

# Dissimilarities of reduced density matrices and eigenstate thermalization hypothesis

Song He<sup>1,2\*</sup>, Feng-Li Lin<sup>3†</sup> and Jia-ju Zhang<sup>4,5‡</sup>

<sup>1</sup>*Max Planck Institute for Gravitational Physics (Albert Einstein Institute),  
Am Mühlenberg 1, 14476 Golm, Germany*

<sup>2</sup>*CAS Key Laboratory of Theoretical Physics, Institute of Theoretical Physics, Chinese Academy of  
Sciences, 55 Zhong Guan Cun East Road, Beijing 100190, China*

<sup>3</sup>*Department of Physics, National Taiwan Normal University, Taipei 11677, Taiwan*

<sup>4</sup>*Dipartimento di Fisica, Università degli Studi di Milano-Bicocca,  
Piazza della Scienza 3, I-20126 Milano, Italy*

<sup>5</sup>*INFN, Sezione di Milano-Bicocca, Piazza della Scienza 3, I-20126 Milano, Italy*

## Abstract

We calculate various quantities that characterize the dissimilarity of reduced density matrices for a short interval of length  $\ell$  in a two-dimensional (2D) large central charge conformal field theory (CFT). These quantities include the Rényi entropy, entanglement entropy, relative entropy, Jensen-Shannon divergence, as well as the Schatten 2-norm and 4-norm. We adopt the method of operator product expansion of twist operators, and calculate the short interval expansion of these quantities up to order of  $\ell^9$  for the contributions from the vacuum conformal family. The formal forms of these dissimilarity measures and the derived Fisher information metric from contributions of general operators are also given. As an application of the results, we use these dissimilarity measures to compare the excited and thermal states, and examine the eigenstate thermalization hypothesis (ETH) by showing how they behave in high temperature limit. This would help to understand how ETH in 2D CFT can be defined more precisely. We discuss the possibility that all the dissimilarity measures considered here vanish when comparing the reduced density matrices of an excited state and a generalized Gibbs ensemble thermal state. We also discuss ETH for a microcanonical ensemble thermal state in a 2D large central charge CFT, and find that it is approximately satisfied for a small subsystem and violated for a large subsystem.

---

\*hesong17@gmail.com

†fengli.lin@gmail.com

‡jiaju.zhang@mib.infn.it

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Prescriptions of the method</b>	<b>3</b>
2.1	CFT basics . . . . .	3
2.2	OPE of twist operators . . . . .	4
2.3	Partition function from twist operators . . . . .	8
2.4	The $n \rightarrow 1$ limit . . . . .	10
2.5	Rényi and entanglement entropies on various Riemann surfaces . . . . .	11
<b>3</b>	<b>Dissimilarities of reduced density matrices</b>	<b>13</b>
3.1	Relative entropy . . . . .	13
3.2	Jensen-Shannon divergence . . . . .	18
3.3	Schatten 2-norm and 4-norm . . . . .	19
<b>4</b>	<b>ETH for canonical ensemble and GGE thermal states</b>	<b>21</b>
<b>5</b>	<b>ETH for microcanonical ensemble thermal state</b>	<b>24</b>
<b>6</b>	<b>Conclusion and discussion</b>	<b>26</b>
<b>A</b>	<b>Relative entropy from modular Hamiltonian</b>	<b>27</b>
<b>B</b>	<b>Contributions from general operators</b>	<b>28</b>

## 1 Introduction

Motivated by the eigenstate thermalization hypothesis (ETH) [1,2] or its generalization, the subsystem ETH [3,4], it is important to characterize quantitatively the difference between the excited state and the thermal state. One such characterization is to quantify the difference between reduced density matrices over a local regions of these two states. This is also an interesting question by itself in quantum information theory. For two-dimensional (2D) conformal field theory (CFT), many other quantities of examining ETH have been adopted, such as correlation functions [5,6], entanglement entropy, Rényi entropy, relative entropy [3,4,7–10], trace square [11], etc. Due to the infinite number of degrees of freedom in CFT, not every quantity is good for the use of examining the ETH [3,4], unless its behaviors for both excited and thermal states are known precisely.

It was proposed in [12] to use correlation functions of twist operators to calculate the Rényi entropy in a 2D CFT, i.e., the partition function of the Riemann surface resulting from the replica trick. When there is no compact form for these twist-operator correlation functions, one can use operator product expansion (OPE) of twist operators to calculate the short interval expansion of Rényi entropy [13–17]. Following this method, in this paper we will calculate various quantities which are just the sums of

some partition functions, and moreover can be used to characterize the dissimilarity of the reduced density matrices of thermal and excited states, and other states on various Riemann surfaces.

Our results can be used to examine ETH. The ETH and subsystem ETH are originally defined by comparing the highly excited state with the microcanonical ensemble thermal state [1–4]. Motivated by [18, 19], as well as [5–8], we compare in [10] the excited state with the canonical ensemble thermal state, and adopt the so-called weak ETH [18, 19]. In [10] the short-interval  $\ell$  expansions of the entanglement entropies for the excited state and canonical ensemble thermal state are calculated to order  $\ell^8$ , and it was found that their difference, which is just the relative entropy, is only suppressed by the powers of large central charge  $c$ , instead of exponential suppression. In this paper we show that there are similar behaviors for the Jensen-Shannon divergence and Schatten 2-norm. For the more refined consideration, one should compare the excited state with the generalized Gibbs ensemble (GGE) thermal state [20–25]. We will discuss the possibility that all the dissimilarities considered in this paper vanish when comparing the reduced density matrices of an excited state and a suitably defined GGE thermal state. As a by-product, we also check ETH for the microcanonical ensemble thermal state with the dissimilarity measures of comparing with the energy eigenstate.

The rest of this paper is arranged as follows. In section 2 we give prescriptions of the method and show how to get the partition functions from OPE of twist operators. Moreover, in subsection 2.5 we apply the prescriptions to evaluate the Rényi and entanglement entropies. In section 3 we calculate the various dissimilarity measures between reduced density matrices. In section 4 we apply our results to examine ETH and the possible scenarios of GGE. In section 5 we discuss ETH for a microcanonical ensemble thermal state in a 2D large central charge CFT, and find that it is approximately satisfied for a small subsystem and violated for a large subsystem. We conclude with discussion in section 6. In appendix A we calculate the relative entropy from modular Hamiltonian as a consistent check. In appendix B we consider the contributions from general operators, and get the formal forms of the various dissimilarity measures and the Fisher information metric.

## 2 Prescriptions of the method

In this section we first give the useful basics of the vacuum conformal family in two-dimensional large central charge CFT and then show how we calculate the partition functions on various Riemann surfaces using OPE of the twist operators.

### 2.1 CFT basics

In this paper we only consider the contributions from the holomorphic sector of the vacuum conformal family in a two-dimensional large central charge CFT, and the generalization to antiholomorphic sector can be figured out easily. We need the quasiprimary operators to level 9, i.e.,  $T$ ,  $\mathcal{A}$ ,  $\mathcal{B}$ ,  $\mathcal{D}$ ,  $\mathcal{E}$ ,  $\mathcal{H}$ ,  $\mathcal{I}$  and  $\mathcal{J}$  as shown in table 1. The definitions, normalization factors, and conformal transformations of the quasiprimary operators up to level 8, as well as some useful structure constants, can be found in [10, 16, 17, 26].

level	0	2	4	6	8	9
operator	1	$T$	$\mathcal{A}$	$\mathcal{B}, \mathcal{D}$	$\mathcal{E}, \mathcal{H}, \mathcal{I}$	$\mathcal{J}$

Table 1: The holomorphic quasiprimary operators to level 9 in vacuum conformal family of a two-dimensional large central charge CFT.

In this paper, we need the additional structure constants

$$\begin{aligned}
C_{TTE} &= \frac{20c(105c+11)}{63}, \quad C_{TTH} = C_{TTI} = 0, \\
C_{AAE} &= \frac{8c(5c+22)(525c+2419)}{315}, \\
C_{AAH} &= -\frac{8c(5c+22)(8400c^2+44575c-6961)}{125(105c+11)}, \\
C_{AAI} &= \frac{3c(2c-1)(3c+46)(5c+3)(5c+22)(7c+68)}{2(1050c^2+3305c-251)}.
\end{aligned} \tag{2.1}$$

Furthermore, at level 9 we have the operator and its normalization

$$\begin{aligned}
-i\mathcal{J} &= (\partial T(\partial T\partial T)) - \frac{6}{5}(\partial^2 T(\partial TT)) + \frac{4}{15}(\partial^3 T(TT)) - \frac{1}{10}(\partial^4 T\partial T) + \frac{1}{100}(\partial^5 TT) - \frac{1}{3150}\partial^7 T, \\
\alpha_{\mathcal{J}} &= \frac{224c(2c-1)(5c+22)}{25},
\end{aligned} \tag{2.2}$$

with  $(\mathcal{X}\mathcal{Y})$  denoting normal ordering of two operators  $\mathcal{X}$  and  $\mathcal{Y}$ . Under a general conformal transformation  $z \rightarrow f(z)$  it transforms as

$$\mathcal{J}(z) = f'^9 \mathcal{J}(f) + \dots + \frac{c(2c-1)(5c+22)(4s^2 s''' + 15s'^3 - 18ss's'')}{259200}, \tag{2.3}$$

where  $s$  denotes the Schwarzian derivative

$$s(x) = \frac{f'''(x)}{f'(x)} - \frac{3}{2} \left( \frac{f''(x)}{f'(x)} \right)^2, \tag{2.4}$$

and  $\dots$  represents the omitted terms that are proportional to  $T, \mathcal{A}, \mathcal{B}, \mathcal{D}$  and their derivatives.

## 2.2 OPE of twist operators

For one short interval  $A = [0, \ell]$  on a Riemann surface  $\mathcal{R}$ , replica trick leads to a CFT on an  $n$ -fold Riemann surface  $\mathcal{R}^n$ . The partition function on  $\mathcal{R}^n$  can be written as a two-point function of twist operators  $\mathcal{T}$  and  $\tilde{\mathcal{T}}$  in an  $n$ -fold CFT on  $\mathcal{R}$  [12]

$$\text{tr}_{\rho_A^n} = \langle \mathcal{T}(\ell) \tilde{\mathcal{T}}(0) \rangle_{\mathcal{R}}, \quad h_{\mathcal{T}} = h_{\tilde{\mathcal{T}}} = \frac{c(n^2-1)}{24n}, \tag{2.5}$$

and the  $n$  folds of the CFT, which we call  $\text{CFT}^n$ , are independent except the connection by the twist operators. In this paper we only consider Riemann surface  $\mathcal{R}$  with translation symmetry, and so the one-point functions are all constants. Using OPE of twist operators [13–17], we may get

$$\langle \mathcal{T}(\ell) \tilde{\mathcal{T}}(0) \rangle_{\mathcal{R}} = \frac{c_n}{\ell^{2h_{\sigma}}} \sum_K d_K \ell^{h_K} \langle \Phi_K(0) \rangle_{\mathcal{R}}, \tag{2.6}$$

and in the summation we only need to consider the quasiprimary operators  $\Phi_K$  in  $\text{CFT}^n$  that are the direct products of the quasiprimary operators in different replicas of the CFT. Only considering the contributions from the vacuum conformal family, we list the quasiprimary operators in  $\text{CFT}^n$  to level 9 in table 2. To level 8, the coefficients  $d_K$  can be found in [16, 26], and using the method in [15] and (2.2), (2.3) we can easily get

$$d_{\mathcal{J}} = 0. \quad (2.7)$$

Interestingly, there is no contribution from level 9 operators, which consist of  $\mathcal{J}$  only.

level	operator	level	operator	level	operator	level	operator
2	$T$	6	$\mathcal{B}, \mathcal{D}$	8	$\mathcal{E}, \mathcal{H}, \mathcal{I}$	8	$TTA$
4	$\mathcal{A}$		$TA$		$T\mathcal{B}, T\mathcal{D}$		$TTTT$
	$TT$		$TTT$		$\mathcal{A}\mathcal{A}$	9	$\mathcal{J}$

Table 2: The holomorphic nonidentity quasiprimary operators to be considered in this paper for  $\text{CFT}^n$  and up to level 9. We have omitted the replica indices and their constraints, which can be easily figured out and can also be found in [26].

Each of the  $\text{CFT}^n$  quasiprimary operator  $\Phi_K$  in (2.6) has the form

$$\Phi_K^{j_1, j_2, \dots, j_k} = \mathcal{X}_1^{j_1} \mathcal{X}_2^{j_2} \dots \mathcal{X}_k^{j_k}, \quad (2.8)$$

with  $\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_k$  being nonidentity quasiprimary operators in table 1 and there are also some constraints for the  $k$  replica indices  $j_1, j_2, \dots, j_k$ . We have the one-point functions that are independent of the replica indices

$$\langle \Phi_K^{j_1 j_2 \dots j_k} \rangle_{\mathcal{R}} = \langle \mathcal{X}_1 \rangle_{\mathcal{R}} \langle \mathcal{X}_2 \rangle_{\mathcal{R}} \dots \langle \mathcal{X}_k \rangle_{\mathcal{R}}, \quad (2.9)$$

and so we can define  $b_K$  from the OPE coefficient  $d_K^{j_1 j_2 \dots j_k}$  by summing over the replica indices [17]

$$b_K = \sum_{j_1, j_2, \dots, j_k} d_K^{j_1 j_2 \dots j_k} \text{ with some constraints for } 0 \leq j_1, j_2, \dots, j_k \leq n-1. \quad (2.10)$$

To level 8 the form of  $b_K$  can be found in [10, 17], and from (2.7) we know

$$b_{\mathcal{J}} = 0. \quad (2.11)$$

Then we write (2.5) explicitly as

$$\begin{aligned} \text{tr}_{A\rho_A^n} = \frac{c_n}{\ell^{2h_\sigma}} & \left[ 1 + b_T \langle T \rangle_{\mathcal{R}} \ell^2 + (b_{\mathcal{A}} \langle \mathcal{A} \rangle_{\mathcal{R}} + b_{TT} \langle T \rangle_{\mathcal{R}}^2) \ell^4 + (b_{\mathcal{B}} \langle \mathcal{B} \rangle_{\mathcal{R}} + b_{\mathcal{D}} \langle \mathcal{D} \rangle_{\mathcal{R}} \right. \\ & + b_{TA} \langle T \rangle_{\mathcal{R}} \langle \mathcal{A} \rangle_{\mathcal{R}} + b_{TTT} \langle T \rangle_{\mathcal{R}}^3) \ell^6 + (b_{\mathcal{E}} \langle \mathcal{E} \rangle_{\mathcal{R}} + b_{\mathcal{H}} \langle \mathcal{H} \rangle_{\mathcal{R}} + b_{\mathcal{I}} \langle \mathcal{I} \rangle_{\mathcal{R}} \\ & + b_{T\mathcal{B}} \langle T \rangle_{\mathcal{R}} \langle \mathcal{B} \rangle_{\mathcal{R}} + b_{T\mathcal{D}} \langle T \rangle_{\mathcal{R}} \langle \mathcal{D} \rangle_{\mathcal{R}} + b_{\mathcal{A}\mathcal{A}} \langle \mathcal{A} \rangle_{\mathcal{R}}^2 + b_{TTA} \langle T \rangle_{\mathcal{R}}^2 \langle \mathcal{A} \rangle_{\mathcal{R}} \\ & \left. + b_{TTTT} \langle T \rangle_{\mathcal{R}}^4) \ell^8 + O(\ell^{10}) \right]. \quad (2.12) \end{aligned}$$

Due to the absence of level 9 contribution, in the above the unknown terms start from  $O(\ell^{10})$ .

In this paper we consider several different Riemann surfaces that are environments of a short interval  $A = [0, \ell]$ , and they are shown in figure 1. Note that the complex plane case figure 1(a) can be got as limits of other six cases.

- In figure 1(a), the interval is on an infinite straight line in ground state of the CFT. It is just a complex plane  $\mathcal{R}(\emptyset)$ , and we denote the total system density matrix as  $\rho(\emptyset)$  and reduced density matrix as  $\rho_A(\emptyset)$ .
- In figure 1(b), the interval is on a length  $L$  circle in ground state, and it is a vertical cylinder  $\mathcal{R}(L)$ . We have the density matrix  $\rho(L)$  and reduced density matrix  $\rho_A(L)$ .
- In figure 1(c), the interval is on a circle in excited state  $|\phi\rangle$  of a primary operator  $\phi$  with conformal weight  $h_\phi$  and normalization  $\alpha_\phi = 1$ . The manifold is a vertical cylinder capped with an operator inserted at each of the two ends, and we denote it as  $\mathcal{R}(L, \phi)$ . We have the density matrix  $\rho(L, \phi)$  and reduced density matrix  $\rho_A(L, \phi)$ .
- In figure 1(d), the interval is on an infinite straight line in thermal state with inverse temperature  $\beta$ . The manifold is a horizontal cylinder  $\mathcal{R}(\beta)$ , and it is the modular transformation of  $\mathcal{R}(L)$ . We have the density matrix  $\rho(\beta)$  and reduced density matrix  $\rho_A(\beta)$ .
- Figure 1(e) is the modular transformation of figure 1(c). The interval is on an infinite straight line in thermal state with inverse temperature  $\beta$ , and also there are boundary conditions imposed on both ends of the horizontal cylinder. Each boundary condition is effectively represented by insertion of a primary operator  $\phi$ . We have the Riemann surface  $\mathcal{R}(\beta, \phi)$ , the density matrix  $\rho(\beta, \phi)$  and reduced density matrix  $\rho_A(\beta, \phi)$ .
- In figure 1(f), the interval is on a length  $L$  circle in thermal state with inverse temperature  $\beta$ . The temperature is low  $\beta \gg L$ , and the manifold is a fat torus. In limit  $\beta/L \rightarrow \infty$ , it becomes a vertical cylinder figure 1(b). We have the Riemann surface  $\mathcal{R}(L, q)$ , the density matrix  $\rho(L, q)$  and reduced density matrix  $\rho_A(L, q)$ , with definition  $q = e^{-2\pi\beta/L}$ .
- In figure 1(g), the interval is on a length  $L$  circle in thermal state with inverse temperature  $\beta$ . The temperature is high  $L \gg \beta$ , the manifold is a thin torus, and it is the modular transformation of the fat torus figure 1(f). In limit  $L/\beta \rightarrow \infty$ , it becomes the horizontal cylinder figure 1(d). We have the Riemann surface  $\mathcal{R}(\beta, p)$ , the density matrix  $\rho(\beta, p)$  and reduced density matrix  $\rho_A(\beta, p)$ , with  $p = e^{-2\pi L/\beta}$ .

We need the one-point functions  $\langle \mathcal{X} \rangle_{\mathcal{R}}$  with  $\mathcal{X} = T, \mathcal{A}, \mathcal{B}, \mathcal{D}, \mathcal{E}, \mathcal{H}, \mathcal{I}$  for  $\mathcal{R}$  being each of these Riemann surfaces in figure 1. In practice, we only need to consider the cases of  $\mathcal{R}(L, \phi)$  and  $\mathcal{R}(L, q)$ , and the other cases can be got from them by some simple substitutes and/or limits. For the case  $\mathcal{R}(L, \phi)$  one can find the results in [10]. For the case  $\mathcal{R}(L, q)$  one can find the results to level 6 in [17]. Using the method in appendix B of [17], the conformal transformations of  $\mathcal{E}, \mathcal{H}, \mathcal{I}$  in [10], as well as

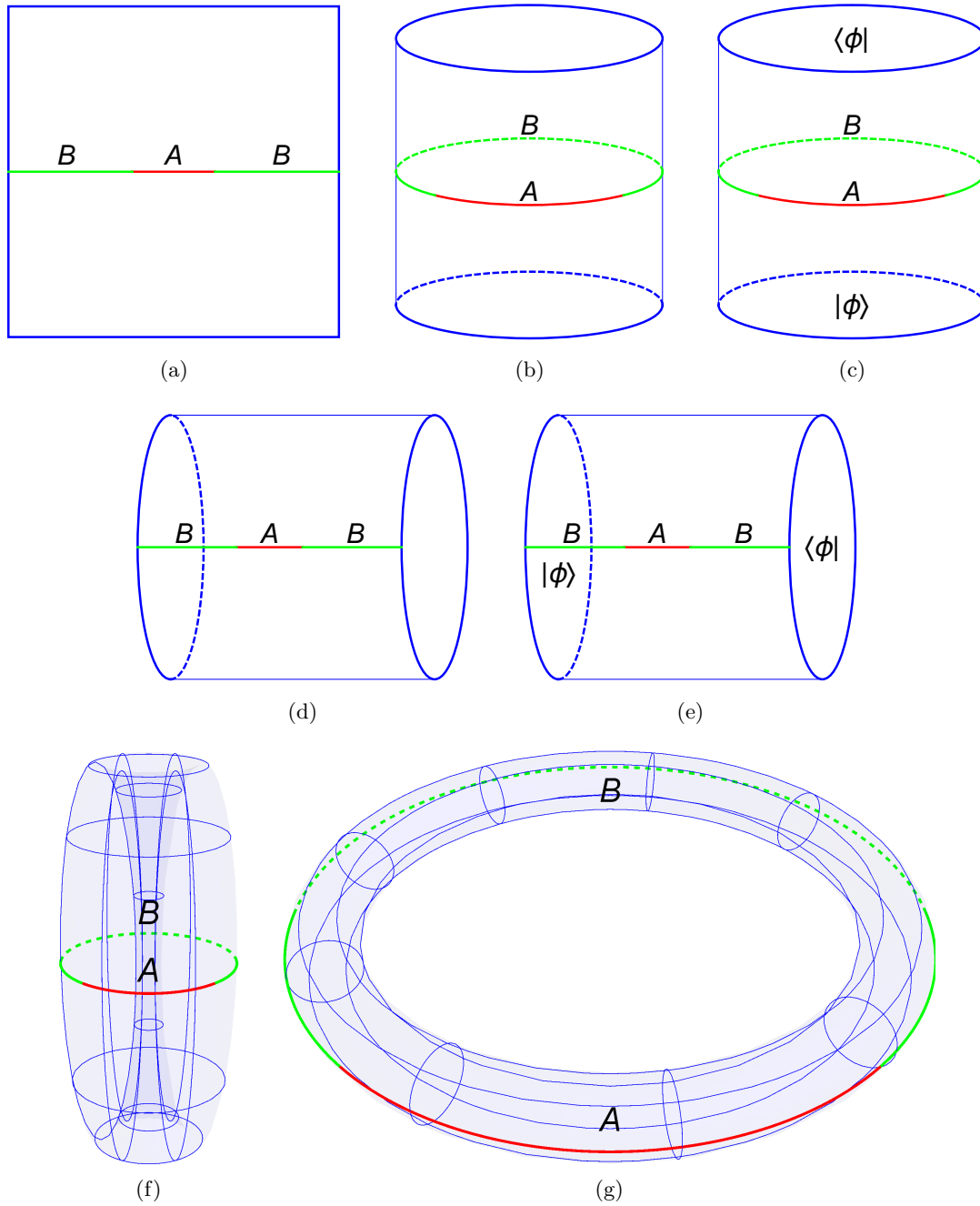


Figure 1: The Riemann surfaces as environments for the interval  $A = [0, \ell]$  we consider in this paper. (a) A complex plane  $\mathcal{R}(\emptyset)$ . (b) A vertical cylinder  $\mathcal{R}(L)$ . (c) A vertical cylinder capped with operators  $\mathcal{R}(L, \phi)$ . (d) A horizontal cylinder  $\mathcal{R}(\beta)$ . (e) A horizontal cylinder capped with operators  $\mathcal{R}(\beta, \phi)$ . (f) A fat torus  $\mathcal{R}(L, q = e^{-2\pi\beta/L})$ . (g) A thin torus  $\mathcal{R}(\beta, p = e^{-2\pi L/\beta})$ .

the structure constants in [17] and (2.1), we get the one-point functions

$$\begin{aligned}
\langle \mathcal{E} \rangle_{\mathcal{R}(L,q)} &= \frac{23452\pi^8 c}{59535L^8} + \frac{36608\pi^8(1008c-1)q^2}{3969L^8} + \frac{18304\pi^8(13608c+335)q^3}{1323L^8} \\
&\quad + \frac{36608\pi^8(54096c+5795)q^4}{1323L^8} + O(q^5), \\
\langle \mathcal{H} \rangle_{\mathcal{R}(L,q)} &= -\frac{13\pi^8 c(5c+22)(465c-127)}{10125(105c+11)L^8} + \frac{1664\pi^8(5c+22)(945c^2+2184c-10)q^2}{675(105c+11)L^8} \\
&\quad + \frac{416\pi^8(5c+22)(11340c^2+31323c-1517)q^3}{225(105c+11)L^8} \\
&\quad + \frac{1664\pi^8(64575c^3+334935c^2+226879c+26048)q^4}{225(105c+11)L^8} + O(q^5), \\
\langle \mathcal{I} \rangle_{\mathcal{R}(L,q)} &= \frac{\pi^8 c(2c-1)(3c+46)(5c+3)(5c+22)(7c+68)}{1296(1050c^2+3305c-251)L^8} \left(1 + \frac{3264q^2}{c} + \frac{13536q^3}{c}\right. \\
&\quad \left. + \frac{576(325c+4814)q^4}{c(5c+22)} + O(q^5)\right).
\end{aligned} \tag{2.13}$$

### 2.3 Partition function from twist operators

Gluing  $n$  reduced density matrices  $\rho_{A,j}$  on  $n$  different Riemann surface  $\mathcal{R}_j$  with  $j = 0, 1, \dots, n-1$ , one gets a CFT on the Riemann surface  $\mathcal{R}^n = \mathcal{R}_0 \oplus \dots \oplus \mathcal{R}_{n-1}$ . This suggests to assume that the partition function on  $\mathcal{R}^n$  can still be written as a two-point function of twist operators

$$\text{tr}_A(\rho_{A,0} \cdots \rho_{A,n-1}) = \langle \mathcal{T}(\ell) \tilde{\mathcal{T}}(0) \rangle_{\mathcal{R}_0 \oplus \dots \oplus \mathcal{R}_{n-1}}. \tag{2.14}$$

Each replica of the CFT lives on one of the Riemann surfaces, and different replicas are connected only by twist operators. For the  $n=2$  and  $n=3$  cases one can see, for examples, [11, 21, 22, 27–33], but we are not sure if it is applicable for general  $n$  when  $Z_n$  replica symmetry is lost. Actually, in this paper we only use a relaxed relation

$$\frac{1}{n!} [\text{tr}_A(\rho_{A,0} \cdots \rho_{A,n-1}) + \text{permutations}] = \frac{1}{n!} [\langle \mathcal{T}(\ell) \tilde{\mathcal{T}}(0) \rangle_{\mathcal{R}_0 \oplus \dots \oplus \mathcal{R}_{n-1}} + \text{permutations}], \tag{2.15}$$

and  $Z_n$  replica symmetry is recovered after permutations. Thus when we write (2.14), we actually mean (2.15), and there is caveat that (2.15) basically is an assumption that we have no concrete proof.

For two different Riemann surfaces  $\mathcal{R}$  and  $\mathcal{S}$ , we may define respectively two reduced density matrices  $\rho_A$  and  $\sigma_A$ . In this paper, we need to calculate the partition function

$$\text{tr}_A(\rho_A^m \sigma_A^{n-m}), \tag{2.16}$$

with  $n$  being an integer and  $m = 0, 1, \dots, n$ . Using (2.15), we see that it is just the right-hand side of



(2.12) with the substitutes of the forms

$$\begin{aligned}
b_{\mathcal{X}}\langle\mathcal{X}\rangle_{\mathcal{R}} &\rightarrow \frac{b_{\mathcal{X}}}{n} [m\langle\mathcal{X}\rangle_{\mathcal{R}} + (n-m)\langle\mathcal{X}\rangle_{\mathcal{S}}], \\
b_{\mathcal{X}\mathcal{X}}\langle\mathcal{X}\rangle_{\mathcal{R}}^2 &\rightarrow \frac{b_{\mathcal{X}\mathcal{X}}}{n(n-1)} [m(m-1)\langle\mathcal{X}\rangle_{\mathcal{R}}^2 + 2m(n-m)\langle\mathcal{X}\rangle_{\mathcal{R}}\langle\mathcal{X}\rangle_{\mathcal{S}} + (n-m)(n-m-1)\langle\mathcal{X}\rangle_{\mathcal{S}}^2], \\
b_{\mathcal{X}\mathcal{Y}}\langle\mathcal{X}\rangle_{\mathcal{R}}\langle\mathcal{Y}\rangle_{\mathcal{R}} &\rightarrow \frac{b_{\mathcal{X}\mathcal{Y}}}{n(n-1)} [m(m-1)\langle\mathcal{X}\rangle_{\mathcal{R}}\langle\mathcal{Y}\rangle_{\mathcal{R}} + m(n-m)(\langle\mathcal{X}\rangle_{\mathcal{R}}\langle\mathcal{Y}\rangle_{\mathcal{S}} + \langle\mathcal{X}\rangle_{\mathcal{S}}\langle\mathcal{Y}\rangle_{\mathcal{R}}) \\
&\quad + (n-m)(n-m-1)\langle\mathcal{X}\rangle_{\mathcal{S}}\langle\mathcal{Y}\rangle_{\mathcal{S}}], \\
b_{\mathcal{X}\mathcal{X}\mathcal{X}}\langle\mathcal{X}\rangle_{\mathcal{R}}^3 &\rightarrow \frac{b_{\mathcal{X}\mathcal{X}\mathcal{X}}}{n(n-1)(n-2)} [m(m-1)(m-2)\langle\mathcal{X}\rangle_{\mathcal{R}}^3 + 3m(m-1)(n-m)\langle\mathcal{X}\rangle_{\mathcal{R}}^2\langle\mathcal{X}\rangle_{\mathcal{S}} \\
&\quad + 3m(n-m)(n-m-1)\langle\mathcal{X}\rangle_{\mathcal{R}}\langle\mathcal{X}\rangle_{\mathcal{S}}^2 + (n-m)(n-m-1)(n-m-2)\langle\mathcal{X}\rangle_{\mathcal{S}}^3], \\
b_{\mathcal{X}\mathcal{X}\mathcal{Y}}\langle\mathcal{X}\rangle_{\mathcal{R}}^2\langle\mathcal{Y}\rangle_{\mathcal{R}} &\rightarrow \frac{b_{\mathcal{X}\mathcal{X}\mathcal{Y}}}{n(n-1)(n-2)} [m(m-1)(m-2)\langle\mathcal{X}\rangle_{\mathcal{R}}^2\langle\mathcal{Y}\rangle_{\mathcal{R}} \\
&\quad + m(m-1)(n-m)(\langle\mathcal{X}\rangle_{\mathcal{R}}^2\langle\mathcal{Y}\rangle_{\mathcal{S}} + 2\langle\mathcal{X}\rangle_{\mathcal{R}}\langle\mathcal{X}\rangle_{\mathcal{S}}\langle\mathcal{Y}\rangle_{\mathcal{R}}) \\
&\quad + m(n-m)(n-m-1)(2\langle\mathcal{X}\rangle_{\mathcal{R}}\langle\mathcal{X}\rangle_{\mathcal{S}}\langle\mathcal{Y}\rangle_{\mathcal{S}} + \langle\mathcal{X}\rangle_{\mathcal{S}}^2\langle\mathcal{Y}\rangle_{\mathcal{R}}) \\
&\quad + (n-m)(n-m-1)(n-m-2)\langle\mathcal{X}\rangle_{\mathcal{S}}^2\langle\mathcal{Y}\rangle_{\mathcal{S}}], \\
b_{\mathcal{X}\mathcal{X}\mathcal{X}\mathcal{X}}\langle\mathcal{X}\rangle_{\mathcal{R}}^4 &\rightarrow \frac{b_{\mathcal{X}\mathcal{X}\mathcal{X}\mathcal{X}}}{n(n-1)(n-2)(n-3)} [m(m-1)(m-2)(m-3)\langle\mathcal{X}\rangle_{\mathcal{R}}^4 \\
&\quad + 4m(m-1)(m-2)(n-m)\langle\mathcal{X}\rangle_{\mathcal{R}}^3\langle\mathcal{X}\rangle_{\mathcal{S}} \\
&\quad + 6m(m-1)(n-m)(n-m-1)\langle\mathcal{X}\rangle_{\mathcal{R}}^2\langle\mathcal{X}\rangle_{\mathcal{S}}^2 \\
&\quad + 4m(n-m)(n-m-1)(n-m-2)\langle\mathcal{X}\rangle_{\mathcal{R}}\langle\mathcal{X}\rangle_{\mathcal{S}}^3 \\
&\quad + (n-m)(n-m-1)(n-m-2)(n-m-3)\langle\mathcal{X}\rangle_{\mathcal{S}}^4],
\end{aligned} \tag{2.17}$$

with  $\mathcal{X}, \mathcal{Y}$  denoting general quasiprimary operators. A general substitute takes the form

$$b_{\mathcal{X}_1\mathcal{X}_2\cdots\mathcal{X}_k}\langle\mathcal{X}_1\rangle_{\mathcal{R}}\langle\mathcal{X}_2\rangle_{\mathcal{R}}\cdots\langle\mathcal{X}_k\rangle_{\mathcal{R}} \rightarrow \frac{b_{\mathcal{X}_1\mathcal{X}_2\cdots\mathcal{X}_k}}{C_n^k} (C_m^k\langle\mathcal{X}_1\rangle_{\mathcal{R}}\langle\mathcal{X}_2\rangle_{\mathcal{R}}\cdots\langle\mathcal{X}_k\rangle_{\mathcal{R}} + \cdots), \tag{2.18}$$

with  $C_n^k$  and  $C_m^k$  being the binomial coefficients, and in the right hand side we have omitted various terms with some  $\mathcal{R}$ 's being replaced by  $\mathcal{S}$ 's.

In section 3.2, we need to calculate the partition function

$$\mathrm{tr}_A\left(\frac{\rho_A + \sigma_A}{2}\right)^n = \frac{1}{2^n} \sum_{m=0}^n C_n^m \mathrm{tr}_A(\rho_A^m \sigma_A^{n-m}), \tag{2.19}$$

with  $\mathrm{tr}_A(\rho_A^m \sigma_A^{n-m})$  being understood as the left-hand side of (2.15). Using the summation formulas

$$\begin{aligned}
\sum_{m=0}^n C_n^m m &= 2^{n-1}n, \quad \sum_{m=0}^n C_n^m m(m-1) = \sum_{m=0}^n C_n^m m(n-m) = 2^{n-2}n(n-1), \\
\sum_{m=0}^n C_n^m m(m-1)(m-2) &= \sum_{m=0}^n C_n^m m(m-1)(n-m) = 2^{n-3}n(n-1)(n-2), \\
\sum_{m=0}^n C_n^m m(m-1)(m-2)(m-3) &= \sum_{m=0}^n C_n^m m(m-1)(m-2)(n-m) \\
&= \sum_{m=0}^n C_n^m m(m-1)(n-m)(n-m-1) = 2^{n-4}n(n-1)(n-2)(n-3),
\end{aligned} \tag{2.20}$$

we get that (2.19) is just the right-hand side of (2.12) with the substitutes

$$\begin{aligned}
b_{\mathcal{X}}\langle\mathcal{X}\rangle_{\mathcal{R}} &\rightarrow \frac{b_{\mathcal{X}}}{2}(\langle\mathcal{X}\rangle_{\mathcal{R}} + \langle\mathcal{X}\rangle_{\mathcal{S}}), \\
b_{\mathcal{X}\mathcal{X}}\langle\mathcal{X}\rangle_{\mathcal{R}}^2 &\rightarrow \frac{b_{\mathcal{X}\mathcal{X}}}{4}(\langle\mathcal{X}\rangle_{\mathcal{R}} + \langle\mathcal{X}\rangle_{\mathcal{S}})^2, \\
b_{\mathcal{X}\mathcal{Y}}\langle\mathcal{X}\rangle_{\mathcal{R}}\langle\mathcal{Y}\rangle_{\mathcal{R}} &\rightarrow \frac{b_{\mathcal{X}\mathcal{Y}}}{4}(\langle\mathcal{X}\rangle_{\mathcal{R}} + \langle\mathcal{X}\rangle_{\mathcal{S}})(\langle\mathcal{Y}\rangle_{\mathcal{R}} + \langle\mathcal{Y}\rangle_{\mathcal{S}}), \\
b_{\mathcal{X}\mathcal{X}\mathcal{X}}\langle\mathcal{X}\rangle_{\mathcal{R}}^3 &\rightarrow \frac{b_{\mathcal{X}\mathcal{X}\mathcal{X}}}{8}(\langle\mathcal{X}\rangle_{\mathcal{R}} + \langle\mathcal{X}\rangle_{\mathcal{S}})^3, \\
b_{\mathcal{X}\mathcal{X}\mathcal{Y}}\langle\mathcal{X}\rangle_{\mathcal{R}}^2\langle\mathcal{Y}\rangle_{\mathcal{R}} &\rightarrow \frac{b_{\mathcal{X}\mathcal{X}\mathcal{Y}}}{8}(\langle\mathcal{X}\rangle_{\mathcal{R}} + \langle\mathcal{X}\rangle_{\mathcal{S}})^2(\langle\mathcal{Y}\rangle_{\mathcal{R}} + \langle\mathcal{Y}\rangle_{\mathcal{S}}), \\
b_{\mathcal{X}\mathcal{X}\mathcal{X}\mathcal{X}}\langle\mathcal{X}\rangle_{\mathcal{R}}^4 &\rightarrow \frac{b_{\mathcal{X}\mathcal{X}\mathcal{X}\mathcal{X}}}{16}(\langle\mathcal{X}\rangle_{\mathcal{R}} + \langle\mathcal{X}\rangle_{\mathcal{S}})^4.
\end{aligned} \tag{2.21}$$

In section 3.3, we need to calculate

$$\mathrm{tr}_A(\rho_A - \sigma_A)^n = \sum_{m=0}^n C_n^m (-)^{n-m} \mathrm{tr}_A(\rho_A^m \sigma_A^{n-m}). \tag{2.22}$$

Using the fact that

$$\begin{aligned}
\sum_{m=0}^n C_n^m (-)^{n-m} m^k &= 0 \text{ for } k = 0, 1, \dots, n-1, \\
\sum_{m=0}^n C_n^m (-)^{n-m} m^n &= n!,
\end{aligned} \tag{2.23}$$

we get

$$\mathrm{tr}_A(\rho_A - \sigma_A)^n = \sum_{\{\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_n\}} b_{\mathcal{X}_1 \mathcal{X}_2 \dots \mathcal{X}_n} (\langle\mathcal{X}_1\rangle_{\mathcal{R}} - \langle\mathcal{X}_1\rangle_{\mathcal{S}}) (\langle\mathcal{X}_2\rangle_{\mathcal{R}} - \langle\mathcal{X}_2\rangle_{\mathcal{S}}) \cdots (\langle\mathcal{X}_n\rangle_{\mathcal{R}} - \langle\mathcal{X}_n\rangle_{\mathcal{S}}). \tag{2.24}$$

Note that the summation of  $\{\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_n\}$  is over different sets of nonidentity quasiprimary operators and the order of the operators in each set does not matter. For  $n = 2$  it is just the result in [11]. Note that for general  $n$ ,  $b_{\mathcal{X}_1 \mathcal{X}_2 \dots \mathcal{X}_n}$  is complex and has no universal form, and it is related to the  $n$ -point correlation function on complex plane  $\langle\mathcal{X}_1(z_1) \mathcal{X}_2(z_2) \cdots \mathcal{X}_n(z_n)\rangle_{\mathcal{C}}$ .

## 2.4 The $n \rightarrow 1$ limit

If we are only interested in the  $n \rightarrow 1$  limit

$$-\left. \frac{\log \mathrm{tr}_A \rho_A^n}{n-1} \right|_{n \rightarrow 1}, \tag{2.25}$$

instead of the general  $n$  result, there can be a simpler calculation [34, 35]. For each CFT <sup>$n$</sup>  operator  $\Phi_K$ , we may define

$$a_K = -\lim_{n \rightarrow 1} \frac{b_K}{n-1}, \tag{2.26}$$

with  $b_K$  being defined in (2.10). Using the results of  $b_K$  in [10, 17], we get the relevant results of  $a_K$

$$\begin{aligned}
a_T &= -\frac{1}{6}, \quad a_{TT} = -\frac{1}{30c}, \quad a_{TTT} = -\frac{4}{315c^2}, \\
a_{AA} &= -\frac{1}{126c(5c+22)}, \quad a_{TTA} = \frac{1}{315c^2}, \quad a_{TTTT} = -\frac{c+8}{630c^3}.
\end{aligned} \tag{2.27}$$

For the reduced density matrix  $\rho_A$  on Riemann surface  $\mathcal{R}$ , we get

$$-\frac{\log \text{tr}_A \rho_A^n}{n-1} \Big|_{n \rightarrow 1} = \frac{c}{6} \log \frac{\ell}{\epsilon} + a_T \langle T \rangle_{\mathcal{R}} \ell^2 + a_{TT} \langle T \rangle_{\mathcal{R}}^2 \ell^4 + a_{TTT} \langle T \rangle_{\mathcal{R}}^3 \ell^6 + (a_{AA} \langle \mathcal{A} \rangle_{\mathcal{R}})^2 + a_{TTA} \langle T \rangle_{\mathcal{R}}^2 \langle \mathcal{A} \rangle_{\mathcal{R}} + a_{TTTT} \langle T \rangle_{\mathcal{R}}^4 \ell^8 + O(\ell^{10}). \quad (2.28)$$

For the reduced density matrix  $\rho_A, \sigma_A$ , defined respectively on Riemann surface  $\mathcal{R}, \mathcal{S}$ , we get that

$$-\frac{\log \text{tr}_A (\rho_A \sigma_A^{n-1})}{n-1} \Big|_{n \rightarrow 1} \quad (2.29)$$

equals right-hand side of (2.28) with the substitutes

$$\begin{aligned} a_{\mathcal{X}\mathcal{X}} \langle \mathcal{X} \rangle_{\mathcal{R}}^2 &\rightarrow a_{\mathcal{X}\mathcal{X}} (2 \langle \mathcal{X} \rangle_{\mathcal{R}} - \langle \mathcal{X} \rangle_{\mathcal{S}}) \langle \mathcal{X} \rangle_{\mathcal{S}}, \\ a_{\mathcal{X}\mathcal{X}\mathcal{X}} \langle \mathcal{X} \rangle_{\mathcal{R}}^3 &\rightarrow a_{\mathcal{X}\mathcal{X}\mathcal{X}} (3 \langle \mathcal{X} \rangle_{\mathcal{R}} - 2 \langle \mathcal{X} \rangle_{\mathcal{S}}) \langle \mathcal{X} \rangle_{\mathcal{S}}^2, \\ a_{\mathcal{X}\mathcal{X}\mathcal{Y}} \langle \mathcal{X} \rangle_{\mathcal{R}}^2 \langle \mathcal{Y} \rangle_{\mathcal{R}} &\rightarrow a_{\mathcal{X}\mathcal{X}\mathcal{Y}} (2 \langle \mathcal{X} \rangle_{\mathcal{R}} \langle \mathcal{X} \rangle_{\mathcal{S}} \langle \mathcal{Y} \rangle_{\mathcal{S}} + \langle \mathcal{X} \rangle_{\mathcal{S}}^2 \langle \mathcal{Y} \rangle_{\mathcal{R}} - 2 \langle \mathcal{X} \rangle_{\mathcal{S}}^2 \langle \mathcal{Y} \rangle_{\mathcal{S}}), \\ a_{\mathcal{X}\mathcal{X}\mathcal{X}\mathcal{X}} \langle \mathcal{X} \rangle_{\mathcal{R}}^4 &\rightarrow a_{\mathcal{X}\mathcal{X}\mathcal{X}\mathcal{X}} (4 \langle \mathcal{X} \rangle_{\mathcal{R}} - 3 \langle \mathcal{X} \rangle_{\mathcal{S}}) \langle \mathcal{X} \rangle_{\mathcal{S}}^3. \end{aligned} \quad (2.30)$$

Similarly, we get that

$$-\frac{\log \text{tr}_A \left( \frac{\rho_A + \sigma_A}{2} \right)^n}{n-1} \Big|_{n \rightarrow 1} \quad (2.31)$$

equals right-hand side of (2.28) with the substitutes (2.21).

## 2.5 Rényi and entanglement entropies on various Riemann surfaces

Using the above prescriptions, we can evaluate the entanglement and Rényi entropies on various Riemann surfaces, some of which have been obtained before. The results will then serve in the next section for calculating the dissimilarity measures between reduced density matrices.

For a reduced density matrix  $\rho_A$ , the Rényi entropy is defined as

$$S_n = -\frac{1}{n-1} \log \text{tr}_A \rho_A^n, \quad (2.32)$$

and taking the  $n \rightarrow 1$  limit one can get the entanglement entropy

$$S = -\text{tr}_A (\rho_A \log \rho_A). \quad (2.33)$$

The Rényi entropy can be calculated from (2.12), and the entanglement entropy can be calculated from the  $n \rightarrow 1$  limit of the Rényi entropy or directly from (2.28).

We calculate the Rényi entropies and entanglement entropies for the seven Riemann surfaces in figure 1. The seven Rényi entropies are shown in figure 2. In practice we only need to calculate  $S_n(L, \phi)$  and  $S_n(L, q)$ , as marked in blue in the figure, and the other cases can be obtained easily from them. Note that most of the results in this section are not new, and just serves as a check of the OPE coefficients and the one-point functions.

Rényi entropy  $S_n(L, \phi)$  and entanglement entropy  $S(L, \phi)$  have been calculated in [10]<sup>1</sup>, and we will not repeat the results here. Since now at level 9 we have (2.11), the unknown terms  $O(\ell^9)$  in

<sup>1</sup>One can also follow different approach [36–38] to obtain Rényi entropy  $S_n(L, \phi)$  with finite size system in 2D rational CFTs.

$$\begin{array}{cccc}
S_n(L, \phi) & S_n(L, q) & S_n(\beta, \phi) & S_n(\beta, \rho) \\
S_n(L) & S_n(\emptyset) & S_n(\beta) & 
\end{array}$$

Figure 2: The seven Rényi entropies we can calculate using OPE of the twist operators. In practice, we only need to calculate  $S_n(L, \phi)$  and  $S_n(L, q)$ , as marked in blue, and the other cases can be obtained easily from them.

results of [10] are actually of order  $O(\ell^{10})$ . For the reduced density matrix  $\rho_A(\beta, \phi)$ , we have the Rényi entropy and entanglement entropy

$$S_n(\beta, \phi) = S_n(L, \phi)|_{L \rightarrow i\beta}, \quad S(\beta, \phi) = S(L, \phi)|_{L \rightarrow i\beta}. \quad (2.34)$$

For  $\rho_A(L, q)$ , the Rényi entropy and entanglement entropy have been calculated using OPE of the twist operators to order  $\ell^7$  in [17], and here we calculate the results to order  $\ell^9$ . In large  $c$  limit we write the Rényi entropy as the leading part, the next-to-leading part, the next-to-next-to-leading part, and etc,

$$S_n(L, q) = S_n^L(L, q) + S_n^{\text{NL}}(L, q) + S_n^{\text{NNL}}(L, q) + \dots, \quad (2.35)$$

and to order  $\ell^9$  only the first three parts are non-vanishing. Explicitly, we have the leading part

$$\begin{aligned}
S_n^L(L, q) = & \frac{c(n+1)}{12n} \log \frac{\ell}{\epsilon} - \frac{\pi^2 c(n+1)}{72n} \frac{\ell^2}{L^2} + \left( -\frac{\pi^4 c(n+1)}{2160n} - \frac{\pi^4 c(n-1)(n+1)^2 q^2}{18n^3} \right. \\
& - \frac{2\pi^4 c(n-1)(n+1)^2 q^3}{9n^3} - \frac{11\pi^4 c(n-1)(n+1)^2 q^4}{18n^3} + O(q^5) \Big) \frac{\ell^4}{L^4} \\
& + \left( -\frac{\pi^6 c(n+1)}{34020n} + \frac{\pi^6 c(n-1)(n+1)^2 q^2}{27n^3} + \frac{10\pi^6 c(n-1)(n+1)^2 q^3}{27n^3} \right. \\
& + \frac{\pi^6 c(n+1)(48n^4 - 49n^2 + 1)q^4}{27n^5} + O(q^5) \Big) \frac{\ell^6}{L^6} + \left( -\frac{\pi^8 c(n+1)}{453600n} \right. \\
& - \frac{\pi^8 c(n-1)(n+1)^2 q^2}{90n^3} - \frac{4\pi^8 c(n-1)(n+1)^2 q^3}{15n^3} \\
& \left. - \frac{\pi^8 c(n-1)(n+1)^2 (3727n^4 - 62n^2 - 11)q^4}{1620n^7} + O(q^5) \right) \frac{\ell^8}{L^8} + O(\ell^{10})
\end{aligned} \quad (2.36)$$

the next-to-leading part

$$\begin{aligned}
S_n^{\text{NL}}(L, q) = & \left( \frac{2\pi^2(n+1)q^2}{3n} + \frac{\pi^2(n+1)q^3}{n} + \frac{2\pi^2(n+1)q^4}{n} + O(q^5) \right) \frac{\ell^2}{L^2} \\
& + \left( -\frac{\pi^4(n+1)(9n^2 - 11)q^2}{45n^3} - \frac{\pi^4(n+1)(41n^2 - 44)q^3}{45n^3} \right. \\
& - \frac{\pi^4(n+1)(49n^2 - 51)q^4}{15n^3} + O(q^5) \Big) \frac{\ell^4}{L^4} + \left( \frac{2\pi^6(n+1)(17n^4 - 46n^2 + 31)q^2}{945n^5} \right. \\
& + \frac{\pi^6(n+1)(492n^4 - 1013n^2 + 527)q^3}{945n^5} + \frac{2\pi^6(n+1)(1654n^4 - 2903n^2 + 1255)q^4}{945n^5} \\
& \left. + O(q^5) \right) \frac{\ell^6}{L^6} + \left( -\frac{\pi^8(n+1)(62n^6 - 278n^4 + 415n^2 - 205)q^2}{14175n^7} \right. \\
& - \frac{\pi^8(n+1)(866n^6 - 2694n^4 + 2850n^2 - 1025)q^3}{4725n^7} \\
& \left. - \frac{\pi^8(n+1)(66439n^6 - 163681n^4 + 139223n^2 - 36257)q^4}{28350n^7} + O(q^5) \right) \frac{\ell^8}{L^8} + O(\ell^{10}),
\end{aligned} \quad (2.37)$$

and the next-to-next-to-leading part

$$\begin{aligned}
S_n^{\text{NNL}}(L, q) &= \left( -\frac{4\pi^4(n+1)(n^2+11)q^4}{45n^3c} + O(q^5) \right) \frac{\ell^4}{L^4} \\
&+ \left( \frac{4\pi^6(n+1)(26n^4+271n^2-345)q^4}{945n^5c} + O(q^5) \right) \frac{\ell^6}{L^6} \\
&+ \left( -\frac{8\pi^8(n+1)(116n^6+1141n^4-3017n^2+3398)q^4}{14175n^7c} + O(q^5) \right) \frac{\ell^8}{L^8} + O(\ell^{10}).
\end{aligned} \tag{2.38}$$

The leading and next-to-leading parts match the results in [39–41], which are calculated in another method. The  $\ell^8$  order of the next-to-next-to-leading part is a new result. Taking  $n \rightarrow 1$  limit we get the entanglement entropy

$$\begin{aligned}
S(L, q) &= \frac{c}{6} \log \frac{\ell}{\epsilon} + \left( -\frac{c\pi^2}{36} + \frac{4\pi^2q^2}{3} + 2\pi^2q^3 + 4\pi^2q^4 + O(q^5) \right) \frac{\ell^2}{L^2} \\
&+ \left( -\frac{c\pi^4}{1080} + \frac{4\pi^4q^2}{45} + \frac{2\pi^4q^3}{15} + \frac{4(c-8)\pi^4q^4}{15c} + O(q^5) \right) \frac{\ell^4}{L^4} \\
&+ \left( -\frac{c\pi^6}{17010} + \frac{8\pi^6q^2}{945} + \frac{4\pi^6q^3}{315} + \frac{8(c-16)\pi^6q^4}{315c} + O(q^5) \right) \frac{\ell^6}{L^6} \\
&+ \left( -\frac{c\pi^8}{226800} + \frac{4\pi^8q^2}{4725} + \frac{2\pi^8q^3}{1575} - \frac{4\pi^8(159c+728)q^4}{1575c} + O(q^5) \right) \frac{\ell^8}{L^8} + O(\ell^{10}).
\end{aligned} \tag{2.39}$$

The Rényi entropy and entanglement entropy for  $\rho_A(\beta, p)$  are just the modular transformation of those for  $\rho_A(L, q)$ , i.e.,

$$S_n(\beta, p) = S_n(L, q)|_{L \rightarrow i\beta, q \rightarrow p}, \quad S(\beta, p) = S(L, q)|_{L \rightarrow i\beta, q \rightarrow p}. \tag{2.40}$$

Without considering the subtlety of boundary conditions at the entangling surface [42, 43], the Rényi entropy and entanglement entropy for  $\rho_A(\emptyset)$ ,  $\rho_A(L)$  and  $\rho_A(\beta)$  are of universal forms and depend only on the central charge [12]

$$\begin{aligned}
S_n(\emptyset) &= \frac{c(n+1)}{12n} \log \frac{\ell}{\epsilon}, \quad S(\emptyset) = \frac{c}{6} \log \frac{\ell}{\epsilon}, \\
S_n(L) &= \frac{c(n+1)}{12n} \log \left( \frac{L}{\pi\epsilon} \sin \frac{\pi\ell}{L} \right), \quad S(L) = \frac{c}{6} \log \left( \frac{L}{\pi\epsilon} \sin \frac{\pi\ell}{L} \right), \\
S_n(\beta) &= \frac{c(n+1)}{12n} \log \left( \frac{\beta}{\pi\epsilon} \sinh \frac{\pi\ell}{\beta} \right), \quad S(\beta) = \frac{c}{6} \log \left( \frac{\beta}{\pi\epsilon} \sinh \frac{\pi\ell}{\beta} \right).
\end{aligned} \tag{2.41}$$

To order  $\ell^9$  the above results can be obtained easily as the limits and/or substitutes of  $S_n(L, q)$ ,  $S(L, q)$ .

### 3 Dissimilarities of reduced density matrices

In this section we evaluate various dissimilarity measures between reduced density matrices, which include relative entropy, Jensen-Shannon divergence, Schatten 2-norm and 4-norm.

#### 3.1 Relative entropy

The relative entropy is also called Kullback-Leibler divergence. For two reduced density matrices  $\rho_A$  and  $\sigma_A$ , the relative entropy is defined as

$$S(\rho_A \parallel \sigma_A) = \text{tr}_A(\rho_A \log \rho_A) - \text{tr}_A(\rho_A \log \sigma_A). \tag{3.1}$$

To calculate the relative entropy, one may first calculate the  $n$ -th relative entropy

$$S_n(\rho_A\|\sigma_A) = \frac{1}{n-1}[\log \text{tr}_A \rho_A^n - \log \text{tr}_A(\rho_A \sigma_A^{n-1})], \quad (3.2)$$

and then takes the  $n \rightarrow 1$  limit. The relative entropy is not symmetric for its two arguments, and one may define the symmetrized relative entropy

$$S(\rho_A, \sigma_A) = S(\rho_A\|\sigma_A) + S(\sigma_A\|\rho_A). \quad (3.3)$$

To calculate the symmetrized relative entropy, one can first calculate the  $n$ -th symmetrized relative entropy

$$S_n(\rho_A, \sigma_A) = S_n(\rho_A\|\sigma_A) + S_n(\sigma_A\|\rho_A), \quad (3.4)$$

and then takes the  $n \rightarrow 1$  limit. It turns out that

$$\begin{aligned} S(\rho_A\|\sigma_A) &= -a_{TT}(\langle T \rangle_{\mathcal{R}} - \langle T \rangle_{\mathcal{S}})^2 \ell^4 - a_{TTT}(\langle T \rangle_{\mathcal{R}} - \langle T \rangle_{\mathcal{S}})^2(\langle T \rangle_{\mathcal{R}} + 2\langle T \rangle_{\mathcal{S}}) \ell^6 \\ &\quad - [a_{\mathcal{A}\mathcal{A}}(\langle \mathcal{A} \rangle_{\mathcal{R}} - \langle \mathcal{A} \rangle_{\mathcal{S}})^2 + a_{TT\mathcal{A}}(\langle T \rangle_{\mathcal{R}} - \langle T \rangle_{\mathcal{S}})(\langle T \rangle_{\mathcal{R}} \langle \mathcal{A} \rangle_{\mathcal{R}} + \langle T \rangle_{\mathcal{S}} \langle \mathcal{A} \rangle_{\mathcal{R}} - 2\langle T \rangle_{\mathcal{S}} \langle \mathcal{A} \rangle_{\mathcal{S}}) \\ &\quad + a_{TTTT}(\langle \mathcal{A} \rangle_{\mathcal{R}} - \langle \mathcal{A} \rangle_{\mathcal{S}})^2(\langle T \rangle_{\mathcal{R}}^2 + 2\langle T \rangle_{\mathcal{R}} \langle T \rangle_{\mathcal{S}} - 3\langle T \rangle_{\mathcal{S}}^2)] \ell^8 + O(\ell^{10}), \end{aligned} \quad (3.5)$$

$$\begin{aligned} S(\rho_A, \sigma_A) &= -2a_{TT}(\langle T \rangle_{\mathcal{R}} - \langle T \rangle_{\mathcal{S}})^2 \ell^4 - 3a_{TTT}(\langle T \rangle_{\mathcal{R}} - \langle T \rangle_{\mathcal{S}})^2(\langle T \rangle_{\mathcal{R}} + \langle T \rangle_{\mathcal{S}}) \ell^6 \\ &\quad - [2a_{\mathcal{A}\mathcal{A}}(\langle \mathcal{A} \rangle_{\mathcal{R}} - \langle \mathcal{A} \rangle_{\mathcal{S}})^2 + a_{TT\mathcal{A}}(\langle T \rangle_{\mathcal{R}} - \langle T \rangle_{\mathcal{S}})(3\langle T \rangle_{\mathcal{R}} \langle \mathcal{A} \rangle_{\mathcal{R}} - \langle T \rangle_{\mathcal{R}} \langle \mathcal{A} \rangle_{\mathcal{S}} \\ &\quad + \langle T \rangle_{\mathcal{S}} \langle \mathcal{A} \rangle_{\mathcal{R}} - 3\langle T \rangle_{\mathcal{S}} \langle \mathcal{A} \rangle_{\mathcal{S}}) + 4a_{TTTT}(\langle \mathcal{A} \rangle_{\mathcal{R}} - \langle \mathcal{A} \rangle_{\mathcal{S}})^2(\langle T \rangle_{\mathcal{R}}^2 + \langle T \rangle_{\mathcal{R}} \langle T \rangle_{\mathcal{S}} + \langle T \rangle_{\mathcal{S}}^2)] \ell^8 \\ &\quad + O(\ell^{10}). \end{aligned} \quad (3.6)$$

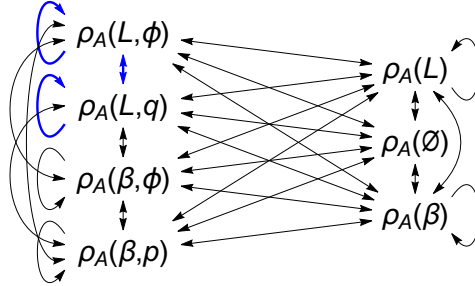


Figure 3: The 48 relative entropies we can calculate using OPE of the twist operators. By  $\rho_A \rightarrow \sigma_A$  we mean the relative entropy  $S(\rho_A\|\sigma_A)$ , and by  $\rho_A \leftrightarrow \sigma_A$  we mean the relative entropies  $S(\rho_A\|\sigma_A)$  and  $S(\sigma_A\|\rho_A)$ . Note that  $q = e^{-2\pi\beta/L}$  and  $p = e^{-2\pi L/\beta}$  depend on both  $L$  and  $\beta$ . In the figure  $\dots L \dots \rightarrow \dots L \dots$  actually means  $\dots L_1 \dots \rightarrow \dots L_2 \dots$  with generally  $L_1 \neq L_2$ ,  $\dots \phi \dots \rightarrow \dots \phi \dots$  means  $\dots \phi_1 \dots \rightarrow \dots \phi_2 \dots$ , and  $\dots \beta \dots \rightarrow \dots \beta \dots$  means  $\dots \beta_1 \dots \rightarrow \dots \beta_2 \dots$ . In practice, we only need to calculate the four relative entropies as marked in blue.

As shown in figure 3, we use OPE of twist operators as described in section 2 to calculate four

relative entropies. For  $\rho_A(L_1, \phi_1)$  and  $\rho_A(L_2, \phi_2)$  we have the relative entropy<sup>2</sup>

$$\begin{aligned}
S(\rho_A(L_1, \phi_1) \parallel \rho_A(L_2, \phi_2)) &= \frac{\pi^4 [L_1^2(c - 24h_{\phi_2}) - L_2^2(c - 24h_{\phi_1})]^2 \ell^4}{1080cL_1^4L_2^4} \\
&+ \frac{\pi^6 [L_1^2(c - 24h_{\phi_2}) - L_2^2(c - 24h_{\phi_1})]^2 [2L_1^2(c - 24h_{\phi_2}) + L_2^2(c - 24h_{\phi_1})] \ell^6}{17010c^2L_1^6L_2^6} \\
&+ \dots \ell^8 + O(\ell^{10}).
\end{aligned} \tag{3.7}$$

For the special case  $L_1 = L_2$  in (3.7), it matches the result in [10]. For  $\rho_A(L_1, \phi)$  and  $\rho_A(L_2, q)$  we have the relative entropy

$$\begin{aligned}
S(\rho_A(L_1, \phi) \parallel \rho_A(L_2, q)) &= \left[ \frac{\pi^4 [cL_1^2 - (c - 24h_\phi)L_2^2]^2}{1080cL_1^4L_2^4} + \frac{4\pi^4 [(c - 24h_\phi)L_2^2 - cL_1^2]q^2}{45cL_1^2L_2^4} \right. \\
&+ \frac{2\pi^4 [(c - 24h_\phi)L_2^2 - cL_1^2]q^3}{15cL_1^2L_2^4} + \frac{4\pi^4 [(c - 24h_\phi)L_2^2 - (c - 8)L_1^2]q^4}{15cL_1^2L_2^4} + O(q^5) \Big] \ell^4 \\
&+ \left[ \frac{\pi^6 [cL_1^2 - (c - 24h_\phi)L_2^2]^2 (2cL_1^2 + (c - 24h_\phi)L_2^2)}{17010c^2L_1^6L_2^6} + \frac{16\pi^6 [(c - 24h_\phi)L_2^2 - cL_1^2]q^2}{945cL_1^2L_2^6} \right. \\
&+ \frac{8\pi^6 [(c - 24h_\phi)L_2^2 - cL_1^2]q^3}{315cL_1^2L_2^6} + \frac{16\pi^6 [(c - 8)(c - 24h_\phi)L_2^2 - c(c - 16)L_1^2]q^4}{315c^2L_1^2L_2^6} + O(q^5) \Big] \ell^6 \\
&+ \dots \ell^8 + O(\ell^{10}).
\end{aligned} \tag{3.8}$$

For  $\rho_A(L_1, q)$  and  $\rho_A(L_2, \phi)$  we have the relative entropy

$$\begin{aligned}
S(\rho_A(L_1, q) \parallel \rho_A(L_2, \phi)) &= \left[ \frac{\pi^4 [(c - 24h_\phi)L_1^2 - cL_2^2]^2}{1080cL_1^4L_2^4} + \frac{4\pi^4 [(c - 24h_\phi)L_1^2 - cL_2^2]q^2}{45cL_1^4L_2^2} \right. \\
&+ \frac{2\pi^4 [(c - 24h_\phi)L_1^2 - cL_2^2]q^3}{15cL_1^4L_2^2} + \frac{4\pi^4 [(c - 24h_\phi)L_1^2 - (c - 8)L_2^2]q^4}{15cL_1^4L_2^2} + O(q^5) \Big] \ell^4 \\
&+ \left[ \frac{\pi^6 [(c - 24h_\phi)L_1^2 - cL_2^2]^2 [2(c - 24h_\phi)L_1^2 + cL_2^2]}{17010c^2L_1^6L_2^6} + \frac{8\pi^6 [(c - 24h_\phi)^2L_1^4 - c^2L_2^4]q^2}{945c^2L_1^6L_2^4} \right. \\
&+ \frac{4\pi^6 [(c - 24h_\phi)^2L_1^4 - c^2L_2^4]q^3}{315c^2L_1^6L_2^4} + \frac{8\pi^6 [(c - 24h_\phi)^2L_1^4 - (c - 16)cL_2^4]q^4}{315c^2L_1^6L_2^4} + O(q^5) \Big] \ell^6 \\
&+ \dots \ell^8 + O(\ell^{10}).
\end{aligned} \tag{3.9}$$

Note that

$$S(\rho_A(L_1, \phi) \parallel \rho_A(L_2, q)) \neq S(\rho_A(L_2, q) \parallel \rho_A(L_1, \phi)). \tag{3.10}$$

---

<sup>2</sup>Note that the  $\ell^8$  order in (3.7) is too complex and so we just omit it. One can see its explicit form in the Mathematica notebook in the attachment of the paper in arXiv. In the following part of the paper, all the missing terms represented by  $\dots$  can be found in the attached Mathematica notebook in arXiv.

For  $\rho_A(L_1, q_1)$  and  $\rho_A(L_2, q_2)$  we have the relative entropy

$$\begin{aligned}
S(\rho_A(L_1, q_1) \parallel \rho_A(L_2, q_2)) &= \left[ \frac{\pi^4 c (L_1^2 - L_2^2)^2}{1080 L_1^4 L_2^4} - \frac{4\pi^4 (L_1^2 - L_2^2)(L_1^2 q_2^2 - L_2^2 q_1^2)}{45 (L_1^4 L_2^4)} \right. \\
&\quad - \frac{2\pi^4 (L_1^2 - L_2^2)(L_1^2 q_2^3 - L_2^2 q_1^3)}{15 (L_1^4 L_2^4)} - \frac{4\pi^4 [(c-8)(L_1^4 q_2^4 + L_2^4 q_1^4) - L_1^2 L_2^2 (c q_1^4 - 16 q_2^2 q_1^2 + c q_2^4)]}{15 (c L_1^4 L_2^4)} \\
&\quad + O(q_1^5, q_2^5) \Big] \ell^4 + \left[ \frac{\pi^6 c (2L_1^6 - 3L_2^2 L_1^4 + L_2^6)}{17010 L_1^6 L_2^6} - \frac{8\pi^6 (L_1^2 - L_2^2) [2L_1^4 q_2^2 - L_2^2 (L_1^2 + L_2^2) q_1^2]}{945 L_1^6 L_2^6} \right. \\
&\quad + \frac{4\pi^6 [L_1^4 L_2^2 (q_1^3 + 2q_2^3) - 2L_1^6 q_2^3 - L_2^6 q_1^3]}{315 L_1^6 L_2^6} \\
&\quad - \left. \frac{8\pi^6 [(c-16)(2L_1^6 q_2^4 + L_2^6 q_1^4) - L_1^4 L_2^2 (c q_1^4 - 32 q_1^2 q_2^2 + 2(c-8) q_2^4)]}{315 c L_1^6 L_2^6} + O(q_1^5, q_2^5) \right] \ell^6 \\
&\quad + \dots \ell^8 + O(\ell^{10}).
\end{aligned} \tag{3.11}$$

For general  $n \neq 1$ , the  $n$ -th relative entropy  $S_n(\rho_A \parallel \sigma_A)$  and  $n$ -th symmetrized relative entropy  $S_n(\rho_A, \sigma_A)$  have no obvious physical meaning because they are not positive definite. However, the 2nd symmetrized relative entropy, which is defined as

$$S_2(\rho_A, \sigma_A) = \log \frac{(\text{tr}_A \rho_A^2)(\text{tr}_A \sigma_A^2)}{[\text{tr}_A(\rho_A \sigma_A)]^2}, \tag{3.12}$$

is positive definite and can be used to characterize the dissimilarity of  $\rho_A, \sigma_A$ . In fact, it is directly related to the overlap of the two reduced density matrices

$$\mathcal{F}(\rho_A, \sigma_A) = \frac{[\text{tr}_A(\rho_A \sigma_A)]^2}{(\text{tr}_A \rho_A^2)(\text{tr}_A \sigma_A^2)}. \tag{3.13}$$

As shown in figure 4, we calculate three symmetrized relative entropies using OPE of twist operators. Explicitly, we have

$$\begin{aligned}
S(\rho_A(L_1, \phi_1), \rho_A(L_2, \phi_2)) &= \frac{\pi^4 [L_1^2 (c - 24h_{\phi_2}) - L_2^2 (c - 24h_{\phi_1})]^2 \ell^4}{540 c L_1^4 L_2^4} \\
&\quad + \frac{\pi^6 [L_1^2 (c - 24h_{\phi_2}) - L_2^2 (c - 24h_{\phi_1})]^2 [L_1^2 (c - 24h_{\phi_2}) + L_2^2 (c - 24h_{\phi_1})] \ell^6}{5670 c^2 L_1^6 L_2^6} + \dots \ell^8 + O(\ell^{10}),
\end{aligned} \tag{3.14}$$

$$\begin{aligned}
S(\rho_A(L_1, \phi), \rho_A(L_2, q)) &= \left[ \frac{\pi^4 [c L_1^2 - (c - 24h_\phi) L_2^2]^2}{540 c L_1^4 L_2^4} + \frac{8\pi^4 [(c - 24h_\phi) L_2^2 - c L_1^2] q^2}{45 c L_1^2 L_2^4} \right. \\
&\quad + \frac{4\pi^4 [(c - 24h_\phi) L_2^2 - c L_1^2] q^3}{15 c L_1^2 L_2^4} + \frac{8\pi^4 [(c - 24h_\phi) L_2^2 - (c - 8) L_1^2] q^4}{15 c L_1^2 L_2^4} + O(q^5) \Big] \ell^4 \\
&\quad + \left[ \frac{\pi^6 [c L_1^2 - (c - 24h_\phi) L_2^2]^2 [c L_1^2 + (c - 24h_\phi) L_2^2]}{5670 c^2 L_1^6 L_2^6} \right. \\
&\quad - \frac{8\pi^6 [c L_1^2 - (c - 24h_\phi) L_2^2] [3c L_1^2 + (c - 24h_\phi) L_2^2] q^2}{945 c^2 L_1^4 L_2^6} \\
&\quad - \frac{4\pi^6 [c L_1^2 - (c - 24h_\phi) L_2^2] [3c L_1^2 + (c - 24h_\phi) L_2^2] q^3}{315 c^2 L_1^4 L_2^6} \\
&\quad - \left. \frac{8\pi^6 [3c(c-16) L_1^4 - 2(c-8)(c-24h_\phi) L_1^2 L_2^2 - (c-24h_\phi)^2 L_2^4] q^4}{315 c^2 L_1^4 L_2^6} + O(q^5) \right] \ell^6 \\
&\quad + \dots \ell^8 + O(\ell^{10}),
\end{aligned} \tag{3.15}$$



$$\begin{aligned}
S(\rho_A(L_1, q_1), \rho_A(L_2, q_2)) &= \left[ \frac{c\pi^4(L_1^2 - L_2^2)^2}{540L_1^4L_2^4} - \frac{8\pi^4(L_1^2 - L_2^2)(L_1^2q_2^2 - L_2^2q_1^2)}{45L_1^4L_2^4} \right. \\
&\quad - \frac{4\pi^4(L_1^2 - L_2^2)(L_1^2q_2^3 - L_2^2q_1^3)}{15L_1^4L_2^4} + \frac{8\pi^4[8(L_2^2q_1^2 - L_1^2q_2^2)^2 + c(L_1^2 - L_2^2)(L_2^2q_1^4 - L_1^2q_2^4)]}{15cL_1^4L_2^4} \\
&\quad \left. + O(q_1^5, q_2^5) \right] \ell^4 + \left[ \frac{c\pi^6(L_1^2 - L_2^2)^2(L_1^2 + L_2^2)}{5670L_1^6L_2^6} - \frac{8\pi^6(L_1^2 - L_2^2)[3q_2^2L_1^4 + L_2^2(q_2^2 - q_1^2)L_1^2 - 3L_2^4q_1^2]}{945L_1^6L_2^6} \right. \\
&\quad \left. - \frac{4\pi^6(L_1^2 - L_2^2)[3q_2^3L_1^4 + L_2^2(q_2^3 - q_1^3)L_1^2 - 3L_2^4q_1^3]}{315L_1^6L_2^6} + \dots + O(q_1^5, q_2^5) \right] \ell^6 + \dots \ell^8 + O(\ell^{10}).
\end{aligned} \tag{3.16}$$

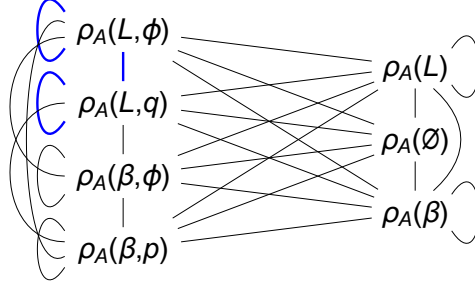


Figure 4: The 27 symmetrized relative entropies we can calculate using OPE of the twist operators. We only need to calculate the three ones marked in blue. This figure also applies to the 2nd symmetrized relative entropy, Jensen-Shannon divergence, as well as the Schatten 2-norm and 4-norm in the following subsections.

We have the 2nd symmetrized relative entropies

$$\begin{aligned}
S_2(\rho_A(L_1, \phi_1), \rho_A(L_2, \phi_2)) &= \frac{\pi^4[L_1^2(c - 24h_{\phi_2}) - L_2^2(c - 24h_{\phi_1})]^2 \ell^4}{4608cL_1^4L_2^4} \\
&\quad + \frac{\pi^6[L_1^2(c - 24h_{\phi_2}) - L_2^2(c - 24h_{\phi_1})][L_1^4(11c - 240h_{\phi_2}) - L_2^4(11c - 240h_{\phi_1})] \ell^6}{552960cL_1^6L_2^6} \\
&\quad + \dots \ell^8 + O(\ell^{10}),
\end{aligned} \tag{3.17}$$

$$\begin{aligned}
S_2(\rho_A(L_1, \phi), \rho_A(L_2, q)) &= \left[ \frac{\pi^4[cL_1^2 - (c - 24h_\phi)L_2^2]^2}{4608cL_1^4L_2^4} + \frac{\pi^4[(c - 24h_\phi)L_2^2 - cL_1^2]q^2}{48cL_1^2L_2^4} \right. \\
&\quad + \frac{\pi^4[(c - 24h_\phi)L_2^2 - cL_1^2]q^3}{32cL_1^2L_2^4} + \frac{\pi^4[(c - 24h_\phi)L_2^2 - (c - 8)L_1^2]q^4}{16cL_1^2L_2^4} + O(q^5) \left. \right] \ell^4 \\
&\quad + \left[ \frac{\pi^6[cL_1^2 - (c - 24h_\phi)L_2^2][11cL_1^4 + (240h_\phi - 11c)L_2^4]}{552960cL_1^6L_2^6} \right. \\
&\quad + \frac{\pi^6[3c(10c + 33)L_1^4 - 10(3c + 11)(c - 24h_\phi)L_2^2L_1^2 + (11c - 240h_\phi)L_2^4]q^2}{11520cL_1^4L_2^6} \\
&\quad + \frac{\pi^6[c(80c + 319)L_1^4 - 10(8c + 33)(c - 24h_\phi)L_2^2L_1^2 + (11c - 240h_\phi)L_2^4]q^3}{7680cL_1^4L_2^6} \\
&\quad + \frac{\pi^6[11(c(10c + 9) - 160)L_1^4 - 10(11c + 59)(c - 24h_\phi)L_2^2L_1^2 + (11c - 240h_\phi)L_2^4]q^4}{3840cL_1^4L_2^6} \\
&\quad \left. + O(q^5) \right] \ell^6 + \dots \ell^8 + O(\ell^{10}),
\end{aligned} \tag{3.18}$$

$$\begin{aligned}
S_2(\rho_A(L_1, q_1), \rho_A(L_2, q_2)) &= \left[ \frac{c\pi^4(L_1^2 - L_2^2)^2}{4608L_1^4L_2^4} - \frac{\pi^4(L_1^2 - L_2^2)(L_1^2q_2^2 - L_2^2q_1^2)}{48L_1^4L_2^4} \right. \\
&\quad - \frac{\pi^4(L_1^2 - L_2^2)(L_1^2q_2^3 - L_2^2q_1^3)}{32L_1^4L_2^4} + \frac{\pi^4[8(L_2^2q_1^2 - L_1^2q_2^2)^2 + c(L_1^2 - L_2^2)(L_2^2q_1^4 - L_1^2q_2^4)]}{16cL_1^4L_2^4} \\
&\quad + O(q_1^5, q_2^5) \Big] \ell^4 + \left[ \frac{11\pi^6c(L_1^2 - L_2^2)^2(L_1^2 + L_2^2)}{552960L_1^6L_2^6} \right. \\
&\quad + \frac{\pi^6(L_1^2 - L_2^2)[3(10c + 33)q_2^2L_1^4 + 11L_2^2(q_1^2 - q_2^2)L_1^2 - 3(10c + 33)L_2^4q_1^2]}{11520L_1^6L_2^6} \\
&\quad + \frac{\pi^6(L_1^2 - L_2^2)[(80c + 319)q_2^3L_1^4 + 11L_2^2(q_1^3 - q_2^3)L_1^2 - (80c + 319)L_2^4q_1^3]}{7680L_1^6L_2^6} \\
&\quad \left. + \dots + O(q_1^5, q_2^5) \Big] \ell^6 + \dots \ell^8 + O(\ell^{10}). \tag{3.19}
\end{aligned}$$

### 3.2 Jensen-Shannon divergence

The Jensen-Shannon divergence of two reduced density matrices  $\rho_A$  and  $\sigma_A$  are defined as

$$JS(\rho_A, \sigma_A) = S\left(\frac{\rho_A + \sigma_A}{2}\right) - \frac{1}{2}S(\rho_A) - \frac{1}{2}S(\sigma_A), \tag{3.20}$$

with  $S\left(\frac{\rho_A + \sigma_A}{2}\right)$ ,  $S(\rho_A)$ ,  $S(\sigma_A)$  being the von Neumann entropies. By definition

$$0 \leq JS(\rho_A, \sigma_A) \leq \log 2. \tag{3.21}$$

One can also define the Jensen-Shannon distance

$$\sqrt{JS(\rho_A, \sigma_A)}. \tag{3.22}$$

To calculate the Jensen-Shannon divergence, we first calculate the Jensen-Rényi divergence

$$JR_n(\rho_A, \sigma_A) = S_n\left(\frac{\rho_A + \sigma_A}{2}\right) - \frac{1}{2}S_n(\rho_A) - \frac{1}{2}S_n(\sigma_A), \tag{3.23}$$

with  $S_n\left(\frac{\rho_A + \sigma_A}{2}\right)$ ,  $S_n(\rho_A)$ ,  $S_n(\sigma_A)$  being the Rényi entropies, and then take the  $n \rightarrow 1$  limit. We then get

$$\begin{aligned}
JS(\rho_A, \sigma_A) &= -\frac{1}{4}a_{TT}(\langle T \rangle_{\mathcal{R}} - \langle T \rangle_{\mathcal{S}})^2 \ell^4 - \frac{3}{8}a_{TTT}(\langle T \rangle_{\mathcal{R}} - \langle T \rangle_{\mathcal{S}})^2(\langle T \rangle_{\mathcal{R}} + \langle T \rangle_{\mathcal{S}}) \ell^6 \\
&\quad - \left[ \frac{1}{4}a_{AA}(\langle \mathcal{A} \rangle_{\mathcal{R}} - \langle \mathcal{A} \rangle_{\mathcal{S}})^2 + \frac{1}{8}a_{TTA}(\langle T \rangle_{\mathcal{R}} - \langle T \rangle_{\mathcal{S}})(3\langle T \rangle_{\mathcal{R}}\langle \mathcal{A} \rangle_{\mathcal{R}} - \langle T \rangle_{\mathcal{R}}\langle \mathcal{A} \rangle_{\mathcal{S}} \right. \\
&\quad + \langle T \rangle_{\mathcal{S}}\langle \mathcal{A} \rangle_{\mathcal{R}} - 3\langle T \rangle_{\mathcal{S}}\langle \mathcal{A} \rangle_{\mathcal{S}} + \frac{1}{16}a_{TTTT}(\langle \mathcal{A} \rangle_{\mathcal{R}} - \langle \mathcal{A} \rangle_{\mathcal{S}})^2(7\langle T \rangle_{\mathcal{R}}^2 + 10\langle T \rangle_{\mathcal{R}}\langle T \rangle_{\mathcal{S}} + 7\langle T \rangle_{\mathcal{S}}^2) \Big] \ell^8 \\
&\quad + O(\ell^{10}). \tag{3.24}
\end{aligned}$$

Explicitly, we have

$$\begin{aligned}
JS(\rho_A(L_1, \phi_1), \rho_A(L_2, \phi_2)) &= \frac{\pi^4[L_1^2(c - 24h_{\phi_2}) - L_2^2(c - 24h_{\phi_1})]^2 \ell^4}{4320cL_1^4L_2^4} \\
&\quad + \frac{\pi^6[L_1^2(c - 24h_{\phi_2}) - L_2^2(c - 24h_{\phi_1})]^2[L_1^2(c - 24h_{\phi_2}) + L_2^2(c - 24h_{\phi_1})] \ell^6}{45360c^2L_1^6L_2^6} \\
&\quad + \dots \ell^8 + O(\ell^{10}), \tag{3.25}
\end{aligned}$$

$$\begin{aligned}
JS(\rho_A(L_1, \phi), \rho_A(L_2, q)) &= \left[ \frac{\pi^4 [cL_1^2 - (c - 24h_\phi)L_2^2]^2}{4320cL_1^4L_2^4} + \frac{\pi^4 [(c - 24h_\phi)L_2^2 - cL_1^2]q^2}{45cL_1^2L_2^4} \right. \\
&+ \frac{\pi^4 [(c - 24h_\phi)L_2^2 - cL_1^2]q^3}{30cL_1^2L_2^4} + \frac{\pi^4 [(c - 24h_\phi)L_2^2 - (c - 8)L_1^2]q^4}{15cL_1^2L_2^4} + O(q^5) \left. \right] \ell^4 \\
&+ \left[ \frac{\pi^6 [cL_1^2 - (c - 24h_\phi)L_2^2]^2 [cL_1^2 + (c - 24h_\phi)L_2^2]}{45360c^2L_1^6L_2^6} \right. \\
&- \frac{\pi^6 [cL_1^2 - (c - 24h_\phi)L_2^2] [3cL_1^2 + (c - 24h_\phi)L_2^2] q^2}{945c^2L_1^4L_2^6} \\
&- \frac{\pi^6 [cL_1^2 - (c - 24h_\phi)L_2^2] [3cL_1^2 + (c - 24h_\phi)L_2^2] q^3}{630c^2L_1^4L_2^6} \\
&- \frac{\pi^6 [3(c - 16)cL_1^4 - 2(c - 8)(c - 24h_\phi)L_1^2L_2^2 - (c - 24h_\phi)^2L_2^4] q^4}{315c^2L_1^4L_2^6} \\
&\left. + O(q^5) \right] \ell^6 + \dots \ell^8 + O(\ell^{10}), \tag{3.26}
\end{aligned}$$

$$\begin{aligned}
JS(\rho_A(L_1, q_1), \rho_A(L_2, q_2)) &= \left[ \frac{c\pi^4 (L_1^2 - L_2^2)^2}{4320L_1^4L_2^4} - \frac{\pi^4 (L_1^2 - L_2^2)(L_1^2q_2^2 - L_2^2q_1^2)x^2}{45L_1^4L_2^4} \right. \\
&- \frac{\pi^4 (L_1^2 - L_2^2)(L_1^2q_2^3 - L_2^2q_1^3)}{30L_1^4L_2^4} + \frac{\pi^4 [8(L_2^2q_1^2 - L_1^2q_2^2)^2 + c(L_1^2 - L_2^2)(L_2^2q_1^4 - L_1^2q_2^4)]}{15cL_1^4L_2^4} \\
&+ O(q_1^5, q_2^5) \left. \right] \ell^4 + \left[ \frac{c\pi^6 (L_1^2 - L_2^2)^2 (L_1^2 + L_2^2)}{45360L_1^6L_2^6} - \frac{\pi^6 (L_1^2 - L_2^2)(3q_2^2L_1^4 + L_2^2(q_2^2 - q_1^2)L_1^2 - 3L_2^4q_1^2)}{945L_1^6L_2^6} \right. \\
&- \frac{\pi^6 (L_1^2 - L_2^2) [3q_2^3L_1^4 + L_2^2(q_2^3 - q_1^3)L_1^2 - 3L_2^4q_1^3]}{630L_1^6L_2^6} + \dots + O(q_1^5, q_2^5) \left. \right] \ell^6 + \dots \ell^8 + O(\ell^{10}). \tag{3.27}
\end{aligned}$$

### 3.3 Schatten 2-norm and 4-norm

For a general matrix  $\rho$ , the Schatten  $n$ -norm is defined as

$$\|\rho\|_n = (\text{tr}|\rho|^n)^{1/n}, \tag{3.28}$$

with  $|\rho| = \sqrt{\rho^\dagger \rho}$ . For  $n = 1$  it is just the trace norm, and for  $n = 2$  it is just the Hilbert-Schmidt norm. For two reduced density matrices  $\rho_A, \sigma_A$ , we just calculate

$$\|\rho_A - \sigma_A\|_n^n = \text{tr}_A |\rho_A - \sigma_A|^n. \tag{3.29}$$

For  $n = 1$  it is just the trace distance, and for  $n = 2$  it is just trace square. Since the reduced density matrices are hermitian, when  $n := 2p$  is an even integer we have a simpler expression

$$\|\rho_A - \sigma_A\|_{2p}^{2p} = \text{tr}_A (\rho_A - \sigma_A)^{2p}. \tag{3.30}$$

In this paper we use (2.24) to calculate the Schatten 2-norms and 4-norms. We have

$$\begin{aligned}
\|\rho_A(L_1, \phi_1) - \rho_A(L_2, \phi_2)\|_2^2 &= \left( \frac{\ell}{\epsilon} \right)^{-\frac{\epsilon}{8}} \left\{ \frac{\pi^4 (c + 2)\ell^4 [L_1^2(c - 24h_{\phi_2}) - L_2^2(c - 24h_{\phi_1})]^2}{9216cL_1^4L_2^4} \right. \\
&+ \frac{\pi^6 (c + 4)\ell^6}{4423680cL_1^6L_2^6} [L_1^2(c - 24h_{\phi_2}) - L_2^2(c - 24h_{\phi_1})] \{ L_1^4 [c(5c + 22) - 240h_{\phi_2}(c - 12h_{\phi_2} + 2)] \\
&- L_2^4 [c(5c + 22) - 240h_{\phi_1}(c - 12h_{\phi_1} + 2)] \} + \dots \ell^8 + O(\ell^{10}) \left. \right\}, \tag{3.31}
\end{aligned}$$

$$\begin{aligned}
\|\rho_A(L_1, \phi) - \rho_A(L_2, q)\|_2^2 &= \left(\frac{\ell}{\epsilon}\right)^{-\frac{5c}{8}} \left\{ \pi^4(c+2) \left\{ \frac{[cL_1^2 - (c-24h_\phi)L_2^2]^2}{9216cL_1^4L_2^4} \right. \right. \\
&+ \frac{[(c-24h_\phi)L_2^2 - cL_1^2]q^2}{96cL_1^2L_2^4} + \frac{[(c-24h_\phi)L_2^2 - cL_1^2]q^3}{64cL_1^2L_2^4} + \frac{[(c-24h_\phi)L_2^2 - (c-8)L_1^2]q^4}{32cL_1^2L_2^4} \\
&+ O(q^5) \left. \right\} \ell^4 + \pi^6(c+4) \left\{ \frac{[cL_1^2 - (c-24h_\phi)L_2^2]}{4423680cL_1^6L_2^6} \{c(5c+22)L_1^4 \right. \\
&+ [240(c-12h_\phi+2)h_\phi - c(5c+22)]L_2^4 \left. \right\} + \frac{q^2}{92160cL_1^4L_2^6} \{9c(5c+22)L_1^4 \\
&- 10(5c+22)(c-24h_\phi)L_1^2L_2^2 + [c(5c+22) - 240(c-12h_\phi+2)h_\phi]L_2^4 \left. \right\} \\
&+ \frac{q^3}{61440cL_1^4L_2^6} \{29c(5c+22)L_1^4 - 30(5c+22)(c-24h_\phi)L_2^2L_1^2 \\
&+ [c(5c+22) - 240(c-12h_\phi+2)h_\phi]L_2^4 \left. \right\} + \dots q^4 + O(q^5) \left. \right\} \ell^6 + \dots \ell^8 + O(\ell^{10}), \tag{3.32}
\end{aligned}$$

$$\begin{aligned}
\|\rho_A(L_1, q_1) - \rho_A(L_2, q_2)\|_2^2 &= \left(\frac{\ell}{\epsilon}\right)^{-\frac{5c}{8}} \left\{ \pi^4(c+2) \left[ \frac{c(L_1^2 - L_2^2)^2}{9216L_1^4L_2^4} \right. \right. \\
&- \frac{(L_1^2 - L_2^2)(L_1^2q_2^2 - L_2^2q_1^2)}{96L_1^4L_2^4} - \frac{(L_1^2 - L_2^2)(L_1^2q_2^3 - L_2^2q_1^3)}{64L_1^4L_2^4} \\
&- \frac{(c-8)(L_1^4q_2^4 + L_2^4q_1^4) - L_1^2L_2^2(cq_1^4 - 16q_2^2q_1^2 + cq_2^4)}{32cL_1^4L_2^4} + O(q_1^5, q_2^5) \left. \right] \ell^4 \\
&+ \pi^6c(c+4)(5c+22) \left[ \frac{L_1^4 - L_2^4}{4423680L_1^6L_2^6} + \frac{9L_1^4q_2^2 + L_1^2L_2^2(q_1^2 - q_2^2) - 9L_2^4q_1^2}{92160L_1^6L_2^6} \right. \\
&+ \left. \frac{29L_1^4q_2^3 + L_1^2L_2^2(q_1^3 - q_2^3) - 29L_2^4q_1^3}{61440L_1^6L_2^6} + \dots + O(q_1^5, q_2^5) \right] \ell^6 + \dots \ell^8 + O(\ell^{10}) \left. \right\}. \tag{3.33}
\end{aligned}$$

We also have

$$\begin{aligned}
\|\rho_A(L_1, \phi_1) - \rho_A(L_2, \phi_2)\|_4^4 &= \left(\frac{\ell}{\epsilon}\right)^{-\frac{5c}{16}} \left[ \frac{\pi^8(c+2)[25c(25c-14) + 192]\ell^8}{21743271936c^3L_1^8L_2^8} [L_1^2(c-24h_{\phi_2}) \right. \\
&- \left. L_2^2(c-24h_{\phi_1})]^4 + O(\ell^{10}) \right], \tag{3.34}
\end{aligned}$$

$$\begin{aligned}
\|\rho_A(L_1, \phi) - \rho_A(L_2, q)\|_4^4 &= \left(\frac{\ell}{\epsilon}\right)^{-\frac{5c}{16}} \left\{ \pi^8(c+2)[25c(25c-14) + 192] \left[ \frac{[cL_1^2 - L_2^2(c-24h_\phi)]^4}{21743271936c^3L_1^8L_2^8} \right. \right. \\
&- \frac{[cL_1^2 - L_2^2(c-24h_\phi)]^3q^2}{113246208c^3L_1^6L_2^8} - \frac{[cL_1^2 - L_2^2(c-24h_\phi)]^3q^3}{75497472c^3L_1^6L_2^8} \\
&- \left. \frac{[(c-24)L_1^2 - L_2^2(c-24h_\phi)][cL_1^2 - L_2^2(c-24h_\phi)]^2q^4}{37748736c^3L_1^6L_2^8} + O(q^5) \right] \ell^8 + O(\ell^{10}) \left. \right\}, \tag{3.35}
\end{aligned}$$

$$\begin{aligned}
\|\rho_A(L_1, q_1) - \rho_A(L_2, q_2)\|_4^4 &= \left(\frac{\ell}{\epsilon}\right)^{-\frac{5c}{16}} \left\{ \pi^8(c+2)[25c(25c-14) + 192] \left[ \frac{c(L_1^2 - L_2^2)^4}{21743271936L_1^8L_2^8} \right. \right. \\
&- \frac{(L_1^2 - L_2^2)^3(L_1^2q_2^2 - L_2^2q_1^2)}{113246208L_1^8L_2^8} - \frac{(L_1^2 - L_2^2)^3(L_1^2q_2^3 - L_2^2q_1^3)}{75497472L_1^8L_2^8} \\
&- \frac{(L_1^2 - L_2^2)^2[(c-24)(L_1^4q_2^4 + L_2^4q_1^4) - L_1^2L_2^2(cq_1^4 + cq_2^4 - 48q_2^2q_1^2)]}{37748736cL_1^8L_2^8} \\
&+ \left. O(q_1^5, q_2^5) \right] \ell^8 + O(\ell^{10}) \left. \right\}. \tag{3.36}
\end{aligned}$$

## 4 ETH for canonical ensemble and GGE thermal states

As a check of ETH for canonical ensemble thermal state, we calculate various quantities to characterize the dissimilarity of the reduced density matrix  $\rho_A(L, \phi)$  for the excited state and  $\rho_A(\beta)$  for the thermal state. Note that ETH is for comparing a highly excited state and a high temperature state, so that we use  $\rho_A(\beta)$  to approximate  $\rho_A(\beta, p)$ . The excited state  $|\phi\rangle$  is heavy and we write the conformal weight as

$$h_\phi = c\epsilon_\phi, \quad (4.1)$$

and by requiring

$$\langle T \rangle_{\mathcal{R}(L, \phi)} = \langle T \rangle_{\mathcal{R}(\beta)}, \quad (4.2)$$

we get the identification [5, 6]

$$\beta = \frac{L}{\sqrt{24c\epsilon_\phi - 1}}. \quad (4.3)$$

We have the difference of Rényi entropy

$$\begin{aligned} S_n(L, \phi) - S_n(\beta) &= \frac{\pi^4 c(n-1)(n+1)^2 \epsilon_\phi (22\epsilon_\phi - 1) \ell^4}{90n^3 L^4} \\ &- \frac{\pi^6 c(n-1)(n+1)^2 \epsilon_\phi [8(145n^2 + 188)\epsilon_\phi^2 - 3(46n^2 + 37)\epsilon_\phi + 4n^2 + 2] \ell^6}{2835n^5 L^6} + \dots \ell^8 + O(\ell^{10}), \end{aligned} \quad (4.4)$$

and it has been calculated in [3, 9, 10]. The difference of entanglement entropy is

$$S(L, \phi) - S(\beta) = -\frac{128\pi^8 c\epsilon_\phi^2 (22\epsilon_\phi - 1)^2 \ell^8}{1575(5c + 22)L^8} + O(\ell^{10}), \quad (4.5)$$

and it has been calculated in [10]. We have the relative entropies

$$\begin{aligned} S(\rho_A(L, \phi) \parallel \rho_A(\beta)) &= \frac{128\pi^8 c\epsilon_\phi^2 (22\epsilon_\phi - 1)^2 \ell^8}{1575(5c + 22)L^8} + O(\ell^{10}), \\ S(\rho_A(\beta) \parallel \rho_A(L, \phi)) &= \frac{128\pi^8 c\epsilon_\phi^2 (22\epsilon_\phi - 1)^2 \ell^8}{1575(5c + 22)L^8} + O(\ell^{10}), \end{aligned} \quad (4.6)$$

and the first one has been calculated in [10] by a different method. Note that  $S(\rho_A(L, \phi) \parallel \rho_A(\beta))$  and  $S(\rho_A(\beta) \parallel \rho_A(L, \phi))$  happen to be the same at order  $\ell^8$ , and we expect they will be different at higher orders. We have the symmetrized relative entropy and the 2nd symmetrized relative entropy

$$\begin{aligned} S(\rho_A(L, \phi), \rho_A(\beta)) &= \frac{256\pi^8 c\epsilon_\phi^2 (22\epsilon_\phi - 1)^2 \ell^8}{1575(5c + 22)L^8} + O(\ell^{10}), \\ S_2(\rho_A(L, \phi), \rho_A(\beta)) &= \frac{\pi^8 c(5c + 27)\epsilon_\phi^2 (22\epsilon_\phi - 1)^2 \ell^8}{3200(5c + 22)L^8} + O(\ell^{10}). \end{aligned} \quad (4.7)$$

The Jensen-Rényi divergence and Jensen-Shannon divergence are respectively

$$\begin{aligned} JR_n(\rho_A(L, \phi), \rho_A(\beta)) &= -\frac{\pi^8 (n+1)c\epsilon_\phi^2 (22\epsilon_\phi - 1)^2 \ell^8}{2268000(5c + 22)n^7 L^8} [175c^2(n^2 - 1)^3 + 70c(7n^2 - 55)(n^2 - 1)^2 \\ &- 8(n^2 + 11)(157n^4 - 298n^2 + 381)] + O(\ell^{10}), \\ JS(\rho_A(L, \phi), \rho_A(\beta)) &= \frac{32\pi^8 c(22\epsilon_\phi - 1)^2 \epsilon_\phi^2 \ell^8}{1575(5c + 22)L^8} + O(\ell^{10}). \end{aligned} \quad (4.8)$$

We also have

$$\|\rho_A(L, \phi) - \rho_A(\beta)\|_2^2 = \left(\frac{\ell}{\epsilon}\right)^{-\frac{c}{8}} \left[ \frac{\pi^8 c [c(5c + 62) + 216] \epsilon_\phi^2 (22\epsilon_\phi - 1)^2 \ell^8}{25600(5c + 22)L^8} + O(\ell^{10}) \right]. \quad (4.9)$$

All the above dissimilarities between the excited and thermal state originate from the fact that the level 4 operator  $\mathcal{A}$  has different expectation values [9, 10]

$$\langle \mathcal{A} \rangle_{\mathcal{R}(L, \phi)} \neq \langle \mathcal{A} \rangle_{\mathcal{R}(\beta)}. \quad (4.10)$$

A more refined consideration is that one should not compare the excited state and the canonical ensemble thermal state, instead one need to consider the generalized Gibbs ensemble (GGE) thermal state [20–25]. The GGE state has the density matrix

$$\rho_{\text{GGE}} = e^{-\beta H - \sum_i \beta_i J_i}, \quad (4.11)$$

with  $J_i$  being some conserved charges and  $\beta_i$  being the corresponding chemical potentials. By requiring the ETH comparison is done for the same macroscopic super-selection sector, we should impose

$$\langle H \rangle_{\mathcal{R}(L, \phi)} = \langle H \rangle_{\text{GGE}}, \quad \langle J_i \rangle_{\mathcal{R}(L, \phi)} = \langle J_i \rangle_{\text{GGE}}, \quad (4.12)$$

so that one can get the relation of  $h_\phi$  with the GGE parameters  $\beta, \mu_i$ . In the vacuum conformal family, there are an infinite number of commuting conserved charges  $I_{2k+1}$  with  $k = 0, 1, \dots$  [44, 45]. For examples, one has

$$\begin{aligned} I_1 &= -\frac{1}{2\pi} \int_{-L/2}^{L/2} dw T(w) = \frac{2\pi}{L} \left( L_0 - \frac{c}{24} \right) = H, \\ I_3 &= \frac{1}{2\pi} \int_{-L/2}^{L/2} dw \mathcal{A}(w) = \left( \frac{2\pi}{L} \right)^3 \left[ \mathcal{A}_0 - \frac{5c + 22}{60} \left( L_0 - \frac{c}{48} \right) \right]. \end{aligned} \quad (4.13)$$

We may choose the GGE state

$$\rho_{\text{GGE}} = e^{-\beta H - \sum_{k=1}^{\infty} \beta_{2k+1} I_{2k+1}}. \quad (4.14)$$

Then we have the requirement

$$\langle \mathcal{X} \rangle_{\mathcal{R}(L, \phi)} = \langle \mathcal{X} \rangle_{\text{GGE}}, \quad (4.15)$$

for all vacuum conformal family quasiprimary operator  $\mathcal{X}$ . Since there are more equations than the unknown chemical potentials, we do not know if there is a unique solution for all  $\beta, \beta_{2k+1}, k = 1, 2, \dots$ . If this is the case, all the dissimilarities considered in this paper vanish so that there is no difference between the reduced density matrices of the excited state and GGE thermal state.

Furthermore, in GGE it is not necessarily that all the conserved charges commute with each other [22]. For each nonidentity quasiprimary operator in vacuum conformal family, say  $\mathcal{X}$ , we may define a conserved charge

$$I_{\mathcal{X}} \propto \int_{-L/2}^{L/2} dw \mathcal{X}(w). \quad (4.16)$$

Then we may define the GGE state

$$\rho_{\text{GGE}} = e^{-\sum_{\mathcal{X}} \beta_{\mathcal{X}} I_{\mathcal{X}}}, \quad (4.17)$$

with which there are the same number of equations and the unknown chemical potentials. However, we still do not know if there is any solution to the equations (4.15).

To be more concrete, we consider a toy model of GGE

$$\rho_{\text{GGE}} = e^{-\beta H - \frac{\mu}{2\pi} \int_{-L/2}^{L/2} dw \mathcal{A}(w)}. \quad (4.18)$$

For an arbitrary operator  $\mathcal{X}$  we have

$$\frac{\text{tr}(\mathcal{X}(w_0)\rho_{\text{GGE}})}{\text{tr}e^{-\beta H}} = \left\langle \mathcal{X}(w_0) e^{-\frac{\mu}{2\pi} \int_{-L/2}^{L/2} dw \mathcal{A}(w)} \right\rangle_{\mathcal{R}(\beta,p)} \approx \left\langle \mathcal{X}(w_0) e^{-\frac{\mu}{2\pi} \int_{-\infty}^{\infty} dw \mathcal{A}(w)} \right\rangle_{\mathcal{R}(\beta)}. \quad (4.19)$$

We get the expectation value of GGE in expansion of the small chemical potential  $\mu$

$$\begin{aligned} \langle \mathcal{X}(w_0) \rangle_{\text{GGE}} &= \frac{\text{tr}(\mathcal{X}(w_0)\rho_{\text{GGE}})}{\text{tr}\rho_{\text{GGE}}} \approx \langle \mathcal{X}(w_0) \rangle_{\mathcal{R}(\beta)} - \frac{\mu}{2\pi} \int_{-\infty}^{\infty} dw [\langle \mathcal{A}(w)\mathcal{X}(w_0) \rangle_{\mathcal{R}(\beta)} \\ &\quad - \langle \mathcal{A}(w) \rangle_{\mathcal{R}(\beta)} \langle \mathcal{X}(w_0) \rangle_{\mathcal{R}(\beta)}] + O(\mu^2). \end{aligned} \quad (4.20)$$

The correlation functions on the cylinder  $\mathcal{R}(\beta)$  can be calculated by mapping the cylinder to a complex plane by the conformal transformation  $z = e^{\frac{2\pi w}{\beta}}$ . Note that the above expectation value should be independent of the position  $w_0$ . Using the integral<sup>3</sup>

$$\int_0^{\infty} \frac{dx}{\sinh^S x} = \frac{\Gamma(\frac{S}{2})\Gamma(\frac{1-S}{2})}{2\sqrt{\pi}}, \quad (4.21)$$

with  $S = 4$  and  $S = 8$ , we finally get

$$\begin{aligned} \langle T \rangle_{\text{GGE}} &= -\frac{\pi^2 c}{6\beta^2} + \frac{\pi^4 c(5c + 22)\mu}{45\beta^5} + O(\mu^2), \\ \langle \mathcal{A} \rangle_{\text{GGE}} &= \frac{\pi^4 c(5c + 22)}{180\beta^4} - \frac{\pi^6 c(5c + 22)(7c + 74)\mu}{945\beta^7} + O(\mu^2). \end{aligned} \quad (4.22)$$

In the excited state  $|\phi\rangle$  of a holomorphic primary operator  $\phi$  with conformal weight  $h_\phi = c\epsilon_\phi$ , there are expectation values [9, 10]

$$\langle T \rangle_\phi = \frac{\pi^2 c(1 - 24\epsilon_\phi)}{6L^2}, \quad \langle \mathcal{A} \rangle_\phi = \frac{\pi^4 c((5c + 22) - 240(c + 2)\epsilon_\phi + 2880c\epsilon_\phi^2)}{180L^4}. \quad (4.23)$$

To consider ETH comparison for the same super-selection sector, we equate (4.22) and (4.23)

$$\langle T \rangle_{\text{GGE}} = \langle T \rangle_\phi, \quad \langle \mathcal{A} \rangle_{\text{GGE}} = \langle \mathcal{A} \rangle_\phi, \quad (4.24)$$

and solve the inverse temperature  $\beta$  and chemical potential  $\mu$  in terms of  $\epsilon_\phi$ ,  $c$ ,  $L$ . As known that the ETH for canonical ensemble works well in the leading order of large  $c$  limit [7, 8], we should then expect

$$\lim_{c \rightarrow \infty} \beta = \frac{L}{\sqrt{24\epsilon_\phi - 1}}, \quad \lim_{c \rightarrow \infty} \mu = 0. \quad (4.25)$$

On the other hand, the finite  $c$  correction causes the mismatch between excited state and the canonical thermal state by power suppression of  $1/c$  [10], we then need to find the solution of (4.24) for GGE

<sup>3</sup>Note that the integral is only convergent for  $0 < \text{Re}S < 1$ , and it is analytically continued to other values of  $S$ . The results are the same as these from more delicate calculations in [46–48].

with power correction of  $1/c$  to (4.25) as follows. To make the  $1/c$  expansions in (4.22) well-defined, we need the leading order  $\mu \sim 1/c^\alpha$  with  $\alpha > 1$ . Since there is no subleading term in  $\langle T \rangle_\phi$ , we need the leading order correction to  $\beta$  of order  $1/c^{\alpha-1}$ . We then make the following ansatz for the solution to equations (4.24)

$$\begin{aligned}\beta &= \frac{L}{\sqrt{24\epsilon_\phi - 1}} + \frac{aL}{c^{\alpha-1}} + o\left(\frac{1}{c^{\alpha-1}}\right), \\ \mu &= \frac{bL^3}{c^\alpha} + o\left(\frac{1}{c^\alpha}\right),\end{aligned}\tag{4.26}$$

with the constants  $\alpha, a, b$  to be determined. It is easy to see that  $\langle \mathcal{A} \rangle_{\text{GGE}} = \langle \mathcal{A} \rangle_\phi$  cannot be satisfied for  $\alpha \geq 2$ . Thus, we have  $1 < \alpha < 2$  in ansatz (4.26). However, we cannot determine the coefficient  $b$  in ansatz (4.26) at the present expansion order of (4.22), but might be determined uniquely at the higher expansion orders.

## 5 ETH for microcanonical ensemble thermal state

The ETH [1, 2] and its corollaries such as the subsystem ETH [3, 4] are originally considered for comparing the energy eigenstate and the microcanonical (ensemble) thermal state. Despite that the difference between canonical and microcanonical thermal states is power-law negligible in the limit of large number of degrees of freedom, it is still interesting to check ETH directly for microcanonical thermal state. In this appendix we will do this using OPE of twist operators as described in section 2.

The microcanonical thermal state to be considered is the equal-weight sum of the pure states  $|\phi_i\rangle := \phi_i|0\rangle$   $i = 1, 2, \dots, \Omega$ , i.e., its density matrix is given by

$$\rho_{\text{me}} = \frac{1}{\Omega} \sum_{i=1}^{\Omega} |\phi_i\rangle\langle\phi_i| \tag{5.1}$$

where  $\phi_i$ 's are nonidentity primary operators of conformal weights  $(h_{\phi_i}, \bar{h}_{\phi_i})$ . For the microcanonical thermal states, we should require for all  $i = 1, 2, \dots, \Omega$

$$h_{\phi_i} \simeq h_\phi, \quad \bar{h}_{\phi_i} \simeq \bar{h}_\phi \tag{5.2}$$

where  $(h_\phi, \bar{h}_\phi)$  is the conformal weight of the excited state  $\phi$  with which we will compare for checking ETH.

For simplicity, we can choose an orthonormal set of  $\phi_i$ 's, i.e.,

$$\langle\phi_i|\phi_j\rangle = \delta_{ij}. \tag{5.3}$$

We also choose  $\phi$  as one of the  $\Omega$  operators  $\phi_i$ , i.e.,  $\phi \in \{\phi_i\}$ .

Globally, the pure excited state density matrix  $\rho_\phi = |\phi\rangle\langle\phi|$  and the microcanonical thermal state density matrix  $\rho_{\text{me}} = \frac{1}{\Omega} \sum_{i=1}^{\Omega} \rho_{\phi_i}$  are very different. This can be seen from various dissimilarity measures, i.e., starting from their von-Neumann entropies,

$$S(\rho_\phi) = 0, \quad S(\rho_{\text{me}}) = \log \Omega, \tag{5.4}$$



and then the relative entropy

$$S(\rho_\phi \parallel \rho_{\text{me}}) = \log \Omega, \quad (5.5)$$

and the Jensen-Shannon divergence

$$JS(\rho_\phi, \rho_{\text{me}}) = \log 2 + \frac{1}{2} \log \Omega - \frac{\Omega + 1}{2\Omega} \log(\Omega + 1). \quad (5.6)$$

Instead, the ETH should be explored by the local observables. If ETH holds, for arbitrary local observable  $\mathcal{X}$  we should have

$$\text{tr}(\rho_\phi \mathcal{X}) \simeq \text{tr}(\rho_{\text{me}} \mathcal{X}). \quad (5.7)$$

If  $\mathcal{X}$  is the operator in the vacuum conformal family, it is easy to see that (5.7) holds by the fact (5.2). On the other hand, if  $\mathcal{X}$  is some nonidentity primary operator or its descendants, then the ETH imposes constraints on OPE coefficients  $C_{\phi_i \phi_i \mathcal{X}}$ :

$$C_{\phi \phi \mathcal{X}} \simeq \frac{1}{\Omega} \sum_{i=1}^{\Omega} C_{\phi_i \phi_i \mathcal{X}}. \quad (5.8)$$

This implies that not every CFT satisfies ETH.

However, in a large  $c$  CFT, it is often a good approximation to consider contributions only from the vacuum conformal family, and this is what we adopt in this paper. We now consider to divide the circle of length  $L$ , on which the large  $c$  CFT lives, into a small subsystem  $A$  of length  $\ell$  and its large complement  $B$  of length  $L - \ell$ . We can define the reduced density matrices  $\rho_{A,\phi}$  and  $\rho_{B,\phi}$  for the excited state  $\rho_\phi$ , and  $\rho_{A,\text{me}}$  and  $\rho_{B,\text{me}}$  for the microcanonical thermal state  $\rho_{\text{me}}$ . We then use OPE of twist operators to calculate dissimilarity measures for comparing  $\rho_{A,\phi}$ ,  $\rho_{A,\text{me}}$ , and for comparing  $\rho_{B,\phi}$ ,  $\rho_{B,\text{me}}$ .

We only include contributions from the vacuum conformal family in the following calculation. For the small subsystem  $A$ , from (5.7) we get

$$\text{tr}_A(\rho_{A,\phi}^m \rho_{A,\text{me}}^{n-m}) \simeq \text{tr}_A \rho_{A,\phi}^n, \quad m = 0, 1, \dots, n, \quad (5.9)$$

and we further get the entanglement entropy, relative entropy, and Jensen-Shannon divergence

$$S(\rho_{A,\phi}) \simeq S(\rho_{A,\text{me}}), \quad S(\rho_{A,\phi} \parallel \rho_{A,\text{me}}) \simeq 0, \quad JS(\rho_{A,\phi}, \rho_{A,\text{me}}) \simeq 0. \quad (5.10)$$

For the large subsystem  $B$ , we use (5.3) and [49]

$$\text{tr}_B[\rho_B(\phi_{i_1}, \phi_{i_1}) \rho_B(\phi_{i_2}, \phi_{i_2}) \cdots \rho_B(\phi_{i_n}, \phi_{i_n})] = \text{tr}_A[\rho_A(\phi_{i_1}, \phi_{i_2}) \rho_A(\phi_{i_2}, \phi_{i_3}) \cdots \rho_A(\phi_{i_n}, \phi_{i_1})], \quad (5.11)$$

and get

$$\text{tr}_B \rho_{B,\text{me}}^n \simeq \Omega^{1-n} \text{tr}_A \rho_{A,\phi}^n, \quad \text{tr}_B(\rho_{B,\phi}^m \rho_{B,\text{me}}^{n-m}) \simeq \Omega^{m-n} \text{tr}_A \rho_{A,\phi}^n, \quad m = 1, 2, \dots, n. \quad (5.12)$$

Then we get

$$\begin{aligned} S(\rho_{B,\phi}) &= S(\rho_{A,\phi}), \quad S(\rho_{B,\text{me}}) \simeq S(\rho_{A,\phi}) + \log \Omega, \\ S(\rho_{B,\phi} \parallel \rho_{B,\text{me}}) &\simeq \log \Omega, \\ JS(\rho_{B,\phi}, \rho_{B,\text{me}}) &\simeq \log 2 + \frac{1}{2} \log \Omega - \frac{\Omega + 1}{2\Omega} \log(\Omega + 1). \end{aligned} \quad (5.13)$$

The above result agrees with the expectation from ETH, which states that the energy eigenstate approximates the microcanonical thermal state only for a small enough subsystem but not for a large one. This is also verified by the numerical simulations for lattice models done in [50] as long as the size of subsystem is smaller than the half of the total system size. When the size of the subsystem  $A$  becomes as large as half the total system size, the trace square distance starts to deviate from zero, and the behavior indicates that one may be able to extract some critical exponents from the behavior around  $\ell = L/2$ .

As a byproduct, from (5.10), (5.13), we get the approximate saturation of the microcanonical ensemble version of the Araki-Lieb inequality [51]

$$S(\rho_{\text{me}}) - S(\rho_{B,\text{me}}) + S(\rho_{A,\text{me}}) \simeq 0. \quad (5.14)$$

The saturation of canonical ensemble version of the Araki-Lieb inequality and its holography have been studied in [49, 52–58]. We expect the saturation in (5.14) is lifted if the approximation (5.2) is scrutinized carefully and/or contributions from nonvacuum conformal families are included, i.e., that

$$S(\rho_{\text{me}}) - S(\rho_{B,\text{me}}) + S(\rho_{A,\text{me}}) > 0. \quad (5.15)$$

## 6 Conclusion and discussion

We have used the OPE of the twist operators to calculate various quantities that can be used to characterize the dissimilarity of two reduced density matrices, and these quantities include the Rényi entropy, entanglement entropy, relative entropy, Jensen-Shannon divergence, as well as the Schatten 2-norm and 4-norm. We first consider contributions from only the holomorphic sector of the vacuum conformal family, and make expansion of all the quantities by the length of short interval  $\ell$  to order  $\ell^9$ . As an application of the results, for ETH we show how these dissimilarity measures behave for the excited and thermal states in the high temperature limit. As we have showed in this paper, all these quantities can capture the dissimilarity of the two reduced density matrices. Furthermore, we also discuss the possibility to define ETH with GGE thermal state. By using GGE, we provide a possible scenario to define ETH and resolve the mismatch between ETH and highly excited states in large  $c$  CFT. We also discuss ETH for microcanonical ensemble thermal state. In the appendix we give the formal forms of the entanglement entropy, relative entropy, Jensen-Shannon divergence, and Fisher quantum metric with contributions from general operators.

In the method of twist operators we cannot calculate the trace distance, which is essential for the definition of subsystem ETH [3, 4]. Trace distance is just the Schatten  $n$ -norm with  $n = 1$ , and the absolute value in the definition makes it hard to evaluate when  $n$  is an odd integer. It would be nice if the trace distance can be calculated in CFT.

## Acknowledgement

We would like to thank Alexandre Belin, Xi Dong, Nabil Iqbal, Guojing Liu, and Gábor Sárosi for helpful discussions. We thank the anonymous JHEP referee of our previous paper [10] for discussions

about higher order conserved charges the generalized Gibbs ensemble. SH is supported by Max-Planck fellowship in Germany and the National Natural Science Foundation of China Grant No. 11305235. FLL is supported by Taiwan Ministry of Science and Technology through Grant No. 103-2112-M-003-001-MY3, No. 106-2112-M-003-004-MY3 and No. 103-2811-M-003-024. JJZ is supported by the ERC Starting Grant 637844-HBQFTNCER and in part by Italian Ministero dell’Istruzione, Università e Ricerca (MIUR) and Istituto Nazionale di Fisica Nucleare (INFN) through the “Gauge Theories, Strings, Supergravity” (GSS) research project.

## A Relative entropy from modular Hamiltonian

We calculate the relative entropies using modular Hamiltonian as shown in figure 5, and some of them have been calculated from the same method in [59,60]. This appendix serves as a check of the relative entropies from twist operators in section 3.1.

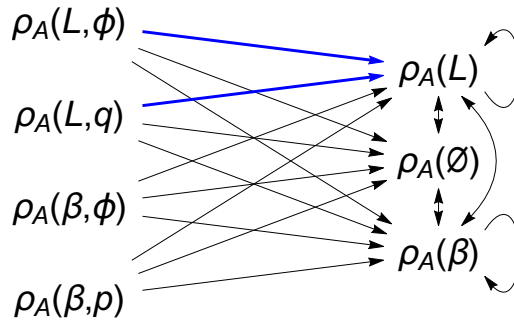


Figure 5: The 20 relative entropies we can calculate using modular Hamiltonian and entanglement entropy. We only need to calculate the two relative entropies  $S(\rho_A(L_1, \phi) \|\rho_A(L_2))$  and  $S(\rho_A(L_1, q) \|\rho_A(L_2))$  as marked in blue.

For a reduced density matrix  $\rho_A$ , the modular Hamiltonian  $H(\rho_A)$  is defined as

$$\rho_A = \frac{e^{-H(\rho_A)}}{\text{tr}_A e^{-H(\rho_A)}}. \quad (\text{A.1})$$

For two reduced density matrices  $\rho_A, \sigma_A$ , the relative entropy can be written as

$$S(\rho_A \|\sigma_A) = \langle H(\sigma_A) \rangle_\rho - \langle H(\sigma_A) \rangle_\sigma - S(\rho_A) + S(\sigma_A), \quad (\text{A.2})$$

with  $H(\sigma_A)$  being the modular Hamiltonian of  $\sigma_A$ . The modular Hamiltonian is known only for cases of  $\rho_A(\emptyset)$ ,  $\rho_A(L)$  and  $\rho_A(\beta)$ , and one has [43, 61, 62]<sup>4</sup>

$$\begin{aligned} H_A(\emptyset) &= - \int_0^\ell \frac{x(\ell-x)}{\ell} T(x) dx, \\ H_A(L) &= - \frac{L}{\pi} \int_0^\ell \frac{\sin \frac{\pi x}{L} \sin \frac{\pi(\ell-x)}{L}}{\sin \frac{\pi \ell}{L}} T(x) dx, \\ H_A(\beta) &= - \frac{\beta}{\pi} \int_0^\ell \frac{\sinh \frac{\pi x}{\beta} \sinh \frac{\pi(\ell-x)}{\beta}}{\sinh \frac{\pi \ell}{\beta}} T(x) dx. \end{aligned} \quad (\text{A.3})$$

<sup>4</sup>One can see modular Hamiltonian for excited states in [63,64].

We have only incorporated contributions from the holomorphic sector. They satisfy the relations

$$H_A(\emptyset) = \lim_{L \rightarrow \infty} H_A(L) = \lim_{\beta \rightarrow \infty} H_A(\beta), \quad H_A(\beta) = H_A(L)|_{L \rightarrow i\beta}. \quad (\text{A.4})$$

As shown in figure 5, we use the entanglement entropy and modular Hamiltonian to calculate the relative entropies  $S(\rho_A(L_1, \phi) \| \rho_A(L_2))$  and  $S(\rho_A(L_1, q) \| \rho_A(L_2))$ . We have

$$\begin{aligned} S(\rho_A(L_1, \phi) \| \rho_A(L_2)) &= \frac{\pi^4 [cL_1^2 - L_2^2(c - 24h_\phi)]^2 \ell^4}{1080cL_1^4 L_2^4} \\ &+ \frac{\pi^6 [cL_1^2 - L_2^2(c - 24h_\phi)]^2 [2cL_1^2 + L_2^2(c - 24h_\phi)] \ell^6}{17010c^2 L_1^6 L_2^6} + \dots \ell^8 + O(\ell^{10}), \end{aligned} \quad (\text{A.5})$$

and this is in accord with (3.7) and (3.8). We have

$$\begin{aligned} S(\rho_A(L_1, q) \| \rho_A(L_2)) &= \left[ \frac{\pi^4 c(L_1^2 - L_2^2)^2}{1080L_1^4 L_2^4} + \frac{4\pi^4 (L_1^2 - L_2^2)q^2}{45L_1^4 L_2^2} + \frac{2\pi^4 (L_1^2 - L_2^2)q^3}{15L_1^4 L_2^2} \right. \\ &+ \left. \frac{4\pi^4 [cL_1^2 - (c-8)L_2^2]q^4}{15cL_1^4 L_2^2} + O(q^5) \right] \ell^4 + \left[ \frac{\pi^6 c(2L_1^6 - 3L_2^2 L_1^4 + L_2^6)}{17010L_1^6 L_2^6} + \frac{8\pi^6 (L_1^4 - L_2^4)q^2}{945L_1^6 L_2^4} \right. \\ &+ \left. \frac{4\pi^6 (L_1^4 - L_2^4)q^3}{315L_1^6 L_2^4} + \frac{8\pi^6 [cL_1^4 - (c-16)L_2^4]q^4}{315cL_1^6 L_2^4} + O(q^5) \right] \ell^6 + \left[ \frac{\pi^8 c(3L_1^8 - 4L_2^2 L_1^6 + L_2^8)}{226800L_1^8 L_2^8} \right. \\ &+ \left. \frac{4\pi^8 (L_1^6 - L_2^6)q^2}{4725L_1^8 L_2^6} + \frac{2\pi^8 (L_1^6 - L_2^6)q^3}{1575L_1^8 L_2^6} + \frac{4\pi^8 [cL_1^6 + (159c + 728)L_2^6]q^4}{1575cL_1^8 L_2^6} + O(q^5) \right] \ell^8 + O(\ell^{10}), \end{aligned} \quad (\text{A.6})$$

and this is in accord with (3.9) and (3.11).

## B Contributions from general operators

In the main text, we only consider the contributions from the holomorphic part of the vacuum conformal family to order  $\ell^9$ . In this appendix we consider the contributions from general holomorphic and antiholomorphic operators, and we get closed forms of the entanglement entropy, relative entropy, and Jensen-Shannon divergence.

For a short interval  $A = [0, \ell]$  on a Riemann surface  $\mathcal{R}$  that has translational symmetry, we have the reduced density matrix  $\rho_A$  and get

$$\text{tr}_A \rho_A^n = \left( \frac{\epsilon}{\ell} \right)^{4h_\sigma} \left( 1 + \sum_{k=1}^n \sum_{\{\mathcal{X}_1, \dots, \mathcal{X}_k\}} \ell^{\Delta_{\mathcal{X}_1} + \dots + \Delta_{\mathcal{X}_k}} b_{\mathcal{X}_1 \dots \mathcal{X}_k} \langle \mathcal{X}_1 \rangle_{\mathcal{R}} \dots \langle \mathcal{X}_k \rangle_{\mathcal{R}} \right), \quad (\text{B.1})$$

with the summation  $\{\mathcal{X}_1 \dots \mathcal{X}_k\}$  being over different sets of all the nonidentity holomorphic and antiholomorphic quasiprimary operators. For a quasiprimary operator  $\mathcal{X}$ , we use  $\Delta_{\mathcal{X}}$  to denote its scaling dimension. Then we get the entanglement entropy

$$S(\rho_A) = \frac{c}{3} \log \frac{\ell}{\epsilon} + \sum_{k=1}^{\infty} \sum_{\{\mathcal{X}_1, \dots, \mathcal{X}_k\}} \ell^{\Delta_{\mathcal{X}_1} + \dots + \Delta_{\mathcal{X}_k}} a_{\mathcal{X}_1 \dots \mathcal{X}_k} \langle \mathcal{X}_1 \rangle_{\mathcal{R}} \dots \langle \mathcal{X}_k \rangle_{\mathcal{R}}. \quad (\text{B.2})$$

For the same short interval  $A = [0, \ell]$  on another Riemann surface  $\mathcal{S}$  that also has translation symmetry, we have the reduced density matrix  $\sigma_A$  and similar expression for entanglement entropy  $S(\sigma_A)$ . The difference of entanglement entropies is

$$S(\rho_A) - S(\sigma_A) = \sum_{k=1}^{\infty} \sum_{\{\mathcal{X}_1, \dots, \mathcal{X}_k\}} \ell^{\Delta_{\mathcal{X}_1} + \dots + \Delta_{\mathcal{X}_k}} a_{\mathcal{X}_1 \dots \mathcal{X}_k} (\langle \mathcal{X}_1 \rangle_{\mathcal{R}} \dots \langle \mathcal{X}_k \rangle_{\mathcal{R}} - \langle \mathcal{X}_1 \rangle_{\mathcal{S}} \dots \langle \mathcal{X}_k \rangle_{\mathcal{S}}). \quad (\text{B.3})$$

For  $k$  quasiprimary operators  $\mathcal{X}_1, \dots, \mathcal{X}_k$ , and two translation invariant Riemann surfaces  $\mathcal{R}, \mathcal{S}$ , we may define

$$\mathcal{F}_i(\mathcal{X}_1, \dots, \mathcal{X}_k | \mathcal{R}, \mathcal{S}) = \frac{1}{n!} (\langle \mathcal{X}_1 \rangle_{\mathcal{R}} \cdots \langle \mathcal{X}_i \rangle_{\mathcal{R}} \langle \mathcal{X}_{i+1} \rangle_{\mathcal{S}} \cdots \langle \mathcal{X}_k \rangle_{\mathcal{S}} + \text{permutations}), \quad (\text{B.4})$$

with  $0 \leq i \leq k$ . Note that the above definition is normalized such that  $\mathcal{F}_i(\mathcal{X}, \dots, \mathcal{X} | \mathcal{R}, \mathcal{S}) = \langle \mathcal{X} \rangle_{\mathcal{R}}^i \langle \mathcal{X} \rangle_{\mathcal{S}}^{k-i}$ ,  $\mathcal{F}_i(\mathcal{X}_1, \dots, \mathcal{X}_k | \mathcal{R}, \mathcal{R}) = \langle \mathcal{X}_1 \rangle_{\mathcal{R}} \cdots \langle \mathcal{X}_k \rangle_{\mathcal{R}}$ . For  $0 \leq m \leq n$ , we have

$$\text{tr}_A(\rho_A^m \sigma_A^{n-m}) = \left(\frac{\epsilon}{\ell}\right)^{4h_\sigma} \left(1 + \sum_{k=1}^n \sum_{\{\mathcal{X}_1, \dots, \mathcal{X}_k\}} \ell^{\Delta_{\mathcal{X}_1} + \dots + \Delta_{\mathcal{X}_k}} b_{\mathcal{X}_1 \dots \mathcal{X}_k} \sum_{i=0}^k \frac{C_m^i C_{n-m}^{k-i}}{C_n^k} \mathcal{F}_i(\mathcal{X}_1, \dots, \mathcal{X}_k | \mathcal{R}, \mathcal{S})\right). \quad (\text{B.5})$$

Then we get the relative entropy

$$S(\rho_A || \sigma_A) = \sum_{k=2}^{\infty} \sum_{\{\mathcal{X}_1, \dots, \mathcal{X}_k\}} \left\{ \ell^{\Delta_{\mathcal{X}_1} + \dots + \Delta_{\mathcal{X}_k}} a_{\mathcal{X}_1 \dots \mathcal{X}_k} [k \mathcal{F}_1(\mathcal{X}_1, \dots, \mathcal{X}_k | \mathcal{R}, \mathcal{S}) - \langle \mathcal{X}_1 \rangle_{\mathcal{R}} \cdots \langle \mathcal{X}_k \rangle_{\mathcal{R}} - (k-1) \langle \mathcal{X}_1 \rangle_{\mathcal{S}} \cdots \langle \mathcal{X}_k \rangle_{\mathcal{S}}] \right\}, \quad (\text{B.6})$$

and the symmetrized relative entropy

$$S(\rho_A, \sigma_A) = \sum_{k=2}^{\infty} k \sum_{\{\mathcal{X}_1, \dots, \mathcal{X}_k\}} \left\{ \ell^{\Delta_{\mathcal{X}_1} + \dots + \Delta_{\mathcal{X}_k}} a_{\mathcal{X}_1 \dots \mathcal{X}_k} [\mathcal{F}_1(\mathcal{X}_1, \dots, \mathcal{X}_k | \mathcal{R}, \mathcal{S}) + \mathcal{F}_1(\mathcal{X}_1, \dots, \mathcal{X}_k | \mathcal{S}, \mathcal{R}) - \langle \mathcal{X}_1 \rangle_{\mathcal{R}} \cdots \langle \mathcal{X}_k \rangle_{\mathcal{R}} - \langle \mathcal{X}_1 \rangle_{\mathcal{S}} \cdots \langle \mathcal{X}_k \rangle_{\mathcal{S}}] \right\}. \quad (\text{B.7})$$

Using the summation

$$\sum_{m=0}^n C_n^m C_m^i C_{n-m}^{k-i} = 2^{n-k} C_n^k C_k^i, \quad (\text{B.8})$$

and (B.5), we get

$$\text{tr}_A\left(\frac{\rho_A + \sigma_A}{2}\right)^n = \left(\frac{\epsilon}{\ell}\right)^{4h_\sigma} \left(1 + \sum_{k=1}^n \sum_{\{\mathcal{X}_1, \dots, \mathcal{X}_k\}} \ell^{\Delta_{\mathcal{X}_1} + \dots + \Delta_{\mathcal{X}_k}} b_{\mathcal{X}_1 \dots \mathcal{X}_k} \sum_{i=0}^k \frac{C_k^i}{2^k} \mathcal{F}_i(\mathcal{X}_1, \dots, \mathcal{X}_k | \mathcal{R}, \mathcal{S})\right). \quad (\text{B.9})$$

Then we get the Jensen-Shannon divergence

$$JS(\rho_A, \sigma_A) = \sum_{k=2}^{\infty} \sum_{\{\mathcal{X}_1, \dots, \mathcal{X}_k\}} \left\{ \ell^{\Delta_{\mathcal{X}_1} + \dots + \Delta_{\mathcal{X}_k}} a_{\mathcal{X}_1 \dots \mathcal{X}_k} \left[ -\frac{1}{2} \langle \mathcal{X}_1 \rangle_{\mathcal{R}} \cdots \langle \mathcal{X}_k \rangle_{\mathcal{R}} - \frac{1}{2} \langle \mathcal{X}_1 \rangle_{\mathcal{S}} \cdots \langle \mathcal{X}_k \rangle_{\mathcal{S}} + \sum_{i=0}^k \frac{C_k^i}{2^k} \mathcal{F}_i(\mathcal{X}_1, \dots, \mathcal{X}_k | \mathcal{R}, \mathcal{S}) \right] \right\}. \quad (\text{B.10})$$

With the above results, we can also calculate short interval expansion of the Fisher information metric. We parameterize the states of the CFT by  $\theta^\alpha$ , and we have the density matrix  $\rho(\theta)$ , and formally the Riemann surface  $\mathcal{R}(\theta)$ . For the reduced density matrix  $\rho_A(\theta)$ , the Fisher information metric is defined as

$$g_{\alpha\beta} = \frac{1}{2} \left\{ \text{tr}_A[\rho_A(\theta) (\partial_\alpha \log \rho_A(\theta)) (\partial_\beta \log \rho_A(\theta))] + (\alpha \leftrightarrow \beta) \right\}. \quad (\text{B.11})$$

It is related to the relative entropy and Jensen-Shannon divergence as [65, 66]

$$\begin{aligned} S(\rho_A(\theta + \delta\theta) || \rho_A(\theta)) &= \frac{1}{2} g_{\alpha\beta} \delta\theta^\alpha \delta\theta^\beta + O[(\delta\theta)^3], \\ JS(\rho_A(\theta + \delta\theta), \rho_A(\theta)) &= \frac{1}{8} g_{\alpha\beta} \delta\theta^\alpha \delta\theta^\beta + O[(\delta\theta)^3]. \end{aligned} \quad (\text{B.12})$$

From (B.6) or (B.10) we get short interval expansion of the Fisher information metric

$$g_{\alpha\beta} = - \sum_{k=2}^{\infty} \sum_{\{\mathcal{X}_1, \dots, \mathcal{X}_k\}} \ell^{\Delta_{\mathcal{X}_1} + \dots + \Delta_{\mathcal{X}_k}} a_{\mathcal{X}_1 \dots \mathcal{X}_k} \mathcal{G}_{\alpha\beta}(\mathcal{X}_1, \dots, \mathcal{X}_k | \mathcal{R}(\theta)), \quad (\text{B.13})$$

with the definition

$$\mathcal{G}_{\alpha\beta}(\mathcal{X}_1, \dots, \mathcal{X}_k | \mathcal{R}(\theta)) = \sum_{1 \leq i < j \leq k} \{ \langle \mathcal{X}_1 \rangle_{\mathcal{R}(\theta)} \dots [\partial_{\theta^\alpha} \langle \mathcal{X}_i \rangle_{\mathcal{R}(\theta)}] \dots [\partial_{\theta^\beta} \langle \mathcal{X}_j \rangle_{\mathcal{R}(\theta)}] \dots \langle \mathcal{X}_k \rangle_{\mathcal{R}(\theta)} + (\alpha \leftrightarrow \beta) \}. \quad (\text{B.14})$$

In principle, the Fisher information metric can be used to define the distance on the state space, i.e., all the thermal and quasi-primary states of 2D CFTs as considered in this paper. Though we do not know at present how to efficiently characterize the state space by this metric, we expect it may help to visualize the ETH geometrically for the future studies.

## References

- [1] J. M. Deutsch, “Quantum statistical mechanics in a closed system,” *Phys. Rev.* **A43** (1991) 2046–2049.
- [2] M. Srednicki, “Chaos and quantum thermalization,” *Phys. Rev.* **E50** (1994) 888–901.
- [3] N. Lashkari, A. Dymarsky, and H. Liu, “Eigenstate Thermalization Hypotehsis in Conformal Field Theory,” [arXiv:1610.00302](#) [[hep-th](#)].
- [4] A. Dymarsky, N. Lashkari, and H. Liu, “Subsystem ETH,” [arXiv:1611.08764](#) [[cond-mat.stat-mech](#)].
- [5] A. L. Fitzpatrick, J. Kaplan, and M. T. Walters, “Universality of Long-Distance AdS Physics from the CFT Bootstrap,” *JHEP* **1408** (2014) 145, [arXiv:1403.6829](#) [[hep-th](#)].
- [6] A. L. Fitzpatrick, J. Kaplan, and M. T. Walters, “Virasoro Conformal Blocks and Thermalities from Classical Background Fields,” *JHEP* **1511** (2015) 200, [arXiv:1501.05315](#) [[hep-th](#)].
- [7] C. T. Asplund, A. Bernamonti, F. Galli, and T. Hartman, “Holographic Entanglement Entropy from 2d CFT: Heavy States and Local Quenches,” *JHEP* **1502** (2015) 171, [arXiv:1410.1392](#) [[hep-th](#)].
- [8] P. Caputa, J. Simón, A. Štikonas, and T. Takayanagi, “Quantum Entanglement of Localized Excited States at Finite Temperature,” *JHEP* **1501** (2015) 102, [arXiv:1410.2287](#) [[hep-th](#)].

- [9] F.-L. Lin, H. Wang, and J.-j. Zhang, “Thermality and excited state Rényi entropy in two-dimensional CFT,” *JHEP* **1611** (2016) 116, [arXiv:1610.01362 \[hep-th\]](#).
- [10] S. He, F.-L. Lin, and J.-j. Zhang, “Subsystem eigenstate thermalization hypothesis for entanglement entropy in CFT,” [arXiv:1703.08724 \[hep-th\]](#).
- [11] P. Basu, D. Das, S. Datta, and S. Pal, “On thermality of CFT eigenstates,” [arXiv:1705.03001 \[hep-th\]](#).
- [12] P. Calabrese and J. L. Cardy, “Entanglement entropy and quantum field theory,” *J. Stat. Mech.* **0406** (2004) P06002, [arXiv:hep-th/0405152 \[hep-th\]](#).
- [13] P. Calabrese, J. Cardy, and E. Tonni, “Entanglement entropy of two disjoint intervals in conformal field theory,” *J. Stat. Mech.* **0911** (2009) P11001, [arXiv:0905.2069 \[hep-th\]](#).
- [14] M. Headrick, “Entanglement Rényi entropies in holographic theories,” *Phys. Rev.* **D82** (2010) 126010, [arXiv:1006.0047 \[hep-th\]](#).
- [15] P. Calabrese, J. Cardy, and E. Tonni, “Entanglement entropy of two disjoint intervals in conformal field theory II,” *J. Stat. Mech.* **1101** (2011) P01021, [arXiv:1011.5482 \[hep-th\]](#).
- [16] B. Chen and J.-j. Zhang, “On short interval expansion of Rényi entropy,” *JHEP* **1311** (2013) 164, [arXiv:1309.5453 \[hep-th\]](#).
- [17] B. Chen, J.-B. Wu, and J.-j. Zhang, “Short interval expansion of Rényi entropy on torus,” *JHEP* **1608** (2016) 130, [arXiv:1606.05444 \[hep-th\]](#).
- [18] E. Iyoda, K. Kaneko, and T. Sagawa, “Fluctuation Theorem for Many-Body Pure Quantum States,” [arXiv:1603.07857 \[cond-mat.stat-mech\]](#).
- [19] H. Tasaki, “On the local equivalence between the canonical and the microcanonical distributions for quantum spin systems,” [arXiv:1609.06983 \[cond-mat.stat-mech\]](#).
- [20] M. Rigol, V. Dunjko, V. Yurovsky, and M. Olshanii, “Relaxation in a Completely Integrable Many-Body Quantum System: An AbInitio Study of the Dynamics of the Highly Excited States of 1D Lattice Hard-Core Bosons,” *Phys. Rev. Lett.* **98** (2007) 050405, [arXiv:cond-mat/0604476 \[cond-mat\]](#).
- [21] G. Mandal, R. Sinha, and N. Sorokhaibam, “Thermalization with chemical potentials, and higher spin black holes,” *JHEP* **1508** (2015) 013, [arXiv:1501.04580 \[hep-th\]](#).
- [22] J. Cardy, “Quantum Quenches to a Critical Point in One Dimension: some further results,” *J. Stat. Mech.* **1602** (2016) 023103, [arXiv:1507.07266 \[cond-mat.stat-mech\]](#).
- [23] G. Mandal, S. Paranjape, and N. Sorokhaibam, “Thermalization in 2D critical quench and UV/IR mixing,” [arXiv:1512.02187 \[hep-th\]](#).

- [24] L. Vidmar and M. Rigol, “Generalized Gibbs ensemble in integrable lattice models,” *J. Stat. Mech.* **6** (2016) 064007, [arXiv:1604.03990](#) [[cond-mat.stat-mech](#)].
- [25] J. de Boer and D. Engelhardt, “Remarks on thermalization in 2D CFT,” *Phys. Rev.* **D94** (2016) 126019, [arXiv:1604.05327](#) [[hep-th](#)].
- [26] B. Chen, J. Long, and J.-j. Zhang, “Holographic Rényi entropy for CFT with  $W$  symmetry,” *JHEP* **1404** (2014) 041, [arXiv:1312.5510](#) [[hep-th](#)].
- [27] M. R. Gaberdiel, C. A. Keller, and R. Volpato, “Genus Two Partition Functions of Chiral Conformal Field Theories,” *Commun. Num. Theor. Phys.* **4** (2010) 295–364, [arXiv:1002.3371](#) [[hep-th](#)].
- [28] J. Cardy, “Thermalization and Revivals after a Quantum Quench in Conformal Field Theory,” *Phys. Rev. Lett.* **112** (2014) 220401, [arXiv:1403.3040](#) [[cond-mat.stat-mech](#)].
- [29] M. Headrick, A. Maloney, E. Perlmutter, and I. G. Zadeh, “Rényi entropies, the analytic bootstrap, and 3D quantum gravity at higher genus,” *JHEP* **1507** (2015) 059, [arXiv:1503.07111](#) [[hep-th](#)].
- [30] A. Belin, C. A. Keller, and I. G. Zadeh, “Genus Two Partition Functions and Renyi Entropies of Large  $c$  CFTs,” [arXiv:1704.08250](#) [[hep-th](#)].
- [31] J. Cardy, A. Maloney, and H. Maxfield, “A new handle on three-point coefficients: OPE asymptotics from genus two modular invariance,” [arXiv:1705.05855](#) [[hep-th](#)].
- [32] C. A. Keller, G. Mathys, and I. G. Zadeh, “Bootstrapping Chiral CFTs at Genus Two,” [arXiv:1705.05862](#) [[hep-th](#)].
- [33] M. Cho, S. Collier, and X. Yin, “Genus Two Modular Bootstrap,” [arXiv:1705.05865](#) [[hep-th](#)].
- [34] M. Beccaria and G. Macorini, “On the next-to-leading holographic entanglement entropy in  $AdS_3/CFT_2$ ,” *JHEP* **1404** (2014) 045, [arXiv:1402.0659](#) [[hep-th](#)].
- [35] Z. Li and J.-j. Zhang, “On one-loop entanglement entropy of two short intervals from OPE of twist operators,” *JHEP* **1605** (2016) 130, [arXiv:1604.02779](#) [[hep-th](#)].
- [36] S. He, T. Numasawa, T. Takayanagi, and K. Watanabe, “Quantum dimension as entanglement entropy in two dimensional conformal field theories,” *Phys. Rev.* **D90** (2014) 041701, [arXiv:1403.0702](#) [[hep-th](#)].
- [37] W.-Z. Guo and S. He, “Rényi entropy of locally excited states with thermal and boundary effect in 2D CFTs,” *JHEP* **1504** (2015) 099, [arXiv:1501.00757](#) [[hep-th](#)].
- [38] B. Chen, W.-Z. Guo, S. He, and J.-q. Wu, “Entanglement Entropy for Descendent Local Operators in 2D CFTs,” *JHEP* **1510** (2015) 173, [arXiv:1507.01157](#) [[hep-th](#)].



- [39] J. Cardy and C. P. Herzog, “Universal Thermal Corrections to Single Interval Entanglement Entropy for Two Dimensional Conformal Field Theories,” *Phys. Rev. Lett.* **112** (2014) 171603, [arXiv:1403.0578 \[hep-th\]](#).
- [40] B. Chen and J.-q. Wu, “Single interval Rényi entropy at low temperature,” *JHEP* **1408** (2014) 032, [arXiv:1405.6254 \[hep-th\]](#).
- [41] B. Chen, J.-q. Wu, and Z.-c. Zheng, “Holographic Rényi entropy of single interval on torus: with W symmetry,” *Phys. Rev.* **D92** (2015) 066002, [arXiv:1507.00183 \[hep-th\]](#).
- [42] K. Ohmori and Y. Tachikawa, “Physics at the entangling surface,” *J. Stat. Mech.* **1504** (2015) P04010, [arXiv:1406.4167 \[hep-th\]](#).
- [43] J. Cardy and E. Tonni, “Entanglement hamiltonians in two-dimensional conformal field theory,” *J. Stat. Mech.* **1612** (2016) 123103, [arXiv:1608.01283 \[cond-mat.stat-mech\]](#).
- [44] R. Sasaki and I. Yamanaka, “Virasoro Algebra, Vertex Operators, Quantum Sine-Gordon and Solvable Quantum Field Theories,” *Adv. Stud. Pure Math.* **16** (1988) 271–296.
- [45] V. V. Bazhanov, S. L. Lukyanov, and A. B. Zamolodchikov, “Integrable structure of conformal field theory, quantum KdV theory and thermodynamic Bethe ansatz,” *Commun. Math. Phys.* **177** (1996) 381–398, [arXiv:hep-th/9412229 \[hep-th\]](#).
- [46] S. Datta, J. R. David, M. Ferlino, and S. P. Kumar, “Higher spin entanglement entropy from CFT,” *JHEP* **1406** (2014) 096, [arXiv:1402.0007 \[hep-th\]](#).
- [47] S. Datta, J. R. David, M. Ferlino, and S. P. Kumar, “Universal correction to higher spin entanglement entropy,” *Phys. Rev.* **D90** (2014) 041903, [arXiv:1405.0015 \[hep-th\]](#).
- [48] S. Datta, J. R. David, and S. P. Kumar, “Conformal perturbation theory and higher spin entanglement entropy on the torus,” *JHEP* **1504** (2015) 041, [arXiv:1412.3946 \[hep-th\]](#).
- [49] B. Chen and J.-q. Wu, “Universal relation between thermal entropy and entanglement entropy in conformal field theories,” *Phys. Rev.* **D91** (2015) 086012, [arXiv:1412.0761 \[hep-th\]](#).
- [50] J. R. Garrison and T. Grover, “Does a single eigenstate encode the full Hamiltonian?,” [arXiv:1503.00729 \[cond-mat.str-el\]](#).
- [51] H. Araki and E. H. Lieb, “Entropy inequalities,” *Commun. Math. Phys.* **18** (1970) 160–170.
- [52] S. Ryu and T. Takayanagi, “Holographic derivation of entanglement entropy from AdS/CFT,” *Phys. Rev. Lett.* **96** (2006) 181602, [arXiv:hep-th/0603001 \[hep-th\]](#).
- [53] M. Headrick and T. Takayanagi, “A Holographic proof of the strong subadditivity of entanglement entropy,” *Phys. Rev.* **D76** (2007) 106013, [arXiv:0704.3719 \[hep-th\]](#).

- [54] T. Azeyanagi, T. Nishioka, and T. Takayanagi, “Near Extremal Black Hole Entropy as Entanglement Entropy via AdS(2)/CFT(1),” *Phys. Rev.* **D77** (2008) 064005, [arXiv:0710.2956 \[hep-th\]](#).
- [55] D. D. Blanco, H. Casini, L.-Y. Hung, and R. C. Myers, “Relative Entropy and Holography,” *JHEP* **08** (2013) 060, [arXiv:1305.3182 \[hep-th\]](#).
- [56] V. E. Hubeny, H. Maxfield, M. Rangamani, and E. Tonni, “Holographic entanglement plateaux,” *JHEP* **1308** (2013) 092, [arXiv:1306.4004 \[hep-th\]](#).
- [57] T. Faulkner, A. Lewkowycz, and J. Maldacena, “Quantum corrections to holographic entanglement entropy,” *JHEP* **1311** (2013) 074, [arXiv:1307.2892](#).
- [58] B. Chen, Z. Li, and J.-j. Zhang, “Corrections to holographic entanglement plateau,” [arXiv:1707.07354 \[hep-th\]](#).
- [59] G. Sárosi and T. Ugajin, “Relative entropy of excited states in two dimensional conformal field theories,” *JHEP* **1607** (2016) 114, [arXiv:1603.03057 \[hep-th\]](#).
- [60] G. Sárosi and T. Ugajin, “Relative entropy of excited states in conformal field theories of arbitrary dimensions,” *JHEP* **1702** (2017) 060, [arXiv:1611.02959 \[hep-th\]](#).
- [61] H. Casini, M. Huerta, and R. C. Myers, “Towards a derivation of holographic entanglement entropy,” *JHEP* **1105** (2011) 036, [arXiv:1102.0440 \[hep-th\]](#).
- [62] G. Wong, I. Klich, L. A. Pando Zayas, and D. Vaman, “Entanglement Temperature and Entanglement Entropy of Excited States,” *JHEP* **1312** (2013) 020, [arXiv:1305.3291 \[hep-th\]](#).
- [63] N. Lashkari, “Modular Hamiltonian for Excited States in Conformal Field Theory,” *Phys. Rev. Lett.* **117** (2016) 041601, [arXiv:1508.03506 \[hep-th\]](#).
- [64] G. Sárosi and T. Ugajin, “Modular Hamiltonians of excited states, OPE blocks and emergent bulk fields,” [arXiv:1705.01486 \[hep-th\]](#).
- [65] P. Sánchez-Moreno, A. Zarzo, and J. S. Dehesa, “Jensen divergence based on Fisher’s information,” *J. Phys.* **A45** (2012) 125305, [arXiv:1012.5041 \[cs.IT\]](#).
- [66] G. E. Crooks, “Measuring Thermodynamic Length,” *Phys. Rev. Lett.* **99** (2007) 100602, [arXiv:0706.0559 \[cond-mat.stat-mech\]](#).