

# An entropy current for dynamical black holes in four-derivative theories of gravity

Jyotirmoy Bhattacharya,<sup>a</sup> Sayantani Bhattacharyya,<sup>b</sup> Anirban Dinda<sup>b</sup> and Nilay Kundu<sup>c</sup>

<sup>a</sup>*Department of Physics, Indian Institute of Technology Kharagpur, Kharagpur 721302, India*

<sup>b</sup>*School of Physical Sciences, National Institute of Science Education and Research, HBNI, Bhubaneswar, Khurda 752050, Odisha, India*

<sup>c</sup>*Department of Physics, Indian Institute of Technology Kanpur, Kalyanpur, Kanpur 208016, India*

*E-mail:* [jyoti@phy.iitkgp.ac.in](mailto: jyoti@phy.iitkgp.ac.in), [sayanta@niser.ac.in](mailto: sayanta@niser.ac.in), [anirban.dinda@niser.ac.in](mailto: anirban.dinda@niser.ac.in), [nilayhep@iitk.ac.in](mailto: nilayhep@iitk.ac.in)

**ABSTRACT:** We propose an entropy current for dynamical black holes in a theory with arbitrary four derivative corrections to Einstein's gravity, linearized around a stationary black hole. The Einstein-Gauss-Bonnet theory is a special case of the class of theories that we consider. Within our approximation, our construction allows us to write down a completely local version of the second law of black hole thermodynamics, in the presence of the higher derivative corrections considered here. This ultra-local, stronger form of the second law is a generalization of a weaker form, applicable to the total entropy, integrated over a compact 'time-slice' of the horizon, a proof of which has been recently presented in [1]. We also provide a general algorithm to construct the entropy current for the four derivative theories, which may be straightforwardly generalized to arbitrary higher derivative corrections to Einstein's gravity. This algorithm highlights the possible ambiguities in defining the entropy current.

**KEYWORDS:** Black Holes, Black Holes in String Theory

**ARXIV EPRINT:** [1912.11030](https://arxiv.org/abs/1912.11030)

---

**Contents**

<b>1</b>	<b>Introduction and summary</b>	<b>2</b>
<b>2</b>	<b>A comprehensive review of [1]</b>	<b>7</b>
2.1	Strategy of the proof of second law of black hole thermodynamics	8
2.2	An entropy for non-stationary horizons obeying the second law	12
2.2.1	A residual coordinate redefinition freedom	12
2.2.2	Structural form of $E_{vv}^{\text{HD}}$	13
2.2.3	The physical process version of first law and its implications	16
<b>3</b>	<b>An entropy current for four derivative theories of gravity</b>	<b>20</b>
3.1	Explicit calculation of $E_{vv}^{\text{HD}}$ and the entropy current for theories with four derivative corrections to Einstein gravity	21
3.1.1	Ricci scalar square theory	23
3.1.2	Ricci tensor square theory	24
3.1.3	Riemann tensor square theory	25
3.2	The most general structure of the ‘zero boost term’ in $E_{vv}^{\text{HD}}$	25
3.3	Constraints on the ‘zero boost terms’ in $E_{vv}^{\text{HD}}$	33
3.4	The general strategy for constructing the entropy current maintaining the boost symmetry	36
3.5	Einstein-Gauss-Bonnet gravity in $d \geq (4 + 1)$	43
3.6	The Einstein-Gauss-Bonnet theory in $d = 3 + 1$	44
3.7	Comments on entropy current for higher boost terms in $E_{vv}^{\text{HD}}$	46
<b>4</b>	<b>Discussions and future directions</b>	<b>48</b>
<b>A</b>	<b>A general stationary metric can have <math>v</math> dependent components</b>	<b>53</b>
<b>B</b>	<b>Arguments leading to vanishing of <math>T_{vv}</math> on any Killing horizon</b>	<b>54</b>
<b>C</b>	<b>Conventions, notations and useful formulae</b>	<b>56</b>
<b>D</b>	<b>Detailed expressions</b>	<b>58</b>
D.1	Expressions of Riemann tensors and Ricci tensors off the horizon	58
D.2	Relevant terms on the horizon $\mathcal{H}$ , to compute $E_{vv}^{\text{HD}}$ for different theories	59
D.3	Ricci scalar square theory	59
D.4	Ricci tensor squared theory	60
D.5	Riemann tensor squared theory	61

---

## 1 Introduction and summary

It is widely believed that Einstein's theory of gravity must admit an adequate UV completion when we approach length scales comparable to Planck length. Such a putative UV complete theory of quantum gravity, at large length scales, must reduce to the weakly coupled, two derivative Einstein's theory, which has been exhaustively verified in the IR, by several experiments. However, at intermediate length scales, we may encounter a regime, where gravity is still weakly coupled and the quantum corrections are suppressed, but higher derivative corrections to Einstein equations cannot be ignored.<sup>1</sup> In this regime, gravity would be described by an arbitrary diffeomorphism invariant classical theory

$$\mathcal{L} = \mathcal{L}_g(g_{\mu\nu}, R_{\mu\nu\lambda\sigma}, \nabla_\alpha R_{\mu\nu\lambda\sigma}, \nabla_\beta \nabla_\alpha R_{\mu\nu\lambda\sigma} + \dots) + \mathcal{L}_m, \quad (1.1)$$

where, the gravity part of the Lagrangian would admit a derivative expansion, as follows

$$\mathcal{L}_g = R + \left( \alpha_1 R^2 + \alpha_2 R_{\mu\nu} R^{\mu\nu} + \alpha_3 R_{\mu\nu\alpha\beta} R^{\mu\nu\alpha\beta} \right) + \text{higher derivatives} \quad (1.2)$$

Here, in (1.1),  $\mathcal{L}_m$  represents the matter part the theory.<sup>2</sup> Any non-minimal coupling to gravity, i.e., the terms involving curvature and the matter fields are also included in  $\mathcal{L}_m$ . Note that, the specific values of the coefficients  $\alpha_1$ ,  $\alpha_2$  and  $\alpha_3$  depend on the details of the UV complete theory. These are dimension full constants and therefore must be proportional to the (square of) some fundamental length scale of the UV complete theory. For example, in the case of string theory, all of these coefficients will have the form  $\alpha_i \sim l_s^2 \tilde{\alpha}_i$  where  $l_s$  is the string length and  $\tilde{\alpha}_i$ s are some numbers. The limit that we shall consider here, is the one, where the length scale associated with the curvatures of the space-time and those associated with the variations of the matter fields, are much larger compared to  $l_s$ . In other words, all the higher derivative corrections will be more and more suppressed, with the increase in the number of derivatives.

We know that the two derivative Einstein's theory admits black holes solutions with Killing event horizons. These solutions can be understood as macroscopic manifestation of an ensemble of many microscopic degrees of freedom of the more fundamental theory of gravity, in thermodynamic equilibrium, at finite temperature. If this statistical picture of the black hole is correct, then stationary black hole solutions should exist even after we add higher derivative corrections to the gravity action. Also, within these higher derivative

---

<sup>1</sup>Such an intermediate regime exists, for instance, in string theory, which is a prominent candidate for the UV complete theory of quantum gravity. In string theory, within this regime, the string coupling  $g_s \rightarrow 0$ , implying the quantum corrections are suppressed. While, the ratio  $l_s/\mathfrak{R}$  is non-negligible,  $\mathfrak{R}$  being the length scale associated with space-time curvatures, while  $l_s$  is the string length. Whether such an intermediate regime exists in the real world, or whether Einstein's description is a good description right up to the Planck scale, can only be answered with precision experiments of the future.

<sup>2</sup>Note that, there would certainly be some phenomenological restrictions on  $\mathcal{L}_m$ . For example, it should reduce to the standard model Lagrangian at large length scales. Besides this,  $\mathcal{L}_m$  here also incorporates all the matter fields, that may be required to make the higher derivative theory of gravity (1.2) well defined. For instance, It was pointed out in [2], that causality constraints on tree level graviton three point functions imply that  $\mathcal{L}_m$  should incorporate higher spin fields.

theories of gravity, we should be able to construct macroscopic quantities, such as entropy, for the black hole solutions, which will satisfy the laws of thermodynamics.

In two derivative theory of Einstein's gravity, we have a candidate for entropy that satisfies both the first and second laws of thermodynamics. The entropy, in this case, is given by the area of a 'time-slice' of the event horizon [3–5] (also see [6–8]). We shall denote the event horizon with  $\mathcal{H}$  and the time-slice of it with  $\mathcal{H}_v$ . The second law, for this entropy, followed from the famous area increase theorem for black holes [3, 6], which assumes that the matter energy-momentum tensor obeys the null-energy condition. Throughout our discussion in this note, we shall assume this condition to be valid for the matter part of the Lagrangian  $\mathcal{L}_m$ .

This concept of entropy was generalized to stationary black hole solutions in higher derivative theories of gravity in [9, 10], in such a way that it satisfied the first law of thermodynamics.<sup>3</sup> We shall refer to this construction of entropy as Wald entropy  $S_W$ . Now, the first law of thermodynamics relates the infinitesimal shifts in the parameters of two different but nearby equilibrium configurations. Therefore, the Wald entropy, whose construction was solely based on consideration of first law alone, does not unambiguously extend to dynamical situations. Indeed, as it was pointed out in [13–15], there were ambiguities associated with Wald entropy, for non-stationary black hole solutions with dynamical event horizons. All these ambiguities vanished for stationary solutions. We shall refer to these ambiguities as the JKM ambiguities.

Unlike, two derivative Einstein's theory, it is not a priori clear whether Wald entropy satisfies the second law of thermodynamics. In this note, we would like to explore this question further.

In the weakest version, the second law could be stated as follows. Consider two equilibrium configurations (in our case two black hole solutions, not necessarily close by in any sense)  $B_1$  and  $B_2$  such that if one perturbs  $B_1$  in certain ways it is possible to reach  $B_2$  eventually. Then the entropy evaluated on the solution  $B_2$  must be strictly greater than the entropy of  $B_1$ .

Though the above formulation of the second law does not really need a definition of entropy away from stationarity, clearly it refers to dynamics. One way towards a proof would be to show that there exists some extension of Wald entropy to dynamical situations so that the second law is satisfied. It is natural to expect that this extension (if at all possible) might fix those ambiguities related to the definition of entropy, which only arises in non-stationary situations (i.e. the JKM ambiguities).

Now it is very difficult to analyze dynamical black hole solutions even in two derivative Einstein's theory of gravity. People usually take recourse to several perturbation schemes, around the stationary solutions that are known exactly. In this context, the simplest situation that comes to mind is the case where some stationary black hole is slightly perturbed by some external agent so that the resulting dynamical black hole metric could be decomposed as a sum of stationary part and the time-dependent part with small amplitude. Then one can analyze the equations in an expansion in terms of the amplitude. Note that

---

<sup>3</sup>See [11, 12] for the latest reviews of black hole thermodynamics in higher derivative theories of gravity.

under this approximation, it is not possible to study violent processes such as the formation of black holes or the merger of two black holes. Throughout this note, we shall work only up to linear order in the amplitude of fluctuations, and therefore, our results would not directly apply to these violent scenarios. See [16–18], where similar questions relating to entropy and the second law, for processes involving the merger of black holes have been addressed.

In [1] the author has used this expansion to construct one ‘out of equilibrium’ extension of Wald entropy, which, up to linear order in amplitude expansion, is monotonically increasing at every instant of ‘time’ and therefore satisfies the second law in a stronger sense (also see [19–22]). This locality in time is not entirely unexpected in this type of set-up where the space-time is ‘near’ some equilibrium or stationary solution at every instant of time. Following the same intuition, we could also say that for such slow enough, ‘near-equilibrium’ time evolution, where we could assume that different sub-regions of a large macroscopic system are in approximate equilibrium with its immediate neighbourhood, at every instant of time, we should also expect a spatial locality, in the formulation of second law. This expectation is completely consistent with the scenario in Einstein’s theory of gravity, where the area increase theorem is valid locally, for every infinitesimal area element of a ‘time-slice’ of the horizon. Our expectation for a local, stronger form of second law, is very much motivated by this example of Einstein’s theory. Now, in a more general setting involving higher derivative corrections to Einstein’s equation, during slow time evolutions, besides entropy production in every infinitesimal sub-region, we have to be also open to the possibility that entropy could be redistributed between the neighbouring regions, by flowing in or out via some spatial current. The necessity of having such non-zero spatial current for entropy, and the existence of a strong ultra-local form of the second law of thermodynamics, in higher derivative theories of gravity as well, are the key points of our investigation here.

In this note, we shall demonstrate that it is possible to formulate the second law in its strongest form, so that at least for ‘slow enough’ dynamical situations, entropy is produced at every point of the evolving space-time, up to a possible inflow and outflow via some spatial current.<sup>4</sup> We will explicitly construct this spatial current for entropy flow, in the most general four derivative theory of gravity.

Let us now outline the organization of this note, along with a brief summary of the key arguments and results in the various sections.

At first, in section 2, we shall review the paper [1] in detail. The author in [1] has shown that at the leading order in amplitude expansion, certain ‘time-time’ component of the equation of motion of any higher derivative theories of gravity could always be written as two ‘time derivatives’ acting on some quantity. Then he could further argue that if one identifies the integral of this quantity (the expression on which the two time derivatives are

---

<sup>4</sup>An entropy current with non-negative divergence certainly exists in near-equilibrium states for theories that do not include dynamical gravity [23–26]. In the case of gravity, where the space-time itself becomes the fundamental dynamical object, the concept of locality might become a bit confusing. The locality in space-time in some sense becomes analogous to some form of locality in the space of fundamental fields of non-gravitational theories. However, the kind of perturbation that we are considering here, there is always a stationary base metric which could play the role of the background and the above-mentioned issues could be avoided.

acting) over  $\mathcal{H}_v$ , with the entropy of the gravitational theory, then it will satisfy the second law at least at the leading order within this approximation. Following [1] we have first set up an appropriate coordinate system, thus defining the ‘time’ mentioned above. Next, by using a symmetry of the horizon geometry (referred to as ‘boost symmetry’ in [1]), we have classified the terms that can appear in that particular ‘time-time’ component of the equations of motion, according to their weight under this boost transformation. We shall see that the argument and construction in [1] smoothly goes through for all the higher weight terms except the one that appears at zero boost weight.

This point was noted in [1] and it has been argued that if these ‘zero boost terms’ are not of the correct form (i.e., two time derivatives acting on some quantity) it would amount to the violation of the first law itself, once viewed in the ‘physical version’ formulation of it [15, 27] (also see [28–32]). Therefore, though the central argument in [1] naively break down for these special ‘zero boost terms’, it must work out in actual theories, where the physical process version of the first law is valid.

Note that the formulation of the ‘physical process version’ of the first law uses exactly the same setup as the one used in [1]. Here also one perturbs the stationary black hole out of equilibrium and lets it settle to another nearby stationary solution with slightly shifted parameters. The first law is a relation between these shifts of parameters, which characterize the two equilibrium solutions.<sup>5</sup> In the arguments leading to this physical version of the first law, the external agent which drives the system out of equilibrium is a very specific one — some small matter (associated with a small shift in matter stress tensor) entering the system through asymptotic infinity. The similarity between the two set-up of the second law and the physical process version of the first law is very suggestive of the fact that the structural nature of the terms in entropy, which play a major role in the proof of the physical process version of first law, would also be extremely important in the proof of the second law. We shall refer to such terms as ‘zero boost terms’, the justification of such terminology would be explained later in the main text.

After this extensive review of [1], we shall closely study how the physical version of the first law constrains these ‘zero boost terms’. We shall find that locally the required ‘time-time’ component of the equation of motion (let us denote this ‘time as  $v$  and the relevant component of the equation of motion as  $E_{vv}$ ) need not have the form specified form which is naively implied by the physical process version of the first law. This naive expectation would be that the zero boost terms in  $E_{vv}$ , has two ‘time’ derivatives acting on some local quantity (let us denote it as  $J^v$ ) defined on the horizon. This naive expectation is not accurate, since any term that could be expressed as a single ‘time’ derivative acting on the spatial divergence of some space current (denoted here as  $J^i$ ) may also be present in  $E_{vv}$ , without affecting the first law. This is because the physical process version of the first law deals with a total change in entropy ( along with the charge and mass) as the black hole evolves from one equilibrium to another. The total entropy always comes with

---

<sup>5</sup>We would like to emphasize that though the proofs of both first law and second law use the same set-up, they are very different in terms of details. In particular, Wall’s construction could fix many more terms in entropy (usually denoted as JKM ambiguities in literature) that do not contribute to the first law at all.

integration overall spatial section of the horizon  $\mathcal{H}_v$ . In that case, any such total divergence term would just integrate to zero.

It should be noted that such a term will not affect the argument of [1], which proves a weaker version of the second law, in which the total entropy (integrated over  $\mathcal{H}_v$ ) has been considered. This weaker form is a local statement in time (i.e, total entropy increases at every instant of time) but not in space. At every stage of the arguments in [1], the integration over the spatial sections of the horizon played an important role.

Once we have realized that, it is possible to introduce the notion of a spatial entropy current, without affecting the proof of both the first law and the second law (even a strong ultra-local form of it), the next immediate question is whether such a spatial entropy current is necessary. In other words, we should investigate that, if we were to write down an ultra-local version of the second law, largely following the procedure of [1], can we do it without introducing the entropy current, in any higher derivative theory of gravity. In more practical terms, we need to check whether the relevant ‘zero boost terms’ in the equation of motion  $E_{vv}$  for a given higher derivative theory of gravity, does indeed have terms which give rise to the spatial entropy current  $J^i$ . To answer this question, we specialize to four-derivative theories of gravity. In section 3.1, we explicitly compute the relevant component of the equations of motion and we see that there has to be a spatial entropy current in some of these four derivative theories, if we want a completely local version of the second law to be true. This is the central result of our note. Besides achieving manifest locality, our procedure of constructing the entropy current might play a crucial role in providing an alternative proof of the second law, without invoking the ‘physical process’ version of the first law, the use of which has been a necessary input for the proof presented in [1]. We would like to investigate this possibility further in future work.

In this context, the four-dimensional Gauss-Bonnet theory requires a special mention. We have discussed this case in details in section 3.6. It is well known that in four dimensions Gauss-Bonnet action is a total derivative and therefore does not contribute to the equation of motion. However, from Wald’s analysis, we know that the entropy of the black holes in Gauss-Bonnet theory does receive correction which is proportional to the intrinsic Ricci-scalar evaluated on the two-dimensional spatial section of the horizon. Integration of this quantity over a compact two dimensional manifold results in a topological quantity, the Euler characteristics, that does not change under small continuous deformation of the horizon caused due to the perturbation. This is perfectly consistent with the fact that 4-d Gauss-Bonnet term does not introduce any correction to the equation of motion and so (following the argument of [1]) no correction to the change in total entropy during time evolution. However, if we are thinking of in terms of the entropy density (i.e., the same intrinsic Ricci scalar without the integration over all spatial sections of the horizons), it does evolve with time. But, this entropy density, clearly, would not satisfy the ultra-local version of the second law. However, the validity of the local version of the second law is restored, if we also consider a spatial entropy current. It is satisfying to check that, for the 3 + 1 dimension the  $v$ -derivative of the Ricci scalar is identical to the time derivative of the divergence of a spatial current (given in terms of the extrinsic curvatures of  $\mathcal{H}_v$ ). These two contributions from the entropy density and the spatial entropy current, to the

equation of motion, cancel out each other. This cancellation is off-shell and specific to  $3 + 1$  dimensions only. Thus, in this simple example, it is easy to recognize the necessity of the entropy current, even before performing the detailed calculation. As we will see in more detail in section 3.6, this example also helps us to identify an ambiguity present in the definition of the entropy current.

In section 3.2, section 3.3 and section 3.4, we go on to develop a general algorithm for constructing the spatial components of the entropy current for arbitrary four derivative theory. From this exercise, we learn that the most general form of the relevant equation of motion  $E_{vv}$ , which is consistent with the boost symmetry, has a structure that is more general compared to what would be essential for defining the entropy current. In other words, the fact we have an entropy current and consequently a local second law, puts very non-trivial constraints on the most general possible structure of  $E_{vv}$ .<sup>6</sup> Although at the moment we do not have a precise explanation regarding the physical origin of these constraints, we believe that these may arise due to some residual gauge freedom. We think understanding the exact mathematical reason behind these constraints, would lead us to a proof of the local second law through the construction of the entropy current, without invoking the first law at all.

This general algorithm has also helped us to understand the ambiguities related to the construction of the current more clearly. In section 3.7, we report on one of the primary sources of such ambiguities. Finally in section 4, we conclude and discuss possible future directions.

Before concluding this section, we would like to mention that, the notion of an entropy current for black hole dynamics in higher derivative theories of gravity, is not completely new in this note. This idea has been previously introduced in [33–35]. In [33, 34] it was primarily motivated by the entropy current, constructed in the context of the fluid gravity correspondence [23]. While in [35], the entropy current was constructed exploiting the membrane-gravity duality, using an expansion in inverse powers of space-time dimension. Although, the exact context of these constructions are different from our considerations here, but there are some similarities in the basic idea (see section 4 for further discussions on this). The exact relation between our construction and that reported in these papers is a topic of our current investigation and we hope to report on it in the near future.

## 2 A comprehensive review of [1]

As we have discussed in section 1 in an arbitrary diffeomorphism invariant theory of gravity, a proof of the second law for dynamical black holes was provided in [1]. Let us review the details of the proof here, which would serve as a useful prelude to the subsequent discussion of our entropy current.

Let us first choose a coordinate system. Let  $\partial_v$  be the null generators of the event horizon, where  $v$  is the affine parameter. Let  $\partial_i$ s denote the rest of the spatial tangents of the horizon. Integral curves of  $\partial_i$ s are the spatial coordinates along the constant  $v$  slices

---

<sup>6</sup>As we have mentioned above, an explicit calculation for the four derivative theories demonstrates that, these constraints are automatically met for these theories.



of the horizon. Then from every point on the horizon, we shoot off a set of null geodesics, making a precise angle with the coordinates on the horizon. We label each of these new sets of null geodesics (null everywhere) by the coordinates of the point at which it intersects the horizon. We denote the affine parameter along the null geodesics to be  $r$  which is the coordinate, away from the horizon. The most general metric with this choice of coordinates would have the following structure (see appendix of [36])

$$ds^2 = 2dv dr - r^2 X(r, v, x^i) dv^2 + 2r \omega_i(r, v, x^i) dv dx^i + h_{ij}(r, v, x^i) dx^i dx^j \quad (2.1)$$

Here we have chosen the horizon to be at  $r = 0$  (a choice for the origin of the affine parameter along each null geodesic  $\partial_r$ ). Note that this choice of gauge is slightly different from that of [1]. We have set  $g_{rv} = 1$  throughout space-time, but in [1] this condition was set only on the future horizon  $\mathcal{H}$ . In appendix A of [36], it was demonstrated that this choice of metric (2.1), is possible without any loss of any generality, even for dynamical black holes. We would also like to emphasize that this difference in gauge choice, do not affect the arguments in [1], in any way. We shall work with this slight difference in this note, simply because we prefer to work with a metric where the gauge fixing is more complete.

Given the form of the metric (2.1), let us now outline the broad strategy of the proof of second law provided in [1].

### 2.1 Strategy of the proof of second law of black hole thermodynamics

The general strategy of the proof follows that of area increase theorem for dynamical black holes in Einstein gravity [3, 6]. Following [1] we shall consider small time-dependent fluctuations about stationary black holes. Let us denote the amplitude of the fluctuation to be  $\epsilon$ . All the analysis would be linear in the amplitude (denoted by  $\epsilon$ ) of this fluctuation.

Let us denote a  $v$ -slice of the horizon of the dynamical black hole to be  $\mathcal{H}_v$ . In [1]  $\mathcal{H}_v$  has been considered to be compact, an assumption which played an extremely important role in the proof. This assumption ensured certain boundary terms to vanish. Therefore, even if the horizon was non-compact, but those boundary terms continued to vanish, the proof of [1] is completely valid. However, in the present paper, our local statement of the second law should not be sensitive to the compact nature of the horizon or depend on the vanishing of such boundary terms. Then, under the approximations considered here, let us schematically write down the entropy of the black hole, in an out of equilibrium scenario, to be

$$S = \int_{\mathcal{H}_v} \sqrt{h}(1 + s_n) \quad (2.2)$$

where

$$s_n = s_w^{\text{HD}} + s_c.$$

Here  $s_w^{\text{HD}}$  are the corrections to the area law, coming from the Wald entropy formula due to the presence of the higher derivative corrections to the Einstein-Hilbert action. Note that, in this notation the Wald entropy is given by<sup>7</sup>

$$S_W = \int_{\mathcal{H}_v} \sqrt{h}(1 + s_w^{\text{HD}}) = \int_{\mathcal{H}_v} \frac{\partial \mathcal{L}}{\partial R_{\mu\nu\rho\sigma}} \epsilon_{\mu\nu} \epsilon_{\rho\sigma}, \quad (2.3)$$

---

<sup>7</sup>Here, we have treated the area term corresponding to Einstein theory separately, to facilitate comprehension for our readers who are familiar with the area increase theorem.

where  $\epsilon_{\mu\nu}$  are the bi-normal to  $\mathcal{H}_v$ , the co-dimension-2 spatial slicing of the horizon. Here,  $s_c$  are further corrections to the Wald entropy, which are a part of the JKM ambiguity. One of the central idea of the proof, is to choose an appropriate  $s_c$  so as to ensure  $\partial_v S \geq 0$ . This in turn, therefore, fixes the ambiguity.

Now let us act (2.2) with a  $v$ -derivative, which can be moved inside the integral in the r.h.s., since the integral is over a  $v$ -slice of the horizon. We have

$$\partial_v S = \int_{\mathcal{H}_v} \partial_v \left( \sqrt{h}(1 + s_n) \right) \equiv \int_{\mathcal{H}_v} \sqrt{h} \vartheta, \tag{2.4}$$

where  $h_{ij}$  is the induced metric on  $\mathcal{H}_v$ , and

$$\vartheta = \vartheta_E + \frac{1}{\sqrt{h}} \partial_v \left( \sqrt{h} s_n \right),$$

with  $\vartheta_E$  being the contribution coming from the area form, which is present even in pure Einstein gravity without any higher derivative corrections. We can show that  $\vartheta_E = \frac{1}{2} h^{ij} \partial_v h_{ij}$  is the expansion of the congruence of the null generators of the horizon [8].

Following the proof of the black hole area increase theorem [3, 6] for Einstein gravity, the general strategy for proving  $\partial_v S \geq 0$ , is to demonstrate that  $\partial_v \vartheta \leq 0$ . This, together with the additional physical expectation<sup>8</sup>  $\vartheta|_{v \rightarrow +\infty} \rightarrow 0$ , implies that  $\vartheta \geq 0$ , for all  $v \geq 0$ .

Now at linear order in amplitude both inequalities (i.e.,  $\partial_v S \geq 0$  and  $\partial_v \vartheta \leq 0$ ), must be some equality relation since terms linear in amplitude ( $\epsilon$ ) could have any sign depending on the sign of  $\epsilon$ . In other words, the only way the inequalities could be satisfied is to set them to zero at linear order in  $\epsilon$ .

$$\partial_v \vartheta = \mathcal{O}(\epsilon^2) \tag{2.5}$$

We shall try to choose  $s_n$  such that  $\partial_v \vartheta = \mathcal{O}(\epsilon^2)$  is ensured. More precisely, by equation (2.5), what we mean is the following.

- We shall consider only those dynamics where every metric component  $G_{AB}$ , that are not already fixed by our gauge choice, could be decomposed as

$$G_{AB} = G_{AB}^{(0)} + \epsilon \delta G_{AB}$$

where  $G_{AB}^{(0)}$  is the non-dynamical part of the metric and has a time-like Killing vector with a Killing horizon.  $\delta G_{AB}$  is time-dependent.  $\epsilon$  is the small parameter encoding the amplitude of dynamics, which could be of either sign but always small. All terms quadratic or higher-order in  $\epsilon$  would be neglected.

- We further demand that  $G_{AB}^{(0)}$  is an exact solution of the Einstein equation with appropriate higher derivative corrections and also relevant matter stress tensor. Additionally,  $G_{AB}$  also solves the same equations but up to corrections of order  $\mathcal{O}(\epsilon^2)$ .

---

<sup>8</sup>Here, the physical expectation is that the dynamical black hole will settle to a stationary metric with a Killing horizon at  $v \rightarrow \infty$ , leading to the vanishing of  $\partial_v S$  at  $v \rightarrow \infty$ .

- Now our goal is to construct an  $s_n$  out of  $G_{AB}^{(0)}$  and  $\delta G_{AB}$  such that if we just blindly evaluate  $\partial_v \vartheta$  and impose equations of motion it turns out to be order  $\mathcal{O}(\epsilon^2)$  (or just vanishes within our approximation).

One of the key points of [1] is to provide an algorithm to construct such an  $s_n$  in all possible higher derivative theories of gravity.

At this stage we would like to emphasize that equation (2.5) is a necessary condition for second law, but certainly not sufficient, even within this perturbative treatment. Sufficiency would demand a particular sign for the coefficient of the  $\mathcal{O}(\epsilon^2)$  term and in those special space-time points where this coefficient also vanishes, one has to keep track of even the higher-order terms. However, as it is the case in [1], in this note we shall confine ourselves to computations only up to order  $\mathcal{O}(\epsilon)$ . They themselves turn out to be constraining enough to fix a large part of ambiguities that are there in the form of gravitational entropy for higher derivative theories.

Now let us process equation (2.5) little further, which will finally tell us how, manipulating a particular component of equations of motion, we could construct some  $s_n$  that satisfies equation (2.5).

$$\begin{aligned}
 \partial_v \vartheta &= \partial_v \vartheta_E + \partial_v \left( \frac{1}{\sqrt{h}} \partial_v (\sqrt{h} s_n) \right) \\
 &= -R_{vv} + \partial_v \left( \frac{1}{\sqrt{h}} \partial_v (\sqrt{h} s_n) \right) + \mathcal{O}(\epsilon^2) \\
 &= -T_{vv} + E_{vv}^{\text{HD}} + \partial_v \left( \frac{1}{\sqrt{h}} \partial_v (\sqrt{h} s_n) \right) + \mathcal{O}(\epsilon^2)
 \end{aligned}
 \tag{2.6}$$

Here  $T_{vv}$  denotes the  $vv$  component of the matter stress tensor and  $E_{vv}^{\text{HD}}$  is the  $vv$  component of the higher derivative corrections to the gravity part of the equations of motion. In the second line we have used the fact

$$\partial_v \vartheta_E = -R_{vv} + \mathcal{O}(\epsilon^2)$$

This is essentially the Raychaudhuri equation for the congruence of null geodesics and this is an off-shell equation — it does not require the metric to satisfy any particular equation of motion. We have used the equation of motion while going from second to the last line of equation (2.6)

$$R_{vv} + E_{vv}^{\text{HD}} = T_{vv}
 \tag{2.7}$$

In all of our analysis this the only place where we shall use the on-shell condition on the metric components.

Now let us analyze the  $\epsilon$  dependence of  $T_{vv}$ . We would like to argue that  $T_{vv}$  is also of order  $\mathcal{O}(\epsilon^2)$  and therefore does not contribute within our approximation.

In the case of higher derivative theory, the definition of matter stress tensor might become a bit confusing. Our convention is the following. If we vary the action with respect to the metric fluctuation, the resultant two-indexed tensors could be categorized in two different classes; terms that depend only on the metric components and terms that along

with the metric components, also depend on the matter fields. All the higher derivative terms that are of the first category, are together called as  $E_{AB}^{\text{HD}}$  and the matter stress tensor  $T_{AB}$  consists of all the terms in the second category.

Clearly, if we want to know the  $\epsilon$  dependence of  $T_{vv}$ , we need to fix the  $\epsilon$  dependence of the matter fields. Let  $\Phi$  denotes all the matter fields (collectively) and let's assume that it also admits the following expansion.

$$\Phi = \Phi^{(0)} + \epsilon \delta\Phi$$

Here  $\delta\Phi$  encodes the dynamics and  $\Phi^{(0)}$  is the value of  $\Phi$  on the stationary situation i.e., when all field configurations, including both metric and the matter fields, admit a Killing vector. As in the case of metric, we want  $\Phi^{(0)}$  to satisfy the equations of motion (on the background of stationary metric  $G_{AB}^{(0)}$ ) for the matter field exactly and  $\delta\Phi$  up to linear order in  $\epsilon$ .

We shall consider only those matter stress tensors that satisfy the null energy condition. In our context, it implies that as long as the matter fields satisfy their equations of motion (in any smooth background geometry that need not be dynamical), the  $vv$  component of the stress tensor is always non-negative.

$$T_{vv} \geq 0$$

We would like to stress again that the validity of the above condition requires only the matter fields to be on-shell, but the metric need not be. Now one can argue that in a stationary situation (i.e., in the limit of  $\epsilon \rightarrow 0$ )  $T_{vv}$  simply vanishes (see appendix B for the details). Any quantity that satisfies some positivity condition and also vanishes in a stationary situation, must be quadratic in the amplitude of dynamics since the linear term could have either sign. It follows (exactly for the same reason as in equation (2.5)) that  $T_{vv}$  is also of order  $\mathcal{O}(\epsilon^2)$  at every order in derivatives.

Now equation (2.6) could be simply satisfied for some choice of  $s_n$ , provided  $E_{vv}^{\text{HD}}$  has the following off-shell form<sup>9</sup>

$$E_{vv}^{\text{HD}} \Big|_{\text{offshell}} = \partial_v \left( \frac{1}{\sqrt{h}} \partial_v \left( \sqrt{h} \varsigma \right) \right) + \mathcal{O}(\epsilon^2). \tag{2.8}$$

If equation (2.8) is true, then one could just choose  $\left[ \int_{\mathcal{H}_v} \sqrt{h} s_n \right]$  to be minus of  $\left[ - \int_{\mathcal{H}_v} \sqrt{h} \varsigma \right]$  up to correction of order  $\mathcal{O}(\epsilon^2)$ .

Let us pause here for a moment, to make an important observation about a very special situation. Imagine a situation in which, a fluctuation in the matter field with a very small amplitude sources the metric, through its energy-momentum tensor. The first correction to

---

<sup>9</sup>Note it is very important that the form predicted in (2.8) is an off-shell requirement on  $E_{vv}^{\text{HD}}$ . Since we have already argued that  $T_{vv}$  is of order  $\mathcal{O}(\epsilon^2)$ , equations of motion for the metric (2.7) ensures that on-shell  $E_{vv}^{\text{HD}}$  must be of order  $\mathcal{O}(\epsilon^2)$ . In other words, just like the solutions of any other differential equations, on-shell we do not have the freedom of determining the  $\epsilon$  dependence of terms involving large number of derivatives, once the lower derivatives are fixed. However, our final goal is to construct an expression for  $s_n$  and we can actually achieve this goal by treating  $E_{vv}^{\text{HD}}$  off-shell, where the naive  $\epsilon$  counting works.

the zeroth-order stationary metric is entirely due to back-reaction from this source. In such a situation, the first change in the matter fields is of order  $\epsilon$ , and the energy-momentum tensor is of order  $\epsilon^2$ ; but, the first correction to the metric would be of order  $\epsilon^2$ . The  $\mathcal{O}(\epsilon)$  piece for the metric would be trivially zero, in this special case, when the boundary conditions are chosen suitably. So everything we have said so far, for the  $\mathcal{O}(\epsilon)$  coefficient in the metric remains true, but trivially true.

Now, in this special case, since the first correction to the metric occurs at  $\mathcal{O}(\epsilon^2)$ , all our conclusions in this note, regarding the linearized corrections to metric, would then be applicable to the  $\mathcal{O}(\epsilon^2)$  terms. The only crucial difference would be that, instead of the equality (2.5), we would now get an inequality for the coefficient of  $\epsilon^2$  in  $\partial_v \vartheta$ , i.e.  $\partial_v \vartheta|_{\epsilon^2} \leq 0$ . This happens because of (2.6), where, due to the null-energy condition,  $T_{vv}$  now contributes positively, at the same order at which the metric receives its first corrections. The cancellation between  $E_{vv}^{\text{HD}}$  and  $\partial_v \left( \frac{1}{\sqrt{h}} \partial_v \left( \sqrt{h} s_n \right) \right)$  is realized in exactly the same way as discussed earlier. This, in turn, implies that  $\partial_v S \geq 0$ , the inequality being important even for the linear (first non-trivial) corrections to the metric. Note that, this situation is perhaps physically important, since this is one of the simplest situations where we can realize a dynamical event horizon, by throwing a tiny amount of matter towards the black hole.

In [1], author has explicitly shown that equation (2.8) is true. As we have mentioned in the introduction, he has used the transformation property of  $E_{vv}^{\text{HD}}$  under certain boost symmetry for his proof. His key argument works barring few ‘leading terms’ in  $E_{vv}^{\text{HD}}$ , for which the author has used the ‘physical process formulation’ of the first law as an extra input. As we shall see below, for particularly these terms the integration over the constant  $v$  slices of the horizon turns out to be very crucial.

## 2.2 An entropy for non-stationary horizons obeying the second law

In this subsection, we shall review the arguments in [1], which establishes that most of the terms in  $E_{vv}^{\text{HD}}$  could be recast in the form (2.8).

### 2.2.1 A residual coordinate redefinition freedom

Let us recall that in (2.1), the coordinate  $v$  constitutes an affine parameter along the null generators of the horizon, while the coordinate  $x^i$  labels the individual generators. But this definition does not completely fix the coordinates on the horizon  $\mathcal{H}$ . We still have the freedom of the following two classes of coordinate redefinition.

1. We can perform an affine re-parametrization of the generators of horizon  $\mathcal{H}$ , through the transformation

$$v \rightarrow \tilde{v} = v p_1(x^j) + p_2(x^j). \tag{2.9}$$

Note that, an arbitrary re-parametrization of this form (2.9), may not be compatible with the gauge choice as used in (2.1). Therefore, along with the transformation (2.9), it may be required to transform the  $r$  coordinate as well, so that the gauge choice of (2.1) is retained, even after the coordinate transformation (2.9). This point is explicated further below, in a special case of (2.9).

2. We can also relabel the generators as follows

$$x^i \rightarrow \tilde{x}^i = f^i(x^j). \tag{2.10}$$

The transformation (2.10) does not change the constant  $v$  slices of the horizon; consequently  $\mathcal{H}_v$  is invariant under it. Also, for the choice of the metric (2.1), covariance under (2.10) may be implemented by ensuring that all the spatial  $i$ -indices are covariant; especially, the covariant derivatives along  $x^i$ , should be compatible with the metric  $h_{ij}$ . For this reason, ensuring invariance (covariance) of entropy (or the second law) under (2.10) is relatively easy. However, covariance under (2.9) is extremely non-trivial and leads to constraints, that were exploited in [1] to fix the form of  $E_{vv}^{\text{HD}}$  and hence the correction to the entropy.

In [1] only a special case of (2.9) was considered under which  $p_1 = a$ ,  $p_2 = 0$ , so that  $v$  is re-scaled as  $v \rightarrow \tilde{v} = av$ ,  $a$  being a constant. Now, in order to ensure that our coordinate redefinition is compatible with the gauge choice of (2.1), we must rescale the  $r$  coordinate suitably. For instance, in order to ensure that  $g_{rv} = 1$  everywhere, even after rescaling  $v$ , we must simultaneously rescale<sup>10</sup>

$$v \rightarrow \tilde{v} = av, \quad r \rightarrow \tilde{r} = \frac{1}{a}r. \tag{2.11}$$

In the new coordinates (2.11) the metric takes the following form

$$\begin{aligned} ds^2 = & 2 d\tilde{v} d\tilde{r} - \tilde{r}^2 X \left( \lambda\tilde{r}, \frac{\tilde{v}}{\lambda}, x^i \right) d\tilde{v}^2 \\ & + 2\tilde{r} \omega_i \left( \lambda\tilde{r}, \frac{\tilde{v}}{\lambda}, x^i \right) d\tilde{v} dx^i + h_{ij} \left( \lambda\tilde{r}, \frac{\tilde{v}}{\lambda}, x^i \right) dx^i dx^j \end{aligned} \tag{2.12}$$

Note that, for the parametrization (2.1), the metric looks almost invariant under this coordinate transformation (2.11), however, the arguments of the metric functions are appropriately scaled. In particular, on the horizon  $\mathcal{H}$ , the induced metric in the new coordinates takes the following form

$$ds_{\mathcal{H}}^2 = 2 d\tilde{v} d\tilde{r} + h_{ij} dx^i dx^j, \tag{2.13}$$

which has an identical structure as compared to that in the old coordinates.

### 2.2.2 Structural form of $E_{vv}^{\text{HD}}$

At first, let us enlist the various derivatives and functions that may occur in  $E_{vv}^{\text{HD}}$ , for any general diffeomorphism invariant theory of gravity (1.1). These building blocks for constructing  $E_{vv}^{\text{HD}}$  include

1. The metric functions  $X$ ,  $\omega_i$  and  $h_{ij}$ .
2. The covariant derivative  $\nabla_i$  with respect to  $x^i$ , compatible with the metric  $h_{ij}$ , which can act on the above metric functions.
3. The partial derivatives  $\partial_r$  and  $\partial_v$ , which may also act on the metric functions.

---

<sup>10</sup>In [1], this rescaling has been referred to as ‘boosts’, while the quantities invariant under this rescaling has been referred to as ‘boost invariant’.

Let us immediately note that among these building blocks, it is only  $\partial_v$  and  $\partial_r$  that transform non-trivially under the coordinate rescaling (2.11). These transform as

$$\partial_r \rightarrow \partial_{\tilde{r}} = \lambda \partial_r, \quad \partial_v \rightarrow \partial_{\tilde{v}} = \frac{1}{\lambda} \partial_v \quad (2.14)$$

As is apparent from (2.12), the rest of the building blocks of  $E_{vv}^{\text{HD}}$ , which include the metric functions and the covariant derivative  $\nabla_i$ , remain invariant under (2.11).

Let us now note that under the coordinate rescaling (2.11),  $E_{vv}^{\text{HD}}$ , being the  $vv$ -component of a covariant tensor, must transform as

$$E_{vv}^{\text{HD}} \rightarrow E_{\tilde{v}\tilde{v}}^{\text{HD}} = \frac{1}{\lambda^2} E_{vv}^{\text{HD}}. \quad (2.15)$$

Let us define the weight of a quantity to be the power of  $\lambda$  by which the quantity rescales under the transformation (2.11). In this sense, the weight of  $E_{vv}^{\text{HD}}$  is  $-2$ .

Now, from the transformation property of the building blocks, it is clear that only  $\partial_v$  has a negative weight under (2.11). Hence, it follows that, every term in  $E_{vv}^{\text{HD}}$  must have at least two  $\partial_v$ . At linear order in  $\epsilon$ , the most general schematic structure of any term in  $E_{vv}^{\text{HD}}$  would be

$$E^{(m,n,k)} = \partial_r^k [(\partial_v \partial_r)^m P] \partial_v^{k+2} [(\partial_v \partial_r)^n Q] + \mathcal{O}(\epsilon^2), \quad (2.16)$$

where  $m, n$  and  $k$  are positive integers including zero. Here we have kept all the  $\partial_r$  and  $\partial_v$  derivatives explicit.<sup>11</sup>  $P$  and  $Q$  are appropriate structures built out of rest of the building blocks, which consist of the metric functions and  $\nabla_i$  acting on them. They do not contain any further  $\partial_v$  or  $\partial_r$  derivatives. Thus the most general structure of  $E_{vv}^{\text{HD}}$  would be

$$E_{vv}^{\text{HD}} = \sum_{m,n,k} E^{(m,n,k)} \quad (2.17)$$

The upper limits of these sum would be fixed by the number of derivative on the metric in the gravity Lagrangian (1.1).

Now we shall manipulate  $E^{(m,n,k)}$  to demonstrate that  $E_{vv}^{\text{HD}}$  as given by (2.17) can be cast into the form (2.8). At first, we shall consider  $E^{(m,n,k)}$  for  $k \neq 0$ , and derive a recursion relation for this quantity. This recursion relation would be used to derive a general structure for  $E^{(m,n,k)}$  and hence for  $E_{vv}^{\text{HD}}$ . Certain terms corresponding to the case  $k = 0$  would require special treatment, and after invoking the ‘physical process’ version of the first law, we shall demonstrate that entire sum (2.17), and hence the most general form of  $E_{vv}^{\text{HD}}$ , can be cast into the form (2.8). This would complete our objective as laid out in section 2.1, thus proving the second law of thermodynamics in the linearized case. This would also provide us with an explicit construction to compute the corrections to the Wald entropy, for an arbitrary theory (1.1).

At first, let us note that, any term of the form  $\partial_v^m \partial_r^n P$ , where  $n \geq m$  and  $P$  does not contain any further  $\partial_r$  or  $\partial_v$  derivative, would generically be non-zero once evaluated

---

<sup>11</sup>It may seem at first glance that the term in (2.16) is already second order in  $\epsilon$  for any non-zero  $m$ , since in that case, it becomes a product of two terms, on which some  $v$ -derivative has acted. However, it should be noted that when both the derivatives ( $\partial_v \partial_r$ ) act together on  $P$ , it can be non-zero on  $\mathcal{H}$  even in equilibrium, i.e. it can be an  $\mathcal{O}(\epsilon^0)$  term (see appendix A).

on the Killing horizon of the stationary solution. On the other hand, such terms would vanish on the Killing horizon if  $m > n$  (see appendix A for a demonstration of this fact). Therefore, it follows that, in a dynamical situation, when the amplitude of the time-dependent perturbation  $\epsilon$  is small,  $(\partial_v^m \partial_r^n P)$  must be of order  $\mathcal{O}(\epsilon)$ , whenever  $m > n$ . Consequently, all terms of the product form  $(\partial_v^m \partial_r^n P)(\partial_v^{m'} \partial_r^{n'} Q)$  with  $m > n$  and  $m' > n'$ , are of order  $\mathcal{O}(\epsilon^2)$ , and therefore, can be neglected in our linearized analysis.

Now let us turn our attention back to the expression (2.16), for  $k \neq 0$ . In (2.16), we can take the  $v$ -derivatives acting on the term involving  $Q$  and transfer them onto the term involving  $P$  at the expense of a minus sign and a total derivative term. By repeating this operation, even on the terms that are generated due to previous such operations, it is possible to reduce (2.16) to the following recurrence relation

$$E^{(m,n,k)} = \partial_v^2 \left( \sum_{p=0}^{k-1} (-1)^p \partial_r^{k-p} [(\partial_v \partial_r)^{m+p} P] \partial_v^{k-p} [(\partial_v \partial_r)^n Q] \right) + (-1)^k \partial_v \left( [(\partial_v \partial_r)^{m+k} P] \partial_v [(\partial_v \partial_r)^n Q] \right) - E^{(m+1,n,k-1)} \quad (2.18)$$

Now, we can use this recursion relation (2.18) itself, to evaluate  $E^{(m+1,n,k-1)}$

$$E^{(m+1,n,k-1)} = \sum_{p=0}^{k-2} (-1)^p \partial_v^2 \left( \partial_r^{k-p-1} [(\partial_v \partial_r)^{m+p+1} P] \partial_v^{k-p-1} [(\partial_v \partial_r)^n Q] \right) + (-1)^{k-1} \partial_v \left( [(\partial_v \partial_r)^{m+k} P] \partial_v [(\partial_v \partial_r)^n Q] \right) - E^{(m+2,n,k-2)} \quad (2.19)$$

Using the recursion relation repeatedly, we can recast  $E^{(m,n,k)}$  into the form

$$E^{(m,n,k)} = \sum_{q=0}^{k-1} \sum_{p=0}^{k-q-1} (-1)^{q+p} \partial_v^2 \left( \partial_r^{k-q-p} [(\partial_v \partial_r)^{m+q+p} P] \partial_v^{k-q-p} [(\partial_v \partial_r)^n Q] \right) + \underbrace{\left( (-1)^k - (-1)^{k-1} + (-1)^{k-2} - \dots \right)}_{k\text{-terms}} \partial_v \left( [(\partial_v \partial_r)^{m+k} P] \partial_v [(\partial_v \partial_r)^n Q] \right) \quad (2.20)$$

Hence, after performing the sum in the second term, we have

$$E^{(m,n,k)} = \sum_{q=0}^{k-1} \sum_{p=0}^{k-q-1} (-1)^{q+p} \partial_v^2 \left( \partial_r^{k-q-p} [(\partial_v \partial_r)^{m+q+p} P] \partial_v^{k-q-p} [(\partial_v \partial_r)^n Q] \right) + k(-1)^k \partial_v \left( [(\partial_v \partial_r)^{m+k} P] \partial_v [(\partial_v \partial_r)^n Q] \right) \quad (2.21)$$

Note that, inside the sum, the first term has the structure  $\partial_v^2 (\partial_r^\ell X^{(\ell)}) (\partial_v^\ell Y^{(\ell)})$ , where the lowest value of  $\ell = 1$ . If we now, plug in (2.21) back into the sum (2.17), a similar structure of this first term would be maintained, again with  $\ell$  starting from 1. Note that, here, we are only considering values of  $k$  starting from 1, in the sum (2.17). As we mentioned previously, the  $k = 0$  terms needs to be treated separately.



But now, let us also note from (2.16), that the  $k = 0$  term can always be recast into the same form as the second term in (2.21). This is true only in the linearized approximation.

Thus, from (2.17) and (2.21) it is clear that, at linear order in amplitude,  $E_{vv}^{\text{HD}}$  could always be written in the following form

$$\begin{aligned}
 E_{vv}^{\text{HD}} &= \partial_v [A \partial_v B] + \partial_v^2 \left[ \sum_{k=1} \left( \partial_r^k A^{(k)} \right) \left( \partial_v^k B^{(k)} \right) \right] + \mathcal{O}(\epsilon^2) \\
 &= \partial_v [A \partial_v B] + \partial_v \left[ \left( \frac{1}{\sqrt{h}} \right) \partial_v \sum_{k=1} \sqrt{h} \left( \partial_r^k A^{(k)} \right) \left( \partial_v^k B^{(k)} \right) \right] + \mathcal{O}(\epsilon^2)
 \end{aligned}
 \tag{2.22}$$

where  $A$ ,  $B$ ,  $A^{(k)}$  and  $B^{(k)}$  are appropriate structures as implied by (2.17) and (2.21), which do not transform under rescaling (2.11). Let us re-emphasize that, although (2.21) has been derived under the assumption  $k \neq 0$ , the form of  $E_{vv}^{\text{HD}}$  in (2.22), also incorporates the  $k = 0$  term, in the sum (2.17).

As argued above (also see appendix A), we know that the action of a  $v$ -derivative, which does not appear with a compensating  $r$ -derivative, on a quantity that is invariant under the rescaling (2.11), must be of  $\mathcal{O}(\epsilon)$  on  $\mathcal{H}$ , in the amplitude expansion. This is because, such a quantity should vanish on the Killing horizon, and so must be at least  $\mathcal{O}(\epsilon)$  for dynamical horizons. Consequently, whenever one or more  $\partial_v$  act on any one of these  $A$ ,  $B$ ,  $A^{(k)}$  or  $B^{(k)}$ , it must be  $\mathcal{O}(\epsilon)$ . This also justifies appropriate incorporation of the factors of  $\sqrt{h}$  in (2.22). We do not get any additional terms due to these factors of  $\sqrt{h}$ , since we are working in the linearized approximation in  $\epsilon$ .

Thus to conclude, we have obtained a precise structural form of  $E_{vv}^{\text{HD}}$  in (2.22). At this stage, we observe that the second term of  $E_{vv}^{\text{HD}}$  in (2.22) is already in the desired form (2.8). So our objective would be accomplished, if we are able to argue that the first term in (2.22), can also be written in the form (2.8), i.e. as two  $v$ -derivatives acting on a quantity which is invariant under the rescaling (2.11).

### 2.2.3 The physical process version of first law and its implications

In [1] it was argued that the structure of the quantities  $A$  and  $B$  in (2.22), must be such, that  $E_{vv}^{\text{HD}}$  has the form (2.8). This conclusion followed from the *physical process* version of the first law of black hole thermodynamics [15, 27] (also see [30, 37]), which was assumed to be applicable to the theory of gravity under consideration. It was demonstrated in [1], that if  $A$  and  $B$  in (2.22) did not have the requisite structures, then the physical process version of the first law would be invalidated. Let us now review this argument, as presented in [1].

Let us consider a stationary black hole, which is perturbed by small fluctuations in the matter sector. The amplitude of such perturbations is assumed to be small. For instance, this could be some small amount of matter falling into the black hole. The matter stress tensor would back-react onto the metric and produces fluctuations in it, which would also be small. These fluctuations would result in a non-stationary fluctuating black hole. However, in a physical situation, it may be expected that, at late times, these fluctuations would die down and the black hole would again become stationary. Such a dynamical process, where the black hole is stationary, both at early and late times, is referred to as a ‘physical process’.

The new stationary black hole at late times would have slightly different parameters compared to the one at early times (such as mass or angular momentum). The overall shift in the mass (energy) of the black hole would be given by integrating specific component of the energy-momentum tensor over the horizon. This shift in mass, must be related to the shift in entropy of the black hole through the first law of black hole thermodynamics  $T\Delta S = \Delta\mathcal{E}$ . Therefore, we can express the change in entropy during this physical process  $\Delta S$ , to the integrated energy-momentum tensor, in the following way

$$\Delta S = -\frac{2\pi}{\kappa} \int_{\mathcal{H}} \Delta T_{ab} \xi^a d\Sigma^b \tag{2.23}$$

Here,  $\Delta T_{ab}$  is the part of the energy-momentum tensor that has initiated the dynamics of the black hole horizon.  $\xi^a$  is the generator of the future horizon  $\mathcal{H}$ ; it is a Killing generator at early and late times, when the black hole is stationary. Also,  $d\Sigma^b$  is the area element along the horizon. The parameter  $\kappa$  is the surface gravity of the black hole and is proportional to the temperature of the black hole.<sup>12</sup>

The equation (2.23) is referred to as the physical process version of the first law of black hole thermodynamics. For a more complete and detailed discussion of this, see section (2) of [15]. We should note that because the initial and final states are stationary, the  $\Delta S$  in (2.23) is expected to be given by the change in Wald entropy, which, by construction, satisfies the *usual* form of the first law for stationary black holes. However, whether Wald entropy does satisfy the physical process version of the first law (2.23) does not immediately follow from its construction, and we require additional arguments to establish this, see [27, 30] for recent developments.

This version of the first law (2.23) now enables us to make further deductions regarding the structural form of  $A$  and  $B$  appearing in (2.22). With our choice of coordinates (2.1),  $\xi^a$  is related to the affinely parametrized null generators of the horizon  $\partial_v$ , in the following way

$$\xi^a \partial_a = \kappa v \partial_v \tag{2.24}$$

While, in our coordinates (2.1), the area element on the horizon  $\mathcal{H}$  is given by

$$d\Sigma^b \partial_b = -\sqrt{h} d^{d-2}x dv \partial_v \tag{2.25}$$

Using (2.24) and (2.25) back in (2.23), we have

$$\Delta S = 2\pi \int_{\mathcal{H}} \sqrt{h} d^{d-2}x dv v \Delta T_{vv} = 2\pi \int_{\mathcal{H}} \sqrt{h} d^{d-2}x dv v (R_{vv} + E_{vv}^{\text{HD}}) \tag{2.26}$$

Here, we have used the equation of motion to rewrite the stress tensor in terms of geometric quantities. Now, if entropy  $S$  has the form (2.2),  $\Delta S$  in (2.26) can be split into two parts  $\Delta S = \Delta S_E + \Delta S_{\text{HD}}$ .  $\Delta S_E$  being the change in the integrated area of  $\mathcal{H}_v$ , responsible for the

---

<sup>12</sup>It turns out that the combination  $\Delta T_{ab} \xi^a d\Sigma^b$  itself is of the order of amplitude of the perturbation. Therefore, as long as we are working at linear order in the amplitude of perturbation, the difference in the value of  $\kappa$  for the initial and final stationary black holes is negligible. Thus, within this approximation,  $\kappa$  can be taken to be constant throughout the duration of the physical process.

change in entropy in two derivative Einstein theory, while  $\Delta S_{HD}$  is the change in entropy due to higher derivative terms. Clearly, the terms proportional to  $R_{vv}$  on the r.h.s. of (2.26), must be equal to  $\Delta S_E$  on the l.h.s. of (2.26). This is manifest in the limit when the higher derivative corrections to Einstein's gravity vanish. Using (2.22), we can therefore write

$$\begin{aligned} \Delta S_{HD} &= 2\pi \int_{\mathcal{H}} \sqrt{h} d^{d-2}x dv v (E_{vv}^{HD}) \\ &= 2\pi \int_H \sqrt{h} d^{d-2}x dv v \partial_v \left[ A \partial_v B + \left( \frac{1}{\sqrt{h}} \right) \sum_{k=1} \partial_v \left( \sqrt{h} \partial_r^k A^{(k)} \partial_v^k B^{(k)} \right) \right] \\ &\quad + \mathcal{O}(\epsilon^2) \end{aligned} \tag{2.27}$$

It is extremely important for the subsequent arguments to realize that  $\Delta S_{HD}$  must be non-zero in general, and should be expressible in terms of some geometrical quantity integrated over  $\mathcal{H}_v$ . This is clear from the fact that for arbitrary higher derivative corrections to Einstein's gravity  $s_n$  in (2.2) is non-zero even for stationary black holes. In the stationary case, it is expected to be given by Wald entropy (2.3), which in general is different than area of the horizon. Thus, if  $s_n$  is non-trivial, it is expected that, in general, in a dynamical scenario,  $\Delta S_{HD}$  must be non-trivial. In fact, as pointed out earlier,  $\Delta S_{HD}$  should be given by change in the corresponding Wald entropy, since the state both at early and late times are stationary states.

Now the contribution from the second term in (2.27) vanishes. This can be seen by manipulating the term as follows.

$$\begin{aligned} &\int_{\mathcal{H}} \sqrt{h} d^{d-2}x dv v \partial_v \left[ \frac{1}{\sqrt{h}} \partial_v (\sqrt{h} X) \right] \\ &= \int_{\mathcal{H}} d^{d-2}x dv \partial_v \left[ v \partial_v (\sqrt{h} X) - (\sqrt{h} X) \right] + \mathcal{O}(\epsilon^2) \\ &= \int_{v=-\infty}^{v=\infty} dv \partial_v \left( \int_{\mathcal{H}_v} d^{d-2}x \left[ v \partial_v (\sqrt{h} X) - (\sqrt{h} X) \right] \right) + \mathcal{O}(\epsilon^2) \\ &= \left[ \int_{\mathcal{H}_v} d^{d-2}x v \partial_v (\sqrt{h} X) - \int_{\mathcal{H}_v} d^{d-2}x (\sqrt{h} X) \right]_{v=-\infty}^{v=\infty} \end{aligned} \tag{2.28}$$

where

$$X = \sum_{k=1}^N \left[ \left( \partial_r^k A^{(k)} \right) \left( \partial_v^k B^{(k)} \right) \right] + \mathcal{O}(\epsilon^2) \tag{2.29}$$

Note that both the terms in the last line of equation (2.28) contain more than one  $\partial_v$  derivatives on expressions that are invariant under the rescaling (2.11). Therefore, they must vanish in the two limits of far past and far future, where we have a stationary black hole with a Killing horizon. It follows that, these terms do not contribute to  $\Delta S_{HD}$  in (2.27). Hence, (2.27) can be always reduced to an integral of the form

$$\Delta S_{HD} = 2\pi \int_{\mathcal{H}} \sqrt{h} d^{d-2}x dv v \partial_v [A \partial_v B] \tag{2.30}$$

Again by performing an integration by parts, we obtain

$$\Delta S_{HD} = 2\pi \left[ \int_{\mathcal{H}_v} \sqrt{h} d^{d-2}x v (A \partial_v B) \right]_{v=-\infty}^{v=\infty} - 2\pi \int_{\mathcal{H}_v} \sqrt{h} d^{d-2}x dv [A \partial_v B] \tag{2.31}$$

However, to have a consistent first law, the second term also should be write-able as a total  $\partial_v$  derivative so that the  $v$  integration of this term from infinite past to infinite future finally would give the net change of some geometric quantity — entropy, defined on the constant  $v$  slices of the horizon.<sup>13</sup> Naively this is possible if  $A$  is of the form  $\left[ \text{some constant}/\sqrt{h} \right]$  and  $B$  has the form  $\left[ \sqrt{h} \tilde{B} \right]$  where  $\tilde{B}$  is a scalar under  $\{x^i\} \rightarrow \{y^i\} = y^i(\{\tilde{x}\})$ .

If it is indeed the case that  $A$  is always a constant times  $(1/\sqrt{h})$  (let's choose the constant to be one without any loss of generality), then the schematic form of  $E_{vv}^{\text{HD}}$

$$E_{vv}^{\text{HD}} = \partial_v \left( \frac{\partial_v (\sqrt{h} \tilde{B})}{\sqrt{h}} \right) + \partial_v \left[ \left( \frac{1}{\sqrt{h}} \right) \partial_v \left( \sqrt{h} \sum_{k=1}^N \partial_r^k A^{(k)} \partial_v^k B^{(k)} \right) \right] + \mathcal{O}(\epsilon^2) \quad (2.32)$$

If  $E_{vv}^{\text{HD}}$  does have the form of equation (2.32), not only the ‘physical process version’ of the first law but also the second law as argued in [1] will be true with the following identification for correction to the total entropy (see equation (2.8)).

$$\delta S_{\text{HD}} = \int_{\mathcal{H}_v} \left[ \tilde{B} + \sum_{k=1}^N \partial_r^k A^{(k)} \partial_v^k B^{(k)} \right] \quad (2.33)$$

Note that  $\tilde{B}$  is the only term that is non-zero even in equilibrium. This term must match with Wald entropy.<sup>14</sup> Rest of the terms (for  $k \geq 1$ ) vanish on stationary metric and therefore are part of JKM ambiguities.

So in summary, [1] has given a constructive proof for the second law for dynamical black hole solutions in higher derivative theories of gravity provided the physical process formulation of the first law is true for these solutions. The validity of the ‘physical process formulation of the first law’ requires a very specific structure for a certain term in the equation of motion (the first term in (2.22) must take the form of the first term in (2.32)), which does not follow from the boost-symmetry (2.11) alone (the only symmetry that is considered in [1]).<sup>15</sup> For the convenience of reporting, we shall refer to the first term in (2.32) (or the first term in (2.22)) as the ‘zero boost term’. This nomenclature is inspired by the fact that  $\tilde{B}$  in (2.32) (or  $A$  and  $B$  in (2.22)) does not have any  $v$  or  $r$  derivative. Apart from this, there is absolutely no other physical motivation behind this nomenclature. The reader must not confuse the phrase ‘zero boost term’ to be a synonym for boost invariant

<sup>13</sup>At this stage, by ‘geometric quantity’ we simply mean some expression in terms of the metric components and their derivatives that is invariant under any diffeomorphism, mixing only the spatial coordinates of the constant  $v$  slices of the horizon.

<sup>14</sup>Though we have explained the argument here, specializing to higher derivatives theories, it is trivially true for two derivative theories of gravity where  $\tilde{B}$  is simply 0, and entropy is simply given by  $S = \int_{\mathcal{H}_v} \sqrt{h}$ .

<sup>15</sup>This special structure of the first term in (2.32) has only been verified in specific theories of gravity where the physical process version of the first law has been proven (for instance, see [15]). To our knowledge, a complete proof demonstrating this special structure of the zero boost term in a general higher derivative theory of gravity does not exist. It would be interesting to explore, if it is possible to arrive at such a proof using the residual gauge transformations (2.9), which is more general than the boost symmetry (2.11) (see section 4 for further discussion).

term. By boost invariant terms we shall continue to mean such terms which are invariant under the rescaling symmetry (2.11).

### 3 An entropy current for four derivative theories of gravity

In the previous section 2, we have reviewed the proof of the second law for linearized fluctuations, following [1]. As emphasized earlier, this proof is designed to prove a second law for the ‘total entropy’ of the system. The proof crucially involves an integration over the full spatial slice of the horizon, which defines the ‘total entropy’. Therefore, it is insensitive to any total (spatial) derivative term, that may be present in the integrand, which is derived from the equation of motion. This drawback exists even in the proof for the physical version of the first law.

In this section, we shall carefully re-examine this particular subtlety. In explicit examples of four-derivative theories of gravity, we shall demonstrate that such total derivative terms do exist if we follow the algorithm of [1], and their inclusion would naturally lead to a construction of an entropy current. With the help of this entropy current, we can immediately prove an ultra-local version of second law, associated with any dynamical horizon  $\mathcal{H}$ .

Let us now elaborate this point further. In (2.32), we have shown that a special structure for  $E_{vv}^{\text{HD}}$  is necessary for the validity of both the second law, as well as the physical process version of the first law. As we have explained in section 2, the structure of the second term in (2.32) is fixed by the ‘boost symmetry’ (2.11), up to higher-order corrections in the amplitude of fluctuations  $\epsilon$ . But the same is not true for the first term in (2.32), which we have named as ‘zero-boost terms’ of  $E_{vv}^{\text{HD}}$ .

In section 2, we have argued that the physical version of the first law, and consequently, the second law, would be true if the ‘zero boost term’ in  $E_{vv}^{\text{HD}}$  has the following schematic structure (see (2.32))

$$E_{vv}^{\text{HD}}|_{\text{zero boost}} \sim \partial_v \left( \frac{1}{\sqrt{h}} \partial_v \left( \sqrt{h} \tilde{B} \right) \right), \tag{3.1}$$

where  $\tilde{B}$  is some scalar, which is invariant under spatial diffeomorphism.

However, the above form of the zero-boost term, though sufficient for the validity of the physical process version of the first law<sup>15</sup> and the second law, it is neither necessary, nor does it follow in any way, from the boost-symmetry (2.11). In this section, to begin with, our goal is just to verify (3.1). We shall explicitly compute the equation of motion, and in particular the zero-boost terms, in all possible four-derivative theories of gravity. From this explicit computation in four derivative theories of gravity we will show that (3.1) is not true in general. There exist cases where the zero boost terms in  $E_{vv}^{\text{HD}}$  could not be recast in the above form. In fact, the zero boost terms in  $E_{vv}^{\text{HD}}$  consists of additional terms, which can never be cast into the form (3.1).

Being motivated by this observation, we investigate the structural nature of the zero boost terms of  $E_{vv}^{\text{HD}}$ , to understand why such additional terms do not affect validity of the first and the second law. The possibility of non-zero spatial components of the entropy current arises here very naturally. Finally, through a general algorithm, we shall establish

that, the zero boost terms of  $E_{vv}^{\text{HD}}$  for every four derivative theories of gravity, could be rendered into a form, which guarantees an ultra-local version of the second law, in terms of an entropy current with non-zero spatial components.

### 3.1 Explicit calculation of $E_{vv}^{\text{HD}}$ and the entropy current for theories with four derivative corrections to Einstein gravity

In this subsection, we shall compute the ‘ $vv$ ’-component of the equation of motion,  $E_{vv}$ , for all possible four derivative theories of gravity. We shall immediately find that it is possible to rearrange the terms so that up to corrections of order  $\mathcal{O}(\epsilon^2)$  it takes the form

$$E_{vv} = -\partial_v \left[ \frac{1}{\sqrt{h}} \partial_v \left( \sqrt{h} J^v \right) + \nabla_i J^i \right] + \mathcal{O}(\epsilon^2). \tag{3.2}$$

Once we could rewrite  $E_{vv}$  in this form,<sup>16</sup> it is very natural to identify  $J^v$  with the entropy density and  $J^i$  as the spatial entropy current, capturing the in-flow and out-flow of entropy. Vanishing of  $E_{vv}$  at order  $\mathcal{O}(\epsilon)$  would then correspond to a locally conserved entropy current and therefore an ultra-local version of the second law (see section 2 for details of this argument).<sup>17</sup> More explicitly, once  $E_{vv}$  has the form (3.2), the standard arguments outlined in section 2 would imply that<sup>18</sup>

$$\frac{1}{\sqrt{h}} \partial_v \left( \sqrt{h} J^v \right) + \nabla_i J^i = \mathcal{O}(\epsilon^2). \tag{3.4}$$

There are only three possible covariant terms which can appear in the gravity Lagrangian with 4-derivatives on the metric. These are given by:  $R^2$ ,  $R_{\mu\nu}R^{\mu\nu}$ ,  $R_{\mu\nu\sigma\lambda}R^{\mu\nu\sigma\lambda}$ .

---

<sup>16</sup>Note that for Einstein gravity  $E_{vv}$  takes the simple form

$$E_{vv}^{\text{Einstein}} = -\partial_v \left[ \frac{1}{\sqrt{h}} \partial_v \left( \sqrt{h} \right) \right].$$

Hence, when we consider higher derivative corrections to Einstein’s equations, the terms in this equation arising out of these corrections also has a similar form

$$E_{vv}^{\text{HD}} = -\partial_v \left[ \frac{1}{\sqrt{h}} \partial_v \left( \sqrt{h} \tilde{J}^v \right) + \nabla_i J^i \right] + \mathcal{O}(\epsilon^2),$$

where  $J^v - \tilde{J}^v = 1$ . For most of our analysis, especially in the abstract manipulations, we have used  $E_{vv}^{\text{HD}}$ , instead of  $E_{vv}$ .

<sup>17</sup>The calculations here clearly suggest about the existence of an entropy current for some of these higher derivative theories. However, it requires a bit of clever manipulation. See the following subsections for a more algorithmic method which clearly exhibits that we need the spatial entropy current, which in turn provides us with an ultra-local version of the second law.

<sup>18</sup>If we consider special processes where the metric is entirely sourced by a small matter energy-momentum tensor, so that both the first correction to the metric as well as the matter energy-momentum tensor are of  $\mathcal{O}(\epsilon^2)$  (the  $\mathcal{O}(\epsilon)$  correction to the metric being zero), then for the  $\epsilon^2$  coefficient, (3.4) would be modified to the inequality

$$\left( \frac{1}{\sqrt{h}} \partial_v \left( \sqrt{h} J^v \right) + \nabla_i J^i \right) \Big|_{\epsilon^2} \geq 0. \tag{3.3}$$

Note that, while deducing this inequality, we have assumed that there exist other matter fields satisfying the null energy condition. See section 2.1 for a more detailed discussion of this point.

In the following subsections, we shall separately consider three different four derivative theories of gravity

1. Ricci scalar squared theory:  $\mathcal{I}^{(1)} = \int d^d x \sqrt{-g} (R + a_1 R^2),$
2. Ricci tensor squared theory:  $\mathcal{I}^{(2)} = \int d^d x \sqrt{-g} (R + a_2 R_{\mu\nu} R^{\mu\nu}),$
3. Riemann tensor squared theory:  $\mathcal{I}^{(3)} = \int d^d x \sqrt{-g} (R + a_3 R_{\mu\nu\rho\sigma} R^{\mu\nu\rho\sigma}),$

and explicitly compute the ‘ $vv$ ’-component of the respective equations of motion,  $E_{vv}$ , for each of them. After some algebraic manipulations on  $E_{vv}$ , in each of these cases, we shall write down the entropy current, It is then trivial to combine these results, to give us the entropy current for any arbitrary four derivative theory of gravity. The final result is tabulated in table 1.

For each of the three four-derivative theories mentioned above, if we just evaluate the equation of motion on our gauge fixed metric (2.1), it turns out to be an extremely complicated expression, even after we restrict it to the horizon. In general, just by inspection, it is quite difficult to rearrange the terms to arrive at the form (3.2). However, we know that in stationary situations, at least  $J^v$  should reduce to the well-known form of Wald entropy and the rest of the terms must be such that they vanish in a stationary situation. We shall use this fact to guide our intuition about the form of the entropy density and then finally deduce the form of the entropy current. More precisely, we shall obtain the following constituents for  $J^v$

$$J^v = \sqrt{h} (s_w + s_c) \tag{3.5}$$

where  $s_w$  is the Wald entropy density for the stationary black holes defined as

$$s_w = \frac{\partial \mathcal{L}}{\partial R_{\mu\nu\rho\sigma}} \epsilon_{\mu\nu} \epsilon_{\rho\sigma}, \tag{3.6}$$

where  $\epsilon_{\mu\nu}$  are the bi-normal to  $\mathcal{H}_v$ , the co-dimension–2 spatial slicing of the horizon. Also,  $s_c$  is the non-stationary correction to  $s_w$ . As we have argued before,  $s_c$  will vanish once we take stationary limit. Let us also define the contribution to  $s_w$  from the higher derivative part of the action as  $s_w^{\text{HD}}$

$$s_w^{\text{HD}} = \frac{\partial \mathcal{L}^{\text{HD}}}{\partial R_{\mu\nu\rho\sigma}} \epsilon_{\mu\nu} \epsilon_{\rho\sigma}. \tag{3.7}$$

At this point, let us clarify one subtlety regarding the split mentioned in (3.5). It turns out that if we evaluate  $s_w$  on any dynamical metric, along with the terms that contribute in stationary situation, it will also have terms that vanish in the stationary limit. For convenience, let us name such terms as ‘off-equilibrium’ structures. Such off-equilibrium terms in the entropy suffer from the well-known class of JKM ambiguities, which arises as soon as we try to extrapolate Wald’s formalism to non-stationary solutions. In our identification of the entropy density, we have used the fact that, in the stationary limit, it should reduce to Wald entropy. As we will see in the later sections, this requirement also fixes one class of ambiguities, in defining the entropy current.

For convenience, let us separate out the contribution of Wald entropy density  $s_w^{\text{HD}}$  to  $E_{vv}^{\text{HD}}$  and define the quantity  $\mathbb{E}_{vv}^{\text{HD}*}$  as follows

$$\mathbb{E}_{vv}^{\text{HD}*} \equiv E_{vv}^{\text{HD}} + \partial_v \left[ \frac{1}{\sqrt{h}} \partial_v \left( \sqrt{h} s_w^{\text{HD}} \right) \right] \quad (3.8)$$

Then, from the definition (3.5) it follows that

$$\mathbb{E}_{vv}^{\text{HD}*} = -\partial_v \left[ \frac{1}{\sqrt{h}} \left( \partial_v \sqrt{h} s_c \right) + \nabla_i J^i \right] + \mathcal{O}(\epsilon^2). \quad (3.9)$$

It turns out that, in the examples that we consider, algebraically it is comparatively easier to recast  $\mathbb{E}_{vv}^{\text{HD}*}$  in the form (3.9), instead of dealing with the full  $E_{vv}^{\text{HD}}$ .

We would like to emphasize here that, the procedure adopted in this subsection, is a set of intuitive manipulations and educated guess-work. It gives us an explicit demonstration that for the theories that we consider here, it is possible to lift both the first and the second law to an ultra-local form, by entertaining the possibility of non-zero spatial components of the entropy current, which captures the effect of the inflow and outflow of the entropy from any arbitrary local sub-region. However, at this stage, we would not be able to say, whether the spatial components of the current is an absolute necessity, or there exist other possible rearrangements of terms, such that we can avoid the spatial components of the current altogether. In the later subsections, we shall repeat the same analysis more systematically, and for the four derivative theories, we shall be able to quantify these ambiguities involved in defining the entropy current more precisely. We shall conclude that, although there are some ambiguities in defining the entropy current, its non-zero spatial components are an unavoidable feature of the ultra-local form of the second law.

### 3.1.1 Ricci scalar square theory

The action for Ricci scalar square theory is

$$\mathcal{I}^{(1)} = \int d^d x \sqrt{-g} (R + a_1 R^2) \quad (3.10)$$

where  $a_1$  is an arbitrary constant. The equations of motion which follows from the action (3.10), is given by

$$E_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R + E_{\mu\nu}^{\text{HD}} = 0, \quad (3.11)$$

where

$$E_{\mu\nu}^{\text{HD}} = a_1 \left( 2RR_{\mu\nu} - 2D_\mu D_\nu R + 2g_{\mu\nu} D^\rho D_\rho R - \frac{1}{2} g_{\mu\nu} R^2 \right) \quad (3.12)$$

are the higher derivative corrections to the Einstein equation. The explicit form of the vv-component of the equations of motion, on the horizon, is

$$E_{vv} = R_{vv} + E_{vv}^{\text{HD}}, \quad (3.13)$$

where

$$E_{vv}^{\text{HD}} = a_1 (2RR_{vv} - 2D_v D_v R). \quad (3.14)$$



The Wald entropy for this theory happens to be

$$S_w = \int_{\mathcal{H}_v} d^{d-2}x \sqrt{h} (1 + 2 a_1 R). \quad (3.15)$$

Once we have the Wald entropy, we could compute  $\mathbb{E}_{vv}^{\text{HD}*}$ . In this case it simply vanishes implying that we do not need to add any current, neither do we get any correction to entropy density, beyond what is given by the Wald entropy.

### 3.1.2 Ricci tensor square theory

In this theory, the Ricci tensor square is added to the Einstein-Hilbert action, as a higher derivative correction. We have

$$I = \int d^d x \sqrt{-g} (R + a_2 R_{\mu\nu} R^{\mu\nu}) \quad (3.16)$$

The equations of motion, for this theory are given by

$$\begin{aligned} E_{\mu\nu} &= R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R + E_{\mu\nu}^{\text{HD}} = 0, \quad \text{where} \\ E_{\mu\nu}^{\text{HD}} &= a_2 \left( 2R^{\alpha\beta} R_{\mu\alpha\nu\beta} - D_\mu D_\nu R + D^\alpha D_\alpha R_{\mu\nu} + \frac{1}{2} g_{\mu\nu} D^\alpha D_\alpha R - \frac{1}{2} g_{\mu\nu} R_{\alpha\beta} R^{\alpha\beta} \right) \end{aligned} \quad (3.17)$$

The explicit form of the vv-component of the equations of motion on the horizon, is as given below

$$\begin{aligned} E_{vv} &= R_{vv} + E_{vv}^{\text{HD}} = 0, \\ E_{vv}^{\text{HD}} &= a_2 (2R^{\alpha\beta} R_{v\alpha v\beta} - D_v D_v R + D^\alpha D_\alpha R_{vv}). \end{aligned} \quad (3.18)$$

The Wald entropy for this theory is given by

$$S_w = \int_{\mathcal{H}_v} d^{d-2}x \sqrt{h} (1 + 2 a_2 R_{rv}), \quad (3.19)$$

so that  $s_w^{\text{HD}} = 2 a_2 R_{rv}$ . Once we have obtained the Wald entropy, we can compute  $\mathbb{E}_{vv}^{\text{HD}*}$ . Using the form of the metric (2.1) and the formulae provided in appendix C, we evaluate  $\mathbb{E}_{vv}^{\text{HD}*}$  explicitly in terms of metric functions and their derivatives.

$$\mathbb{E}_{vv}^{\text{HD}*} = a_2 \partial_v \left[ \frac{1}{\sqrt{h}} \partial_v (\sqrt{h} K \bar{K}) \right] + a_2 \partial_v [\nabla_i (\nabla^i K + h^{ij} \partial_v \omega_j - 2 \nabla_j K^{ij})]. \quad (3.20)$$

Now, we could easily re-express  $\mathbb{E}_{vv}^{\text{HD}*}$  in the form of (3.9). Subsequently, it is straightforward to identify the current as

$$\begin{aligned} J^v &= -s_w^{\text{HD}} - a_2 K \bar{K}, \\ J^i &= a_2 (2 \nabla_j K^{ij} - \nabla^i K - h^{ij} \partial_v \omega_j). \end{aligned} \quad (3.21)$$

### 3.1.3 Riemann tensor square theory

The action for Riemann tensor square theory is

$$I = \int d^d x \sqrt{-g} (R + a_3 R_{\mu\nu\rho\sigma} R^{\mu\nu\rho\sigma})$$

The corresponding equations of motion are

$$E_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R + E_{\mu\nu}^{\text{HD}} = 0, \quad (3.22)$$

where

$$E_{\mu\nu}^{\text{HD}} = a_3 \left( 4R^{\alpha\beta} R_{\mu\alpha\nu\beta} - 2D_\mu D_\nu R + 4D^\alpha D_\alpha R_{\mu\nu} - 4R_\mu^\alpha R_{\nu\alpha} - \frac{1}{2} g_{\mu\nu} R_{\alpha\beta\gamma\sigma} R^{\alpha\beta\gamma\sigma} + 2R_\mu^{\alpha\beta\sigma} R_{\nu\alpha\beta\sigma} \right). \quad (3.23)$$

The vv-component of the equations of motion is

$$E_{vv} = R_{vv} + E_{vv}^{\text{HD}} = 0, \\ E_{vv}^{\text{HD}} = a_3 \left( 4R^{\alpha\beta} R_{v\alpha v\beta} - 2D_v D_v R + 4D^\alpha D_\alpha R_{vv} - 4R_v^\alpha R_{v\alpha} + 2R_v^{\alpha\beta\sigma} R_{v\alpha\beta\sigma} \right). \quad (3.24)$$

The Wald entropy for this theory will be

$$s_w = \int_{\mathcal{H}_v} d^{d-2} x \sqrt{h} (1 - 4 a_3 R_{rvrv}),$$

such that  $s_w^{\text{HD}} = -4 a_3 R_{rvrv}$ .

Once we have Wald entropy, it is easy to compute  $\mathbb{E}_{vv}^{\text{HD}*}$ . Using the form of the metric (2.1), and the formulae provided in appendix C, we can evaluate  $\mathbb{E}_{vv}^{\text{HD}*}$ , explicitly in terms of metric functions and their derivatives. We find that

$$\mathbb{E}_{vv}^{\text{HD}*} = 4 a_3 \partial_v \left[ \frac{1}{\sqrt{h}} \partial_v \left( \sqrt{h} K_{ij} \bar{K}^{ij} \right) \right] + 4 a_3 \partial_v \left[ \nabla_i \left( h^{ij} \partial_v \omega_j - \nabla_j K^{ij} \right) \right], \quad (3.25)$$

which has been expressed in the structural form (3.9). From this, it is again straightforward to read off the entropy current to be

$$J^v = -s_w^{\text{HD}} - 4 a_3 K_{ij} \bar{K}^{ij}, \\ J^i = -4 a_3 \left( h^{ij} \partial_v \omega_j - \nabla_j K^{ij} \right). \quad (3.26)$$

### 3.2 The most general structure of the ‘zero boost term’ in $E_{vv}^{\text{HD}}$

Determining the equation of motion and in particular, its ‘vv’-component, given the coordinate choice in (2.1), are, in principle, a straightforward task. But, it becomes increasingly tedious with the number of derivatives present in the action. Also, as we will see in section 3.4, the unambiguous definition of the spatial components of the entropy current arises

1	Ricci scalar square theory:  $I = \int d^d x \sqrt{-g} (R + a_1 R^2)$	$E_{vv}^{\text{HD}} = -\partial_v \Theta + \mathcal{O}[\epsilon]^2$  $\Theta = \frac{2a_1}{\sqrt{h}} \partial_v (\sqrt{h} R)$
2	Ricci tensor square theory:  $I = \int d^d x \sqrt{-g} (R + a_2 R_{\mu\nu} R^{\mu\nu})$	$E_{vv}^{\text{HD}} = -\partial_v \Theta - \partial_v (\nabla_i J^i) + \mathcal{O}[\epsilon]^2$  $\Theta = \frac{a_2}{\sqrt{h}} \partial_v [\sqrt{h} (2R_{rv} - \bar{K}K)]$  $J_i = a_2 [2\nabla^j K_{ij} - \nabla_i K - \partial_v \omega_i]$
3	Riemann tensor square theory:  $I = \int d^d x \sqrt{-g} (R + a_3 R_{\mu\nu\sigma\lambda} R^{\mu\nu\sigma\lambda})$	$E_{vv}^{\text{HD}} = -\partial_v \Theta - \partial_v (\nabla_i J^i) + \mathcal{O}[\epsilon]^2$  $\Theta = \frac{4a_3}{\sqrt{h}} \partial_v (\sqrt{h} (-R_{rvrv} + \bar{K}_{ij} K^{ij}))$  $J_i = 4a_3 [\nabla^j K_{ij} - \partial_v \omega_i]$

**Table 1.** Table showing the higher derivative corrections to Einstein's equations, for all possible 4-derivative theories of gravity.

out of the zero boost term in  $E_{vv}^{\text{HD}}$ . This implies that, for the construction of the entropy current, we do not need the equation of motion in its every detail. What we need is a very specific set of terms in  $E_{vv}^{\text{HD}}$ , namely the terms that could be written in the form of the first term in (2.22). For convenience, we are re-writing (2.22) here again

$$E_{vv}^{\text{HD}} = \underbrace{\partial_v [A \partial_v B]}_{(=E_{vv}^{\text{HD}}|_{\text{zero boost}}} + \underbrace{\frac{1}{\sqrt{h}} \partial_v^2 \left[ \sum_{k \geq 1} \sqrt{h} \left( \partial_r^k A^{(k)} \right) \left( \partial_v^k B^{(k)} \right) \right]}_{(=E_{vv}^{\text{HD}}|_{\text{higher boost}}} + \mathcal{O}(\epsilon^2). \quad (3.27)$$

In this subsection, we would like to develop an algorithm that would isolate out these zero-boost terms in  $E_{vv}^{\text{HD}}$ . The most important feature of these terms is that at linear order in amplitude expansion of the perturbations, it is always possible to rewrite them as

$$E_{vv}^{\text{HD}} \Big|_{\text{zero boost}} = \partial_v [A \partial_v B] \sim A \partial_v^2 B + \mathcal{O}(\epsilon^2), \quad (3.28)$$

where both  $A$  and  $B$  are boost invariant quantities and they are non-vanishing on the stationary solutions. Hence here our main focus would be to search for terms of the form  $A \partial_v^2 B$  in the ‘vv’-component of the linearized equation of motion,  $E_{vv}^{\text{HD}}$ . However, before proceeding to extract the zero boost terms from  $E_{vv}^{\text{HD}}$ , let first point out an important ambiguity in defining the zero boost terms, through the structure (3.28).

**Generating terms like  $A \partial_v^2 B$  from the  $k = 1$  terms in (3.27).** Before proceeding further with the zero boost terms, we would like to discuss one subtle point that will be important in our attempt to separate out the  $k = 0$  terms (i.e. the zero boost sector) from the  $k \neq 0$  ones in (3.27). Recall that our final goal is to determine the form of the boost invariant terms  $A$  and  $B$  in (3.27) and we plan to do that by keeping track of the terms of form  $A \partial_v^2 B$  in the linearized  $E_{vv}^{\text{HD}}$ . However, the strategy mentioned above to uniquely extract out the zero boost terms from linearized  $E_{vv}^{\text{HD}}$  would be unsuccessful if there is a possibility of generating terms of the form  $A \partial_v^2 B$  (with  $A$  and  $B$  being boost invariant) from the second term in equation (3.27). As we will see now, there is indeed such a possibility of contamination arising from the term  $k = 1$  in the summation on the r.h.s. of (3.27). Let us analyze this term more carefully

$$\begin{aligned} E_{vv}^{\text{HD}} \Big|_{k=1} &= \frac{1}{\sqrt{h}} \partial_v^2 \left[ \sqrt{h} \left( \partial_r A^{(1)} \right) \left( \partial_v B^{(1)} \right) \right] + \mathcal{O}(\epsilon^2) \\ &= 2 \left( \partial_v \partial_r A^{(1)} \right) \left( \partial_v^2 B^{(1)} \right) + \left( \partial_r A^{(1)} \right) \left( \partial_v^3 B^{(1)} \right) + \mathcal{O}(\epsilon^2). \end{aligned} \quad (3.29)$$

In (3.29) above the first term is precisely of the form  $\sim X \partial_v^2 Y$ , where  $(\partial_v \partial_r A^{(1)})$  and  $B^{(1)}$  respectively can be added to  $A$  and  $B$  of the zero boost terms. Thus, the terms of our interest  $A \partial_v^2 B$ , which we are looking for in  $E_{vv}^{\text{HD}}$  can be contaminated by terms generated from  $\partial_v [B^{(1)} \partial_v (\partial_v \partial_r A^{(1)})]$ . This clearly demonstrates that it is impossible to uniquely determine  $A$  and  $B$ , appearing in the first term of (3.27), just by looking at the terms of the form  $A \partial_v^2 B$  alone in  $E_{vv}^{\text{HD}}$ ; it would be difficult to know if they arise from  $k = 0$  or  $k = 1$  terms in our classification (3.27) for the terms in  $E_{vv}^{\text{HD}}$ .

	Candidate terms	Derivative counting	Boost weight
1.	$h_{ij}$	zero	zero
2.	$\omega_i$	one	zero
3.	$X$	two	zero

**Table 2.** The basic building blocks.

With this subtlety in mind, let us also comment on the way to tackle this issue. We can subtract off the contributions coming from the  $k = 1$  terms, that are of the same form as the  $k = 0$  terms in (3.27). This could be done easily by noting that whenever such a term is generated from  $k = 1$  piece, it will also generate the second term in equation (3.29). Hence to determine  $A$  and  $B$  unambiguously and construct  $E_{vv}^{\text{HD}}|_{\text{zero boost}}$ , we will have to isolate out few special terms of the form  $(\partial_r X)(\partial_v^3 Y)$  in  $E_{vv}^{\text{HD}}$ , with  $X$  and  $Y$  being boost invariant.

Note that, due to the structural nature of the terms, a similar issue may also arise from the  $k = 2$  term in (3.27). However, we are not discussing the  $k = 2$  case in greater detail here, because such terms would not arise in four derivative theories of gravity. This is because, there are a total of six derivatives in the  $k = 2$  terms.

**Algorithm to uniquely extract the terms like  $A \partial_v^2 B$  from linearized  $E_{vv}^{\text{HD}}$ .** Our job will now be to develop an algorithm, to determine the most general structure of this  $k = 0$  ‘zero boost term’ appearing in (3.29), keeping in mind the above-mentioned subtlety.

It is clear from our previous discussions that for constructing the entropy current, which satisfies the strongest form of the second law, we need the knowledge of the zero boost term in  $E_{vv}^{\text{HD}}$ , only on the horizon  $\mathcal{H}$ . This in turn means that, the  $g_{vi} = r \omega_i$  component of the metric (2.1) can appear in the zero boost term, only after differentiation with one  $\partial_r$ , and the  $g_{vv} = r^2 X$  component can appear only after the action of two  $\partial_r$ . The spatial components of the metric  $h_{ij}$  can appear without any derivative acting on it. So the basic building blocks, for constructing the zero boost term on the horizon, are given in table 2.

Let us first concentrate on terms of the form  $(X \partial_v^2 Y)$  and isolate such terms in  $E_{vv}^{\text{HD}}$  when we have a four derivative theory of gravity. These terms in  $E_{vv}^{\text{HD}}$  can be constructed by applying  $\partial_r$ ,  $\partial_v$  and  $\nabla_i$  on these building blocks,<sup>19</sup> so that the total number of derivatives are always equal to four,<sup>20</sup> when we restrict to the four derivative theories of gravity.

---

<sup>19</sup>To begin with the structures that appear in  $E_{vv}^{\text{HD}}$  will have only  $\partial_i$ . However, we know that  $E_{vv}^{\text{HD}}$  is a scalar with respect to the coordinate transformation that only mixes the  $\{x^i\}$  coordinates among themselves. If we want to construct scalars out of the horizon data with spatial derivatives on the three building blocks, it must be combined with appropriate spatial derivatives of  $h_{ij}$  so that it finally becomes a covariant derivative with respect to  $h_{ij}$ . This covariance with respect to the mixing of  $\{x^i\}$  tells us that just spatial derivatives of  $h_{ij}$  need not be taken as any independent data.

Also note that in our set-up  $r$  and  $v$  are genuinely distinguished coordinates and we do not demand any covariance with respect to the transformation that mixes these two coordinates among themselves and others. Therefore the derivatives with respect to  $r$  and  $v$  would remain as simple  $\partial_r$  and  $\partial_v$ .

<sup>20</sup>In this derivative counting  $\omega_i$  and  $X$  must be taken as one derivative and two derivative data, respectively.

	Candidate structures	Derivative counting	Boost weight
1.	$(\nabla_{j_1} \cdots \nabla_{j_p}) (\partial_r \partial_v)^{m_1} h_{ij}$	$p + 2 m_1$	zero
2.	$(\partial_r \partial_v)^{m_2} (\nabla_{j_1} \cdots \nabla_{j_q}) \omega_i$	$q + 2 m_2$	zero
3.	$(\partial_r \partial_v)^{m_3} (\nabla_{j_1} \cdots \nabla_{j_r}) X$	$r + 2 m_3$	zero

**Table 3.** ‘Equilibrium’ and ‘boost-invariant’ structures built out of the basic building blocks.

	Candidate structures	Derivative counting	Boost weight
1.	$(\nabla_{j_1} \cdots \nabla_{j_p}) \partial_v^2 (\partial_r \partial_v)^{m_1} h_{ij}$	$p + 2 m_1 + 2$	two
2.	$\partial_v^2 (\partial_r \partial_v)^{m_2} (\nabla_{j_1} \cdots \nabla_{j_q}) \omega_i$	$q + 2 m_2 + 2$	two
3.	$\partial_v^2 (\partial_r \partial_v)^{m_3} (\nabla_{j_1} \cdots \nabla_{j_r}) X$	$r + 2 m_3 + 2$	two

**Table 4.** The list of ‘off-equilibrium’ and ‘boost-weight= 2’ data built out of the basic building blocks.

For convenience, we shall now classify the data in two categories:

1. *equilibrium* data and 2. *off-equilibrium* data.

As it is clear from the names, ‘equilibrium data’ are those structures that are non-vanishing even in a stationary situation, whereas ‘off-equilibrium data’ vanishes when stationary limits are taken. Now, from the discussion in appendix A it follows that ‘equilibrium data’ must be ‘boost-invariant’ and therefore could have definite structures and their appropriate products as listed in table 3.

On the other hand, the ‘off-equilibrium data’ are not boost-invariant, i.e., the total number of  $\partial_v$  should be more than the total number of  $\partial_r$ , when we consider these two derivatives as operators acting on the three basic building blocks listed above. In general there are many possibilities for such ‘off-equilibrium data’. However, here we are interested in a very specific term in  $E_{vv}^{\text{HD}}$ , where the total number of  $\partial_v$ ’s is exactly two more than the number of  $\partial_r$ ’s ( again considering them as operators on the basic building blocks and not directly on the metric components). Also both of these two extra  $\partial_v$ ’s must be acting on the same structure, otherwise it would generate a term which is second order in terms of the amplitude expansion we are considering here. ‘Off-equilibrium data’ with this property could have the following structures in general as given in table 4.

Finally, we have to contract the ‘equilibrium data’ and ‘off-equilibrium data’ appropriately to get the scalar term in  $E_{vv}^{\text{HD}}$ . Since in this note we are focusing only on the four-derivative theories of gravity, every term in  $E_{vv}^{\text{HD}}$  contains four-derivatives on the metric components. So the relevant equilibrium data can have a maximum of two derivatives acting on the metric components, the possible structures are listed in table 5. Following the same argument to maintain the derivative counting, the relevant ‘off-equilibrium data’ are listed below in table 6.

Now our job is to contract these two sets of data as given in table 5 and table 6, to get the candidate scalar terms in  $E_{vv}^{\text{HD}}$ , maintaining the count of total number of derivatives

Equilibrium and boost-invariant data		Number of derivatives	
1.	Tensor structures:	$T_{ij}^{(1)} \equiv \partial_r \partial_v h_{ij}$	2
		$T_{ij}^{(2)} \equiv \nabla_i \omega_j$	2
		$T_{ij}^{(3)} \equiv \mathcal{R}_{ij}$	2
2.	Vector structure:	$V_i^{(1)} \equiv \omega_i$	1
3.	Scalar Structure:	$S^{(1)} \equiv X$	2

**Table 5.** Relevant equilibrium and boost invariant data with maximum number of derivatives= 2, in four-derivative theories of gravity.

Off-equilibrium and boost-weight= 2 data		Number of derivatives	
1.	Tensor structures:	$T_{ij}^{(4)} \equiv \nabla_i \nabla_j (\partial_v^2 h_{kl})$	4
		$T_{ij}^{(5)} \equiv \nabla_i (\partial_v^2 h_{jk})$	3
		$T_{ij}^{(6)} \equiv \partial_v^2 (\partial_r \partial_v h_{ij})$	4
		$T_{ij}^{(7)} \equiv \partial_v^2 (\nabla_i \omega_j)$	4
		$T_{ij}^{(8)} \equiv \partial_v^2 h_{ij}$	2
2.	Vector structure:	$V_i^{(2)} \equiv \partial_v^2 \omega_i$	3
3.	Scalar Structure:	$S^{(2)} \equiv \partial_v^2 X$	4

**Table 6.** Relevant off-equilibrium data with maximum number of derivatives= 4 and boost-weight= 2, in four-derivative theories of gravity. Let us emphasize that within the tensor structures there are three types of terms: (i) 4-index structure:  $T_{ij}^{(4)}$ , (ii) 3-index structure:  $T_{ij}^{(5)}$ , (iii) 2-index structure:  $T_{ij}^{(6)}$ ,  $T_{ij}^{(7)}$  and  $T_{ij}^{(8)}$ .

equal to four. This could be done systematically as outlined below:

- The four-indexed tensor structure  $T_{ij}^{(4)}$  itself has four derivatives. Therefore the free indices have to be contracted with zero derivative ‘equilibrium-data’ or just among themselves. Now, there is no ‘equilibrium-data’ that has zero derivatives, see table 5. Therefore, self contraction of the indices in  $T_{ij}^{(4)}$  is the only possibility here and it could be done in two ways leading to two different scalar structures:

$$T_1 = h^{ij} h^{kl} \nabla_i \nabla_j (\partial_v^2 h_{kl}), \quad T_2 = h^{ik} h^{jl} \nabla_i \nabla_j (\partial_v^2 h_{kl})$$

- The three-indexed ‘off-equilibrium data’  $T_{ij}^{(5)}$  has three derivatives and therefore it has to be contracted with one derivative ‘equilibrium data’ ( $V_i^{(1)} = \omega_i$ ). Here also, two different types of contractions are possible leading to two different scalars:

$$T_3 = h^{ij} h^{kl} \omega_i \nabla_j (\partial_v^2 h_{kl}), \quad T_4 = h^{ik} h^{jl} \omega_i \nabla_j (\partial_v^2 h_{kl})$$

- The two ‘off-equilibrium’ tensor structures with 2 indices,  $T_{ij}^{(6)}$  and  $T_{ij}^{(7)}$ , themselves have four derivatives and therefore the free indices have to be contracted among

themselves. In each case there is only one way the contraction could be done. The resultant scalars are

$$T_5 = h^{ij} \partial_v^2 (\partial_r \partial_v h_{ij}), \quad T_6 = h^{ij} \partial_v^2 (\nabla_i \omega_j)$$

- The last ‘off equilibrium’ tensor structure  $T_{ij}^{(8)}$  has two derivatives. It has to be contracted with two derivative ‘equilibrium-data’ and also the equilibrium data must have even number (in this case it could be either zero or two) of free indices so that contraction is possible. Here we get the following structures:

$$\begin{aligned} T_7 &= X h^{ij} (\partial_v^2 h_{ij}), \\ T_8 &= h^{ij} h^{kl} (\nabla_i \omega_j) (\partial_v^2 h_{kl}), \quad T_9 = h^{ik} h^{jl} (\nabla_i \omega_j) (\partial_v^2 h_{kl}), \\ T_{10} &= h^{ij} h^{kl} (\omega_i \omega_j) (\partial_v^2 h_{kl}), \quad T_{11} = h^{ik} h^{jl} (\omega_i \omega_j) (\partial_v^2 h_{kl}), \\ T_{12} &= h^{ij} h^{kl} (\partial_r \partial_v h_{ij}) (\partial_v^2 h_{kl}), \quad T_{13} = h^{ik} h^{jl} (\partial_r \partial_v h_{ij}) (\partial_v^2 h_{kl}) \end{aligned}$$

- The ‘off-equilibrium’ vector data,  $V_i^{(2)}$ , is a three-derivative structure therefore it has to be contracted with one derivative ‘equilibrium data’  $V_i^{(1)} = \omega_i$ , leading to the following scalar structure

$$T_{14} = h^{ij} \omega_i \partial_v^2 \omega_j$$

- The ‘off-equilibrium’ scalar data,  $S^{(2)}$ , itself is a four-derivative and no contraction is needed.

$$T_{15} = \partial_v^2 X$$

- Considering possible contractions between the ‘equilibrium data’  $T_{ij}^{(3)}$  (given in terms of the intrinsic curvature of  $\mathcal{H}_v$ ), and the ‘off equilibrium’ tensor structure  $T_{ij}^{(8)}$ , we can also get two more terms as given below

$$T_{16} = h^{ik} h^{jl} \mathcal{R}_{kl} \partial_v^2 h_{ij}; \quad T_{17} = h^{ij} h^{kl} \mathcal{R}_{kl} \partial_v^2 h_{ij}.$$

- Finally, as we have already mentioned in the beginning of this subsection, to determine the boost-invariant  $A$  and  $B$  in (3.27) unambiguously, we also need to keep track of terms of the form  $(\partial_r X)(\partial_v^3 Y)$  where  $X$  and  $Y$  are boost invariant. These are the terms which will contribute to the  $k = 1$  sector of linearized  $E_{vv}^{\text{HD}}$ . Although, we are interested in finding out the  $k = 0$  zero-boost sector of the same, we need to track these specific  $k = 1$  terms (see (3.29)) as they will be needed to separate out the boost-invariant  $A$  and  $B$  in (3.27). In case of four-derivative theories we have only two possibilities for these terms as listed below

$$\tilde{T}_1 = h^{ij} h^{kl} (\partial_r h_{ij}) (\partial_v^3 h_{kl}), \quad \tilde{T}_2 = h^{ik} h^{jl} (\partial_r h_{ij}) (\partial_v^3 h_{kl}). \quad (3.30)$$

It is important to note that in this list of structures we have not counted  $h_{ij}$  and the determinant of  $h_{ij}$  as independent structures. All possible occurrences of these two pieces of data are automatically taken care of in the way we have listed our data. For example  $h_{ij}$



$T_1 = h^{ij}h^{kl}\nabla_i\nabla_j(\partial_v^2 h_{kl})$	$T_2 = h^{ik}h^{jl}\nabla_i\nabla_j(\partial_v^2 h_{kl})$
$T_3 = h^{ij}h^{kl}\omega_i\nabla_j(\partial_v^2 h_{kl})$	$T_4 = h^{ik}h^{jl}\omega_i\nabla_j(\partial_v^2 h_{kl})$
$T_5 = h^{ij}\partial_v^2\partial_r\partial_v h_{ij}$	$T_6 = h^{ij}\partial_v^2(\nabla_i\omega_j)$
$T_7 = X h^{ij}(\partial_v^2 h_{ij})$	$T_8 = h^{ij}h^{kl}(\nabla_i\omega_j)(\partial_v^2 h_{kl})$
$T_9 = h^{ik}h^{jl}(\nabla_i\omega_j)(\partial_v^2 h_{kl})$	$T_{10} = h^{ij}h^{kl}(\omega_i\omega_j)(\partial_v^2 h_{kl})$
$T_{11} = h^{ik}h^{jl}(\omega_i\omega_j)(\partial_v^2 h_{kl})$	$T_{12} = h^{ij}h^{kl}(\partial_r\partial_v h_{ij})(\partial_v^2 h_{kl})$
$T_{13} = h^{ik}h^{jl}(\partial_r\partial_v h_{ij})(\partial_v^2 h_{kl})$	$T_{14} = h^{ij}\omega_i\partial_v^2\omega_j$
$T_{15} = \partial_v^2 X$	$T_{16} = h^{ik}h^{jl}\mathcal{R}_{kl}\partial_v^2 h_{ij}$
$T_{17} = h^{ij}h^{kl}\mathcal{R}_{kl}\partial_v^2 h_{ij}$	
$\tilde{T}_1 = h^{ij}h^{kl}(\partial_r h_{ij})(\partial_v^3 h_{kl})$	$\tilde{T}_2 = h^{ik}h^{jl}(\partial_r h_{ij})(\partial_v^3 h_{kl})$

**Table 7.** Listing the seventeen  $T_i$ 's and two  $\tilde{T}_i$ 's, the possible 4-derivative scalar data with boost weight = 2. They are candidate terms that appear in  $E_{vv}^{\text{HD}}$  for 4-derivative theories of gravity. The seventeen  $T_i$  terms will contribute to  $k = 0$  sector, and the two  $\tilde{T}_i$  terms will contribute to  $k = 1$  sector of  $E_{vv}^{\text{HD}}$ .

could only occur in contraction of other indices and all possible contractions of indices are already counted in our listing. Finally, all the nineteen possible candidate terms (seventeen of the  $T_i$ 's and two of the  $\tilde{T}_i$ 's) to appear in  $E_{vv}^{\text{HD}}|_{\text{zero boost}}$ , are listed in table 7.

At this stage our claim is that the first term in (3.27), i.e., the term of the form  $\partial_v(A\partial_v B) \sim A\partial_v^2 B + \mathcal{O}(\epsilon^2)$ , for any four-derivative theory could always be expressed as a sum of these seventeen terms listed in table 7 with constant coefficients. Further, we claim that the contribution of the  $k = 1$  piece from the second term of (3.27) (written in the form of a sum over several  $k$  values) could also be expressed in terms of these seventeen structures plus two more, listed in equation (3.30)

$$E_{vv}^{\text{HD}} = -\sum_{i=1}^{17} a_i T_i - \sum_{i=1}^2 \tilde{a}_i \tilde{T}_i + \dots, \quad (3.31)$$

where  $\dots$  denote the terms that do not matter for the proof of the physical process version of the first law. The negative sign on the r.h.s. of (3.31) is chosen for convenience. The specific values of these seventeen  $a_i$  and two  $\tilde{a}_i$  coefficients appearing in (3.31), will of course vary from theory to theory. As we have mentioned before, the above classification of terms have been done keeping in mind the four derivative theories of gravity. The most general four derivative theory of pure gravity could have three more terms apart from the standard two derivative term in Einstein gravity. In table 8 we are listing the values of  $a_i$ 's and  $\tilde{a}_i$ 's for each of these three cases. These set of values of the  $a_i$  coefficients are obtained by comparing (3.31) with the explicit calculation of  $E_{vv}^{\text{HD}}$  for each of the three four derivative theories of gravity, which was performed in section 3.1.1, section 3.1.2 and section 3.1.3.

	Different theories ( $\mathcal{I}^{(i)} = \int d^d x \sqrt{-g} \mathcal{L}^{(i)}$ )	The calculated values of the coefficients $a_i$ and $\tilde{a}_i$ 's
1.	$\mathcal{L}^{(1)} = R^2$ (Ricci scalar squared)	$a_1 = 2, a_2 = -2, a_3 = 0, a_4 = 0,$ $a_5 = 4, a_6 = -4, a_7 = 1,$ $a_8 = -2, a_9 = 4, a_{10} = 3/2,$ $a_{11} = -3, a_{12} = 4,$ $a_{13} = -10, a_{14} = 6, a_{15} = 2,$ $a_{16} = 2, a_{17} = -1,$ $\tilde{a}_1 = 1, \tilde{a}_2 = -3.$
2.	$\mathcal{L}^{(2)} = R_{\mu\nu} R^{\mu\nu}$ (Ricci tensor squared)	$a_1 = 1/2, a_2 = -1, a_3 = -(1/2),$ $a_4 = 1, a_5 = 1, a_6 = 0,$ $a_7 = 1/2, a_8 = -(1/2), a_9 = 1,$ $a_{10} = 1/2, a_{11} = -1, a_{12} = 1,$ $a_{13} = -2, a_{14} = 2, a_{15} = 1,$ $a_{16} = 0, a_{17} = 0,$ $\tilde{a}_1 = 1/4, \tilde{a}_2 = -(1/2).$
3.	$\mathcal{L}^{(3)} = R_{\mu\nu\rho\sigma} R^{\mu\nu\rho\sigma}$ (Riemann tensor squared)	$a_1 = 0, a_2 = -2, a_3 = -2,$ $a_4 = 4, a_5 = 0, a_6 = 4, a_7 = 1,$ $a_8 = 0, a_9 = 0, a_{10} = 1/2,$ $a_{11} = -1, a_{12} = 0, a_{13} = 2,$ $a_{14} = 2, a_{15} = 2, a_{16} = 0, a_{17} = 0,$ $\tilde{a}_1 = 0, \tilde{a}_2 = 1..$

**Table 8.** Explicit calculation for each of the three theories produces these values of the coefficients  $a_i$ , appearing in  $E_{vv}^{\text{HD}} = -\sum_{i=1}^{17} a_i T_i - \sum_{i=1}^2 \tilde{a}_i \tilde{T}_i$ , for 4-derivative theories of gravity.

### 3.3 Constraints on the ‘zero boost terms’ in $E_{vv}^{\text{HD}}$

A very specific structure for the zero boost terms in  $E_{vv}^{\text{HD}}$  is predicted in (3.1). This structure does not follow automatically just from the boost transformation property, which we have used to classify terms in the previous subsection. Clearly, imposing (3.1) would impose further constraints on the seventeen coefficients mentioned above. In this subsection, we shall first find those constraints. We shall list the most general possible structure for  $\tilde{B}$ , defined in (3.1), which is a two-derivative scalar with vanishing boost weight. According to the terminology of the previous subsection it must be an ‘equilibrium data’. It turns out

Candidate terms for $\tilde{B}$	
Equilibrium data:	1. $h^{ij} (\partial_r \partial_v h_{ij})$ , 2. $h^{ij} \nabla_i \omega_j$ , 3. $h^{ij} \omega_i \omega_j$ , 4. $\mathcal{R}$ , 5. $X$ ,
Off-equilibrium data:	6. $h^{ij} h^{kl} (\partial_v h_{ij}) (\partial_r h_{kl})$ , 7. $h^{ik} h^{jl} (\partial_v h_{ij}) (\partial_r h_{kl})$

**Table 9.** Possible structures that can appear in  $\tilde{B}$ : each of them has two derivatives and boost weight= 0.

that  $\tilde{B}$  could have only five independent structures which are non-vanishing at equilibrium (see table 9).

Using linear combinations of the independent structures presented in table 9 we can now write down the most general structure of  $\tilde{B}$ , if it exists, as follows

$$\tilde{B} = A_1 h^{ij} \partial_r \partial_v h_{ij} + A_2 h^{ij} \nabla_i \omega_j + A_3 h^{ij} \omega_i \omega_j + A_4 X + A_5 \mathcal{R}. \quad (3.32)$$

The first term in  $\tilde{B}$  needs a special attention. This is the term whose contribution to  $E_{vv}^{\text{HD}}$  could get mixed with some the  $k = 1$  term (see equation (3.27) and the discussion after that). To see this more explicitly, let us write down the contribution to  $E_{vv}^{\text{HD}}$  coming from the term  $\tilde{B} = A_1 h^{ij} \partial_r \partial_v h_{ij}$

$$\begin{aligned} \partial_v \left( \frac{1}{\sqrt{h}} \partial_v \left( \sqrt{h} \tilde{B} \right) \right) &\sim A_1 \partial_v \left( \frac{1}{\sqrt{h}} \partial_v \left( \sqrt{h} h^{ij} \partial_r \partial_v h_{ij} \right) \right) \\ &= A_1 \left( T_5 + \frac{T_{12}}{2} - T_{13} \right). \end{aligned} \quad (3.33)$$

It can be easily checked that the terms  $T_{12}$  and  $T_{13}$  could also be generated as  $k = 1$  terms in  $E_{vv}^{\text{HD}}$  from the following two off-equilibrium candidates for  $\tilde{B}$  (see the list of off-equilibrium data in table 9)

$$(i) h^{ij} h^{kl} (\partial_v h_{ij}) (\partial_r h_{kl}), \quad (ii) h^{ik} h^{jl} (\partial_v h_{ij}) (\partial_r h_{kl}).$$

We assume that in  $E_{vv}^{\text{HD}}$ , these two terms mentioned above contribute with coefficients  $A_6$  and  $A_7$  respectively as written below<sup>21</sup>

$$E_{vv}^{\text{HD}}|_{k=1} = -A_6 \left( 2 T_{12} + \tilde{T}_1 \right) - A_7 \left( 2 T_{13} + \tilde{T}_2 \right). \quad (3.34)$$

<sup>21</sup>To obtain the expressions in (3.36) we have used the following relations

$$\begin{aligned} \partial_v \left( \frac{1}{\sqrt{h}} \partial_v \left( \sqrt{h} h^{ij} h^{kl} (\partial_v h_{ij}) (\partial_r h_{kl}) \right) \right) &= 2 T_{12} + \tilde{T}_1, \\ \partial_v \left( \frac{1}{\sqrt{h}} \partial_v \left( \sqrt{h} h^{ik} h^{jl} (\partial_v h_{ij}) (\partial_r h_{kl}) \right) \right) &= 2 T_{13} + \tilde{T}_2. \end{aligned}$$

From the above equation it is clear that  $A_6$  and  $A_7$  could be simply fixed by comparing the coefficients of  $\tilde{T}_1$  and  $\tilde{T}_2$  respectively in the  $k = 1$  sector of (3.31) and (3.34),

$$A_6 = \tilde{a}_1, \quad A_7 = \tilde{a}_2. \tag{3.35}$$

We have now extracted out the  $k = 1$  part of  $E_{vv}^{\text{HD}}$  in (3.34) which has the form of  $A \partial_v^2 B$ , as desired from (3.27). Next, we subtract off (3.34) from (3.31) and obtain the part of  $E_{vv}^{\text{HD}}$  that is entirely generated from zero boost sector. This could be written as

$$E_{vv}^{\text{HD}}|_{\text{zero boost}} = - \sum_{i=1}^{11} a_i T_i - (a_{12} - 2\tilde{a}_1) T_{12} - (a_{13} - 2\tilde{a}_2) T_{13} - \sum_{i=14}^{17} a_i T_i. \tag{3.36}$$

At this point, it is important to note that although in (3.31) there were nineteen terms to begin with, the zero boost sector  $E_{vv}^{\text{HD}}|_{k=0}$  is constructed out of seventeen terms  $T_i$ 's appearing on the r.h.s. of (3.36). We, therefore, have to deal with seventeen coefficients as well. The easiest way to understand this is by realizing that the coefficients  $\tilde{a}_1$  and  $\tilde{a}_2$  do not count as additional ones since they will always appear in the combination  $(a_{12} - 2\tilde{a}_1)$  and  $(a_{13} - 2\tilde{a}_2)$  respectively.<sup>22</sup>

On the other hand, from (3.32) we know that the number of free coefficients in  $\tilde{B}$  thus turns out to be five, namely the  $A_i$ 's (for  $i = 1, \dots, 5$ ). As of now the  $A_i$ 's are free coefficients and we want to solve them in terms of the  $a_i$ 's appearing in (3.36). To do that, we first substitute  $\tilde{B}$  from (3.32) in (3.1) and write it in terms of the basis of  $T_i$  structures, as listed in table 7 and obtain the following

$$E_{vv}^{\text{HD}}|_{\text{zero boost}} \sim - \partial_v \left( \frac{1}{\sqrt{h}} \partial_v (\sqrt{h} \tilde{B}) \right) = - \left[ A_1 \left( T_5 + \frac{T_{12}}{2} - T_{13} \right) + A_2 \left( T_6 + \frac{T_8}{2} - T_9 \right) + A_3 \left( 2T_{14} + \frac{T_{10}}{2} - T_{11} \right) + A_4 \left( T_{15} + \frac{T_7}{2} \right) - A_5 \left( T_{16} - \frac{T_{17}}{2} + T_1 - T_2 \right) \right]. \tag{3.37}$$

---

<sup>22</sup>In what follows, whenever we refer to the seventeen coefficients  $a_i$ 's, it will be implied that there are actually nineteen coefficients (the  $a_i$ 's and the  $\tilde{a}_i$ 's) but the two coefficients  $\tilde{a}_1$  and  $\tilde{a}_2$  will always appear being paired with  $a_{12}$  and  $a_{13}$  respectively, see (3.36), and hence the independent coefficients will be counted as seventeen.

In deriving (3.37), we have used the following relations

$$\begin{aligned}
 \partial_v \left( \frac{1}{\sqrt{h}} \partial_v \left( \sqrt{h} h^{ij} \partial_r \partial_v h_{ij} \right) \right) &= T_5 + \frac{T_{12}}{2} - T_{13}, \\
 \partial_v \left( \frac{1}{\sqrt{h}} \partial_v \left( \sqrt{h} h^{ij} \nabla_i \omega_j \right) \right) &= T_6 + \frac{T_8}{2} - T_9, \\
 \partial_v \left( \frac{1}{\sqrt{h}} \partial_v \left( \sqrt{h} h^{ij} \omega_i \omega_j \right) \right) &= 2T_{14} + \frac{T_{10}}{2} - T_{11}, \\
 \partial_v \left( \frac{1}{\sqrt{h}} \partial_v \left( \sqrt{h} X \right) \right) &= T_{15} + \frac{T_7}{2}, \\
 \partial_v \left( \frac{1}{\sqrt{h}} \partial_v \left( \sqrt{h} \mathcal{R} \right) \right) &= - \left( T_{16} - \frac{T_{17}}{2} + T_1 - T_2 \right).
 \end{aligned} \tag{3.38}$$

We now compare (3.37) with (3.36) and equate the coefficients of  $T_i$ 's on both sides, which gives us seventeen relations between the  $A_i$  ( $i = 1, \dots, 5$ ) and  $a_j$  ( $j = 1, \dots, 17$ ),  $\tilde{a}_i$  ( $i = 1, 2$ ).<sup>23</sup> We can solve the five  $A_i$ 's in terms of the  $a_i$ 's and then we will be left with twelve constraints on the coefficients  $a_i$ 's which ensure the consistency of (3.1). These twelve constraints on  $a_i$ 's are listed below,

$$\begin{aligned}
 a_1 &= a_{16}, & 2a_{10} &= -a_{11}, & a_{14} &= 4a_{10}, & a_{15} &= 2a_7, & 2a_{17} &= -a_1, \\
 a_2 &= -a_1, & a_3 &= 0, & a_4 &= 0, & a_6 &= 2a_8, & 2a_8 &= -a_9, \\
 2(a_{12} - 2\tilde{a}_1) &= -(a_{13} - 2\tilde{a}_2), & a_5 &= 2(a_{12} - 2\tilde{a}_1).
 \end{aligned} \tag{3.39}$$

Finally, we would like to check whether the  $a_i$ 's as given in table 8 satisfy the constraints given in (3.39). Remember that in the previous subsection, we have already calculated the allowed values of the  $a_i$ 's for each of the three different 4-derivative theory of gravity, see table 8. Upon inspection, we can convince ourselves that for Ricci scalar squared theory the constraints in (3.39) are satisfied, where as for both of the other two four derivative theories of gravity, namely the Ricci tensor squared and the Riemann tensor squared theories, the constraints in (3.39) are simply not satisfied. Therefore, we convince ourselves that the constraints obtained in (3.39) are not correct, as they are not satisfied by the results obtained by explicit calculation of  $E_{vv}^{\text{HD}}$  which is the content of table 8 for the most general four derivative theory of gravity. We should keep this in mind that these constraints were derived from demanding the consistency of (3.1). As a result, we are led to the conclusion that the general structure of  $E_{vv}^{\text{HD}}$  in the zero boost sector, as predicted in (3.1), is not generically true for the most general four derivative theory of gravity.

### 3.4 The general strategy for constructing the entropy current maintaining the boost symmetry

From the analysis of the previous subsection, we have established the fact that the zero boost terms in  $E_{vv}^{\text{HD}}$  does not always follow the structure predicted in (3.1). Motivated by this observation, in this subsection our goal will therefore be to explore what are the

---

<sup>23</sup>As we have mentioned before, we have seventeen coefficients on the r.h.s. of (3.36), and not nineteen, because the coefficients  $(a_{12} - 2\tilde{a}_1)$  and  $(a_{13} - 2\tilde{a}_2)$  always comes in this particular combination.

further structures, if any, that we could allow for the zero boost terms in  $E_{vv}^{\text{HD}}$  without affecting the proofs for the physical version of the first law and second law.

As we have explained before, the first law is a statement about the total change in the thermodynamic parameters like entropy, energy etc., characterizing two nearby equilibrium solutions connected by dynamics. Hence its formulation always involves an integration over all space and therefore is usually insensitive to any boundary terms. The same is true for black hole mechanics. The total change in entropy as described in (2.31) has an integration over the spatial slices of the horizon. If the horizon is compact, this integration would be insensitive to any boundary term that appears in  $E_{vv}^{\text{HD}}$ . It follows that the zero boost term in  $E_{vv}^{\text{HD}}$ , in addition to the term already mentioned in (3.1), could also have a structure of the form

$$E_{vv}^{\text{HD}}|_{\text{zero boost}} \sim -\partial_v (\nabla_i J^i) = -\partial_v \left( \frac{1}{\sqrt{h}} \partial_i (\sqrt{h} J^i) \right) \quad (3.40)$$

where  $J^i$  is some spatial current with boost weight 1 (i.e., it must contain an explicit  $\partial_v$  that could not be paired up with any  $\partial_r$ ). On compact horizons such a term would clearly integrate to zero and therefore will not contribute to the total change in entropy (see the derivation of (2.31)).

It is worth noting that the compatibility with the first law also allows a term, generically of the form  $\nabla_i Y^i$  in  $E_{vv}^{\text{HD}}$  where  $Y^i$  is some arbitrary vector quantity, i.e. a spatial current with boost weight equal to 2. However, the manipulation that follows from (2.16) shows that working up to linear order of amplitude perturbations we could always re-arrange the terms in  $E_{vv}^{\text{HD}}$  (including the possible  $\nabla_i Y^i$  term) in a form where there is an overall  $\partial_v$  outside.<sup>24</sup> It is important to stress that although the first law itself does not require this rearrangement as in (3.41), it is a must to proceed towards an argument for the second law. Therefore, in our classification, we shall not consider such terms for which this rearrangement is not true. This, in particular, allows us not to consider the term that we have just mentioned above in (3.40) as a possible term in  $E_{vv}^{\text{HD}}$ .

Combining equations (3.1) and (3.40), it follows that, both the first and the second law would be satisfied, at least at the linear order in amplitude of time-dependent perturbations, provided the zero boost terms in  $E_{vv}^{\text{HD}}$  has the following form

$$E_{vv}^{\text{HD}}|_{\text{zero boost}} \sim -\partial_v \left( \frac{1}{\sqrt{h}} \partial_v (\sqrt{h} \tilde{B}) + \nabla_i J^i \right). \quad (3.42)$$

Interestingly, we should note that on the r.h.s. of (3.42), the term inside the parenthesis (i.e. ignoring the overall  $\partial_v$ ), looks exactly like the divergence of a ‘four-current’, let us call it  $S^A$ , such that it’s  $v$  and  $i$  components are respectively given by

$$\begin{aligned} E_{vv}^{\text{HD}}|_{\text{zero boost}} &\sim -\partial_v (\nabla_A S^A), \\ \text{such that } S_{(k=0)}^v &= \tilde{B}, & S_{(k=0)}^i &= J^i, \end{aligned} \quad (3.43)$$

<sup>24</sup>This can be schematically presented as

$$E_{vv}^{\text{HD}}|_{\text{zero boost}} \sim \nabla_i Y^i \sim \partial_v (\nabla_i \tilde{Y}^i) + \mathcal{O}(\epsilon^2), \quad (3.41)$$

where  $\tilde{Y}^i$  is some spatial vector with boost weight equal to one, since one  $\partial_v$  is extracted from  $Y^i$  which has boost weight equal to one.

Candidate terms for $J_i$	
Off-equilibrium data:	<ol style="list-style-type: none"> <li>1. <math>\partial_v \omega_i</math></li> <li>2. <math>h^{jk} \nabla_j (\partial_v h_{ki})</math></li> <li>3. <math>h^{jk} \nabla_i (\partial_v h_{jk})</math></li> <li>4. <math>h^{jk} \omega_j (\partial_v h_{ki})</math></li> <li>5. <math>h^{jk} \omega_i (\partial_v h_{jk})</math></li> </ol>

**Table 10.** Possible structures that can appear in  $J_i$ : each one of them has two derivatives and boost weight= 1.

where, the index  $A = v, x^i$  and we have also used  $k = 0$  as a subscript in  $S_{(k=0)}^A$  to denote the fact that we are only looking at the zero boost terms in  $E_{vv}^{\text{HD}}$ .

Next, we would like to see how the seventeen structures, listed in table 7 in section 3.2, should combine so that the zero boost terms in  $E_{vv}^{\text{HD}}$  could be recast in the form of (3.42). In other words, if the form of  $E_{vv}^{\text{HD}}$  as proposed in (3.42) is correct, we will be using that to derive the constraints that the seventeen coefficients  $a_i$  should satisfy. As it appears in (3.42),  $\tilde{B}$  is a boost invariant scalar data and  $J^i$  is a vector data with boost weight one. In the previous subsection, we have already argued the most general structure of  $\tilde{B}$  in (3.32). Now  $J^i$  is an off-equilibrium data and from the counting of boost-weight we could see it must have exactly one  $\partial_v$  derivative, which is not paired with an  $\partial_r$  provided we are considering them as operators acting on the three basic building blocks, namely  $h_{ij}$ ,  $\omega_i$  and  $X$ . Taking all these facts into account, we could construct the five possible structures for a candidate term in  $J_i$ , as listed in table 10 below.

It is now straightforward to write down the most general form of  $J^i$  using linear combinations of the structures written in table 10

$$\begin{aligned}
 J^i = & B_1 h^{ij} \partial_v \omega_j + B_2 h^{il} h^{jk} \nabla_j (\partial_v h_{kl}) + B_3 h^{il} h^{jk} \nabla_l (\partial_v h_{jk}) \\
 & + B_4 h^{il} h^{jk} \omega_j (\partial_v h_{kl}) + B_5 h^{il} h^{jk} \omega_l (\partial_v h_{jk}),
 \end{aligned}
 \tag{3.44}$$

where the coefficients  $B_i$  for  $i = 1, \dots, 5$ , are, as of now, arbitrary constant coefficients. Our aim will now be to fix them in terms of the coefficients  $a_i$ 's ( $i = 1, \dots, 17$ ), just like the coefficients  $A_i$ 's, appearing in (3.32), were fixed in the previous subsection. To achieve this we will calculate the second term on the r.h.s. of (3.42), with  $J^i$  being substituted from (3.44). We express the resulting expression in terms of the  $T_i$ 's, listed in table 7, and obtain the following relation

$$\begin{aligned}
 E_{vv}^{\text{HD}} \Big|_{\text{zero boost}}^{(J^i \text{ part})} \sim -\partial_v (\nabla_i J^i) = & - \left[ B_1 \left( T_6 + T_4 - \frac{T_3}{2} \right) + B_2 T_2 \right. \\
 & \left. + B_3 T_1 + B_4 (T_4 + T_9) + B_5 (T_3 + T_8) \right].
 \end{aligned}
 \tag{3.45}$$

In deriving (3.45) we have used the following relations

$$\begin{aligned}
 \partial_v [\nabla_i (h^{ij} \partial_v \omega_j)] &= T_6 + T_4 - (T_3/2), & \partial_v [\nabla_i (h^{il} h^{jk} \nabla_j (\partial_v h_{kl}))] &= T_2, \\
 \partial_v [\nabla_i (h^{il} h^{jk} \nabla_l (\partial_v h_{jk}))] &= T_1, & \partial_v [\nabla_i (h^{il} h^{jk} \omega_j (\partial_v h_{kl}))] &= T_4 + T_9, \\
 \partial_v [\nabla_i (h^{il} h^{jk} \omega_l (\partial_v h_{jk}))] &= T_3 + T_8.
 \end{aligned} \tag{3.46}$$

Once we have obtained (3.37) and (3.45), we shall combine them to obtain a complete expression for the zero boost part (i.e.  $k = 0$ ) of  $E_{vv}^{\text{HD}}$  in terms of the  $T_i$ 's as follows

$$\begin{aligned}
 E_{vv}^{\text{HD}}|_{\text{zero boost}} &= - \left[ A_1 \left( T_5 + \frac{T_{12}}{2} - T_{13} \right) + A_2 \left( T_6 + \frac{T_8}{2} - T_9 \right) \right. \\
 &\quad + A_3 \left( 2T_{14} + \frac{T_{10}}{2} - T_{11} \right) + A_4 \left( T_{15} + \frac{T_7}{2} \right) \\
 &\quad - A_5 \left( T_{16} - \frac{T_{17}}{2} + T_1 - T_2 \right) + B_1 \left( T_6 + T_4 - \frac{T_3}{2} \right) + B_2 T_2 + B_3 T_1 \\
 &\quad \left. + B_4 (T_4 + T_9) + B_5 (T_3 + T_8) \right].
 \end{aligned} \tag{3.47}$$

It is obvious from the r.h.s. of (3.47) above that we still have ten undetermined coefficients, five of the  $A_i$ 's and five of the  $B_i$ 's. We therefore conclude that, if we want the first term in (3.27), to have a form such that it is compatible with the physical process version of the first law, then it can have twelve independent coefficients ( $A_i, B_i$ ) for any four derivative theories of gravity. On the other hand, just from the consideration of boost symmetry, a total of seventeen terms are allowed in  $E_{vv}^{\text{HD}}$ , see (3.31). Clearly, even after the inclusion of the spatial current in (3.42), the compatibility with the physical version of the first law would imply some constraints between  $a_i$ 's (though it would certainly be less in number than what we have derived in the previous subsection). A naive counting suggests that there must be  $(17 - 10) = 7$  relations among the seventeen possible coefficients  $a_i$ . However, as it turns out, there is a redundancy in our counting of independent structures that could appear in the expression of entropy density ( $\tilde{B}$ ) and spatial entropy current ( $J^i$ ). In other words, not all of the ten  $A_i, B_i$ 's are independently and one of them, the term with  $A_2$  as coefficient, can be absorbed into others by redefining some of the  $B_i$  coefficients. It is easy to check that if we redefine the coefficients  $B_1, B_4$  and  $B_5$  in the following way,

$$\hat{B}_1 = B_1 + A_2, \quad \hat{B}_4 = B_4 - A_2, \quad \hat{B}_5 = B_5 + \frac{A_2}{2}, \tag{3.48}$$

the term with coefficient  $A_2$  in (3.47) disappears and we are left with

$$\begin{aligned}
 E_{vv}^{\text{HD}}|_{\text{zero boost}} &= - \left[ A_1 \left( T_5 + \frac{T_{12}}{2} - T_{13} \right) + A_3 \left( 2T_{14} + \frac{T_{10}}{2} - T_{11} \right) \right. \\
 &\quad + A_4 \left( T_{15} + \frac{T_7}{2} \right) - A_5 \left( T_{16} - \frac{T_{17}}{2} + T_1 - T_2 \right) \\
 &\quad + \hat{B}_1 \left( T_6 + T_4 - \frac{T_3}{2} \right) + B_2 T_2 + B_3 T_1 + \hat{B}_4 (T_4 + T_9) \\
 &\quad \left. + \hat{B}_5 (T_3 + T_8) \right].
 \end{aligned} \tag{3.49}$$



As the independent terms on the r.h.s. of (3.49) has now been reduced to nine, we should obtain eight relations among the coefficients  $a_i$ , which are given by

$$\begin{aligned} a_4 &= a_6 + a_9, & a_3 &= -\frac{a_6}{2} + a_8, & a_{16} &= -2a_{17}, & a_{15} &= 2a_7, \\ a_{11} &= -\frac{a_{14}}{2}, & a_{10} &= \frac{a_{14}}{4}, & 2(a_{12} - 2\tilde{a}_1) &= a_5, & a_{13} - 2\tilde{a}_2 &= a_5. \end{aligned} \quad (3.50)$$

Furthermore, once the  $a_i$ 's satisfy the identities given in (3.50), we can solve the  $A_i$ 's and  $B_i$ 's in terms of the  $a_i$ 's, as given below<sup>25</sup>

$$\begin{aligned} A_1 &= a_5, & A_3 &= \frac{a_{14}}{2}, & A_4 &= 2a_7, & A_5 &= 2a_{17}, \\ A_2 &= \text{free/undetermined}, \\ B_1 &= a_6 - A_2, & B_2 &= a_2 - 2a_{17}, & B_3 &= a_1 + 2a_{17}, & B_4 &= a_9 + A_2, \\ B_5 &= a_8 - \frac{A_2}{2}. \end{aligned} \quad (3.51)$$

It is worth mentioning that in deriving the identities in (3.50) and the solutions in (3.51) we have not assumed any particular form of the four derivative gravity Lagrangian. In other words these relations are true for any four derivative theory of gravity.

Once we have obtained the coefficients  $A_i$  and  $B_i$ , one can readily derive the entropy density  $\tilde{B}$  and the entropy currents  $J^i$  in terms of the coefficients  $a_i$ . Since specific values for the set of coefficients  $a_i$  corresponds to specific four derivative theories of gravity, (see table 8), we can substitute them for  $a_i$ 's in (3.51) to obtain the specific values of  $A_i$  and  $B_i$  for each of the three individual four derivative theories of gravity. We present them in table 11.

Finally, we conclude this sub-section with the following remarks:

- **More details on the redundancy in the parameter  $A_2$ :** we have already mentioned before that our analysis in this subsection to classify possible candidate terms in the zero boost sector of  $E_{vv}^{\text{HD}}$  solely based on boost symmetry, cannot fix the coefficient  $A_2$  in (3.47). As a result it remained undetermined in (3.51). We have also seen that this redundancy in fixing  $A_2$  is actually related to a proper count of the independent data in  $\tilde{B}$  and  $J^i$ .

In order to make it explicitly manifest, let us now consider the specific terms written below and their combinations as candidates for  $\tilde{B}$  and  $J^i$ <sup>26</sup>

$$\tilde{B}_{(*)} = h^{ij}\nabla_i\omega_j, \quad J_{(*)}^i = -h^{ij}\partial_v\omega_j + h^{il}h^{jk}\omega_j\partial_v h_{kl} - \frac{1}{2}h^{il}h^{jk}\omega_l\partial_v h_{jk},$$

and with this choices it can be shown that

$$\frac{1}{\sqrt{h}}\partial_v\left(\sqrt{h}\tilde{B}_{(*)}\right) + \nabla_i J_{(*)}^i = 0. \quad (3.52)$$

<sup>25</sup>In (3.51) we are still writing in terms of the coefficients  $B_1, B_2, B_3$ , instead of writing them in terms of the redefined  $\tilde{B}_1, \tilde{B}_2, \tilde{B}_3$ . This makes the appearance of the undetermined coefficient  $A_2$  explicit, and is just a matter of convenient choice for us.

<sup>26</sup>Note that this  $\tilde{B}_{(*)}$  appears in the expression of  $\tilde{B}$  in (3.32) with the coefficient  $A_2$ .

	Different theories ( $\mathcal{I}^{(i)} = \int d^d x \sqrt{-g} \mathcal{L}^{(i)}$ )	Values of the coefficients $A_i, B_i$
1	$\mathcal{L}^{(1)} = R^2$ (Ricci scalar squared)	$A_1 = 4, A_2 = \text{undetermined},$ $A_3 = 3, A_4 = 4, A_5 = -2,$ $B_1 = -4 - A_2, B_2 = 0, B_3 = 0,$ $B_4 = 4 + A_2, B_5 = -2 - (A_2/2).$
2	$\mathcal{L}^{(2)} = R_{\mu\nu}R^{\mu\nu}$ (Ricci tensor squared)	$A_1 = 1, A_2 = \text{undetermined},$ $A_3 = 1, A_4 = 2, A_5 = 0,$ $B_1 = -A_2, B_2 = -1, B_3 = 1/2,$ $B_4 = 1 + A_2, B_5 = -(1/2) - (A_2/2).$
3	$\mathcal{L}^{(3)} = R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma}$ (Riemann tensor squared)	$A_1 = 0, A_2 = \text{undetermined},$ $A_3 = 1, A_4 = 4, A_5 = 0,$ $B_1 = 4 - A_2, B_2 = -2, B_3 = 0,$ $B_4 = A_2, B_5 = -(A_2/2).$

**Table 11.**  $A_i, B_i$ 's for different 4-derivative theories of gravity.

The interesting thing to note about the combination written in (3.52) is that it identically vanishes without any use of the gravity equations of motion and therefore we could add the  $v$ -derivative of this combination (so that it has the appropriate boost weight= 2) to any expression for  $E_{vv}^{\text{HD}}$ , without affecting the equation of motion and dynamics. Because of this, among the twelve terms that appeared on the r.h.s. of (3.47) above we could hope to fix only eleven of them by comparing with the  $E_{vv}^{\text{HD}}$  of a given four derivative theory, (3.31).<sup>27</sup> Also, for the same reason, we have seen that in each of the three cases tabulated in table 11, the coefficient  $A_2$  could not be fixed as it could combine with few spatial currents to give vanishing contribution to  $E_{vv}^{\text{HD}}$ .

- **The redundancy in  $A_2$  is fixed by matching the equilibrium limit of  $\tilde{B}$  with the equilibrium Wald entropy density:** having realized the fact that only boost symmetry alone can not fix the coefficient  $A_2$ , let us now focus on the implications

<sup>27</sup>Actually, we can make use of this redundancy to reorganize (3.32) and (3.44) with the following redefinition of  $\tilde{B}$  and  $J^i$

$$\tilde{B} \rightarrow \tilde{B}; \quad J^i \rightarrow J^i + \alpha_* A_2 J_{(*)}^i, \tag{3.53}$$

where  $\alpha_*$  being a tunable free parameter and thus enabling us to fix the value of the coefficient  $A_2$  to any specific number. In particular by making the choice of  $\alpha_* = 1$ , we can even make the coefficient  $A_2$  not contributing to (3.47), as in that case, as  $A_2$  disappears from  $E_{vv}^{\text{HD}}$ .

of this redundancy in the coefficient  $A_2$  beyond boost symmetry and try to explore if there is any other principle that can fix it. Looking at the table 9 and (3.32), we remind ourselves that, by construction, the scalar structures appearing in  $\tilde{B}$  does not vanish when evaluated on a stationary solution. Most importantly, the term that appears in  $\tilde{B}$ , (3.32), with coefficient  $A_2$  is generically non-zero in the equilibrium limit. Therefore the redundancy in the coefficient  $A_2$  discussed in detail above, implies that possible different choices of  $A_2$  would amount to having different expressions for the equilibrium entropy density,  $s_w$ , of the same configuration. Though the difference does not persist in the expression of total entropy  $S_W$ , since this density turns out to be a total derivative term:  $(\nabla \cdot \omega)$  in this case.

Motivated by the arguments given above and based on general grounds, we should, therefore, also require that once the equilibrium limit is considered, the entropy density  $\tilde{B}$  in (3.32), should reproduce the appropriate Wald entropy density. This should be satisfied by the  $\tilde{B}$  apart from being constructed following the boost symmetry. As we will see now, at least for the cases that we are studying in this note, this additional requirement uniquely fixes the ambiguity related to the coefficient  $A_2$ . Thus, the important point to note here is that the Wald's formula (3.6) picks up a very specific value for  $A_2$  for every case that we have discussed here, and in some sense fixes this ambiguity which clearly could not be fixed just by imposing first or second law of thermodynamics even in its ultra-local version. For example, in  $R^2$  theory, once we demand matching with (3.6),  $A_2$  gets fixed to a specific numerical value  $A_2 = -4$ , implying that there is no spatial current, which is actually consistent with what we have found in subsection 3.1.<sup>28</sup>

The consistency with Wald's formula in the equilibrium limit, forces the entropy density  $\tilde{B}$ , that we have obtained in this subsection, to reduce to the stationary limit of  $s_{Wald}^{HD}$  (see (3.7)), which we derived in subsection 3.1, up to the ambiguity of  $A_2$ . More precisely, if we take the expressions of  $s_{Wald}^{HD}$  as computed in subsections 3.1.1, section 3.1.2 and section 3.1.3 and simply remove the terms that would vanish in stationary situations (for example, a term like  $\mathcal{K}\bar{\mathcal{K}}$  would be ignored), the resultant expressions should exactly match with the corresponding  $\tilde{B}$ 's derived in this subsection with a specific choice of the coefficient  $A_2$  for every case.<sup>29</sup> It turns out that they indeed match provided we choose the coefficient  $A_2$  to be as follows:

1. for  $R^2$  theory:  $A_2 = -4$ ,
2. for  $R_{\mu\nu}R^{\mu\nu}$  theory:  $A_2 = -1$ ,
3. for  $R_{\mu\nu\alpha\beta}R^{\mu\nu\alpha\beta}$  theory:  $A_2 = 0$ .

---

<sup>28</sup>A first glance at the non-zero values of the coefficients  $B_1$ ,  $B_4$  and  $B_5$  for the  $R^2$  theory in table 11 might naively suggest that there is a non-zero current for the  $R^2$  theory. However once we make the choice of  $A_2 = -4$  in order to match with the equilibrium Wald entropy density  $s_w^{HD}$ , it can be verified that there is no spatial current in this case, but a finite non-equilibrium correction  $s_{cor}$  to  $s_w^{HD}$ , see (3.5).

<sup>29</sup>Though a mismatch at this stage would have been a serious contradiction with the existing literature and Wald's formalism, we still do not have any abstract proof for it, applicable to any higher derivative theories of gravity. According to our understanding, this would essentially amount to showing a step by step equivalence between the proof of physical version of the first law and the Wald formalism. We could not find it in literature and leave it for future work.

	Different theories ( $\mathcal{I}^{(i)} = \int d^d x \sqrt{-g} \mathcal{L}^{(i)}$ )	Expressions for $\tilde{B}$ , $J^i$
1.	$\mathcal{L}^{(1)} = R^2$	$\tilde{B} = 4 h^{ij} \partial_r \partial_v h_{ij} - 4 h^{ij} \nabla_i \omega_j$ $+ 3 h^{ij} \omega_i \omega_j + 4 X - 2 \mathcal{R}$
		$J^i = 0$
2.	$\mathcal{L}^{(2)} = R_{\mu\nu} R^{\mu\nu}$	$\tilde{B} = h^{ij} \partial_r \partial_v h_{ij} + h^{ij} \omega_i \omega_j + 2 X$
		$J^i = -h^{il} h^{jk} \nabla_j (\partial_v h_{kl})$ $+ \frac{1}{2} h^{il} h^{jk} \nabla_l (\partial_v h_{jk}) + h^{il} h^{jk} \omega_j (\partial_v h_{kl})$ $- \frac{1}{2} h^{il} h^{jk} \omega_l (\partial_v h_{jk})$
3.	$\mathcal{L}^{(3)} = R_{\mu\nu\rho\sigma} R^{\mu\nu\rho\sigma}$	$\tilde{B} = h^{ij} \omega_i \omega_j + 4 X$
		$J^i = 4 h^{ij} \partial_v \omega_j - 2 h^{il} h^{jk} \nabla_j (\partial_v h_{kl})$

**Table 12.**  $\tilde{B}$  and  $J^i$ 's for different four derivative theory of gravity. While writing the expressions we have used the values for the coefficient  $A_2$  in each of the three cases as following: (i) for  $R^2$  theory:  $A_2 = -4$ , (ii) for  $R_{\mu\nu} R^{\mu\nu}$  theory:  $A_2 = 0$ , and (iii) for  $R_{\mu\nu\alpha\beta} R^{\mu\nu\alpha\beta}$  theory:  $A_2 = 0$ .

This matching serves as a consistency check for our results. Therefore, once we use the values of the coefficient  $A_2$  for different cases, as written above, in table 11 and further using (3.32) and (3.44) the specific expressions for  $\tilde{B}$  and  $J^i$  can be derived as listed in table 12.

- **Constraints on  $a_i$ 's satisfied:** in section 3.2 we computed the specific values of the coefficients  $a_i$ 's and tabulated them in the table 8 for three different four derivative theories of gravity. It is now straightforward to check that the constraints derived in (3.50) are indeed satisfied by all of the four derivative theories of gravity. In other words, the physical process version of the first law holds for all of these theories once we allow for the spatial current term in  $E_{vv}^{\text{HD}}$ , (3.42).

### 3.5 Einstein-Gauss-Bonnet gravity in $d \geq (4 + 1)$

The Einstein-Gauss-Bonnet theory has been extensively studied as a prototype of higher derivative corrections to Einstein's gravity and has been accorded significant importance in the relevant literature. It is also a theory with 4-derivative correction to Einstein's gravity, where the 4-derivative term is a specific combination of the three terms, that has been discussed in section 3.1.1, section 3.1.2 and section 3.1.3. This linear combination is such that, although the Einstein-Hilbert action has 4-derivatives corrections, the equations of motion that follow from it, only have two derivatives on the metric, just like Einstein equations. Since the Gauss-Bonnet term is simply a specific linear combination of the four derivative terms discussed in the previous sections, the analysis for the Einstein-Gauss-Bonnet theory can be done quickly by considering the same linear combination of the results we obtained before. In this section, we state our results explicitly for this theory.

The Einstein-Gauss-Bonnet theory is non-trivial in any dimensions greater than  $3 + 1$ . In  $3 + 1$  dimension the 4-derivative term is a total derivative (and is, therefore, a topological

surface term). In lower dimensions, it vanishes as an identity. Let us first consider this theory in space-time dimensions  $d \geq 4 + 1$ ; we shall discuss the special case of  $d = 3 + 1$  in the next subsection.

The action for Einstein-Gauss-Bonnet theory is given by

$$I = \int d^d x \sqrt{-g} (R + a_{gb} (R^2 - 4R_{\mu\nu}R^{\mu\nu} + R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma})), \quad (3.54)$$

where  $a_{gb}$  is a constant Gauss-Bonnet parameter. The corresponding equations of motion are

$$E_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R + E_{\mu\nu}^{HD} = 0, \quad (3.55)$$

where

$$E_{\mu\nu}^{HD} = a_{gb} \left( 2RR_{\mu\nu} - 4R^{\alpha\beta}R_{\mu\alpha\nu\beta} - 4R_{\mu}^{\alpha}R_{\nu\alpha} + 2R_{\mu}^{\alpha\beta\sigma}R_{\nu\alpha\beta\sigma} - \frac{1}{2}g_{\mu\nu}(R^2 - 4R_{\alpha\beta}R^{\alpha\beta} + R_{\alpha\beta\gamma\rho}R^{\alpha\beta\gamma\rho}) \right). \quad (3.56)$$

The explicit vv-component of the equations of motion is

$$\begin{aligned} E_{vv} &= R_{vv} + E_{vv}^{HD} = 0, \\ E_{vv}^{HD} &= a_{gb} \left( 2RR_{vv} - 4R^{\alpha\beta}R_{v\alpha v\beta} - 4R_v^{\alpha}R_{v\alpha} + 2R_v^{\alpha\beta\sigma}R_{v\alpha\beta\sigma} \right) \end{aligned} \quad (3.57)$$

By explicitly computing  $E_{vv}^{HD}$  in terms of the metric components (2.1) and their derivatives, it is possible to rewrite  $E_{vv}$  for the Einstein-Gauss-Bonnet theory into the form (3.2). Subsequently, we can read off the entropy current from it and we have

$$\begin{aligned} J^v &= (1 + 2a_{gb}(\mathcal{R} - 2\bar{K}_{AB}K^{AB} + 2K\bar{K})) \\ J^i &= -4a_{gb}\nabla^j(Kh_{ij} - K_{ij}) \end{aligned} \quad (3.58)$$

Note that, this entropy density and spatial entropy current for the Einstein-Gauss-Bonnet theory has been constructed following the philosophy of section 3.1. In the next subsection we shall do a systematic study of this entropy current, concentrating particularly in  $d = (3 + 1)$  space-time dimensions, where the Gauss-Bonnet term becomes topological.

### 3.6 The Einstein-Gauss-Bonnet theory in $d = 3 + 1$

The Gauss-Bonnet theory in  $(3 + 1)$  space-time dimensions needs a separate discussion. In this case, the Gauss-Bonnet term becomes a total derivative term and therefore it does not contribute to the equations of motion, i.e.  $E_{vv}^{HD} = 0$  identically. However, if one uses the Wald entropy as the equilibrium definition of black hole entropy (2.3), there is a finite non-vanishing contribution to it even from the topological Gauss-Bonnet part of the Lagrangian. The Wald entropy density  $s_w^{HD}$  (see (3.7)) for this case, is given by the Ricci scalar of the co-dimension-2 spatial slice of the horizon  $\mathcal{H}_v$ ,

$$s_w^{HD} = 2a_{gb}\mathcal{R}, \quad (3.59)$$

where  $a_{gb}$  is the Gauss-Bonnet parameter appearing in (3.54). Since  $\mathcal{H}_v$ , in this case, is a 2-dimensional manifold, the integrated total entropy  $S_W$  becomes the topological Euler number of  $\mathcal{H}_v$ .

Once we consider dynamical black hole solutions in this theory and restrict ourselves to consider perturbations characterized by small amplitudes around a stationary configuration, the total integrated Wald entropy  $S_W$  doesn't change with time as long as the perturbation is small and therefore, does not affect the topology of  $\mathcal{H}_v$ . However, if we consider the local Wald entropy density  $s_w$  (without being integrated on the spatial slice  $\mathcal{H}_v$ ), it does indeed change with time and therefore has a non-zero contribution to  $\partial_v(\sqrt{h} \tilde{B})$ . With these in mind let us look at (3.42), which is the main result of this note and rewrite it here again for convenience

$$E_{vv}^{\text{HD}}|_{\text{zero boost}} \sim -\partial_v \left( \frac{1}{\sqrt{h}} \partial_v \left( \sqrt{h} \tilde{B} \right) \right) - \partial_v \left( \nabla_i J^i \right). \tag{3.60}$$

From the above discussion it is clear that for  $(3 + 1)$  space-time dimensions, the l.h.s. of (3.60) vanishes identically. However, the first term on the r.h.s. is non-zero, making us wonder how to make sense of this equation if we had not included the second term on r.h.s. involving the spatial entropy current. As we will see, both the terms on the r.h.s. of the above equation are non-zero but they will precisely cancel each other, and that is how this equation will be satisfied. In other words, we are left with verifying that the r.h.s. of (3.60) vanishes identically without using any on-shell gravity equations of motion, up to  $\mathcal{O}(\epsilon^2)$  corrections.

We start by noting that the values of the coefficients  $a_i$  presented in table 8 are achieved by explicitly computing the  $E_{vv}^{\text{HD}}$  for different four derivative gravity theories and for our metric choice (2.1), but most importantly, the results are not limited to the space-time dimensions we are working in. Therefore the same results (presented in table 8) holds for  $(3 + 1)$ -dimensional space-time as well. The specific values of these coefficients for Gauss-Bonnet theory turns out to be the following:

$$a_{16} = 2, a_{17} = -1, a_i = 0 \text{ (for all } i = 1, \dots, 15),$$

$$\text{such that, } E_{vv}^{\text{HD}} = a_{16} T_{16} + a_{17} T_{17} = 2 \left( \mathcal{R}^{ij} - \frac{1}{2} h^{ij} \mathcal{R} \right) \partial_v^2 h_{ij}. \tag{3.61}$$

However, for 2-dimensional space-time one can show that the following relation is identically true,

$$\mathcal{R}^{ij} - \frac{1}{2} h^{ij} \mathcal{R} = 0. \tag{3.62}$$

This is true because, in 2-dimensional space we could always choose a coordinate system where the metric is conformally flat and Einstein tensor vanishes on any 2-dimensional conformally flat space-time.

Let us now consider the expression  $(1/\sqrt{h}) \partial_v(\sqrt{h} \mathcal{R})$ . It is well-known that the linear variation of Ricci scalar around any metric generates a term proportional to the Einstein tensor plus a total derivative term. Because of the fact mentioned above, without doing

any further calculation, we could say

$$\frac{1}{\sqrt{h}}\partial_v\left(\sqrt{h}\mathcal{R}\right)=\left(\mathcal{R}^{ij}-\frac{1}{2}h^{ij}\mathcal{R}\right)(\partial_v h_{ij})+\nabla_i Z^i, \quad (3.63)$$

where  $Z^i$  is some spatial current characterizing the total derivative term, which could be easily fixed as follows. Using table 11 we could find the list of values for  $A_i, B_i$  for Gauss-Bonnet theory

$$A_5=-2, B_2=2, B_3=-2, \quad (3.64)$$

leading to the following expression for  $Z^i$

$$Z^i=2\nabla_j(K^{ij}-h^{ij}K) \quad (3.65)$$

Again in  $(3+1)$ -dimensions, where  $\{i, j\}$  indices run over  $\{1, 2\}$ , the first term in the r.h.s. of (3.63) identically vanishes. Therefore, we can immediately rewrite (3.63) as

$$\frac{1}{\sqrt{h}}\partial_v\left(\sqrt{h}\mathcal{R}\right)-\nabla_i Z^i=0, \quad (\text{only in } (3+1)\text{-dimensions})$$

This looks exactly like a divergence of a four-current and identically vanishes in  $(3+1)$ -dimensions. In this particular case of  $(3+1)$ -dimensional space-time, the above expression has exactly the same status as that of (3.52), or the structure multiplying the coefficient  $A_2$  in the expression of  $E_{vv}^{\text{HD}}$  (3.47), see also (3.37) and (3.38). In other words, in  $(3+1)$ -dimension we are free to add  $\{\mathcal{R}, Z^i\}$  to the expression of entropy density and spatial entropy current respectively, with any arbitrary overall coefficient. Such an addition will not affect the ultra-local version of the second law or the physical process version of the first law and this is true for all theories as long as we are restricting ourselves to  $(3+1)$  dimensions. However, just like in case of  $A_2$ , the Wald entropy formalism fixes that arbitrary coefficient to a very specific value.

To summarize, the main physical interpretation that one can draw from the arguments presented above is the following. For Gauss-Bonnet theory in  $(3+1)$  dimensions  $E_{vv}^{\text{HD}}$  vanishes identically and that is related to the fact that the total integrated Wald entropy  $S_W$  is not changing due to time-dependent perturbations. This is because the Wald entropy  $S_W$  in this case is given by topological Euler number of the 2-dimensional  $\mathcal{H}_v$  and we are considering small amplitude approximation for the perturbations, which are too weak to change the topology of  $\mathcal{H}_v$ . However, even in that approximation, the local change of entropy density is not vanishing. This necessitates the introduction of the idea of a spatial entropy current, that quantifies the inflow or outflow of local entropy density and cancels the change in local entropy density, within any infinitesimal region in  $\mathcal{H}_v$ . This analysis, at least for the situation considered in this subsection, therefore plays an important role in motivating the need for a spatial entropy current.

### 3.7 Comments on entropy current for higher boost terms in $E_{vv}^{\text{HD}}$

Once we have analyzed the zero-boost terms, the next immediate question is to analyze the contribution of higher boost terms of  $E_{vv}^{\text{HD}}$ , to the entropy current (3.43). As we

have observed in section 2, the arguments in [1] for second law, works smoothly, for all the higher-boost terms in  $E_{vv}^{\text{HD}}$ . The contribution from these higher boost terms, to the total entropy falls within the class of JKM ambiguities, and they do not contribute to the physical process version of the first law. Unlike the zero-boost terms, nothing necessitates the existence of a spatial component of the current for these higher boost terms. Both the first and second law would remain valid, if we simply declare that these terms would just modify the entropy density as in (2.33), and they do not affect the spatial components of the current. However, the spatial components of the current could still exist, even for the higher boost terms as we now demonstrate.

Before we proceed it is worth clarifying that we will not be doing an exhaustive classification of all such possible higher boost terms in  $E_{vv}^{\text{HD}}$ . Our aim here is just to present an argument based on analyzing a candidate term as an example that justifies the above-mentioned statement. We postpone a more detailed study of this aspect to future work.

Schematically, the higher boost terms have the following structure (see (2.22) or (3.27))

$$E_{vv}^{\text{HD}}|_{\text{higher boost}} \sim \partial_v^2 \left[ \partial_r^k A^{(k)} \partial_v^k B^{(k)} \right] + \mathcal{O}(\epsilon^2), \quad (3.66)$$

where  $A^{(k)}$  and  $B^{(k)}$  are boost-invariant. Now it turns out that the same higher boost term could be recast in different ways, up to corrections that are quadratic or higher-order, in the amplitude of the dynamics. This allows us to absorb certain higher boost terms (the ones that have at least one  $\nabla_i$ ) either entirely within the correction to entropy density, or partially in entropy density and partially in the spatial components of the current. Let us explain this ambiguity more specifically.

Consider a typical higher boost term in  $E_{vv}^{\text{HD}}$ , as in (3.66), where the term  $\partial_v^k B^{(k)}$  could be expressed as divergence of a spatial current with boost-weight  $k > 1$ ,

$$\partial_v^k B^{(k)} \sim \vec{\nabla} \cdot \vec{J}^{(k)},$$

and substituting it in the expression of  $E_{vv}^{\text{HD}}$  we find

$$E_{vv}^{\text{HD}}|_{\text{higher boost}} \sim \partial_v^2 \left[ \partial_r^k A^{(k)} \vec{\nabla} \cdot \vec{J}^{(k)} \right] \quad (3.67)$$

$$= \partial_v \left( \vec{\nabla} \cdot \left[ \partial_v \left( \vec{J}^{(k)} \partial_r^k A^{(k)} \right) \right] \right) + \partial_v^2 \left[ -\vec{J}^{(k)} \cdot \vec{\nabla} \left( \partial_r^k A^{(k)} \right) \right] + \mathcal{O}(\epsilon^2). \quad (3.68)$$

On one hand, from the first line of (3.67), we can conclude that the contribution of this term to entropy current is simply<sup>30</sup>

$$\text{From (3.67): } S_{k \geq 1}^v = -\partial_r^k A^{(k)} \vec{\nabla} \cdot \vec{J}^{(k)}, \quad S_{k \geq 1}^i = 0. \quad (3.69)$$

with no spatial current. On the other hand, from the second line (3.68), we may infer that, this term contributes to the spatial components of the entropy current, apart from

---

<sup>30</sup>The quantity  $S^A$  has been introduced in (3.43). The subscript  $k \geq 1$  in  $S^A$  is to denote that this is the contribution from the higher boost terms in  $E_{vv}^{\text{HD}}$ .



the contribution to the entropy density, which is different from the previous case (3.69). That is, we can write the contribution to entropy current also in the following way

$$\text{From (3.68): } S_{k \geq 1}^v = \vec{J}^{(k)} \cdot \vec{\nabla} \left( \partial_r^k A^{(k)} \right), \quad \vec{S}_{k \geq 1} = -\partial_v \left( \vec{J}^{(k)} \partial_r^k A^{(k)} \right). \quad (3.70)$$

We would like to emphasize that the above manipulation, which is essentially an interchange of  $\partial_v$  and  $\vec{\nabla}$ , crucially uses the fact that any term, generated due to the non zero commutator of these two types of derivatives, would be of higher-order in amplitude. This is because the commutator itself is of boost weight one, for higher boost terms,

$$\partial_v(\vec{\nabla} \cdot \vec{J}^k) \sim \vec{\nabla} \cdot (\partial_v \vec{J}^k) + (\partial_v \Gamma_{ij}^i) J^{(k)j} \sim \vec{\nabla} \cdot (\partial_v \vec{J}^k) + \mathcal{O}(\epsilon^2). \quad (3.71)$$

Also note that, this is true only for the higher-boost terms, and in particular, not true for the boost invariant terms, for which the presence non-zero spatial entropy current was unambiguous.

The two different choices of entropy density in (3.69) and (3.70), are related by a total spatial derivative, as expected. This ensures that, in the integrated (weak) version of the second law, this difference would have no impact. However, in the ultra-local version, where we demand the entropy to be produced at every point in space and time, this difference is significant. This leads to an ambiguity in the definition of our entropy current, which cannot be fixed, merely from the transformation property of  $E_{vv}^{\text{HD}}$  under boost (2.11).

It is possible that, if we keep track of the higher order terms in amplitude expansion, this ambiguity may be removed. Alternatively, it is also possible that some suitable extension of the boost symmetry, like (2.9), which preserves our global choice of coordinates, might constraint the structure of our entropy current further, and consequently fix this ambiguity. We would like to explore this point further in our future work.

#### 4 Discussions and future directions

In this note, we have demonstrated that the intricacies in the arguments involved in the proof of the physical process version of the first law, and the second law, naturally lead us to the notion of a spatial entropy current on the horizon. This spatial entropy current captures the inflow or outflow of entropy from any sub-region of  $\mathcal{H}_v$  — the horizon  $v$ -slice. For most of our analysis in this note, we consider dynamical black holes which can be treated within the linearized approximation, where the amplitude of the ‘time’ dependent metric fluctuations, about a given stationary black hole solution, is small. Under this approximation, we are able to establish that the entropy density and the spatial components of the entropy current, constructed through our algorithm, satisfy an ultra-local stronger version of the second law of black hole thermodynamics. The validity of this local form of the second law is ensured by the equations of motion for the higher derivative theories of gravity, and therefore, true for any metric that solves these classical equations, at the linearized level.

The construction of our entropy current is not unique. All the ambiguities in defining the current can be traced back to the fact that there exist certain terms  $\mathcal{T}_{\text{amb}}$  which can be simultaneously written in two ways. We can write  $\mathcal{T}_{\text{amb}} = \frac{1}{\sqrt{h}} \partial_v \left( \sqrt{h} \mathcal{J}^v \right)$ , but also we

can write the same term as  $\mathcal{T}_{\text{amb}} = -\nabla_i \mathcal{J}^i$ , for some choice of  $\mathcal{J}^v$  and some choice of  $\mathcal{J}^i$ . Obviously, it follows that  $\mathcal{J}^v$  must have at least one spatial derivative, while  $\mathcal{J}^i$  must have at least one  $v$ -derivative. If we have such terms, appearing in the equation of motion as  $\partial_v \mathcal{T}_{\text{amb}}$ , then it becomes unclear whether to write it as  $\mathcal{T}_{\text{amb}} = \frac{1}{\sqrt{h}} \partial_v (\sqrt{h} \mathcal{J}^v)$  and consider it to be a part of the entropy density. Or to write it as  $\mathcal{T}_{\text{amb}} = -\nabla_i \mathcal{J}^i$  and interpret it to be being a part of the spatial components of the entropy current. A third possibility is to split this term up, into the entropy density and the entropy current. In section 3.7, we have discussed these kind of ambiguities in detail.

Again, if we indeed have terms like  $\mathcal{T}_{\text{amb}}$  which can be written in both these ways, we can always add a  $0 = \partial_v \left( \frac{1}{\sqrt{h}} \partial_v (\sqrt{h} \mathcal{J}^v) + \nabla_i \mathcal{J}^i \right)$ , to the equation of motion, and subsequently include  $\mathcal{J}^v$  and  $\mathcal{J}^i$  into the definition of the entropy density and the spatial entropy current. Neither the equation of motion nor any of the laws of thermodynamics would be affected by this operation. This kind of ambiguity arises, for example, in the Einstein-Gauss-Bonnet theory in 3+1 dimensions, discussed in detail in section 3.6.<sup>31</sup> If a term like  $\mathcal{T}_{\text{amb}}$  is such that,  $\mathcal{J}^v$  is non-zero on stationary solutions, then it would contribute to Wald entropy as well. In such a case, the ambiguity corresponding to this term may be removed by demanding that our entropy reduces to Wald entropy on stationary solutions. But if  $\mathcal{J}^v$  vanishes in equilibrium then this additional criterion would remain ineffective in fixing it.

It should also be noted that it may be possible to write down a particular term simultaneously in both the forms, only at the linearized order in perturbations. Such an equivalence may cease to be true once we proceed to consider corrections which are higher-order in amplitudes. In that case, these ambiguities would only be a linear order artifact and would disappear once we are able to construct the full non-linear current. However, some of these ambiguities of the entropy current may remain, even in the full non-linear construction.

Having highlighted the ambiguities of the entropy density and the corresponding current, we must point out that, every member of this ambiguous class, have the property that the total entropy reduces to Wald entropy for stationary black holes. For the non-stationary dynamical black holes, all these entropy density and currents also satisfy a local second law. Hence, every such entropy density and entropy current are perfectly well defined macroscopic entities that can provide excellent effective thermodynamics description of the system. Some additional microscopic information is likely to make one of them special, and it can stand out as the correct definition of entropy density and entropy current away from equilibrium. Therefore, despite these ambiguities, it appears to us, that the notion of the spatial components of the entropy current and a local second law on a dynamical horizon is a concept of significant importance in the thermodynamic description of black holes.

This note is essentially a series of observations, on the evolution of black hole entropy in dynamical scenarios, in a specific set of examples of higher derivative theories of gravity. Through explicit calculations, we have been able to test our hypothesis about the spatial components of the entropy current, only in four derivative theories of gravity. This is a small step towards formulating an ultra-local version of the second law in gravitational theories (if it exists in a full non-perturbative sense) and deciphering all its physical ramifications.

---

<sup>31</sup>Also see the ambiguity related to the parameter  $A_2$ , discussed in detail in section 3.4.

Clearly, there are several directions in which this work needs to be extended, so that a more complete picture of this whole mechanism may emerge. Here is a brief list of related questions, which we would like to investigate in the near future.

1. The reader may have noticed that, throughout this note, we have used the word ‘time’ under a quotation mark. This is because our ‘time’ here is not really a parameter along a time-like vector field; rather, it is the affine parameter along a ‘distinguished’ null direction that generates the event horizon. Therefore, the expression appearing in the local version of the second law is not the  $d + 1$  dimensional<sup>32</sup> covariant divergence of a covariant current. This is quite unlike the standard way in which the local version of the second law is expressed, for near-equilibrium dynamics of non-gravitational theories, where the  $d + 1$  dimensional Lorentz covariance is manifestly maintained.

On the event horizon, we do not have a time-like direction, so the question of Lorentz invariance does not arise here. However, our construction has used a specific choice of coordinates and physically we expect some form of invariance should exist once we choose a different coordinate system — for example, a different spatial slicing of the null generators.

It would be extremely important to explore whether any such invariance exists and if it exists, then how does it control our construction.

2. Another question related to the above is as follows. We have seen that our construction of the spatial entropy current mainly involves the ‘zero boost terms’ in the equation of motion. For this construction to work, these ‘zero boost terms’ were required to have a specific form. This requirement may be viewed as a set of constraints on the most general structure of the relevant component of the equation of motion (see section 3.3 and section 3.4). The physical origin of these constraints is, at the moment, unclear. We suspect that the reason behind these constraints could be the set of residual gauge invariance, expressed in (2.9), which is a generalization of the boost symmetry (2.11). Whether this suspicion is true, or there is a completely different reason for these constraints must be investigated through explicit computations.
3. The ‘zero boost terms’ in the equation of motion, which are central to our analysis in this note, are also relevant for the physical process version of the first law, and hence, control the definition of entropy in stationary situations.

Now, it is well known that Wald’s formalism [9, 10] also determines this same equilibrium entropy in a covariant fashion using the conserved Noether current corresponding to the diffeomorphism symmetry. It would be extremely interesting to clearly establish a connection between these two methods. In particular, if it is possible to identify our spatial current within Wald’s construction, it would probably lead to a more satisfying covariant construction of the entropy current. This may help us

---

<sup>32</sup>Remember if we are working in  $D + 1$  dimensional space-time, then the current is expected to have  $d + 1$  components, with  $d = D - 1$ . This is because this current would be defined on the event horizon, which is a co-dimension one surface.

arrive at an abstract proof for the existence of this entropy current and the local second law, for any higher derivative theory of gravity.

In absence of any such concrete proof, it would be quite useful to gather more data, simply by repeating the exercise presented in this note, for theories of gravity with 6 or more derivatives.

4. Another obvious generalization would be to extend our construction to non-linear order in amplitude  $\epsilon$ . This can potentially fix the ambiguity related to the construction of the spatial components of the entropy current, which arises for higher boost ( $k \geq 0$ ) terms (see the discussion in section 3.7).

But more importantly, it might provide us with further insights, which can help us formulate a non-perturbative proof of second law for higher derivative theories. For Einstein's theory, entropy production is ensured by the famous 'horizon area increase theorem', which is proved for any dynamical situation, in full non-perturbative way. It would be nice to have a similar proof (or a clean counter-example) for higher derivative theories of gravity.

5. Naively, it might seem that, at non-linear order, we do not have to worry about the second law, since for Einstein's theory itself, the entropy production takes place at quadratic order in amplitude. Now because higher derivative corrections are always suppressed compared to the leading order piece corresponding to Einstein's theory, they cannot reverse the sign which guarantees entropy production.

However, if we are interested in an ultra-local form, then during a non-trivial 'time' evolution, the contribution to entropy due to Einstein's theory could vanish locally at a given point. Then, for the question of entropy production and the second law, we must take the higher-derivative corrections seriously. See [36] for the construction of entropy in dynamical black holes for the Einstein-Gauss-Bonnet theory, where these subtle issues have been addressed. The construction of entropy in [36] did not yield a second law, for the most generic dynamical situation. But, in [36] this idea of a spatial entropy current was not used. It would be very interesting to revisit [36], and check if the obstruction is resolved when the spatial entropy current is incorporated into the statement of the second law.

6. Within the framework of gauge gravity duality, a precise correspondence exists between slowly varying fluctuations of a black hole and the hydrodynamic fluctuations of the boundary fluid. Since the boundary fluid dynamics comes equipped with a local entropy current, there exists a dual of this current, for the black hole in the bulk [23, 33, 34]. This dual also constitutes a gravitational entropy current, in this particular context. In [23], the construction has been done for two derivative Einstein's theories. While in [33, 34], it has been extended to higher derivative theories of gravity, following Wald's formalism of Noether current. All these constructions use the derivative expansion extensively and their validity is restricted to this particular case of fluid-gravity correspondence. Therefore, although these constructions of the

entropy current relate to horizon dynamics, it subtly uses the asymptotic AdS conditions, which ensures black-brane solutions exist and the fluid-gravity correspondence could be formulated in a clean fashion. For example, the entropy current constructed in these papers is clearly a (3+1) dimensional current (for 5 bulk dimension) with one component (the entropy density) clearly along with a time-like direction. This is achieved by lifting the null coordinate along the horizon, to the time-like direction of the boundary through the fluid-gravity map. This time-like direction also serves to formulate an unambiguous statement of the second law, in terms of the divergence of this entropy current.

On the other hand, our construction is completely confined to the horizon, it does not have any time-like direction, to begin with. As we have explained before, in absence of any Lorentz symmetry it is not straight-forward to interpret our result as a covariant ‘four’-current. Also, we do not need any assumption about the asymptotic structure of spatial infinity.

Our construction looks quite different from what has been done in [23, 33, 34]. But it is also clear that there must be some relation between these two constructions. This question is a topic of our ongoing investigation.

7. Very recently, one candidate entropy current has been constructed in [35], for Gauss-Bonnet theory within the framework of membrane-gravity duality, in an expansion in inverse powers of space-time dimension  $D$ . This is a duality that gives a precise correspondence between the dynamics of a membrane (a time-like hyper-surface, embedded in flat space-time) and that of the horizon, in the large  $D$  expansion. Unlike our construction, which does not rely on any duality, in [35] the entropy current has been constructed in the dual picture of the membrane. Their entropy current is entirely confined within the membrane and has the usual property of Lorentz invariance. In their case, the non-negative divergence of the entropy current follows from the membrane dynamics governed by those membrane equations, which have been derived from the dual gravity picture.

Moreover, within their approximation, the authors of [35] have also shown that the existence of a Killing vector is a consequence of no entropy production. They have also demonstrated that the charge corresponding to this conserved entropy current reduces to the well-known expression of Wald entropy, in a stationary situation.

It would be extremely interesting to see how the entropy current of [35] compares with ours. In particular, we would like to explore, if this membrane-gravity duality can be used to formulate a principle which can fix the ambiguities of the entropy density and spatial entropy current.

## Acknowledgments

We would like to thank Parthajit Biswas, Milan Patra, and Shuvayu Roy for initial collaboration, several useful discussions and many important inputs. We would also like to thank

Arjun Bagchi, Shamik Banerjee, Pallab Basu, Diptarka Das, Sayan Kar, Loganayagam R., Mangesh Mandlik, Shiraz Minwalla, Sudipta Mukherji, Arunabha Saha, Sudipta Sarkar, Tarun Sharma, Nilakash Sorokhaibam and Yogesh Kumar Srivastava for several useful discussions. NK and JB would like to acknowledge hospitality during an academic visit at NISER Bhubaneswar, where much of this work had been carried out.

## A A general stationary metric can have $v$ dependent components

In a black hole usually, the Killing generators of the horizon cannot be affinely parametrized maintaining the Killing conditions. In other words, the components of the stationary metric are independent of the Killing coordinate —  $\tau$ , but they are not independent of the affine parameter  $v$  along the generators of the Killing horizons. Though in a stationary metric with a Killing horizon, the Killing vector field —  $\partial_\tau$  and the affinely parametrized null generators  $\partial_v$  are proportional to each other and there exists a precise relation between them. In this appendix, we shall use this relation to fix the  $v$  dependence of the stationary metric.

More precisely, we would like to determine how the components of a generic stationary metric, written in the gauge of (2.1), could depend on the  $v$ -coordinate.

Consider a generic stationary black hole with a Killing horizon, i.e, there exists a coordinate  $\tau$  such that

1. All metric components are independent of  $\tau$
2.  $\partial_\tau$  is time-like everywhere outside the horizon.
3.  $\partial_\tau$  becomes null on the event horizon.

Now we could do exactly the same construction as in case of the metric (2.1), the only difference being that now the coordinates on the horizon would be  $\partial_\tau$  and  $\partial_i$ , instead of the affinely parametrized  $\partial_v$ . Let  $\rho$  be the coordinate that denotes distances away from the horizon. Now also we could choose  $\rho$  to be the affine parameter along the set of null geodesics, intersecting the horizon at fixed angles with  $\partial_\tau$  and  $\partial_i$  and labelled by the coordinates of the intersection point. Following the same logic as before, the metric in  $\tau$ ,  $x^i$  and  $\rho$  coordinate will have almost the same structure as that of (2.1). The  $\tau\tau$  and  $\tau i$  components of the metric will again vanish on the horizon ( $\rho = 0$ ) owing to the fact that it is a null hyper-surface. But since  $\partial_\tau$  is not affinely parametrized, unlike (2.1), the first  $\rho$  derivative of the  $(\tau\tau)$  component of the metric (let us denote it by  $g_{\tau\tau}(\rho, x^i)$ ) will not vanish on the horizon. However, for stationary black holes,  $\partial_\rho g_{\tau\tau}$  is related to the temperature of the black hole and the zeroth law of Black hole mechanics ensures  $\left[ \partial_\rho g_{\tau\tau} |_{\rho=0} \equiv C \right]$  is a constant, i.e., independent of the spatial coordinates  $x^i$ s. Putting all these facts together we finally write the most general stationary metric in our gauge.

$$ds^2 = 2 d\tau d\rho - (\rho C + \rho^2 X(\rho)) d\tau^2 + 2\rho \omega_i(\rho) d\tau dx^i + h_{ij}(\rho) dx^i dx^j \quad (\text{A.1})$$

Now we have to transform this metric to the gauge of (2.1) where the null coordinate along the horizon is an affine parameter of the null generators  $v$ . Our final goal is to find out how the metric components of an arbitrary stationary metric will depend on  $v$ .

The coordinate transformation which fulfills this objective is given by

$$\rho = \frac{C}{2} r v, \quad \tau = \frac{2}{C} \log\left(\frac{C v}{2}\right) \quad (\text{A.2})$$

The metric in the new coordinate takes the following form

$$ds^2 = 2 dv dr - r^2 X(Crv/2) dv^2 + r\omega_i(Crv/2) dv dx^i + h_{ij}(Crv/2) dx^i dx^j \quad (\text{A.3})$$

To get the above metric, we have crucially used the fact that  $C$  is independent of  $v$  and  $x^i$  ( $C$  is independent of  $\rho$  by construction).

The most important noteworthy feature of this metric (A.3) is that the metric components are explicitly dependent on the  $v$ -coordinate, although it describes a stationary black hole because it is a mere coordinate transformation of the most general stationary metric (A.1). However, though imposing the condition of stationarity on the general form of the metric (2.1), *does not* imply that the metric functions  $X$ ,  $\omega_i$  and  $h_{ij}$  should be independent of the  $v$  coordinate, there are some constraints on the  $v$  dependence of the stationary metrics. Here the metric components never depend on  $r$  and  $v$  independently, but always on the product  $rv$ . In other words, on any metric of the form (A.3), with components depending only on the product  $rv$ , we could always apply the inverse of the coordinate transformation (A.2) to take it to a form where redefined coordinate —  $\tau$  is manifestly the Killing coordinate.

Now for the proof of second law, it is crucial that the  $\partial_v$  of entropy vanishes on stationary black holes attained at  $v \rightarrow \infty$ . Naively the form of the stationary metric (A.3) contradicts this step of the argument. But note that any  $\partial_v$  derivative on the metric components in (A.3), will also bring down a factor of  $r$  and therefore will vanish on  $\mathcal{H}$  (the hyper-surface at  $r = 0$ ), unless there is also one  $\partial_r$  derivative present along with every  $\partial_v$  derivative. Thus we may conclude that, the terms of the form  $((\partial_r \partial_v)^m P)$ , where  $P$  is a function of the metric components in (2.1) and their  $\nabla_i$  derivatives (without any  $\partial_r$  or  $\partial_v$  derivatives), can be non-zero on a generic Killing horizon. Note that, all such terms are invariant under the  $\lambda$  scaling (2.11). It also implies that the terms of the form  $(\partial_r^n \partial_v^m P)$ , with  $m > n$  must vanish on a Killing horizon. This is because, as is apparent from (A.3), the higher number of  $v$ -derivatives would give rise to factors of  $r$ , which will force the entire term to zero on the  $r = 0$  hyper-surface.

## B Arguments leading to vanishing of $T_{vv}$ on any Killing horizon

In this section we would like to argue that the  $vv$  component of the matter stress tensor vanishes on Killing horizons.

We shall use the boost transformation property of  $T_{vv}$  to reach this conclusion. Note that just like the gravity part of the equation of motion (i.e.,  $E_{vv}$  or  $E_{vv}^{\text{HD}}$ ), matter stress tensor itself is a covariant with nice transformation properties under any coordinate transformation, in particular the  $\lambda$  scaling described in equation (2.11).  $T_{vv}$  should transform exactly the way  $E_{vv}$  or  $E_{vv}^{\text{HD}}$  transforms, namely

$$T_{vv} \rightarrow T_{\tilde{v}\tilde{v}} = \frac{1}{\lambda^2} T_{vv} \quad (\text{B.1})$$

Now we shall consider only those stress tensors that are regular on the event horizons at least in those coordinate systems where the full dynamical metric is regular, everywhere apart from the black hole singularity. This is certainly the case in the coordinate system we have chosen in our metric (2.1). It follows that  $T_{vv}$  must admit a Taylor series expansion around the horizon at  $r = 0$ . Equation (2.11) and equation (B.1) together suggest the following expansion for  $T_{vv}$

$$T_{vv} = \frac{1}{v^2} \sum_{k=0}^{\infty} (rv)^k w^{(k)}(\vec{x}) \tag{B.2}$$

where  $w^{(k)}$  s are scalar functions of only the spatial coordinates  $\{x^i\}$ , (i.e., they are both boost invariant and also invariant under any coordinate transformation that mixes only the  $\{x^i\}$  coordinates among themselves). Exactly on the horizon only the leading terms of the above expansion will contribute.

$$T_{vv}|_{\text{horizon}} = \frac{w^{(0)}(\vec{x})}{v^2} \tag{B.3}$$

Note that both (B.2) and (B.3) do not need any stationarity for their validity.

Now let us specialize to stationary cases. Here we have a Killing vector ( $\partial_\tau$ ). All relevant fields including the matter fields are independent of this  $\tau$  coordinate and the same is true for  $T_{vv}$ , as well. In terms of equation it implies

$$\partial_\tau T_{vv}|_{\text{stationary}} = 0 \tag{B.4}$$

From equation (A.2) it follows

$$\partial_\tau = \frac{C}{2} (v \partial_v - r \partial_r) \tag{B.5}$$

Equation (B.5) clearly contradicts equation (B.4) unless  $w^{(k)} = 0$  for every  $k$ . It follows that  $T_{vv}$  vanishes identically on any configuration with a Killing vector.

We would like to emphasize that in the above arguments the key elements are

1. The existence of the event horizon (or more precisely a null hyper-surface at  $r = 0$ ) so that the horizon-adapted coordinate choice in the metric (2.1) and consequently the boost symmetry is meaningful.
2. Stationarity or the existence of a Killing vector, which is proportional to the null generators of the horizon.

We have not used the fact that  $T_{vv}$  is stress tensor, neither the fact that the field configuration (including the metric) satisfies any particular equation. What we have argued is that whenever there is one Killing vector field, the  $vv$  component of any covariant tensor identically vanishes in the vicinity of the horizon (where the Taylor expansion in equation (B.2) makes sense) and it is a completely off-shell statement.



## C Conventions, notations and useful formulae

In this appendix we summarize our conventions, write down various notations and collectively represent several important formulae that we have used in the note.

- The coordinate choice:

$x^\mu$  = The full space-time coordinates in  $(d + 1)$ -dimensions :  $\{v, r, x^i\}$ ,  
 $v$  = The Eddington-Finkelstein type time coordinate,  
 $r$  = The radial coordinate,  
 $x^i$  = The  $(d - 1)$  spatial coordinates,

- The choice for the space-time metric:

$$\begin{aligned}
 ds^2 = & 2 dv dr - r^2 X(r, v, x^i) dv^2 + 2 r \omega_i(r, v, x^i) dv dx^i \\
 & + h_{ij}(r, v, x^i) dx^i dx^j
 \end{aligned}
 \tag{C.1}$$

- Useful notations and conventions:

1.  $\mathcal{H}$  = The co-dimension one horizon, which we choose to be at the radial coordinate  $r = 0$ ,
2.  $\mathcal{H}_v$  = The co-dimension two, constant  $v$ -slice of the horizon,
3.  $h$  = Determinant of the induced metric,  $h_{ij}$ , on  $\mathcal{H}$ ,
4. The total integrated Wald entropy at equilibrium is defined as

$$S_W = -2\pi \int_{\mathcal{H}_v} d^{d-2}x \sqrt{h} \frac{\partial \mathcal{L}}{\partial R_{\mu\nu\rho\sigma}} \epsilon_{\mu\nu} \epsilon_{\rho\sigma} = -2\pi \int_{\mathcal{H}_v} d^{d-2}x \sqrt{h} s_w,$$

where  $\epsilon_{\mu\nu}$  = Bi-normal to  $\mathcal{H}_v$ ,

5.  $s_w = \frac{\partial \mathcal{L}}{\partial R_{\mu\nu\rho\sigma}} \epsilon_{\mu\nu} \epsilon_{\rho\sigma}$  = The Wald entropy density,
6.  $S_W^{\text{HD}}, s_w^{\text{HD}}$  = Contributions to integrated Wald entropy ( $S_W$ ) and Wald entropy density ( $s_w$ ) from the higher derivative part of the gravity Lagrangian  $\mathcal{L}^{\text{HD}}$ .  
 It can be shown that:  $s_w = 1 + s_w^{\text{HD}}$ , such that for Einstein gravity one obtains  $s_w = 1$ .
7. The density of time variation of Wald entropy is denoted by  $\vartheta$ , defined as

$$\partial_v S_W = \int_{\mathcal{H}_v} d^{d-2}x \sqrt{h} \vartheta; \quad \vartheta = \vartheta_E + \vartheta^{\text{HD}},$$

such that  $\vartheta_E = \frac{1}{\sqrt{h}} \partial_v (\sqrt{h})$  = contribution from Einstein gravity, and  $\vartheta^{\text{HD}} = \frac{1}{\sqrt{h}} \partial_v (\sqrt{h} s_w^{\text{HD}})$  = contribution from higher derivative part of the Lagrangian  $\mathcal{L}^{\text{HD}}$ ,

8.  $s_c$  = correction to the entropy density which vanish on stationary solutions,

9.  $E_{vv}^{\text{HD}}$  = ‘vv’ component of the equation of motion, getting contribution only from the higher derivative part of the Lagrangian  $\mathcal{L}^{\text{HD}}$ ,

- Useful definitions:

1. The extrinsic curvatures of the horizon  $\mathcal{H}$ :

$$\begin{aligned} \text{(a). } \mathcal{K}_{ij} &= \frac{1}{2} \partial_v h_{ij}; \quad \mathcal{K}^{ij} = -\frac{1}{2} \partial_v h^{ij}, \\ \text{(b). } \bar{\mathcal{K}}_{ij} &= \frac{1}{2} \partial_r h_{ij}; \quad \bar{\mathcal{K}}^{ij} = -\frac{1}{2} \partial_r h^{ij}. \end{aligned}$$

2. The trace of the extrinsic curvatures:

$$\begin{aligned} \text{(a). } \mathcal{K} &= \frac{1}{2} h^{ij} \partial_v h_{ij} = \frac{1}{\sqrt{h}} \partial_v \sqrt{h}, \\ \text{(b). } \bar{\mathcal{K}} &= \frac{1}{2} h^{ij} \partial_r h_{ij} = \frac{1}{\sqrt{h}} \partial_r \sqrt{h}. \end{aligned}$$

- Expressions for the components of Riemann tensors, Ricci tensors and Ricci scalar on the horizon:

$$\begin{aligned} R_{rvrv} &= X + \frac{1}{4} \omega^2 \\ R_{rvri} &= -\partial_r \omega_i + \frac{1}{2} \omega^j \bar{\mathcal{K}}_{ij} \\ R_{rvvi} &= -\frac{1}{2} (\partial_v \omega_i + \omega^j \mathcal{K}_{ij}) \\ R_{rirj} &= -\partial_r \bar{\mathcal{K}}_{ij} + \bar{\mathcal{K}}_{ik} \bar{\mathcal{K}}_j^k \\ R_{rivj} &= -\partial_r \mathcal{K}_{ij} + \frac{1}{2} \nabla_j \omega_i - \frac{1}{4} \omega_i \omega_j + \bar{\mathcal{K}}_{jk} \mathcal{K}_i^k \\ R_{vivi} &= -\partial_v \mathcal{K}_{ij} + \mathcal{K}_{ik} \mathcal{K}_j^k \\ R_{ijvk} &= \nabla_j \mathcal{K}_{ik} - \nabla_i \mathcal{K}_{jk} - \frac{1}{2} \omega_i \mathcal{K}_{jk} + \frac{1}{2} \omega_j \mathcal{K}_{ik} \\ R_{ijrk} &= \left( \nabla_j - \frac{1}{2} \omega_j \right) \bar{\mathcal{K}}_{ik} - \left( \nabla_i - \frac{1}{2} \omega_i \right) \bar{\mathcal{K}}_{jk} \\ R_{ijkl} &= \mathcal{R}_{ijkl} - \bar{\mathcal{K}}_{ik} \mathcal{K}_{jl} - \mathcal{K}_{ik} \bar{\mathcal{K}}_{jl} + \bar{\mathcal{K}}_{il} \mathcal{K}_{jk} + \mathcal{K}_{il} \bar{\mathcal{K}}_{jk} \end{aligned} \tag{C.2}$$

where  $\nabla_i$  is the covariant derivative with respect to the induced metric  $h_{ij}$ .

- Expressions for the components of Ricci tensors on the horizon:

$$\begin{aligned} R_{rr} &= -\partial_r \bar{\mathcal{K}} - \bar{\mathcal{K}}_{ij} \bar{\mathcal{K}}^{ij} \\ R_{rv} &= -X - \frac{1}{2} \omega^2 - \partial_r \mathcal{K} - \bar{\mathcal{K}}_{ij} \mathcal{K}^{ij} + \frac{1}{2} \nabla^i \omega_i \\ R_{ri} &= \partial_r \omega_i - \frac{1}{2} \bar{\mathcal{K}}_i^j \omega_j + \left( \nabla_j - \frac{1}{2} \omega_j \right) \bar{\mathcal{K}}_i^j - \left( \nabla_i - \frac{1}{2} \omega_i \right) \bar{\mathcal{K}} \\ R_{vv} &= -\partial_v \mathcal{K} - \mathcal{K}_{ij} \mathcal{K}^{ij} \end{aligned} \tag{C.3}$$

$$\begin{aligned}
 R_{vi} &= -\frac{1}{2} \partial_v \omega_i - \frac{1}{2} \mathcal{K}_i^j \omega_j + \left( \nabla_j + \frac{1}{2} \omega_j \right) \mathcal{K}_i^j - \left( \nabla_i + \frac{1}{2} \omega_i \right) \mathcal{K} \\
 R_{ij} &= \mathcal{R}_{ij} - 2 \partial_r \mathcal{K}_{ij} + \frac{1}{2} (\nabla_j \omega_i + \nabla_i \omega_j - \omega_i \omega_j) - \bar{\mathcal{K}} \mathcal{K}_{ij} - \mathcal{K} \bar{\mathcal{K}}_{ij} \\
 &\quad + 2 \left( \bar{\mathcal{K}}_{ik} \mathcal{K}_j^k + \bar{\mathcal{K}}_{jk} \mathcal{K}_i^k \right)
 \end{aligned}$$

- Expressions for the Ricci scalar on the horizon:

$$R = \mathcal{R} - 2X - \frac{3}{2} \omega^2 - 4 \partial_r \mathcal{K} + 2(\nabla \cdot \omega) - 2 \bar{\mathcal{K}}_{ij} \mathcal{K}^{ij} - 2 \mathcal{K} \bar{\mathcal{K}} \quad (\text{C.4})$$

## D Detailed expressions

### D.1 Expressions of Riemann tensors and Ricci tensors off the horizon

As we will compute the ‘vv’-component of the equations of motion  $E_{vv}^{\text{HD}}$ , we will need the following expressions for the components of Riemann tensors and Ricci tensor calculated off the horizon, i.e. without imposing  $r = 0$ ,

$$\begin{aligned}
 R_{rvrv} &= X + \frac{1}{4} \omega^2 \\
 R_{rviv} &= \frac{1}{2} \nabla_i (2rX + r^2 \partial_r X) + \frac{1}{2} \partial_v (\omega_i + r \partial_r \omega_i) + \frac{r^2}{2} (2X + r \partial_r X) \omega^j \bar{\mathcal{K}}_{ij} \\
 &\quad - r (\partial_v \omega_j) \bar{\mathcal{K}}_i^j - \frac{r^2}{2} (\nabla_j X) \bar{\mathcal{K}}_i^j + \frac{r}{4} (\omega^j (\omega_j + r \partial_r \omega_j)) (\omega_i + r \partial_r \omega_i) \\
 &\quad - \frac{r}{4} (\omega^j + r \partial_r \omega^j) (\nabla_j \omega_i - \nabla_i \omega_j - \mathcal{K}_{ij}) \\
 4 R_{vivi} &= r^2 (\nabla^m \omega_i) (\nabla_m \omega_j) + r^2 (\nabla_i \omega^m) (\nabla_j \omega_m) - r^2 (\nabla_m \omega_i) (\nabla_j \omega^m) \\
 &\quad - r^2 (\nabla_m \omega_j) (\nabla_i \omega^m) - 2r (\nabla_m \omega_i) \mathcal{K}_j^m - 2r (\nabla_m \omega_j) \mathcal{K}_i^m + 4 \mathcal{K}_{im} \mathcal{K}_j^m \\
 &\quad + r^2 (\nabla_i \omega^m) \mathcal{K}_{jm} + 2r (\nabla_j \omega^m) \mathcal{K}_{im} + 2 \bar{\mathcal{K}}_{ij} [r^2 \omega^2 (2rX + r^2 \partial_r X) \\
 &\quad + r^2 X (2rX + r^2 \partial_r X) - r^3 (\omega \cdot \nabla) X - 2 r^2 \omega^m (\partial_v \omega_m) - (\partial_v r^2 X)] \\
 &\quad + (\omega_j + r \partial_r \omega_j) [2 r \omega^m \mathcal{K}_{im} - r^2 (\omega \cdot \nabla) \omega_i + r^2 \omega^m (\nabla_i \omega_m) + (\nabla_i r^2 X)] \\
 &\quad + 4r^2 (r^2 X + r^r \omega^2) (\omega_i + r \partial_r \omega_i) (\omega_j + r \partial_r \omega_j) \\
 &\quad + (\omega_i + r \partial_r \omega_i) [2 r \omega^m \mathcal{K}_{jm} - r^2 (\omega \cdot \nabla) \omega_j + r^2 \omega^m (\nabla_j \omega_m) + (\nabla_j r^2 X)] \\
 &\quad - r^2 (2X + r \partial_r X) [\nabla_i \omega_j + \nabla_j \omega_i] \\
 &\quad + 2r^2 \nabla_i \nabla_j X + 2(2rX + r^2 \partial_r X) \mathcal{K}_{ij} - 4 \partial_v \mathcal{K}_{ij} \\
 &\quad + 2r \nabla_i (\partial_v \omega_j) + 2 r \nabla_j (\partial_v \omega_i) \\
 R_{vv} &= r^2 (X + \omega^2) \left[ \frac{1}{2} (2X + 4rX + r^2 \partial_r^2 X) + \frac{1}{2} (\omega^i + r \partial_r \omega^i) (\omega_i + r \partial_r \omega_i) \right. \\
 &\quad \left. + \frac{1}{2} \bar{\mathcal{K}} (2rX + r^2 \partial_r X) \right] - r^2 (\omega \cdot \nabla) (2X + r \partial_r X) - r \omega^i (\partial_v (\omega_i + r \partial_r \omega_i)) \\
 &\quad - r^3 (2X + r \partial_r X) (\omega^i \bar{\mathcal{K}}_{ij} \omega^j) + 2 r^2 (\partial_v \omega_i) \omega_j \bar{\mathcal{K}}^{ij} + r^3 (\nabla_j X) \omega_i \bar{\mathcal{K}}^{ij} \\
 &\quad - \frac{r^2}{2} (\omega^i (\omega_i + r \partial_r \omega_i))^2 + r^2 \omega^i ((\omega + r \partial_r \omega) \cdot \nabla) \omega_i \quad (\text{D.2})
 \end{aligned}$$

$$\begin{aligned}
 & -r^2(\omega^i + r\partial_r\omega^i)(\omega \cdot \nabla)\omega_i + \frac{r^2}{2}(\nabla^i\omega^j)(\nabla_i\omega_j) - \frac{r^2}{2}(\nabla_j\omega^i)(\nabla_i\omega^j) \\
 & -\partial_v\mathcal{K} - \mathcal{K}_{ij}\mathcal{K}^{ij} + \frac{r^2}{2}(\nabla^2 X) + \frac{1}{2}(2rX + r^2\partial_r X)\mathcal{K} \\
 & + r\nabla^i(\partial_v\omega_i) - \frac{r^2}{2}\bar{\mathcal{K}}(\omega \cdot \nabla)X - \frac{r^2}{2}\bar{\mathcal{K}}(\partial_v X) - r^2\bar{\mathcal{K}}\omega^i(\partial_v\omega_i) \\
 & + \frac{r^2}{2}((\omega + r\partial_r\omega) \cdot \nabla)X - \frac{r^2}{2}(2X + r\partial_r X)(\nabla \cdot \omega)
 \end{aligned}$$

## D.2 Relevant terms on the horizon $\mathcal{H}$ , to compute $E_{vv}^{\text{HD}}$ for different theories

$$\begin{aligned}
 R_v^\alpha R_{\alpha v} &= (\partial_v\mathcal{K})(2X + \omega^2 - \nabla \cdot \omega + 2\partial_v\bar{\mathcal{K}}) \\
 R_v^{\alpha\beta\gamma} R_{v\alpha\beta\gamma} &= \omega^i\omega^j\partial_v\mathcal{K}_{ij} - 2\nabla^i\omega^j\partial_v\mathcal{K}_{ij} + 4\partial_v\bar{\mathcal{K}}_{ij}\partial_v\mathcal{K}^{ij} \\
 R^{\alpha\beta} R_{v\alpha v\beta} &= -\left(X + \frac{\omega^2}{4}\right)(\partial_v\mathcal{K}) - \mathcal{R}^{ij}\partial_v\mathcal{K}_{ij} - (\nabla^i\omega^j)\partial_v\mathcal{K}_{ij} \\
 & \quad + \frac{1}{2}\omega^i\omega^j\partial_v\mathcal{K}_{ij} + 2\partial_r\mathcal{K}_{ij}\partial_v\mathcal{K}^{ij} \\
 D_v D_v R &= \partial_v^2\mathcal{R} - 2\partial_v^2 X - 3\omega^i\partial_v^2\omega_i + 3\omega^i\omega^j\partial_v\mathcal{K}_{ij} \\
 & \quad + 2\partial_v^2(\nabla \cdot \omega) - 4\partial_r\partial_v^2\mathcal{K} - 2\bar{\mathcal{K}}_{ij}\partial_v^2\mathcal{K}^{ij} \\
 & \quad - 2\bar{\mathcal{K}}\partial_v^2\mathcal{K} - 4\partial_v\bar{\mathcal{K}}_{ij}\partial_v\mathcal{K}^{ij} - 4\partial_v\mathcal{K}\partial_v\bar{\mathcal{K}} \\
 D_\alpha D^\alpha R_{vv} &= \left(\frac{3}{2}\omega^2 - \nabla \cdot \omega + 4X\right)(\partial_v\mathcal{K}) + (\omega \cdot \nabla)\partial_v\mathcal{K} \\
 & \quad - 2\omega^i\nabla^j(\partial_v\mathcal{K}_{ij}) - 2\partial_r\partial_v^2\mathcal{K} - \omega^i\partial_v^2\omega_i - \nabla^2(\partial_v\mathcal{K}) \\
 & \quad + 2\nabla^i(\partial_v^2\omega_i) - \bar{\mathcal{K}}\partial_v^2\mathcal{K} - 4\partial_r\mathcal{K}_{ij}\partial_v\mathcal{K}^{ij}.
 \end{aligned} \tag{D.3}$$

Note that, to obtain the last two expressions above we need to first evaluate  $D_\mu D_\nu R_{\alpha\beta}$  where the indices  $\mu, \nu$  runs over the full space-time coordinates:  $v, r, x^i$  and  $D_\mu$  is covariant derivative with respect to the full space-time metric  $g_{\mu\nu}$ .

## D.3 Ricci scalar square theory

Following the discussions in sections 3.1 and section 3.1.1 here we write down the detailed expressions for various quantities for the Ricci Scalar squared theory.

The ‘vv’-component of  $E_{\mu\nu}^{\text{HD}}$  from (3.14)

$$\begin{aligned}
 E_{vv}^{\text{HD}} &= a_1 \left[ \partial_v \left( \frac{1}{\sqrt{h}} \partial_v (\sqrt{h} (3\omega^2 - 4(\nabla \cdot \omega) + 4X)) \right) - 2\mathcal{R}\partial_v\mathcal{K} - 2\partial_v^2\mathcal{R} \right. \\
 & \quad \left. + 8\partial_r\partial_v^2\mathcal{K} + 4\bar{\mathcal{K}}_{ij}\partial_v^2\mathcal{K}^{ij} + 4\bar{\mathcal{K}}\partial_v^2\mathcal{K} + 8\partial_v\bar{\mathcal{K}}_{ij}\partial_v\mathcal{K}^{ij} + 16\partial_v\mathcal{K}\partial_v\bar{\mathcal{K}} \right] + \mathcal{O}[\epsilon^2].
 \end{aligned} \tag{D.4}$$

From (3.15) we know the Wald entropy density as

$$s_w^{\text{HD}} = 2a_1 R, \tag{D.5}$$

and therefore, we immediately obtain

$$\begin{aligned}
 \partial_v \left( \frac{1}{\sqrt{h}} \partial_v (\sqrt{h} s_w^{\text{HD}}) \right) &= 2a_1 \partial_v \left( \frac{1}{\sqrt{h}} \partial_v \left( \sqrt{h} \left( \mathcal{R} - 2X - \frac{3}{2}\omega^2 - 4\partial_r\mathcal{K} \right. \right. \right. \\
 & \quad \left. \left. \left. + 2(\nabla \cdot \omega) - 2\bar{\mathcal{K}}_{ij}\mathcal{K}^{ij} - 2\mathcal{K}\bar{\mathcal{K}} \right) \right) \right).
 \end{aligned} \tag{D.6}$$

We can now use (3.8) and after some algebraic manipulation we obtain

$$\mathbb{E}_{vv}^{\text{HD}*} = \mathcal{O}[\epsilon^2]. \quad (\text{D.7})$$

Finally, comparing with (3.9) we see that for Ricci scalar squared theory there is no spatial entropy current

$$J^i = 0. \quad (\text{D.8})$$

#### D.4 Ricci tensor squared theory

The ‘vv’-component of  $E_{\mu\nu}^{\text{HD}}$  for Ricci tensor squared theory, following (3.14) as discussed in section 3.1.2, comes out to be

$$\begin{aligned} E_{vv}^{\text{HD}} = a_2 \left[ \partial_v \left( \frac{1}{\sqrt{h}} \partial_v \left( \sqrt{h} (\omega^2 + 2X - 2\nabla \cdot \omega + 2\partial_v \bar{\mathcal{K}} + \bar{\mathcal{K}} \mathcal{K} + 2\bar{\mathcal{K}}_{ij} \mathcal{K}^{ij}) \right) \right) \right. \\ \left. + \partial_v (\nabla_i (2h^{ij} \partial_v \omega_j + \omega^i \mathcal{K} - h^{ij} \nabla_j \mathcal{K} - 2\omega_j \mathcal{K}^{ij})) \right. \\ \left. - 2\partial_v (\nabla_i \nabla_j (\mathcal{K}^{ij} - \mathcal{K} h^{ij})) \right] \end{aligned} \quad (\text{D.9})$$

From (3.19) we recognize that the Wald entropy density coming from the higher derivative part of the Lagrangian is

$$s_w^{\text{HD}} = 2a_2 R_{rv} \quad (\text{D.10})$$

and therefore we compute

$$\begin{aligned} \partial_v \left( \frac{1}{\sqrt{h}} \partial_v (\sqrt{h} s_w^{\text{HD}}) \right) = 2a_2 \partial_v \left( \frac{1}{\sqrt{h}} \partial_v \left( \sqrt{h} \left( -X - \frac{1}{2} \omega^2 - \partial_r \mathcal{K} \right. \right. \right. \\ \left. \left. \left. - \bar{\mathcal{K}}_{ij} \mathcal{K}^{ij} + \frac{1}{2} \nabla^i \omega_i \right) \right) \right). \end{aligned} \quad (\text{D.11})$$

Using the definition as given in (3.8) we calculate the following

$$\mathbb{E}_{vv}^{\text{HD}*} = a_2 \partial_v \left[ \frac{1}{\sqrt{h}} \partial_v (\sqrt{h} \bar{\mathcal{K}} \mathcal{K}) + \nabla_i (h^{ij} \nabla_j \mathcal{K} + h^{ij} \partial_v \omega_j - 2\nabla_j \mathcal{K}^{ij}) \right], \quad (\text{D.12})$$

where to derive this we have used the identity

$$\partial_v \left[ \frac{1}{\sqrt{h}} \partial_v (\sqrt{h} \nabla \cdot \omega) \right] = \nabla_i (h^{ij} \partial_v \omega_j - 2\mathcal{K}^{ij} \omega_j + \omega^i \mathcal{K}) \quad (\text{D.13})$$

Therefore the spatial entropy current turns out to be

$$\begin{aligned} J^v &= -s_w^{\text{HD}} - a_2 \bar{\mathcal{K}} \mathcal{K}, \\ J^i &= a_2 (2\nabla_j \mathcal{K}^{ij} - h^{ij} \nabla_j \mathcal{K} - h^{ij} \partial_v \omega_j). \end{aligned} \quad (\text{D.14})$$

## D.5 Riemann tensor squared theory

Following the same steps as followed in the previous subsections for the cases of Ricci scalar squared and Ricci tensor squared theory, we compute the ‘vv’-component of equations of motion for Riemann tensor squared theory, previously discussed in section 3.1.3, as follows

$$E_{vv}^{\text{HD}} = a_3 \left[ \partial_v \left( \frac{1}{\sqrt{h}} \partial_v \left( \sqrt{h} (\omega^2 + 4X - 4\nabla \cdot \omega + 4\bar{\mathcal{K}}_{ij} \mathcal{K}^{ij}) \right) \right) + 4 \partial_v (\nabla_i (2h^{ij} \partial_v \omega_j + \omega^i \mathcal{K} - h^{ij} \nabla_j \mathcal{K} - 2\omega_j \mathcal{K}^{ij})) - 4 \partial_v (\nabla_i \nabla_j (\mathcal{K}^{ij} - \mathcal{K} h^{ij})) \right] \quad (\text{D.15})$$

We take note of the fact that in this case the Wald entropy density for the higher derivative part of the Lagrangian is

$$s_w^{\text{HD}} = -4 a_3 R_{rrvv} \quad (\text{D.16})$$

and using this we compute

$$\partial_v \left( \frac{1}{\sqrt{h}} \partial_v \left( \sqrt{h} s_w^{\text{HD}} \right) \right) = -4 a_3 \partial_v \left( \frac{1}{\sqrt{h}} \partial_v \left( \sqrt{h} \left( X + \frac{1}{4} \omega^2 \right) \right) \right). \quad (\text{D.17})$$

Next we compute  $\mathbb{E}_{vv}^{\text{HD}*}$  defined in (3.8) as given below,

$$\mathbb{E}_{vv}^{\text{HD}*} = 4 a_3 \partial_v \left[ \frac{1}{\sqrt{h}} \partial_v (\sqrt{h} \bar{\mathcal{K}}^{ij} \mathcal{K}_{ij}) + \nabla_i (h^{ij} \partial_v \omega_j - \nabla_j \mathcal{K}^{ij}) \right]. \quad (\text{D.18})$$

Finally we are now at a stage to write down the expressions for the components of the entropy current

$$\begin{aligned} J^v &= -s_w^{\text{HD}} - 4 a_3 \bar{\mathcal{K}}^{ij} \mathcal{K}_{ij}, \\ J^i &= 4 a_3 (\nabla_j \mathcal{K}^{ij} - h^{ij} \partial_v \omega_j) \end{aligned} \quad (\text{D.19})$$

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

## References

- [1] A.C. Wall, *A second law for higher curvature gravity*, *Int. J. Mod. Phys. D* **24** (2015) 1544014 [[arXiv:1504.08040](https://arxiv.org/abs/1504.08040)] [[INSPIRE](#)].
- [2] X.O. Camanho, J.D. Edelstein, J. Maldacena and A. Zhiboedov, *Causality constraints on corrections to the graviton three-point coupling*, *JHEP* **02** (2016) 020 [[arXiv:1407.5597](https://arxiv.org/abs/1407.5597)] [[INSPIRE](#)].
- [3] S.W. Hawking, *Gravitational radiation from colliding black holes*, *Phys. Rev. Lett.* **26** (1971) 1344 [[INSPIRE](#)].
- [4] J.D. Bekenstein, *Black holes and entropy*, *Phys. Rev. D* **7** (1973) 2333 [[INSPIRE](#)].
- [5] S.W. Hawking and J.B. Hartle, *Energy and angular momentum flow into a black hole*, *Commun. Math. Phys.* **27** (1972) 283 [[INSPIRE](#)].

- [6] R.M. Wald, *General relativity*, Chicago Univ. Press, Chicago, IL, U.S.A. (1984).
- [7] R.M. Wald, *Quantum field theory in curved space-time and black hole thermodynamics*, Chicago Lectures in Physics, University of Chicago Press, Chicago, IL, U.S.A. (1995).
- [8] M. Blau, *Lecture notes on general relativity*, <http://www.blau.itp.unibe.ch/Lecturenotes.html>.
- [9] R.M. Wald, *Black hole entropy is the Noether charge*, *Phys. Rev. D* **48** (1993) R3427 [[gr-qc/9307038](#)] [[INSPIRE](#)].
- [10] 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](#)].
- [11] A.C. Wall, *A survey of black hole thermodynamics*, [arXiv:1804.10610](#) [[INSPIRE](#)].
- [12] S. Sarkar, *Black hole thermodynamics: general relativity and beyond*, *Gen. Rel. Grav.* **51** (2019) 63 [[arXiv:1905.04466](#)] [[INSPIRE](#)].
- [13] T. Jacobson and R.C. Myers, *Black hole entropy and higher curvature interactions*, *Phys. Rev. Lett.* **70** (1993) 3684 [[hep-th/9305016](#)] [[INSPIRE](#)].
- [14] T. Jacobson, G. Kang and R.C. Myers, *On black hole entropy*, *Phys. Rev. D* **49** (1994) 6587 [[gr-qc/9312023](#)] [[INSPIRE](#)].
- [15] 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](#)].
- [16] T. Liko, *Topological deformation of isolated horizons*, *Phys. Rev. D* **77** (2008) 064004 [[arXiv:0705.1518](#)] [[INSPIRE](#)].
- [17] S. Sarkar and A.C. Wall, *Second law violations in Lovelock gravity for black hole mergers*, *Phys. Rev. D* **83** (2011) 124048 [[arXiv:1011.4988](#)] [[INSPIRE](#)].
- [18] S. Chatterjee and M. Parikh, *The second law in four-dimensional Einstein-Gauss-Bonnet gravity*, *Class. Quant. Grav.* **31** (2014) 155007 [[arXiv:1312.1323](#)] [[INSPIRE](#)].
- [19] S. Sarkar and A.C. Wall, *Generalized second law at linear order for actions that are functions of Lovelock densities*, *Phys. Rev. D* **88** (2013) 044017 [[arXiv:1306.1623](#)] [[INSPIRE](#)].
- [20] 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](#)].
- [21] 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](#)].
- [22] A.C. Wall, *A proof of the generalized second law for rapidly changing fields and arbitrary horizon slices*, *Phys. Rev. D* **85** (2012) 104049 [[arXiv:1105.3445](#)] [[INSPIRE](#)].
- [23] S. Bhattacharyya et al., *Local fluid dynamical entropy from gravity*, *JHEP* **06** (2008) 055 [[arXiv:0803.2526](#)] [[INSPIRE](#)].
- [24] S. Bhattacharyya, *Constraints on the second order transport coefficients of an uncharged fluid*, *JHEP* **07** (2012) 104 [[arXiv:1201.4654](#)] [[INSPIRE](#)].
- [25] S. Bhattacharyya, *Entropy current from partition function: one example*, *JHEP* **07** (2014) 139 [[arXiv:1403.7639](#)] [[INSPIRE](#)].
- [26] S. Bhattacharyya, *Entropy current and equilibrium partition function in fluid dynamics*, *JHEP* **08** (2014) 165 [[arXiv:1312.0220](#)] [[INSPIRE](#)].

- [27] S. Gao and R.M. Wald, *The ‘physical process’ version of the first law and the generalized second law for charged and rotating black holes*, *Phys. Rev. D* **64** (2001) 084020 [[gr-qc/0106071](#)] [[INSPIRE](#)].
- [28] A.J. Amsel, D. Marolf and A. Virmani, *The physical process first law for bifurcate Killing horizons*, *Phys. Rev. D* **77** (2008) 024011 [[arXiv:0708.2738](#)] [[INSPIRE](#)].
- [29] S. Bhattacharjee and S. Sarkar, *Physical process first law and caustic avoidance for Rindler horizons*, *Phys. Rev. D* **91** (2015) 024024 [[arXiv:1412.1287](#)] [[INSPIRE](#)].
- [30] A. Mishra, S. Chakraborty, A. Ghosh and S. Sarkar, *On the physical process first law for dynamical black holes*, *JHEP* **09** (2018) 034 [[arXiv:1709.08925](#)] [[INSPIRE](#)].
- [31] 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](#)].
- [32] S. Kolekar, T. Padmanabhan and S. Sarkar, *Entropy increase during physical processes for black holes in Lanczos-Lovelock gravity*, *Phys. Rev. D* **86** (2012) 021501 [[arXiv:1201.2947](#)] [[INSPIRE](#)].
- [33] S. Chapman, Y. Neiman and Y. Oz, *Fluid/gravity correspondence, local Wald entropy current and gravitational anomaly*, *JHEP* **07** (2012) 128 [[arXiv:1202.2469](#)] [[INSPIRE](#)].
- [34] C. Eling, A. Meyer and Y. Oz, *Local entropy current in higher curvature gravity and Rindler hydrodynamics*, *JHEP* **08** (2012) 088 [[arXiv:1205.4249](#)] [[INSPIRE](#)].
- [35] Y. Dandekar and A. Saha, *Large D membrane for higher derivative gravity and black hole second law*, *JHEP* **02** (2020) 083 [[arXiv:1910.10964](#)] [[INSPIRE](#)].
- [36] 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](#)].
- [37] S. Kolekar, T. Padmanabhan and S. Sarkar, *Entropy increase during physical processes for black holes in Lanczos-Lovelock gravity*, *Phys. Rev. D* **86** (2012) 021501 [[arXiv:1201.2947](#)] [[INSPIRE](#)].