PHYSICAL REVIEW D **82,** 126010 (2010)

Entanglement Rényi entropies in holographic theories

Matthew Headrick*

Martin Fisher School of Physics, Brandeis University, Waltham Massachusetts 02453, USA

(Received 31 August 2010; published 22 December 2010)

Ryu and Takayanagi conjectured a formula for the entanglement (von Neumann) entropy of an arbitrary spatial region in an arbitrary holographic field theory. The von Neumann entropy is a special case of a more general class of entropies called Rényi entropies. Using Euclidean gravity, Fursaev computed the entanglement Rényi entropies (EREs) of an arbitrary spatial region in an arbitrary holographic field theory, and thereby derived the RT formula. We point out, however, that his EREs are incorrect, since his putative saddle points do not in fact solve the Einstein equation. We remedy this situation in the case of two-dimensional conformal field theories (CFTs), considering regions consisting of one or two intervals. For a single interval, the EREs are known for a general CFT; we reproduce them using gravity. For two intervals, the RT formula predicts a phase transition in the entanglement entropy as a function of their separation, and that the mutual information between the intervals vanishes for separations larger than the phase transition point. By computing EREs using gravity and CFT techniques, we find evidence supporting both predictions. We also find evidence that large N symmetric product theories have the same EREs as holographic ones.

DOI: 10.1103/PhysRevD.82.126010 PACS numbers: 11.25.Tq, 03.67.Mn, 11.25.Hf

I. INTRODUCTION

The concept of holography originated as an idea about quantum information, that the number of qubits that can be stored in a region of space is fundamentally limited by its surface area in Planck units. Modern holographic theories go beyond a mere counting of states, and posit that the physics governing certain spacetimes can be fully described by a quantum field theory residing on its boundary. However, the way that those qubits are organized remains unclear on both sides of the correspondence. On one side, we do not yet understand how the states are organized in quantum gravity; on the other, despite an in-principle understanding of the state space of quantum field theories, in practice we have to deal with a strongly coupled theory with a large number of degrees of freedom. And, of course, the map between the two descriptions remains deeply mysterious.

A useful probe of physical information in quantum systems is the entanglement entropy (EE). Here we imagine decomposing a system into two subsystems, A, A^c, with a corresponding decomposition of the Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_{A^c}$. Given a density matrix ρ for the full system, the reduced density matrix ρ_A, which acts on $\mathcal{H}_A$, is defined by tracing ρ over $\mathcal{H}_{A^c}$ and represents the effective density matrix for an observer who has access only to the subsystem A. The EE for A is then the von Neumann entropy of ρ_A: $S_A \equiv -\mathrm{tr}(\rho_A \ln \rho_A)$. A nonzero EE may be due to the full system being in a mixed state, to information about the state being lost by the inability to observe the rest of the system, or to a combination of the two effects. The degree of correlation (both classical and quantum) between disjoint subsystems may be quantified by their mutual information $I_{A,B} \equiv S_A + S_B - S_{A\cup B}$, which puts an upper bound on correlators between operators in A and in B [1].

In a quantum field theory, it is natural to consider subsystems that are spatial regions. Their EEs and mutual informations then tell us about the spatial distribution and correlations of quantum information in a given state. Unfortunately, EEs in quantum field theories are notoriously difficult to calculate, mainly because one does not have a good way to represent the operator $\ln \rho_A$. On the other hand, if the density matrix for the full system can be represented by a path integral (as in the vacuum or a thermal ensemble, for example), then both the reduced density matrix ρ_A and its positive integer powers ρ_A^n can also be represented in a fairly simple way by path integrals. If those path integrals can be computed explicitly for all n, then one can obtain the EE indirectly as follows. Defining the *entanglement Rényi entropy* (ERE) $S_A^{(n)} \equiv (\ln \mathrm{tr} \rho_A^n)/(1-n)$ for $n > 1$, one analytically continues $S_A^{(n)}$ in n and takes the limit $n \to 1$ to obtain the EE. This procedure is called the replica trick. Aside from being easier to calculate than the EE, the EREs are of interest in their own right, as a more refined characterization of the reduced density matrix ρ_A. In fact, knowing $S_A^{(n)}$ for all n is equivalent to knowing the full eigenvalue distribution of ρ_A. In Sec. II, we review the basic properties of entanglement and Rényi entropies.

Even given the replica trick, exact results for the EE in field theories are known only in very simple cases, such as a single interval in the vacuum of an arbitrary two-dimensional conformal field theory [2]. For two disjoint intervals, the EE, and hence mutual information, remain unknown even for a theory as simple as that of a compact free scalar [3]. It might therefore seem hopeless to dream of knowing the EE in a strongly coupled, large N field theory. Remarkably, however, Ryu and Takayanagi (RT)

*mph@brandeis.edu

proposed a simple, elegant, and universal formula for the EE of an arbitrary spatial region in an arbitrary holographic field theory [4,5]. Their formula, which applies to any state described by a static classical geometry, says simply that the EE equals one-quarter the area in Planck units of the minimal surface in the bulk ending on the boundary of the region A. If correct, the RT formula is not only very useful as a calculational tool, but also a significant hint regarding quantum information in holographic theories, and probably in quantum gravity more generally (see for example [6]).

The RT formula passes several nontrivial checks. For example, it correctly reproduces the EE for a single interval in a two-dimensional CFT. A general derivation, using the replica trick, was offered by Fursaev [7]. He found that the ERE $S_A^{(n)}$ equaled one-quarter the minimal-surface area, independent of n. The analytic continuation in n was thus trivial, giving agreement between the resulting value of S_A and the RT formula. In computing the ERE, Fursaev performed the necessary path integrals using Euclidean quantum gravity. Unfortunately, as we show, the bulk geometries that he used to evaluate the partition function are not actually saddle points of the gravitational action. As a result, the ERE he derived is incorrect, as we can see by comparing it to the known exact result in the case of a single interval in a two-dimensional CFT. We show how the latter result can be reproduced using the correct saddle point action. The RT formula and Fursaev's proof are reviewed and discussed in Sec. III.

The question thus arises of whether, in cases where the correct value is not already known, we can compute the ERE in a holographic theory, both for its own sake and in order to confirm or refute the RT conjecture. Unfortunately, to do so in complete generality, as Fursaev attempted, appears to be quite difficult. Therefore, in Sec. IV, we focus on a simple but nontrivial case: two disjoint intervals in a two-dimensional CFT. The RT formula predicts a rather interesting phase transition for the mutual information between the two intervals as a function of their separation. In particular, for separations larger than a certain critical value, the mutual information vanishes, implying a decoupling between the degrees of freedom in the two regions. (This behavior of the mutual information is a completely general prediction of the RT formula, applying essentially to any two regions in any state of any holographic theory. It is closely analogous to the factorization property for disconnected Wilson loops [8].)

The ERE for two disjoint intervals can be expressed in terms of the partition function on a certain Riemann surface of genus $n-1$. For $n=2$, we thus need the torus partition function, which fortunately is known for a general holographic CFT [9]. Indeed, as we show, it exhibits a phase transition at precisely the same separation as that predicted for the EE by the RT formula. For higher values of n, while the partition function is not known explicitly, we show using symmetry arguments that the ERE continues to have a phase transition at the same separation. This strongly suggests that the same will hold for $n=1$, confirming this prediction of the RT formula.

The fact that we can compute the ERE explicitly only for $n=2$ precludes analytically continuing it to $n=1$, to directly confirm or refute the full EE predicted by the RT formula. We therefore pursue a different strategy. Using the operator product expansion (OPE), we expand the ERE, for any given n, in powers of the inverse separation between the intervals. The coefficient of any given power can be computed explicitly for all n using formulas for conformal blocks, and analytically continued to $n=1$. We carry this out for a number of coefficients, finding that, thanks to a rather intricate pattern of cancellations, in each case the continuation to $n=1$ vanishes, precisely as predicted by the RT formula.

As a by-product of our analysis of the ERE for two disjoint intervals, we find that the result for certain nonholographic CFTs with large central charges, such as large N symmetric product theories, is precisely the same as for holographic ones. It seems that there is some form of large c universality operating here, with a large class of such CFTs having identical EREs. This possible feature of the ERE deserves further study.

We conclude in Sec. V with a list of open questions and possible generalizations of our work, and some remarks concerning our current understanding of the RT formula.

An Appendix contains some calculations in symmetric product theories whose results are used in the main text.

II. ENTANGLEMENT RÉNYI ENTROPY: REVIEW

In Subsection II A we briefly motivate, define, and state (without proof) the important properties of the entanglement Rényi entropy and mutual Rényi information. In Subsection II B, we illustrate these ideas in the simple example of two subsystems that are weakly coupled to each other. In Subsection II C we then briefly review the replica trick for computing the entanglement Rényi entropy, and in Subsection II D apply it to the simplest field theory example, a single interval in a two-dimensional conformal field theory. For more details, we refer the reader to the books [10,11] and the review [12]; the latter provides a comprehensive introduction to Rényi and entanglement entropies in two-dimensional CFTs.

A. Basic definitions and properties

Given a density matrix ρ and a positive real number $\alpha \neq 1$, the Rényi entropy is defined as[1]

[1]Writing $e^{-S^{(\alpha)}} = \langle \rho^{\alpha-1} \rangle_\rho^{1/(\alpha-1)}$, we see that the definition of the Rényi entropy is similar to that of the L^p norm of a positive function, $\|f\|_p \equiv (\int f^p)^{1/p}$. The difference is that, whereas in the L^p norm we evaluate the integral with respect to a fixed measure, in the Rényi entropy we evaluate it with respect to the very density matrix whose entropy we are computing.

$$S^{(\alpha)} \equiv \frac{1}{1-\alpha} \ln \mathrm{tr} \rho^{\alpha}. \tag{2.1}$$

At $\alpha = 1$ the Rényi entropy is defined by taking the limit, and equals the von Neumann entropy:

$$S = S^{(1)} \equiv \lim_{\alpha \to 1} S^{(\alpha)} = -\mathrm{tr}(\rho \ln \rho). \tag{2.2}$$

Two other interesting limits are $\lim_{\alpha \to 0} S^{(\alpha)} = \ln \dim \mathcal{H}_{\text{occupied}}$, where $\mathcal{H}_{\text{occupied}}$ is the image of ρ, and $\lim_{\alpha \to \infty} S^{(\alpha)} = -\ln \rho_{\max}$, called the min-entropy, where $\rho_{\max}$ is the largest eigenvalue of ρ. The following properties of $S^{(\alpha)}$ are straightforward to prove: (1) $S^{(\alpha)} \geq 0$, with equality if and only if ρ represents a pure state; (2) $S^{(\alpha)}$ is constant if and only if ρ is proportional to the identity on $\mathcal{H}_{\text{occupied}}$, and is otherwise a decreasing function of α; (3) for $\alpha > 1$ it satisfies $S^{(\alpha)} \leq \alpha(\alpha-1)^{-1} S^{(\infty)}$.

If the system contains a subsystem A—for example, in a field theory, A could be a spatial region[2]—then the Hilbert space $\mathcal{H}$ can be expressed as the tensor product of Hilbert spaces corresponding to A and to its complement A^c: $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_{A^c}$. Let $\rho_A \equiv \mathrm{tr}_{\mathcal{H}_{A^c}} \rho$ be the reduced density matrix, defined in $\mathcal{H}_A$, obtained by tracing ρ over $\mathcal{H}_{A^c}$; this is the effective density matrix for an observer who has access only to A. Its Rényi entropy $S_A^{(\alpha)} = (1-\alpha)^{-1} \times \ln \mathrm{tr}_A \rho_A^{\alpha}$ is called the *entanglement Rényi entropy* (ERE) of A, with the special case $S_A \equiv S_A^{(1)}$ simply called the *entanglement entropy* (EE). It can be shown that, if the full theory is in a pure state, then $S_A^{(\alpha)} = S_{A^c}^{(\alpha)}$.

The EE (but not the ERE for $\alpha \neq 1$) satisfies an important property called *strong subadditivity* [13,14], namely, for any two subsystems (or spatial regions) C and D,

$$S_C + S_D \geq S_{C \cup D} + S_{C \cap D}, \qquad S_C + S_D \geq S_{C \setminus D} + S_{D \setminus C}. \tag{2.3}$$

Strong subadditivity in fact characterizes EE, in the sense that any measure of entanglement that satisfies (2.3) for all subsystems C and D (as well as certain basic reasonableness requirements) must equal the EE [15,16]. As a special case, strong subadditivity implies *subadditivity*, namely, for disjoint subsystems A and B,[3]

$$|S_A - S_B| \leq S_{A \cup B} \leq S_A + S_B. \tag{2.4}$$

The second inequality is saturated if and only if the density matrix $\rho_{A \cup B}$ factorizes: $\rho_{A \cup B} = \rho_A \otimes \rho_B$. Motivated partly by this fact, the *mutual information* (MI) is defined by

$$I_{A,B} = S_A + S_B - S_{A \cup B}, \tag{2.5}$$

which quantifies the extent to which the degrees of freedom of A and B are correlated with each other, including both quantum entanglement and classical correlations. For example, the MI puts an upper bound on the correlator between any operator $\mathcal{A}_A$ in subsystem A and any operator $\mathcal{A}_B$ in subsystem B [1]:

$$\frac{(\langle \mathcal{A}_A \mathcal{A}_B \rangle - \langle \mathcal{A}_A \rangle \langle \mathcal{A}_B \rangle)^2}{2 \langle \mathcal{A}_A^2 \rangle \langle \mathcal{A}_B^2 \rangle} \leq I_{A,B}. \tag{2.6}$$

As a consequence of strong subadditivity, the mutual information is monotone under restriction: if $B' \subset B$ then $I_{A,B'} \leq I_{A,B}$.

The natural generalization of the mutual information defines the *mutual Rènyi information* (MRI):

$$I_{A,B}^{(\alpha)} \equiv S_A^{(\alpha)} + S_B^{(\alpha)} - S_{A \cup B}^{(\alpha)}. \tag{2.7}$$

Unlike the MI, the MRI is not necessarily positive. However, it is nonzero only when $\rho_{A \cup B} \neq \rho_A \otimes \rho_B$, and in this sense still quantifies the extent of correlation between A and B.

Another reason to study the MRI (including the MI) is that, when we are considering a field theory, it is universal, whereas the ERE (including the EE) is cutoff- or regulator-dependent. Specifically, when A is a spatial region, $S_A^{(\alpha)}$ usually suffers from an ultraviolet divergence proportional to the area of the boundary of A. However, if the two regions A and B are disjoint and mutually disconnected, then those divergences cancel in the MRI. Since the UV regulator generally violates conformal invariance, it follows that in conformal field theories the MRI is generally conformally invariant while the ERE is not. Also, in the CFT case the ERE suffers from an infrared divergence when one of the regions is infinite in size, but this cancels in the MRI (although not when both A and B are infinite). We will see explicit examples of these statements throughout this paper.

In the field theory context, the restriction monotonicity of the MI implies that the bound (2.6) will be strongest when the operators $\mathcal{A}_A$, $\mathcal{A}_B$ "cover" the regions A, B, respectively. In particular, the bound is trivially satisfied for local operators, since then the denominator on the left-hand side is generally divergent.[4]

B. Perturbative MRI

Before tackling the computation of entanglement entropies in field theories, as a warm-up we first consider the

[2]By *region* we technically mean codimension zero submanifold (possibly with boundary). In this paper we will not consider other types of sets, such as single points, fractals, etc.

[3]The ERE satisfies $|S_A^{(\alpha)} - S_B^{(\alpha)}| \leq S_{A \cup B}^{(\alpha)}$ for classical distributions, but not in general for quantum density matrices.

[4]One could worry about the case where the field theory has a UV cutoff—so the denominator on the left-hand side of (2.6) is not strictly infinite—but the MI vanishes exactly. In this case the bound is satisfied for a different reason, namely, that the density matrix factorizes, so all correlators factorize and the numerator vanishes. As far as we know, such a situation can only occur if the space where the field theory lives is disconnected. We thank M. van Raamsdonk for pointing out this possibility.

perturbative computation for subsystems that are weakly coupled to each other. We will see that in this case the $\alpha \neq 1$ MRI between the subsystems is parametrically larger than the MI, a result that foreshadows the results of Sec. IV concerning holographic systems.

We begin by considering a single system, and the effect on its Rényi entropy of a small perturbation to its density matrix:

$$\rho = \rho_{(0)} + \lambda \rho_{(1)}, \tag{2.8}$$

where $\mathrm{tr}\rho_{(0)} = 1$, $\mathrm{tr}\rho_{(1)} = 0$, and λ is a small parameter. To first order in λ we have

$$S^{(\alpha)} = \frac{1}{1-\alpha} \ln \mathrm{tr} \rho_{(0)}^{\alpha} + \lambda \frac{\alpha}{1-\alpha} \frac{\mathrm{tr}(\rho_{(1)} \rho_{(0)}^{\alpha-1})}{\mathrm{tr}\rho_{(0)}^{\alpha}} + O(\lambda^2) \tag{2.9}$$

$$S = -\mathrm{tr}(\rho_{(0)} \ln \rho_{(0)}) - \lambda \, \mathrm{tr}(\rho_{(1)} \ln \rho_{(0)}) + O(\lambda^2). \tag{2.10}$$

Now suppose our system is composed of two subsystems, and the unperturbed density matrix factorizes:

$$\rho_{(0)} = \hat{\rho}_A \otimes \hat{\rho}_B. \tag{2.11}$$

(For example, A and B could represent distinct regions of momentum space, and λ the coupling constant, in a weakly-coupled field theory.) At zeroth order in λ the MRI of course vanishes. To first order we have

$$I_{A,B}^{(\alpha)} = \lambda \frac{\alpha}{\alpha-1} \mathrm{tr}\left(\rho_{(1)} \left(\frac{\hat{\rho}_A^{\alpha-1}}{\mathrm{tr}\hat{\rho}_A^{\alpha}} \otimes I_B + I_A \otimes \frac{\hat{\rho}_B^{\alpha-1}}{\mathrm{tr}\hat{\rho}_B^{\alpha}} - \frac{\hat{\rho}_A^{\alpha-1}}{\mathrm{tr}\hat{\rho}_A^{\alpha}} \otimes \frac{\hat{\rho}_B^{\alpha-1}}{\mathrm{tr}\hat{\rho}_B^{\alpha}} \right) \right) + O(\lambda^2). \tag{2.12}$$

For $\alpha \approx 1$, the operator in the inner parentheses is proportional to $(\alpha-1)^2$. Hence, to first order in λ, the MI vanishes:

$$I_{A,B} = O(\lambda^2). \tag{2.13}$$

It can be shown that the order λ^2 term generically does not vanish. Note that the left-hand side of the inequality (2.6) is at most of order λ^2, so it is reasonable that the right-hand side would be of the same order.

C. Replica trick

Unfortunately, in practice there are very few known methods for computing EREs (or EEs) in field theories. One of the most useful is the *replica trick*, which we will review below, that allows one to compute the ERE $S^{(n)}$ for integer $n > 1$ [2]. In favorable circumstances a simple analytic form for $S^{(\alpha)}$ for general real α can be found which fits those data points, and from this form the EE can be read off by setting $\alpha = 1$.[5] It is important to say at the outset that in proceeding this way we are merely presuming to have guessed the ERE $S^{(\alpha)}$ correctly; first, nothing guarantees (in an infinite-dimensional Hilbert space) that the ERE is analytic, and, second, the values of a function on a countable infinite set (in this case, the integers larger than 1) are not sufficient to fix a unique analytic continuation. (There exist analytic functions, such as $(1-\alpha)^{-1}\sin\pi\alpha$, that vanish for all integer $\alpha > 1$ but not elsewhere, including at $\alpha = 1$.) Having stated this caveat, for the rest of the paper we will assume that all EREs we consider are indeed analytic functions of α. We will find nothing inconsistent with this assumption.

The replica trick applies when the theory is in a state, such as the vacuum or a thermal state, whose partition function can be obtained by a path integral over some Euclidean spacetime E (possibly with some operator insertions, which for the purposes of this discussion we will consider to be part of E). Let A be a spatial region, and E_n the n-sheeted cover of E with the sheets connected along branch cuts placed at A on a constant Euclidean-time surface. Then $\mathrm{tr}_A \rho_A^n = Z_n / Z_1^n$, where Z_n is the partition function of the theory on E_n (and, in particular, Z_1 is the partition function for the original theory).[6] Hence we have (for $n > 1$)

$$S_A^{(n)} = \frac{1}{1-n} \ln\left(\frac{Z_n}{Z_1^n} \right). \tag{2.14}$$

To be more concrete, let us further specialize to a two-dimensional conformal field theory $\mathcal{C}$,[7] and let A be the union of N disjoint intervals $[u_i, v_i]$, where $u_i < v_i < u_{i+1}$. Then we can rewrite the expression (2.14) in terms of correlators of twist operators in the symmetric product theory $\mathcal{C}^n/S_n$ (or equivalently $\mathcal{C}^n/\mathbf{Z}_n$), computed on E:

[5] Throughout this paper α will lie in the interval $[0, \infty]$, while n will be a positive integer.

[6] If the theory contains fermions then one needs to specify their boundary conditions across the constant Euclidean-time surface where the sheets are sewn together, which we will call S. The original partition function Z_1 is computed with a sign flip on the fermionic fields across S. More generally, Z_n is computed with a flip on $S \cap A^c$ on each sheet of E_n, along with one on $S \cap A$ where the nth sheet connects to the first sheet (but not on the other $n-1$ copies of $S \cap A$). Hence a curve that winds around a branch point n times, returning to the same point on E_n, crosses $n+1$ sign flips. The resulting overall flip for even n is part of the definition of the twist operators σ_1 and σ_{-1} of the next paragraph. When Z_n is evaluated by passing to a coordinate system that is single-valued on E_n, this overall flip is canceled by the branch cut in the coordinate transformation for the fermionic field. (For example, in complex coordinates if z is a local coordinate on E with the branch point at $z = 0$, and $t = z^{1/n}$ is a single-valued local coordinate on E_n, then $\psi_t = (dz/dt)^{1/2}\psi_z = n^{-1/2} t^{(1-n)/2} \psi_z$. For even n the factor $t^{(n-1)/2}$ has a branch cut with a sign flip.) Hence there is no operator insertion in the new coordinate system. However, there may still be sign flips around noncontractible cycles. (See, for example, the case of the torus in footnote 14 below.)

[7] All CFTs will be assumed compact, unitary, and modular-invariant in this paper.

$$S_A^{(n)} = \frac{1}{1-n} \ln\langle \sigma_1^\epsilon(u_1)\sigma_{-1}^\epsilon(v_1)\dots\sigma_1^\epsilon(u_N)\sigma_{-1}^\epsilon(v_N)\rangle. \tag{2.15}$$

The correlator of twist operators is divergent, due to the singular geometry of E_n at the branch point. It can be regularized by regularizing each twist operator separately; hence the notation σ_1^ϵ and σ_{-1}^ϵ, where ϵ is a UV cutoff length, for the regularized twist operators [17].

D. Single interval in a CFT

As an example of the application of (2.15), the ERE for a single interval, in the vacuum, is

$$\begin{aligned} S_{[u,v]}^{(n)} &= \frac{1}{1-n} \ln\langle \sigma_1^\epsilon(u)\sigma_{-1}^\epsilon(v)\rangle \\ &= \frac{c}{6}\left(1+\frac{1}{n}\right)\ln\left(\frac{v-u}{\epsilon}\right) + c_n, \end{aligned} \tag{2.16}$$

where c is the central charge of $\mathcal{C}$ and c_n is a scheme-dependent quantity. Here we used the fact that the twist operators have scaling dimension

$$d_\sigma = \frac{c}{12}\left(n - \frac{1}{n}\right). \tag{2.17}$$

The (simplest) analytic continuation of (2.16) to noninteger α is[8]

$$S_{[u,v]}^{(\alpha)} = \frac{c}{6}\left(1+\frac{1}{\alpha}\right)\ln\left(\frac{v-u}{\epsilon}\right) + c_\alpha, \tag{2.20}$$

which yields the EE [2]

$$S_{[u,v]} = \frac{c}{3}\ln\left(\frac{v-u}{\epsilon}\right) + c_1. \tag{2.21}$$

[8]It is interesting to ask what eigenvalue distribution for $\rho_{[u,v]}$ gives rise to the α-dependence $S_{[u,v]}^{(\alpha)} = (1+1/\alpha)C$ seen in (2.20) (neglecting the subleading and scheme-dependent quantity c_α). This question can be answered by defining a fictional "Hamiltonian" $\hat{H} \equiv -\ln\rho_{[u,v]}$ acting on $\mathcal{H}_{[u,v]}$. Then $S_{[u,v]}^{(\alpha)}$ is related to the free energy of $\hat{H}$ at the temperature α^{-1}:

$$F = -\frac{1}{\alpha}\ln \mathrm{tr} e^{-\alpha\hat{H}} = \left(1-\frac{1}{\alpha}\right)S_A^{(\alpha)} = \left(1-\frac{1}{\alpha^2}\right)C. \tag{2.18}$$

(The first two equalities apply to the Rényi entropy of any system.) The density of states that gives rise to this temperature dependence for the free energy is easily found, in the saddle point approximation, by performing a Legendre transform:

$$\rho(\hat{E}) = \begin{cases} 0, & \hat{E} < C \\ \exp(2C^{1/2}(\hat{E}-C)^{1/2}), & \hat{E} > C \end{cases}, \tag{2.19}$$

where $\hat{E}$ is the eigenvalue of $\hat{H}$. (See [18] or [12] for the form of the full inverse Laplace transform.) It is interesting that (up to a shift of $\hat{E}$ by C) $\rho(\hat{E})$ has the same form as the Cardy formula for the asymptotic density of states in a CFT on a circle. Note that we have not determined which physical observable $\hat{H}$ represents—it is not necessarily related to a physical energy.

Note that $S_{[u,v]}^{(\alpha)}$ indeed satisfies the properties (1), (2), (3) mentioned below Eq. (2.2).

It is also possible to obtain the result (2.16) (and thereby derive the scaling dimension (2.17)) by computing Z_n and applying (2.14). The computation of Z_n is done as follows [17]. We are in the vacuum, so the Euclidean spacetime E is simply the plane, to which we add a point at infinity to make it topologically a sphere. The multisheeted surface E_n is then also topologically a sphere. A Weyl transformation maps the metric ds^2 on E_n to a fiducial metric $d\hat{s}^2 = e^{-\phi}ds^2$ on the sphere. We then have $Z_n = e^{S_L}\hat{Z}$, where $\hat{Z}$ is the partition function of $\mathcal{C}$ on the sphere with the fiducial metric, and S_L is the Liouville action:

$$S_L = \frac{c}{96\pi}\int \hat{g}^{1/2}\left(\hat{g}^{\mu\nu}\partial_\mu\phi\partial_\nu\phi + 2\hat{R}\phi\right) \tag{2.22}$$

(which depends on $\mathcal{C}$ only through its central charge). For $n > 1$ the metric on E_n has a conical singularity at each branch point u, v, so the Liouville action is divergent. The divergence can be regulated by replacing a disc of radius ϵ about each branch point with a smooth metric, which defines the regularized twist operators $\sigma_{\pm 1}^\epsilon$.

III. HOLOGRAPHIC ENTANGLEMENT ENTROPIES

A. Ryu-Takayanagi formula

In this subsection we will provide a brief summary of Ryu and Takayanagi's proposal for the entanglement entropy (EE) in field theories with holographic duals [4,5], along with some of the evidence supporting it. A more complete review can be found in [19].

The Ryu-Takayanagi (RT) conjecture is a proposed formula for the EE of a given spatial region A in certain states of holographic field theories whose dual gravitational theory is classical Einstein gravity (possibly with matter). Specifically, the proposal concerns states that are described in the dual theory by static classical solutions; this includes, for example, the vacuum and thermal states.[9] We work in a fixed constant-time (i.e. timelike-Killing-field orthogonal) slice of the bulk. The conjecture states that

$$S_A = \frac{\mathrm{area}(m_A)}{4G_N}, \tag{3.1}$$

where m_A is the minimal-area surface in the bulk that is homologous to A, i.e. such that there exists a region r_A with $\partial r_A = A \cup m_A$. (We use the term "homologous" loosely, given that A is not generally closed. This topological condition was originally suggested by Fursaev [7], and plays a crucial role in several checks of the proposal.) The area is evaluated with respect to the Einstein-frame metric.

[9]Possible generalizations to time-dependent states were proposed in [20].

An interesting question, assuming the RT formula is valid, is how it gets corrected by quantum effects and by higher-derivative (e.g. α') corrections to the classical action in the bulk. Quantum effects presumably lead to G_N corrections to (3.1) (starting at order $G_{\rm N}^0$), although a specific form has not been proposed. In the presence of higher-derivative corrections to the classical bulk action, it is expected, based on consistency with black-hole entropy (discussed below), that in the formula (3.1) the area is replaced by Wald's black-hole entropy formula [21] (or at least some functional that coincides with it on horizons).

The RT proposal passes several basic checks. For example, if A is the entire boundary, then S_A should simply be the statistical entropy of the state. Indeed, according to the RT proposal we should take m_A to be the minimal surface in the bulk that is homologous to the boundary; this will generally be the horizon, if there is one, giving agreement with the Bekenstein-Hawking entropy. If there is no horizon, then the boundary is homologically trivial in the bulk (i.e. the topological boundary of the bulk is precisely the boundary where the field theory lives); hence the minimal surface is the empty set, giving $S_A = 0$.[10] (Again, this is the order $G_{\rm N}^{-1}$ entropy—the RT formula does not capture the entropy due, for example, to a gas of gravitons, which is of order $G_{\rm N}^0$.) Furthermore, when the total entropy is zero (or of order $G_{\rm N}^0$), then if we instead take A to be a subset of the boundary, we expect from (2.4) that $S_A = S_{A^c}$. Indeed, in this case the entire boundary is homologically trivial in the bulk, so A and A^c are homologous, implying $m_A = m_{A^c}$.

Another important check on the RT proposal is that it satisfies the strong subadditivity (SSA) property (2.3) for any regions C and D, as can be shown by a simple geometrical argument [23]. (Interestingly, the proof of SSA based on the RT formula is far simpler than the general proof.) Since, as mentioned in Subsection II A, SSA characterizes the EE, this is quite strong evidence in favor of the RT formula. (However, it is not sufficient to prove its correctness, since it only shows that (2.3) is satisfied for subsystems corresponding to geometrical regions, whereas for the characterization proof one needs it to hold for *all* subsystems.) The proof extends trivially to the inclusion of higher-derivative corrections, as long as they are extensive.

As a final check, let us see how the RT formula reproduces the EE (2.21) of a single interval $[u, v]$, in the vacuum of a two-dimensional CFT. The vacuum is described holographically by AdS_3, whose metric on a constant-time slice is

$$ds^2 = \frac{\ell_{\rm AdS}^2}{z^2}(dz^2 + dy^2); \tag{3.2}$$

here y is the coordinate along the boundary and z is the radial coordinate, with the boundary being at $z = 0$. We employ a simple UV cutoff in which we shift the boundary to $z = \epsilon$. The minimal-surface $m_{[u,v]}$ is a geodesic connecting the points on the boundary $(y, z) = (u, \epsilon), (v, \epsilon)$, which is an arc of a circle (almost a semicircle) with center $((u + v)/2, 0)$. Applying (3.1) and using the standard holographic relation $\ell_{\rm AdS}/G_{\rm N} = 2c/3$, one finds [4,5]

$$S_{[u,v]} = \frac{\ell_{\rm AdS}}{2G_{\rm N}} \ln\left(\frac{v-u}{\epsilon}\right) = \frac{c}{3} \ln\left(\frac{v-u}{\epsilon}\right), \tag{3.3}$$

matching (2.21). (In this scheme, the finite part c_1 vanishes.) In higher dimensional CFTs, although one does not have exact formulas for the EEs even of simple regions, the leading UV divergence is known and matches that predicted by the RT formula [4,5].

For a time it was believed that the RT formula should only apply to connected regions. (See, for example, the paper [24].) The reason was that, when applied to the union of two intervals (a case that will be considered in detail in the next section), it disagreed with a calculation by Calabrese and Cardy [25] which (like the formula (2.21) for a single interval) was supposed to be valid in any two-dimensional CFT. However, those same authors have since shown that their original calculation was incorrect. At present, there is no reason to believe that the RT formula, if it is valid at all, would not apply equally well to connected and to disconnected regions. For example, all the checks discussed above apply to both cases (including the last one, which can be considered a computation of the EE of the disconnected region $(-\infty, u] \cup [v, \infty)$).

When applied to a disconnected region, the RT formula makes a fascinating prediction for the mutual information (MI) between its components, similar to the phase transition for disconnected Wilson loops found by Gross and Ooguri [8]. For simplicity, let us consider two disjoint and mutually disconnected regions A, B. Each has a corresponding minimal-surface m_A, m_B and region r_A, r_B. (We assume the generic situation that r_A, r_B are disjoint and mutually disconnected.) When we consider the region $A \cup B$, the disconnected surface $m_A \cup m_B$ is topologically allowed and locally minimal. Assuming that the full bulk spacetime is itself connected, surfaces will also exist that connect A and B. However, if the separation between A and B is sufficiently large compared to their sizes (and any

[10] Since we are considering static spacetimes, any black holes in the spacetime should be eternal, so the maximally extended spacetime may include other, topologically disconnected boundaries. Consider, for example, the maximally extended spacetime of the AdS-Schwarzschild black hole, whose boundary has two connected components. It is believed [22] that the field theory defined on *both* boundaries represents the thermofield double of the field theory defined on only one boundary. In this picture, the black-hole spacetime, which represents a thermal and therefore mixed state in the original field theory, represents a pure state in the thermofield double. This result is faithfully reproduced by the RT prescription; the full boundary (including both components) is homologically trivial in the bulk, giving $S_A = 0$.

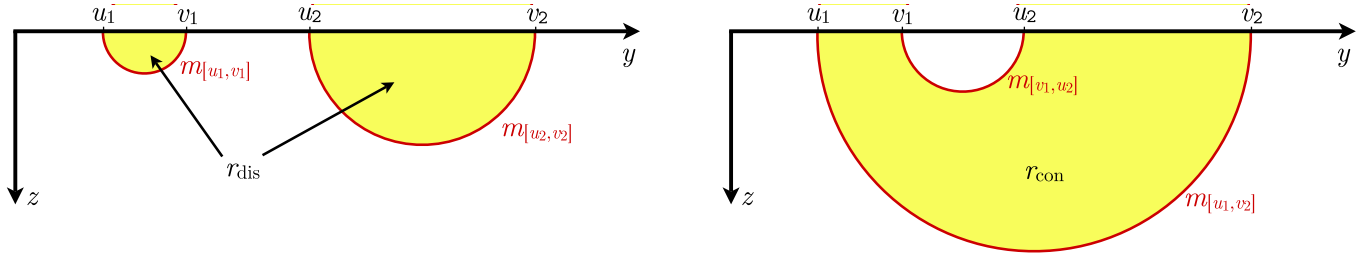

FIG. 1 (color online). The two locally minimal surfaces for the boundary region $[u_1, v_1] \cup [u_2, v_2]$. The global minimum is the one on the left is when $x < 1/2$, and the one on right when $x > 1/2$, where x is the cross-ratio defined in (4.5).

other scales defined in the theory or state), then $m_A \cup m_B$ will necessarily be the globally minimal surface. Then we have $S_{A\cup B} = S_A + S_B$, so the MI $I_{A,B}$ vanishes. More precisely, $I_{A,B}$ is of order $G_{\rm N}^0$, rather than $G_{\rm N}^{-1}$.[11] This implies that, from a quantum information point of view, the two regions are approximately decoupled from each other. (See, for example, the bound (2.6) on correlators between A and B. Note however that this bound does not give us new information about correlators of local operators, for which the left-hand side is always of order $G_{\rm N}^0$.) This prediction is striking in its generality, applying as it does to all holographic theories. (In Sec. IV we will find evidence that in fact it applies even more generally, to a large class of large N field theories.) If we then imagine bringing A and B closer to each other, then it may happen that, at some critical separation, the minimal surface will switch from $m_A \cup m_B$ to one that connects ∂A and ∂B (see, for example, Fig. 1). In this case, the MI will (in the thermodynamic/classical limit $G_{\rm N} \to 0$) undergo a first-order phase transition; it will become nonzero, with a continuous value but discontinuous first derivative as a function of the separation between A and B. Section IV will be devoted to a detailed study of these phenomena in the simplest example, namely, two intervals in the vacuum of a two-dimensional CFT.

B. Fursaev's ERE calculation

In the paper [7], Fursaev gave a derivation, based on the replica trick, of the RT formula. In this subsection, we will briefly summarize his argument, and then point out a flaw that results in an incorrect value for the entanglement Rényi entropy (ERE).

In our sketch of Fursaev's argument, for simplicity we will take the bulk action to be pure Einstein gravity; matter fields and higher-derivative (e.g. α') corrections are straightforwardly incorporated, as he discusses. We will also assume that the ultraviolet divergence in the field theory is cut off in some manner whose details will not concern us. Fursaev's starting point is (2.14), where Z_n is the partition function on the n-sheeted Euclidean spacetime E_n. Recall that, if A is the spatial region whose EE we are computing, then the sheets of E_n are connected by a branch cut along A on a constant-time slice. In a holographic theory, this partition function is given by the gravitational path integral over Euclidean geometries whose conformal boundary is E_n. In the classical limit, this path integral goes over to its saddle point approximation $e^{-S_{\rm min}}$, where $S_{\rm min}$ is the minimal value of the Euclidean Einstein-Hilbert action among extrema obeying the boundary conditions. Fursaev constructs a set of geometries with boundary E_n, then minimizes the Euclidean action within that set. He takes as given the bulk Euclidean spacetime $\tilde{E}$ representing the original state of the system; its boundary is E and its Euclidean action is $-\ln Z_1$. He takes n copies of $\tilde{E}$ and connects them along a branch cut r_A, which is a spatial region in $\tilde{E}$ lying in the same constant-time slice as A. In order for this n-sheeted bulk geometry to have boundary E_n, the part of the boundary of r_A that lies in E must coincide with A (i.e. $\partial r_A \cap E = A$); apart from this condition, the choice of r_A is at this point arbitrary. The branch "point" is m_A, the rest of the boundary of r_A ($m_A = \partial r_A \setminus A$ and $\partial r_A = A \cup m_A$). He now evaluates the Euclidean Einstein-Hilbert action for this geometry. There are two contributions. First, the geometry is made up of n copies of $\tilde{E}$, so there is a contribution $-n \ln Z_1$, which is independent of the choice of r_A. In addition, the Ricci scalar has a delta function along the branch "point" m_A, which is codimension 2 and hosts a conical singularity with excess angle $2\pi(n-1)$. It therefore contributes a term $(n-1)\text{area}(m_A)/(4G_{\rm N})$ to the action. Minimizing this action over all possible choices of r_A, he obtains the minimal-surface m_A, and (from (2.14)) the ERE

$$S_A^{(n)} = \frac{\text{area}(m_A)}{4G_{\rm N}}. \tag{3.4}$$

Since there is no n dependence, the analytic continuation is particularly simple: $S_A^{(\alpha)} = \text{area}(m_A)/4G_{\rm N}$. Finally, setting $\alpha = 1$, he obtains the RT formula (3.1).

The problem with this derivation is that the action has been extremized only with respect to a subset of the degrees of freedom in the metric, namely, the choice

[11] A closely related phenomenon, in which the A and B are held fixed but the bulk spacetime is deformed, was discussed in the paper [6].

of r_A. The resulting field configuration is therefore not guaranteed to be a true saddle point, and in fact it does not solve the Einstein equation: the Einstein tensor has a delta function supported on m_A due to the conical singularity, with no corresponding source.

We can confirm that the ERE (3.4) is incorrect by comparing it to the exact result in a case where the latter is known. For example, when A is a single interval in the vacuum of a two-dimensional CFT, the exact ERE (2.16) depends on n (by the factor $1 + 1/n$), whereas the Fursaev result (3.4) is independent of n. What is the true saddle point in this case? The Euclidean space E is a plane, and the corresponding bulk geometry $\tilde{E}$ is hyperbolic 3-space H^3 (a.k.a. Euclidean AdS_3). The saddle point corresponding to the n-sheeted cover E_n is also H^3. The easiest way to see this is to add a point at infinity to E to make it a sphere; then its n-sheeted cover E_n is also a sphere, so the corresponding bulk geometry is H^3. This geometry is smooth, in contrast to Fursaev's, which is n copies of H^3 glued together in such a way as to create a conical singularity along the geodesic connecting the endpoints of A. Given that the bulk geometry is H^3 for all n, why does its action depend on n? The bulk action is divergent due to the infinite volume near the boundary; while the full bulk geometry is H^3 for any n, the cutoff geometry depends on n. (The full geometry depends only on the Weyl class of the boundary metric, which is the same for all n, since the sphere admits a unique Weyl class. On the other hand, the cutoff geometry is sensitive to the actual boundary metric. This is the holographic manifestation of the Weyl anomaly [26].) The n dependence can most easily be calculated by performing a Weyl transformation to put the metric on E_n into a standard form and taking into the account the resulting change in the partition function due to the Liouville action, as described at the end of Subsection II D, or equivalently by the holographic renormalization procedure [26].

It is worth noting that the true H^3 saddle point with boundary E_n can be obtained *topologically* by gluing n copies of $\tilde{E} = H^3$ together in precisely the manner described by Fursaev. This will continue to be the case in the more complicated examples we will study in the next section, suggesting that, while it carries the wrong metric, Fursaev's construction may be topologically correct in general. This would explain why the topological condition on the minimal-surface m_A that he suggested—that m_A should be homologous to A—appears to be correct.

Finally, it is intriguing that, while Fursaev's value (3.4) for the ERE is incorrect for $n > 1$, it somehow manages to give the right answer for the EE ($n = 1$), assuming that the RT conjecture holds. We can only speculate that, if there is some sense in which the spacetimes E_n and their bulk duals can be defined for noninteger values of n, then his construction may be correct ''at linear order'' in a neighborhood of $n = 1$

IV. MUTUAL RÉNYI INFORMATION BETWEEN TWO INTERVALS

As we saw in Subsection II D, the entanglement entropy for a single interval in the vacuum of a two-dimensional CFT depends only on the theory's central charge. The fact that the Ryu-Takayanagi formula correctly reproduces this entropy, as shown in Subsection III A, is an important check on the proposal, but does not give us any new information. The next simplest configuration we can consider in such a theory consists of two disjoint intervals. As suggested by the fact that the Rényi entropies (2.15) depend in this case on four-point rather than two-point functions of twist operators, we would expect the EE to depend on the full operator content of the theory, rather than simply its central charge. As we will see in this section, the RT formula can give us significant new physical information in this case. The new predictions in turn give us the opportunity to subject the formula to new and highly nontrivial quantitative tests.

We begin by reviewing the necessary formulas and setting up the basic properties of the ERE for two intervals.

A. General properties

We consider two separated intervals $[u_1, v_1]$, $[u_2, v_2]$ ($u_1 < v_1 < u_2 < v_2$) in the vacuum of a conformal field theory $\mathcal{C}$ with central charge c. As discussed in Subsection II A, it is convenient to consider the mutual Rényi information (MRI) between the two intervals,

$$I^{(\alpha)}_{[u_1,v_1],[u_2,v_2]} = S^{(\alpha)}_{[u_1,v_1]} + S^{(\alpha)}_{[u_2,v_2]} - S^{(\alpha)}_{[u_1,v_1]\cup[u_2,v_2]}, \tag{4.1}$$

which measures the extent to which the degrees of freedom of the two intervals are entangled with each other (including both classical correlations and quantum entanglement).

We first consider the integer case $\alpha = n > 1$. Using (2.15), the MRI is given in terms of a finite ratio of four-point and two-point functions in the symmetric product theory $\mathcal{C}^n/S_n$:

$$I^{(n)}_{[u_1,v_1],[u_2,v_2]} = \frac{1}{n-1} \times \ln\left(\frac{\langle \sigma^\epsilon_1(u_1)\sigma^\epsilon_{-1}(v_1)\sigma^\epsilon_1(u_2)\sigma^\epsilon_{-1}(v_2)\rangle}{\langle \sigma^\epsilon_1(u_1)\sigma^\epsilon_{-1}(v_1)\rangle\langle \sigma^\epsilon_1(u_2)\sigma^\epsilon_{-1}(v_2)\rangle}\right) \tag{4.2}$$

$$= \frac{1}{n-1}\ln\left(\frac{\langle \sigma_1(u_1)\sigma_{-1}(v_1)\sigma_1(u_2)\sigma_{-1}(v_2)\rangle}{\langle \sigma_1(u_1)\sigma_{-1}(v_1)\rangle\langle \sigma_1(u_2)\sigma_{-1}(v_2)\rangle}\right), \tag{4.3}$$

where we have defined the renormalized twist operators:

$$\sigma_{\pm 1} \equiv \frac{\sigma^\epsilon_{\pm 1}}{\langle \sigma^\epsilon_1(0)\sigma^\epsilon_{-1}(1)\rangle^{1/2}}. \tag{4.4}$$

This is an example of the UV divergences in the EREs, which occur at the endpoints of the intervals, cancelling in the MRI, as discussed at the end of Subsection II A.

Since the twist operators are primaries, the transformation law for the four- and two-point function implies that $I_{[u_1,v_1],[u_2,v_2]}$ is conformally invariant, and therefore depends only on the cross-ratio

$$x \equiv \frac{(v_1 - u_1)(v_2 - u_2)}{(u_2 - u_1)(v_2 - v_1)}, \tag{4.5}$$

which lies in the interval $0 < x < 1$. By a conformal transformation, the four points u_1, v_1, u_2, v_2 can be brought to 0, x, 1, ∞, respectively, so we have

$$I^{(n)}_{[u_1,v_1],[u_2,v_2]} = I^{(n)}(x) \equiv I^{(n)}_{[0,x],[1,\infty]} = \frac{1}{n-1} \ln(x^{2d_\sigma} \langle \sigma_1(0)\sigma_{-1}(x)\sigma_1(1)\sigma'_{-1}(\infty)\rangle), \tag{4.6}$$

where $\sigma'_{-1}(\infty) \equiv \lim_{z\to\infty} z^{2d_\sigma}\sigma_{-1}(z)$. (The scaling dimensions d_σ of the twist operators are given by (2.17).) Notice that, like the UV divergence, the IR divergence in the ERE cancels in the MRI. The four-point function, and therefore $I^{(n)}(x)$, is an analytic function of x in the interval $0 < x < 1$.

It is useful to note that the four-point function in (4.6) can be expanded as a power series in x, where the powers are the dimensions d_m of operators $\mathcal{A}'_m$ in the orbifold theory,[12] and the coefficients are given in terms of OPE coefficients:

$$I^{(n)}(x) = \frac{1}{n-1} \ln\left(\sum_m c^{\sigma_1}{}_{\sigma_1 m} c^m{}_{\sigma_1\sigma_{-1}} x^{d_m}\right). \tag{4.7}$$

Note that only untwisted operators contribute to the sum. Assuming we are dealing with a unitary theory, the operator with lowest scaling dimension is the unit operator, for which (by the normalization of the twist operators) the OPE coefficients are 1. Hence $I^{(n)}(x)$ goes to 0 as $x \to 0$, as we would expect on physical grounds. For example, if we fix the sizes of the intervals and take their separation to infinity, we would expect all correlations between them to go to zero. We will study the higher-order terms in the expansion (4.7) in Subsections IV D and IV Fand in the Appendix.

A final important property of the MRI, implied by the invariance of the four-point function in (4.6) under $x \to 1 - x$, is

$$I^{(n)}(1-x) = I^{(n)}(x) + \frac{c}{6}\left(1 + \frac{1}{n}\right)\ln\frac{1-x}{x}. \tag{4.8}$$

At the level of the definition (4.1) of the MRI, this relation is due to the fact that, in a pure state (in this case, the vacuum), $S^{(\alpha)}_A = S^{(\alpha)}_{A^c}$, so $S^{(\alpha)}_{[0,x]\cup[1,\infty]} = S^{(\alpha)}_{[-\infty,0]\cup[x,1]} = S^{(\alpha)}_{[0,1-x]\cup[1,\infty]}$.

We have listed five general properties that the MRI satisfies for integer $\alpha > 1$, but for the reasons given we either know or expect each to hold for general values of α:

(1) UV finiteness, and IR finiteness when one of the intervals is semi-infinite;

(2) conformal invariance, implying

$$I^{(\alpha)}_{[u_1,v_1],[u_2,v_2]} = I^{(\alpha)}(x) \equiv I^{(\alpha)}_{[0,x],[1,\infty]} \tag{4.9}$$

(where $0 < x < 1$);

(3)

$$\lim_{x\to 0} I^{(\alpha)}(x) = 0; \tag{4.10}$$

(4) for all x,

$$I^{(\alpha)}(1-x) = I^{(\alpha)}(x) + \frac{c}{6}\left(1 + \frac{1}{\alpha}\right)\ln\frac{1-x}{x}; \tag{4.11}$$

(5) analyticity of $I^{(\alpha)}(x)$ as a function of x.

So far we have not assumed anything about the theory $\mathcal{C}$ (other than unitary and compactness). In the rest of this section, we will study the function $I^{(\alpha)}(x)$ in holographic CFTs, as well as certain other theories with large central charge.

B. Prediction from Ryu-Takayanagi formula

As in the holographic derivation of the EE for a single interval, reviewed in Subsection III A, we use the fact that the holographic dual of the vacuum is AdS_3, with $\ell_{AdS}/G_N = 2c/3$, and we cut off integrals near the boundary at radial coordinate value $z = \epsilon$.

The RT formula is straightforward to apply to the union of two intervals $[u_1, v_1]\cup[u_2, v_2]$. There are two locally minimal surfaces in the bulk that are homologous to this boundary region, as shown in Fig. 1. The first is the union of the minimal surfaces for the two intervals separately, $m_{dis} = m_{[u_1,v_1]} \cup m_{[u_2,v_2]}$ (similarly for the corresponding bulk region $r_{dis} = r_{[u_1,v_1]} \cup r_{[u_2,v_2]}$). This has "area" (i.e. length)

$$\text{area}\,(m_{dis}) = \text{area}(m_{[u_1,v_1]}) + \text{area}(m_{[u_2,v_2]}) \tag{4.12}$$

$$= 2\ell_{AdS}\ln\left(\frac{v_1 - u_1}{\epsilon}\right) + 2\ell_{AdS}\ln\left(\frac{v_2 - u_2}{\epsilon}\right) \tag{4.13}$$

(see (3.3)). The other locally minimal surface connects u_1 to v_2 and u_2 to v_1: $m_{con} = m_{[u_1,v_2]} \cup m_{[v_1,u_2]}$. (The

[12] Throughout the paper, we use primes on (untwisted) operators of $\mathcal{C}^n/S_n$, to distinguish them from operators of $\mathcal{C}$.

corresponding bulk region is a semiannulus connecting the two intervals: $r_{\rm con} = r_{[u_1,v_2]} \setminus r_{[v_1,u_2]}$.) Its area is

$$\text{area}\,(m_{\rm con}) = 2\ell_{\rm AdS} \ln\left(\frac{v_2 - u_1}{\epsilon}\right) + 2\ell_{\rm AdS} \ln\left(\frac{u_2 - v_1}{\epsilon}\right). \tag{4.14}$$

It is easy to see that $m_{\rm dis}$ is the globally minimal surface when $x < 1/2$, and $m_{\rm con}$ otherwise (see (4.5)), so

$$\begin{aligned} S_{[u_1,v_1]\cup[u_2,v_2]} &= \frac{1}{4G_{\rm N}} \min(\text{area}(m_{\rm dis}), \text{area}(m_{\rm con})) \\ &= \frac{c}{3} \times \begin{cases} \ln((v_1 - u_1)(v_2 - u_2)/\epsilon^2), & x \le 1/2 \\ \ln((v_2 - u_1)(u_2 - v_1)/\epsilon^2), & x \ge 1/2 \end{cases}. \end{aligned} \tag{4.15}$$

Combining (4.15) with (3.3), we obtain the following mutual information:

$$\begin{aligned} I_{[u_1,v_1],[u_2,v_2]} &= I^{(1)}(x) \\ &= \begin{cases} 0, & x \le 1/2 \\ (c/3)\ln(x/(1-x)), & x \ge 1/2 \end{cases}. \end{aligned} \tag{4.16}$$

Of the five properties of the MI listed at the end of the previous subsection, this formula obeys the first four. It does not obey the last—analyticity—as it has a discontinuous first derivative at $x = 1/2$ and vanishes for $x \le 1/2$. These two features were anticipated in the discussion at the end of Subsection III A. The discontinuity in the first derivative occurs because the global minimum switches between the local minima as we vary x, and is reminiscent of phase transitions due to competing saddle points of the Euclidean action, such as the Hawking-Page transition. As in that case, the transition is presumably sharp only in the classical limit in the bulk, which corresponds to the thermodynamic ($c \to \infty$) limit of the CFT, and gets smoothed out by finite c effects. Similarly, the vanishing of the MI for $x \le 1/2$ is presumably true only at order c; if the MI vanished exactly for $x \le 1/2$, then the reduced density matrix for the two intervals would factorize, $\rho_{[u_1,v_1]\cup[u_2,v_2]} = \rho_{[u_1,v_1]} \otimes \rho_{[u_2,v_2]}$, implying that the two intervals are completely decoupled from each other, a rather unphysical situation (in particular, it would violate the inequality (2.6), as we will discuss below). Thus we should expect both perturbative and nonperturbative corrections to (4.16) in $G_{\rm N} \sim c^{-1}$, with the first perturbative correction at order c^0. Nonetheless, since the MI is apparently parametrically small for $x \le 1/2$—smaller than the EE for either interval separately or for their union, and smaller than the MI for $x > 1/2$—it appears that the density matrix factorizes approximately.

Unlike quantum corrections, we do not expect higher-derivative (e.g. α') corrections to the classical bulk action to change the result (4.16), for the following reason. As discussed in Subsection III A, such corrections are believed to correct the area functional appearing in the RT formula without changing the basic prescription of minimizing over topologically allowed surfaces. The symmetries of AdS_3 guarantee that the minimal surfaces shown in Fig. 1 remain uncorrected; furthermore, the corrected "area" of each curve is unchanged when written as a function of c, since we know that the EE is always given by (2.21). In fact, this argument applies for any bulk gravitational action, not just Einstein-Hilbert with small higher-derivative corrections.

It is interesting to consider the result (4.16) in the context of the general inequality (2.6). Local operators in holographic field theories are dual to bulk fields. Smearing such operators over the two intervals will give a result of order c^0—no matter how the operators are normalized, the numerator and denominator will be of the same order in c. This is consistent with the result (4.16) (and shows that it must be corrected at order c^0). However, for $x > 1/2$, the right-hand side of (2.6) is of order c. It would be quite interesting to find examples of operators $\mathcal{A}_{[u_1,v_1]}$, $\mathcal{A}_{[u_2,v_2]}$ that saturate (2.6), or at least for which the left-hand side is of order c. It is then interesting to ask what happens to these operators for $x < 1/2$. Do they simply cease to exist? Or does their two-point function become parametrically smaller? Finding explicit examples of operators with such behavior would not just constitute strong evidence in favor of the RT formula, but also clarify in what sense the density matrix factorizes.

C. Universality in the large c limit?

Given any family of CFTs $\mathcal{C}$ that admit a large c limit, such as holographic ones, we can consider the expansion of the MRI in powers of c^{-1}. Since the number of degrees of freedom is of order c, the leading term will be at most of that order, so we have

$$I^{(\alpha)}(x) = I_1^{(\alpha)}(x)c + I_0^{(\alpha)}(x) + O(c^{-1}). \tag{4.17}$$

In particular, we focus our attention on the leading function $I_1^{(\alpha)}(x)$. In the previous subsection, we used the RT formula to compute, for holographic theories,

$$I_1^{(1)}(x) = \begin{cases} 0, & x \le 1/2 \\ (1/3)\ln(x/(1-x)), & x \ge 1/2 \end{cases}, \tag{4.18}$$

and argued that this result should hold no matter what the bulk gravitational theory is. As discussed, (4.18) has two striking qualitative features, namely, its discontinuous first derivative at $x = 1/2$ and the fact that it vanishes for $x \le 1/2$. In the rest of this section, we will study $I_1^{(\alpha)}(x)$

using the replica trick, and find independent evidence for both phenomena. In the next subsection, we will compute $I_1^{(2)}(x)$ in holographic theories, and show that the result is independent of the details of the bulk theory (e.g. the presence of higher-curvature corrections), and applies also to symmetric product theories $\mathcal{C} = \mathcal{C}_0^N/S_N$, even though their large c limit is not described by classical gravity. Like (4.18), the result will have a phase transition at $x = 1/2$. In Subsection IV E we will argue that this phase transition occurs also in $I_1^{(n)}(x)$ for $n > 2$, at least in holographic theories. Then in the last subsection we will use CFT techniques to study the expansion of $I_1^{(\alpha)}(x)$ in powers of x for general α, and find evidence that every coefficient in this expansion goes to 0 in the limit $\alpha \to 1$. That analysis will assume very little about the CFT $\mathcal{C}$, essentially just that the number of operators below any given dimension stays finite as c goes to infinity, a condition that holds for both holographic and symmetric product theories (but not, for example, in the power theory $\mathcal{C}_0^N$ without the orbifold).

These results not only give strong quantitative support to the RT formula, but point to a broader picture, namely, that a large class of large c CFTs—including holographic and symmetric product theories—share the same leading MRI $I_1^{(\alpha)}(x)$, as a function of both α and x. Although we do not know the explicit form of this function except for $\alpha = 1, 2$, we can deduce that it is analytic in x except at $x = 1/2$, where it has a discontinuous first derivative, and satisfies the following properties:

$$\lim_{x\to 0} I_1^{(\alpha)}(x) = 0, \tag{4.19}$$

$$I_1^{(\alpha)}(1-x) = I_1^{(\alpha)}(x) + \frac{1}{6}\left(1 + \frac{1}{\alpha}\right)\ln\frac{1-x}{x}. \tag{4.20}$$

Based on these considerations, it appears that in the range $0 < x \leq 1/2$ the MRI is parametrically larger for $\alpha \neq 1$ (where it is of order c) than for $\alpha = 1$ (where it is of order 1). This is similar to what we found in the perturbative calculation of Subsection II B. It would be interesting to find a simple toy model of a system with N degrees of freedom, in which the MRI between two subsystems is of order N, but the MI is only of order 1.

D. MRI for $n = 2$

In this subsection we will begin by expressing the mutual Rényi information $I^{(2)}(x)$ (sometimes called the *mutual collision information*) in a general CFT in terms of its torus partition function. Using this expression, we will calculate the order c part $I_1^{(2)}(x)$ in a general holographic CFT, finding that—like the RT prediction (4.18) for $I_1^{(1)}(x)$—it is analytic except at $x = 1/2$, where it has a discontinuous first derivative. We will then show that $I_1^{(2)}(x)$ is precisely the same function in large N symmetric product theories, supporting the idea of universality (i.e. theory independence in the large c limit) proposed in the previous subsection.

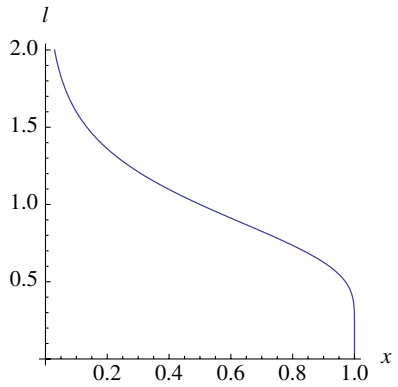

FIG. 2 (color online). Modular parameter $\tau = il$ for the two-sheeted Riemann surface with branch points at 0, x, 1, ∞. The relation between x and l is given by Eq. (4.22).

We begin by applying the formula (4.6) for $n = 2$. In the $\mathcal{C}^2/Z_2$ orbifold theory, there is a unique twist operator $\sigma \equiv \sigma_1 = \sigma_{-1}$. Lunin and Mathur [17] showed that its four-point function is given by[13]

$$\langle\sigma(0)\sigma(x)\sigma(1)\sigma'(\infty)\rangle = (2^8 x(1-x))^{-c/12} Z_{il}, \tag{4.21}$$

where Z_{il} is the partition function for $\mathcal{C}$ on a flat rectangular torus[14] with modular parameter $\tau = il$; x and l are related by

$$x = \frac{\theta_2^4(il)}{\theta_3^4(il)}. \tag{4.22}$$

As x goes from 0 to 1, l goes from ∞ to 0, with $x = 1/2$ corresponding to $l = 1$ (see Fig. 2). Since

$$1 - x = \frac{\theta_4^4(il)}{\theta_3^4(il)} = \frac{\theta_2^4(i/l)}{\theta_3^4(i/l)}, \tag{4.23}$$

the invariance of the four-point function (4.21), and hence of the ERE, can be traced to the modular invariance of the

[13]In terms of Lunin and Mathur's variables, $x = 1/w$ and $il = \tau = -1/\tau_{\text{Lunin-Mathur}}$. The four-point function of twist fields was computed in [27] in the case where the underlying CFT is a free scalar field.

[14]The fermion sign flips explained in footnote 6 imply that fermionic fields should have antiperiodic (NS) boundary conditions on both cycles of the torus. The reason is that, on the double-sheeted plane, in going around either cycle one encounters two sign flips, so no overall flip. The coordinate transformation to the torus introduces a flip, just as when passing from the plane to the cylinder. Hence the partition function is not invariant under the full modular group, but it is invariant under $\tau \to -1/\tau$.

torus partition function, $Z_{il} = Z_{i/l}$. The reason for the appearance of the torus partition function of $\mathcal{C}$ is that the four-point function of twist operators is the (renormalized) zero-point function on the two-sheeted Riemann surface E_2 with a branch cut connecting 0 to x and another one connecting 1 to ∞, which is a torus with complex structure $\tau = il$. The Weyl transformation that flattens it, when plugged into the Liouville action, leads to the prefactor $(2^8 x(1-x))^{-c/12}$.

Plugging (4.21) into (4.6), we obtain

$$I^{(2)}(x) = \ln Z_{il} - \frac{c}{12}\ln\left(\frac{2^8(1-x)}{x^2}\right) = \ln Z_{il} - \frac{c}{3}\ln\left(\frac{4\theta_4(il)\theta_3(il)}{\theta_2^2(il)}\right). \tag{4.24}$$

The first term in (4.24) is ($-l$ times) the free energy of $\mathcal{C}$ on a circle of unit circumference at temperature l^{-1}. (Note however that the basic cycles of this torus, which we interpret as space and Euclidean-time directions when we speak of the free energy, are not the same as the space and time directions of the Euclidean plane E where the theory was originally defined, whose double cover is E_2. Rather, the spatial circle of the torus encircles the points 0 and x, staying on one sheet, while its Euclidean-time circle encircles x and 1, crossing each branch cut once.)

Let us consider the expansion of (4.24) for small values of x, where $x \approx 16e^{-\pi l}$.[15] The behavior of the torus partition function is universal in this limit (for compact unitary CFTs), $\ln Z_{il} \approx 2\pi c l/12$, which precisely cancels the leading behavior of the second term in (4.24), giving a vanishing MRI as expected (Eq. (4.10)). The leading x dependence depends on the gap in the operator spectrum of $\mathcal{C}$. If the lowest nonunit operator $\hat{\mathcal{A}}$ has dimension $\hat{d}$ and multiplicity $\hat{m}$ (where fermionic operators are counted negatively), then

$$\ln Z_{il} = \frac{2\pi c l}{12} + \hat{m}e^{-2\pi\hat{d}l} + \cdots, \tag{4.29}$$

so

$$I^{(2)}(x) \sim \left\{\begin{array}{ll} 2ce^{-2\pi l} \sim 2^{-7}cx^2, & \hat{d} > 1 \\ (2c+\hat{m})e^{-2\pi l} \sim 2^{-8}(2c+\hat{m})x^2, & \hat{d} = 1 \\ \hat{m}e^{-2\pi\hat{d}l} \sim \hat{m}(x/16)^{2\hat{d}}, & \hat{d} < 1 \end{array}\right\} \quad (x \ll 1). \tag{4.30}$$

This term can be matched onto the leading term in the expansion in intermediate states (4.7), by noting that the lowest-dimension operator of $\mathcal{C}^2/Z_2$ appearing in the $\sigma\sigma$ OPE, other than the unit operator, is $\hat{\mathcal{A}} \otimes \hat{\mathcal{A}}$ if $\hat{d} \leq 1$, and the stress tensor if $\hat{d} \geq 1$. The OPE coefficients are computed and matched to (4.30) in Appendix A 1 (see also the discussion around (4.39)), where we also consider more generally the matching between the Lunin-Mathur formula for the four-point function (4.21) and its expansion in intermediate states.

A simple example of the application of (4.24) is to a free scalar compactified on a circle of radius R. The torus partition function is

$$Z_{il} = \frac{\theta_3(il/R^2)\theta_3(ilR^2)}{\eta^2(il)}, \tag{4.31}$$

so [3,28]

$$I^{(2)}(x) = \ln\left(\frac{\theta_3(il/R^2)\theta_3(ilR^2)\theta_2^{2/3}(il)}{2^{2/3}\eta^2(il)\theta_4^{1/3}(il)\theta_3^{1/3}(il)}\right). \tag{4.32}$$

This is plotted against x for several values of R in Fig. 3. The small x behavior is as predicted by (4.30), with the

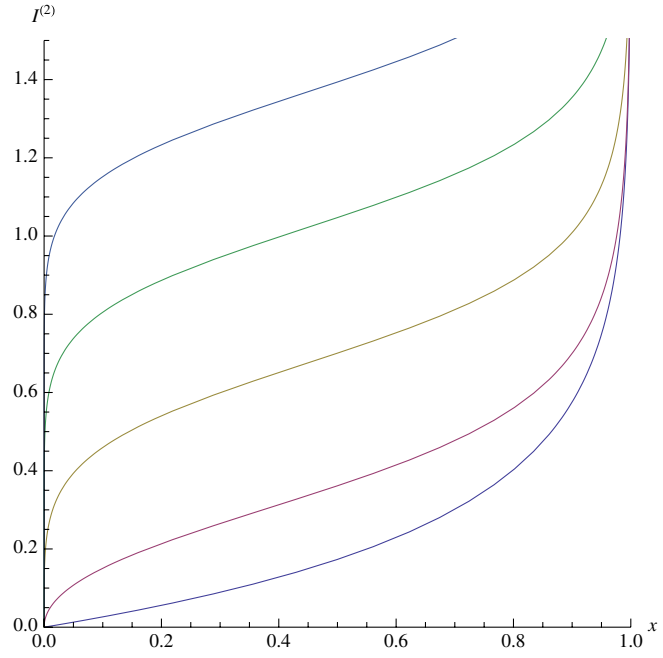

FIG. 3 (color online). Mutual Rényi information of intervals $[0, x]$ and $[1, \infty]$ for a free scalar field of radius R, with (from bottom to top) $R^2 = 1, 2, 4, 8, 16$.

[15] Defining $q = e^{2\pi i\tau} = e^{-2\pi l}$, the expansions of the theta functions for small q are as follows:

$$\theta_2(\tau) = 2q^{1/8}(1 + q + O(q^3)) \tag{4.25}$$

$$\theta_3(\tau) = 1 + 2q^{1/2} + O(q^2) \tag{4.26}$$

Hence

$$\theta_4(\tau) = 1 - 2q^{1/2} + O(q^2). \tag{4.27}$$

$$\frac{4\theta_4(\tau)\theta_3(\tau)}{\theta_2^2(\tau)} = q^{-1/4}(1 - 6q + O(q^2)). \tag{4.28}$$

lowest nonunit operator having dimension $\hat{d} = 1/(2R^2)$ (for $R^2 \geq 1$) and multiplicity $\hat{m} = 2$ (except for $R = 1$, where the lightest winding and momentum modes are degenerate, so $\hat{m} = 4$) [3,28]. We now turn to holographic CFTs, briefly reviewing Maldacena and Strominger's result for the torus partition function [9]. Expanding the free energy in powers of $c^{-1} \sim G_N$, the leading term is of order c and is given by the Euclidean action of the dominant saddle point. Here the boundary condition is simply that the conformal boundary should be the torus with $\tau = il$; there are no operators inserted in the path integral so no fields other than the metric are sourced. For $l > 1$ ($x < 1/2$) the dominant saddle point is the Euclidean BTZ black hole, which is topologically a solid torus in which the Euclidean-time circle (the circle of length l) is contractible. For $l < 1$ ($x > 1/2$) the dominant saddle point is Euclidean AdS_3 with the Euclidean-time direction periodically identified; the topology is a solid torus in which the spatial circle is contractible.[16] The phase transition between the two saddles, the Hawking-Page transition, is first order, so the free energy, and hence $I_1^{(2)}(x)$, has a discontinuous first derivative. Specifically, the Euclidean actions of the two saddle points yield [9]

$$\ln Z_{il} = \begin{cases} 2\pi c/(12l) + O(c^0), & l < 1 \\ 2\pi c l/12 + O(c^0), & l > 1 \end{cases}, \tag{4.33}$$

so

$$I_1^{(2)}(x) = -\frac{1}{3}\ln\left(\frac{4\theta_4(il)\theta_3(il)}{\theta_2^2(il)}\right) + \begin{cases} 2\pi/(12l), & l < 1 \\ 2\pi l/12, & l > 1 \end{cases}, \tag{4.34}$$

which is plotted in Fig. 4. Note that, although $I_1^{(2)}(x)$ does not vanish in the region $x < 1/2$, it is numerically quite small—smaller than $I^{(2)}(x)$ for the free scalar by two orders of magnitude or more. The expansion for small x is $2^{-7}x^2$; comparing to (4.30), it is as if the lowest-dimension operator of $\mathcal{C}^2/\mathbf{Z}_2$ is the stress tensor. In fact, there are other operators, but since their multiplicity is finite in the limit $c \to \infty$, they do not contribute to $I_1^{(2)}$. We will discuss this expansion in detail in Subsection IV F.

If we compare the bulk saddle point geometries used to derive (4.33) to the ones obtained from Fursaev's construction, we see that they are topologically identical but metrically different. To describe the geometry obtained from Fursaev's construction, we add a Euclidean-time direction, coming out of the page, to each diagram in Fig. 1, and consider the double cover of the resulting three-dimensional geometry, branched over r_{dis} and r_{con}

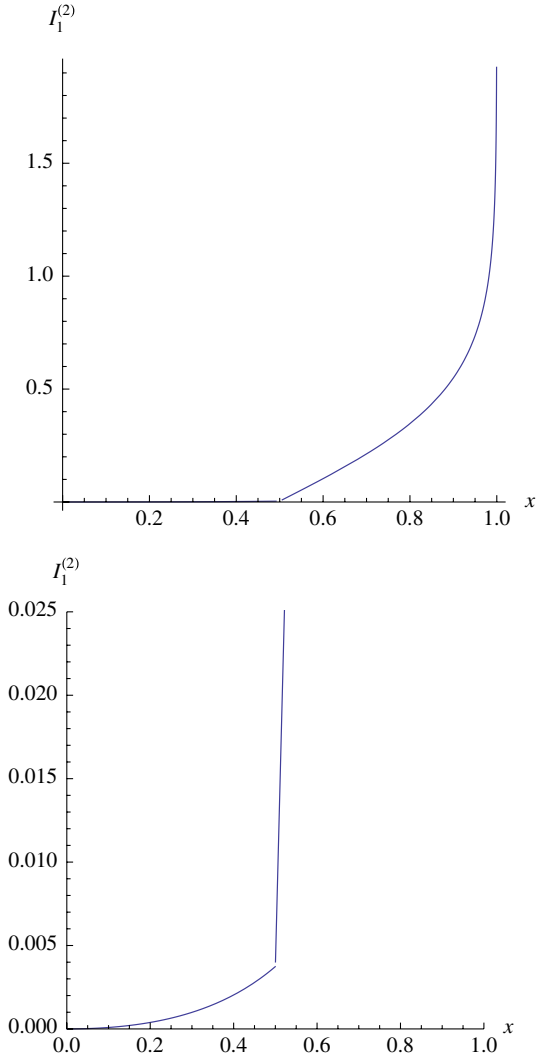

FIG. 4 (color online). Coefficient of c in the mutual Rényi information of intervals $[0, x]$ and $[1, \infty]$ in a general holographic CFT. The two plots differ only in the scale of the vertical axis. In particular, the plot on the right shows that $I_1^{(2)}$ is nonzero (although quite small) for $x < 1/2$.

respectively. On the left-hand diagram, relevant when $x < 1/2$, the cycle on the boundary that encircles u_1 and v_1, staying on one sheet, is contractible through the bulk; this is the spatial circle of the torus. On the right-hand diagram, relevant when $x > 1/2$, the boundary cycle that encircles v_1 and u_2, crossing both branch cuts, is contractible through the bulk; this is the time circle of the torus. Hence in each case the topology is precisely the same as that of the true saddle point geometry. However, their metrics are different; in particular, while the former are singular, the latter are smooth.

The formula (4.33) for the torus partition function, and therefore the formula (4.34) for the MRI, applies not only to holographic CFTs but also to symmetric product theories $\mathcal{C} = \mathcal{C}_0^N/S_N$ at large N, where $\mathcal{C}_0$ is any (compact unitary) CFT [29]. The basic reason is that, for $l < 1$, the

[16]Note that the contractibility of the two cycles of the boundary torus requires antiperiodicity of fermions on both. As explained in footnote 14, these are precisely the boundary conditions we have in this case.

effective temperature for the long strings, which dominate the partition function, is enhanced by a factor of N, so the theory is effectively always in the high-temperature limit; the partition function for $l > 1$ is then given by modular invariance. An example is the supersymmetric $(T^4)^N/S_N$ theory, which is conjectured to be connected in a moduli space to type IIB string theory on $AdS_3 \times S^3 \times T^4$. It would seem reasonable then to guess that the torus partition function is given by (4.33) for all theories on this moduli space. In other words, the MRI appears to enjoy a nonrenormalization theorem.

The phase transition at $x = 1/2$ can be understood in terms of the expansion $Z_{il} = \sum_i e^{-2\pi(d_i + c/12)l}$, where the d_i are the scaling dimensions of the operators of $\mathcal{C}$. In any fixed theory, with finite c, this expansion converges and is analytic for all $x \in [0, 1)$. In the large c limit, the operators of $\mathcal{C}$ can (roughly speaking) be divided into those with scaling dimensions of order 1 (''short strings'') and those with scaling dimensions of order c (''long strings''). Each long-string operator makes a contribution to the sum that is exponentially suppressed in c. However, the number of long-string operators is exponentially large in c, so they may actually dominate the sum. In fact, whether the short strings or the long strings dominate depends on the value of l, and hence of x. In both holographic and large N symmetric product theories, short strings dominate for $x < 1/2$ and long strings for $x > 1/2$. Thus the order c part of the free energy is due entirely to short strings for $x < 1/2$ and to long strings for $x > 1/2$.

E. MRI for $n > 2$

In this subsection we will extend our study of $I^{(n)}(x)$ to larger values of n. Although we will not be able to give explicit formulas, we will argue that all the main qualitative features carry through. In particular, the existence for all $n > 1$ of a discontinuous first derivative in $I_1^{(n)}(x)$ at $x = 1/2$ constitutes significant evidence in favor of the RT formula, which predicts precisely such a phase transition for $n = 1$.

According to Eq. (4.6), the MRI $I^{(n)}(x)$ for general n is given in terms of the four-point function of twist operators in the symmetric product theory $\mathcal{C}^n/S_n$. This four-point function is in turn equal to the (renormalized) zero-point function of $\mathcal{C}$ on the surface E_n, made of n sheets connected by a branch cut running from 0 to x and another from 1 to ∞. This is the Riemann surface for the equation $y^n = z(z-1)/(z-x)$, which has genus $n-1$ and a complex structure that depends on x. As in the genus-1 case studied in the previous subsection, the surface E_n can be taken by a Weyl transformation to a fiduciary metric with the same complex structure, for example, the constant-curvature metric. Hence there is an analogue of (4.6), in which $I^{(n)}$ is written as a sum two terms:

$$I^{(n)}(x) = \ln Z_x^{(n)} + c I^{(n)}_{1,\text{geometric}}(x). \tag{4.35}$$

$Z_x^{(n)}$ is the partition function of $\mathcal{C}$ on the surface carrying the fiduciary metric. The second term is derived from the Liouville action for the Weyl transformation from E_n to the fiduciary metric; aside from the overall coefficient c, it is independent of the particular theory $\mathcal{C}$, giving a universal contribution to $I_1^{(n)}$.

The geometrical term $I^{(n)}_{1,\text{geometric}}$ has not been explicitly computed for $n > 2$. However, assuming that the fiduciary metric is chosen to depend smoothly on x (as does, for example, the constant-curvature metric), the Weyl transformation and hence $I^{(n)}_{1,\text{geometric}}$ will be smooth functions of x.

Meanwhile, the genus $(n-1)$ partition function $Z_x^{(n)}$ is known explicitly for $n > 2$ only in a small number of CFTs. For holographic theories, despite considerable progress (especially in the context of pure gravity theories), explicit formulas are not available, even in the large c limit; see for example [30–34]. Even in the absence of an explicit formula, however, we can argue that the partition function is smooth except at $x = 1/2$, where it has a discontinuous first derivative, just as we saw for $n = 2$ in the last subsection. It is known that phase transitions, analogous to the Hawking-Page transition, occur at fixed points of the mapping-class group (the group of large diffeomorphisms of the Riemann surface). The reason is that different saddle points of the bulk gravitational action are effectively mapped onto each other by the action of the mapping-class group, and therefore at a fixed point they necessarily have the same action.[17] On the surface E_n, there is an element of the mapping-class group that effectively takes x to $1 - x$; for example, the surfaces E_n with $x = 1/3$ and $x = 2/3$ have the same complex structure up to the action of an element of the mapping-class group. That element permutes the cycle that encircles 0 and x with the one that encircles x and 1. It has a unique fixed point, namely $x = 1/2$. Hence we expect a phase transition in (the order c part of) $\ln Z_x^{(n)}$, and therefore in $I_1^{(n)}(x)$, at $x = 1/2$, and only there. This is precisely the property predicted by the RT formula for $I_1^{(1)}(x)$ (see (4.18)). (Furthermore, it seems likely that, as we saw for $n = 2$, the dominant saddle has the same topology as predicted by Fursaev, i.e for $x < 1/2$ the cycle encircling 0 and x is contractible in the bulk, while for $x > 1/2$ the cycle encircling x and 1 is contractible in the bulk.)

In the previous subsection we used the fact that the torus partition function is the same (at leading order in c) for large N symmetric product theories $(\mathcal{C}_0^N/S_N)$ as for holographic ones, to support the claim of universality (theory independence) for $I_1^{(2)}(x)$. It would be interesting to investigate whether the same holds for $Z_x^{(n)}$ for $n > 2$. Our results in the next subsection, which apply for all n, support such universality.

[17] We thank A. Maloney for helpful discussions on this point.

F. Expansion in x

In Subsection IV D we derived an explicit expression for $I_1^{(2)}(x)$ that applied to both holographic and large N symmetric product theories. Unfortunately, as we discussed in the last subsection, the computation of higher-genus partition functions, and therefore $I_1^{(n)}(x)$, remains out of reach technically in such theories. Even if we could find explicit expressions, their analytic continuation to general α, and, in particular, to $\alpha = 1$, may not be feasible.[18] (For example, Calabrese, Cardy, and Tonni were able to compute $I^{(n)}(x)$ for $n > 1$ for a compactified free boson [3]. However, the analytic continuation of the resulting expression, which involves a Riemann-Siegel theta function, is unknown.) In this subsection, therefore, we will take a different approach, and compute $I_1^{(n)}(x)$ order by order in x for general $n > 1$. The coefficient of each power of x will be a simple enough function of n to allow us to analytically continue it straightforwardly.[19] As discussed in the previous subsection, for each $n > 1$, $I_1^{(n)}(x)$ is analytic on $[0, 1/2]$, but not on larger intervals. Assuming that this property continues to hold for $I_1^{(\alpha)}(x)$ for general α, the formulas we derive will be valid on that interval. We are particularly interested in testing two hypotheses concerning the coefficient of each term in the power series expansion of $I_1^{(\alpha)}(x)$: that it is "universal" in the sense of Subsection IV C, i.e. the same for all theories in the class we are considering; and that it goes to zero in the limit $\alpha \to 1$, in agreement with the RT prediction (4.18). We will find significant evidence in favor of both hypotheses.

We are considering theories $\mathcal{C}$ for which the number of operators with any given dimension is finite in the large c limit. This condition applies to holographic and $\mathcal{C}_0^N/S_N$ theories (where $\mathcal{C}_0$ is held fixed as we take $c \sim N \to \infty$), but not, for example, to the theory $\mathcal{C}_0^N$ without the orbifold. We will also assume that the n-point functions of primaries in $\mathcal{C}$ do not diverge in the large c limit; again, this property holds for holographic and $\mathcal{C}_0^N/S_N$ theories.

We begin with the relation (4.7), in the form

$$\exp((n-1)I^{(n)}(x)) = \sum_m c^{\sigma_1}{}_{\sigma_1 m} c^m{}_{\sigma_1\sigma_{-1}} x^{d_m}. \tag{4.36}$$

At this stage we have not taken the large c limit, and the convergence of the OPE dictates that the sum on the right-hand side converges and is analytic for all $x \in [0, 1)$. Only untwisted-sector operators occur in the $\sigma_1\sigma_{-1}$ OPE; these are of the form $\mathcal{A}'_m = (\mathcal{A}_{i_1} \otimes \cdots \otimes \mathcal{A}_{i_n})_{\text{sym}}$, where the $\mathcal{A}_i$ are operators of $\mathcal{C}$ ("sym" means average over all permutations; if some of the $\mathcal{A}_i$ are identical, then this is the same as the average over distinct permutations).

Before considering the large c limit, it is interesting to ask what the assumed analyticity in α of $I^{(\alpha)}(x)$ implies about the coefficients on the right-hand side of (4.36). First, except for the leading term, 1, the total coefficient of each power of x must be a multiple of $n-1$ (i.e. when analytically continued must have a root at $n = 1$). Although we do not know a CFT proof of this statement, it can be tested to any given order. For example, for the holomorphic and antiholomorphic parts of the stress tensor, T' and $\tilde{T}'$, we have

$$c^{\sigma_1}{}_{\sigma_1 T'} = c^{\sigma_1}{}_{\sigma_1 \tilde{T}'} = \frac{d_\sigma}{2}. \tag{4.37}$$

The Zamolodchikov metric on these operators is $\mathcal{G}_{T'T'} = \mathcal{G}_{\tilde{T}'\tilde{T}'} = nc/2$, since the central charge of the orbifold theory is nc; they do not mix with each other or with other operators, so $\mathcal{G}^{T'T'} = \mathcal{G}^{\tilde{T}'\tilde{T}'} = 2/(nc)$, and

$$c^{T'}{}_{\sigma_1\sigma_{-1}} = c^{\tilde{T}'}{}_{\sigma_1\sigma_{-1}} = \frac{d_\sigma}{nc}. \tag{4.38}$$

Hence their contribution to the sum on the right-hand side of (4.36) is $2ax^2$, where

$$a \equiv \frac{d_\sigma^2}{2nc} = \frac{(n^2-1)^2 c}{288 n^3}, \tag{4.39}$$

which has the required factor of $n-1$. The contribution of the operator $(\mathcal{O}_i \otimes \mathcal{O}_i \otimes I \otimes \cdots \otimes I)_{\text{sym}}$, where $\mathcal{O}_i$ is a primary of $\mathcal{C}$, is computed in Appendix A 2 (for integer values of the scaling dimension of $\mathcal{O}_i$); the result again has a zero at $n = 1$. More generally, analyticity of $I^{(\alpha)}(x)$ demands that the contribution of an operator made up of k nonunit operators of $\mathcal{C}$, $(\mathcal{A}_{i_1} \otimes \cdots \otimes \mathcal{A}_{i_k} \otimes I \otimes \cdots \otimes I)_{\text{sym}}$, should contain a factor $(n-1)(n-2)\cdots(n-k+1)$, simply because that operator only exists for $n \geq k$.

We now consider the large c limit. Since we are working order by order in x, we will consider only "short-string" operators of $\mathcal{C}$, i.e. those whose scaling dimensions are finite in the large c limit. A convenient machinery for systematically computing terms in the sum (4.36) is provided by the conformal blocks. We thus write it as a sum over primaries $\mathcal{O}'_m$ (of $\mathcal{C}^n/S_n$):

$$\exp((n-1)I^{(n)}(x)) = \sum_m C_m \mathcal{F}(h_m, x)\mathcal{F}(\tilde{h}_m, x) x^{h_m + \tilde{h}_m}, \tag{4.40}$$

where $C_m \equiv c^{\sigma_1}{}_{\sigma_1 m} c^m{}_{\sigma_1\sigma_{-1}}$, h_m and $\tilde{h}_m$ are the weights of $\mathcal{O}'_m$, and $\mathcal{F}(h_m, x)$ is (up to a factor of $x^{2d_\sigma - h_m}$) the conformal block with all 4 external operators of weight $h_\sigma = \tilde{h}_\sigma = d_\sigma/2$. ($C_m$ and the conformal blocks also depend

[18] We remind the reader that throughout this paper we use n to denote a positive integer and α a non-negative real number. We assume that all quantities are analytic functions of α.

[19] This procedure was applied to the compactified free boson in [3].

implicitly on c and n.) The conformal blocks can be computed straightforwardly (albeit somewhat tediously) to any desired order in x. Using *Mathematica*,[20] we computed them to order x^5. In the following expression, the first two terms are exact, while the subsequent ones have been expanded in powers of $a^{-1} \sim c^{-1}$:

$$\begin{aligned}\mathcal{F}(h_m, x) = 1 &+ \frac{h_m}{2}x + (a + f_{20} + O(a^{-1}))x^2 \\ &+ \left(\left(1 + \frac{h_m}{2}\right)a + f_{30} + O(a^{-1})\right)x^3 \\ &+ \left(\frac{1}{2}a^2 + f_{41}a + f_{40} + O(a^{-1})\right)x^4 \\ &+ \left(\left(1 + \frac{h_m}{4}\right)a^2 + f_{51}a + f_{50} + O(a^{-1})\right)x^5 \\ &+ O(x^6). \end{aligned} \tag{4.41}$$

The f_{ij} are rational functions of h_m and n (regular at $n = 1$); their precise form is not important for us, except for one feature we will point out below. We now factor out the positive powers of a (i.e. of c), since those determine $I_1^{(n)}(x)$. It turns out that they organize themselves naturally into an exponential:

$$\begin{aligned}\mathcal{F}(h_m, x) = F(h_m, x)\exp(&ax^2 + ax^3 + g_4 ax^4 \\ &+ g_5 ax^5 + O(x^6)), \end{aligned} \tag{4.42}$$

where, by definition, F contains only nonpositive powers of a:

$$\begin{aligned}F(h_m, x) = 1 &+ \frac{h_m}{2}x + (f_{20} + O(a^{-1}))x^2 \\ &+ (f_{30} + O(a^{-1}))x^3 + O(x^4). \end{aligned} \tag{4.43}$$

Remarkably, thanks to some cancellations among the f_{ij}, the coefficients of ax^4 and ax^5 turn out to be independent of h_m:

$$\begin{aligned} g_4 &= f_{41} - f_{20} - \frac{h_m}{2} = \frac{1309n^4 - 2n^2 - 11}{1440n^4} \\ g_5 &= f_{51} - f_{30} - f_{20} - g_4\frac{h_m}{2} = \frac{589n^4 - 2n^2 - 11}{720n^4}. \end{aligned} \tag{4.44}$$

Assuming that this pattern continues to higher orders, it allows us to pull the exponential out of the sum (4.40), and write

$$\begin{aligned}I_1^{(n)}(x) = J^{(n)}(x) &+ \frac{(n-1)(n+1)^2}{144n^3}(x^2 + x^3 + g_4 x^4 \\ &+ g_5 x^5 + O(x^6)), \end{aligned} \tag{4.45}$$

[20]For these computations we used the package Virasoro.nb, available at http://people.brandeis.edu/~headrick/physics/.

where $J^{(n)}(x)$ is the contribution to $I_1^{(n)}(x)$ (if any) from the OPE coefficients C_m:

$$J^{(n)}(x) \equiv \frac{1}{n-1}\lim_{c\to\infty}\frac{1}{c}\ln\left(\sum_m C_m F(h_m, x)F(\tilde{h}_m, x)x^{h_m + \tilde{h}_m}\right). \tag{4.46}$$

Before discussing $J^{(n)}(x)$, let us point out several noteworthy features of the second term of (4.45). First, it does not depend at all on the particular theory, supporting the universality proposed in Subsection IV C; this is a consequence of the cancellation of the h_m-dependence in g_4 and g_5, (4.44). Second, if we set $n = 2$, it agrees with the expansion to fifth order of (4.34); hence $J^{(2)}(x)$ vanishes at least to fifth order. Third, it can be straightforwardly continued to noninteger values of α, and vanishes at $\alpha = 1$:

$$\begin{aligned}I_1^{(\alpha)}(x) = J^{(\alpha)}(x) &+ \frac{(\alpha-1)(\alpha+1)^2}{144\alpha^3}(x^2 + x^3 \\ &+ g_4 x^4 + g_5 x^5 + O(x^6)). \end{aligned} \tag{4.47}$$

The fact that the second term vanishes at $\alpha = 1$, which provides strong quantitative evidence in favor of the RT formula, can be traced to the fact that the conformal block (4.41) depends on c through $a \sim (n-1)^2 c$.

It remains to ask what we can say about $J^{(n)}(x)$. Since we are disallowing theories in which the number of primaries of a given dimension is proportional to c, $J^{(n)}(x)$ will be nonzero if and only if some of the coefficients C_m contain positive powers of c. Some (but not all) of the primaries of $\mathcal{C}^n/S_n$ are products of primaries of $\mathcal{C}$: $\mathcal{O}'_m = (\mathcal{O}_{i_1} \otimes \cdots \otimes \mathcal{O}_{i_n})_{\text{sym}}$. For these, as we show in Appendix A 2, the OPE coefficient $c^{\sigma_1}{}_{\sigma_1 m}$ is given by the (symmetrized) n-point function in $\mathcal{C}$ of $\mathcal{O}_{i_1}, \ldots, \mathcal{O}_{i_n}$, where the operators are placed at distinct nth roots of unity (see (A19)). In holographic and $\mathcal{C}^N/S_N$ theories, these n-point functions go like $c^{1-k/2}$, where k is the number of nonidentity operators, so $C_m \sim c^{2-k}$ (this is for $k > 1$; for $k = 0$, i.e. the identity of $\mathcal{C}^N/S_N$, $C_1 = 1$, while for $k = 1$, $C_m = 0$, since the one-point function of a nonidentity operator vanishes). For example, for $k = 2$ we have a two-point function, which is clearly independent of c (C_m is computed in this case in Appendix A 2). Hence primaries of $\mathcal{C}^n/S_n$ that are products of primaries of $\mathcal{C}$ do not contribute to $J^{(n)}(x)$. However, there are other primaries of $\mathcal{C}^n/S_n$ that are made up of descendants of $\mathcal{C}$. The simplest example is

$$\begin{aligned}\Big(L_{-1}L_{-1}\mathcal{O}_i \otimes \mathcal{O}_i \otimes I \otimes \cdots \otimes I - \left(1 + \frac{1}{2h_i}\right) \\ \times L_{-1}\mathcal{O}_i \otimes L_{-1}\mathcal{O}_i \otimes I \otimes \cdots \otimes I\Big)_{\text{sym}}. \end{aligned} \tag{4.48}$$

The computation of C_m for such operators is more involved, because of the more complicated transformation law of the $\mathcal{C}$-descendants ($L_{-1}\mathcal{O}_i$ and $L_{-1}L_{-1}\mathcal{O}_i$, in the above example) in going from the n-sheeted plane to the standard plane. It is possible that such operators contribute to $J^{(n)}(x)$. If so, the conjecture is then that $J^{(\alpha)}(x)$ is independent of the particular theory $\mathcal{C}$, and vanishes at $\alpha = 1$. It should be straightforward to compute C_m and test these conjectures in specific examples, such as (4.48). We leave this to future work.

In this subsection, we have provided nontrivial evidence, based on the expansion (4.36), that $I_1^{(\alpha)}(x)$ is theory independent and, for $0 \leq x \leq 1/2$, vanishes at $\alpha = 1$. In view of the pattern we have found, it would clearly be desirable to have some general understanding of the structure of the OPE coefficients at large c that leads to these properties. We leave the exploration of this structure to future work.

V. GENERALIZATIONS, OPEN QUESTIONS, AND DISCUSSION

In the previous section, through the study of Rényi entropies, we provided strong evidence in favor of the Ryu-Takayanagi formula. We focused on one of the simplest nontrivial field theory examples, namely, two disjoint intervals in the vacuum of a two-dimensional CFT. Along the way, we found evidence that a large class of large c theories share the same entanglement (Rényi) entropies. It would be interesting to extend our analysis to more general situations, including: more than two intervals; states other than the vacuum, such as thermal states; CFTs on the circle rather than the line; CFTs in more than two dimensions; and nonconformal field theories. In particular, it is clear that the two key qualitative predictions of the RT formula persist in all these examples, namely, that there is a phase transition in the mutual information between two regions as a function of their sizes and separations, and that it vanishes on one side of the phase transition. One should be able to test these predictions using similar techniques to the ones used in this paper, namely, classical gravity and the OPE. One should also be able to test whether the EREs are the same for nonholographic theories with large central charges.

Our analysis leaves a number of open questions. We will start with the more technical ones, and move towards the more conceptual. First, our calculation of $I_1^{(n)}(x)$ in Subsection IV F left out the term $J^{(n)}(x)$, which comes from primary operators of $\mathcal{C}^n/S_n$, such as (4.48), that are composed of descendants of $\mathcal{C}$. It would be useful to evaluate this term, at least up to some power of x, to confirm both its theory independence and that it vanishes at $n = 1$. More generally, it should be possible to understand on general CFT grounds the pattern found in Subsection IV F that, in the four-point function of twist operators, every factor of c is accompanied by a factor of $(n-1)^2$.

Second, it would be very interesting to compute the MRI $I_1^{(n)}(x)$ explicitly for $n > 2$ in holographic and large N symmetric product theories, to see, first, if they agree, and second, if they indeed have a phase transition at $x = 1/2$. Better yet would be to analytically continue the resulting expressions to general α, and directly confirm or refute the RT formula in this case.

Third, as discussed in Subsection IV B, the RT formula predicts that the mutual information between the two intervals is of order c when they are close together, but only of order 1 when they are far apart. It would be quite illuminating to find operators $\mathcal{A}_{[u_1,v_1]}$, $\mathcal{A}_{[u_2,v_2]}$ that saturate the inequality for $x > 1/2$ (2.6) (or at least such that the left-hand side is of order c), and to understand what happens to them for $x \leq 1/2$. A related question is raised by the fact that, for $x \leq 1/2$, the MRI is of order c (for $\alpha \neq 1$). In order to understand the behavior better, it would be interesting to find a simple toy model system with a large number of degrees of freedom, in which the MRI between two subsystems is of order of the number of degrees of freedom while the MI is only of order 1.

Lastly, and perhaps most importantly, we should ask what the status of the RT formula is, given the results of this paper. On the one hand, we have provided strong evidence that it is correct. On the other hand, have we understood any better *why* it should be true? In particular, why does the minimal surface play a critical role in the entanglement entropy, and what is the physical significance of the bulk region r_A that it bounds? Fursaev's proof [7], though incorrect, had the advantage of explaining in a simple and elegant manner the role of the minimal surface. On the other hand, in the Rényi entropy calculations we have performed in this paper, this role is not so clear. Rather, the agreement between the Rényi entropies and the RT formula appeared to be almost fortuitous. Clearly, while the RT formula provides a tantalizing hint about the structure of quantum information in holographic theories, most of that structure still remains hidden from view.

ACKNOWLEDGMENTS

First and foremost, M. H. would like to thank S. Minwalla for initial collaboration on this project; essentially all of the new ideas in this paper originated during discussions with him. It is also a pleasure to thank the following people for helpful discussions: M. Gaberdiel, S. Hartnoll, V. Hubeny, A. Lawrence, L. Levitov, H. Liu, A. Maloney, J. McGreevy, M. Rangamani, T. Takayanagi, E. Tonni, M. van Raamsdonk, and T. Wiseman; and to thank T. Takayanagi, E. Tonni, and M. van Raamsdonk for useful comments on a draft. Finally, M. H. would like to thank the Tata Institute of Fundamental Research, where this project was initiated, and Korea Institute for Advanced Study, where it was completed, for their hospitality. This research was supported in part by DOE Grant No. DE-FG02-92ER40706.

APPENDIX A: SYMMETRIC PRODUCT CFT COMPUTATIONS

1. Analysis of four-point function of twist operators in $\mathcal{C}^2/\mathbf{Z}_2$

In this Appendix, we consider a general modular-invariant, compact, unitary CFT $\mathcal{C}$ with central charge c, and its symmetric product $\mathcal{C}^2/\mathbf{Z}_2$. (At the end we also make some comments about the general symmetric product $\mathcal{C}^n/S_n$.) The orbifold theory has central charge $2c$, and contains a single twist operator σ with conformal weights $h_\sigma = \tilde{h}_\sigma = c/16$. Lunin and Mathur [17] computed the four-point function of these operators, showing that it is determined by the partition function Z_τ of $\mathcal{C}$ on a flat torus with modular parameter τ,

$$\langle\sigma(0)\sigma(x)\sigma(1)\sigma'(\infty)\rangle = |2^8 x(1-x)|^{-c/12} Z_\tau, \tag{A1}$$

where τ and x are related by[21]

$$x = \frac{\theta_2^4(\tau)}{\theta_3^4(\tau)}. \tag{A2}$$

The reason for the appearance of the torus partition function of $\mathcal{C}$ is that the four-point function of twist operators is the (renormalized) zero-point function on the two-sheeted Riemann surface E_2 with a branch cut connecting 0 to x and another one connecting 1 to ∞. E_2 is a torus with complex structure τ. The Weyl transformation that flattens it leads to the prefactor $|2^8 x(1-x)|^{-c/12}$.

Both sides of (A1) can be decomposed into a sum of states, and we would like to understand the relationship between these two decompositions. The torus partition function is a sum over states $\mathcal{A}_m$ in $\mathcal{C}$[22]:

$$Z_\tau = \sum_m q^{h_m - c/24} \bar{q}^{\tilde{h}_m - c/24} = \sum_i \chi_{c,h_i}(q)\chi_{c,\tilde{h}_i}(\bar{q}), \tag{A3}$$

where $q \equiv e^{2\pi i\tau}$. In the second equality, we have grouped the states into conformal families. Each family is labeled by its primary operator $\mathcal{O}_i$, and χ_{c,h_i} is its Virasoro character

$$\chi_{c,h_i}(q) = q^{-c/24+h_i} \sum_{N=0}^{\infty} d(N) q^N, \tag{A4}$$

where $d(N)$ is the number of descendants of $\mathcal{O}_i$ at level N. The decomposition (A3) can be obtained by cutting the torus along a cycle and inserting a complete set of states. In the usual presentation of the torus as $\mathbf{C}/(\mathbf{Z} + \tau\mathbf{Z})$, that cycle should be horizontal.

[21] In terms of Lunin and Mathur's variables, $x = 1/w$ and $\tau = -1/\tau_{\text{Lunin-Mathur}}$. In the bulk of the paper we consider x to be real and lying in the interval $0 < x < 1$, but in this Appendix we will let x be a general complex number. As in the main text, $\sigma_1'(\infty) \equiv \lim_{z\to\infty} z^{d_\sigma}\sigma_1(z)$.

[22] For simplicity we are taking all states to be bosonic.

Meanwhile, the left-hand side of (A1) can be written as a sum over intermediate states $\mathcal{A}_l'$ of $\mathcal{C}^2/\mathbf{Z}_2$, with weights $(h_l', \tilde{h}_l')$:

$$\langle\sigma(0)\sigma(x)\sigma(1)\sigma'(\infty)\rangle = \sum_l c_{\sigma\sigma l} c^l{}_{\sigma\sigma} x^{h_l' - c/8} \bar{x}^{\tilde{h}_l' - c/8}, \tag{A5}$$

where $c_{\sigma\sigma l} = \langle\sigma'(\infty)\sigma(1)\mathcal{A}_l'(0)\rangle$ and $c^l{}_{\sigma\sigma}$ is the coefficient of $\mathcal{A}_l'$ in the σ-σ OPE. Assuming for clarity that $|x| < 1$, this decomposition is obtained by cutting the sphere on a circle of radius r ($|x| < r < 1$) around the origin, which separates the twist operators located at 0 and x from those located at 1 and ∞, and inserting a complete set of states.

The intermediate states in (A5) can also be organized into conformal families, leading to a sum of conformal blocks. However, since the conformal families of $\mathcal{C}^2/\mathbf{Z}_2$ are not in one-to-one correspondence with the conformal families of $\mathcal{C}$, and we are trying to reproduce the sum (A3) which is over the latter, we will organize the intermediate states slightly differently. First we note that only untwisted states appear in the sum, and these are of the form $\mathcal{A}_l' = \mathcal{A}_m \otimes \mathcal{A}_n + \mathcal{A}_n \otimes \mathcal{A}_m$, where $\mathcal{A}_m$, $\mathcal{A}_n$ are states of $\mathcal{C}$. We are inserting this state on the circle of radius r mentioned in the previous paragraph. In the presence of the twist operators, we can consider that we are working in the theory $\mathcal{C}$ on the Riemann surface E_2, where the circle is two circles, one on each sheet; we are inserting $\mathcal{A}_n$ on one circle and $\mathcal{A}_m$ on the other. These two circles both represent the same cycle of the torus, namely, the horizontal cycle mentioned below (A4). In other words we have cut the torus into two finite cylinders. Each cylinder has $\mathcal{A}_n$ inserted on one boundary and $\mathcal{A}_m$ inserted on the other.

We now perform the Weyl transformation that turns E_2 into the flat torus. Two things will happen. First, we get the geometrical factor $|2^8 x(1-x)|^{-c/12}$, as computed by Lunin and Mathur, which is independent of the states. Second, each state gets mapped by the action of the conformal group to a linear combination of states. By definition, this group acts within conformal families. Hence if $\mathcal{A}_n$ and $\mathcal{A}_m$ are not in the same family, then the cylinder amplitude vanishes. So we can gather the terms in (A5) into conformal families of $\mathcal{C}$:

$$\langle\sigma(0)\sigma(x)\sigma(1)\sigma'(\infty)\rangle = \sum_i K_i(x,\bar{x}), \tag{A6}$$

where

$$K_i(x,\bar{x}) = \sum_{\mathcal{A}_m,\mathcal{A}_n \text{ descendants of } \mathcal{O}_i} c_{\sigma\sigma(m,n)} c^{(m,n)}{}_{\sigma\sigma} x^{h_m+h_n-c/8} \bar{x}^{\tilde{h}_m+\tilde{h}_n-c/8}. \tag{A7}$$

(The set of operators in $\mathcal{C}^2/\mathbf{Z}_2$ of the form $\mathcal{A}_m \otimes \mathcal{A}_n + \mathcal{A}_n \otimes \mathcal{A}_m$ where $\mathcal{A}_m$ and $\mathcal{A}_n$ are both descendants of the primary $\mathcal{O}_i$ in $\mathcal{C}$, is the union of several conformal

families of $\mathcal{C}^2/\mathbf{Z}_2$. Hence K_i includes several conformal blocks of $\mathcal{C}^2/\mathbf{Z}_2$.) Each term of (A6) corresponds to precisely one term in the sum on the right-hand side of (A3), and the Lunin-Mathur formula tells us that

$$K_i(x,\bar{x}) = |2^8 x(1-x)|^{-c/12}\chi_{c,h_i}(q)\chi_{c,\tilde{h}_i}(\bar{q}). \tag{A8}$$

It is interesting that K_i factorizes as a holomorphic times an antiholomorphic function.

Each state in the sum (A7) contributes to K_i a monomial in x, $\bar{x}$, while each state in the sum (A4) contributes to $\chi_{c,h_i}(q)$ a monomial in q. The complicated mixing between states due to the action of the conformal group is reflected in the complicated relationship between x and q. However, the leading terms for small x on the two sides of (A8) can be matched easily. On the right-hand side the leading term is due to the primary $\mathcal{O}_i$ itself, so we have

$$|2^8 x|^{-c/12} q^{-c/24+h_i}\bar{q}^{-c/24+\tilde{h}_i} \approx 2^{-8h_i-8\tilde{h}_i}x^{-c/8+2h_i}\bar{x}^{-c/8+2\tilde{h}_i}, \tag{A9}$$

where we used the expansion for small x, $q \approx 2^{-8}x^2$. The leading term on the left-hand side is due to the operator $\mathcal{O}_i' = \mathcal{O}_i \otimes \mathcal{O}_i$, which has weights $(h_i', \tilde{h}_i') = (2h_i, 2\tilde{h}_i)$. $\mathcal{O}_i'$ is primary, so (taking it to be normalized in the Zamolodchikov metric) we have $c_{\sigma\sigma i'} = c^{i'}{}_{\sigma\sigma} = \langle\sigma(0)\mathcal{O}_i'(1)\sigma'(\infty)\rangle$. To evaluate this three-point function, we consider the theory $\mathcal{C}$ on the two-sheeted Riemann surface with a branch cut extending from 0 to ∞, and with $\mathcal{O}_i$ inserted at the point $z = 1$ on both sheets. We can use the map $z = t^2$ to relate this to the two-point function $\langle\mathcal{O}_i(-1)\mathcal{O}_i(1)\rangle$ in the t frame. (The factor arising from the Weyl transformation is absorbed in the renormalization of the twist fields.) All in all we find

$$c_{\sigma\sigma i'} = c^{i'}{}_{\sigma\sigma} = 2^{-4h_i-4\tilde{h}_i}, \tag{A10}$$

which leads immediately to agreement with (A9).

In principle Eq. (A8) can be checked to higher orders. Consider, for example, the conformal family of the identity. For convenience, let us divide both sides of (A8) by the leading term

$$|x|^{c/4}K_1(x,\bar{x}) = \left|\left(2^8\frac{1-x}{x^2}\right)^{-c/24}\chi_{c,0}(q)\right|^2. \tag{A11}$$

Generically, the conformal family of the identity is a full Verma module except the states $L_{-1}|0\rangle$ and $\tilde{L}_{-1}|0\rangle$ and their would-be descendants, which vanish. (For the minimal models there are also other missing states.) In that case the character is

$$\chi_{c,0}(q) = q^{-c/24}\prod_{n=2}^{\infty}\frac{1}{1-q^n} = q^{-c/24}\frac{q^{1/24}(1-q)}{\eta(q)}, \tag{A12}$$

so the holomorphic part of (A11) is

$$\left(2^8\frac{1-x}{x^2}q\right)^{-c/24}\prod_{n=2}^{\infty}\frac{1}{1-q^n}. \tag{A13}$$

The first few terms in the expansion in powers of x are

$$1 + 2^{-8}cx^2 + 2^{-8}cx^3 + 2^{-17}(c^2+465c+2)x^4. \tag{A14}$$

In the expansion in states of $\mathcal{C}^2/\mathbf{Z}_2$, (A7), the quadratic term is due to the stress tensor, while the cubic term is due to $L_{-3}|0\rangle \cong \partial T$. The correct matching of the coefficient for the former can be seen by setting $n = 2$ in (4.39).

Modular invariance means that the torus partition function can be written as a sum of characters in a different way, namely

$$Z_\tau = \sum_i \chi_{c,h_i}(\hat{q})\chi_{c,\tilde{h}_i}(\hat{\bar{q}}), \tag{A15}$$

where $\hat{q} \equiv e^{-2\pi i/\tau}$. This decomposition is produced by cutting the torus along its "vertical" cycle. Meanwhile, associativity of the OPE means that the four-point function of twist operators can be decomposed in intermediate states with each state contributing a power of $1-x$ (instead of x as in (A5)), by cutting along a circle centered on 1 that separates 1 and x from 0 and ∞. That circle corresponds to two circles on E_2, both representing the vertical cycle. Thus the two decompositions can be mapped to each other just as we did above. It is interesting that the associativity of the OPE in $\mathcal{C}^2/\mathbf{Z}_2$ is directly related to the modular invariance of $\mathcal{C}$.

If we attempt to generalize the above analysis to the analogous four-point function of twist operators

$$\langle\sigma_1(0)\sigma_{-1}(x)\sigma_1(1)\sigma'_{-1}(\infty)\rangle \tag{A16}$$

in the $\mathcal{C}^n/S_n$ orbifold theory, the following structure emerges. The Riemann surface E_n has n sheets joined by a branch cut extending from 0 to x and another one extending from 1 to ∞. This surface has genus $n-1$, and the circle centered on 0, that separates the twist operators located at 0 and x from those located at 1 and ∞ decomposes into n circles, which separate E_n into two n-punctured spheres. Again, only untwisted states, which are of the form $(\mathcal{A}_{m_1}\otimes\cdots\otimes\mathcal{A}_{m_n})_{\mathrm{sym}}$, enter in the sum we insert on that circle. We are left with a sum of squares of n-point functions of $\mathcal{C}$ (to be contrasted with (A5), which is a sum of squares of three-point functions of $\mathcal{C}^2/\mathbf{Z}_2$, or in this case $\mathcal{C}^n/S_n$). Unlike in the $\mathcal{C}^2/\mathbf{Z}_2$ case, the $\mathcal{A}_{m_i}$ do not all have to belong to the same conformal family of $\mathcal{C}$ to contribute to this sum. For this reason, this decomposition is less immediately useful than for the case $n = 2$. There will also be an overall geometrical factor coming from the appropriate Weyl transformation.

2. Computation of certain OPE coefficients in $\mathcal{C}^n/S_n$

In this Appendix we will consider primary operators in the symmetric product theory $\mathcal{C}^n/S_n$, of the form

$$\mathcal{O}'_m = (\mathcal{O}_{i_1} \otimes \cdots \otimes \mathcal{O}_{i_n})_{\text{sym}}, \tag{A17}$$

where the $\mathcal{O}_i$ are primaries of $\mathcal{C}$. We will first show that $c^{\sigma_1}{}_{\sigma_1 m}$, the OPE coefficient with the twist operator σ_1, is given in terms of the n-point function (in $\mathcal{C}$) of the component operators $\mathcal{O}_i$. We will then focus on the simplest nontrivial case, with only two nonidentity operators (necessarily the same, for $c^{\sigma_1}{}_{\sigma_1 m}$ to be nonzero):

$$\mathcal{O}'_m = (\mathcal{O}_i \otimes \mathcal{O}_i \cdots \otimes I \otimes \cdots \otimes I)_{\text{sym}}. \tag{A18}$$

We calculate $C_m = c^{\sigma_1}{}_{\sigma_1 m} c^m{}_{\sigma_1 \sigma_{-1}}$ in the case where the scaling dimension d_i of $\mathcal{O}_i$ is an integer, showing that its analytic continuation to noninteger n vanishes in the limit $n \to 1$.

We begin with the more general operator (A17). We compute

$$\begin{aligned} c^{\sigma_1}{}_{\sigma_1 m} &= c_{\sigma_{-1} m \sigma_1} \\ &= \langle \sigma_{-1}(0) \mathcal{O}'_m(1) \sigma'_1(\infty) \rangle_{\mathcal{C}^n/S_n} \\ &= \langle \sigma^\epsilon_{-1}(0) \sigma^\epsilon_1(1) \rangle^{-1}_{\mathcal{C}^n/S_n} \langle \sigma^\epsilon_{-1}(0) \mathcal{O}'_m(1) \sigma^{\epsilon\prime}_1(\infty) \rangle_{\mathcal{C}^n/S_n} \\ &= \langle \sigma^\epsilon_{-1}(0) \sigma^\epsilon_1(1) \rangle^{-1}_{\mathcal{C}^n/S_n} \\ &\quad \times \left(\langle \mathcal{O}_{i_1}(e^{2\pi i}) \mathcal{O}_{i_2}(e^{4\pi i}) \cdots \mathcal{O}_{i_n}(e^{2\pi i n}) \rangle_{\mathcal{C} \text{ on } E_n} \right)_{\text{sym}} \\ &= n^{-\sum_j d_{i_j}} (\langle \mathcal{O}_{i_1}(e^{2\pi i/n}) \mathcal{O}_{i_2}(e^{4\pi i/n}) \cdots \mathcal{O}_{i_n}(1) \rangle_{\mathcal{C}})_{\text{sym}}. \end{aligned} \tag{A19}$$

(All correlators except the one marked "$\mathcal{C}$ on E_n" are evaluated on the Riemann sphere.) In the first line we used the fact that the twist operators are normalized, and both they and $\mathcal{O}'_m$ are primary. In the fourth we used the definition of the twist operators to move to the original theory $\mathcal{C}$ on the n-sheeted surface E_n, where the sheets are connected by a branch cut running from 0 to ∞ (the positions of the twist operators). The operator $\mathcal{O}_{i_j}$ is positioned at 1 on the jth sheet, denoted $e^{2\pi i j}$. In the last line we conformally mapped E_n to the plane by $t = z^{1/n}$. The geometrical factor from the associated Weyl tranformation is independent of the operator insertions, cancelling the factor $\langle \sigma^\epsilon_{-1}(0) \sigma^\epsilon_1(1) \rangle^{-1}$. The operator positions are mapped to the nth roots of unity. Since each $\mathcal{O}_{i_j}$ is primary, under the conformal transformation it becomes, in the t frame, $|\partial z/\partial t|^{-d_{i_j}} \mathcal{O}_{i_j} = n^{-d_{i_j}} \mathcal{O}_{i_j}$ (we are assuming for simplicity that the operators are scalars; including spin there are additional phase factors, which depend on the position and therefore have to be included before the symmetrization over permutations).

To compute C_m one also needs to know the Zamolodchikov metric for $\mathcal{O}'_m$. A bit of combinatorics shows that this is

$$\mathcal{G}_{mm} = \frac{n_1! \cdots n_k!}{n!}, \tag{A20}$$

where k is the number of types of operators $\mathcal{O}_i$ in $\mathcal{O}'_m$ and the n_l are the number of operators of each type (so $\sum_l n_l = n$).

We now specialize to the operator (A18). From (A19) we obtain

$$\begin{aligned} c^{\sigma_1}{}_{\sigma_1 m} &= n^{-2d_i} \frac{2}{n(n-1)} \sum_{\substack{j,k=1 \\ j<k}}^{n} \langle \mathcal{O}_i(e^{2\pi i j/n}) \mathcal{O}_i(e^{2\pi i k/n}) \rangle \\ &= n^{-2d_i} \frac{1}{n-1} \sum_{j=1}^{n-1} |1 - e^{2\pi i j/n}|^{-2d_i}. \end{aligned} \tag{A21}$$

We wish to analytically continue this expression in n.[23] We were not able to do this for general dimension d_i, but in the next paragraph we will show that, when d_i is a (positive) integer, the sum in (A21) is a polynomial in n of degree $2d_i$, with a root at $n = 1$. Hence $c^{\sigma_1}{}_{\sigma_1 m}$ is a rational function of n that is regular at $n = 1$. Since $\mathcal{G}_{mm} = 2/(n(n-1))$, $C_m = \mathcal{G}^{-1}_{mm} (c^{\sigma_1}{}_{\sigma_1 m})^2$ is a rational function with a zero at $n = 1$. This was the statement that was used in Subsection IV F.

In order to analytically continue the sum in (A21), we note that the summand equals the reside of the pole at $t = e^{2\pi i j/n}$ of the function

$$f(t) = \frac{n}{z(z^n - 1)(1 - z)^{d_i}(1 - 1/z)^{d_i}}. \tag{A22}$$

We are assuming that d_i is a positive integer, so $f(t)$ is single-valued. It is easy to show that it is regular everywhere on the Riemann sphere except for a pole at each nth root of unity. In particular, at $t = 1$ there is a pole of order $2d_i + 1$, and the sum in (A21) equals minus its residue. Writing $u = t - 1$, this is the coefficient of $u^{2d_i - 1}$ in the expansion of the function

$$n(-1)^{d_i + 1} \frac{(1 + u)^{d_i - 1}}{(1 + u)^n - 1}. \tag{A23}$$

Now, it is clear that this coefficient is zero for $n = 1$, since $u^{-1}(1 + u)^{d_i - 1}$ has no term of order $u^{2d_i - 1}$. It remains to show that it is a polynomial of degree $2d_i$. To do this we rewrite (A23) as

$$(-1)^{d_i + 1} (1 + u)^{d_i - 1} \left(\sum_{k=0}^{\infty} \frac{n^k}{(k+1)!} (\ln(1 + u))^{k+1} \right)^{-1}. \tag{A24}$$

[23] The analytic continuation of the sum in (A21) was also considered in [3]. In particular, an expression was derived that allowed numerical approximations to be computed.

When we expand the sum in large parentheses in powers of u, the leading term is u, and after that the coefficient of u^m is a polynomial in n of degree $m-1$. It follows that, when we expand the whole expression in powers of u, the leading term is u^{-1}, and after that the coefficient of u^m is a polynomial in n of degree $m+1$. So, in particular, the coefficient of u^{2d_i-1} is a polynomial of degree $2d_i$.

[1] M. M. Wolf, F. Verstraete, M. B. Hastings, and J. I. Cirac, Phys. Rev. Lett. **100**, 070502 (2008).
[2] C. Holzhey, F. Larsen, and F. Wilczek, Nucl. Phys. **B424**, 443 (1994).
[3] P. Calabrese, J. Cardy, and E. Tonni, J. Stat. Mech. (2009) P11001.
[4] S. Ryu and T. Takayanagi, Phys. Rev. Lett. **96**, 181602 (2006).
[5] S. Ryu and T. Takayanagi, J. High Energy Phys. 08 (2006) 045.
[6] M. Van Raamsdonk, arXiv:0907.2939.
[7] D. V. Fursaev, J. High Energy Phys. 09 (2006) 018.
[8] D. J. Gross and H. Ooguri, Phys. Rev. D **58**, 106002 (1998).
[9] J. M. Maldacena and A. Strominger, J. High Energy Phys. 12 (1998) 005.
[10] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
[11] D. Petz, *Quantum Information Theory and Quantum Statistics*, Theoretical and Mathematical Physics (Springer-Verlag, Berlin, 2008).
[12] P. Calabrese and J. Cardy, J. Phys. A **42**, 504005 (2009).
[13] E. H. Lieb and M. B. Ruskai, Phys. Rev. Lett. **30**, 434 (1973).
[14] E. H. Lieb and M. B. Ruskai, J. Math. Phys. (N.Y.) **14**, 1938 (1973), with an Appendix by B. Simon.
[15] J. Aczél, B. Forte, and C. T. Ng, Adv. Appl. Probab. **6**, 131 (1974).
[16] W. Ochs, Rep. Math. Phys. **8**, 109 (1975).
[17] O. Lunin and S. D. Mathur, Commun. Math. Phys. **219**, 399 (2001).
[18] P. Calabrese and A. Lefevre, Phys. Rev. A **78**, 032329 (2008).
[19] T. Nishioka, S. Ryu, and T. Takayanagi, J. Phys. A **42**, 504008 (2009).
[20] V. E. Hubeny, M. Rangamani, and T. Takayanagi, J. High Energy Phys. 07 (2007) 062.
[21] R. M. Wald, Phys. Rev. D **48**, R3427 (1993).
[22] J. M. Maldacena, J. High Energy Phys. 04 (2003) 021.
[23] M. Headrick and T. Takayanagi, Phys. Rev. D **76**, 106013 (2007).
[24] V. E. Hubeny and M. Rangamani, J. High Energy Phys. 03 (2008) 006.
[25] P. Calabrese and J. L. Cardy, J. Stat. Mech. (2004) P002.
[26] M. Henningson and K. Skenderis, J. High Energy Phys. 07 (1998) 023.
[27] L. J. Dixon, D. Friedan, E. J. Martinec, and S. H. Shenker, Nucl. Phys. **B282**, 13 (1987).
[28] S. Furukawa, V. Pasquier, and J. Shiraishi, Phys. Rev. Lett. **102**, 170602 (2009).
[29] R. Dijkgraaf, J. M. Maldacena, G. W. Moore, and E. P. Verlinde, arXiv:hep-th/0005003.
[30] K. Krasnov, Adv. Theor. Math. Phys. **4**, 929 (2000).
[31] E. Witten, arXiv:0706.3359.
[32] X. Yin, arXiv:0710.2129.
[33] X. Yin, J. High Energy Phys. 09 (2008) 120.
[34] A. Maloney and E. Witten, J. High Energy Phys. 02 (2010) 029.