemma 2.4 can be defined so that $\boldsymbol{I}_{m-1}$ is written in terms of $\mathrm{d}\boldsymbol{\beta}$ and characteristic classes built from the curvature of μ_{AB} as follows:*

$$\boldsymbol{I}_{m-1} = \sum_j \left(\bigwedge^j \mathrm{d}\boldsymbol{\beta}\right) \wedge p_{m-1-2j}(\underbrace{\boldsymbol{R}[\mu],\dots,\boldsymbol{R}[\mu]}_{(m-1-2j)/2}), \tag{2.45}$$

where $\boldsymbol{R}[\mu]_A{}^B = R[\mu]_{CDA}{}^B\mathrm{d}x^C\wedge\mathrm{d}x^D$ and p_i is an invariant symmetric polynomial of the Lie-algebra $\mathfrak{so}(n-2)$. Such terms can only appear when m is odd.

Proof. Given our metric corresponding to the $\psi = (\alpha, \beta_A, \mu_{AB})$ in its GNC system, let ψ_0 be such that $\alpha_0 = (\beta_0)_A = 0$ and $(\mu_0)_{AB}(v, r, x^A) = \mu_{AB}(0, 0, x^A)$. We can construct a 1-parameter family ψ_λ with $\psi_1 = \psi$ by setting $\alpha_\lambda = \lambda\alpha$, $(\beta_\lambda)_A = \lambda\beta_A$ and $(\mu_\lambda)_{AB} = (1-\lambda)(\mu_0)_{AB} + \lambda\mu_{AB}$. From the algebraic Poincaré lemma we have $\boldsymbol{I}_{m-1}[\psi] = \boldsymbol{I}_{m-1}[\psi_0] + \mathrm{d}\boldsymbol{J}_{m-2}[\psi]$ where $\boldsymbol{J}_{m-2}$ is a local form such that $\mathrm{d}\boldsymbol{J}_{m-2}$ is gauge invariant. The latter condition gives $\gamma\mathrm{d}\boldsymbol{J}_{m-2} = 0$ hence $\mathrm{d}\gamma\boldsymbol{J}_{m-2} = 0$ but $\gamma\boldsymbol{J}_{m-2}$ has ghost number 1 so we can use the algebraic Poincaré lemma interpolating Λ to 0 (treating other fields as background fields) and deduce that $\gamma\boldsymbol{J}_{m-2}$ is exact. We can now apply lemma 2.4 to deduce $\boldsymbol{J}_{m-2} = \boldsymbol{\beta}\wedge\boldsymbol{I}_{m-3} + \boldsymbol{I}_{m-2} - \mathrm{d}\boldsymbol{b}_{m-3}$ where $\boldsymbol{I}_{m-3}$ is gauge-invariant and closed, and $\boldsymbol{I}_{m-2}$ is gauge-invariant. This gives

$$\boldsymbol{I}_{m-1}[\psi] = \boldsymbol{I}_{m-1}[\psi_0] + \mathrm{d}\boldsymbol{\beta}\wedge\boldsymbol{I}_{m-3}[\psi] + \mathrm{d}\boldsymbol{I}_{m-2}[\psi]\,. \tag{2.46}$$

JHEP08(2022)258

We now repeat the process starting from $\boldsymbol{I}_{m-3}$ to obtain

$$\boldsymbol{I}_{m-3}[\psi] = \boldsymbol{I}_{m-3}[\psi_0] + \mathrm{d}\boldsymbol{\beta} \wedge \boldsymbol{I}_{m-5}[\psi] + \mathrm{d}\boldsymbol{I}_{m-4}[\psi] \tag{2.47}$$

and hence

$$\boldsymbol{I}_{m-1}[\psi] = \boldsymbol{I}_{m-1}[\psi_0] + \mathrm{d}\boldsymbol{\beta} \wedge \boldsymbol{I}_{m-3}[\psi_0] + \mathrm{d}\boldsymbol{\beta} \wedge \mathrm{d}\boldsymbol{\beta} \wedge \boldsymbol{I}_{m-5}[\psi] + \mathrm{d}\boldsymbol{I}'_{m-2}[\psi] \tag{2.48}$$

where $\boldsymbol{I}'_{m-2} = \mathrm{d}\boldsymbol{\beta} \wedge \boldsymbol{I}_{m-4} + \boldsymbol{I}_{m-2}$ is gauge-invariant. Proceeding inductively we reach

$$\boldsymbol{I}_{m-1}[\psi] = \sum_j \left(\bigwedge^j \mathrm{d}\boldsymbol{\beta} \right) \wedge \boldsymbol{I}_{m-1-2j}[\psi_0] + \mathrm{d}\boldsymbol{I}''_{m-2}[\psi]\,. \tag{2.49}$$

On the r.h.s., $\boldsymbol{I}_k[\psi_0]$ is an identically closed k-form constructed locally and covariantly only from $(\mu_0)_{AB} \equiv \mu_{AB}|_C$: these are the characteristic classes built from $\boldsymbol{R}[\mu]_A{}^B = R[\mu]_{CDA}{}^B \mathrm{d}x^C \wedge \mathrm{d}x^D$ [see for example eq. (6.32) of [27], replacing spacetime by (C, μ_{AB})]. The result now follows upon substituting the above equation into our formula for $\boldsymbol{w}_m$ in lemma 2.4 and setting $\boldsymbol{I}'_m = \boldsymbol{I}_m + \mathrm{d}\boldsymbol{\beta} \wedge \mathrm{d}\boldsymbol{I}''_{m-2}$ (which is gauge invariant) and $\boldsymbol{b}'_{m-1} = \boldsymbol{b}_{m-1} + \boldsymbol{\beta} \wedge \mathrm{d}\boldsymbol{I}''_{m-2}$. □

3 Entropy current associated with a null surface

3.1 Covariant phase space formalism

We shall use the covariant phase space formalism of Iyer and Wald [2] and so we summarize the main results of that formalism that we shall need. Consider a Lagrangian n-form $\boldsymbol{L}$ depending locally and covariantly on some fields $\Psi^i = (g_{\mu\nu}, \Phi_I, A_{J\mu})$ including a Lorentzian metric g. When the fields include 1-form fields $\boldsymbol{A}_J$, we also require that $\boldsymbol{L}$ is gauge invariant. Local and covariant means by definition that in any coordinate system and at any point x, $\boldsymbol{L}|_x$ is expressible in terms of a finite number of derivatives $\partial_{\alpha_1} \dots \partial_{\alpha_d} \Psi|_x$, and that, for every diffeomorphism f, $f^*\boldsymbol{L}[\Psi] = \boldsymbol{L}[f^*\Psi]$. By the Thomas replacement lemma [2], $\boldsymbol{L}$ can be written as $L\boldsymbol{\epsilon}$, with $\boldsymbol{\epsilon}$ the positively oriented volume form defined from $g_{\mu\nu}$, where L can be expressed entirely in terms of $g_{\mu\nu}, g^{\mu\nu}, \nabla_{\alpha_1} \cdots \nabla_{\alpha_k} R_{\mu\nu\alpha\beta}$, $\epsilon_{\alpha_1 \dots \alpha_n}$, $\nabla_{\alpha_1} \cdots \nabla_{\alpha_k} \Phi_I$, and $\nabla_{\alpha_1} \cdots \nabla_{\alpha_k} F_{\mu_1 \dots \mu_{p+1} J}$ when p-form fields with curvatures $\boldsymbol{F}_J$ are present. However, for most parts of this paper we will not consider matter fields and treat (local and covariant) Lagrangians depending only on the metric. Note that anti-symmetrized combinations of covariant derivatives can be rewritten in terms of curvature tensors, so without loss of generality we will usually consider symmetrized derivatives as the variable fields in our Lagrangians and other functionals. There are further dependencies among the derivatives of the curvature components arising from the Bianchi identities which are usually understood to be eliminated by choosing a suitably linearly independent set, see below.

In such a case, the only equation of motion is the Einstein equation $E^{\mu\nu} = 0$ where

$$E^{\mu\nu}\boldsymbol{\epsilon} \equiv \boldsymbol{E}_g^{\mu\nu} = \sum_{j \geq 0} (-1)^j \partial_{\alpha_1} \dots \partial_{\alpha_j} \frac{\partial \boldsymbol{L}}{\partial(\partial_{\alpha_1} \dots \partial_{\alpha_j} g_{\mu\nu})}. \tag{3.1}$$

JHEP08(2022)258

$E^{\mu\nu}$ specifies the classical dynamics so any classical physical quantity should be specified only in terms of this dynamical content plus potentially some global quantities of topological nature. The symplectic $(n-1)$ potential is defined by $\boldsymbol{\theta}[\Psi;\delta\Psi]$

$$\delta\boldsymbol{L}[\Psi] = \boldsymbol{E}_{\Psi^i}[\Psi]\cdot\delta\Psi^i + \mathrm{d}\boldsymbol{\theta}[\Psi;\delta\Psi]. \tag{3.2}$$

We use the general notation

$$\delta\Psi = \frac{\mathrm{d}}{\mathrm{d}\lambda}\Psi(\lambda), \quad \delta^2\Psi = \frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}\lambda^2}\Psi(\lambda) \tag{3.3}$$

for a 1-parameter family of dynamical fields. More generally, we write $\delta_1\Psi, \delta_2\Psi, \delta_1\delta_2\Psi$ for families depending on multiple parameters. The definition of $\boldsymbol{\theta}[\Psi;\delta\Psi]$ leaves some ambiguities because we may add a d-exact covariant term or topological term to the Lagrangian $\boldsymbol{L} \to \boldsymbol{L} + \mathrm{d}\boldsymbol{\mu} + \boldsymbol{L}^{\mathrm{top}}$ without changing the equations of motion, $E_g^{\mu\nu}$. Furthermore, we may add a local and covariant d-exact term to $\boldsymbol{\theta}$ itself, $\boldsymbol{\theta} \to \boldsymbol{\theta} + \mathrm{d}\boldsymbol{Y}$. Altogether, these ambiguities result in the freedom to modify $\boldsymbol{\theta} \to \boldsymbol{\theta} + \delta\boldsymbol{\mu} + \boldsymbol{\theta}^{\mathrm{top}} + \mathrm{d}\boldsymbol{Y}$ without changing the local dynamical content of the theory.

The symplectic $(n-1)$-form is, for a purely gravitational theory,

$$\boldsymbol{\omega}[g;\delta_1 g,\delta_2 g] = \delta_1\boldsymbol{\theta}[g;\delta_2 g] - \delta_2\boldsymbol{\theta}[g;\delta_1 g]. \tag{3.4}$$

The Noether current is

$$\boldsymbol{J}_X[g] = \boldsymbol{\theta}[g;\mathcal{L}_X g] - X\cdot\boldsymbol{L}[g], \tag{3.5}$$

again for a purely gravitational theory, and the above ambiguities of $\boldsymbol{\theta}$ and $\boldsymbol{L}$ propagate into ambiguities of $\boldsymbol{J}_X$ and of $\boldsymbol{\omega}$. One can establish [2] the existence of a non-unique Noether-charge $(n-2)$-form $\boldsymbol{Q}_X$ satisfying

$$\mathrm{d}\boldsymbol{Q}_X[g] = \boldsymbol{J}_X[g] + \boldsymbol{C}_X[g], \tag{3.6}$$

where $\boldsymbol{C}_X$, which we call the constraint, does not depend on derivatives of X, and $\boldsymbol{C}_X = 0$ when the equations of motion are fulfilled. For a purely gravitational theory as considered here, the constraint is [12]

$$(\boldsymbol{C}_X)_{\alpha_1\ldots\alpha_{n-1}} = 2E^{\mu\nu}X_\nu\epsilon_{\mu\alpha_1\ldots\alpha_{n-2}}. \tag{3.7}$$

To see this, note that equation (3.2) and the (off-shell) Bianchi identity $\nabla_\mu E^{\mu\nu} = 0$ give

$$\begin{aligned}
\mathrm{d}\boldsymbol{J}_X[g] &= \mathrm{d}\boldsymbol{\theta}[g,\mathcal{L}_X g] - \mathrm{d}(X\cdot\boldsymbol{L}[g]) \\
&= \mathcal{L}_X\boldsymbol{L}[g] - \boldsymbol{E}_g^{\mu\nu}[g]\mathcal{L}_X g_{\mu\nu} - (\mathrm{d}(X\cdot\boldsymbol{L}[g]) + X\cdot\mathrm{d}\boldsymbol{L}[g]) \\
&= -2\boldsymbol{\epsilon}E^{\mu\nu}\nabla_\mu X_\nu \\
&= -2\boldsymbol{\epsilon}\nabla_\mu(E^{\mu\nu}X_\nu) \\
&= -\mathrm{d}\boldsymbol{C}_X[g],
\end{aligned} \tag{3.8}$$

using Cartan's magic formula for the Lie derivative. Hence $\mathrm{d}(\boldsymbol{J}_X[g] + \boldsymbol{C}_X[g]) = 0$. Since this is true for any X and any $g_{\mu\nu}$, the algebraic Poincare lemma 2.6 yields the existence

JHEP08(2022)258

of a corresponding local $\boldsymbol{Q}_X$. Note also that the expression (3.7) is unaffected by the above ambiguity to change $\boldsymbol{L} \to \boldsymbol{L} + \mathrm{d}\boldsymbol{\mu} + \boldsymbol{L}_{\mathrm{top}}$ and the corresponding ambiguity $\boldsymbol{\theta} \to \boldsymbol{\theta} + \delta\boldsymbol{\mu} + \boldsymbol{\theta}^{\mathrm{top}} + \mathrm{d}\boldsymbol{Y}$, even though $\boldsymbol{Q}_X$ itself is affected by this ambiguity.

A general result from [2] is that, up to the ambiguities affecting $\boldsymbol{\theta}$ and $\boldsymbol{L}$ just described, $\boldsymbol{Q}_X$ may be written as

$$Q_{\mu_1\ldots\mu_{n-2}}[g] = -E_R^{\alpha\beta\mu\nu}[g](\nabla_\mu X_\nu)\epsilon_{\alpha\beta\mu_1\ldots\mu_{n-2}} + W^{\alpha\beta\nu}[g]X_\nu\epsilon_{\alpha\beta\mu_1\ldots\mu_{n-2}}, \tag{3.9}$$

where $E_R^{\alpha\beta\mu\nu}$ is

$$E_R^{\alpha\beta\mu\nu}\boldsymbol{\epsilon} := \sum_{k\geq 0}(-1)^k\nabla_{(\alpha_1}\cdots\nabla_{\alpha_k)}\frac{\partial\boldsymbol{L}}{\partial(\nabla_{(\alpha_1}\cdots\nabla_{\alpha_k)}R_{\mu\nu\alpha\beta})}. \tag{3.10}$$

The ambiguities affecting $\boldsymbol{\theta}$ and $\boldsymbol{L}$ result in the change $\boldsymbol{Q}_X[g] \to \boldsymbol{Q}_X[g] + X\cdot\boldsymbol{\mu}[g] + \boldsymbol{Y}[g;\mathcal{L}_X g] + \mathrm{d}\boldsymbol{Z}_X[g] + \boldsymbol{Q}_X^{\mathrm{top}}[g]$, which represents the total ambiguity in defining the Noether charge.

Another useful general identity which holds for general 1-parameter families of fields which are not necessarily solutions is obtained starting from

$$\begin{aligned}\delta\boldsymbol{C}_X[g] &= \mathrm{d}\delta\boldsymbol{Q}_X[g] - \delta\boldsymbol{J}_X[g]\\ &= \mathrm{d}\delta\boldsymbol{Q}_X[g] - \delta(\boldsymbol{\theta}[g;\mathcal{L}_X g] - X\cdot\boldsymbol{L}[g])\\ &= \mathrm{d}\delta\boldsymbol{Q}_X[g] - \delta\boldsymbol{\theta}[g;\mathcal{L}_X g] + X\cdot(\boldsymbol{E}_g[g]\,\delta g + \mathrm{d}\boldsymbol{\theta}[g;\delta g])\\ &= \mathrm{d}(\delta\boldsymbol{Q}_X[g] - X\cdot\boldsymbol{\theta}[g;\delta g]) - \boldsymbol{\omega}[g;\delta g,\mathcal{L}_X g] - X\cdot(\boldsymbol{E}_g[g]\,\delta g),\end{aligned} \tag{3.11}$$

using Cartan's magic formula $\mathcal{L}_X\boldsymbol{\theta} = \mathrm{d}(X\cdot\boldsymbol{\theta}) + X\cdot\mathrm{d}\boldsymbol{\theta}$ and the definition of $\boldsymbol{\omega}$ in the last step. Now take a C^1-codimension 1 surface $\mathcal{S}$ in spacetime, take X tangent to $\mathcal{S}$, and pull back the above equation for forms to $\mathcal{S}$. Since the pull back of $X\cdot(\boldsymbol{E}_g\,\delta g)$ vanishes, we get:

$$\mathrm{d}(\delta\boldsymbol{Q}_X[g] - X\cdot\boldsymbol{\theta}[g;\delta g]) - \delta\boldsymbol{C}_X[g] = \boldsymbol{\omega}[g;\delta g,\mathcal{L}_X g] \quad \text{on } \mathcal{S}. \tag{3.12}$$

It is important to note that this is an off-shell identity: neither g nor δg have to satisfy any equations.

3.2 Covariant phase space derivation of IWW entropy current

The aim of this section is to demonstrate the relation between the entropy determined by Wall's procedure and the Iyer-Wald entropy. We will apply the results of the previous section to the case where $\mathcal{S}$ is a smooth null hypersurface $\mathcal{N}$ ruled by affinely parameterized null geodesics. We choose an arbitrary cut C and introduce GNCs as described in section 1.3. We now define the vector field

$$K = v\partial_v - r\partial_r \tag{3.13}$$

generating a boost in the coordinates r, v. We take $X = K$ and consider the constraint $\boldsymbol{C}_K$, which we write

$$(\boldsymbol{C}_K)_{\mu_1\ldots\mu_{n-1}} = 2K^\nu E_\nu{}^\alpha\epsilon_{\alpha\mu_1\ldots\mu_{n-1}}, \tag{3.14}$$

The pull-back of $\boldsymbol{C}_K$ to $\mathcal{N}$ is $2vE_{vv}\sqrt{\mu}\mathrm{d}v\wedge\mathrm{d}^{n-2}x$ in GNCs. To analyze the structure of E_{vv} it is useful to introduce an equivalence relation on functionals on $\mathcal{N}$:

JHEP08(2022)258

Definition 3.1. *Two functionals A, B each of which is a sum of primitive monomials (definition 2.1) with definite boost weight p are said to be equivalent $A \sim B$ if $A - B$ contains only monomials having at least two primitive factors with positive boost weight.*

In the following, functionals of boost weight p will be denoted with symbols such as $A^{(p)}, X^{(p)}, \ldots$. Consider a functional $A^{(2)}$ of boost weight $+2$ (later we will take this to be E_{vv}). One can use a "partial integration" argument to show that there exist $X^{(-k)}, W^{(1)A}, k \geq 0$ such that (i) $X^{(-k)}$ is a sum of primitive monomials for which *each* primitive factor has non-positive boost weight, such that (ii) $W^{(1)A}$ is a sum of primitive monomials for which precisely one primitive factor has positive boost weight, and such that (iii)

$$\begin{aligned} A^{(2)} \sim \partial_v \Bigg\{ & \frac{1}{\sqrt{\mu}} \sum_{j,p,\psi} X_{\psi}^{(0)A_1 \ldots A_j} \partial_v (\sqrt{\mu} D_{(A_1} \cdots D_{A_j)} \partial_v^p \partial_r^p \psi) \\ & + \frac{1}{\sqrt{\mu}} \sum_j X_R^{(0)A_1 \ldots A_j ABCD} \partial_v (\sqrt{\mu} D_{(A_1} \cdots D_{A_j)} R[\mu]_{ABCD}) \\ & + \frac{1}{\sqrt{\mu}} \partial_v \sum_{j,p-q>0,\psi} \sqrt{\mu} X_{q,\psi}^{(-p+q)A_1 \ldots A_j} D_{(A_1} \cdots D_{A_j)} \partial_v^p \partial_r^q \psi + D_A W^{(1)A} \Bigg\}. \end{aligned} \tag{3.15}$$

Note that (i) implies that each primitive monomial in an $X^{(0)}$ term must be a product of primitive factors with zero boost weight. A derivation of the above result was sketched in [4] and explained in more detail in [12].

The key trick is now to define a 1-parameter family of metrics interpolating between a stationary "background" metric (similarly as in [2]) and some other metric. Let $g_{\mu\nu}(x^\alpha)$ be the given metric expressed in GNCs and read off $\psi \in \{\alpha, \beta_A, \mu_{AB}\}$. The background fields are defined by

$$\bar{\psi} := \sum_{p \geq 0} \frac{1}{(p!)^2} \chi(c_p r v)(rv)^p (\partial_r \partial_v)^p \psi|_{r=v=0} \tag{3.16}$$

where χ is smooth and of compact support and identically equal to 1 in a neighborhood of the origin and $\{c_p\}$ is a sequence of positive numbers increasing sufficiently fast in order that the series converges.[17] From $\bar{\psi}$ we define a corresponding background metric $\bar{g}_{\mu\nu}$ in GNCs. This background metric has the properties:

- $\mathcal{L}_K \bar{g} = 0$. In particular, the background metric $\bar{g}_{\mu\nu}$ is such that $\mathcal{N}$ is a Killing horizon with bifurcation surface C. Note that this background metric depends on the original choice of GNCs.

- In the background, any primitive factor with positive boost weight vanishes on $\mathcal{N}$ and any primitive factor with negative boost weight vanishes on C. This implies that when $X^{(k)}$ has positive boost weight k, then $X^{(k)}[\bar{g}] = 0$ on $\mathcal{N}$, and when $X^{(-k)}$ has negative boost weight $-k$, then $X^{(-k)}[\bar{g}] = 0$ on C.

[17] This is just a convenient trick to avoid having to consider Taylor series with remainder, which would be fine too but notationally cumbersome.

JHEP08(2022)258

- If $X^{(0)}$ is a sum of primitive monomials where each monomial involves only primitive factors of boost weight 0 then $X^{(0)}[\bar{g}] = X^{(0)}[g]$ on C.

Next for $\lambda \in [0,1]$ (the interpolation parameter), and for an arbitrary smooth $\delta\psi$ we set

$$\psi(\lambda) := \bar{\psi} + \lambda\delta\psi \tag{3.17}$$

and from $\psi(\lambda)$, we define a corresponding 1-parameter family of metrics $g_{\mu\nu}(\lambda, x^\alpha)$ in GNCs.[18] This 1-parameter family has the properties

- $g|_{\lambda=0} = \bar{g}$
- If we choose $\delta\psi = \psi - \bar{\psi}$ then $g|_{\lambda=1}$ is the original metric g. Note that we do not make this choice at the beginning of the following argument.

We now take $A^{(2)} = E_{vv}$ in (3.15) and consider the first variation of this equation within our 1-parameter family. Using the properties just described, we get on $\mathcal{N}$ (schematically denoting by $Y^{(k)}$ the monomials $D_{(A_1}\cdots D_{A_j)}\partial_v^p\partial_r^q\psi$ with $k = p - q$ or the monomials $D_{(A_1}\cdots D_{A_j)}R[\mu]_{ABCD}$ with $k = 0$)

$$\frac{\partial}{\partial\lambda}E_{vv}|_{\lambda=0} = \partial_v^2\left\{X^{(0)}\frac{\partial}{\partial\lambda}Y^{(0)} + \sum_{k>0}X^{(-k)}\frac{\partial}{\partial\lambda}Y^{(k)}\right\}_{\lambda=0} + \partial_v D_A\frac{\partial}{\partial\lambda}W^{(1)A}|_{\lambda=0} \tag{3.18}$$

because all other terms have an unvaried monomial with positive boost weight and such a monomial vanishes in the background $\bar{g}$. The terms with $k > 0$ in curly brackets can be rewritten as $\frac{\partial}{\partial\lambda}\sum_{k>0}X^{(-k)}Y^{(k)}|_{\lambda=0}$ because $Y^{(k)}[\bar{g}] = 0$ for $k > 0$.

We will now show that the terms with $k = 0$ are related to the Noether charge using the restriction of eq. (3.12) (with $X = K$) to $\mathcal{N}$,[19] and evaluated on the background $\bar{g}$:

$$\partial_v\left(\frac{1}{\sqrt{\mu}}\frac{\partial}{\partial\lambda}(\star Q_K^{vr}\sqrt{\mu}) - v\,(\star\theta)^r\left[g;\frac{\partial}{\partial\lambda}g\right]\right)_{\lambda=0} + D_A\frac{\partial}{\partial\lambda}(\star Q_K^{Ar})|_{\lambda=0} = 2v\,\frac{\partial}{\partial\lambda}E_{vv}|_{\lambda=0}, \tag{3.19}$$

using that

$$\boldsymbol{\omega}\left[g;\frac{\partial}{\partial\lambda}g, \mathcal{L}_K g\right]_{\lambda=0} = 0, \tag{3.20}$$

as $\mathcal{L}_K g|_{\lambda=0} = 0$ by construction. We now assume that $\delta\psi$ has compact support so $\frac{\partial}{\partial\lambda}g|_{\lambda=0}$ also has compact support. If we integrate the above equation over the $v > 0$ portion of $\mathcal{N}$, use the structure (3.18), and perform integrations by part in v, we obtain

$$\int_C\left(\frac{\partial}{\partial\lambda}(\star Q_K^{vr}\sqrt{\mu}) + 2\sqrt{\mu}X^{(0)}\frac{\partial}{\partial\lambda}Y^{(0)}\right)_{\lambda=0} = 0. \tag{3.21}$$

[18] We assume that $\delta\mu_{AB}$ is chosen such that $\mu_{AB}(\lambda)$ is Riemannian for $\lambda \in [0,1]$ which ensures that $g_{\mu\nu}(\lambda)$ is Lorentzian near $\mathcal{N}$.

[19] When considering the restriction of forms to $\mathcal{N}$ it is useful to work with the Hodge duals of the forms; for example we write $(\boldsymbol{Q}_X)_{\mu_1\ldots\mu_{n-2}} = \frac{1}{2}\epsilon_{\mu_1\ldots\mu_{n-2}\alpha\beta}(\star\boldsymbol{Q}_X)^{\alpha\beta}$.

JHEP08(2022)258

In deriving this equation it has been used that $X^{(-k)}(g|_{\lambda=0}) = 0$ for a negative boost weight $-k$ quantity on C, we have used that all terms with an explicit dependence on v vanish at C, and that divergence terms (i.e. $D_A \frac{\partial}{\partial\lambda}(\star Q_K^{Ar})$ and $D_A \frac{\partial}{\partial\lambda} W^{(1)A}$) integrate to zero.

On $\mathcal{N}$, we can write

$$(\star Q_K)^{vr} = q^{(0)} + \dots \tag{3.22}$$

where $q^{(0)}$ is a sum of primitive monomials for which *all* primitive factors have boost weight 0,[20] and the ellipsis represents a sum of primitive monomials, each containing at least one primitive factors of strictly positive boost weight and at least one primitive factor of strictly negative boost weight (since the total boost weight is 0). The Iyer-Wald entropy is defined as the integral over a horizon cross-section of $2\pi\sqrt{\mu}q^{(0)}$ [2]. The terms in the ellipsis of (3.22) are $O(\lambda^2)$ on C and therefore their λ-derivative vanishes on C for $\lambda = 0$, so $q^{(0)}$ is the only surviving term on C after we take a variation off our stationary background $g|_{\lambda=0}$. Therefore we have derived

$$\int_C \left(\frac{1}{2}\frac{\partial}{\partial\lambda}(q^{(0)}\sqrt{\mu}) + \sqrt{\mu} X^{(0)} \frac{\partial}{\partial\lambda} Y^{(0)} \right)_{\lambda=0} = 0. \tag{3.23}$$

By construction, the integrand here has the general form

$$\left(\frac{1}{2\sqrt{\mu}}\frac{\partial}{\partial\lambda}(q^{(0)}\sqrt{\mu}) + X^{(0)} \frac{\partial}{\partial\lambda} Y^{(0)} \right)_{\lambda=0} = \sum_{p,q} A_p^{(0)B_1\dots B_q}[\bar{\psi}] \bar{D}_{(B_1} \cdots \bar{D}_{B_q)} (\partial_v \partial_r)^p \frac{\partial}{\partial\lambda}\psi|_{\lambda=0}, \tag{3.24}$$

where $\bar{D}_A$ is the covariant derivative of $\bar{\mu}_{AB} = \mu_{AB}|_{\lambda=0}$. The integral over C of this quantity vanishes for completely arbitrary $\delta\psi = \frac{\partial}{\partial\lambda}\psi|_{\lambda=0}$, and in particular the restrictions $(\partial_v\partial_r)^p \delta\psi|_C$ can be chosen independently. Therefore we can write (still on C)

$$\left(\frac{1}{2\sqrt{\mu}}\frac{\partial}{\partial\lambda}(q^{(0)}\sqrt{\mu}) + X^{(0)} \frac{\partial}{\partial\lambda} Y^{(0)} \right)_{\lambda=0} = \bar{D}_A \bar{W}^{(0)A} \tag{3.25}$$

with

$$\bar{W}^{(0)C} = \sum_{j,p,q} (-1)^j \bar{D}_{(B_{j+1}} \cdots \bar{D}_{B_q} A_p^{(0)CB_1\dots B_q}[\bar{\psi}] \bar{D}_{B_1} \cdots \bar{D}_{B_j)} (\partial_v\partial_r)^p \frac{\partial}{\partial\lambda}\psi|_{\lambda=0}, \tag{3.26}$$

We will now argue that (3.25) holds not just on C but on any constant v cross section $C(v)$ of the horizon. To this end we note that evaluating the integrals over $C(v)$ instead of C is equivalent to evaluating them over the variation $T_v^* \delta\psi$ instead of $\delta\psi$, where T_v is a translation by v. This is because, on $\mathcal{N}$, all boost weight zero quantities are invariant under T_v in the background $\bar{\psi}$. But (3.25) is an identity that holds for arbitrary variations $\delta\psi$ of compact support and in particular also for $T_v^*\delta\psi$. Hence we conclude that this equation holds on $C(v)$. Then we can take a derivative with respect to v and note that this equation becomes

$$\frac{\partial}{\partial\lambda}\left[\frac{1}{2\sqrt{\mu}}\partial_v(q^{(0)}\sqrt{\mu}) + X^{(0)}\partial_v Y^{(0)} \right]_{\lambda=0} = \frac{\partial}{\partial\lambda}(D_A V^{(1)A})_{\lambda=0} \tag{3.27}$$

[20] By (3.9), we may determine $q^{(0)}$ by considering all contributions to E_R^{rvrv} made up exclusively of boost weight 0 primitive factors [2].

JHEP08(2022)258

where

$$V^{(1)C} = \sum_{j,p,q} (-1)^j D_{(B_{j+1}} \cdots D_{B_p} A_p^{(0)CB_1\ldots B_q}[\psi] D_{B_1} \cdots D_{B_j)} (\partial_v \partial_r)^p \partial_v \psi. \tag{3.28}$$

Thus, in a GNC system, the first variation of E_{vv} around a background $\bar{\psi}$ of the form described above can be written on $\mathcal{N}$ as

$$\frac{\partial}{\partial\lambda} E_{vv}|_{\lambda=0} = \frac{\partial}{\partial\lambda} \left(\partial_v \left\{ \frac{1}{\sqrt{\mu}} \partial_v (\sqrt{\mu} s^v) + D_A s^A \right\} \right)_{\lambda=0} \tag{3.29}$$

where

$$s^v := -\frac{1}{2} q^{(0)} + \sum_{j\geq 0, p-q>0, \psi} X_{q,\psi}^{(-p+q)A_1\ldots A_j} D_{(A_1} \cdots D_{A_j)} \partial_v^p \partial_r^q \psi \tag{3.30}$$

with $q^{(0)}$ as in (3.22), $X_{q,\psi}^{(k)}$ as in (3.15) for $A^{(2)} = E_{vv}$, and where

$$s^A := W^{(1)A} - V^{(1)A}, \tag{3.31}$$

with $V^{(1)}$ as in (3.28) and $W^{(1)}$ as in (3.15) for $A^{(2)} = E_{vv}$. In deriving this result we assumed a perturbation of compact support. If we have instead a perturbation of non-compact support then, for any compact subset of $\mathcal{N}$, we can define a perturbation of compact support that agrees with the original perturbation on this subset and therefore the above result holds for the original perturbation on this subset. Since the subset is arbitrary, it follows that this result holds on all of $\mathcal{N}$ even for perturbations of non-compact support.

If $\mathcal{N}$ is a black hole event horizon then, when substituted into (1.13), the first term of (3.30) gives the Iyer-Wald (IW) entropy [2]. The other terms in (3.30) are those predicted by the argument of Wall [4]. We will therefore refer to (3.30) as the Iyer-Wald-Wall (IWW) entropy density. The need for the term in (3.29) involving s^A was pointed out in [12].

We will now investigate E_{vv} at the fully non-linear level. Consider a metric $g_{\mu\nu}(x^\alpha)$ (not necessarily a solution) in GNCs with corresponding $\psi \in \{\alpha, \beta_A, \mu_{AB}\}$, define $\bar{\psi}$ as the series in (3.16), choose $\delta\psi$ to be $\psi - \bar{\psi}$ and define $g_{\mu\nu}(\lambda, x^\alpha)$ to be the 1-parameter family of metrics in GNCs corresponding to $\psi(\lambda)$ as in (3.16). By construction, this family interpolates between the stationary metric $g|_{\lambda=0}$ and $g|_{\lambda=1}$ which is equal to the given metric g. Now let

$$F[\psi] \equiv E_{vv} - \partial_v \left\{ \frac{1}{\sqrt{\mu}} \partial_v (\sqrt{\mu} s^v) + D_A s^A \right\}. \tag{3.32}$$

Using (3.15) and the expressions for s^v and s^A gives (recall definition 3.1)

$$F \sim \partial_v \left[\frac{1}{2\sqrt{\mu}} \partial_v (q^{(0)} \sqrt{\mu}) + X^{(0)} \partial_v Y^{(0)} - D_A V^{(1)A} \right]. \tag{3.33}$$

We expand F to all orders in λ around the background $\lambda = 0$. The zeroth order term must vanish on $\mathcal{N}$ since all positive boost weight quantities vanish on $\mathcal{N}$ in the background. The linear term must vanish on $\mathcal{N}$ by (3.27) or (3.29). Hence F contains only terms of quadratic or higher order in λ.

JHEP08(2022)258

If we consider a typical monomial in F then it is a product of factors with different boost weights. If we evaluate this on C then the zero boost weight factors coincide with those of $\bar{\psi}$. For a non-zero boost weight factor we have $D_{A_1}\dots D_{A_j}\partial_v^p\partial_r^q\psi(\lambda)|_C = \lambda\bar{D}_{A_1}\dots\bar{D}_{A_j}\partial_v^p\partial_r^q\delta\psi|_C$ if $p\neq q$. Thus the power of λ in any given monomial equals the number of factors in that monomial that have non-zero boost weight. Since non-zero boost weight arises from v or r derivatives, the number of such factors will be bounded if the total number of derivatives in each monomial of E_{vv} is bounded above (as is the case if we consider an EFT in which we include terms in the Lagrangian only up to some fixed total number of derivatives). In such a case, the expansion in λ terminates on C.

We will now show that $F\sim 0$. Note that the expression in square brackets in (3.33) involves no terms of boost weight higher than 1. Therefore expanding out the v-derivative gives

$$F\sim\sum_{n,q,\psi}A^{(0)A_1\dots A_n}_{np\psi}[\psi]D_{(A_1}\dots D_{A_n)}\partial_v^{q+2}\partial_r^q\psi \tag{3.34}$$

where each $A^{(0)}$ is a sum of monomials within which each factor has zero boost weight. Quantities that are ignored in the $\sim$ relation involve monomials with at least two positive boost weight factors, and such terms are of quadratic order, or higher, in λ. Hence

$$F|_C=\lambda\sum_{n,q,\psi}A^{(0)A_1\dots A_n}_{np\psi}[\bar{\psi}](\bar{D}_{(A_1}\dots\bar{D}_{A_n)}\partial_v^{q+2}\partial_r^q\delta\psi)_C+O(\lambda^2)\,. \tag{3.35}$$

However, as we have explained above, the terms in F linear in λ must cancel for *any* ψ. In particular, if we fix $\bar{\psi}$ and then construct ψ by specifying $\delta\psi$ then the linear terms above must cancel. Since this $\delta\psi$ is arbitrary, this implies $A^{(0)A_1\dots A_n}_{np\psi}[\bar{\psi}]=0$ on C. Our choice of $\bar{\psi}$ here was arbitrary so this must hold for any $\bar{\psi}$. Now given ψ, let $\psi_v=T_v^*\psi$ where T_{v_0} is the diffeomorphism corresponding to translation in v by v_0. We then have

$$0=A^{(0)A_1\dots A_n}_{np\psi}[\overline{\psi_v}]|_C=A^{(0)A_1\dots A_n}_{np\psi}[\psi_v]|_C=A^{(0)A_1\dots A_n}_{np\psi}[\psi]|_{C(v)} \tag{3.36}$$

where $C(v)$ is a constant v cut of $\mathcal{N}$. But v is arbitrary here so it follows that $A^{(0)A_1\dots A_n}_{np\psi}[\psi]$ vanishes on all of $\mathcal{N}$. This implies $F\sim 0$ as claimed.

We summarize the discussion so far in the following lemma:

Lemma 3.1. *Consider a null surface $\mathcal{N}$ ruled by affinely parameterized null geodesics and a generally covariant Lagrangian $\boldsymbol{L}$. Define a foliation $\{C(v)\}$ adapted to GNCs. Then we can define s^v, s^A and F satisfying (3.32), where:*

- *Each s^v, s^A, F is a sum of primitive monomials.*
- *Each primitive monomial in s^v, s^A, F has boost weight $0, 1, 2$, respectively.*
- *Each primitive monomial in F contains at least two primitive factors with strictly positive boost weight.*
- *s^v contains those primitive monomials in E_R^{rvrv} which are a product of primitive factors with zero boost weight (Iyer-Wald dynamical entropy), and all other primitive*

JHEP08(2022)258

monomials in s^v contain precisely one primitive factor with positive boost weight. s^A consists of a sum of primitive monomials of boost weight 1, each of which has precisely one primitive factor of boost weight 1.

This justifies equation (1.9). The main novelty of our approach compared to earlier work [4, 11, 12] has been to demonstrate the connection to the Iyer-Wald entropy. Note that we could add extra terms to s^v and s^A that are of quadratic (or higher) order in primitive factors of positive boost weight without affecting the above results. This will be important below.

3.3 Transformation of IWW entropy under a change of GNCs

The above procedure for determining s^v and s^A worked with the GNC components of the metric. Even though GNCs are defined geometrically starting from a cross section C of $\mathcal{N}$, their construction requires apart from the chosen cut C a choice of affine parameter for each null generator of $\mathcal{N}$. Under the remaining reparameterization freedom, $\alpha, \beta_A, \mu_{AB}$ and their derivatives in general transform in a very complicated way (e.g. equation (2.12)), so it is not clear that for a given cut C, s^v is fully covariant, i.e. a functional of $g_{\mu\nu}$ that is independent of arbitrary choices apart from the cut C. Thus, we must investigate if, and how, s^v transforms under a change of GNCs.

Consider a change of coordinates $x^\mu = x^\mu(x'^\alpha)$ in a neighborhood of $\mathcal{N}$ preserving the Gaussian Null Form, see (2.4). Thus, on $\mathcal{N}$, we have $(v, x^A) = (a(x'^C)v', x'^A)$, and off of $\mathcal{N}$, the — in general very complicated — form of the coordinate transformation is described in section 2.1. The GNC components of g in the primed coordinates are denoted by $\psi' \in \{\mu'_{A'B'}, \beta'_{A'}, \alpha'\}$. We want to understand the behavior of s^v under such a transformation. If the multiplication factor a in (2.4) is constant, then this behavior is trivially determined by the boost weight of each quantity, so we get $s^v[\psi] = s^v[\psi']$ (a constant), because s^v has boost weight 0. But for non-constant a, the transformation is not at all evident. In this subsection, we shall show (proposition 1) that s^v can be adjusted by terms quadratic in positive boost weight quantities, and a total divergence (which does not affect the total entropy), such that the modified s^v is manifestly invariant under a change of GNCs.

To investigate this, we first define $\boldsymbol{l} = g_{\mu\nu} l^\nu \mathrm{d}x^\mu$, and define, uniquely, an $(n-1)$ form $\boldsymbol{\epsilon}_l$ on $\mathcal{N}$ by demanding that $\boldsymbol{l} \wedge \boldsymbol{\epsilon}_l = \boldsymbol{\epsilon}$ and $n \cdot \boldsymbol{\epsilon}_l = 0$. In our GNCs this gives $\boldsymbol{\epsilon}_l = \sqrt{\mu}\mathrm{d}v \wedge \mathrm{d}x^1 \ldots \wedge \mathrm{d}x^{n-2}$. The constraint $(n-1)$ form $\boldsymbol{C}_l$ is defined as in (3.14), giving $\boldsymbol{C}_l = 2E_{vv}\boldsymbol{\epsilon}_l$. Similarly we define $\boldsymbol{F}_l = F\boldsymbol{\epsilon}_l$. We now define the entropy current as a $(n-2)$ form $\boldsymbol{s}$ on $\mathcal{N}$ by

$$\begin{aligned} \boldsymbol{s} &:= (s^v\partial_v + s^A\partial_A) \cdot \boldsymbol{\epsilon}_l \\ &= s^v\sqrt{\mu}\mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^{n-2} + \sum_{A=1}^{n-2}(-1)^A s^A \sqrt{\mu}\mathrm{d}v \wedge \mathrm{d}x^1 \wedge \cdots \widehat{\mathrm{d}x^A} \wedge \cdots \mathrm{d}x^{n-2}. \end{aligned} \tag{3.37}$$

In forms notation, (3.32) can be rewritten as

$$\mathcal{L}_l \mathrm{d}\boldsymbol{s} - \frac{1}{2}\boldsymbol{C}_l = \boldsymbol{F}_l \quad \text{on } \mathcal{N}. \tag{3.38}$$

JHEP08(2022)258

In our second set of GNCs we have $l'^\mu = al^\mu$ on $\mathcal{N}$ and so the primed version of the above equation gives

$$\mathcal{L}_{al}\mathrm{d}\boldsymbol{s}' - \frac{1}{2}\boldsymbol{C}_{al} = \boldsymbol{F}'_{al} \quad \text{on } \mathcal{N}. \tag{3.39}$$

In this formula, $\boldsymbol{s}'$ is the same functional as $\boldsymbol{s}$ but evaluated on the transformed $\psi' \in \{\mu'_{AB}, \beta'_A, \alpha'\}$ and in the transformed coordinates. $\boldsymbol{s}'$ can be transformed back to the original coordinates in the usual way using the Jacobian of the coordinate transformation.

Since $\boldsymbol{C}_l$ is defined by contracting l with a local, covariant tensor, we have $\boldsymbol{C}_{al} = a\boldsymbol{C}_l$. We also have $\mathcal{L}_{al}\mathrm{d}\boldsymbol{s}' = \mathrm{d}(al \cdot \mathrm{d}\boldsymbol{s}') = a\mathcal{L}_l\mathrm{d}\boldsymbol{s}' + \mathrm{d}a \wedge l \cdot \mathrm{d}\boldsymbol{s}' = a\mathcal{L}_l\mathrm{d}\boldsymbol{s}'$ where the final equality uses $l \cdot \mathrm{d}\boldsymbol{s}' \propto \mathrm{d}x'^1 \wedge \ldots \wedge \mathrm{d}x'^{(n-2)}$ and $\mathrm{d}a = (\partial_{A'}a)\mathrm{d}x^{A'}$. Combining these results gives

$$\mathcal{L}_l\mathrm{d}(\boldsymbol{s} - \boldsymbol{s}') = \boldsymbol{F}_l - \frac{1}{a}\boldsymbol{F}'_{l'} \quad \text{on } \mathcal{N}. \tag{3.40}$$

We shall now show that this equation implies that the IWW entropy is gauge invariant. In our first set of GNCs, let $\tilde{\psi}$ arise from a "background" metric (not necessarily a solution) for which all positive boost weight primitive factors vanish on $\mathcal{N}$. This need not be a background constructed as in the previous subsection (e.g. negative boost weight quantities need not vanish on C). The most interesting case is when $\tilde{\psi}$ corresponds to a stationary black hole solution with event horizon $\mathcal{N}$. Now we perturb this background: let $\psi(\lambda)$ be a 1-parameter family such that $\psi(\lambda = 0) = \tilde{\psi}$. Let the corresponding metric be $g_{\alpha\beta}(\lambda, x^\mu)$.

Consider transforming this metric to the "primed" set of GNCs. These are defined by a function $a > 0$ on the cut C. On $\mathcal{N}$ the coordinate transformation does not depend on λ. However, since the definition of GNCs depends on the metric, the change of coordinates does depend on λ away from $\mathcal{N}$: $x^\mu = x^\mu(\lambda, x'^\alpha)$. Under this coordinate transformation, we obtain $g = g'_{\alpha\beta}(\lambda, x'^\mu)\mathrm{d}x'^\alpha\mathrm{d}x'^\beta$, corresponding to $\psi'(\lambda, x'^\mu) \in \{\mu'_{A'B'}(\lambda, x'^\mu), \beta'_{A'}(\lambda, x'^\mu), \alpha'(\lambda, x'^\mu)\}$. The expression for derivatives $D'_{(A_1} \cdots D'_{A_j)}\partial^p_{v'}\partial^q_{r'}\psi'(\lambda, x'^\mu)$ in terms of $D_{(A_1} \cdots D_{A_j)}\partial^p_v\partial^q_r\psi(\lambda, x^\mu)$ will in general be very complicated on $\mathcal{N}$ (with explicit factors of v but not r), but it follows from lemma 2.1 that a positive boost weight primitive factor always transforms to an expression only involving positive boost weight primitive factors times appropriate powers of v (which has boost weight -1). Hence $\psi'(\lambda = 0)$ has the property that positive boost weight primitive factors vanish on $\mathcal{N}$. As a consequence, because $\boldsymbol{F}_l$ is at least quadratic in positive boost weight primitive factors, we obtain from (3.40)

$$\mathcal{L}_l\mathrm{d}\frac{\partial}{\partial\lambda}(\boldsymbol{s} - \boldsymbol{s}')|_{\lambda=0} = 0. \tag{3.41}$$

Now assume that the perturbation $\partial\psi(\lambda, x^\mu)/\partial\lambda|_{\lambda=0}$ has compact support (we relax this below). Integrating up the above equation gives

$$\mathrm{d}\frac{\partial}{\partial\lambda}(\boldsymbol{s} - \boldsymbol{s}')|_{\lambda=0} = 0. \tag{3.42}$$

We now use the algebraic Poincaré lemma (lemma 2.6) taking $\Psi_\lambda = \lambda(\partial_\lambda\psi|_{\lambda=0})$ and $\Phi = (\tilde{\psi}, a)$. This gives

$$\frac{\partial}{\partial\lambda}(\boldsymbol{s} - \boldsymbol{s}')|_{\lambda=0} = \mathrm{d}\boldsymbol{b}_{n-3}, \tag{3.43}$$

JHEP08(2022)258

where $\boldsymbol{b}_{n-3}$ is a local form depending on $\partial_\lambda \psi|_{\lambda=0}$, $\tilde{\psi}$ and a. Next we investigate $(\boldsymbol{s}-\boldsymbol{s}')|_{\lambda=0}$. Because all positive boost quantities built from $\psi(\lambda=0,x^\mu)$ vanish on $\mathcal{N}$, we have $s^A|_{\lambda=0} = 0$ and $s^v|_{\lambda=0} = (1/2)q^{(0)}[\psi|_{\lambda=0}]$ which is proportional to $E_R^{rvrv}[\psi|_{\lambda=0}]$ (as $q^{(0)}$ only contains zero boost weight primitive factors), Then, because $E_R^{\mu\nu\sigma\rho}$ is a covariant tensor we easily see

$$(\boldsymbol{s}-\boldsymbol{s}')|_{\lambda=0} = 0 \qquad \text{on } \mathcal{N}. \tag{3.44}$$

Combining (3.43) and (3.44) we find

$$\boldsymbol{s}-\boldsymbol{s}' = \lambda \mathrm{d}\boldsymbol{b}_{n-3} + O(\lambda^2) \qquad \text{on } \mathcal{N}. \tag{3.45}$$

Integrating this equation over C gives

$$S_{\mathrm{IWW}}[C] - S'_{\mathrm{IWW}}[C] = O(\lambda^2) \tag{3.46}$$

where the l.h.s. refers to the IWW entropy defined w.r.t. the two different sets of GNCs. Hence we have shown that *the IWW entropy is gauge invariant to linear order* in perturbations around our background metric. Note that this is an off-shell result: neither the background nor the perturbation need satisfy any equations of motion. In deriving this result we assumed that the perturbation has compact support. For a non-compactly supported perturbation we can pick a perturbation of compact support that agrees with the given perturbation in a neighbourhood of C and then apply the above result.

We shall now extend this to a fully nonlinear result, eliminating any reference to the background metric. We assume that we have some given metric $g_{\alpha\beta}$ and a set of GNCs, with corresponding quantities ψ. We now *define* a background metric $\tilde{\psi}$ such that, on C, primitive factors of zero *or negative* boost weight agree with those of our original metric. We do this by setting

$$\tilde{\psi} := \sum_{q\geq p} \frac{1}{p!q!}\chi(c_p v)\chi(c_q r) v^p r^q \partial_v^p \partial_r^q \psi|_{r=v=0} \tag{3.47}$$

where χ is of compact support, with χ identically equal to 1 in a neighbourhood of 0, and where c_p goes to infinity suitably rapidly so as to make the sums convergent. The corresponding metric $\tilde{g}_{\alpha\beta}$ is our background metric. We now define the one parameter family ($\lambda \in [0,1]$)

$$\psi(\lambda) := \tilde{\psi} + \lambda \delta\psi, \tag{3.48}$$

where $\delta\psi$ is of compact support. We let $g_{\alpha\beta}(\lambda, x^\mu)$ denote the corresponding family of metrics in GNCs. We choose $\delta\psi$ such that $\delta\psi = \psi - \tilde{\psi}$ in a neighbourhood of C so, on C, all derivatives of $g_{\alpha\beta}(\lambda=1,x^\mu)$ coincide with those of the original metric. Since we will in the end be interested in the behavior of s^v on the chosen C, we may therefore replace $g_{\alpha\beta}(x^\mu)$ with $g_{\alpha\beta}(\lambda=1,x^\mu)$.

We now apply (3.45) to $g_{\alpha\beta}(\lambda,x^\mu)$. By construction, the $O(\lambda^2)$ terms must on C be at least quadratic in positive boost weight quantities. This is because each factor of λ must be accompanied by $\delta\psi$, and on C the only non-vanishing quantities linear in $\delta\psi$ are those with positive boost weight (e.g. $D^k \partial_v^p \partial_r^q \delta\psi|_C = 0$ if $q \geq p$). Furthermore, in

JHEP08(2022)258

such quantities we can replace $\delta\psi$ with ψ (e.g. $D^k\partial_v^p\partial_r^q\delta\psi|_C = D^k\partial_v^p\partial_r^q\psi|_C$ for $p > q$). Similarly, on C we can write terms involving $\tilde{\psi}$ in terms of ψ using the definition of $\tilde{\psi}$ (e.g. $D^k\partial_v^p\partial_r^q\tilde{\psi}|_C = D^k\partial_v^p\partial_r^q\psi|_C$ for $q \geq p$). Thus we can write everything in (3.45) locally in terms of ψ (and a): the dependence on the background has been eliminated.

Now setting $\lambda = 1$, so that (3.45) is being evaluated for our original metric, we learn that there exists a local form $\boldsymbol{b}_{n-3}[\psi, a]$ on C such that $\boldsymbol{s} - \boldsymbol{s}' - \mathrm{d}\boldsymbol{b}_{n-3}$ is at least quadratic in positive boost weight quantities. We now dualize $\boldsymbol{b}_{n-3}$ by setting $(\boldsymbol{b}_{n-3})_{A_1\ldots A_{n-3}} = \epsilon[\mu]_{CA_1\ldots A_{n-3}}B^{vC}$, $(\boldsymbol{b}_{n-3})_{vA_1\ldots A_{n-4}} = \frac{1}{2}\epsilon[\mu]_{CDA_1\ldots A_{n-4}}B^{CD}$. We then have, on C,

$$s^{v'} - s^v = D_A B^{vA} + \ldots \qquad \frac{1}{a}s^{A'} - s^A = \frac{1}{\sqrt{\mu}}\partial_v(\sqrt{\mu}B^{vA}) + D_C B^{AC} + \ldots \tag{3.49}$$

where the ellipses represent terms that, when expressed w.r.t. the first set of GNCs, are at least quadratic in positive boost weight primitive factors. These terms depend on a in an extremely complicated way, but we will now show by a general argument using lemmas 2.2–2.5 that they can always be reabsorbed in a redefinition of s^v, thereby rendering this quantity invariant (up to total derivative) under a change of GNCs. For this purpose, it is useful to pass to an infinitesimal version of the transformation (3.49), which in view of (2.14) is fully equivalent to transformation law (3.49) under finite GNC coordinate transformations. Consider an a of the form $e^{t\Lambda/2}$ and take a derivative of the above equation with respect to t at $t = 0$. Calling the corresponding infinitesimal transformation γ (given in the basis of monomials of lemma 2.3 concretely by (2.34)), we see that, on C,

$$\gamma s^v[\psi] = D_A X^A[\psi, \Lambda] + Y[\psi, \Lambda], \tag{3.50}$$

where X^A, Y are local and Y is linear in Λ and at least quadratic in positive boost weight factors. Taking γ of this equation, we find a consistency condition on Y, which is, on C,

$$\gamma Y[\psi, \Lambda] = -D_A(\gamma X^A[\psi, \Lambda]). \tag{3.51}$$

We can use lemma 2.5 to characterize the solutions to this cohomological problem, identifying $Y\sqrt{\mu}\mathrm{d}^{n-2}x =: \boldsymbol{w}_{n-2}$, $-\gamma(\star X_A \mathrm{d}x^A) =: \boldsymbol{w}_{n-3}$. Lemma 2.5 made no assumption that $\boldsymbol{w}_{n-2}$ be quadratic in positive boost weight factors, but the method for constructing the solutions to the cohomological problem given in the proof is constructive, and we can see from the proof that the solution provided by the lemma can be chosen to be at least quadratic in positive boost weight: firstly, the forms in the descent equations constructed in the proof do not leave the space of local forms at least quadratic in positive boost weight factors and with overall boost weight zero. Indeed, the BRST operator γ preserves this space, and applying the algebraic Poincare lemma to construct the subsequent rungs in the descent equations, we do not leave this space either, as we can see from its proof. Secondly, the arguments used to analyze the bottom rung of the descent equations which relies on the method of contractible pairs can be carried out in that space, too, because the BRST operator γ, the exterior differential d on C, and the operators N, ρ [see (2.36), (2.37)] all commute with the boost weight number operator.[21]

[21] That is, the operator $N_b = \sum_{N,\bar{N},\psi}(N - \bar{N})\partial_v^N\partial_r^{\bar{N}}D_{(A_1}\cdots D_{A_j)}\psi\frac{\partial}{\partial(\partial_v^N\partial_r^N D_{(A_1}\cdots D_{A_j)}\psi)}$.

JHEP08(2022)258

Thus, lemma 2.5 gives, on C,

$$Y[\psi,\Lambda] = \gamma Z[\psi] + D_A U^A[\psi,\Lambda] + \Lambda V[\psi], \tag{3.52}$$

for local forms Z, U^A, V at least quadratic in positive boost weight and V invariant under changes of GNCs, and all quantities have overall boost weight zero. Therefore, on C,

$$\gamma s^v[\psi] = \gamma Z[\psi] + D_A(X^A[\psi,\Lambda] + U^A[\psi,\Lambda]) + \Lambda V[\psi], \tag{3.53}$$

which must hold for all Λ, ψ. The terms containing undifferentiated Λ in the BRST transformation γ (2.34) must cancel in the above expression because Z, s^v have boost weight 0, so the explicit term $\Lambda V[\psi]$ with Λ in undifferentiated form must be cancelled by the total divergence term. Thus, V has to be a total divergence of a local quantity of boost weight 0 which is at least quadratic in positive boost weight monomials, $V = D_A W^A$. Therefore

$$\gamma s^v[\psi] = \gamma(Z[\psi] - \beta_A W^A[\psi]) + D_A(X^A[\psi,\Lambda] + U^A[\psi,\Lambda] + \Lambda W^A[\psi]). \tag{3.54}$$

Now we redefine $s^v[\psi] \to s^v[\psi] - Z[\psi] + \beta_A W^A[\psi]$, which still satisfies (3.32) for a new $F[\psi]$ that is at least quadratic in positive boost weight, and which satisfies additionally $\gamma s^v[\psi] = D_A(X^A[\psi,\Lambda] + U^A[\psi,\Lambda] + \Lambda W^A[\psi])$. This already shows that our modified IWW entropy $\int_C s^v \sqrt{\mu}\mathrm{d}^{n-2}x$ is gauge invariant.

We can say more about the structure of the redefined s^v using lemmas 2.4 and 2.7 for $m = n-2$, setting $\boldsymbol{w}_{n-2} = s^v\sqrt{\mu}\mathrm{d}^{n-2}x$. The non-covariant terms $\boldsymbol{\beta} \wedge \boldsymbol{I}_{n-3}$ in lemma 2.4 involving the characteristic classes $\boldsymbol{I}_{n-3}$ as described in lemma 2.7 (which can only appear if n is odd) consist exclusively of monomials in the primitive basis such that each primitive factor has zero boost weight. Such terms are contained in the IW-part of the entropy — i.e. the term $q^{(0)}$ in (3.30) — which in turn is given by that part the manifestly covariant expression E_R^{rvrv} consisting only of zero boost weight terms. Thus, non-covariant terms $\boldsymbol{\beta} \wedge \boldsymbol{I}_{n-3}$ must in fact be absent in s^v and we have shown:

Proposition 1 (invariance of modified IWW-entropy density). *s^v can be modified by terms at least quadratic in positive boost weight [so it still satisfies the properties claimed in lemma 3.1 and (3.32)] in such a way that it is a sum of:*

- *A local functional of boost weight 0 that is a contraction of the following factors: $\mathcal{D}_{(A_1}\cdots\mathcal{D}_{A_j}K_{A)B}$, $\mathcal{D}_{(A_1}\cdots\mathcal{D}_{A_j}\bar{K}_{A)B}$, GNC components of $\nabla_{(\alpha_1}\cdots\nabla_{\alpha_j)}R[g]_{\mu\nu\sigma\rho}$ [by the Bianchi-identities, not all these components are independent, see lemma 2.3], μ^{AB}, or $\epsilon[\mu]_{A_1\ldots A_{n-2}}$.*
- *A total divergence $D_A\xi^A$, where ξ^A is a local functional of boost weight 0.*

In particular, on C, s^v is invariant up to at total divergence under a change of the affine parameter in GNCs.

Remark 2. Proposition 1 uses that $\boldsymbol{L}$ is locally and covariantly constructed out of the metric $g_{\mu\nu}$, meaning that it is a functional of $g_{\mu\nu}$, $\epsilon_{\mu_1\ldots\mu_n}$ and covariant derivatives of the Riemann tensor (in a purely gravitation theory). Hence proposition 1 does not cover the

JHEP08(2022)258

case of gravitational Lagrangians that are covariant only up to d-exact terms like Chern-Simons terms in odd dimensions n, such as $\mathrm{tr}\left(\boldsymbol{\Gamma}\wedge \mathrm{d}\boldsymbol{\Gamma}+\frac{2}{3}\boldsymbol{\Gamma}\wedge\boldsymbol{\Gamma}\wedge\boldsymbol{\Gamma}\right)$ in $n=3$, where $\boldsymbol{\Gamma}$ is the spin-connection associated with a 3-bein of $g_{\mu\nu}$. Such terms will result in terms of the type (2.45) in s^v. For example from a gravitational Chern-Simons term in $\boldsymbol{L}$ in $n=5$, we can see from the formulas for Christoffel symbols in GNCs in appendix A that we would get a Chern-Simons term of the form $\boldsymbol{\beta}\wedge\mathrm{d}\boldsymbol{\beta}$ in $s^v\sqrt{\mu}\mathrm{d}^3x$, consistent with lemmas 2.4 and 2.7.

On the other hand, fully covariant topological terms in $\boldsymbol{L}$ are covered by our analysis. For example, in any even dimension an Euler density in $\boldsymbol{L}$, i.e. $\boldsymbol{e}_n[g]:=\boldsymbol{R}[g]^{\mu_1\mu_2}\wedge\cdots\wedge\boldsymbol{R}[g]^{\mu_{n-1}\mu_n}\epsilon_{\mu_1\ldots\mu_n}$ will result in an Euler density term $\boldsymbol{e}_{n-2}[\mu]$ in $s^v\sqrt{\mu}\mathrm{d}^{n-2}x$, whereas a Chern class in $\boldsymbol{L}$, i.e. $\boldsymbol{c}_n[g]:=\boldsymbol{R}[g]^{\mu_1}{}_{\mu_2}\wedge\cdots\wedge\boldsymbol{R}[g]^{\mu_{n-1}}{}_{\mu_1}$ will result in a term of the form $\boldsymbol{c}_{n-4}[\mu]\wedge\mathrm{d}\boldsymbol{\beta}$ in $s^v\sqrt{\mu}\mathrm{d}^{n-2}x$. (Recall that $\mathrm{d}\boldsymbol{\beta}$ can be eliminated in favour of Riemann components using (2.25).) All of these terms are built from boost weight zero quantities and are already present in the IW-part of s^v.

4 Improved entropy current

4.1 Vacuum gravity

For simplicity we shall consider vacuum gravity in this subsection, i.e., no matter fields. (In the next subsection we shall consider gravity coupled to a scalar field.) If the spacetime dimensionality is odd then we assume parity symmetry, again for simplicity. Relaxing this assumption should be straightforward. We assume the Lagrangian takes the form[22]

$$\boldsymbol{L}=\left[-2\Lambda+R+\sum_{n\geq 2}\ell^n L_n\right]\boldsymbol{\epsilon} \tag{4.1}$$

with a single UV length scale ℓ. L_n is a local covariant scalar, depending only on the metric and its derivatives, and of dimension $n+2$, i.e., it involves $n+2$ derivatives of the metric. We define the dimension of ℓ to be -1 and the dimension of Λ to be $+2$. Each derivative ∂_μ carries one index, and $g_{\mu\nu}$ has two indices, so non-zero L_n is possible only for even n.[23] It will be convenient in this section to imagine that the L_n are known for all n. It is then not hard to obtain results for which only finitely many of the L_n are known, as we will discuss at the end of this section.

The Einstein equation has the form

$$R_{\mu\nu}-\frac{1}{2}Rg_{\mu\nu}+\Lambda g_{\mu\nu}=\sum_{n\geq 2}\ell^n H_{n\,\mu\nu} \tag{4.2}$$

where again only even n appears in the sum and $H_{n\,\mu\nu}$ has dimension $n+2$.

Recall from section 1.7 that we are not interested in all solutions of (4.2). Instead, we only consider solutions that fall within the regime of validity of EFT. This means that

[22] In this section we no longer use n to denote spacetime dimension.

[23] In an even number of dimensions, there may be dependence on the volume form (for a parity violating theory) but this has an even number of indices, which does not affect the argument. In an odd number of dimensions our assumption of parity symmetry excludes dependence on the volume form.

JHEP08(2022)258

we restrict attention to solutions that, in some set of GNCs near $\mathcal{N}$, are slowly varying compared to the scale ℓ, i.e., $\ell/L \ll 1$ where L is the length scale of variation of the fields. In the case of a dynamical black hole, L will be the minimum of the size of the black hole and any length/time scales associated with the dynamics. Writing this out formally we have

Definition 4.1 (validity of EFT condition). *We consider a 1-parameter family of smooth solutions of (4.2) with parameter L, such that $\mathcal{N}$ is a smooth null hypersurface for all members of the family. We assume that there exist GNCs defined near $\mathcal{N}$ such that, near $\mathcal{N}$, if T_n is a quantity of dimension n (in the sense of definition 2.3) constructed from $\{\alpha, \beta_A, \mu_{AB}\}$ and their derivatives then there is a dimensionless constant C_n (independent of L) such that $|T_n| \leq C_n/L^n$. We also require $|\Lambda|L^2 \leq 1$. Then EFT is valid for sufficiently small ℓ/L (i.e. large enough L).*

For example, on the r.h.s. of (4.2), $|H_{n\,\mu\nu}|$ is bounded above by C_n/L^{n+2} so $\ell^n H_{n\,\mu\nu} = \mathcal{O}(\ell^n/L^{n+2})$. As in section 1.7, from now on we shall not indicate the L-dependence explicitly below so we would write $\ell^n H_{n\mu\nu} = \mathcal{O}(\ell^n)$. Factors of L can be reinstated by dimensional analysis.

Recall the definition of F in equation (3.32). We can decompose this into a sum of terms arising from the various terms in (4.2)

$$F = F_0 + \sum_{n\geq 2} \ell^n F_n \tag{4.3}$$

where F_n has dimension $n+2$ and has the same general structure as described in lemma 3.1. As above, only even n can appear in the sum. This is because only even powers of ℓ appear in the Einstein equation.

F_0 is the expression coming from the Einstein-Hilbert part of the Lagrangian and reads

$$F_0 = -K_{AB}K^{AB}, \tag{4.4}$$

In the EFT spirit, we might expect F_0, which is sign definite, to dominate the higher order terms in ℓ. This would require in particular that if K_{AB} vanishes, then so should $F_n, n > 0$. However, this is not true in general. Nevertheless, we can still make progress with this idea. To do this we need to understand the structure of the terms F_n. We start by using the equation of motion (4.2) to obtain the following result:

Lemma 4.1. *On-shell, F can be written as in (4.3) where each F_n is a polynomial in $D_{(A_1}\cdots D_{A_j)}\partial_v^N K_{AB}$, $D_{(A_1}\cdots D_{A_j)}\partial_r^{\bar{N}}\bar{K}_{AB}$, $D_{(A_1}\cdots D_{A_j)}\beta_B$, $D_{(A_1}\cdots D_{A_j)}R[\mu]_{ABCD}$, μ^{AB}, $\epsilon[\mu]_{A_1\ldots A_{n-2}}$ and Λ. Each term in this polynomial contains at most $n+2$ derivatives.*

Proof. By lemma 2.2, we can write F in terms of $D_{(A_1}\cdots D_{A_j)}\partial_v^N K_{AB}$, $D_{(A_1}\cdots D_{A_j)}\partial_r^{\bar{N}}\bar{K}_{AB}$, $D_{(A_1}\cdots D_{A_j)}\beta_B$, $D_{(A_1}\cdots D_{A_j)}R[\mu]_{ABCD}$, μ^{AB}, $\epsilon[\mu]_{A_1\ldots A_{n-2}}$, and GNC components of covariant derivatives of the Ricci tensor. The idea now is to use the Einstein equation (4.2) to eliminate the terms involving the Ricci tensor. In doing this we replace each occurrence of a Ricci tensor either with a multiple of $\Lambda g_{\mu\nu}$ or with powers of ℓ^2 times curvature terms of

JHEP08(2022)258

higher dimensions. Then we decompose the GNC components of the latter terms again as in lemma 2.2, eliminate any occurrence of a Ricci tensor again and so forth. At each step, any Ricci terms arise with explicit factors of ℓ^2 and therefore get pushed to higher order in ℓ. The end result is that, on-shell, F can be written in terms of the quantities stated in the lemma, up to terms vanishing to infinite order in ℓ. Finally we can express this on-shell result in the form (4.3) where each F_n can be written in terms of the quantities just listed, i.e., explicit occurrence of Ricci tensors and their covariant derivatives have been eliminated.

Since F_n has dimension $n+2$, each term in F_n contains at most $n+2$ derivatives. Terms with fewer than $n+2$ derivatives must have enough factors of Λ to make the dimension up to $n+2$. □

So far we assumed that our EFT is defined to all orders in n. If only terms with $n \leq N-2$ are known then the Einstein equation will contain unknown "errors terms" of order ℓ^N (equation (1.14)). The off-shell expression for F will have a similar form. The above argument still works and any terms of order ℓ^N or higher can be absorbed into the error term in F. The on-shell expression for F will be a sum of terms F_n with $n \leq N-2$, each of the form just described, and an error term of order ℓ^N.

We will now argue that the on-shell structure of F can be further rearranged in a "nice" way. The argument involves induction on the order in ℓ:

Induction hypothesis. For each even $n \geq 0$, there are local tensors X_{nAB}, ς_n^v having boost weights $(1,0)$ and local tensors y_n^A, O_n, each depending polynomially on the quantities listed in lemma 4.1, such that on-shell we have

$$F = -(K_{AB} + X_{nAB})(K^{AB} + X_n^{AB}) - \sum_{j=0}^{n} \partial_v \left[\frac{1}{\sqrt{\mu}} \partial_v \left(\sqrt{\mu} \varsigma_j^v \right) \right] - \sum_{j=2}^{n} D_A y_j^A + O_{n+2}, \tag{4.5}$$

such that X_{nAB} is symmetric, $O_{n+2}, D_A y_j^A$, ς_n^v has the same general structure as F described in lemma 3.1 (in particular they are at least quadratic in positive boost weight quantities) and

$$O_{n+2} = \mathcal{O}(\ell^{n+2}) \tag{4.6}$$

as well as $X_{nAB} = \mathcal{O}(\ell^2)$. Furthermore ς_j^v and y_j^A depend explicitly on ℓ only through an overall factor ℓ^j.

Induction start ($n = 0$). In this case, $X_{0AB} = y_0^A = \varsigma_0^v = 0$ and $O_2 = \sum_{n \geq 2} \ell^n F_n = \mathcal{O}(\ell^2)$ so the induction hypothesis is true for $n = 0$.

Induction step ($n-2 \to n, n \geq 2$). Consider O_n written in terms of the quantities described in lemma 4.1. The induction hypothesis gives $O_n = \mathcal{O}(\ell^n)$ and that O_n is a sum of terms where each term is at least quadratic in terms of positive boost weight, and so must contain at least two factors of the form $D_{(A_1} \cdots D_{A_j)} \partial_v^{p-1} K_{AB}$ $(p \geq 1)$ (as other possible

JHEP08(2022)258

factors listed in lemma 4.1 have non-positive boost weight). In terms of μ_{AB} this means that O_n must take the schematic form

$$O_n = \sum_{k,k',p,p'} A_{n,k,p,k',p'} D^k \partial_v^p \mu D^{k'} \partial_v^{p'} \mu + O_{n+2} \tag{4.7}$$

where $p, p' > 0$ in all terms, the coefficient $A_{n,k,p,k',p'}$ has boost weight $2-p-p'$ and, aside from an overall factor of ℓ^n, depends only on β, $D^k\beta$, Λ, μ and derivatives of μ. O_{n+2} satisfies (4.6). Here and in the following, tensor indices $A, B, C, \ldots$ are suppressed. The total number of derivatives in each term of the above sum is at most $n+2$. Our aim is to show that the above expression for O_n can be rearranged to ensure that the induction hypothesis is true for n.

The next sequence of steps is designed to bring each term in the sum in (4.7) into the form $A_{n,k,p} D^k \partial_v^p \mu \partial_v \mu$, i.e. linear in $\partial_v \mu$ and at least quadratic in positive boost weight quantities, for a new $A_{n,k,p}$, possibly at the expense of further terms to be absorbed in O_{n+2}, or terms involving y_n^A, ς_n^v as in the induction hypothesis. This is done by moving, one by one, the derivatives $D^{k'}$ from the factor $D^{k'} \partial_v^{p'} \mu$ over to the other terms as if we were performing a partial integration, but keeping the "boundary terms". These boundary terms are then absorbed in y_n^A as in the induction hypothesis. As a result,

$$O_n = \sum_{k,p,p'} A_{n,k,p,p'} D^k \partial_v^p \mu \partial_v^{p'} \mu + D_A y_n^A + O_{n+2}, \tag{4.8}$$

with the summation indices subject to $p, p' > 0$. Next, we would like to bring $p'-1$ of the v-derivatives on $\partial_v^{p'} \mu$ in the sum over to the other terms, leaving us finally with expressions that are linear in $\partial_v \mu$ and quadratic in positive boost weight. This is done by a similar "partial integration", moving all boundary terms into a new quantity ς_n^v, as in the induction hypothesis, so in particular quadratic in positive boost weight quantities.[24] We can do this by an induction on $p'+p$ on the terms in the sum, lowering at each step this counter by at least one until we reach terms of the form $A_{n,k,p} D^k \partial_v^p \mu \partial_v \mu$ or $A_{n,k,p} D^k \partial_v \mu \partial_v^p \mu$ (where $p > 0$ in each case). The latter term then can be rewritten in the same form as the former by the same "partial integration" w.r.t. D argument that we just used, which generates further additions to y_n^A. To see how this induction works in more detail, note that the induction hypothesis is true for $p + p' \leq 3$ (since then either $p = 1$ or $p' = 1$) so assume $p + p' > 3$. We claim that there exist numbers a_j such that

$$\partial_v \left[\frac{1}{\sqrt{\mu}} \partial_v \left(\sqrt{\mu} A_{n,k,p,p'} \sum_{j=1}^{p'+p-3} a_j D^k \partial_v^j \mu \partial_v^{p+p'-2-j} \mu \right) \right] = A_{n,k,p,p'} D^k \partial_v^p \mu \partial_v^{p'} \mu + \ldots, \tag{4.9}$$

where the ellipsis represent any terms of the form $A_{n,\bar{k},\bar{p},\bar{k}',\bar{p}'} D^{\bar{k}} \partial_v^{\bar{p}} \mu \partial_v^{\bar{p}'} \mu$ having $\bar{p}+\bar{p}' < p+p'$, $\bar{p}, \bar{p}' > 0$ (which can be dealt with inductively), or terms having $\bar{p}+\bar{p}' = p+p'$, $\bar{p}, \bar{p}' > 0$, but

[24] Note that we might encounter ∂_v acting on a term $D^j \beta_A$ or $D^j \partial_r^q \mu$ contained in $A_{n,k,p,k',p'}$. We deal with such terms as in the proof of lemma 4.1, i.e., converting them to the basis of lemma 2.2 and using the Einstein equation to eliminate any Ricci terms by expressing them in terms of Λ and quantities of order ℓ^2: the latter terms can be absorbed into O_{n+2} (as $A_{n,k,p,k',p'}$ already contains a factor ℓ^n).

JHEP08(2022)258

$\bar{p}' = 1$ or $\bar{p} = 1$. We will justify this claim by showing that in order for the terms represented by the ellipsis to have the desired property, the a_j's must satisfy a linear system of algebraic equations that has a unique solution. Expanding out the l.h.s. of (4.9), we find

$$\begin{aligned}\text{l.h.s. of (4.9)} = A_{n,k,p,p'} \Bigg[& a_1 D^k \partial_v \mu \partial_v^{p'+p-1} \mu + (a_2 + 2a_1) D^k \partial_v^2 \mu \partial_v^{p'+p-2} \mu \\ & + \sum_{j=3}^{p'+p-3} (a_j + 2a_{j-1} + a_{j-2}) D^k \partial_v^j \mu \partial_v^{p'+p-j} \mu \\ & + (2a_{p'+p-3} + a_{p'+p-4}) D^k \partial_v^{p'+p-2} \mu \partial_v^2 \mu \\ & + a_{p'+p-3} D^k \partial_v^{p'+p-1} \mu \partial_v \mu \Bigg] + \dots . \end{aligned} \tag{4.10}$$

Any v-derivatives of $D^j \beta$ or $D^j \partial_r^q \mu$ (arising from v-derivatives of $A_{n,k,p,p'}$) are dealt with as explained in footnote 24, and contribute to the terms of the form $A_{n,\bar{k},\bar{p},\bar{k}',\bar{p}'} D^{\bar{k}} \partial_v^{\bar{p}} \mu \partial_v^{\bar{p}'} \mu$ with $\bar{p} + \bar{p}' < p + p'$ which are denoted by the ellipsis. (The v-derivatives of $\sqrt{\mu}$ also generate such terms, and further such terms arise from commutators of D and ∂_v.) Now we require that

$$\begin{aligned}\text{l.h.s. of (4.9)} = A_{n,k,p,p'} \Big[& a_1 D^k \partial_v \mu \partial_v^{p'+p-1} \mu + D^k \partial_v^p \mu \partial_v^{p'} \mu \\ & + a_{p'+p-3} D^k \partial_v^{p'+p-1} \mu \partial_v \mu \Big] + \dots . \end{aligned} \tag{4.11}$$

Then the $p' + p - 3$ coefficients a_j must satisfy the following $p' + p - 3$ linear equations

$$\begin{aligned} a_2 + 2a_1 &= 0, \\ a_j + 2a_{j-1} + a_{j-2} &= 0, \qquad j \neq p \\ a_p + 2a_{p-1} + a_{p-2} &= 1, \\ 2a_{p'+p-3} + a_{p'+p-4} &= 0 . \end{aligned} \tag{4.12}$$

The coefficient matrix of this linear system has the form

$$M_{p'+p-3} = \begin{pmatrix} 2 & 1 & 0 & 0 & \dots & 0 \\ 1 & 2 & 1 & 0 & \dots & 0 \\ 0 & 1 & 2 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & 0 & 1 & 2 \end{pmatrix} . \tag{4.13}$$

It is easy to check that such a matrix has a nonzero determinant. In fact, the determinant of an $N \times N$ matrix of this form is $N + 1$ which is found by induction on N: expanding the determinant along the first row of the matrix gives the simple recurrence relation $\det M_{N+1} = 2 \det M_N - \det M_{N-1}$. Therefore, $M_{p'+p-3}$ is invertible and the linear system above has a unique solution for the coefficients a_j so we can indeed rewrite terms of the form $A_{n,k,p,p'} D^k \partial_v^p \mu \partial_v^{p'} \mu$ as in (4.9) (with the first and last terms in square brackets in (4.11)

JHEP08(2022)258

contributing to the ellipsis in (4.9)). The terms $a_j A_{n,k,p,p'} D^k \partial_v^j \mu \partial_v^{p+p'-2-j} \mu$ are absorbed in ς_n^v. We repeat the procedure until we have obtained

$$O_n = \sum_{k,p} A_{n,k,p} D^k \partial_v^p \mu \partial_v \mu + \partial_v \left[\frac{1}{\sqrt{\mu}} \partial_v \left(\sqrt{\mu} \varsigma_n^v \right) \right] + D_A y_n^A + O_{n+2}, \tag{4.14}$$

where $p > 0$. Since $A_{n,k,p,k',p'}$ depend only on $D^k \beta$ and derivatives of μ, the quantity ς_n^v have the same dependency on μ and β as in the induction hypothesis.

The first term in (4.14) depends linearly on K_{AB} so, reinstating indices, we can write it as $-2K^{AB} \Delta X_{AB}$ where ΔX_{AB} is symmetric in AB and of order ℓ^n because all of the terms above, except O_{n+2}, depend on ℓ only via an overall factor of ℓ^n. Recalling our original induction hypothesis we can now "recomplete the square" on K_{AB} as follows:

$$-(K_{AB} + X_{n-2\,AB})(K^{AB} + X_{n-2}^{AB}) - 2K^{AB} \Delta X_{AB} = -(K_{AB} + X_{nAB})(K^{AB} + X_n^{AB}) + \mathcal{O}(\ell^{n+2}) \tag{4.15}$$

where $X_n \equiv X_{n-2} + \Delta X$ and we used $X_{n-2} = \mathcal{O}(\ell^2)$ (from our induction hypothesis). The error term above can be absorbed into O_{n+2}. This closes the induction loop.

The inductive structure of the term F can be combined with lemma 3.1 to get the following central result of this section, proposition 2. We set, for each given even $n \geq 2$

$$S_n^v := s^v + \sum_{j=2}^{n} \varsigma_j^v, \quad S_n^A := s^A, \quad Y_n^A = \sum_{j=2}^{n} y_j^A \tag{4.16}$$

and otherwise keep the same notations as in the inductive structure.

Proposition 2. *Consider a generally covariant EFT Lagrangian of the form* (4.1) *and a* 1*-parameter family of smooth solutions (with parameter* L*) satisfying our validity of EFT condition (definition 4.1) with a smooth null hypersurface* $\mathcal{N}$ *ruled by affinely parameterized null geodesics. Then in the GNCs of definition 4.1 on* $\mathcal{N}$ *we have the (on-shell) equation*

$$\partial_v \left[\frac{1}{\sqrt{\mu}} \partial_v \left(\sqrt{\mu} S_n^v \right) + D_A S_n^A \right] = -(K_{AB} + X_{nAB})(K^{AB} + X_n^{AB}) - D_A Y_n^A + O_{n+2}, \tag{4.17}$$

such that S_n^v, S_n^A, X_{nAB}, Y_n^A, O_{n+2} *have the same general structure as follows from lemma 3.1 and the induction hypothesis above. In particular,* $O_{n+2} = \mathcal{O}(\ell^{n+2})$, $X_{nAB} = \mathcal{O}(\ell^2)$ *is symmetric,* $Y_n^A = \mathcal{O}(\ell^2)$.

As discussed in section 1.7, in practice the EFT Lagrangian will not be known to all orders, instead only the terms with $n \leq N-2$ will be known, for some $N \geq 2$. The equation of motion will take the form (1.14). In this case we could repeat all of the above arguments, now with lemma 4.1 only determining F up to unknown terms of order ℓ^N and the above induction stopping when we reach $n = N-2$. Equivalently we could simply appeal to the above results applied to the full EFT, including the unknown terms with $n \geq N$, and note that these unknown terms will only affect the final answer at $\mathcal{O}(\ell^N)$.

JHEP08(2022)258

Either way, we obtain equation (1.16) by setting $n = N - 2$ in the above proposition (and dropping the n index on S^v, S^A, X_{AB} and Y^A).

We can now define the entropy by equation (1.17) and this will obey the second law to quadratic order, in the sense of EFT, as explained in section 1.7.

4.2 Scalar-tensor effective field theories

Now we briefly discuss how the previous construction (valid for vacuum gravity) is modified when we include the simplest form of matter: a real scalar field. We again exclude the case of parity violating theories in odd spacetime dimension. The Lagrangian for a scalar-tensor EFT takes the form

$$\boldsymbol{L} = \left[R - V(\phi) + X + \sum_{n\geq 2} \ell^n L_n \right] \boldsymbol{\epsilon} \tag{4.18}$$

where the scalar potential $V(\phi)$ has dimension $+2$ and the scalar kinetic term is given by $X \equiv -\frac{1}{2}(\partial\phi)^2$. Here, we assumed again that there is a single UV length scale ℓ (of dimension -1) suppressing the terms L_n which are therefore subleading corrections to Einstein's theory with a minimally coupled scalar. We further assume that L_n are diffeomorphism covariant scalars (of dimension $n+2$) that are locally constructed out of the metric g, the scalar field ϕ and derivatives of these fields. The fact that L_n is a local covariant scalar implies that n must be even. Moreover, in each monomial in L_n the number of derivatives acting on the metric and the scalar field must be $n+2$ in total.

The gravitational and scalar equations of motion have the form

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} - \frac{1}{2}\partial_\mu\phi\partial_\nu\phi - \frac{1}{2}(X - V)g_{\mu\nu} = \sum_{n\geq 2} \ell^n H_{n\,\mu\nu} \tag{4.19}$$

$$\Box\phi - V'(\phi) = \sum_{n\geq 2} \ell^n I_n \tag{4.20}$$

where n runs over even numbers and $H_{n\,\mu\nu}$ and I_n are local covariant tensors of dimension $n+2$ constructed out of the metric and the scalar field.

Validity of the EFT requires that we only consider 1-parameter families of smooth solutions to (4.19)–(4.20) that satisfy the conditions of definition 4.1, with the additions that any tensor field T_n of dimension n locally constructed out of α, β_A, μ_{AB}, ϕ and their derivatives must satisfy a bound $|T_n| \leq C_n/L^n$. Note in particular that this means that the scalar potential $V(\phi)$ must obey $L^2 d^k V/d\phi^k \leq 1$ for any $k \geq 0$.

Now we are in the position to state the corresponding result on the existence of an entropy current for theories of the form (4.18).

Proposition 3. *Consider a diffeomorphism-covariant scalar-tensor EFT with a Lagrangian given by (4.18) and a 1-parameter family of smooth solutions (with parameter L) of this theory. Suppose that the 1-parameter family of solutions fall within the regime of validity of the EFT (in the sense of definition 4.1 and the addition provided above) and the solutions possess a smooth null hypersurface $\mathcal{N}$ ruled by affinely parameterized null geodesics (for any L). Then in the GNCs of definition 4.1 on $\mathcal{N}$ there exist quantities $S^v_n, S^A_n, X_{nAB}, P_n, Y^A_n, O_{n+2}$ with the following properties:*

JHEP08(2022)258

(i) they satisfy an on-shell relation of the form

$$\partial_v \left[\frac{1}{\sqrt{\mu}} \partial_v \left(\sqrt{\mu} S_n^v \right) + D_A S_n^A \right] = -(K_{AB} + X_{nAB})(K^{AB} + X_n^{AB}) + \\ -\frac{1}{2}(\partial_v \phi + P_n)^2 - D_A Y_n^A + O_{n+2}; \qquad (4.21)$$

(ii) each of S_n^v, S_n^A, X_{nAB}, P_n, Y_n^A, O_{n+2} is a sum of monomials such that each monomial is a product of factors of $D_{(A_1} \cdots D_{A_j)} \partial_v^N K_{AB}$, $D_{(A_1} \cdots D_{A_j)} \partial_r^{\bar{N}} \bar{K}_{AB}$, $D_{(A_1} \cdots D_{A_j)} \partial_v^N \phi$, $D_{(A_1} \cdots D_{A_j)} \partial_r^{\bar{N}} \phi$, $D_{(A_1} \cdots D_{A_j)} \beta_B$, μ^{AB}, $D_{(A_1} \cdots D_{A_j)} R[\mu]_{ABCD}$, $\epsilon[\mu]_{A_1 \ldots A_{n-2}}$ and $d^k V/d\phi^k$;

(iii) S_n^v, S_n^A, X_{nAB}, P_n, Y_n^A, O_{n+2} have the same general structure as follows from lemma 3.1. In particular, $O_{n+2} = \mathcal{O}(\ell^{n+2})$, $X_{nAB}, P_n, Y_n^A = \mathcal{O}(\ell^2)$.

To prove this statement, one can follow the algorithm devised for vacuum gravity with minor modifications. Hence, to avoid repetition, we merely highlight the key differences in this section rather than giving a complete proof.

We start by noting that a primitive factor $D_{(A_1} \cdots D_{A_j)} \partial_v^p \partial_r^q \phi$ with $p, q \geq 1$ can be written on $\mathcal{N}$ in terms of the factors of μ^{AB}, $D_{(A_1} \cdots D_{A_j)} \partial_v^N K_{AB}$, $D_{(A_1} \cdots D_{A_j)} \partial_r^{\bar{N}} \bar{K}_{AB}$, $D_{(A_1} \cdots D_{A_j)} \partial_v^N \phi$, $D_{(A_1} \cdots D_{A_j)} \partial_r^{\bar{N}} \phi$, $D_{(A_1} \cdots D_{A_j)} \beta_B$, $D_{(A_1} \cdots D_{A_j)} R[\mu]_{ABCD}$, and covariant quantities of the form $\nabla_{(\alpha_1} \ldots \nabla_{\alpha_j)} \Box \phi$, $\nabla_{(\alpha_1} \ldots \nabla_{\alpha_j)} R_{\mu\nu}$ so that the expression has polynomial dependence on each of these quantities. To prove this statement, one starts with the expression of $\Box \phi$ in GNCs:

$$\Box \phi = 2\partial_r \partial_v \phi + \mu^{AB} D_A D_B \phi + K \partial_r \phi + \bar{K} \partial_v \phi + \beta^A D_A \phi + \ldots \qquad (4.22)$$

where the ellipsis stands for terms vanishing at $r = 0$. Taking derivatives of this identity and arguing inductively as in the proof of lemma 2.3 lets us eliminate all mixed $r - v$ derivatives of ϕ, in favour of the quantities listed above, establishing our claim.

The next step is to apply this result to F (defined in equation (3.32)) and use the gravitational and scalar equations of motion (4.19)–(4.20) to eliminate the dependencies on $\nabla_{(\alpha_1} \ldots \nabla_{\alpha_j)} \Box \phi$ and $\nabla_{(\alpha_1} \ldots \nabla_{\alpha_j)} R_{\mu\nu}$. This yields the scalar-tensor version of lemma 4.1: on-shell, F can be written as

$$F = F_0 + \sum_{n \geq 2} \ell^n F_n, \qquad \text{with} \qquad F_0 = -K_{AB} K^{AB} - \frac{1}{2}(\partial_v \phi)^2 \qquad (4.23)$$

where each F_n depends polynomially on μ^{AB}, $D_{(A_1} \cdots D_{A_j)} \partial_v^N K_{AB}$, $D_{(A_1} \cdots D_{A_j)} \partial_r^{\bar{N}} \bar{K}_{AB}$, $D_{(A_1} \cdots D_{A_j)} \partial_v^N \phi$, $D_{(A_1} \cdots D_{A_j)} \partial_r^{\bar{N}} \phi$, $D_{(A_1} \cdots D_{A_j)} \beta_B$, $D_{(A_1} \cdots D_{A_j)} R[\mu]_{ABCD}$, $\epsilon[\mu]_{A_1 \ldots A_{n-2}}$ and $d^k V/d\phi^k$.

To obtain (4.21), one can argue inductively in n. At $n = 0$ (4.21) holds with X_{0AB}, P_0, Y_0^A vanishing, $S_0^v = s^v$, $S_0^A = s^A$ and $O_2 = \sum_{n \geq 2} \ell^n F_n = \mathcal{O}(\ell^2)$. Next, we assume that (4.21) is true for any $n' \leq n - 2$ and argue that it must hold for $n' = n$. This amounts to showing that O_n can be brought to a form that has the same structure as the r.h.s.

JHEP08(2022)258

of (4.21). To do this, we first write O_n in terms of the quantities listed at the end of the previous paragraph. By assumption, O_n must be $\mathcal{O}(\ell^n)$ and it must take the form (tensor indices are suppressed in the rest of the discussion)

$$O_n = \sum_{\substack{k,k',p,p',\\ \psi,\psi'\in\{\mu,\phi\}}} A_{n,k,p,k',p',\psi,\psi'} D^k \partial_v^p \psi D^{k'} \partial_v^{p'} \psi' + O_{n+2} \tag{4.24}$$

where $p, p' > 0$ in all terms, i.e. O_n must contain at least two factors with quadratic boost weight and these factors can only involve μ and ϕ. Now one can proceed as in the case of vacuum gravity and perform a sequence of "integrations by parts" to produce terms that are linear either in $\partial_v \mu$ or in $\partial_v \phi$. This of course comes at the cost of boundary terms to be absorbed in Y_n^A, S_n^v or terms that can be absorbed into O_{n+2}. (The latter class of terms may arise when a v-derivative is moved from one of the two positive boost weight factors to the coefficients $A_{n,k,p,k',p',\psi,\psi'}$. The reason for this is that these coefficients can depend on $D^m\beta$, $D^m\partial_r^q\mu$ or $D^m\partial_r^q\phi$. When ∂_v is relocated to act on such terms, the resulting objects need to be expressed with the desired set of primitive factors and some covariant components involving $R_{\mu\nu}$ and $\Box\phi$. However, the covariant components can be shifted to higher order in ℓ by using the gravitational or scalar equations of motion.) Note in particular that to determine the boundary terms to be added to S_n^v, one needs to use an identity of the form (cf. equation (4.9))

$$\partial_v \left[\frac{1}{\sqrt{\mu}} \partial_v \left(\sqrt{\mu} A_{n,k,p,p',\psi,\psi'} \sum_{j=1}^{p'+p-3} a_j D^k \partial_v^j \psi \partial_v^{p+p'-2-j} \psi' \right) \right] = A_{n,k,p,p',\psi,\psi'} D^k \partial_v^p \psi \partial_v^{p'} \psi' + \dots$$

where $\psi, \psi' \in \{\mu, \phi\}$, and the ellipsis stands for terms of the form $A_{n,\bar{k},\bar{p},\bar{k}',\bar{p}',\psi,\psi'} D^{\bar{k}} \partial_v^{\bar{p}} \psi \partial_v^{\bar{p}'} \psi'$ with either (1) $\bar{p}+\bar{p}' < p+p'$, $\bar{p}, \bar{p}' > 0$, (2) $\bar{p}+\bar{p}' = p+p'$, $\bar{p} > 0$, $\bar{p}' = 1$, or (3) $\bar{p}+\bar{p}' = p+p'$, $\bar{p}' > 0$, $\bar{p} = 1$. The coefficients a_j in this identity are to be determined by solving the linear system (4.12) (for any $\psi, \psi' \in \{\mu, \phi\}$). In the end, this procedure lets us write

$$\begin{aligned} O_n = & \sum_{\substack{k,p>0,\\ \psi\in\{\mu,\phi\}}} \left(A_{n,k,p,\psi} D^k \partial_v^p \psi \partial_v \mu + B_{n,k,p,\psi} D^k \partial_v^p \psi \partial_v \phi \right) \\ & + \partial_v \left[\frac{1}{\sqrt{\mu}} \partial_v \left(\sqrt{\mu} \varsigma_n^v \right) \right] + D_A y_n^A + O_{n+2}. \end{aligned} \tag{4.25}$$

The quantities y_n^A, ς_n^v can be absorbed into Y_n^A, S_n^v, and they have the same dependencies on μ, ϕ and β as required by the proposition. The terms in the first line of the r.h.s. of this equation can be dealt with by "completing the square" on K_{AB} and $\partial_v\phi$, thereby producing terms with definite signs and an error term that is higher order in ℓ and to be absorbed into O_{n+2}, concluding the sketch of the proof of proposition 3.

Acknowledgments

We are very grateful to Iain Davies for suggesting a simplification to the inductive argument of section 4. We are grateful to Bob Wald, Aron Wall and Zihan Yan for helpful discussions.

JHEP08(2022)258

SH thanks Trinity College, Cambridge U. for support as a VFC, and DAMTP, Cambridge U. He is grateful to the Max-Planck Society for supporting the collaboration between MPI-MiS and Leipzig U., grant Proj. Bez. M.FE.A.MATN0003. ADK acknowledges support from the European Union's H2020 ERC Consolidator Grant "GRavity from Astrophysical to Microscopic Scales" (Grant No. GRAMS815673). This work was supported by the EU Horizon 2020 Research and Innovation Programme under the Marie Sklodowska-Curie Grant Agreement No. 101007855. HSR is supported by STFC grant no. ST/T000694/1.

A Curvature components in Gaussian null coordinates

Below we provide the expressions for the curvature components in Gaussian null coordinates. The metric is given by

$$g = 2\mathrm{d}v\mathrm{d}r - r^2\alpha\mathrm{d}v^2 - 2r\beta_A\mathrm{d}x^A\mathrm{d}v + \mu_{AB}\mathrm{d}x^A\mathrm{d}x^B\,. \tag{A.1}$$

The nonzero components of the inverse metric are

$$g^{rr} = r^2(\alpha+\beta^2), \qquad g^{rv} = 1, \qquad g^{rA} = r\beta^A, \qquad g^{AB} = \mu^{AB} \tag{A.2}$$

where μ^{AB} is the inverse of μ_{AB}, $\beta^A \equiv \mu^{AB}\beta_B$ and $\beta^2 \equiv \beta_A\beta^A$. We introduce the following variables

$$K_{AB} \equiv \frac{1}{2}\partial_v h_{AB}, \qquad \bar{K}_{AB} \equiv \frac{1}{2}\partial_r h_{AB}\,. \tag{A.3}$$

The nonzero Christoffel symbols are given by

$$\begin{aligned}
\Gamma^r_{rv} &= -\frac{1}{2}\partial_r(r^2\alpha) - \frac{1}{2}r\beta^A\partial_r(r\beta_A)\\
\Gamma^A_{rv} &= -\frac{1}{2}\mu^{AB}\partial_r(r\beta_B)\\
\Gamma^r_{rA} &= -\frac{1}{2}\partial_r(r\beta_A) + r\beta^B\bar{K}_{AB}\\
\Gamma^A_{rB} &= \bar{K}^A{}_B\\
\Gamma^v_{vv} &= \frac{1}{2}\partial_r(r^2\alpha)\\
\Gamma^r_{vv} &= -\frac{1}{2}r^2\partial_v\alpha + \frac{1}{2}r^2(\alpha+\beta^2)\partial_r(r^2\alpha) - r^2\beta^A\left(\partial_v\beta_A + \frac{1}{2}rD_A\alpha\right)\\
\Gamma^A_{vv} &= -\frac{1}{2}r\beta^A\partial_r(r^2\alpha) + r\mu^{AB}\left(-\partial_v\beta_B + \frac{1}{2}rD_B\alpha\right)\\
\Gamma^v_{vA} &= \frac{1}{2}\partial_r(r\beta_A)\\
\Gamma^r_{vA} &= \frac{1}{2}r^2(\alpha+\beta^2)\partial_r(r\beta_A) - \frac{1}{2}r^2D_A\alpha - r^2\beta^BD_{[A}\beta_{B]} + r\beta^BK_{AB}\\
\Gamma^A_{vB} &= \frac{1}{2}r\beta^A\partial_r(r\beta_B) + r\mu^{AC}D_{[C}\beta_{B]} + K^A{}_B\\
\Gamma^v_{AB} &= -\bar{K}_{AB}
\end{aligned}$$

JHEP08(2022)258

$$
\begin{aligned}
\Gamma^r_{AB} &= -r^2\left(\alpha+\beta^2\right)\bar{K}_{AB} - rD_{(A}\beta_{B)} - K_{AB} \\
\Gamma^C_{AB} &= \Gamma[\mu]^C_{AB} - r\beta^C\bar{K}_{AB}\,.
\end{aligned} \tag{A.4}
$$

Then the components of the Riemann tensor are as follows

$$
\begin{aligned}
R_{rvrv} &= \frac{1}{4}\mu^{AB}\partial_r(r\beta_A)\partial_r(r\beta_B) + \frac{1}{2}\partial_r^2(r^2\alpha) \\
R_{rvrA} &= -\frac{1}{2}\bar{K}_A{}^B\partial_r(r\beta_B) + \frac{1}{2}\partial_r^2(r\beta_A) \\
R_{rvvA} &= \frac{1}{2}\partial_v\partial_r(r\beta_A) + \frac{1}{2}K_A{}^B\partial_r(r\beta_B) - \bar{K}_A{}^B r\partial_v\beta_B - \frac{1}{4}r\beta^B\partial_r(r\beta_A)\partial_r(r\beta_B) + \\
&\quad -\frac{1}{2}r\mu^{BC}\partial_r(r\beta_C)D_{[A}\beta_{B]} - \frac{1}{2}D_A\partial_r(r^2\alpha) + \frac{1}{2}r^2\bar{K}_{AB}D^B\alpha + \frac{1}{4}r\beta^B\bar{K}_{AB}\;\partial_r(r^2\alpha) \\
R_{rArB} &= \bar{K}_A{}^C\bar{K}_{BC} - \partial_r\bar{K}_{AB} \\
R_{rAvB} &= -\partial_v\bar{K}_{AB} + K_A{}^C\bar{K}_{BC} + r\beta^C\bar{K}_{B[C}\partial_r(r\beta_{A]}) + \\
&\quad -r\bar{K}_B{}^C D_{[A}\beta_{C]} - \frac{1}{2}D_B\partial_r(r\beta_A) - \frac{1}{4}\partial_r(r\beta_A)\partial_r(r\beta_B) - \frac{1}{2}\bar{K}_{AB}\partial_r(r^2\alpha) \\
R_{vAvB} &= -\partial_v K_{AB} + K_A{}^C K_{BC} - 2r\beta^C\bar{K}_{A[B}\partial_r(r\beta_{C]}) - 2r\beta^C\bar{K}_{B[A}\partial_r(r\beta_{C]}) \\
&\quad +\frac{1}{4}r^2(\alpha+\beta^2)\partial_r(r\beta_A)\partial_r(r\beta_B) - \frac{1}{2}r^2\partial_r(r\beta_{(A})D_{B)}\alpha + r^2\beta^C\partial_r(r\beta_{(A})D_{B)}\beta_C \\
&\quad +rK_{C(A}D^C\beta_{B)} - rK_{C(A}D_{B)}\beta^C - rD_{(B}\partial_v\beta_{A)} + \frac{1}{2}r^2D_BD_A\alpha \\
&\quad +\frac{1}{2}r^3\bar{K}_{AB}\beta^C D_C\alpha - \frac{1}{2}r^2\beta^C\partial_r(r\beta_{(A|})D_C\beta_{|B)} \\
&\quad +\mu^{CD}r^2D_{[C}\beta_{A]}D_{[D}\beta_{B]} + \frac{1}{2}K_{AB}\partial_r(r^2\alpha) + \frac{1}{2}r^2(\alpha+\beta^2)\bar{K}_{AB}\partial_r(r^2\alpha) + \\
&\quad -\frac{1}{2}rD_{(A}\beta_{B)}\partial_r(r^2\alpha) - \frac{1}{2}r^2\bar{K}_{AB}\partial_v\alpha \\
R_{ABvC} &= -2D_{[A}K_{B]C} + \partial_r(r\beta_{[A})K_{B]C} + 2K_{D[A}\bar{K}_{B]C}r\beta^D + \partial_r(r\beta_{[A})\bar{K}_{B]C}r^2(\alpha+\beta^2) \\
&\quad +r^2\bar{K}_{C[A}D_{B]}\alpha + \frac{1}{2}r\partial_r(r\beta_{[A})D_{B]}\beta_C - \frac{1}{2}rD_C\beta_{[A}\partial_r(r\beta_{B]}) \\
&\quad +rD_CD_{[A}\beta_{B]} + r^2\bar{K}_{C[A}\beta^D D_{B]}\beta_D - r^2\beta_D\bar{K}_{C[A}D^D\beta_{B]} \\
R_{ABrC} &= 2D_{[B}\bar{K}_{A]C} - 2r\bar{K}_{D[A}\bar{K}_{B]C}\beta^D - \partial_r(r\beta_{[A})\bar{K}_{B]C} \\
R_{ABCD} &= R[\mu]_{ABCD} + 2K_{B[C}\bar{K}_{D]A} + 2K_{A[D}\bar{K}_{C]B} + 2r^2(\alpha+\beta^2)\bar{K}_{A[D}\bar{K}_{C]B} \\
&\quad +r\bar{K}_{D[B}D_{A]}\beta_C - r\bar{K}_{C[B}D_{A]}\beta_D + r\bar{K}_{B[D}D_{C]}\beta_A - r\bar{K}_{A[D}D_{C]}\beta_B\,.
\end{aligned}
$$

The components of the Ricci tensor are given by the following expressions

$$
\begin{aligned}
R_{vv} &= -\mu^{AB}\partial_v K_{AB} + K_{AB}K^{AB} - r\bar{K}\beta^A\partial_r(r\beta_A) + 2r\bar{K}^{AB}\beta_A\partial_r(r\beta_B) + \\
&\quad -r^3\bar{K}_{AB}\beta^A D^B\alpha + \frac{1}{2}r^3\bar{K}\beta^A D_A\alpha - \frac{1}{2}r^2\partial_r(r\beta_A)D^A\alpha + r^2\beta^A\partial_r(r\beta_B)D_A\beta^B + \\
&\quad -rD^A\partial_v\beta_A + r\beta^A D_A\partial_r(r^2\alpha) - r^2\beta^A\partial_r(r\beta_B)D^B\beta_A + \frac{1}{2}r^2D^AD_A\alpha - \frac{1}{2}r^2\bar{K}\partial_v\alpha +
\end{aligned}
$$

JHEP08(2022)258

$$
\begin{aligned}
&-r^2 D_{[A}\beta_{B]} D^B \beta^A + \frac{1}{2} K \partial_r (r^2 \alpha) + \frac{1}{2} r^2 (\alpha + \beta^2) \bar{K} \partial_r (r^2 \alpha) - r \beta^A \partial_v \partial_r (r \beta_A) + \\
&-r^2 \bar{K}_{AB} \beta^A \beta^B \partial_r (r^2 \alpha) - \frac{1}{2} r D_A \beta^A \partial_r (r^2 \alpha) + \frac{1}{2} r^2 (\alpha + \beta^2) \partial_r^2 (r^2 \alpha) \\
&+\frac{1}{2} r^2 (\alpha + \beta^2) \partial_r (r \beta_A) \partial_r (r \beta_B) \mu^{AB} - \frac{1}{2} r^2 \beta^A \beta^B \partial_r (r \beta_A) \partial_r (r \beta_B) \\
R_{rv} =& -\mu^{AB} \partial_v \bar{K}_{AB} + K^{AB} \bar{K}_{AB} - \frac{1}{2} r \bar{K} \beta^A \partial_r (r \beta_A) + r \bar{K}^{AB} \beta_A \partial_r (r \beta_B) - \frac{1}{2} \partial_r^2 (r^2 \alpha) + \\
&-\frac{1}{2} \mu^{AB} \partial_r (r \beta_A) \partial_r (r \beta_B) - \frac{1}{2} D^A \partial_r (r \beta_A) - \frac{1}{2} \bar{K} \partial_r (r^2 \alpha) - \frac{1}{2} r \beta^A \partial_r^2 (r \beta_A) \\
R_{rr} =& \bar{K}_{AB} \bar{K}^{AB} - \mu^{AB} \partial_r \bar{K}_{AB} \\
R_{rA} =& D_B \bar{K}_A{}^B - D_A \bar{K} - \frac{1}{2} \partial_r^2 (r \beta_A) - \frac{1}{2} \bar{K} \partial_r (r \beta_A) + \bar{K}_A{}^B \partial_r (r \beta_B) + \\
&-r \beta^B \left(2 \bar{K}_A{}^C \bar{K}_{BC} - \bar{K}_{AB} \bar{K} - \partial_r \bar{K}_{AB} \right) \\
R_{vA} =& D_B K_A{}^B - D_A K + \frac{1}{2} \partial_v \partial_r (r \beta_A) + \frac{1}{2} K \partial_r (r \beta_A) - r \bar{K}_A{}^B \partial_v (\beta_B) + \\
&-2 r^2 \partial_r (r \beta_{[A} \bar{K}_{B]C} \beta^B \beta^C + \frac{1}{2} r \beta^B \partial_r (r \beta_A) \partial_r (r \beta_B) + \\
&-r^2 (\alpha + \beta^2) \bar{K}_A{}^C \partial_r (r \beta_C) - r^2 \bar{K} \beta^B D_{[A} \beta_{B]} + r \partial_r (r \beta_{[A}) D_{B]} \beta^B \\
&+2 r^2 \bar{K}_{C[A} \beta^B D^C \beta_{B]} - 2 r^2 \bar{K}_{C[A} \beta^B D_{B]} \beta^C - 2 r K_A{}^C \bar{K}_{BC} \beta^B + r K_{AB} \bar{K} \beta^B + \\
&-\frac{1}{2} D_A \partial_r (r^2 \alpha) + r D^B D_{[A} \beta_{B]} + r^2 \bar{K}_{AB} D^B \alpha - \frac{1}{2} r^2 \bar{K} D_A \alpha \\
&+r \beta^B \left(D_B \partial_r (r \beta_A) - \frac{1}{2} D_A \partial_r (r \beta_B) \right) + r \bar{K}_{AB} \beta^B \partial_r (r^2 \alpha) \\
&+\frac{1}{2} r^2 (\alpha + \beta^2) \left(\partial_r^2 (r \beta_A) + \bar{K} \partial_r (r \beta_A) \right) + r \beta^B \partial_v \bar{K}_{AB} \\
R_{AB} =& R[\mu]_{AB} - 2 \partial_v \bar{K}_{AB} - K \bar{K}_{AB} + 4 K_{(A}{}^C \bar{K}_{B)C} - K_{AB} \bar{K} \\
&+r^2 (\alpha + \beta^2) \left(2 \bar{K}_A{}^C \bar{K}_{BC} - \bar{K}_{AB} \bar{K} \partial_r \bar{K}_{AB} \right) - 4 r^2 \bar{K}_{C(A} \bar{K}_{B)D} \beta^C \beta^D \\
&+2 r \partial_r (r \beta_{(A}) \bar{K}_{B)C} \beta^C - 2 r \bar{K}_{AB} \beta^C \partial_r (r \beta_C) - \bar{K}_{AB} \partial_r (r^2 \alpha) \\
&+2 r \beta^C D_{[A} \bar{K}_{C]B} + 2 r \beta^C D_{[B} \bar{K}_{C]A} + r \bar{K} D_{(B} \beta_{A)} + \\
&-D_{(B} \partial_r (r \beta_{A)}) - \frac{1}{2} \partial_r (r \beta_A) \partial_r (r \beta_B) - r \bar{K}_{AB} D_C \beta^C + 2 r \bar{K}_{C(A} D^C \beta_{B)} \,.
\end{aligned}
$$

A.1 Expressions on $\mathcal{N}$

The nonzero Christoffel symbols on $\mathcal{N}$ are given by

$$
\begin{aligned}
\Gamma^r_{rA} = -\Gamma^v_{vA} = \mu_{AB} \Gamma^B_{rv} &= -\frac{1}{2} \beta_A \\
\mu_{AC} \Gamma^C_{vB} = -\Gamma^r_{AB} &= K_{AB} \\
\mu_{AC} \Gamma^C_{rB} = -\Gamma^v_{AB} &= \bar{K}_{AB} \\
\Gamma^C_{AB} &= \Gamma[\mu]^C_{AB} \,.
\end{aligned} \tag{A.5}
$$

JHEP08(2022)258

The Riemann components on $\mathcal{N}$ simplify as follows

$$
\begin{aligned}
R_{rvrv} &= \frac{1}{4}\beta^2 + \alpha \\
R_{rvrA} &= -\frac{1}{2}\bar{K}_A{}^B\beta_B + \partial_r\beta_A \\
R_{rvvA} &= \frac{1}{2}\partial_v\beta_A + \frac{1}{2}K_A{}^B\beta_B \\
R_{rArB} &= \bar{K}_A{}^C\bar{K}_{BC} - \partial_r\bar{K}_{AB} \\
R_{rAvB} &= -\partial_v\bar{K}_{AB} + K_A{}^C\bar{K}_{BC} - \frac{1}{2}D_B\beta_A - \frac{1}{4}\beta_A\beta_B \\
R_{vAvB} &= -\partial_v K_{AB} + K_A{}^C K_{BC} \\
R_{ABvC} &= -2D_{[A}K_{B]C} + \beta_{[A}K_{B]C} \\
R_{ABrC} &= 2D_{[B}\bar{K}_{A]C} - \beta_{[A}\bar{K}_{B]C} \\
R_{ABCD} &= R[\mu]_{ABCD} + 2K_{B[C}\bar{K}_{D]A} + 2K_{A[D}\bar{K}_{C]B}\,.
\end{aligned}
$$

The components of the Ricci tensor on $\mathcal{N}$ are given by

$$
\begin{aligned}
R_{vv} &= -\mu^{AB}\partial_v K_{AB} + K_{AB}K^{AB} \\
R_{rv} &= -\mu^{AB}\partial_v\bar{K}_{AB} + K^{AB}\bar{K}_{AB} - \alpha - \frac{1}{2}\mu^{AB}\beta_A\beta_B - \frac{1}{2}D^A\beta_A \\
R_{rr} &= \bar{K}_{AB}\bar{K}^{AB} - \mu^{AB}\partial_r\bar{K}_{AB} \\
R_{rA} &= -\partial_r\beta_A + \frac{1}{2}\beta^B\bar{K}_{AB} + \left(D_B + \frac{1}{2}\beta_B\right)\bar{K}_A{}^B - \left(D_A + \frac{1}{2}\beta_A\right)\bar{K} \\
R_{vA} &= \frac{1}{2}\partial_v\beta_A + \frac{1}{2}K_A{}^B\beta_B + \left(D_B - \frac{1}{2}\beta_B\right)K_A{}^B - \left(D_A - \frac{1}{2}\beta_A\right)K \\
R_{AB} &= R[\mu]_{AB} - 2\partial_v\bar{K}_{AB} - K\bar{K}_{AB} + 4K_{(A}{}^C\bar{K}_{B)C} - K_{AB}\bar{K} + \\
&\quad -D_{(B}\beta_{A)} - \frac{1}{2}\beta_A\beta_B\,.
\end{aligned}
$$

Open Access. This article is distributed under the terms of the Creative Commons Attribution License (CC-BY 4.0), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited. SCOAP^3 supports the goals of the International Year of Basic Sciences for Sustainable Development.

References

[1] R.M. Wald, *Black hole entropy is the Noether charge*, *Phys. Rev. D* **48** (1993) R3427 [gr-qc/9307038] [INSPIRE].

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

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

JHEP08(2022)258

[4] A.C. Wall, *A Second Law for Higher Curvature Gravity*, *Int. J. Mod. Phys. D* **24** (2015) 1544014 [arXiv:1504.08040] [INSPIRE].

[5] R.M. Wald, *General Relativity*, University of Chicago Press, Chicago, IL, U.S.A. (1984) [https://doi.org/10.7208/chicago/9780226870373.001.0001].

[6] P.T. Chrusciel, E. Delay, G.J. Galloway and R. Howard, *The Area theorem*, *Ann. Henri Poincaré* **2** (2001) 109 [gr-qc/0001003] [INSPIRE].

[7] S. Bhattacharyya, P. Biswas, A. Dinda and N. Kundu, *The zeroth law of black hole thermodynamics in arbitrary higher derivative theories of gravity*, arXiv:2205.01648 [INSPIRE].

[8] I. Racz and R.M. Wald, *Extension of space-times with Killing horizon*, *Class. Quant. Grav.* **9** (1992) 2643 [INSPIRE].

[9] A. Chatterjee and S. Sarkar, *Physical process first law and increase of horizon entropy for black holes in Einstein-Gauss-Bonnet gravity*, *Phys. Rev. Lett.* **108** (2012) 091301 [arXiv:1111.3021] [INSPIRE].

[10] S. Bhattacharjee, S. Sarkar and A.C. Wall, *Holographic entropy increases in quadratic curvature gravity*, *Phys. Rev. D* **92** (2015) 064006 [arXiv:1504.04706] [INSPIRE].

[11] J. Bhattacharya, S. Bhattacharyya, A. Dinda and N. Kundu, *An entropy current for dynamical black holes in four-derivative theories of gravity*, *JHEP* **06** (2020) 017 [arXiv:1912.11030] [INSPIRE].

[12] S. Bhattacharyya, P. Dhivakar, A. Dinda, N. Kundu, M. Patra and S. Roy, *An entropy current and the second law in higher derivative theories of gravity*, *JHEP* **09** (2021) 169 [arXiv:2105.06455] [INSPIRE].

[13] S. Hollands and R.M. Wald, *Stability of Black Holes and Black Branes*, *Commun. Math. Phys.* **321** (2013) 629 [arXiv:1201.0463] [INSPIRE].

[14] S. Bhattacharyya, P. Jethwani, M. Patra and S. Roy, *Reparametrization Symmetry of Local Entropy Production on a Dynamical Horizon*, arXiv:2204.08447 [INSPIRE].

[15] S. Bhattacharyya, F.M. Haehl, N. Kundu, R. Loganayagam and M. Rangamani, *Towards a second law for Lovelock theories*, *JHEP* **03** (2017) 065 [arXiv:1612.04024] [INSPIRE].

[16] S. Bhattacharjee, A. Bhattacharyya, S. Sarkar and A. Sinha, *Entropy functionals and c-theorems from the second law*, *Phys. Rev. D* **93** (2016) 104045 [arXiv:1508.01658] [INSPIRE].

[17] Á.D. Kovács and H.S. Reall, *Well-Posed Formulation of Scalar-Tensor Effective Field Theory*, *Phys. Rev. Lett.* **124** (2020) 221101 [arXiv:2003.04327] [INSPIRE].

[18] Á.D. Kovács and H.S. Reall, *Well-posed formulation of Lovelock and Horndeski theories*, *Phys. Rev. D* **101** (2020) 124003 [arXiv:2003.08398] [INSPIRE].

[19] S. Weinberg, *Effective Field Theory for Inflation*, *Phys. Rev. D* **77** (2008) 123541 [arXiv:0804.4291] [INSPIRE].

[20] S. Bhattacharyya et al., *Local Fluid Dynamical Entropy from Gravity*, *JHEP* **06** (2008) 055 [arXiv:0803.2526] [INSPIRE].

[21] P. Romatschke, *Relativistic Viscous Fluid Dynamics and Non-Equilibrium Entropy*, *Class. Quant. Grav.* **27** (2010) 025006 [arXiv:0906.4787] [INSPIRE].

JHEP08(2022)258

[22] S. Bhattacharyya, *Entropy current and equilibrium partition function in fluid dynamics*, *JHEP* **08** (2014) 165 [arXiv:1312.0220] [INSPIRE].

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

[24] G. Barnich, F. Brandt and M. Henneaux, *Local BRST cohomology in gauge theories*, *Phys. Rept.* **338** (2000) 439 [hep-th/0002245] [INSPIRE].

[25] R.M. Wald, *On identically closed forms locally constructed from a field*, *J. Math. Phys* **31** (1990) 2378.

[26] U. Muller, C. Schubert and A.M.E. van de Ven, *A Closed formula for the Riemann normal coordinate expansion*, *Gen. Rel. Grav.* **31** (1999) 1759 [gr-qc/9712092] [INSPIRE].

[27] F. Brandt, N. Dragon and M. Kreuzer, *The Gravitational Anomalies*, *Nucl. Phys. B* **340** (1990) 187 [INSPIRE].

JHEP08(2022)258