Entanglement equilibrium for higher order gravity

Pablo Bueno, Vincent S. Min, Antony J. Speranza, and Manus R. Visser

Phys. Rev. D **95**, 046003 — Published 10 February 2017
DOI: 10.1103/PhysRevD.95.046003

Entanglement equilibrium for higher order gravity

Pablo Bueno,[1,2,*] Vincent S. Min,[1,†] Antony J. Speranza,[3,‡] and Manus R. Visser[2,§]

[1] *Instituut voor Theoretische Fysica, KU Leuven, Celestijnenlaan 200D, B-3001 Leuven, Belgium*
[2] *Institute for Theoretical Physics, University of Amsterdam, 1090 GL Amsterdam, The Netherlands*
[3] *Maryland Center for Fundamental Physics, University of Maryland, College Park, MD 20742, USA*

We show that the linearized higher derivative gravitational field equations are equivalent to an equilibrium condition on the entanglement entropy of small spherical regions in vacuum. This extends Jacobson's recent derivation of the Einstein equation using entanglement to include general higher derivative corrections. The corrections are naturally associated with the subleading divergences in the entanglement entropy, which take the form of a Wald entropy evaluated on the entangling surface. Variations of this Wald entropy are related to the field equations through an identity for causal diamonds in maximally symmetric spacetimes, which we derive for arbitrary higher derivative theories. If the variations are taken holding fixed a geometric functional that we call the generalized volume, the identity becomes an equivalence between the linearized constraints and the entanglement equilibrium condition. We note that the fully nonlinear higher curvature equations cannot be derived from the linearized equations applied to small balls, in contrast to the situation encountered in Einstein gravity. The generalized volume is a novel result of this work, and we speculate on its thermodynamic role in the first law of causal diamond mechanics, as well as its possible application to holographic complexity.

I. INTRODUCTION

Black hole entropy remains one of the best windows into the nature of quantum gravity available to dwellers of the infrared. Bekenstein's original motivation for introducing it was to avoid gross violations of the second law of thermodynamics by sending matter into the black hole, decreasing the entropy of the exterior [1, 2]. The subsequent discovery by Hawking that black holes radiate thermally at a temperature $T = \kappa/2\pi$, with κ the surface gravity, fixed the value of the entropy in terms of the area to be $S_{\rm BH} = A/4G$, and suggested a deep connection to quantum properties of gravity [3].

The appearance of area in $S_{\rm BH}$ is somewhat mysterious from a classical perspective; however, an intriguing explanation emerges by considering the entanglement entropy of quantum fields outside the horizon [4–7]. Entanglement entropy is UV divergent, and upon regulation it takes the form

$$S_{\rm EE} = c_0 \frac{A}{\epsilon^{d-2}} + \{\text{subleading divergences}\} + S_{\rm finite}\,, \quad (1)$$

with ϵ a regulator and c_0 a constant. Identifying the coefficient c_0/ϵ^{d-2} with $1/4G$ would allow $S_{\rm BH}$ to be attributed to the leading divergence in the entanglement entropy. The subleading divergences could similarly be associated with higher curvature gravitational couplings, which change the expression for the black hole entropy to the Wald entropy [8].

To motivate these identifications, one must assume that the quantum gravity theory is UV finite (as occurs in string theory), yielding a finite entanglement entropy, cut off near the Planck length, $\epsilon \sim \ell_P$. Implementing this cutoff would seem to depend on a detailed knowledge of the UV theory, inaccessible from the vantage of low energy effective field theory. Interestingly, this issue can be resolved within the effective theory by the renormalization of the gravitational couplings by matter loop divergences. There is mounting evidence that these precisely match the entanglement entropy divergences, making the *generalized entropy*

$$S_{\rm gen} = S^{(\epsilon)}_{\rm Wald} + S^{(\epsilon)}_{\rm mat} \quad (2)$$

independent of ϵ [9–12]. Here $S^{(\epsilon)}_{\rm Wald}$ is the Wald entropy expressed in terms of the renormalized gravitational couplings and $S^{(\epsilon)}_{\rm mat}$ is a renormalized entanglement entropy of matter fields that is related to $S_{\rm finite}$ in (1), although the precise relation depends on the renormalization scheme.[1] The identification of gravitational couplings with entanglement entropy divergences is therefore consistent with the renormalization group (RG) flow in the low energy effective theory, and amounts to assuming that the bare gravitational couplings vanish [13]. In this picture, $S_{\rm gen} = S_{\rm EE}$, with $S^{(\epsilon)}_{\rm Wald}$ acting as a placeholder for the UV degrees of freedom that have been integrated out.

When viewed as entanglement entropy, it is clear that generalized entropy can be assigned to surfaces other

* pablo@itf.fys.kuleuven.be
† vincent.min@kuleuven.be
‡ asperanz@gmail.com
§ m.r.visser@uva.nl

[1] A covariant regulator must be used to ensure that the subleading divergences appear as a Wald entropy. Also, since power law divergences are not universal, when they are present the same renormalization scheme must be used for the entanglement entropy and the gravitational couplings. Additional subtleties for nonminimally coupled fields, gauge fields, and gravitons are discussed in Section V D 3.

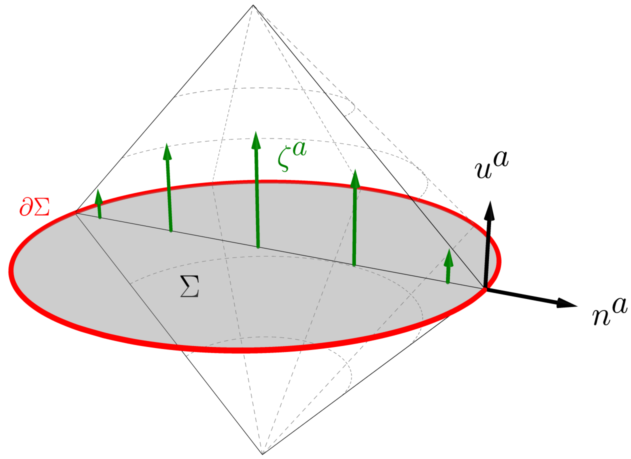

FIG. 1. The causal diamond consists of the future and past domains of dependence of a spatial sphere Σ in a MSS. Σ has a unit normal u^a, induced metric h_{ab}, and volume form η. The boundary $\partial\Sigma$ has a spacelike unit normal n^a, defined to be orthogonal to u^a, and volume form μ. The conformal Killing vector ζ^a generates a flow within the causal diamond, and vanishes on the bifurcation surface $\partial\Sigma$.

than black hole horizon cross sections [12, 14–16]. For example, in holography the generalized entropy of a minimal surface in the bulk is dual via the quantum-corrected Ryu-Takayanagi formula [17, 18] to the entanglement entropy of a region of the boundary CFT.[2] Even without assuming holographic duality, the generalized entropy provides a link between the geometry of surfaces and entanglement entropy. When supplemented with thermodynamic information, this link can give rise to dynamical equations for gravity. The first demonstration of this was Jacobson's derivation of the Einstein equation as an equation of state for local causal horizons possessing an entropy proportional to their area [19]. Subsequent work using entropic arguments [20, 21] and holographic entanglement entropy [22–24] confirmed that entanglement thermodynamics is connected to gravitational dynamics.

Recently, Jacobson has advanced a new viewpoint on the relation between geometry and entanglement that has been dubbed "entanglement equilibrium" [25]. This proposal considers spherical, spatial subregions in geometries that are a perturbation of a maximally symmetric spacetime (MSS). Each such subregion Σ in the maximally symmetric background defines a causal diamond, which admits a conformal Killing vector ζ^a whose flow preserves the diamond (see Figure 1). The entanglement equilibrium hypothesis states that any perturbation of the matter fields and geometry inside the ball leads to a decrease in entanglement, i.e., the vacuum is a maximal entropy state. This hypothesis applies holding the volume of Σ fixed; even so, the introduction of curvature from the geometry variation can lead to a decrease in the area of the boundary $\partial\Sigma$. This affects the divergent terms in the entanglement entropy by changing Wald entropy, which at leading order is simply $A/4G$. The variation of the quantum state contributes a piece δS_{mat}, and maximality implies the total variation of the entanglement entropy vanishes at first order,[3]

$$\delta S_{\mathrm{EE}}\big|_V = \frac{\delta A\big|_V}{4G} + \delta S_{\mathrm{mat}} = 0\,. \tag{3}$$

When applied to small spheres, this maximal entropy condition was shown to be equivalent to imposing the Einstein equation at the center of the ball.

Taken as an effective field theory, gravity is expected to contain higher curvature corrections that arise from matching to its UV completion. An important test of the entanglement equilibrium hypothesis is whether it can consistently accommodate these corrections. It is the purpose of this paper to demonstrate that a generalization to higher curvature theories is possible, and relates to the subleading divergences appearing in (1).

A. Summary of results and outline

It is not *a priori* clear what the precise statement of the entanglement equilibrium condition should be for a higher curvature theory, and in particular what replaces the fixed-volume constraint. The formulation we propose here is advised by the *first law of causal diamond mechanics*, a purely geometrical identity that holds independently of any entanglement considerations. It was derived for Einstein gravity in the supplemental materials of [25], and one of the main results of this paper is to extend it to arbitrary, higher derivative theories. As we show in section II, the first law is related to the off-shell identity

$$\frac{\kappa}{2\pi}\delta S_{\mathrm{Wald}}\big|_W + \delta H^m_\zeta = \int_\Sigma \delta C_\zeta\,, \tag{4}$$

where κ is the surface gravity of ζ^a [28], S_{Wald} is the Wald entropy of $\partial\Sigma$ given in equation (24) [8, 29], H^m_ζ is the matter Hamiltonian for flows along ζ^a, defined in equation (9), and $\delta C_\zeta = 0$ are the linearized constraint

[2] The UV divergences in the CFT entanglement entropy have no relation to the Planck length in the bulk, but instead are related to the infinite area of the minimal surface in AdS, courtesy of the UV/IR correspondence.

[3] The separation of the entanglement entropy into a divergent Wald piece and a finite matter piece is scheme dependent, and can change under the RG flow [26]. Also the matter variation can sometimes produce state-dependent divergences [27], which appear as a variation of the Wald entropy. Since we only ever deal with total variations of the generalized entropy, these subtleties do not affect any results. For simplicity, we will refer to δS_{Wald} as coming from the geometry variation, and δS_{mat} from the matter state variation.

equations of the higher derivative theory. The Wald entropy is varied holding fixed a local geometric functional

$$W = \frac{1}{(d-2)E_0}\int_\Sigma \eta\left(E^{abcd}u_a h_{bc} u_d - E_0\right)\,, \qquad (5)$$

with η, u^a and h_{ab} defined in Figure 1. E^{abcd} is the variation of the gravitational Lagrangian scalar with respect to R_{abcd}, and E_0 is a constant determined by the value of E^{abcd} in a MSS via $E^{abcd} \overset{\text{MSS}}{=} E_0(g^{ac}g^{bd} - g^{ad}g^{bc})$. We refer to W as the "generalized volume" since it reduces to the volume for Einstein gravity.

The Wald formalism contains ambiguities identified by Jacobson, Kang and Myers (JKM) [30] that modify the Wald entropy and the generalized volume by the terms S_{JKM} and W_{JKM} given in (43) and (44). Using a modified generalized volume defined by

$$W' = W + W_{\text{JKM}}\,, \qquad (6)$$

the identity (4) continues to hold with $\delta(S_{\text{Wald}} + S_{\text{JKM}})\big|_{W'}$ replacing $\delta S_{\text{Wald}}\big|_W$. As discussed in section III A, the subleading divergences for the entanglement entropy involve a particular resolution of the JKM ambiguity, while section II D argues that the first law of causal diamond mechanics applies for *any* resolution, as long as the appropriate generalized volume is held fixed.

Using the resolution of the JKM ambiguity required for the entanglement entropy calculation, the first law leads to the following statement of entanglement equilibrium, applicable to higher curvature theories:

Hypothesis (Entanglement Equilibrium)**.** *In a quantum gravitational theory, the entanglement entropy of a spherical region with fixed generalized volume W' is maximal in vacuum.*

This modifies the original equilibrium condition (3) by replacing the area variation with

$$\delta(S_{\text{Wald}} + S_{\text{JKM}})\big|_{W'}\,. \qquad (7)$$

In Section III, this equilibrium condition is shown to be equivalent to the linearized higher derivative field equations in the case that the matter fields are conformally invariant.[4] Facts about entanglement entropy divergences and the reduced density matrix for a sphere in a CFT are used to relate the total variation of the entanglement entropy to the left hand side of (4). Once this is done, it becomes clear that imposing the linearized constraint equations is equivalent to the entanglement equilibrium condition.

In [25], this condition was applied in the small ball limit, in which *any* geometry looks like a perturbation of a MSS. Using Riemann normal coordinates (RNC), the linearized equations were shown to impose the fully nonlinear equations for the case of Einstein gravity. We will discuss this argument in Section IV for higher curvature theories, and show that the nonlinear equations can *not* be obtained from the small ball limit, making general relativity unique in that regard.

In section V, we discuss several implications of this work. First, we describe how it compares to other approaches connecting geometry and entanglement. Following that, we provide a possible thermodynamic interpretation of the first law of causal diamond mechanics derived in section II. We then comment on a conjectural relation between our generalized volume W and higher curvature holographic complexity. Finally, we lay out several future directions for the entanglement equilibrium program.

Note on conventions: we set $\hbar = c = 1$, use metric signature $(-,+,+,\ldots)$, and use d to refer to the spacetime dimension. We write the spacetime volume form as ϵ, and occasionally we will denote it ϵ_a or ϵ_{ab}, suppressing all but its first one or two abstract indices.

II. FIRST LAW OF CAUSAL DIAMOND MECHANICS

Jacobson's entanglement equilibrium argument [25] compares the surface area of a small spatial ball Σ in a curved spacetime to the one that would be obtained in a MSS. The comparison is made using balls of equal volume V, a choice justified by an Iyer-Wald variational identity [29] for the conformal Killing vector ζ^a of the causal diamond in the maximally symmetric background. When the Einstein equation holds, this identity implies the *first law of causal diamond mechanics* [25, 33]

$$-\delta H^m_\zeta = \frac{\kappa}{8\pi G}\delta A - \frac{\kappa k}{8\pi G}\delta V\,, \qquad (8)$$

where k is the trace of the extrinsic curvature of $\partial\Sigma$ embedded in Σ, and the matter conformal Killing energy H^m_ζ is constructed from the stress tensor T_{ab} by

$$H^m_\zeta = \int_\Sigma \eta\, u^a \zeta^b T_{ab}\,. \qquad (9)$$

The purpose of this section is to generalize the variational identity to higher derivative theories, and to clarify its relation to the equations of motion. This is done by focusing on an off-shell version of the identity, which reduces to the first law when the linearized constraint equations for the theory are satisfied. We begin by reviewing the Iyer-Wald formalism in subsection II A, which also serves to establish notation. After describing the geometric setup in subsection II B, we show in subsection II C how the quantities appearing in the identity can be written as variations of local geometric functionals of the surface Σ and its boundary $\partial\Sigma$. As one might expect, the area is

[4] There is a proposal for including nonconformal matter that involves varying a local cosmological constant [25, 31, 32]. If valid, this proposal applies in the higher curvature case as well, since it deals only with the matter variations.

upgraded to the Wald entropy S_{Wald}, and we derive the generalization of the volume given in equation (5). Subsection II D describes how the variational identity can instead be viewed as a variation at fixed generalized volume W, as quoted in equation (4), and describes the effect that JKM ambiguities have on the setup.

A. Iyer-Wald formalism

We begin by recalling the Iyer-Wald formalism [8, 29]. A general diffeomorphism invariant theory may be defined by its Lagrangian $L[\phi]$, a spacetime d-form locally constructed from the dynamical fields ϕ, which include the metric and matter fields. A variation of this Lagrangian takes the form

$$\delta L = E \cdot \delta\phi + \mathrm{d}\theta[\delta\phi]\,, \tag{10}$$

where E collectively denotes the equations of motion for the dynamical fields, and θ is the symplectic potential $(d-1)$-form. Taking an antisymmetric variation of θ yields the symplectic current $(d-1)$-form

$$\omega[\delta_1\phi, \delta_2\phi] = \delta_1\theta[\delta_2\phi] - \delta_2\theta[\delta_1\phi]\,, \tag{11}$$

whose integral over a Cauchy surface Σ gives the symplectic form for the phase space description of the theory. Given an *arbitrary* vector field ζ^a, evaluating the symplectic form on the Lie derivative $\pounds_\zeta\phi$ gives the variation of the Hamiltonian H_ζ that generates the flow of ζ^a

$$\delta H_\zeta = \int_\Sigma \omega[\delta\phi, \pounds_\zeta\phi]\,. \tag{12}$$

Now consider a ball-shaped region Σ, and take ζ^a to be any future-pointed, timelike vector that vanishes on the boundary $\partial\Sigma$. Wald's variational identity then reads

$$\int_\Sigma \omega[\delta\phi, \pounds_\zeta\phi] = \int_\Sigma \delta J_\zeta\,, \tag{13}$$

where the Noether current J_ζ is defined by

$$J_\zeta = \theta[\pounds_\zeta\phi] - i_\zeta L\,. \tag{14}$$

Here i_ζ denotes contraction of the vector ζ^a on the first index of the differential form L. The identity (13) holds when the background geometry satisfies the field equations $E = 0$, and it assumes that ζ^a vanishes on $\partial\Sigma$. Next we note that the Noether current can always be expressed as [34]

$$J_\zeta = \mathrm{d}Q_\zeta + C_\zeta, \tag{15}$$

where Q_ζ is the Noether charge $(d-2)$-form and C_ζ are the constraint field equations, which arise as a consequence of the diffeomorphism gauge symmetry. For non-scalar matter, these constraints are a combination of the metric and matter field equations [35, 36], but, assuming the matter equations are imposed, we can take $C_\zeta = -2\zeta^a E_a{}^b\epsilon_b$, where E^{ab} is the variation of the Lagrangian density with respect to the metric. By combining equations (12), (13) and (15), one finds that

$$-\int_{\partial\Sigma} \delta Q_\zeta + \delta H_\zeta = \int_\Sigma \delta C_\zeta\,. \tag{16}$$

When the linearized constraints hold, $\delta C_\zeta = 0$, the variation of the Hamiltonian is a boundary integral of δQ_ζ. This on-shell identity forms the basis for deriving the first law of causal diamond mechanics. Unlike the situation encountered in black hole thermodynamics, δH_ζ is not zero because below we take ζ^a to be a conformal Killing vector as opposed to a true Killing vector.

B. Geometric setup

Thus far, the only restriction that has been placed on the vector field ζ^a is that it vanishes on $\partial\Sigma$. As such, the quantities δH_ζ and δQ_ζ appearing in the identities depend rather explicitly on the fixed vector ζ^a, and therefore these quantities are not written in terms of only the geometric properties of the surfaces Σ and $\partial\Sigma$. A purely geometric description is desirable if the Hamiltonian and Noether charge are to be interpreted as thermodynamic state functions, which ultimately may be used to define the ensemble of geometries in any proposed quantum description of the microstates. This situation may be remedied by choosing the vector ζ^a and the surface Σ to have special properties in the background geometry. In particular, by choosing ζ^a to be a conformal Killing vector for a causal diamond in the MSS, and picking Σ to lie on the surface where the conformal factor vanishes, one finds that the perturbations δH_ζ and δQ_ζ have expressions in terms of local geometric functionals on the surfaces Σ and $\partial\Sigma$, respectively.

Given a causal diamond in a MSS, there exists a conformal Killing vector ζ^a which generates a flow within the diamond and vanishes at the bifurcation surface $\partial\Sigma$ (see figure 1). The metric satisfies the conformal Killing equation

$$\pounds_\zeta g_{ab} = 2\alpha g_{ab} \quad \text{with} \quad \alpha = \frac{1}{d}\nabla_c\zeta^c\,. \tag{17}$$

and the conformal factor α vanishes on the spatial ball Σ. The gradient of α is hence proportional to the unit normal to Σ,

$$u_a = N\nabla_a\alpha \quad \text{with} \quad N = \|\nabla_a\alpha\|^{-1}. \tag{18}$$

Note the vector u^a is future pointing since the conformal factor α decreases to the future of Σ. In a MSS, the normalization function N has the curious property that it is constant over Σ, and is given by [33]

$$N = \frac{d-2}{\kappa k}, \tag{19}$$

where k is the trace of the extrinsic curvature of $\partial\Sigma$ embedded in Σ, and κ is the surface gravity of the conformal Killing horizon, defined momentarily. This constancy ends up being crucial to finding a local geometric functional for δH_ζ. Throughout this work, N and k will respectively denote constants equal to the normalization function and extrinsic curvature trace, both evaluated in the background spacetime.

Since α vanishes on Σ, ζ^a is instantaneously a Killing vector. On the other hand, the covariant derivative of α is nonzero, so

$$\nabla_d(\pounds_\zeta g_{ab})\big|_\Sigma = \frac{2}{N}u_d g_{ab}\,. \tag{20}$$

The fact that the covariant derivative is nonzero on Σ is responsible for making δH_ζ nonvanishing.

A conformal Killing vector with a horizon has a well-defined surface gravity κ [28], and since α vanishes on $\partial\Sigma$, we can conclude that

$$\nabla_a\zeta_b\big|_{\partial\Sigma} = \kappa n_{ab}\,, \tag{21}$$

where $n_{ab} = 2u_{[a}n_{b]}$ is the binormal for the surface $\partial\Sigma$, and n^b is the outward pointing spacelike unit normal to $\partial\Sigma$. Since $\partial\Sigma$ is a bifurcation surface of a conformal Killing horizon, κ is constant everywhere on it. We provide an example of these constructions in appendix A where we discuss the conformal Killing vector for a causal diamond in flat space.

C. Local geometric expressions

In this subsection we evaluate the Iyer-Wald identity (16) for an arbitrary higher derivative theory of gravity and for the geometric setup described above. The final on-shell result is given in (37), which is the first law of causal diamond mechanics for higher derivative gravity.

Throughout the computation we assume that the matter fields are minimally coupled, so that the Lagrangian splits into a metric and matter piece $L = L^g + L^m$, and we take L^g to be an *arbitrary*, diffeomorphism-invariant function of the metric, Riemann tensor, and its covariant derivatives. The symplectic potential and variation of the Hamiltonian then exhibit a similar separation, $\theta = \theta^g + \theta^m$ and $\delta H_\zeta = \delta H_\zeta^g + \delta H_\zeta^m$, and so we can write equation (16) as

$$-\int_{\partial\Sigma}\delta Q_\zeta + \delta H_\zeta^g + \delta H_\zeta^m = \int_\Sigma \delta C_\zeta\,. \tag{22}$$

Below, we explicitly compute the two terms δH_ζ^g and $\int_{\partial\Sigma}\delta Q_\zeta$ for the present geometric context.

a. Wald entropy. By virtue of equation (21) and the fact that ζ^a vanishes on $\partial\Sigma$, one can show that the integrated Noether charge is simply related to the Wald entropy [8, 29]

$$\begin{aligned}-\int_{\partial\Sigma}Q_\zeta &= \int_{\partial\Sigma}E^{abcd}\,\epsilon_{ab}\nabla_c\zeta_d\\ &= \frac{\kappa}{2\pi}S_{\text{Wald}}\,,\end{aligned} \tag{23}$$

where the Wald entropy is defined as

$$S_{\text{Wald}} = -2\pi\int_{\partial\Sigma}\mu\,E^{abcd}n_{ab}n_{cd}\,. \tag{24}$$

E^{abcd} is the variation of the Lagrangian scalar with respect to the Riemann tensor R_{abcd} taken as an independent field, given in (B2), and μ is the volume form on $\partial\Sigma$, so that $\epsilon_{ab} = -n_{ab}\wedge\mu$ there. The equality (23) continues to hold at first order in perturbations, which can be shown following the same arguments as given in [29], hence,

$$\int_{\partial\Sigma}\delta Q_\zeta = -\frac{\kappa}{2\pi}\delta S_{\text{Wald}}\,. \tag{25}$$

The minus sign is opposite the convention in [29] since the unit normal n^a is outward pointing for the causal diamond.

b. Generalized volume. The gravitational part of δH_ζ is related to the symplectic current $\omega[\delta g, \pounds_\zeta g]$ via (12). The symplectic form has been computed on an arbitrary background for any higher curvature gravitational theory whose Lagrangian is a function of the Riemann tensor, but not its covariant derivatives [37]. Here, we take advantage of the maximal symmetry of the background to compute the symplectic form and Hamiltonian for the causal diamond in any higher order theory, including those with derivatives of the Riemann tensor.

Recall that the symplectic current ω is defined in terms of the symplectic potential θ through (11). For a Lagrangian that depends on the Riemann tensor and its covariant derivatives, the symplectic potential θ^g is given in Lemma 3.1 of [29]

$$\begin{aligned}\theta^g &= 2E^{bcd}\nabla_d\delta g_{bc} + S^{ab}\delta g_{ab}\\ &\quad+\sum_{i=1}^{m-1}T_i^{abcda_1\ldots a_i}\delta\nabla_{(a_1}\cdots\nabla_{a_i)}R_{abcd}\,,\end{aligned} \tag{26}$$

where $E^{bcd} = \epsilon_a E^{abcd}$ and the tensors S^{ab} and $T_i^{abcda_1\ldots a_i}$ are locally constructed from the metric, its curvature, and covariant derivatives of the curvature. Due to the antisymmetry of E^{bcd} in c and d, the symplectic current takes the form

$$\begin{aligned}\omega^g &= 2\delta_1 E^{bcd}\nabla_d\delta_2 g_{bc} - 2E^{bcd}\delta_1\Gamma^e_{db}\delta_2 g_{ec} + \delta_1 S^{ab}\delta_2 g_{ab}\\ &+\sum_{i=1}^{m-1}\delta_1 T_i^{abcda_1\ldots a_i}\delta_2\nabla_{(a_1}\cdots\nabla_{a_i)}R_{abcd} - (1\leftrightarrow 2).\end{aligned} \tag{27}$$

Next we specialize to the geometric setup described in section II B. We may thus employ the fact that we are perturbing around a maximally symmetric background. This means the background curvature tensor takes the form

$$R_{abcd} = \frac{R}{d(d-1)}(g_{ac}g_{bd} - g_{ad}g_{bc}) \tag{28}$$

with a constant Ricci scalar R, so that $\nabla_e R_{abcd} = 0$, and also $\pounds_\zeta R_{abcd}\big|_\Sigma = 0$. Since the tensors E^{abcd}, S^{ab}, and $T_i^{abcda_1\ldots a_i}$ are all constructed from the metric and curvature, they will also have vanishing Lie derivative along ζ^a when evaluated on Σ.

Replacing $\delta_2 g_{ab}$ in equation (27) with $\pounds_\zeta g_{ab}$ and using (20), we obtain

$$\omega^g[\delta g, \pounds_\zeta g]\big|_\Sigma = \frac{2}{N}\left[2g_{bc}u_d\delta E^{bcd} + E^{bcd}(u_d\delta g_{bc} - g_{bd}u^e\delta g_{ec})\right]. \tag{29}$$

We would like to write this as a variation of some scalar quantity. To do so, we split off the background value of E^{abcd} by writing

$$F^{abcd} = E^{abcd} - E_0(g^{ac}g^{bd} - g^{ad}g^{bc}). \tag{30}$$

The second term in this expression is the background value, and, due to maximal symmetry, the scalar E_0 must be a constant determined by the parameters appearing in the Lagrangian. By definition, F^{abcd} is zero in the background, so any term in (29) that depends on its variation may be immediately written as a total variation, since variations of other tensors appearing in the formula would multiply the background value of F^{abcd}, which vanishes. Hence, the piece involving δF^{abcd} becomes

$$\frac{4}{N}g_{bc}u_d\delta(F^{abcd}\epsilon_a) = \frac{4}{N}\delta(F^{abcd}g_{bc}u_d\epsilon_a). \tag{31}$$

The remaining terms simply involve replacing E^{abcd} in (29) with $E_0(g^{ac}g^{bd} - g^{ad}g^{bc})$. These terms then take exactly the same form as the terms that appear for general relativity, which we know from the appendix of [25] combine to give an overall variation of the volume. The precise form of this variation when restricted to Σ is

$$-\frac{4(d-2)}{N}\delta\eta, \tag{32}$$

where η is the induced volume form on Σ. Adding this to (31) produces

$$\omega[\delta g, \pounds_\zeta g]\big|_\Sigma = -\frac{4}{N}\delta\left[\eta(E^{abcd}u_a u_d h_{bc} - E_0)\right], \tag{33}$$

where we used that $\epsilon_a = -u_a \wedge \eta$ on Σ. This leads us to define a generalized volume functional

$$W = \frac{1}{(d-2)E_0}\int_\Sigma \eta(E^{abcd}u_a u_d h_{bc} - E_0), \tag{34}$$

and the variation of this quantity is related to the variation of the gravitational Hamiltonian by

$$\delta H_\zeta^g = -4E_0\kappa k\,\delta W, \tag{35}$$

where we have expressed N in terms of κ and k using (19). We have thus succeeded in writing δH_ζ^g in terms of a local geometric functional defined on the surface Σ.

It is worth emphasizing that N being constant over the ball was crucial to this derivation, since otherwise it could not be pulled out of the integral over Σ and would define a non-diffeomorphism invariant structure on the surface. Note that the overall normalization of W is arbitrary, since a different normalization would simply change the coefficient in front of δW in (35). As one can readily check, the normalization in (34) was chosen so that W reduces to the volume in the case of Einstein gravity. In appendix B we provide explicit expressions for the generalized volume in $f(R)$ gravity and quadratic gravity.

Finally, combining (25), (35) and (22), we arrive at the off-shell variational identity in terms of local geometric quantities

$$\frac{\kappa}{2\pi}\delta S_{\text{Wald}} - 4E_0\kappa k\delta W + \delta H_\zeta^m = \int_\Sigma \delta C_\zeta. \tag{36}$$

By imposing the linearized constraints $\delta C_\zeta = 0$, this becomes the first law of causal diamond mechanics for higher derivative gravity

$$-\delta H_\zeta^m = \frac{\kappa}{2\pi}\delta S_{\text{Wald}} - 4E_0\kappa k\delta W. \tag{37}$$

This reproduces (8) for Einstein gravity with Lagrangian $L = \epsilon R/16\pi G$, for which $E_0 = 1/32\pi G$.

D. Variation at fixed W

We now show that the first two terms in (36) can be written in terms of the variation of the Wald entropy at fixed W, defined as

$$\delta S_{\text{Wald}}\big|_W = \delta S_{\text{Wald}} - \frac{\partial S_{\text{Wald}}}{\partial W}\delta W. \tag{38}$$

Here we must specify what is meant by $\frac{\partial S_{\text{Wald}}}{\partial W}$. We will take this partial derivative to refer to the changes that occur in both quantities when the size of the ball is deformed, but the metric and dynamical fields are held fixed. Take a vector v^a that is everywhere tangent to Σ that defines an infinitesimal change in the shape of Σ. The first order change this produces in S_{Wald} and W can be computed by holding Σ fixed, but varying the Noether current and Noether charge as $\delta J_\zeta = \pounds_v J_\zeta$ and $\delta Q_\zeta = \pounds_v Q_\zeta$. Since the background field equations are satisfied and ζ^a vanishes on $\partial\Sigma$, we have there that $\int_{\partial\Sigma} Q_\zeta = \int_\Sigma J_\zeta^g$, without reference to the matter part of the Noether current. Recall that δW is related to the variation of the gravitational Hamiltonian, which can be expressed in terms of δJ^g_ζ through (12) and (13). Then using the relations (23) and (35) and the fact that the Lie derivative commutes with the exterior derivative, we may compute

$$\frac{\partial S_{\rm Wald}}{\partial W} = \frac{-\frac{2\pi}{\kappa}\int_{\partial\Sigma}\pounds_v Q_\zeta}{-\frac{1}{4E_0\kappa k}\int_\Sigma \pounds_v J^g_\zeta} = 8\pi E_0 k\,. \tag{39}$$

Combining this result with equations (37) and (38) we arrive at the off-shell variational identity for higher derivative gravity quoted in the introduction

$$\frac{\kappa}{2\pi}\delta S_{\rm Wald}\big|_W + \delta H^m_\zeta = \int_\Sigma \delta C_\zeta\,. \tag{40}$$

Finally, we comment on how JKM ambiguities [30] affect this identity. The particular ambiguity we are concerned with comes from the fact that the symplectic potential θ in equation (10) is defined only up to addition of an exact form $\mathrm{d}Y[\delta\phi]$ that is linear in the field variations and their derivatives. This has the effect of changing the Noether current and Noether charge by

$$J_\zeta \to J_\zeta + \mathrm{d}Y[\pounds_\zeta\phi]\,, \tag{41}$$
$$Q_\zeta \to Q_\zeta + Y[\pounds_\zeta\phi]\,. \tag{42}$$

This modifies both the entropy and the generalized volume by surface terms on $\partial\Sigma$ given by

$$S_{\rm JKM} = -\frac{2\pi}{\kappa}\int_{\partial\Sigma} Y[\pounds_\zeta\phi]\,, \tag{43}$$
$$W_{\rm JKM} = -\frac{1}{4E_0\kappa k}\int_{\partial\Sigma} Y[\pounds_\zeta\phi]\,. \tag{44}$$

However, it is clear that this combined change in J_ζ and Q_ζ leaves the left hand side of (40) unchanged, since the Y-dependent terms cancel out. In particular,

$$\delta S_{\rm Wald}\big|_W = \delta(S_{\rm Wald} + S_{\rm JKM})\big|_{W+W_{\rm JKM}}\,, \tag{45}$$

showing that any resolution of the JKM ambiguity gives the same first law, provided that the Wald entropy and generalized volume are modified by the terms (43) and (44). This should be expected, because the right hand side of (40) depends only on the field equations, which are unaffected by JKM ambiguities.

III. ENTANGLEMENT EQUILIBRIUM

The original entanglement equilibrium argument for Einstein gravity stated that the total variation away from the vacuum of the entanglement of a region at fixed volume is zero. This statement is encapsulated in equation (3), which shows both an area variation due to the change in geometry, and a matter piece from varying the quantum state. The area variation at fixed volume can equivalently be written

$$\delta A\big|_V = \delta A - \frac{\partial A}{\partial V}\delta V \tag{46}$$

and the arguments of section II D relate this combination to the terms appearing in the first law of causal diamond mechanics (8). Since δH^m_ζ in (8) is related to $\delta S_{\rm mat}$ in (3) for conformally invariant matter, the first law may be interpreted entirely in terms of entanglement entropy variations.

This section discusses the extension of the argument to higher derivative theories of gravity. Subsection III A explains how subleading divergences in the entanglement entropy are related to a Wald entropy, modified by a particular resolution of the JKM ambiguity. Paralleling the Einstein gravity derivation, we seek to relate variations of the subleading divergences to the higher derivative first law of causal diamond mechanics (37). Subsection III B shows that this can be done as long as the generalized volume W' [related to W by a boundary JKM term as in (6)] is held fixed. Then, using the relation of the first law to the off-shell identity (40), we discuss how the entanglement equilibrium condition is equivalent to imposing the linearized constraint equations.

A. Subleading entanglement entropy divergences

The structure of divergences in entanglement entropy is reviewed in [11] and the appendix of [12]. It is well-known that the leading divergence depends on the area of the entangling surface. More surprising, however, is the fact that this divergence precisely matches the matter field divergences that renormalize Newton's constant. This ostensible coincidence arises because the two divergences have a common origin in the gravitational effective action $I_{\rm eff}$, which includes both gravitational and matter pieces. Its relation to entanglement entropy comes from the replica trick, which defines the entropy as [38, 39]

$$S_{\rm EE} = (n\partial_n - 1)I_{\rm eff}(n)\big|_{n=1}\,, \tag{47}$$

where the effective action $I_{\rm eff}(n)$ is evaluated on a manifold with a conical singularity at the entangling surface whose excess angle is $2\pi(n-1)$.

As long as a covariant regulator is used to define the theory, the effective action will consist of terms that are local, diffeomorphism invariant integrals over the manifold, as well as nonlocal contributions. All UV matter divergences must appear in the local piece of the effective action, and each combines with terms in the classical gravitational part of the action, renormalizing the gravitational coupling constants. Furthermore, each such local term contributes to the entanglement entropy in (47) only at the conical singularity, giving a local integral over the entangling surface [10, 40, 41].

When the entangling surface is the bifurcation surface of a stationary horizon, this local integral is simply the Wald entropy [34, 42]. On nonstationary entangling surfaces, the computation can be done using the squashed cone techniques of [43], which yield terms involving extrinsic curvatures that modify the Wald entropy. In holography, the squashed cone method plays a key role in the proof of the Ryu-Takayanagi formula [17, 44], and its higher curvature generalization [45, 46]. The entropy functionals obtained in these works seem to also apply outside of holography, giving the extrinsic curvature terms in the entanglement entropy for general theories [12, 43].[5]

The extrinsic curvature modifications to the Wald entropy in fact take the form of a JKM Noether charge ambiguity [30, 50, 51]. To see this, note the vector ζ^a used to define the Noether charge vanishes at the entangling surface and its covariant derivative is antisymmetric and proportional to the binormal as in equation (21). This means it acts like a boost on the normal bundle at the entangling surface. General covariance requires that any extrinsic curvature contributions can be written as a sum of boost-invariant products,

$$S_{\text{JKM}} = \int_{\partial\Sigma} \mu \sum_{n\geq 1} B^{(-n)} \cdot C^{(n)} \tag{48}$$

where the superscript (n) denotes the boost weight of that tensor, so that at the surface: $\pounds_\zeta C^{(n)} = nC^{(n)}$. It is necessary that the terms consist of two pieces, each of which has nonzero boost weight, so that they can be written as

$$S_{\text{JKM}} = \int_{\partial\Sigma} \mu \sum_{n\geq 1} \frac{1}{n} B^{(-n)} \cdot \pounds_\zeta C^{(n)}\,. \tag{49}$$

This is of the form of a Noether charge ambiguity from equation (42), with[6] [7]

$$Y[\delta\phi] = \mu \sum_{n\geq 1} \frac{1}{n} B^{(-n)} \delta C^{(n)}\,. \tag{50}$$

The upshot of this discussion is that all terms in the entanglement entropy that are local on the entangling surface, including all divergences, are given by a Wald entropy modified by specific JKM terms. The couplings for the Wald entropy are determined by matching to the UV completion, or, in the absence of the UV description, these are simply parameters characterizing the low energy effective theory. In induced gravity scenarios, the divergences are determined by the matter content of the theory, and the matching to gravitational couplings has been borne out in explicit examples [52–54].

B. Equilibrium condition as gravitational constraints

We can now relate the variational identity (40) to entanglement entropy. The reduced density matrix for the ball in vacuum takes the form

$$\rho_\Sigma = e^{-H_{\text{mod}}}/Z\,, \tag{51}$$

where H_{mod} is the modular Hamiltonian and Z is the partition function, ensuring that ρ_Σ is normalized. Since the matter is conformally invariant, the modular Hamiltonian takes a simple form in terms of the matter Hamiltonian H_ζ^m defined in (9) [55, 56]

$$H_{\text{mod}} = \frac{2\pi}{\kappa} H_\zeta^m\,. \tag{52}$$

Next we apply the first law of entanglement entropy [57, 58], which states that the first order perturbation to the entanglement entropy is given by the change in modular Hamiltonian expectation value

$$\delta S_{\text{EE}} = \delta\langle H_{\text{mod}}\rangle\,. \tag{53}$$

Note that this equation holds for a fixed geometry and entangling surface, and hence coincides with what was referred to as δS_{mat} in section I. When varying the geometry, the divergent part of the entanglement entropy changes due to a change in the Wald entropy and JKM terms of the entangling surface. The total variation of the entanglement entropy is therefore

$$\delta S_{\text{EE}} = \delta(S_{\text{Wald}} + S_{\text{JKM}}) + \delta\langle H_{\text{mod}}\rangle\,. \tag{54}$$

At this point, we must give a prescription for defining the surface Σ in the perturbed geometry. Motivated by the first law of causal diamond mechanics, we require that Σ has the same generalized volume W' as in vacuum, where W' differs from the quantity W by a JKM term, as in equation (6). This provides a diffeomorphism-invariant criterion for defining the size of the ball. It does not fully fix all properties of the surface, but it is enough to derive the equilibrium condition for the entropy. As argued in section II D, the first term in equation (54) can be written instead as $\delta S_{\text{Wald}}\big|_W$ when the variation is taken holding W' fixed. Thus, from equations (40), (52) and (54), we arrive at our main result, the equilibrium condition

$$\frac{\kappa}{2\pi}\delta S_{\text{EE}}\big|_{W'} = \int_\Sigma \delta C_\zeta\,, \tag{55}$$

valid for minimally coupled, conformally invariant matter fields.

The linearized constraint equations $\delta C_\zeta = 0$ may therefore be interpreted as an equilibrium condition on entanglement entropy for the vacuum. Since all first variations of the entropy vanish when the linearized gravitational constraints are satisfied, the vacuum is an extremum of entropy for regions with fixed generalized volume W', which is necessary for it to be an equilibrium

[5] For terms involving four or more powers of extrinsic curvature, there are additional subtleties associated with the so called "splitting problem" [47–49].

[6] This formula defines Y at the entangling surface, and allows for some arbitrariness in defining it off the surface. It is not clear that Y can always be defined as a covariant functional of the form $Y[\delta\phi, \nabla_a\delta\phi, \ldots]$ without reference to additional structures, such as the normal vectors to the entangling surface. It would be interesting to understand better if and when Y lifts to such a spacetime covariant form off the surface.

[7] We thank Aron Wall for this explanation of JKM ambiguities.

state. Alternatively, postulating that entanglement entropy is maximal in vacuum for all balls and in all frames would allow one to conclude that the linearized higher derivative equations hold everywhere.

IV. FIELD EQUATIONS FROM THE EQUILIBRIUM CONDITION

The entanglement equilibrium hypothesis provides a clear connection between the linearized gravitational constraints and the maximality of entanglement entropy at fixed W' in the vacuum for conformally invariant matter. In this section, we will consider whether information about the fully nonlinear field equations can be gleaned from the equilibrium condition. Following the approach taken in [25], we employ a limit where the ball is taken to be much smaller than all relevant scales in the problem, but much larger than the cutoff scale of the effective field theory, which is set by the gravitational coupling constants. By expressing the linearized equations in Riemann normal coordinates, one can infer that the full *nonlinear* field equations hold in the case of Einstein gravity. As we discuss here, such a conclusion can *not* be reached for higher curvature theories. The main issue is that higher order terms in the RNC expansion are needed to capture the effect of higher curvature terms in the field equations, but these contribute at the same order as nonlinear corrections to the linearized equations.

We begin by reviewing the argument for Einstein gravity. Near any given point, the metric looks locally flat, and has an expansion in terms of Riemann normal coordinates that takes the form

$$g_{ab}(x) = \eta_{ab} - \frac{1}{3}x^c x^d R_{acbd}(0) + \mathcal{O}(x^3)\,, \tag{56}$$

where (0) means evaluation at the center of the ball. At distances small compared to the radius of curvature, the second term in this expression is a small perturbation to the flat space metric η_{ab}. Hence, we may apply the off-shell identity (55), using the first order variation

$$\delta g_{ab} = -\frac{1}{3}x^c x^d R_{acbd}(0)\,, \tag{57}$$

and conclude that the linearized constraint δC_ζ holds for this metric perturbation. When restricted to the surface Σ, this constraint in Einstein gravity is [35]

$$C_\zeta\big|_\Sigma = -u^a\zeta^b\left(\frac{1}{8\pi G}G_{ab} - T_{ab}\right)\eta\,. \tag{58}$$

Since the background constraint is assumed to hold, the perturbed constraint is

$$\delta C_\zeta\big|_\Sigma = -u^a\zeta^b\left(\frac{1}{8\pi G}\delta G_{ab} - \delta T_{ab}\right)\eta\,, \tag{59}$$

but in Riemann normal coordinates, we have that the linearized perturbation to the curvature is just $\delta G_{ab} = G_{ab}(0)$, up to terms suppressed by the ball radius. Assuming that the ball is small enough so that the stress tensor may be taken constant over the ball, one concludes that the vanishing constraint implies the nonlinear field equation at the center of the ball[8]

$$u^a\zeta^b(G_{ab}(0) - 8\pi G\delta T_{ab}) = 0\,. \tag{60}$$

The procedure outlined above applies at all points and all frames, allowing us to obtain the full tensorial Einstein equation.

Since we have only been dealing with the linearized constraint, one could question whether it gives a good approximation to the field equations at all points within the small ball. This requires estimating the size of the nonlinear corrections to this field equation. When integrated over the ball, the corrections to the curvature in RNC are of order ℓ^2/L^2, where ℓ is the radius of the ball and L is the radius of curvature. Since we took the ball size to be much smaller than the radius of curvature, these terms are already suppressed relative to the linear order terms in the field equation.

The situation in higher derivative theories of gravity is much different. It is no longer the case that the linearized equations evaluated in RNC imply the full nonlinear field equations in a small ball. To see this, consider an $L[g_{ab}, R_{bcde}]$ higher curvature theory.[9] The equations of motion read

$$-\frac{1}{2}g^{ab}\mathcal{L} + E^{aecd}R^b{}_{ecd} - 2\nabla_c\nabla_d E^{acdb} = \frac{1}{2}T^{ab}\,. \tag{61}$$

In appendix C we show that linearizing these equations around a Minkowski background leads to

$$\frac{\delta G^{ab}}{16\pi G} - 2\partial_c\partial_d\delta E^{acdb}_{\text{higher}} = \frac{1}{2}\delta T^{ab}\,, \tag{62}$$

where we split $E^{abcd} = E^{abcd}_{\text{Ein}} + E^{abcd}_{\text{higher}}$ into its Einstein piece, which gives rise to the Einstein tensor, and a piece coming from higher derivative terms. As noted before, the variation of the Einstein tensor evaluated in RNC gives the nonlinear Einstein tensor, up to corrections that are suppressed by the ratio of the ball size to the radius of curvature. However, in a higher curvature theory of gravity, the equations of motion (61) contain terms that are nonlinear in the curvature. Linearization around a

---

[8] In this equation, δT_{ab} should be thought of as a quantum expectation value of the stress tensor. Presumably, for sub-Planckian energy densities and in the small ball limit, this first order variation approximates the true energy density. However, there exist states for which the change in stress-energy is zero at first order in perturbations away from the vacuum, most notable for coherent states [59]. Analyzing how these states can be incorporated into the entanglement equilibrium story deserves further attention.

[9] Note that an analogous argument should hold for general higher derivative theories, which also involve covariant derivatives of the Riemann tensor.

MSS background of these terms would produce, schematically, $\delta(R^n) = n\bar{R}^{n-1}\delta R$, where $\bar{R}$ denotes evaluation in the MSS background. In Minkowski space, all such terms would vanish. This is not true in a general MSS, but evaluating the curvature tensors in the background still leads to a significant loss of information about the tensor structure of the equation. We conclude that the linearized equations cannot reproduce the full nonlinear field equations for higher curvature gravity, and it is only the linearity of the Einstein equation in the curvature that allows the nonlinear equations to be obtained for general relativity.

When linearizing around flat space, the higher curvature corrections to the Einstein equation are entirely captured by the second term in (62), which features four derivatives acting on the metric, since E^{abcd}_{higher} is constructed from curvatures that already contain two derivatives of the metric. Therefore, one is insensitive to higher curvature corrections unless at least $\mathcal{O}(x^4)$ corrections [60] are added to the Riemann normal coordinates expansion (57)

$$\delta g^{(2)}_{ab} = x^c x^d x^e x^f \left(\frac{2}{45} R_{acd}{}^g R_{befg} - \frac{1}{20} \nabla_c \nabla_d R_{aebf} \right). \tag{63}$$

Being quadratic in the Riemann tensor, this term contributes at the same order as the nonlinear corrections to the linearized field equations. Hence, linearization based on the RNC expansion up to x^4 terms is not fully self-consistent. This affirms the claim that for higher curvature theories, the nonlinear equations at a point cannot be derived by only imposing the linearized equations.

V. DISCUSSION

Maximal entanglement of the vacuum state was proposed in [25] as a new principle in quantum gravity. It hinges on the assumption that divergences in the entanglement entropy are cut off at short distances, so it is ultimately a statement about the UV complete quantum gravity theory. However, the principle can be phrased in terms of the generalized entropy, which is intrinsically UV finite and well-defined within the low energy effective theory. Therefore, if true, maximal vacuum entanglement provides a low energy constraint on any putative UV completion of a gravitational effective theory.

Higher curvature terms arise generically in any such effective field theory. Thus, it is important to understand how the entanglement equilibrium argument is modified by them. As explained in section II, the precise characterization of the entanglement equilibrium hypothesis relies on a classical variational identity for causal diamonds in maximally symmetric spacetimes. This identity leads to equation (40), which relates variations of the Wald entropy and matter energy density of the ball to the linearized constraints. The variations are taken holding fixed a new geometric functional W, defined in (34), which we call the "generalized volume."

We connected this identity to entanglement equilibrium in section III, invoking the fact that subleading entanglement entropy divergences are given by a Wald entropy, modified by specific JKM terms, which also modify W by the boundary term (44). With the additional assumption that matter is conformally invariant, we arrived at our main result (55), showing that the equilibrium condition $\delta S_{\text{EE}}\big|_{W'} = 0$ applied to small balls is equivalent to imposing the linearized constraints $\delta C_\zeta = 0$.

In section IV, we reviewed the argument that in the special case of Einstein gravity, imposing the linearized equations within small enough balls is equivalent to requiring that the fully nonlinear equations hold within the ball [25]. Thus by considering spheres centered at each point and in all Lorentz frames, one could conclude that the full Einstein equations hold everywhere.[10] Such an argument cannot be made for a theory that involves higher curvature terms. One finds that higher order terms in the RNC expansion are needed to detect the higher curvature pieces of the field equations, but these terms enter at the same order as the nonlinear corrections to the linearized equations. This signals a breakdown of the perturbative expansion unless the curvature is small.

The fact that we obtain only linearized equations for the higher curvature theory is consistent with the effective field theory standpoint. One could take the viewpoint that higher curvature corrections are suppressed by powers of a UV scale, and the effective field theory is valid only when the curvature is small compared to this scale. This suppression would suggest that the linearized equations largely capture the effects of the higher curvature corrections in the regime where effective field theory is reliable.

A. Comparison to other "geometry from entanglement" approaches

Several proposals have been put forward to understand gravitational dynamics in terms of thermodynamics and entanglement. Here we will compare the entanglement equilibrium program considered in this paper to two other approaches: the equation of state for local causal horizons, and gravitational dynamics from holographic entanglement entropy (see [61] for a related discussion).

[10] There is a subtlety associated with whether the solutions within each small ball can be consistently glued together to give a solution over all of spacetime. One must solve for the gauge transformation relating the Riemann normal coordinates at different nearby points, and errors in the linearized approximation could accumulate as one moves from point to point. The question of whether the ball size can be made small enough so that the total accumulated error goes to zero deserves further attention.

1. Causal horizon equation of state

By assigning an entropy proportional to the area of local causal horizons, Jacobson showed that the Einstein equation arises as an equation of state [19]. This approach employs a physical process first law for the local causal horizon, defining a heat δQ as the flux of local boost energy across the horizon. By assigning an entropy S to the horizon proportional to its area, one finds that the Clausius relation $\delta Q = T\delta S$ applied to all such horizons is equivalent to the Einstein equation.

The entanglement equilibrium approach differs in that it employs an equilibrium state first law [equation (37)], instead of a physical process one [62]. It therefore represents a different perspective that focuses on the steady-state behavior, as opposed to dynamics involved with evolution along the causal horizon. It is consistent therefore that we obtain constraint equations in the entanglement equilibrium setup, since one would not expect evolution equations to arise as an equilibrium condition.[11] That we can infer dynamical equations from the constraints is related to the fact that the dynamics of diffeomorphism-invariant theories is entirely determined by the constraints evaluated in all possible Lorentz frames.

Another difference comes from the focus on spacelike balls as opposed to local causal horizons. Dealing with a compact spatial region has the advantage of providing an IR finite entanglement entropy, whereas the entanglement associated with local causal horizons can depend on fields far away from the point of interest. This allows us to give a clear physical interpretation for the surface entropy functional as entanglement entropy, whereas such an interpretation is less precise in the equation of state approaches.

Finally, we note that both approaches attempt to obtain fully nonlinear equations by considering ultralocal regions of spacetime. In both cases the derivation of the field equations for Einstein gravity is fairly robust, however higher curvature corrections present some problems. Attempts have been made in the local causal horizon approach that involve modifying the entropy density functional for the horizon [63–71], but they meet certain challenges. These include a need for a physical interpretation of the chosen entropy density functional, and dependence of the entropy on arbitrary features of the local Killing vector in the vicinity of the horizon [71, 72]. While the entanglement equilibrium argument avoids these problems, it fails to get beyond linearized higher curvature equations, even after considering the small ball limit. The nonlinear equations in this case appear to involve information beyond first order perturbations, and hence may not be accessible based purely on an equilibrium argument.

[11] We thank Ted Jacobson for clarifying this point.

2. Holographic entanglement entropy

A different approach comes from holography and the Ryu-Takayanagi formula [17]. By demanding that areas of minimal surfaces in the bulk match the entanglement entropies of spherical regions in the boundary CFT, one can show that the linearized gravitational equations must hold [22–24]. The argument employs an equilibrium state first law for the bulk geometry, utilizing the Killing symmetry associated with Rindler wedges in the bulk.

The holographic approach is quite similar to the entanglement equilibrium argument since both use equilibrium state first laws. One difference is that the holographic argument must utilize minimal surfaces in the bulk, which extend all the way to the boundary of AdS. This precludes using a small ball limit as can be done with the entanglement equilibrium derivation, and is the underlying reason that entanglement equilibrium can derive fully nonlinear field equations in the case of Einstein gravity, whereas the holographic approach has thus far only obtained linearized equations. Some progress has been made to go beyond linear order in the holographic approach by considering higher order perturbations in the bulk [73–75]. Higher order perturbations will prove useful in the entanglement equilibrium program as well, and has the potential to extend the higher curvature derivation to fully nonlinear equations. Due to the similarity between the holographic and entanglement equlibrium approaches, progress in one will complement and inform the other.

B. Thermodynamic interpretation of the first law of causal diamond mechanics

Apart from the entanglement equilibrium interpretation, the first law of causal diamond mechanics could also directly be interpreted as a thermodynamic relation. Note that the identity (8) for Einstein gravity bears a striking resemblance to the fundamental relation in thermodynamics

$$dU = TdS - pdV, \tag{64}$$

where $U(S,V)$ is the internal energy, which is a function of the entropy S and volume V. The first law (8) turns into the thermodynamic relation (64), if one makes the following identifications for the temperature T and pressure p

$$T = \frac{\kappa\hbar}{2\pi k_B c}\,, \qquad p = \frac{c^2\kappa k}{8\pi G}\,. \tag{65}$$

Here we have restored fundamental constants, so that the quantities on the RHS have the standard units of temperature and pressure. The expression for the temperature is the well-known Unruh temperature [76]. The formula for the pressure lacks a microscopic understanding at the moment, although we emphasize the expression follows from consistency of the first law.

The thermodynamic interpretation motivates the name "first law" assigned to (8), and arguably it justifies the terminology "generalized volume" used for W in this paper, since it enters into the first law for higher curvature gravity (37) in the place of the volume. The only difference with the fundamental relation in thermodynamics is the minus sign in front of the energy variation. This different sign also enters into the first law for de Sitter horizons [77]. In the latter case the sign appears because empty de Sitter spacetime has maximal entropy, and adding matter only decreases the horizon entropy. Causal diamonds are rather similar in that respect.

C. Generalized volume and holographic complexity

The emergence of a generalized notion of volume in this analysis is interesting in its own right. We showed that when perturbing around a maximally symmetric background, the variation of the generalized volume is proportional to the variation of the gravitational part of the Hamiltonian. The fact that the Hamiltonian could be written in terms of a local, geometric functional of the surface was a nontrivial consequence of the background geometry being maximally symmetric and ζ^a being a conformal Killing vector whose conformal factor vanishes on Σ. The local geometric nature of W makes it a useful, diffeomorphism invariant quantity with which to characterize the region under consideration, and thus should be a good state function in the thermodynamic description of an ensemble of quantum geometry microstates. One might hope that such a microscopic description would also justify the fixed-W' constraint in the entanglement equilibrium derivation, which was only motivated macroscopically by the first law of causal diamond mechanics.

Volume has recently been identified as an important quantity in holography, where it is conjectured to be related to complexity [78, 79], or fidelity susceptibility [80]. The complexity=volume conjecture states that the complexity of some boundary state on a time slice Ω is proportional to the volume of the extremal codimension-one bulk hypersurface $\mathcal{B}$ which meets the asymptotic boundary on the corresponding time slice.[12]

While volume is the natural functional to consider for Einstein gravity, [81] noted that this should be generalized for higher curvature theories. The functional proposed in that work resembles our generalized volume W, but suffers from an arbitrary dependence on the choice of foliation of the codimension-one hypersurface on which it is evaluated. We therefore suggest that W, as defined in (34), may provide a suitable generalization of volume in the context of higher curvature holographic complexity.

Observe however that our derivation of W using the Iyer-Wald formalism was carried out in the particular case of spherical regions whose causal diamond is preserved by a conformal Killing vector. On more general grounds, one could speculate that the holographic complexity functional in higher derivative gravities should involve contractions of E^{abcd} with the geometric quantities characterizing $\mathcal{B}$, namely the induced metric h_{ab} and the normal vector u^a. The most general functional involving at most one factor of E^{abcd} can be written as

$$\mathcal{W}(\mathcal{B}) = \int_{\mathcal{B}} \eta \left(\alpha E^{abcd} u_a h_{bc} u_d + \beta E^{abcd} h_{ad} h_{bc} + \gamma \right) , \tag{66}$$

for some constants α, β and γ which should be such that $\mathcal{W}(\mathcal{B}) = V(\mathcal{B})$ for Einstein gravity. It would be interesting to explore the validity of this proposal in particular holographic setups, e.g., along the lines of [84].

D. Future work

We conclude by laying out future directions for the entanglement equilibrium program.

1. *Higher order perturbations*

In this work we restricted attention only to first order perturbations of the entanglement entropy and the geometry. Working to higher order in perturbation theory could yield several interesting results. One such possibility would be proving that the vacuum entanglement entropy is maximal, as opposed to merely extremal. The second order change in entanglement entropy is no longer just the change in modular Hamiltonian expectation value. The difference is given by the relative entropy, so a proof of maximality will likely invoke the positivity of relative entropy. On the geometrical side, a second order variational identity would need to be derived, along the lines of [85]. One would expect that graviton contributions would appear at this order, and it would be interesting to examine how they play into the entanglement equilibrium story. Also, by considering small balls and using the higher order terms in the Riemann normal coordinate expansion (63), in addition to higher order perturbations, it is possible that one could derive the fully nonlinear field equations of any higher curvature theory. Finally, coherent states pose a puzzle for the entanglement equilibrium hypothesis, since they change the energy within the ball without changing the entanglement [59]. However, their effect on the energy density only appears at second order in perturbations, so carrying the entanglement equilibrium argument to higher order could shed light on this puzzle.

[12] A similar expression has also been proposed for the complexity of subregions of the boundary time slice. In that case, $\mathcal{B}$ is the bulk hypersurface bounded by the corresponding subregion on the asymptotic boundary and the Ryu-Takayanagi surface [17] in the bulk [81, 82], or, more generally, the Hubeny-Rangamani-Takayanagi surface [83] if the spacetime is time-dependent [84].

2. Nonconformal matter

The arguments deriving the entanglement equilibrium condition in section III B were restricted to matter that is conformally invariant. For nonconformal matter, there are corrections to the modular Hamiltonian that spoil the relation between δS_{mat} and the matter Hamiltonian H^m_ζ. Nevertheless, in the small ball limit these corrections take on a simple form, and one possible solution for extending the entanglement equilibrium argument introduces a local cosmological constant to absorb the effects of the modular Hamiltonian corrections [25, 31, 32]. Allowing variations of the local cosmological constant would result in a modified first law [33], and may have connections to the black hole chemistry program [86, 87]. It is also possible that some other resolution exists to this apparent conflict, perhaps involving the RG properties of the matter field theory when taking the small ball limit.

3. Nonminimal couplings and gauge fields

We restricted attention to minimally coupled matter throughout this work. Allowing for nonminimal coupling can lead to new, state-dependent divergences in the entanglement entropy [27]. As before, these divergences will be localized on the entangling surface, taking the form of a Wald entropy. It therefore seems plausible that an entanglement equilibrium argument will go through in this case, reproducing the field equations involving the nonminimally coupled field. Note the state-dependent divergences could lead to variations of the couplings in the higher curvature theory, which may connect to the entanglement chemistry program, which considers Iyer-Wald first laws involving variations of the couplings [88].

Gauge fields introduce additional subtleties related to the existence of edge modes [89–91], and since these affect the renormalization of the gravitational couplings, they require special attention. Gravitons are even more problematic due to difficulties in defining the entangling surface in a diffeomorphism-invariant manner and in finding a covariant regulator [10, 12, 92, 93]. It would be interesting to analyze how to handle these issues in the entanglement equilibrium argument.

4. Nonspherical subregions

The entanglement equilibrium condition was shown to hold for spherical subregions and conformally invariant matter. One question that arises is whether an analogous equilibrium statement holds for linear perturbations to the vacuum in an arbitrarily shaped region. Nonspherical regions present a challenge because there is no longer a simple relation between the modular Hamiltonian and the matter stress tensor. Furthermore, nonspherical regions do not admit a conformal Killing vector which preserves its causal development. Since many properties of the conformal Killing vector were used when deriving the generalized volume W, it may need to be modified to apply to nonspherical regions and their perturbations.

Adapting the entanglement equilibrium arguments to nonspherical regions may involve shifting the focus to evolution under the modular flow, as opposed to a geometrical evolution generated by a vector field. Modular flows are complicated in general, but one may be able to use general properties of the flow to determine whether the Einstein equations still imply maximality of the vacuum entanglement for the region. Understanding the modular flow may also shed light on the behavior of the entanglement entropy for nonconformal matter, and whether some version of the entanglement equilibrium hypothesis continues to hold.

5. Physical process

As emphasized above, the first law of causal diamond mechanics is an equilibrium state construction since it compares the entropy of $\partial\Sigma$ on two infinitesimally related geometries [62]. One could ask whether there exists a physical process version of this story, which deals with entropy changes and energy fluxes as you evolve along the null boundary of the causal diamond. For this, the notion of quantum expansion for the null surface introduced in [12] would be a useful concept, which is defined by the derivative of the generalized entropy along the generators of the surface. One possible subtlety in formulating a physical process first law for the causal diamond is that the (classical) expansion of the null boundary is nonvanishing, so it would appear that this setup does not correspond to a dynamical equilibrium configuration. Nevertheless, it may be possible to gain useful information about the dynamics of semiclassical gravity by considering these nonequilibrium physical processes. An alternative that avoids this issue is to focus on quantum extremal surfaces [94] whose quantum expansion vanishes, and therefore may lend themselves to an equilibrium physical process first law.

ACKNOWLEDGMENTS

We would like to thank Joan Camps, Ted Jacobson, Arif Mohd, Rob Myers, Erik Verlinde and Aron Wall for helpful discussions, Fernando Rejon-Barrera for an early collaboration on this project, and Ted Jacobson for comments on a draft of this work. AJS is grateful to the Maryland Center for Fundamental Physics and Aron Wall for organizing the "Minicourse on Spacetime Thermodynamics." VSM, AJS and MRV thank the organizers of the "Amsterdam String Workshop," hosted by the Delta Institute for Theoretical Physics, and PB and AJS are grateful to the organizers of the "It from Qubit Summer School" held at the Perimeter Institute for Theoretical Physics. Research at Perimeter Institute is supported by the Government of Canada through Industry Canada and by the Province of Ontario through the Ministry of Research and Innovation. The work of PB is supported by a postdoctoral fellowship from the Fund for Scientific Research - Flanders (FWO). PB also acknowledges support from the Delta ITP Visitors Programme. VSM is supported by a PhD fellowship from the FWO and by the ERC grant 616732-HoloQosmos. AJS is supported by the National Science Foundation under grant No. PHY-1407744. MRV acknowledges support from the ERC Advanced Grant 268088-EMERGRAV, the Spinoza Grant of the Dutch Science Organisation (NWO), and the NWO Gravitation Program for the Delta Institute for Theoretical Physics.

Appendix A: Conformal Killing vector in flat space

Here we make explicit the geometric quantities introduced in section II B in the case of a Minkowski background, whose metric we write in spherical coordinates, i.e., $ds^2 = -dt^2 + dr^2 + r^2 d\Omega^2_{d-2}$. Let Σ be a spatial ball of radius ℓ in the time slice $t = 0$ and with center at $r = 0$. The conformal Killing vector which preserves the causal diamond of Σ is given by [25]

$$\zeta = \left(\frac{\ell^2 - r^2 - t^2}{\ell^2}\right)\partial_t - \frac{2rt}{\ell^2}\partial_r\,, \tag{A1}$$

where we have chosen the normalization in a way such that $\zeta^2 = -1$ at the center of the ball, which then gives the usual notion of energy for H^m_ζ (i.e. the correct units). It is straightforward to check that $\zeta(t = \pm\ell, r = 0) = \zeta(t = 0, r = \ell) = 0$, i.e., the tips of the causal diamond and the maximal sphere $\partial\Sigma$ at its waist are fixed points of ζ, as expected. Similarly, ζ is null on the boundary of the diamond. In particular, $\zeta(t = \ell \pm r) = \mp 2r(\ell \pm r)/\ell^2 \cdot (\partial_t \pm \partial_r)$. The vectors u and n (respectively normal to Σ and to both Σ and $\partial\Sigma$) read $u = \partial_t$, $n = \partial_r$, so that the binormal to $\partial\Sigma$ is given by $n_{ab} = 2\nabla_{[a} r \nabla_{b]} t$. It is also easy to check that $\pounds_\zeta g_{ab} = 2\alpha g_{ab}$ holds, where $\alpha \equiv \nabla_a \zeta^a / d = -2t/\ell^2$. Hence, we immediately see that $\alpha = 0$ on Σ, which implies that the gradient of α is proportional to the unit normal $u_a = -\nabla_a t$. Indeed, one finds $\nabla_a \alpha = -2\nabla_a t/\ell^2$, so in this case $N \equiv \|\nabla_a \alpha\|^{-1} = \ell^2/2$. It is also easy to show that $(\nabla_a \zeta_b)|_{\partial\Sigma} = \kappa n_{ab}$ holds, where the surface gravity reads $\kappa = 2/\ell$.

As shown in [28], given some metric g_{ab} with a conformal Killing field ζ^a, it is possible to construct other metrics $\bar{g}_{ab}$ conformally related to it, for which ζ^a is a true Killing field. More explicitly, if $\pounds_\zeta g_{ab} = 2\alpha g_{ab}$, then $\pounds_\zeta \bar{g}_{ab} = 0$ as long as g_{ab} and $\bar{g}_{ab}$ are related through $\bar{g}_{ab} = \Phi\, g_{ab}$, where Φ satisfies

$$\pounds_\zeta \Phi + 2\alpha\Phi = 0\,. \tag{A2}$$

For the vector (A1), this equation has the general solution

$$\Phi(r,t) = \frac{\psi(s)}{r^2} \quad \text{where} \quad s \equiv \frac{\ell^2 + r^2 - t^2}{r}\,. \tag{A3}$$

Here, $\psi(s)$ can be any function. Hence, ζ in (A1) is a true Killing vector for all metrics conformally related to Minkowski's with a conformal factor given by (A3). For example, setting $\psi(s) = L^2$, for some constant L^2, one obtains the metric of $\mathrm{AdS}_2 \times S_{d-2}$ with equal radii, namely: $ds^2 = L^2/r^2(-dt^2 + dr^2) + L^2 d\Omega^2_{d-2}$. Another simple case corresponds to $\psi(s) = L^2((s^2/(4L^2) - 1)^{-1}$. Through the change of variables [56]: $t = L\sinh(\tau/L)/(\cosh u + \cosh(\tau/L))$, $r = L\sinh u/(\cosh u + \cosh(\tau/L))$, this choice leads to the $\mathbb{R} \times H^{d-1}$ metric (where H^{d-1} is the hyperbolic plane): $ds^2 = -d\tau^2 + L^2(du^2 + \sinh^2 u\, d\Omega^2_{d-2})$.

Appendix B: Generalized volume in higher order gravity

The generalized volume W is defined in (34). We restate the expression here

$$W = \frac{1}{(d-2)E_0}\int_\Sigma \eta\left(E^{abcd}u_a u_d h_{bc} - E_0\right)\,, \tag{B1}$$

where E_0 is a theory-dependent constant defined by the tensor E^{abcd} in a maximally symmetric solution to the field equations through $E^{abcd} \overset{\text{MSS}}{=} E_0(g^{ac}g^{bd} - g^{ad}g^{bc})$. Moreover, E^{abcd} is the variation of the Lagrangian scalar $\mathcal{L}$ with respect to the Riemann tensor R_{abcd} if we were to treat it as an independent field [29],

$$\begin{aligned} E^{abcd} = \frac{\partial\mathcal{L}}{\partial R_{abcd}} - \nabla_{a_1}\frac{\partial\mathcal{L}}{\partial\nabla_{a_1}R_{abcd}} + \dots \\ + (-1)^m \nabla_{(a_1}\cdots\nabla_{a_m)}\frac{\partial\mathcal{L}}{\partial\nabla_{(a_1}\cdots\nabla_{a_m)}R_{abcd}}\,, \end{aligned} \tag{B2}$$

where $\mathcal{L}$ is then defined through $L = \epsilon\mathcal{L}$. In this section we provide explicit expressions for W in $f(R)$ gravity, quadratic gravity and Gauss-Bonnet gravity. Observe that throughout this section we use the bar on $\bar{R}$ to denote evaluation on a MSS. Imposing a MSS to solve the field equations of a given higher derivative theory gives rise to a constraint between the theory couplings and the background curvature $\bar{R}$. This reads [37]

$$E_0 = \frac{d}{4\bar{R}}\mathcal{L}(\bar{R})\,, \tag{B3}$$

where $\mathcal{L}(\bar{R})$ denotes the Lagrangian scalar evaluated on the background.

a. $f(R)$ gravity. A simple higher curvature gravity is obtained by replacing R in the Einstein-Hilbert action by a function of R

$$L_{f(R)} = \frac{1}{16\pi G}\epsilon f(R)\,. \tag{B4}$$

To obtain the generalized volume we need

$$E^{abcd}_{f(R)} = \frac{f'(R)}{32\pi G}\left(g^{ac}g^{bd} - g^{ad}g^{bc}\right)\,, \quad E_0 = \frac{f'(\bar{R})}{32\pi G}\,. \tag{B5}$$

The generalized volume then reads

$$W_{f(R)} = \frac{1}{d-2}\int_\Sigma \eta\left[(d-1)\frac{f'(R)}{f'(\bar{R})} - 1\right]. \qquad \text{(B6)}$$

b. Quadratic gravity. A general quadratic theory of gravity is given by the Lagrangian

$$L_{\text{quad}} = \epsilon\Bigg[\frac{1}{16\pi G}(R-2\Lambda) + \alpha_1 R^2 + \alpha_2 R_{ab}R^{ab} + \alpha_3 R_{abcd}R^{abcd}\Bigg]. \qquad \text{(B7)}$$

Taking the derivative of the Lagrangian with respect to the Riemann tensor leaves us with

$$E^{abcd}_{\text{quad}} = \left(\frac{1}{32\pi G} + \alpha_1 R\right) 2g^{a[c}g^{d]b} + \alpha_2\left(R^{a[c}g^{d]b} + R^{b[d}g^{c]a}\right) + 2\alpha_3 R^{abcd}, \qquad \text{(B8)}$$

and using (28) one finds

$$E_0 = \frac{1}{32\pi G} + \left(\alpha_1 + \frac{\alpha_2}{d} + \frac{2\alpha_3}{d(d-1)}\right)\bar{R}. \qquad \text{(B9)}$$

The generalized volume for quadratic gravity thus reads

$$W_{\text{quad}} = \frac{1}{(d-2)E_0}\int_\Sigma \eta\Bigg[(d-1)\left(\frac{1}{32\pi G} + \alpha_1 R\right) - E_0 + \frac{1}{2}\alpha_2\left(R - R^{ab}u_a u_b(d-2)\right) - 2\alpha_3 R^{ab}u_a u_b\Bigg]. \qquad \text{(B10)}$$

An interesting instance of quadratic gravity is Gauss-Bonnet theory, which is obtained by restricting to $\alpha_1 = -\frac{1}{4}\alpha_2 = \alpha_3 = \alpha$. The generalized volume then reduces to

$$W_{\text{GB}} = \frac{1}{(d-2)E_0}\int_\Sigma \eta\left[\frac{1}{32\pi G}(d-1) - E_0 + (d-3)\alpha\Big(R + 2R^{ab}u_a u_b\Big)\right], \qquad \text{(B11)}$$

with $E_0 = 1/(32\pi G) + \alpha\bar{R}(d-2)(d-3)/(d(d-1))$. Since the extrinsic curvature of Σ vanishes in the background, the structure $R + 2R^{ab}u_a u_b$ is equal to the intrinsic Ricci scalar of Σ, in the background and at first order in perturbations.

Appendix C: Linearized equations of motion for higher curvature gravity using RNC

The variational identity (36) states that the vanishing of the linearized constraint equations δC_ζ is equivalent to a relation between the variation of the Wald entropy, generalized volume, and matter energy density. In [25], Jacobson used this relation to extract the Einstein equations, making use of Riemann normal coordinates. Here we perform a similar calculation for the higher curvature generalization of the first law of causal diamond mechanics which will produce the linearized equations of motion. In this appendix we will restrict to theories whose Lagrangian depends on the metric and the Riemann tensor, $L[g_{ab}, R_{abcd}]$, and to linearization around flat space.

The equations of motion for such a general higher curvature theory read

$$-\frac{1}{2}g^{ab}\mathcal{L} + E^{aecd}R^{b}{}_{ecd} - 2\nabla_c\nabla_d E^{acdb} = \frac{1}{2}T^{ab}. \qquad \text{(C1)}$$

Linearizing the equations of motion around flat space leads to

$$-\frac{1}{32\pi G}\eta^{ab}\delta R + E^{aecd}_{\text{Ein}}\delta R^{b}{}_{ecd} - 2\partial_c\partial_d\delta E^{acdb}_{\text{higher}} = \frac{\delta G^{ab}}{16\pi G} - 2\partial_c\partial_d\delta E^{acdb}_{\text{higher}} = \frac{1}{2}\delta T^{ab}, \qquad \text{(C2)}$$

where we split $E^{abcd} = E^{abcd}_{\text{Ein}} + E^{abcd}_{\text{higher}}$ into an Einstein piece, which goes into the Einstein tensor, and a piece coming from higher derivative terms. We used the fact that many of the expressions in (C2) significantly simplify when evaluated in the Minkowski background because the curvatures vanish. For example, one might have expected additional terms proportional to the variation of the Christoffel symbols coming from $\delta(\nabla_c\nabla_d E^{acdb})$. To see why these terms are absent, it is convenient to split this expression into its Einstein part and a part coming from higher derivative terms. The Einstein piece does not contribute since E^{acdb}_{Ein} is only a function of the metric and therefore its covariant derivative vanishes. The higher derivative piece will give $\partial_c\partial_d\delta E^{acdb}_{\text{higher}}$ as well as terms such as $\delta\Gamma^c_{ce}\nabla_d E^{eadb}_{\text{higher}}$ and $\Gamma^c_{ce}\nabla_d\delta E^{eadb}_{\text{higher}}$. However, the latter two terms are zero because both the Christoffel symbols and E^{eadb}_{higher} vanish when evaluated in the Minkowski background with the standard coordinates.

We now want to evaluate each term in (22) using Riemann normal coordinates. Taking the stress tensor T^{ab} to be constant for small enough balls, the variation of (9) reduces to

$$\delta H^m_\zeta = \frac{\Omega_{d-2}\ell^d}{d^2-1}\kappa u_a u_b \delta T^{ab} + \mathcal{O}\left(\ell^{d+2}\right), \qquad \text{(C3)}$$

where Ω_{d-2} denotes the area of the $(d-2)$-sphere, ℓ is the radius of our geodesic ball and u_a is the future pointing unit normal. As was found in [25], the Einstein piece of the symplectic form will combine with the area term of the entropy to produce the Einstein tensor. Therefore, we focus on the higher curvature part of δH^g_ζ. Combining (12) and (33), we find

$$\delta H^g_{\zeta,\text{higher}} = -\frac{4\kappa}{\ell}\int d\Omega \int dr r^{d-2} u_a u_d \eta_{bc}\Big(\delta E^{abcd}_{\text{higher}}(0) + \partial_i \delta E^{abcd}_{\text{higher}}(0) r n^i + \frac{1}{2}\partial_i\partial_j \delta E^{abcd}_{\text{higher}}(0) r^2 n^i n^j + \mathcal{O}\left(r^3\right)\Big)$$
$$= -4\kappa\Omega_{d-2}\ell^{d-2} u_a u_d \eta_{bc}\Bigg(\frac{\delta E^{abcd}_{\text{higher}}(0)}{(d-1)} + \frac{\ell^2\delta^{ij}\partial_i\partial_j\delta E^{abcd}_{\text{higher}}(0)}{2(d^2-1)}\Bigg) + \mathcal{O}\left(\ell^{d+2}\right) . \tag{C4}$$

Here, n^i is the normal vector to $\partial\Sigma$ and the indices a, b run over space-time directions, while the indices i, j run only over spatial directions, and ∂_i is the derivative operator compatible with the flat background metric on Σ. In the first line, we simply use the formula for the Taylor expansion of a quantity f in the coordinate system compatible with ∂_i,

$$f(x) = f(0) + \partial_a f(0) x^a + \frac{1}{2}\partial_a\partial_b f(0) x^a x^b + \mathcal{O}\left(x^3\right) , \tag{C5}$$

where (0) denotes that a term is evaluated at $r = 0$. Since we evaluate our expressions on a constant timeslice at $t = 0$, we have $x^t = 0$ and $x^i = r\, n^i$, where r is a radial coordinate inside the geodesic ball and the index i runs only over the spatial coordinates. To evaluate the spherical integral, it is useful to note that spherical integrals over odd powers of n^i vanish and furthermore

$$\int d\Omega\, n^i n^j = \frac{\Omega_{d-2}}{d-1}\delta^{ij} , \tag{C6}$$

$$\int d\Omega\, n^i n^j n^k n^l = \frac{\Omega_{d-2}}{d^2-1}\left(\delta^{ij}\delta^{kl} + \delta^{ik}\delta^{jl} + \delta^{il}\delta^{jk}\right) . \tag{C7}$$

Next, we evaluate δS_{higher}, the variation of the higher curvature part of the Wald entropy given in (24), in a similar manner.

$$\delta S_{\text{higher}} = 8\pi\Omega_{d-2}\ell^{d-2} u_a u_d \left(\frac{\eta_{bc}\delta E^{abcd}_{\text{higher}}(0)}{(d-1)} + \frac{\ell^2\left[\eta_{bc}\delta^{ij}\partial_i\partial_j\delta E^{abcd}_{\text{higher}}(0) + 2\partial_b\partial_c\delta E^{abcd}_{\text{higher}}(0)\right]}{2(d^2-1)}\right) + \mathcal{O}\left(\ell^{d+2}\right) , \tag{C8}$$

We are now ready to evaluate the first law of causal diamond mechanics (22). Interestingly, the leading order pieces of the Hamiltonian and Wald entropy exactly cancel against each other. Note that these two terms would have otherwise dominated over the Einstein piece. Furthermore, the second term in the symplectic form and Wald entropy also cancel, leaving only a single term from the higher curvature part of the identity. Including the Einstein piece, we find the first law for higher curvature gravity reads in Riemann normal coordinates

$$-\frac{\kappa\Omega_{d-2}\ell^d}{d^2-1} u_a u_d \Big(\frac{\delta G^{ad}(0)}{8\pi G} - 4\partial_b\partial_c\delta E^{abcd}_{\text{higher}}(0) - \delta T^{ad}\Big) + \mathcal{O}\left(\ell^{d+2}\right) = 0 , \tag{C9}$$

proving equivalence to the linearized equations (C2).

[1] J. D. Bekenstein, Lett. Nuovo Cimento **4**, 737 (1972).
[2] J. D. Bekenstein, Phys. Rev. D **7**, 2333 (1973).
[3] S. W. Hawking, Nature **248**, 30 (1974).
[4] R. D. Sorkin, in *Tenth International Conference on General Relativity and Gravitation, Contributed Papers*, Vol. II, edited by B. Bertotti, F. de Felice, and A. Pascolini (Roma, Consiglio Nazionale Delle Ricerche, 1983) pp. 734–736, arXiv:1402.3589.
[5] L. Bombelli, R. K. Koul, J. Lee, and R. D. Sorkin, Phys. Rev. D **34**, 373 (1986).
[6] M. Srednicki, Phys. Rev. Lett. **71**, 666 (1993), arXiv:hep-th/9303048.
[7] V. Frolov and I. Novikov, Phys. Rev. D **48**, 4545 (1993), arXiv:gr-qc/9309001.
[8] R. M. Wald, Phys. Rev. D **48**, R3427 (1993), arXiv:gr-qc/9307038.
[9] L. Susskind and J. Uglum, Phys. Rev. D **50**, 2700 (1994), arXiv:hep-th/9401070.
[10] J. H. Cooperman and M. A. Luty, J. High Energy Phys. **1412**, 45 (2014), arXiv:1302.1878.
[11] S. N. Solodukhin, Living Rev. Relativ. **14** (2011), 10.12942/lrr-2011-8.
[12] R. Bousso, Z. Fisher, S. Leichenauer, and A. C. Wall, Phys. Rev. D **93**, 064044 (2016), arXiv:1506.02669.
[13] T. Jacobson, (1994), arXiv:gr-qc/9404039 [gr-qc].
[14] T. Jacobson, in *Eighth Canadian conference on general relativity and relativistic astrophysics*, Vol. 493 (ASCE, 1999) pp. 85–97, arXiv:gr-qc/9908031.

[15] T. Jacobson and R. Parentani, Found. Phys. **33**, 323 (2003), arXiv:0302099 [gr-qc].
[16] E. Bianchi and R. C. Myers, Classical Quant. Grav. **31**, 214002 (2012), arXiv:1212.5183.
[17] S. Ryu and T. Takayanagi, Phys. Rev. Lett. **96**, 181602 (2006), arXiv:hep-th/0603001.
[18] T. Faulkner, A. Lewkowycz, and J. Maldacena, J. High Energy Phys. **1311**, 74 (2013), arXiv:1307.2892.
[19] T. Jacobson, Phys. Rev. Lett. **75**, 1260 (1995), arXiv:gr-qc/9504004.
[20] E. P. Verlinde, J. High Energy Phys. **1104**, 29 (2010), arXiv:1001.0785.
[21] E. P. Verlinde, (2016), arXiv:1611.02269 [hep-th].
[22] N. Lashkari, M. B. McDermott, and M. Van Raamsdonk, J. High Energy Phys. **1404**, 195 (2013), arXiv:1308.3716.
[23] T. Faulkner, M. Guica, T. Hartman, R. C. Myers, and M. Van Raamsdonk, J. High Energy Phys. **1403**, 51 (2013), arXiv:1312.7856.
[24] B. Swingle and M. Van Raamsdonk, (2014), arXiv:1405.2933.
[25] T. Jacobson, Phys. Rev. Lett. **116**, 201101 (2016), 1505.04753.
[26] T. Jacobson and A. Satz, Phys. Rev. **D87**, 084047 (2013), arXiv:1212.6824 [hep-th].
[27] D. Marolf and A. C. Wall, J. High Energy Phys. **1610**, 109 (2016), arXiv:1607.01246.
[28] T. Jacobson and G. Kang, Classical Quant. Grav. **10**, L201 (1993), arXiv:gr-qc/9307002.
[29] V. Iyer and R. M. Wald, Phys. Rev. D **50**, 846 (1994), arXiv:gr-qc/9403028.
[30] T. Jacobson, G. Kang, and R. C. Myers, Phys. Rev. D **49**, 6587 (1994), arXiv:gr-qc/9312023.
[31] H. Casini, D. A. Galante, and R. C. Myers, J. High Energy Phys. **2016**, 1 (2016), arXiv:1601.00528 [hep-th].
[32] A. J. Speranza, J. High Energy Phys. **2016**, 1 (2016), arXiv:1602.01380 [hep-th].
[33] T. Jacobson and M. Visser, in preparation.
[34] V. Iyer and R. M. Wald, Phys. Rev. D **52**, 4430 (1995), arXiv:gr-qc/9503052.
[35] M. D. Seifert and R. M. Wald, Phys. Rev. D **75**, 084029 (2007), arXiv:gr-qc/0612121.
[36] T. Jacobson, Classical Quant. Grav. **28**, 245011 (2011), arXiv:1108.1496.
[37] P. Bueno, P. A. Cano, V. S. Min, and M. R. Visser, (2016), arXiv:1610.08519 [hep-th].
[38] C. Callan and F. Wilczek, Phys. Lett. B **333**, 55 (1994), arXiv:hep-th/9401072.
[39] P. Calabrese and J. Cardy, J. Phys. A **42**, 504005 (2009), arXiv:0905.4013.
[40] D. V. Fursaev and S. N. Solodukhin, Phys. Lett. B **365**, 51 (1996), arXiv:hep-th/9412020.
[41] F. Larsen and F. Wilczek, Nucl. Phys. B **458**, 249 (1996), arXiv:hep-th/9506066.
[42] W. Nelson, Phys. Rev. D **50**, 7400 (1994), arXiv:hep-th/9406011.
[43] D. V. Fursaev, A. Patrushev, and S. N. Solodukhin, Phys. Rev. D **88**, 044054 (2013), arXiv:1306.4000.
[44] A. Lewkowycz and J. Maldacena, J. High Energy Phys. **1308**, 90 (2013), arXiv:1304.4926.
[45] X. Dong, J. High Energy Phys. **1401**, 44 (2013), arXiv:1310.5713.
[46] J. Camps, J. High Energy Phys. **1403**, 70 (2013), arXiv:1310.6659.
[47] R.-X. Miao and W.-z. Guo, J. High Energy Phys. **1508**, 31 (2014), arXiv:1411.5579.
[48] R.-X. Miao, J. High Energy Phys. **1510**, 49 (2015), arXiv:1503.05538.
[49] J. Camps, J. High Energy Phys. **09**, 139 (2016), arXiv:1605.08588 [hep-th].
[50] S. Sarkar and A. C. Wall, Phys. Rev. D **88**, 044017 (2013), arXiv:1306.1623.
[51] A. C. Wall, Int. J. Mod. Phys. D **24**, 1544014 (2015), arXiv:1504.08040 [gr-qc].
[52] V. Frolov, D. Fursaev, and A. Zelnikov, Nucl. Phys. B **486**, 339 (1997), arXiv:hep-th/9607104.
[53] R. C. Myers, R. Pourhasan, and M. Smolkin, J. High Energy Phys. **2013**, 13 (2013), arXiv:1304.2030.
[54] R. Pourhasan, J. High Energy Phys. **2014**, 4 (2014), arXiv:1403.0951.
[55] P. D. Hislop and R. Longo, Commun. Math. Phys. **84**, 71 (1982).
[56] H. Casini, M. Huerta, and R. C. Myers, J. High Energy Phys. **1105**, 36 (2011), arXiv:1102.0440.
[57] D. D. Blanco, H. Casini, L.-Y. Hung, and R. C. Myers, J. High Energy Phys. **1308**, 60 (2013), arXiv:1305.3182.
[58] J. Bhattacharya, M. Nozaki, T. Takayanagi, and T. Ugajin, Phys. Rev. Lett. **110**, 091602 (2012), arXiv:1212.1164.
[59] M. Varadarajan, Gen. Relativ. Gravit. **48**, 35 (2016), arXiv:1602.00106.
[60] L. Brewin, Classical Quant. Grav. **26**, 175017 (2009), arXiv:0903.2087.
[61] S. M. Carroll and G. N. Remmen, Phys. Rev. D **93**, 124052 (2016), arXiv:1601.07558 [hep-th].
[62] R. M. Wald, *Quantum Field Theory in Curved Spacetime and Black Hole Thermodynamics* (The University of Chicago Press, 1994).
[63] C. Eling, R. Guedens, and T. Jacobson, Phys. Rev. Lett. **96**, 121301 (2006), arXiv:0602001 [gr-qc].
[64] E. Elizalde and P. J. Silva, Phys. Rev. D **78**, 061501 (2008), arXiv:0804.3721 [hep-th].
[65] G. Chirco, C. Eling, and S. Liberati, Phys. Rev. D **83**, 024032 (2011), arXiv:1011.1405 [gr-qc].
[66] R. Brustein and M. Hadad, Phys. Rev. Lett. **103**, 101301 (2009), arXiv:0903.0823 [hep-th].
[67] M. Parikh and S. Sarkar, Entropy **18**, 119 (2016), arXiv:0903.1176 [hep-th].
[68] R. Dey, S. Liberati, and A. Mohd, Phys. Rev. D **94**, 044013 (2016), arXiv:1605.04789 [gr-qc].
[69] T. Padmanabhan, (2009), arXiv:0903.1254.
[70] T. Padmanabhan, Rep. Prog. Phys. **73**, 046901 (2009), arXiv:0911.5004.
[71] R. Guedens, T. Jacobson, and S. Sarkar, Phys. Rev. D **85**, 064017 (2011), arXiv:1112.6215.
[72] T. Jacobson, Int. J. Mod. Phys. D **21**, 1242006 (2012), arXiv:1204.6349.
[73] T. Faulkner, J. High Energy Phys. **1505**, 33 (2014), arXiv:1412.5648.
[74] N. Lashkari and M. Van Raamsdonk, J. High Energy Phys. **1604**, 153 (2015), arXiv:1508.00897.
[75] M. J. S. Beach, J. Lee, C. Rabideau, and M. Van Raamsdonk, J. High Energy Phys. **1606**, 85 (2016), arXiv:1604.05308.
[76] W. G. Unruh, Phys. Rev. **D14**, 870 (1976).
[77] G. W. Gibbons and S. W. Hawking, Phys. Rev. **D15**, 2738 (1977).
[78] L. Susskind, Fortsch. Phys. **64**, 24 (2014), arXiv:1402.5674.

[79] A. R. Brown, D. A. Roberts, L. Susskind, B. Swingle, and Y. Zhao, Phys. Rev. Lett. **116**, 191301 (2015), arXiv:1509.07876.
[80] M. Miyaji, T. Numasawa, N. Shiba, T. Takayanagi, and K. Watanabe, Phys. Rev. Lett. **115**, 261602 (2015), arXiv:1507.07555.
[81] M. Alishahiha, Phys. Rev. D **92**, 126009 (2015), arXiv:1509.06614.
[82] O. Ben-Ami and D. Carmi, (2016), arXiv:1609.02514.
[83] V. E. Hubeny, M. Rangamani, and T. Takayanagi, JHEP **07**, 062 (2007), arXiv:0705.0016 [hep-th].
[84] D. Carmi, R. C. Myers, and P. Rath, (2016), arXiv:1612.00433.
[85] S. Hollands and R. M. Wald, Commun. Math. Phys. **321**, 629 (2013), arXiv:1201.0463 [gr-qc].
[86] D. Kastor, S. Ray, and J. Traschen, Classical Quant. Grav. **26**, 195011 (2009), arXiv:0904.2765.
[87] D. Kubiznak, R. B. Mann, and M. Teo, (2016), arXiv:1608.06147.
[88] E. Caceres, P. H. Nguyen, and J. F. Pedraza, (2016), arXiv:1605.00595.
[89] W. Donnelly and A. C. Wall, Phys. Rev. D **86**, 064042 (2012), arXiv:1206.5831.
[90] W. Donnelly and A. C. Wall, Phys. Rev. Lett. **114**, 111603 (2015), arXiv:1412.1895 [hep-th].
[91] W. Donnelly and A. C. Wall, (2015), arXiv:1506.05792 [hep-th].
[92] D. V. Fursaev and G. Miele, Nucl. Phys. B **484**, 697 (1997), arXiv:9605153 [hep-th].
[93] S. N. Solodukhin, Phys. Rev. D **91**, 084028 (2015), arXiv:1502.03758.
[94] N. Engelhardt and A. C. Wall, J. High Energy Phys. **1501**, 73 (2015), arXiv:1408.3203 [hep-th].