

# Tomography from Entanglement

Jennifer Lin,<sup>1</sup> Matilde Marcolli,<sup>2</sup> Hiroshi Ooguri,<sup>3,4</sup> and Bogdan Stoica<sup>3</sup>

<sup>1</sup>*Enrico Fermi Institute and Department of Physics, University of Chicago, Chicago, IL 60637*

<sup>2</sup>*Department of Mathematics, California Institute of Technology, 253-37, Pasadena, CA 91125*

<sup>3</sup>*Walter Burke Institute for Theoretical Physics, California Institute of Technology, 452-48, Pasadena, CA 91125*

<sup>4</sup>*Kavli Institute for the Physics and Mathematics of the Universe (WPI), University of Tokyo, Kashiwa 277-8583, Japan*  
(Dated: December 5, 2014)

The Ryu-Takayanagi formula relates the entanglement entropy in a conformal field theory to the area of a minimal surface in its holographic dual. We show that this relation can be inverted for any state in the conformal field theory to compute the bulk stress-energy tensor near the boundary of the bulk spacetime, reconstructing the local data in the bulk from the entanglement on the boundary. We also show that positivity, monotonicity, and convexity of the relative entropy for small spherical domains between the reduced density matrices of any state and of the ground state of the conformal field theory, follow from positivity conditions on the bulk matter energy density. We discuss an information theoretical interpretation of the convexity in terms of the Fisher metric.

PACS numbers: 11.25.Tq

## I. INTRODUCTION

Gauge/gravity duality posits an exact equivalence between certain conformal field theories (CFT's) with many degrees of freedom and higher dimensional theories with gravity. It is of obvious interest to understand how bulk spacetime geometry and gravitational dynamics emerge from a non-gravitating theory. In recent years, there have appeared hints that quantum entanglement plays a key role. One important development in this direction was the proposal of Ryu and Takayanagi [1, 2] that the entanglement entropy (EE) between a spatial domain  $D$  of a CFT and its complement is equal to the area of the bulk extremal surface  $\Sigma$  homologous to it,

$$S_{EE} = \min_{\partial D = \partial \Sigma} \frac{\text{area}(\Sigma)}{4G_N}. \quad (1)$$

Using (1), [3–8] showed the emergence of linearized gravity from entanglement physics of the CFT, as we review below. In this note, we continue this program. We show that certain universal properties of CFT's, namely positivity, monotonicity and convexity of the relative entropy between the reduced density matrix on a small spherical domain of an excited state of the CFT and of its vacuum state, follow from positivity conditions on the matter energy density in the near-AdS region of its dual. Moreover, we show that the bulk stress-energy density in this region can be reconstructed point-by-point from the entanglement on the boundary.

We first set the stage with a review of concepts from quantum information theory before more precisely stating our main claim.

### A. Relative Entropy

Relative entropy (see *e.g.* [4] and references therein) is a measure of distinguishability between two quantum

states in the same Hilbert space. The relative entropy of two density matrices  $\rho_0$  and  $\rho_1$  is defined as

$$S(\rho_1|\rho_0) = \text{tr}(\rho_1 \log \rho_1) - \text{tr}(\rho_1 \log \rho_0). \quad (2)$$

It is positive, and increases with system size:

$$S(\rho_1|\rho_0) \geq 0, \quad (3)$$

$$S(\rho_1^W|\rho_0^W) \geq S(\rho_1^V|\rho_0^V), \quad W \supseteq V. \quad (4)$$

This second property is called monotonicity. When  $\rho_0$  and  $\rho_1$  are reduced density matrices on a spatial domain  $D$  for two states of a quantum field theory (QFT), which is the case we specialize to from this point on, (4) implies that  $S(\rho_1|\rho_0)$  increases with the size of  $D$ . That is, over a family of scalable domains with characteristic size  $R$ ,

$$\partial_R S(\rho_1|\rho_0) \geq 0. \quad (5)$$

Defining the modular Hamiltonian  $H_{mod}$  of  $\rho_0$  implicitly through

$$\rho_0 = \frac{e^{-H_{mod}}}{\text{tr}(e^{-H_{mod}})}, \quad (6)$$

it is easy to see that (3) is equivalent to

$$S(\rho_1|\rho_0) = \Delta \langle H_{mod} \rangle - \Delta S_{EE} \geq 0 \quad (7)$$

where  $\Delta \langle H_{mod} \rangle = \text{tr}(\rho_1 H_{mod}) - \text{tr}(\rho_0 H_{mod})$  is the change in the expectation value of the operator  $H_{mod}$  (6) and  $\Delta S_{EE} = -\text{tr}(\rho_1 \log \rho_1) + \text{tr}(\rho_0 \log \rho_0)$  is the change in the entanglement entropy across  $D$  as one goes between the states.

When the states under comparison are close, the positivity (7) is saturated to leading order [4, 5, 7]:

$$S(\rho_1|\rho_0) = \Delta \langle H_{mod} \rangle - \Delta S_{EE} = 0. \quad (8)$$

To see this, consider a reference state of the QFT characterized by  $\rho_0$ , and another, arbitrary state  $\rho_1$ . One can construct a family of interpolating density matrices

$$\rho(\lambda) = (1 - \lambda)\rho_0 + \lambda\rho_1, \quad (9)$$

where  $\lambda$  can be positive or negative. Because the relative entropy  $S(\rho_0|\rho(\lambda))$  is positive for either sign of  $\lambda$ , the first derivative of this relative entropy with respect to  $\lambda$  vanishes. This implies (8) to first order in  $\lambda$ .

Eq. (8) is called the entanglement first law for its resemblance to the first law of thermodynamics. Indeed, when  $\rho_0$  is a thermal density matrix  $\rho_0 = e^{-\beta H}/\text{tr}(e^{-\beta H})$ , (8) reduces to  $\Delta\langle H \rangle = T\Delta S$ , an exact quantum version of the thermal first law.

In general, the modular Hamiltonian (6) associated to a given density matrix is nonlocal. However, there are a few simple cases where it is explicitly known. When  $\rho_0$  is the reduced density matrix of the vacuum state of a CFT on a disk of radius  $R$  which (without loss of generality) we take to be centered at  $\vec{x}_0 = 0$  [9],

$$H_{\text{mod}} = \pi \int_D d^{d-1}x \frac{R^2 - |\vec{x}|^2}{R} T_{tt}(x), \quad (10)$$

where  $T_{tt}$  is the energy density of the CFT.

## B. Summary and Outline

Our goal in this note is to use the entanglement in the CFT, in particular the relative entropy, to elucidate local physics in the bulk. Related recent works include an attempt to constrain the nonlinear gravitational equations of motion using the positivity of relative entropy [10, 11], as well as the converse scenario of deriving differential equations constraining CFT entanglement from the nonlinear bulk Einstein equations [3, 6, 12].

Our starting point is a CFT whose vacuum state is dual to  $\text{AdS}_{d+1}$ . We consider an arbitrary excited state of the CFT which has a semiclassical holographic bulk dual, whose metric can be parametrized as

$$g_{\text{AdS}} = \frac{\ell_{\text{AdS}}^2}{z^2} [dz^2 + (\eta_{\mu\nu} + h_{\mu\nu}) dx^\mu dx^\nu]. \quad (11)$$

Spacetime indices  $a, b, \dots$  run over  $(t, z, x^i)$  while  $\mu, \nu, \dots$  run over  $(t, x^i)$  and  $i \in 1, \dots, d-1$  are boundary spatial directions. Moreover, we assume that the Ryu-Takayanagi formula holds in the excited state, and the relative entropy between the reduced density matrix  $\rho_1$  of the excited state and  $\rho_0$  of the ground state for the entangling disk  $D$  of radius  $R$  is computable using the formulae in the previous subsection.

To apply a perturbative analysis in the bulk, we assume that the radius  $R$  of the entangling domain is small compared to the typical energy scale  $\mathcal{E} \approx \langle T_{\mu\nu} \rangle^{\frac{1}{d}}$  of the state measured by the boundary stress tensor  $T_{\mu\nu}$ ,

$$\mathcal{E}^d R^d \ll 1. \quad (12)$$

Furthermore, we assume that the energy scales of the bulk fields are of the same order as the boundary stress-energy tensor energy scale, so that the geometry in the bulk region bounded by  $D$  and the associated Ryu-Takayanagi surface is close to that of  $\text{AdS}$ .

In this limit, to order less than  $\mathcal{E}^{2d} R^{2d}$ , we will show that the relative entropy is expressed as

$$S(\rho_1|\rho_0) = 8\pi^2 G_N \int_V \frac{R^2 - (z^2 + x^2)}{R} \varepsilon \sqrt{g_V}, \quad (13)$$

where  $G_N$  is Newton's constant,  $V$  is a  $d$ -dimensional region on a constant-time slice bounded by the domain  $D$  on the boundary and the Ryu-Takayanagi surface  $\Sigma$  in the bulk, and  $\sqrt{g_V}$  is the volume form in the bulk (including the time direction). In particular, the positivity of the relative entropy follows the weak energy condition, i.e. the positivity of the bulk energy density  $\varepsilon$  (see *e.g.* [13]). Though the weak energy condition is not necessarily satisfied in  $\text{AdS}$ , it holds near the boundary of  $\text{AdS}$ , where we are evaluating (13). We also note that the positivity is only required for the integrated quantity.

In [7], it was shown that the first law  $S(\rho_1|\rho_0) = 0$  in the linear approximation is equivalent to the linearized Einstein equation. This holds to the order  $O(\mathcal{E}^d R^d)$ . Our result (13) improves the approximation to the order less than  $\mathcal{E}^{2d} R^{2d}$  by taking into account the backreaction to the bulk stress tensor.

Taking one derivative with respect to  $R$ , the monotonicity of the relative entropy can also be related to a bulk energy condition:

$$\partial_R S(\rho_1|\rho_0) = 8\pi^2 G_N \int_V \left(1 + \frac{z^2 + x^2}{R^2}\right) \varepsilon \sqrt{g_V}. \quad (14)$$

One more derivative relates the relative entropy to the integral of the energy density on  $\Sigma$ ,

$$(\partial_R^2 + R^{-1} \partial_R - R^{-2}) S(\rho_1|\rho_0) = 16\pi^2 G_N \int_{\Sigma} \varepsilon \sqrt{g_{\Sigma}}, \quad (15)$$

where  $\sqrt{g_{\Sigma}}$  is the volume form on the Ryu-Takayanagi surface. We will show that (15) can be inverted to express the bulk stress tensor point-by-point in the near- $\text{AdS}$  region using the entanglement information of the CFT.

The outline of the paper is as follows. In Section II we review how to translate each quantity in (7) to holography, and sketch how to derive the linearized vacuum equations of motion from the entanglement first law using a generalized Stokes theorem argument [7, 14, 15]. In Section III we demonstrate that the positivity, monotonicity and convexity of the relative entropy for small spheres in the CFT are dual to the positivity conditions on the bulk stress tensor, and in Section IV we show how to invert (15) to obtain the bulk stress tensor locally in the near- $\text{AdS}$  region. In Section V, we discuss to what extent we can recover from a general quantum-theoretic analysis the convexity of the relative entropy (15) that was derived holographically. We comment on implications and open problems in Section VI.

## II. PRELIMINARIES

### A. $\Delta H_{mod}$ and $\Delta S_{EE}$ in holography

We first review how each quantity appearing in the definition of the relative entropy (7) is mapped holographically. The modular Hamiltonian  $H_{mod}$  for the reduced density matrix of the vacuum state of a CFT on the entangling disk  $D$  of radius  $R$ , centered at a point on the boundary, is a function of the CFT stress tensor  $\langle T_{tt} \rangle$  (10). It vanishes in the CFT vacuum. In the excited state,  $\langle T_{tt} \rangle$  in the CFT can be expressed as a function of bulk fields by using holographic renormalization (see *e.g.* [16–18]) or the shortcut of [7] to exploit the fact that the relative entropy in the CFT vanishes in the limit that the entangling domain shrinks to zero. As long as the bulk matter fields contributing to  $\langle T_{\mu\nu} \rangle$  are dual to operators with scaling dimension  $\delta > d/2$ , both methods give

$$\Delta \langle H_{mod} \rangle = \lim_{z \rightarrow 0} \frac{d\ell_{AdS}^2}{16G_N} \int_D d^{d-1}x \frac{R^2 - |\vec{x}|^2}{R} z^{-d} \eta^{ij} h_{ij}. \quad (16)$$

In general, the right-hand side is modified by boundary counter terms if it involves operators with  $\delta \leq d/2$ . We hope to generalize our result to such a case in future work.

The holographic EE in Einstein gravity is given by the Ryu-Takayanagi area formula (1). On a constant time slice of pure AdS, the codimension-2 bulk extremal surface  $\Sigma$  ending on a boundary sphere of radius  $R$  is the half-sphere

$$z_0(r) = \sqrt{R^2 - r^2}. \quad (17)$$

The EE of the entangling disk of radius  $R$  in the CFT vacuum is equal to the area functional of pure AdS evaluated on the surface (17). Suppose we perturb the bulk metric away from pure AdS by  $h_{ab}$  which is parametrically small. Because the original surface was extremal, the leading variation in the holographic EE comes from evaluating  $z_0(r)$  (17) on the perturbed area functional. One finds [5]

$$\Delta S_{EE} = \frac{\ell_{AdS}^{d-1}}{8G_N R} \int_{\Sigma} d^{d-1}x (R^2 \eta^{ij} - x^i x^j) z^{-d} h_{ij}. \quad (18)$$

At order  $h^2$ , one must account for corrections to the shape of the Ryu-Takayanagi surface, see *e.g.* [4].

### B. Linearized Einstein Equations

We now summarize the derivation of the linearized gravitational equations of motion from the entanglement first law (8), as presented in [7]. The idea of [7] was to apply the Stokes theorem to the bulk  $d$ -dimensional region  $V$  on a constant-time slice bounded by the entangling

disk  $D$  on the boundary and the extremal surface  $\Sigma$  in the bulk. One can write  $\Delta H_{mod}$  and  $\Delta S_{EE}$  as integrals over  $D$  and  $\Sigma$  respectively of a local  $d-1$  form  $\chi$  that is a functional of the metric fluctuation  $h_{ab}$ . Within Einstein gravity, [7, 14, 15] explicitly construct a  $\chi[h_{ab}]$  that gives (16) and (18) when integrated over  $B$  and  $\tilde{B}$ ,

$$\int_D \chi = \Delta \langle H_{mod} \rangle, \quad \int_{\Sigma} \chi = \Delta S_{EE}. \quad (19)$$

Moreover, the exterior derivative of this  $\chi$  is given by

$$d\chi = 2\xi^t E_{tt}^g[h] g^{tt} \sqrt{g_V} dz \wedge dx^{i_1} \cdots \wedge dx^{i_{d-1}}, \quad (20)$$

where  $\sqrt{g_V}$  is the natural volume form on  $V$  induced from the bulk spacetime metric, and

$$\begin{aligned} \xi^a = & -\frac{2\pi}{R} (t - t_0) [z \partial_z + x^i \partial_i] \\ & + \frac{\pi}{R} [R^2 - z^2 - (t - t_0)^2 - x^2] \partial_t \end{aligned} \quad (21)$$

is the Killing vector associated with  $\Sigma$  (17), which is a bifurcate Killing horizon in pure AdS.  $E_{ab}^g[h]$  are the linear gravitational equations of motion in vacuum.

By the Stokes theorem, the relative entropy is given by

$$S(\rho_1|\rho_0) = \Delta \langle H_{mod} \rangle - \Delta S_{EE} = \int_V d\chi. \quad (22)$$

Considering (22) for every disk on a spatial slice at fixed time  $t = 0$ , the entanglement first law  $S(\rho_1|\rho_0) = 0$  can then be shown to be equivalent to  $E_{tt}^g[h] = 0$ . Considering it for Lorentz-boosted frames gives vanishing of the other boundary components,  $E_{\mu\nu}^g[h] = 0$ . Finally, an argument appealing to the initial-value formulation gives vanishing of the remaining components of the linearized Einstein tensor that carry  $z$  indices.

To summarize, [7] proved the existence of a  $d-1$  form  $\chi$  as a functional of a metric fluctuation  $h_{ab}$ , for which (19) hold off shell and (20) holds with  $E_{ab}^g[h]$  the linearized gravity equations of motion in vacuum.

By accounting for the  $1/N$  correction to the Ryu-Takayanagi formula [19], [8] showed that the entanglement first law (8) implies the bulk linearized Einstein equations sourced by the *quantum* expectation value of the bulk stress-energy tensor,  $\langle t_{ab} \rangle$ . Assuming that the source of the linearized Einstein equation is a local QFT operator, one can then argue that the quantum expectation value appearing in their derivation can be uplifted to the bulk operator  $t_{ab}$ . In what follows, we remain in the large  $N$  classical gravity limit, but assume the linearized Einstein equations sourced by the *classical* value of the bulk stress tensor and use it to derive additional results.

## III. EFFECTS DUE TO BULK STRESS TENSOR

We now evaluate the  $(d-1)$ -form  $\chi$  of [7] on the bulk metric fluctuation  $h_{ab}$  of the dual to an arbitrary excited

state of a CFT, but in the interior of the Ryu-Takayanagi surface for the entangling disk (17), whose radius satisfies (12). As the deviation of the bulk metric in the enclosed volume  $V$  is parametrically small, all results of the above discussion carry over:

$$\Delta S_{EE} - \Delta \langle H_{mod} \rangle = \int_{\Sigma} \chi - \int_D \chi = \int_{\partial V} \chi = \int_V d\chi. \quad (23)$$

Here  $d\chi$  is given by (20), but  $E_{ab}^g[h]$  is now the linearized vacuum Einstein tensor evaluated on the  $h_{ab}$  which is reconstructed from CFT data at non-linear level. This  $E_{ab}^g[h_{ab}]$  is now not identically zero. Rather, the linearized Einstein tensor couples to bulk matter in the form of the bulk stress tensor  $t_{ab}$ . Using

$$E_{ab}^g[h_{ab}] = 8\pi G_N t_{ab}, \quad (24)$$

Eq. (23) becomes (see (20))

$$\begin{aligned} S(\rho_1|\rho_0) &= \Delta H_{mod} - \Delta S_{EE} \\ &= 8\pi G \int_V \xi^t \epsilon \sqrt{g_V}, \end{aligned} \quad (25)$$

where the energy density  $\epsilon$  appearing on the right corresponds to the  $tt$ -component of the stress-energy tensor,  $\epsilon = -t_t^t$ .

For example, a massive scalar field in the bulk can contribute to the metric perturbation  $h_{ab}$  as  $\langle \mathcal{O} \rangle^2 z^{2\Delta}$ , where  $\Delta$  is the scaling dimension of the corresponding operator on the boundary and  $\langle \mathcal{O} \rangle$  is its expectation value, leading to an  $O(\langle \mathcal{O} \rangle^2 R^{2\Delta})$  effect in (25). On the other hand, corrections to the relative entropy by non-linear gravity effects are of the order  $O(\mathcal{E}^{2d} R^{2d})$  or higher, which we ignore. Thus, effects due to relevant operators with  $\Delta < d$  are visible in our approximation.

Now, we can show that the monotonicity (5) of the relative entropy follows from a positivity condition on the bulk energy density. By taking a derivative of (25) with respect to the radius  $R$  of the entangling domain, we find

$$\partial_R S(\rho_1|\rho_0) = 8\pi G_N \int_{\Sigma} \xi^t \epsilon \sqrt{g_V} \quad (26)$$

$$\begin{aligned} &+ 8\pi^2 G_N \int_V \left(1 + \frac{x^2 + z^2}{R^2}\right) \epsilon \sqrt{g_V} \\ &= 8\pi^2 G_N \int_V \left(1 + \frac{x^2 + z^2}{R^2}\right) \epsilon \sqrt{g_V}. \end{aligned} \quad (27)$$

The integral over the Ryu-Takayanagi surface  $\Sigma$  vanishes because  $\xi^t$  (21) vanishes on the surface. Assuming the weak energy condition (*i.e.*, the positivity of the energy density), we find the inequality

$$\partial_R S(\rho_1|\rho_0) \geq 0. \quad (28)$$

Though the weak energy condition is not necessarily satisfied in AdS, it holds near the boundary of AdS. We also note that the positivity is only required for the integrated quantity.

#### IV. INVERTING THE BULK INTEGRAL

We found that  $\partial_R S(\rho_1|\rho_0)$  is given by the integral of the energy density  $\epsilon$  over the region  $V$  inside the Ryu-Takayanagi surface. We can invert this relation to compute  $\epsilon$  point-by-point in the bulk by using the relative entropy  $S(\rho_1|\rho_0)$ .

To show this, note that

$$(\partial_R + R^{-1}) S(\rho_1|\rho_0) = 16\pi^2 G_N \int_V \epsilon \sqrt{g_V} \quad (29)$$

so differentiating again,

$$(\partial_R^2 + R^{-1} \partial_R - R^{-2}) S(\rho_1|\rho_0) = 16\pi^2 G_N \int_{\Sigma} \epsilon \sqrt{g_{\Sigma}} \quad (30)$$

where  $\sqrt{g_{\Sigma}}$  is the natural volume form on the Ryu-Takayanagi surface  $\Sigma$  induced from the bulk spacetime metric. We note that the right-hand side is still non-negative if we assume the positivity of the bulk energy density. Thus,

$$(\partial_R^2 + R^{-1} \partial_R - R^{-2}) S(\rho_1|\rho_0) \geq 0. \quad (31)$$

Here the bulk geometry is the unperturbed AdS, and its space-like section is the  $d$ -dimensional hyperbolic space. The surface  $\Sigma$  is then totally geodesic. In this case, the integral (30) is the Radon transform and its inverse is known. For a smooth function  $f$  on  $d$ -dimensional hyperbolic space, the Radon transform  $\mathcal{R}f$  is an integral of  $f$  over a  $n$ -dimensional geodesically complete submanifold with  $n < d$ . This gives a function on the space of geodesically complete submanifolds. The *dual* Radon transform  $\mathcal{R}^* \mathcal{R}f$  gives back a function on the original hyperbolic space in the following way: pick a point in the hyperbolic space, consider all geodesically complete submanifolds passing through the point, and integrate  $\mathcal{R}f$  over such submanifolds.

It was shown by Helgason [20] that if  $d$  is odd,  $f$  is obtained by applying an appropriate differential operator on  $\mathcal{R}^* \mathcal{R}f$ . We are interested in the case  $n = d - 1$  for which

$$f = \left[ (-4)^{(d-1)/2} \pi^{d/2-1} \Gamma(d/2) \right]^{-1} Q(\Delta) \mathcal{R}^* \mathcal{R}f, \quad (32)$$

where  $Q(\Delta)$  is constructed from the Laplace-Beltrami operator  $\Delta$  on the hyperbolic space as

$$\begin{aligned} Q(\Delta) &= [\Delta + 1 \cdot (d-2)] [\Delta + 2 \cdot (d-3)] \\ &\times \cdots \times [\Delta + (d-2) \cdot 1]. \end{aligned} \quad (33)$$

Applying this to (30), we find

$$\begin{aligned} \epsilon &= \left[ (-4)^{(d+3)/2} \pi^{d/2+1} \Gamma(d/2) G_N \right]^{-1} \times \\ &\times Q(\Delta) \mathcal{R}^* (\partial_R^2 + R^{-1} \partial_R - R^{-2}) S(\rho_1|\rho_0), \end{aligned} \quad (34)$$

when  $d$  is odd. There is a similar formula when  $d$  is even [21]. The energy density is the time-time component of



the stress-energy tensor  $t_{tt}$ . By computing the relative entropy in other Lorentz frames, we can also derive components  $t_{\mu\nu}$  along the boundary. Finally, we can use the conservation law,  $\nabla^a t_{ab} = 0$ , to obtain the remaining components,  $t_{z\mu}, t_{\mu\nu}$ . Thus, we can use the entanglement data on the boundary to reconstruct all components of the bulk stress tensor.

Since the Radon transform preserves positivity, the positivity of the energy density implies the positivity of  $(\partial_R^2 + R^{-1}\partial_R - R^{-2})S(\rho_1|\rho_0)$ . Conversely, the positivity of the latter implies the positivity of its dual Radon transform. It is interesting to note that  $Q(\Delta)$  in (34) is a positive definite operator when acting on normalizable functions on the hyperbolic space, though this does not quite imply the positivity of the energy density.

## V. COMPARISON WITH INFORMATION THEORETIC ANALYSIS

In this section, we discuss to what extent we can recover the monotonicity and convexity (15) of the relative entropy from the following general property of the relative entropy. Consider a density matrix  $\rho$  (with  $\rho^* = \rho$ ,  $\rho \geq 0$ , and  $\text{tr}(\rho) = 1$ ), and two increments  $h, \ell$ , given by matrices with  $h = h^*$ ,  $\ell = \ell^*$  and  $\text{tr}(h) = \text{tr}(\ell) = 0$ . If the matrices  $\rho, h, \ell$  satisfy  $[\rho, h] = [\rho, \ell] = 0$ , then the relative entropy satisfies

$$S(\rho + h|\rho + \ell) \sim \langle (h - \ell), \frac{1}{2}\rho^{-1}(h - \ell) \rangle, \quad (35)$$

where the right-hand-side is the Fisher metric, with the Hilbert–Schmidt inner product  $\langle a, b \rangle = \text{tr}(a^*b)$ . Thus, the second order term is non-negative definite, and the quadratic form only vanishes for  $h = \ell$ .

The entanglement density matrices  $\rho(R)$  and  $\rho_0(R)$  discussed in this paper have additional properties for small  $R$ . Since  $H_{mod}$  is given by the integral (10) of  $\xi^t T_{tt}$  over  $|\vec{x} - \vec{x}_0| < R$ , the Taylor expansion of  $T_{tt}$  around  $\vec{x} = \vec{x}_0$  gives  $H_{mod} = h_0 R^d + \dots$ . Therefore, the density matrix for the vacuum state can be expanded as

$$\rho_0(R) = \frac{1}{\mathcal{N}} - h'_0 R^d + \dots, \quad (36)$$

where  $\text{tr} 1 = \mathcal{N}$  and  $h'_0 = h_0 - \frac{1}{\mathcal{N}}\text{tr} h_0$  so that  $\text{tr} h'_0 = 0$ . For  $\rho(R)$ , we postulate

$$\rho(R) = \frac{1}{\mathcal{N}} + \sum_i \ell_i R^{\delta_i} + h R^d + \dots, \quad (37)$$

so that the small  $R$  expansion of the relative entropy  $S = \sum_i R^{2\delta_i} s_i + \dots$  expected from the holographic computation above is reproduced. Here  $\text{tr} \ell_i = 0$  and  $\delta_i$ 's are scaling dimensions of relevant operators,  $\delta_i < d$ .

The right-hand-side of (35) becomes

$$\sum_{ij} \frac{\mathcal{N}}{2} \langle \ell_i, \ell_j \rangle R^{\delta_i + \delta_j}. \quad (38)$$

Thus, the leading order term of the relative entropy  $S(\rho_1|\rho_0)$  can be estimated as

$$S(\rho_1|\rho_0) \sim \frac{\mathcal{N}}{2} |\ell_1|^2 R^{2\delta_1}, \quad (39)$$

where  $\delta_1 = \min_i \{\delta_i\}$ . Its first and second derivatives in  $R$  have leading term

$$\begin{aligned} \partial_R S(\rho_1|\rho_0) &\sim \mathcal{N} \delta_1 |\ell_1|^2 R^{2\delta_1-1}, \quad (40) \\ (\partial_R^2 + R^{-1}\partial_R - R^{-2}) S(\rho_1|\rho_0) &\sim \frac{\mathcal{N}}{2} |\ell_1|^2 \times \\ &\times (4\delta_1^2 - 1) R^{2\delta_1-2}. \end{aligned} \quad (41)$$

The first is manifestly positive, and the second is non-negative provided  $\delta_1 \geq 1/2$ , which is satisfied by our assumption  $\delta_1 > d/2$  for  $d \geq 2$ .

Our holographic analysis shows that the positivity and the convexity of the relative entropy hold for subleading terms up to  $O(R^{2d})$ . On the other hand, corrections to (39) may involve not only quadratic terms with  $\delta_i + \delta_j < 2d$ , but also cubic terms with  $\delta_i + \delta_j + \delta_k < 2d$ , etc. It appears that additional assumptions on the density matrices are required to explain the convexity from this point of view.

## VI. DISCUSSION

We conclude with a few comments on prospects for future work.

In this paper, we have focused on bulk theories of classical Einstein gravity. However, it is probable that our result can be extended to higher-derivative classical gravities, as [7] has constructed the equivalent of the  $d - 1$  form  $\chi$  in such theories. Another interesting question concerns generalizing away from the classical limit. If we were to add the  $1/N$  correction to the Ryu-Takayanagi formula [19], we would appear to obtain constraints on the quantum energy density from the positivity of the relative entropy. We leave these analyses to future work.

Another obvious question involves going beyond the small  $R$  limit (12) that we have taken in this paper.

## Acknowledgements

We thank N. Bao, D. Harlow, T. Hartman, P. Hayden, N. Hunter-Jones, C. Keller, D. Kutasov, H. Liu, Y. Nakayama, S. Pufu, P. Sulkowski, T. Takayanagi, M. Van Raamsdonk and E. Witten for useful discussion. JL acknowledges support from the Sidney Bloomenthal fellowship at the University of Chicago. MM is currently supported by NSF grants PHY-1205440, DMS-1201512, and DMS-1007207. HO and BS are supported in part by the Walter Burke Institute for Theoretical Physics at Caltech, by U.S. DOE grant DE-SC0011632, and by a Simons Investigator award from the Simons Foundation. The work of HO is also supported in part by the WPI Initiative of MEXT of Japan, and JSPS Grant-in-Aid for Scientific Research C-26400240. He also thanks

the hospitality of the Aspen Center for Physics and the National Science Foundation, which supports the Center under Grant No. PHY-1066293. BS is supported in part by a Dominic Orr Graduate Fellowship. JL, HO and BS

would like to thank the Institute for Advanced Study, Princeton University, and the Simons Center for Geometry and Physics for hospitality. JL also thanks Caltech for hospitality.

- 
- [1] S. Ryu and T. Takayanagi, *Phys.Rev.Lett.* **96**, 181602 (2006), [arXiv:hep-th/0603001 \[hep-th\]](#).
  - [2] S. Ryu and T. Takayanagi, *JHEP* **0608**, 045 (2006), [arXiv:hep-th/0605073 \[hep-th\]](#).
  - [3] M. Nozaki, T. Numasawa, A. Prudenziati, and T. Takayanagi, *Phys.Rev.* **D88**, 026012 (2013), [arXiv:1304.7100 \[hep-th\]](#).
  - [4] D. D. Blanco, H. Casini, L.-Y. Hung, and R. C. Myers, *JHEP* **1308**, 060 (2013), [arXiv:1305.3182 \[hep-th\]](#).
  - [5] N. Lashkari, M. B. McDermott, and M. Van Raamsdonk, *JHEP* **1404**, 195 (2014), [arXiv:1308.3716 \[hep-th\]](#).
  - [6] J. Bhattacharya and T. Takayanagi, *JHEP* **1310**, 219 (2013), [arXiv:1308.3792 \[hep-th\]](#).
  - [7] T. Faulkner, M. Guica, T. Hartman, R. C. Myers, and M. Van Raamsdonk, *JHEP* **1403**, 051 (2014), [arXiv:1312.7856 \[hep-th\]](#).
  - [8] B. Swingle and M. Van Raamsdonk, (2014), [arXiv:1405.2933 \[hep-th\]](#).
  - [9] H. Casini, M. Huerta, and R. C. Myers, *JHEP* **1105**, 036 (2011), [arXiv:1102.0440 \[hep-th\]](#).
  - [10] S. Banerjee, A. Bhattacharyya, A. Kaviraj, K. Sen, and A. Sinha, *JHEP* **1405**, 029 (2014), [arXiv:1401.5089 \[hep-th\]](#).
  - [11] S. Banerjee, A. Kaviraj, and A. Sinha, (2014), [arXiv:1405.3743 \[hep-th\]](#).
  - [12] J. Bhattacharya, M. Nozaki, T. Takayanagi, and T. Ugajin, *Phys.Rev.Lett.* **110**, 091602 (2013), [arXiv:1212.1164](#).
  - [13] S. Hawking and G. Ellis, *The Large Scale Structure of Space-Time*, Cambridge Monographs on Mathematical Physics (Cambridge University Press, 1973).
  - [14] R. M. Wald, *Phys. Rev. D* **48**, 3427 (1993), [gr-qc/9307038](#).
  - [15] V. Iyer and R. M. Wald, *Phys.Rev.* **D50**, 846 (1994), [arXiv:gr-qc/9403028 \[gr-qc\]](#).
  - [16] V. Balasubramanian and P. Kraus, *Commun.Math.Phys.* **208**, 413 (1999), [arXiv:hep-th/9902121 \[hep-th\]](#).
  - [17] S. de Haro, S. N. Solodukhin, and K. Skenderis, *Commun.Math.Phys.* **217**, 595 (2001).
  - [18] K. Skenderis, *Class.Quant.Grav.* **19**, 5849 (2002), [arXiv:hep-th/0209067 \[hep-th\]](#).
  - [19] T. Faulkner, A. Lewkowycz, and J. Maldacena, *JHEP* **1311**, 074 (2013), [arXiv:1307.2892](#).
  - [20] S. Helgason, *Acta Math.* **102**, 239 (1959).
  - [21] B. Rubin, *Adv. Math* **170**, 206 (2002).