

Mixed-state entanglement and information recovery in thermalized states and evaporating black holes

Shreya Vardhan,^{a,1} Jonah Kudler-Flam,^{b,1} Hassan Shapourian^c and Hong Liu^a

^aCenter for Theoretical Physics, Massachusetts Institute of Technology,
Cambridge, MA 02139, U.S.A.

^bKadanoff Center for Theoretical Physics, University of Chicago,
Chicago, IL 60637, U.S.A.

^cMicrosoft Station Q,
Santa Barbara, CA 93109, U.S.A.

E-mail: vardhan@mit.edu, jkudlerflam@uchicago.edu,
hassan.shapp@gmail.com, hong_liu@mit.edu

ABSTRACT: We study the universal behavior of quantum information-theoretic quantities in thermalized isolated quantum many-body systems and evaporating black holes. In particular, we study a genuine mixed-state entanglement measure called the logarithmic negativity, other correlation measures including the Renyi negativities and the mutual information, and a signature of multipartite entanglement called the reflected entropy. We also probe the feasibility of recovering quantum information from subsystems of a thermalized quantum many-body system or from the radiation of an evaporating black hole, using quantities such as relative entropy and Petz map fidelity. A recently developed technique called the equilibrium approximation allows us to probe these quantities at finite temperature. We find striking qualitative differences from the infinite temperature case, which has been the topic of previous studies using Haar-random states. In particular, we find regimes where the logarithmic negativity is extensive but the mutual information is sub-extensive, indicating a large amount of undistillable, bound entanglement in thermalized states. For evaporating black holes at finite temperature, both the logarithmic negativity and the Petz map fidelity reveal an important new time scale t_b , which is earlier than the Page time t_p by a finite fraction of the total evaporation time. We find that t_b , as opposed to t_p , is the time scale at which quantum entanglement between different parts of the radiation becomes extensive, and the fidelity of information recovery for a large diary thrown into the black hole starts to grow.

KEYWORDS: 2D Gravity, AdS-CFT Correspondence, Black Holes, Thermal Field Theory

ARXIV EPRINT: [2112.00020](https://arxiv.org/abs/2112.00020)

¹These authors contributed equally to the work.

Contents

1	Introduction	1
2	Review of background materials	4
2.1	Entanglement in mixed states	5
2.2	Renyi entropies and negativities	7
2.3	Brief review of the equilibrium approximation	9
2.3.1	Pure states	9
2.3.2	Mixed states	13
2.4	The Petz recovery map	14
3	Equilibrium approximation for quantum-informational quantities	16
3.1	Renyi and logarithmic negativities	16
3.2	Relative entropy	19
3.3	The fidelity of the Petz map	21
3.4	Reflected entropy	23
4	Entanglement phase diagram of an equilibrated mixed state	24
4.1	General setup	24
4.2	Entanglement structure at infinite temperature	26
4.3	Entanglement phase structure at a finite temperature: general discussion	29
4.4	Canonical ensemble	32
4.4.1	General setup	32
4.4.2	A and B are uniform	36
4.4.3	A at infinite temperature	39
4.5	Microcanonical ensemble	40
4.5.1	General description	40
4.5.2	B at infinite temperature	42
4.5.3	A_2 at infinite temperature	44
4.6	Numerical examples: negativity spectrum in chaotic spin chains	47
5	Implications for mixed-state entanglement and black hole radiation	49
6	Implications for information transfer from black hole to radiation	51
6.1	From the perspective of mutual information	53
6.2	Recovery channel: bound from relative entropy	55
6.2.1	Infinite temperature	55
6.2.2	Finite temperature	56
6.3	Recovery channel: the Petz map and its fidelity	58
6.3.1	Infinite temperature	58
6.3.2	Finite temperature	62

7	Conclusions and discussions	64
A	Properties of the permutation group	66
B	Estimate of fluctuations in the equilibrium approximation for the negativity	69
C	Resolvent calculations of logarithmic negativity in various cases	71
C.1	General diagrammatic approach	71
C.2	Infinite temperature	73
C.3	Canonical ensemble with A at infinite temperature	74
C.3.1	Toy model for black hole evaporation in JT gravity	76
C.4	Microcanonical ensemble with B at infinite temperature	78
C.5	Microcanonical ensemble with A_2 at infinite temperature	85
D	Maxima of functions of permutations at finite temperature	90
E	$\mathcal{Z}_n^{(\text{PT})}$ in the microcanonical ensemble with energy conservation in AB	91
F	Finite temperature Page curve in inhomogeneous systems	94
F.1	Canonical ensemble with A at infinite temperature and \bar{A} at finite temperature	95
F.2	Microcanonical ensemble with different entropy densities in A and B	95
G	Evolution of Petz map fidelity at finite temperature	99
G.1	Canonical ensemble	99
G.2	Microcanonical ensemble	105

1 Introduction

A chaotic quantum many-body system initially in a far-from-equilibrium pure state should eventually approach a macroscopic equilibrium. In equilibrium, despite the fact that the system is in a pure state, we can use an equilibrium density operator $\rho^{(\text{eq})}$ to characterize its macroscopic properties using quantities such as temperature, thermal entropy, and free energy. Furthermore, expectation values and correlation functions of generic few-body observables can also be reliably calculated using $\rho^{(\text{eq})}$.

Surprisingly, it has been recognized recently that fine-grained quantum-informational quantities such as the Renyi and von Neumann entropies of various subsystems can also be calculated using $\rho^{(\text{eq})}$, in a way which is compatible with unitarity [1]. The method, called the equilibrium approximation, provides a powerful tool for extracting universal quantum-informational properties of a chaotic quantum many-body system. For example, it enables one to make predictions for the entanglement structure of the system at a finite temperature, and can be used to explain certain semi-classical gravity calculations of Renyi and von Neumann entropies for evaporating black holes [2–5].

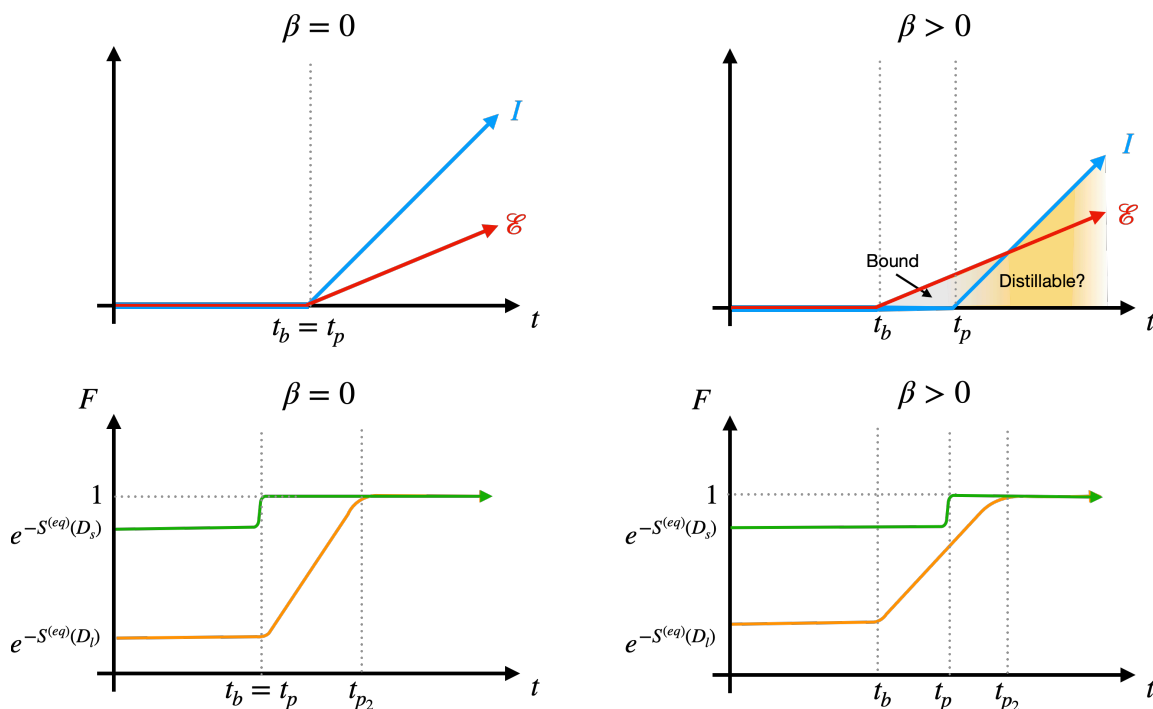


Figure 1. Top row: the general behavior of logarithmic negativity (red) and mutual information (blue) is shown for finite temperature equilibrated pure states, specifically evaporating black holes. While the mutual information within the radiation does not become extensive until the Page time t_p , the negativity starts to become extensive at the earlier time t_b , signaling the existence of quantum correlations in the radiation prior to the Page time that cannot be distilled into EPR pairs. After the Page time, we expect the entanglement to be distillable but do not have a rigorous proof. Bottom row: we also plot the fidelity of the Petz map in the Hayden-Preskill experiment for small (green) and large (orange) diaries. For small diaries, the fidelity increases rapidly from its initial value to 1 at t_p . For sufficiently large diaries, the fidelity begins to increase at the time scale $t_b < t_p$, and reaches a value close to 1 at $t_{p_2} > t_p$. t_{p_2} is defined as the time when the entropy of the radiation is equal to the entropy of the black hole plus the equilibrium entropy of the diary. The cartoon plots in the bottom row are schematic, but should be thought to have a logarithmically scaled y-axis.

In this paper, we generalize the equilibrium approximation to a number of other quantum-informational measures, including Renyi and logarithmic negativities, relative entropy, the fidelity of the Petz map, and reflected entropy. These generalizations enable us to probe and make predictions both for the mixed-state entanglement structure and for information recovery in a system at a *finite* temperature. Besides their implications for general quantum chaotic systems, these issues are also of much interest for probing the quantum nature of black hole evaporation. The results of this paper can be used both to make predictions for the entanglement structure hidden in the Hawking radiation emitted by a black hole, and for understanding how and when quantum information is transferred from a black hole to its radiation. We find various surprising phenomena at finite temperature that have no infinite temperature analog, underscoring the importance of energy conservation.

Our main results can be summarized as follows:¹

1. Consider a system AB in a macroscopically equilibrated pure state $|\Psi\rangle$, and a subsystem $A = A_1 \cup A_2$ whose reduced density operator is given by $\rho_A = \text{Tr}_B |\Psi\rangle \langle \Psi|$. We study the Renyi negativities $R_n(A_1, A_2)$ and the logarithmic negativity $\mathcal{E}(A_1, A_2)$, which encode bi-partite entanglement between A_1 and A_2 in the mixed state ρ_A . The logarithmic negativity \mathcal{E} is non-zero only if ρ_A is not separable, and can be used to lower-bound the PPT entanglement cost, $\mathcal{E}(A_1, A_2) \leq E_c^{\text{(ppt,exact)}}(A_1, A_2)$ [7]. Depending on the sizes of A and A_1 , the behaviors of the negativities give rise to an intricate phase diagram, exhibiting a rich entanglement structure. In particular, finite temperature has a significant effect on the qualitative structure of the entanglement phase diagram.

The behavior of the negativities can further be contrasted with that of the Renyi mutual informations $I_n(A_1, A_2)$ and the mutual information $I(A_1, A_2)$ between A_1 and A_2 . Unlike the logarithmic negativity, the mutual information I can contain information about both classical and quantum correlations in the state, but is nevertheless of importance for understanding the entanglement structure as it upper-bounds the distillable entanglement, $E_d(A_1, A_2) \leq \frac{1}{2}I(A_1, A_2)$ [8]. The phase structure indicated by the mutual information $I(A_1, A_2)$ appears to be insensitive to finite temperature, although the Renyi mutual informations can be sensitive to finite temperature. Comparison between (Renyi) mutual information and negativities indicate that these quantities likely capture complementary aspects of mixed state entanglement.

A particularly surprising result is that in the thermodynamic limit, there can be a finite region in the entanglement phase diagram where the logarithmic negativity is extensive, but the mutual information is sub-extensive, implying that there is a large amount of *bound entanglement*. This phenomenon does not take place in the infinite temperature case previously studied in [9]. This observation also has important implications for black hole physics. It implies the existence of a new time scale t_b before the Page time t_p , after which there is already a significant amount of entanglement correlation within the Hawking radiation. For example, if we take the equilibrium density matrix $\rho^{(\text{eq})}$ to be the canonical ensemble at finite temperature, then t_p is defined as the time scale when $S_B(\rho^{(\text{eq})}) = S_R(\rho^{(\text{eq})})$, where S_B and S_R are the von Neumann entropies in the black hole and radiation subsystems respectively, whereas t_b is the time scale at which $S_{2,B}(\rho^{(\text{eq})}) = S_{\frac{1}{2},R}(\rho^{(\text{eq})})$, where $S_{n,A}$ is the n -th Renyi entropy in subsystem A . In the following point, we give an operational interpretation of t_b . See figure 1.

These results on the behavior of the negativity and mutual information, together with the behavior of a quantity called reflected entropy, also suggest the existence of nontrivial multi-partite entanglement among different parts of the radiation and the black hole.

¹Some of the results in the first point below have been summarized in an earlier paper [6].

2. We consider the Hayden-Preskill thought experiment [10]² for information transfer from a black hole to its Hawking radiation at a finite temperature. We take two perspectives on this process of information transfer, viewing it as (i) growth of the mutual information between an auxiliary reference system and the radiation subsystem, as in [10]; (ii) growth of the fidelity of a recovery map for extracting the information from the radiation. The recovery map is in turn studied from two approaches: from a lower bound on its fidelity provided by relative entropy, and from a direct calculation of the fidelity of an explicit map called the Petz map. We show that in both approaches, the results at finite temperature can be expressed in terms of natural quantum information-theoretic measures of the thermal state. In particular, for a “small” diary whose size does not scale with the total volume of the system, the fidelity rapidly grows from its initial small value to its maximum value of 1 at the standard Page time t_p . For a “large” diary whose size does scale with the volume, the fidelity becomes exponentially close to 1 at a time $t_{p_2} > t_p$, which was previously identified in [11]. But intriguingly, the fidelity of the large diary first starts to grow from its small initial value at the same time scale $t_b < t_p$ when the logarithmic negativity between parts of the radiation starts to grow.³ This result is in contrast to the prior literature based mostly on infinite temperature calculations, which always took t_p and t_{p_2} to be the only relevant time scales for information recovery. These statements can be viewed as giving operational definitions of the time scales t_b , t_p , and t_{p_2} , which apply both at infinite temperature and finite temperature. See figure 1.

The plan of the paper is as follows. In section 2, we review some relevant concepts about mixed-state entanglement, the Renyi and logarithmic negativities, and the Petz recovery map, as well as the equilibrium approximation developed in [1]. In section 3, we explain how to generalize the equilibrium approximation to the various information-theoretic quantities studied in this paper. In section 4, we use these methods to find entanglement phase diagrams for a variety of universality classes of thermalized states. Section 5 discusses the operational implications of these entanglement phase diagrams, and the consequences for black hole dynamics. We study the question of information recovery from the radiation of an evaporating black hole at finite temperature in section 6, and conclude with some discussion of future directions in section 7. Many details are relegated to appendices.

2 Review of background materials

In this section we review topics and technical tools that will be used in subsequent discussions.

²Throw a diary into an evaporating black hole, and see when the information of the diary is recoverable from the radiation.

³As we vary the size of the diary from $O(1)$ to $O(V)$ where V is the volume of the system, the time scale for the beginning of fidelity growth gradually changes from t_p to t_b .

2.1 Entanglement in mixed states

Consider a state ρ in a bipartite system $\mathcal{H}_A = \mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2}$. ρ is said to be a separable state if it can be written as a convex combination of product states,

$$\rho = \sum_{i=1}^q p_i (\rho_i)_{A_1} \otimes (\tilde{\rho}_i)_{A_2}, \quad 0 \leq p_i \leq 1, \quad \sum_{i=1}^q p_i = 1. \quad (2.1)$$

Such a state has no quantum entanglement, as the correlations in it can be given a classical hidden-variable description [12], and it can be prepared using only local operations and classical communications (LOCC) without any need for EPR pairs between A_1 and A_2 . Any state that is not separable is said to be entangled.

If we know that ρ is a pure state, then its entanglement entropy or von Neumann entropy in A_1 , defined as

$$S_{A_1}(\rho) := -\text{Tr}[\rho_{A_1} \log \rho_{A_1}], \quad \rho_{A_1} := \text{Tr}_{A_2}[\rho] \quad (2.2)$$

is sufficient to determine whether it is separable or entangled. ρ is an entangled state if $S_{A_1}(\rho)$ is non-zero, and a separable state with $q = 1$ otherwise.

If ρ is not a pure state, then no easily calculable quantity is known which is zero when the state is unentangled and non-zero when it is entangled. In fact, it is an NP-hard problem to determine whether an arbitrary state is entangled [13].

One familiar quantity that captures correlations between regions A_1 and A_2 in the state ρ is the mutual information,

$$I(A_1, A_2) := S_{A_1}(\rho) + S_{A_2}(\rho) - S_A(\rho), \quad (2.3)$$

where $S_{A_i}(\rho)$ is the von Neumann entropy defined in (2.2). While the mutual information is non-zero for any entangled state, it is also nonzero for a separable state as in (2.1) with $q > 1$, and can hence reflect both classical correlations and quantum entanglement.

Another useful quantity for studying entanglement in mixed states is the logarithmic negativity. Suppose the density matrix ρ has matrix elements $\rho_{a_1 a_2, b_1 b_2}$ in some basis $|a_1 a_2\rangle$ for $\mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2}$. We define the partial transpose ρ^{T_2} of ρ with respect to \mathcal{H}_{A_2} as

$$\rho_{a_1 a_2, b_1 b_2}^{T_2} := \rho_{a_1 b_2, b_1 a_2}. \quad (2.4)$$

The eigenvalues of ρ^{T_2} can now in principle be negative. The presence of negative eigenvalues in ρ^{T_2} reflects that subsystems A_1 and A_2 are entangled; for a separable state (2.1), $\rho^{T_2} = \sum_{i=1}^q p_i (\rho_i)_{A_1} \otimes (\tilde{\rho}_i)_{A_2}^T$ remains a legitimate density operator with non-negative eigenvalues, as the transpose of a positive operator is still positive. This motivates one to define the logarithmic negativity as⁴ [14–20]

$$\mathcal{E}(A_1, A_2) := \log \sum_k |\lambda_k|, \quad (2.5)$$

where λ_k are the eigenvalues of ρ^{T_2} . By definition, ρ^{T_2} has unit trace, and thus $\sum_k |\lambda_k| \geq 1$, which implies that $\mathcal{E} \geq 0$. States with $\mathcal{E} = 0$ are referred to as positive partial transpose

⁴ ρ^{T_1} has the same eigenvalues as ρ^{T_2} , and thus the definition of \mathcal{E} is symmetric between A_1 and A_2 .

(PPT) states, and could still be entangled. Separable states form a proper subset of PPT states. Hence, unlike non-zero mutual information, a non-zero value of the logarithmic negativity implies that ρ is entangled.

Another natural way to quantify entanglement is through hypothetical protocols for interconversion between the state ρ and some number of EPR pairs shared between A_1 and A_2 , using only LOCC. This leads to two operationally motivated measures of entanglement, known as the entanglement cost E_c and the distillable entanglement E_d [21].

To define both of the above quantities, suppose we take n copies of the original system, $A_1^{\otimes n} \otimes A_2^{\otimes n}$. We allow only local operations and classical communication between $A_1^{\otimes n}$ and $A_2^{\otimes n}$, and consider conversions between $\rho^{\otimes n}$ and $(|\text{EPR}\rangle\langle\text{EPR}|)^{\otimes m}$, where

$$|\text{EPR}\rangle = \frac{1}{\sqrt{2}}(|0\rangle_{x_1} |0\rangle_{x_2} + |1\rangle_{x_1} |1\rangle_{x_2}), \quad x_1, x_2 \text{ are qubits in } A_1^{\otimes n}, A_2^{\otimes n}. \quad (2.6)$$

First consider the conversion from $(|\text{EPR}\rangle\langle\text{EPR}|)^{\otimes m}$ to $\rho^{\otimes n}$ under different choices \mathcal{L} of LOCC operations, with vanishing error in the limit $n \rightarrow \infty$. The entanglement cost E_c is defined as the minimum ratio $\frac{m}{n}$ over all choices of \mathcal{L} [22]. Next, consider the conversion from $\rho^{\otimes n}$ to $(|\text{EPR}\rangle\langle\text{EPR}|)^{\otimes m}$ under LOCC operations \mathcal{L} , with vanishing error in the limit $n \rightarrow \infty$. Now the maximum ratio $\frac{m}{n}$ over all choices of \mathcal{L} is defined as the distillable entanglement E_d .

For a pure state ρ , E_c and E_d are both equal to the entanglement entropy $S(\rho_{A_1})$ [23]. This is another justification for seeing the entanglement entropy as a natural measure of entanglement for pure states. For mixed states, in general $E_c \geq E_d$, and neither of these quantities must be equal to $S^{(A_1)}(\rho)$ [21]. In fact, there can be “bound-entangled” states, for which E_c is non-zero while $E_d = 0$ [24].

We can also consider replacing the LOCC operations of the above definitions with a larger set of operations called PPT-preserving transformations, which send any state σ with $\sigma^{T_2} \geq 0$ to another state σ' with $(\sigma')^{T_2} \geq 0$. The entanglement cost and distillable entanglement under such operations, $E_c^{(\text{ppt})}$ and $E_d^{(\text{ppt})}$, are then natural generalizations of the definitions for LOCC given above.

Since LOCC operations are a proper subset of PPT-preserving operations, it is clear that $E_c^{(\text{ppt})} \leq E_c$, and $E_d^{(\text{ppt})} \geq E_d$. It turns out that it is not possible to find states where $E_c^{(\text{ppt})}$ is non-zero while $E_d^{(\text{ppt})} = 0$ [25]. However, there are known examples of states for which the preparation by PPT operations is irreversible, i.e. $E_c^{(\text{ppt})} > E_d^{(\text{ppt})}$ [26].

In the discussion below, we will sometimes refer to “exact” versions of each of the above entanglement costs and distillable entanglements. For these quantities, we modify the above definitions by requiring that the error in the conversion vanishes before taking the $n \rightarrow \infty$ limit. For example, $E_c^{(\text{exact})}$ refers to the minimum ratio $\frac{m}{n}$ in the exact conversion from $(|\text{EPR}\rangle\langle\text{EPR}|)^{\otimes m}$ to $\rho^{\otimes n}$ by LOCC in the $n \rightarrow \infty$ limit. $E_d^{(\text{exact})}$, $E_c^{(\text{ppt,exact})}$, and $E_d^{(\text{ppt,exact})}$ are defined similarly. It is then clear, for instance, that $E_c^{(\text{exact})} \geq E_c$, and $E_d^{(\text{ppt,exact})} \leq E_d^{(\text{ppt})}$.

While the entanglement cost and the distillable entanglement for mixed states are natural generalizations of the entanglement entropy for pure states from an operational perspective, these measures are difficult to compute in practice even for few-qubit systems.

However, they can be related to computable measures such as logarithmic negativity and mutual information through upper and lower bounds. Two inequalities which will be useful in our subsequent discussion are a relation between the mutual information and the distillable entanglement [8],

$$E_d(A_1, A_2) \leq \frac{1}{2}I(A_1, A_2), \tag{2.7}$$

and a relation between the logarithmic negativity and the exact PPT entanglement cost [7],

$$\mathcal{E}(A_1, A_2) \leq E_c^{(\text{ppt,exact})}(A_1, A_2) \leq E_c^{(\text{exact})}(A_1, A_2). \tag{2.8}$$

2.2 Renyi entropies and negativities

Recall that in order to calculate the von Neumann entropy, it is often useful to introduce higher moments of the reduced density matrix, which we will refer to as Renyi partition functions,

$$\mathcal{Z}_{n,A} = \text{Tr}[(\rho_A)^n]. \tag{2.9}$$

These partition functions can be used to define the n -th Renyi entropies

$$S_{n,A}(\rho) = -\frac{1}{n-1} \log \mathcal{Z}_{n,A}, \tag{2.10}$$

which can in turn be used to define the n -th Renyi mutual informations

$$I_n(A_1, A_2) = S_{n,A_1} + S_{n,A_2} - S_{n,A}. \tag{2.11}$$

The Renyi entropies and mutual informations provide further information about the entanglement structure in addition to the von Neumann entropy and mutual information; however, note that in some cases the Renyi mutual information can be negative, unlike the mutual information, which is one reason why the Renyi mutual information should only be considered as proxies for correlation measures. The von Neumann entropy can be written as a limit of the Renyi entropies in the index n ,

$$S_A = \lim_{n \rightarrow 1} S_{n,A}. \tag{2.12}$$

In some cases, such as when the $S_{n,A}$ cannot be written as an analytic function of n , we may not be able to use analytic continuation to find the von Neumann entropy. Another method for obtaining the von Neumann entropy using the Renyi partition functions that turns out to be more generally applicable is through a quantity known as the resolvent,

$$\mathcal{R}(\lambda) = \text{Tr} \left(\frac{1}{\lambda \mathbf{1} - \rho_A} \right) = \frac{1}{\lambda} \sum_{n=0}^{\infty} \frac{1}{\lambda^n} \mathcal{Z}_{n,A}, \tag{2.13}$$

where we take λ to be a general complex parameter. Since the spectrum of ρ_A is bounded, the power series in the second equality is convergent for sufficiently large $|\lambda|$. We can then compute $\mathcal{R}(\lambda)$ for such λ 's and then analytically continue to obtain other values of λ . The

density of eigenvalues of ρ_A , known as the entanglement spectrum, can be obtained from the discontinuities of $\mathcal{R}(\lambda)$ across the real axis,

$$D(\lambda) = \frac{1}{\pi} \text{Im} \mathcal{R}(\lambda - i\epsilon), \quad \lambda \in \mathbb{R}. \quad (2.14)$$

We can then use $D(\lambda)$ to calculate the von Neumann entropy,

$$S_A = - \int d\lambda D(\lambda) \lambda \log \lambda. \quad (2.15)$$

It is similarly useful to define the higher moments of the partially transposed density matrix $\rho_A^{T_2}$ introduced in (2.4), which we refer to as partial transpose partition functions,

$$\mathcal{Z}_n^{(\text{PT})} = \text{Tr} \left(\rho_A^{T_2} \right)^n. \quad (2.16)$$

By definition $\mathcal{Z}_1^{(\text{PT})} = 1$, and it can be readily checked that $\mathcal{Z}_2^{(\text{PT})} = \text{Tr} \rho^2$, so nontrivial moments start with $n = 3$. There are qualitative differences between even and odd n 's, as

$$\mathcal{Z}_{2m}^{(\text{PT})} = \sum_{\lambda_i > 0} |\lambda_i|^{2m} + \sum_{\lambda_i < 0} |\lambda_i|^{2m}, \quad \mathcal{Z}_{2m+1}^{(\text{PT})} = \sum_{\lambda_i > 0} |\lambda_i|^{2m+1} - \sum_{\lambda_i < 0} |\lambda_i|^{2m+1}. \quad (2.17)$$

The logarithmic negativity \mathcal{E} can be obtained by analytically continuing $\mathcal{Z}_{2m}^{(\text{PT})}$ as

$$\mathcal{E} = \lim_{m \rightarrow \frac{1}{2}} \log \mathcal{Z}_{2m}^{(\text{PT})}. \quad (2.18)$$

It can be shown that a PPT state satisfies [27]

$$\mathcal{Z}_3^{(\text{PT})} \geq \left(\mathcal{Z}_2^{(\text{PT})} \right)^2. \quad (2.19)$$

Thus if $\mathcal{Z}_3^{(\text{PT})} < \left(\mathcal{Z}_2^{(\text{PT})} \right)^2$, then it must be that $\mathcal{E} > 0$, which provides a quick diagnostic. The condition (2.19) is weaker than the PPT condition; there can be states that satisfy (2.19) but still have $\mathcal{E} > 0$.

We can also use the partial transpose partition functions to define quantities called the n -th Renyi negativities,

$$R_n(A_1, A_2) = b_n \log \left[\frac{\mathcal{Z}_n^{(\text{PT})}}{\mathcal{Z}_{n,A}} \right], \quad b_n = \begin{cases} \frac{1}{1-n} & n \text{ odd} \\ \frac{1}{2-n} & n \text{ even} \end{cases}. \quad (2.20)$$

$n = 1, 2$ should be understood as being defined through limits. More explicitly,

$$R_1 = - \sum_{\lambda_i > 0} |\lambda_i| \log |\lambda_i| + \sum_{\lambda_i < 0} |\lambda_i| \log |\lambda_i| - S_A(\rho), \quad (2.21)$$

$$R_2 = \frac{1}{\mathcal{Z}_2} \left(- \sum_i |\lambda_i|^2 \log |\lambda_i| + \sum_i |\tilde{\lambda}_i|^2 \log \tilde{\lambda}_i \right), \quad (2.22)$$

where $\tilde{\lambda}_i$ are eigenvalues of ρ . The logarithmic negativity can also be obtained from R_{2m} by analytic continuation

$$\mathcal{E} = \lim_{m \rightarrow \frac{1}{2}} R_{2m}. \quad (2.23)$$

For a product state $\rho = \rho_{A_1} \otimes \rho_{A_2}$, we have $\mathcal{Z}_n^{(\text{PT})} = \mathcal{Z}_n$ and thus $R_n = 0$. But for a general separable state, R_n are in general nonzero. Thus from nonzero R_n , we cannot conclude for certain that there is nonzero entanglement. Even so, there are indications that the Renyi negativities often have the same general behavior as the logarithmic negativity, and these quantities have proven useful in a variety of problems [28–35].

It is often useful to compare the Renyi negativities with the Renyi mutual informations defined in (2.11). In fact, when ρ is a pure state, the Renyi negativities and Renyi mutual informations are related in a simple way. More explicitly, in this case we have

$$R_{2m}(A_1, A_2) = S_{m, A_1} = \frac{1}{2} I_m(A_1, A_2), \quad (2.24)$$

$$R_{2m+1}(A_1, A_2) = S_{2m+1, A_1} = \frac{1}{2} I_{2m+1}(A_1, A_2). \quad (2.25)$$

We also have

$$\mathcal{E}(A_1, A_2) = S_{\frac{1}{2}, A_1} = \frac{1}{2} I_{\frac{1}{2}}(A_1, A_2). \quad (2.26)$$

Equation (2.26) gives a rough intuition on the relative normalization between the logarithmic negativity and mutual information.

In some cases, we will find that the $\mathcal{Z}_n^{(\text{PT})}$ and R_n cannot be written simply as analytic functions of n , so that we will not be able to use (2.18) or (2.23) to find \mathcal{E} . Again, a more generally applicable method is to use the resolvent for $\rho_A^{T_2}$:

$$R_N(\lambda) = \text{Tr} \left(\frac{1}{\lambda \mathbf{1} - \rho_A^{T_2}} \right) = \frac{1}{\lambda} \sum_{n=0}^{\infty} \frac{1}{\lambda^n} \mathcal{Z}_n^{(\text{PT})}. \quad (2.27)$$

The density of eigenvalues of $\rho_A^{T_2}$, known as the negativity spectrum, can be obtained from the discontinuities of $R_N(\lambda)$ across the real axis,

$$D_N(\lambda) = \frac{1}{\pi} \text{Im} R_N(\lambda - i\epsilon), \quad \lambda \in \mathbb{R}, \quad (2.28)$$

and we can use $D_N(\lambda)$ to calculate the logarithmic negativity,

$$\mathcal{E}(A_1, A_2) = \log \left(\int d\lambda D_N(\lambda) |\lambda| \right). \quad (2.29)$$

2.3 Brief review of the equilibrium approximation

In this subsection, we review the equilibrium approximation introduced in [1], which will be the main tool we use in this paper. We first discuss the formulation for a pure state, and then for a mixed state.

2.3.1 Pure states

We consider a system evolving from a far-from-equilibrium pure state $\rho_0 = |\Psi_0\rangle \langle \Psi_0|$ to a state $\rho = |\Psi\rangle \langle \Psi|$ with $|\Psi\rangle = U |\Psi_0\rangle$, which is in equilibrium at a macroscopic level. We assume that macroscopic physical properties of equilibrated pure state ρ can be approximated by an equilibrium density operator

$$\rho^{(\text{eq})} = \frac{\mathcal{I}_\alpha}{Z(\alpha)}, \quad Z(\alpha) = \text{Tr} \mathcal{I}_\alpha, \quad (2.30)$$

where α collectively denotes macroscopic parameters for the equilibrium state.

Consider the n -th Renyi entropy of the equilibrated pure state with respect to a subsystem A

$$Z_{n,A} = e^{-(n-1)S_{n,A}} = \text{Tr}_A \rho_A^n = \text{Tr}_A \left(\text{Tr}_{\bar{A}} U \rho_0 U^\dagger \right)^n = \langle \eta_A \otimes e_{\bar{A}} | (U \otimes U^\dagger)^n | \rho_0, e \rangle, \quad (2.31)$$

where in the last equality we have written it as an amplitude in the replica space $(\mathcal{H} \otimes \mathcal{H})^n$, with various notations defined as follows. For any operator \mathcal{O} acting on \mathcal{H} , the state $|\mathcal{O}, \sigma\rangle \in (\mathcal{H} \otimes \mathcal{H})^n$, where σ is an element of the permutation group \mathcal{S}_n of n objects, is defined as

$$\langle i_1 \bar{i}'_1 i_2 \bar{i}'_2 \cdots i_n \bar{i}'_n | \mathcal{O}, \sigma \rangle = \mathcal{O}_{i_1 i'_{\sigma(1)}} \mathcal{O}_{i_2 i'_{\sigma(2)}} \cdots \mathcal{O}_{i_n i'_{\sigma(n)}}, \quad \mathcal{O}_{ij} = \langle i | \mathcal{O} | j \rangle. \quad (2.32)$$

Here $\{|i_1 \bar{i}'_1 i_2 \bar{i}'_2 \cdots i_n \bar{i}'_n\rangle\}$ is a basis for $(\mathcal{H} \otimes \mathcal{H})^n$ and $\sigma(i)$ denotes the image of i under σ . When \mathcal{O} is given by the identity operator, we will denote the states obtained this way simply as $|\sigma\rangle$. When the system is divided into subsystems, we can similarly define states by associating different permutations to different subsystems. For example, suppose $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_{\bar{A}}$, $|\mathcal{O}, \tau_A \otimes \sigma_{\bar{A}}\rangle$ with $\tau, \sigma \in \mathcal{S}_n$ is defined as

$$\langle i_{1a} i_{1b} \bar{i}'_{1a} \bar{i}'_{1b} \cdots i_{na} i_{nb} \bar{i}'_{na} \bar{i}'_{nb} | \mathcal{O}, \tau_A \otimes \sigma_{\bar{A}} \rangle = \mathcal{O}_{i_{1a} i_{1b}, i'_{\tau(1)a} i'_{\sigma(1)b}} \cdots \mathcal{O}_{i_{na} i_{nb}, i'_{\tau(n)a} i'_{\sigma(n)b}} \quad (2.33)$$

where $|i_{ka}\rangle, |\bar{i}'_{ka}\rangle, |i_{kb}\rangle, |\bar{i}'_{kb}\rangle$ denote respectively basis vectors for subsystem A and \bar{A} in the k -th replica of $\mathcal{H} \otimes \mathcal{H}$. In (2.31), $|\eta_A \otimes e_{\bar{A}}\rangle$ is a state associated with the identity operator, with e representing the identity permutation and η the cyclic permutation $(n, n-1, \dots, 1)$.

We can always decompose the identity on the replica Hilbert space in terms of P_α and its orthogonal projector Q as follows:

$$\mathbb{1} = P_\alpha + Q, \quad P_\alpha Q = Q P_\alpha = 0, \quad Q^2 = Q, \quad (2.34)$$

where P_α is the projector

$$P_\alpha = \frac{1}{Z_2^n} \sum_{\sigma, \tau} g^{\sigma\tau} |\mathcal{I}_\alpha, \sigma\rangle \langle \mathcal{I}_\alpha, \tau|, \quad g_{\tau\sigma} = \frac{\langle \mathcal{I}_\alpha, \tau | \mathcal{I}_\alpha, \sigma \rangle}{\sqrt{\langle \mathcal{I}_\alpha, \tau | \mathcal{I}_\alpha, \tau \rangle \langle \mathcal{I}_\alpha, \sigma | \mathcal{I}_\alpha, \sigma \rangle}}, \quad Z_n := \text{Tr} \mathcal{I}_\alpha^n. \quad (2.35)$$

Inserting the identity twice in the last expression of (2.31) and using the fact that the equilibrium density matrix $\rho^{(\text{eq})}$ commutes with the time-evolution operator U , we get the exact expression

$$Z_{n,A} = \langle \eta_A \otimes e_{\bar{A}} | P_\alpha | \rho_0, e \rangle + \langle \eta_A \otimes e_{\bar{A}} | Q (U \otimes U^\dagger)^n Q | \rho_0, e \rangle. \quad (2.36)$$

The key step in the approximation involves neglecting the second term involving Q , assuming that in a chaotic many-body system it gives a fluctuating contribution which becomes small relative to the time-independent first term at late times. We will be interested in systems with a large number of degrees of freedom, i.e. $Z_1 \gg 1$. For such systems, we then

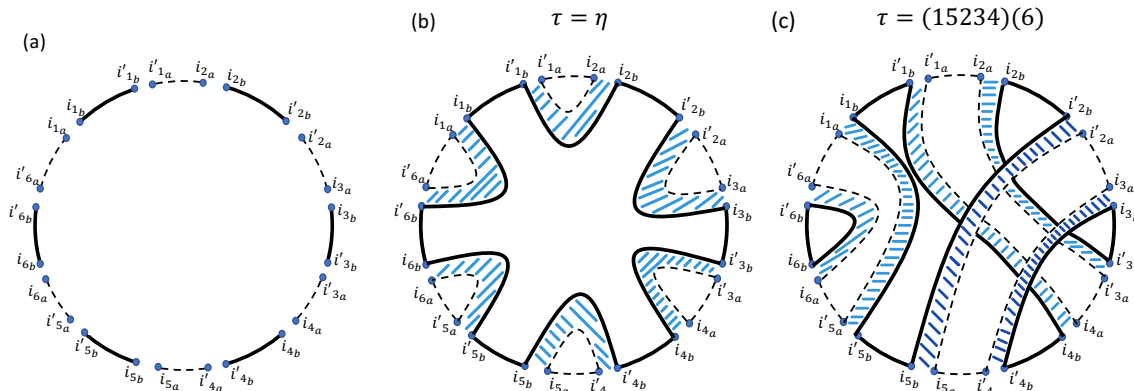


Figure 2. (a) shows the common exterior lines of all diagrams for different terms in (2.39), and (b) and (c) show examples of diagrams for two choices of τ , for the case $n = 6$.

find that $\mathcal{Z}_n^{(A)}$ can be approximated as

$$\mathcal{Z}_{n,A} \approx [\mathcal{Z}_{n,A}]_{\text{eq approx}} := \frac{1}{Z_2^n} \sum_{\sigma, \tau} g^{\tau\sigma} \langle \eta_A \otimes e_{\bar{A}} | \mathcal{I}_\alpha, \tau \rangle \langle \mathcal{I}_\alpha, \sigma | \rho_0, e \rangle \quad (2.37)$$

$$= \frac{1}{Z_1^n} \sum_{\sigma, \tau} g^{\tau\sigma} \langle \eta_A \otimes e_{\bar{A}} | \mathcal{I}_\alpha, \tau \rangle \quad (2.38)$$

$$\approx \frac{1}{Z_1^n} \sum_{\tau \in \mathcal{S}_n} \langle \eta_A \otimes e_{\bar{A}} | \mathcal{I}_\alpha, \tau \rangle. \quad (2.39)$$

In going from (2.37) to (2.38), we use the fact that for (2.37) to be compatible with $\text{Tr} \rho = 1$, \mathcal{I}_α should satisfy a consistency requirement

$$\text{Tr}(\mathcal{I}_\alpha \rho_0) = \frac{Z_2}{Z_1}. \quad (2.40)$$

In going from (2.38) to (2.39), we use the fact that $g^{\tau\sigma}$ is approximately equal to the identity when Z_1 is large.

The size of the terms involving Q that we neglected in reaching (2.39) can be estimated from Δ , defined by

$$\Delta^2 := [(\mathcal{Z}_{n,A})^2]_{\text{eq approx}} - ([\mathcal{Z}_{n,A}]_{\text{eq approx}})^2. \quad (2.41)$$

It was shown in appendix B of [1] that

$$\frac{\Delta}{[\mathcal{Z}_{n,A}]_{\text{eq approx}}} \sim Z_1^{-1/2} \ll 1. \quad (2.42)$$

$\mathcal{Z}_{n,A}$, as given in (2.39), only depends on the equilibrium density operator \mathcal{I}_α , but satisfies the unitarity constraint

$$\mathcal{Z}_{n,A} = \mathcal{Z}_{n,\bar{A}}. \quad (2.43)$$

Each term in the final expression in (2.39) can be given a diagrammatic representation, as shown in figure 2. We can insert the identity to write

$$\langle \eta_A \otimes e_{\bar{A}} | \mathcal{I}_\alpha, \tau \rangle = \sum_{i_1, i'_1, \dots, i_n, i'_n} \langle \eta_A \otimes e_{\bar{A}} | i_1 \bar{i}'_1 \dots i_n \bar{i}'_n \rangle \langle i_1 \bar{i}'_1 \dots i_n \bar{i}'_n | \mathcal{I}_\alpha, \tau \rangle, \quad (2.44)$$

$$|i_m\rangle = |i_{m_a}\rangle_A |i_{m_b}\rangle_{\bar{A}}, \quad |\bar{i}'_m\rangle = |\bar{i}'_{m_a}\rangle_A |\bar{i}'_{m_b}\rangle_{\bar{A}}.$$

The exterior of the diagram, which is the same for all τ , represents $\langle \eta_A \otimes e_{\bar{A}} | i_1 \bar{i}'_1 \dots i_n \bar{i}'_n \rangle$ by connecting i_{m_a} with $i'_{\eta(m)_a}$ using dashed lines, and i_{m_b} with i'_{m_b} using solid lines, as shown in figure 2(a). The interior of the diagram represents $\langle i_1 \bar{i}'_1 \dots i_n \bar{i}'_n | \mathcal{I}_\alpha, \tau \rangle$, by connecting i_m with $i'_{\tau(m)}$, as shown for two examples in figure 2(b) and (c). Roughly, each solid loop in the resulting diagram gives a power of $d_{\bar{A}}$ and each dashed loop gives a power of d_A , where d_A and $d_{\bar{A}}$ are respectively the effective Hilbert space dimensions of A and \bar{A} .⁵ The number of solid and dashed loops in a diagram is respectively equal to $k(\tau)$ (the number of cycles in permutation τ) and $k(\eta^{-1}\tau)$. We therefore find

$$\langle \eta_A \otimes e_{\bar{A}} | \mathcal{I}_\alpha, \tau \rangle \sim d_A^{k(\eta^{-1}\tau)} d_{\bar{A}}^{k(\tau)}. \quad (2.45)$$

For any permutation τ , we have the inequality

$$k(\tau) + k(\eta^{-1}\tau) \leq n + 1, \quad (2.46)$$

and the permutations for which this inequality is saturated are the ones associated with planar diagrams such as figure 2(b).

From (2.45), when A is much smaller than \bar{A} , i.e. $d_A \ll d_{\bar{A}}$, the permutation $\tau = e$, which maximizes $k(\tau)$, dominates, and we have

$$\mathcal{Z}_{n,A} = \mathcal{Z}_{n,A}^{(\text{eq})}, \quad (2.47)$$

where $\mathcal{Z}_{n,A}^{(\text{eq})}$ is the Renyi partition function for A in the state $\rho^{(\text{eq})}$. When A is much larger than \bar{A} , i.e. $d_A \gg d_{\bar{A}}$, $\tau = \eta$, which maximizes $k(\eta^{-1}\tau)$, dominates

$$\mathcal{Z}_{n,A} = \mathcal{Z}_{n,A}^{(\text{eq})}. \quad (2.48)$$

Except for a crossover region around $d_A \sim d_{\bar{A}}$ where the behavior may be more complicated, we then have

$$S_{n,A} = \min(S_{n,A}^{(\text{eq})}, S_{n,\bar{A}}^{(\text{eq})}), \quad n \geq 2 \quad (2.49)$$

where $S_{n,A}^{(\text{eq})}$ denotes the n -th Renyi entropy for subsystem A in the equilibrium density operator $\rho^{(\text{eq})}$. When S_A can be obtained from $S_{n,A}$ by analytic continuation to $n = 1$, (2.49) implies

$$S_A = \min(S_A^{(\text{eq})}, S_{\bar{A}}^{(\text{eq})}), \quad (2.50)$$

⁵When \mathcal{I}_α can be factorized between A and \bar{A} , i.e. $\mathcal{I}_\alpha = \mathcal{I}_A \otimes \mathcal{I}_{\bar{A}}$, we can define the effective dimensions as $d_A = \text{Tr}_A \mathcal{I}_A$ and $d_{\bar{A}} = \text{Tr}_{\bar{A}} \mathcal{I}_{\bar{A}}$. When \mathcal{I}_α cannot be factorized, they can be estimated by counting the numbers of degrees of freedom of the subsystems.

where $S_A^{(\text{eq})}$ is the entanglement entropy for subsystem A in $\rho^{(\text{eq})}$. In cases where the system $A\bar{A}$ is inhomogeneous, in general (2.50) cannot be deduced by analytic continuation. We show in two such examples in appendix F that (2.50) still holds, and will assume in the later discussion in this paper that it holds in general.

When \mathcal{I}_α can be factorized

$$\mathcal{I}_\alpha \approx \mathcal{I}_A \otimes \mathcal{I}_{\bar{A}}, \quad (2.51)$$

equations (2.39) can be written more explicitly in terms of partition functions of A and \bar{A}

$$Z_{m,A} = \text{Tr}_A \mathcal{I}_A^m, \quad Z_{m,\bar{A}} = \text{Tr}_{\bar{A}} \mathcal{I}_{\bar{A}}^m, \quad (2.52)$$

as

$$\mathcal{Z}_{n,A} \approx \frac{1}{Z_1^n} \sum_{\tau} (Z_{m_1,A} \cdots Z_{m_l,A}) (Z_{n_1,\bar{A}} \cdots Z_{n_k,\bar{A}}), \quad (2.53)$$

where k is the number of cycles of τ with n_1, \dots, n_k the lengths of the corresponding cycles, and l is the number of cycles of $\tau\eta^{-1}$ with m_1, \dots, m_l the lengths of the corresponding cycles.

2.3.2 Mixed states

The equilibrium approximation can also be applied to a system $A = A_1 \cup A_2$ in a mixed state ρ_A that is in macroscopic equilibrium, but is far from the thermal density operators [1].

Suppose the system starts with a far-from equilibrium mixed state $\rho_{0,A}$, which evolves under unitary evolution operator U_A to $\rho_A = U_A \rho_{0,A} U_A^\dagger$. All the moments of $\rho_{0,A}$ are also preserved by the time evolution

$$z_{n,A} = \text{Tr} \rho_{0,A}^n = \text{Tr} \rho_A^n = \text{Tr} (U_A \rho_{0,A} U_A^\dagger)^n = \langle \eta | (U_A \otimes U_A^\dagger)^n | \rho_{0,A}, e \rangle, \quad n = 2, \dots \quad (2.54)$$

The statement that $\rho_{0,A}$ is far-from-equilibrium is imposed by requiring that the n -th Renyi entropy of $\rho_{0,A}$ is smaller than the equilibrium entropy of A . For example, this condition is satisfied if we have

$$z_{n,A} \sim Z_A^{-(n-1)f}, \quad 0 \leq f < 1. \quad (2.55)$$

Assuming that ρ_A can be approximated by an equilibrium density operator $\rho_A^{(\text{eq})} = \frac{1}{Z_A} \mathcal{I}_A$ and applying the equilibrium approximation by inserting the projector (2.35) in (2.54) and ignoring terms with $\sigma \neq \tau$, we have

$$z_{n,A} \approx \frac{1}{Z_{2,A}^n} \sum_{\tau} \langle \eta | \mathcal{I}_A, \tau \rangle \langle \mathcal{I}_A, \tau | \rho_{0,A}, e \rangle, \quad (2.56)$$

which can be further simplified to the following constraints on \mathcal{I}_A under the out-of-equilibrium assumption (2.55)

$$\text{Tr} (\mathcal{I}_A \rho_{0,A})^n \approx z_{n,A} \frac{Z_{2,A}^n}{Z_A^n}, \quad Z_{n,A} = \text{Tr}_A \mathcal{I}_A^n, \quad Z_A = Z_{1,A}. \quad (2.57)$$

The Renyi partition functions for A_1 can then be approximated as

$$\mathcal{Z}_{n,A_1} \approx [\mathcal{Z}_{n,A_1}]_{\text{eq approx}} = \frac{1}{Z_A^n} \sum_{\tau} \langle \eta_{A_1} \otimes e_{A_2} | \mathcal{I}_A, \tau \rangle \prod_{i=1}^{k(\tau)} \text{Tr}_A \rho_{0,A}^{n_i}, \quad n = 2, 3, \dots \quad (2.58)$$

where $k(\tau)$ is the number of cycles of τ with $n_1, \dots, n_{k(\tau)}$ the lengths of the cycles.

The above discussion can be further generalized by embedding A in a larger system $S = A \cup B$, with the total system S in an initial pure state $|\Psi_0\rangle$ evolved to $|\Psi\rangle = U |\Psi_0\rangle$ in macroscopic equilibrium.⁶ Suppose $|\Psi\rangle$ can be approximated macroscopically by $\rho^{(\text{eq})} = \frac{1}{Z_1} \mathcal{I}$. We then have

$$\mathcal{Z}_{n,A_1} \approx \frac{1}{Z_1^n} \sum_{\tau} \langle \eta_{A_1} \otimes e_{A_2 B} | \mathcal{I}, \tau \rangle. \quad (2.59)$$

This generalizes (2.58) as under the evolution of U for the full system S , the evolution from the initial density operator $\rho_{0,A}$ to ρ_A is in general not unitary. To recover (2.58) we take $U = U_A \otimes U_B$ to be factorized between A and B , in which case the equilibrium density operator $\mathcal{I} = \mathcal{I}_A \otimes \mathcal{I}_B$ should also factorize, and (2.59) can be written as

$$\mathcal{Z}_{n,A_1} = \frac{1}{Z_A^n} \sum_{\tau} \langle \eta_{A_1} \otimes e_{A_2} | \mathcal{I}_A, \tau \rangle \prod_{i=1}^{k(\tau)} \hat{Z}_{n_i, B}, \quad \hat{Z}_{n, B} := \frac{1}{Z_B^n} \text{Tr}_B \mathcal{I}_B^n. \quad (2.60)$$

Equations (2.58) and (2.60) are equal provided that we choose the initial state $\rho_{0,B}$ such that

$$\text{Tr}_B \rho_{0,B}^m = \frac{1}{Z_B^m} \text{Tr}_B \mathcal{I}_B^m. \quad (2.61)$$

Then since the initial state is pure, $z_{n,A}$ is also given by (2.61). The requirement that $\rho_{0,A}$ is out-of-equilibrium is then equivalent to the requirement that $Z_B \gg Z_A$.

The relation between (2.59) and (2.58) also gives a way to estimate which permutation dominates in (2.58). From (2.59), when A_1 is smaller (larger) than $A_2 B$, the dominant contribution is $\tau = e$ ($\tau = \eta$). Translating these statements to the notation of (2.58), we conclude that

$$S_{n,A_1} = \begin{cases} S_{n,A_1}^{(\text{eq})} & S_{n,A_1}^{(\text{eq})} < S_{n,A_2}^{(\text{eq})} + S_{n,A} \\ S_{n,A_2}^{(\text{eq})} + S_{n,A} & S_{n,A_1}^{(\text{eq})} > S_{n,A_2}^{(\text{eq})} + S_{n,A} \end{cases}. \quad (2.62)$$

In applying (2.59) to explicit calculations, we will need to make assumptions regarding B in the equilibrium density operator $\rho^{(\text{eq})}$ for the full system, which may be considered as specifying different universality classes for ρ_A .

2.4 The Petz recovery map

An important question concerning a quantum channel \mathcal{N} is whether or not it is reversible i.e. whether there exists a recovery channel \mathcal{R} such that for any state ρ , $\mathcal{R} \circ \mathcal{N}(\rho) = \rho$. This question plays a central role, for example, in the theory of quantum error correction as well as quantum thermalization. If the quantum channel is unitary, the initial state is perfectly recovered by acting on the output with the adjoint of the unitary. In the other

⁶ A and B in principle do not have to be in equilibrium with each other.

extreme, a quantum channel could replace all states with the maximally mixed state, in which case an initial state is unrecoverable as the information about it is completely lost. A recovery map that is independent of the initial state ρ is called *universal*.

It follows from a theorem by Petz [36–38] that a quantum channel \mathcal{N} is reversible if and only if the data processing inequality is saturated

$$D(\rho||\sigma) = D(\mathcal{N}(\rho)||\mathcal{N}(\sigma)) \tag{2.63}$$

for all ρ, σ . Here, $D(\rho||\sigma)$ is the relative entropy. Furthermore, suppose we fix some reference state σ . Then, for any state X in the support of $\mathcal{N}(\sigma)$, there exists a recovery channel \mathcal{R} . This channel is given explicitly by the Petz map,

$$\mathcal{P}_{\sigma, \mathcal{N}}(X) = \sigma^{\frac{1}{2}} \mathcal{N}^\dagger \left((\mathcal{N}(\sigma))^{-\frac{1}{2}} X (\mathcal{N}(\sigma))^{-\frac{1}{2}} \right) \sigma^{\frac{1}{2}}, \tag{2.64}$$

where \mathcal{N}^\dagger is the adjoint map of \mathcal{N} . As a basic check, relative entropy is invariant under unitary channels $\rho \rightarrow U\rho U^\dagger$, and we find from (2.64), $\mathcal{P}_{\sigma, \mathcal{N}}(X) = U^\dagger X U$.

Interestingly, it has recently been understood that the change of relative entropy under quantum channels places strict bounds on how well a state can be recovered. The basic idea is intuitive; if two states that were initially easily distinguishable become nearly indistinguishable under a channel, then it should be impossible to identify what the initial states were using only information about the indistinguishable output states. In particular, it can be shown [39] that there exists a recovery map $\mathcal{R}_{\sigma, \mathcal{N}}$ with $\mathcal{R}_{\sigma, \mathcal{N}} \circ \mathcal{N}(\sigma) = \sigma$ satisfying,⁷

$$F(\rho, [\mathcal{R}_{\sigma, \mathcal{N}} \circ \mathcal{N}](\rho)) \geq \exp(D(\mathcal{N}(\rho)||\mathcal{N}(\sigma)) - D(\rho||\sigma)), \tag{2.65}$$

where F is the fidelity, defined as

$$F(\rho, \sigma) := \left(\text{Tr} \left[\sqrt{\sqrt{\rho} \sigma \sqrt{\rho}} \right] \right)^2. \tag{2.66}$$

For example, the bound (2.65) holds for an explicit but complicated recovery channel called the twirled Petz map [39].

For the quantum channel that we are interested in this paper, consider a system $D \cup B = R \cup B'$ in the initial state $\rho_D \otimes \rho_B$, evolve it for a while, and then trace out a portion B' of the full system. If we take ρ_B to be some fixed state, this gives a quantum channel from D to R ,

$$\mathcal{N}(\rho_D) = \text{Tr}_{B'} \left[U (\rho_D \otimes \rho_B) U^\dagger \right]. \tag{2.67}$$

The corresponding \mathcal{N}^\dagger from R to D is then given by

$$\mathcal{N}^\dagger(\phi_R) = \text{Tr}_B \left(\rho_B U^\dagger (\phi_R \otimes \mathbb{1}_{B'}) U \right). \tag{2.68}$$

Applying (2.64) to this case and imposing a replica trick, we have

$$\mathcal{P}_{\sigma_D, \mathcal{N}} \circ \mathcal{N}(\rho_D) = \lim_{m \rightarrow -\frac{1}{2}} \sigma_D^{\frac{1}{2}} \text{Tr}_B \left[\rho_B U^\dagger (\mathcal{N}(\sigma_D)^m \mathcal{N}(\rho_D) \mathcal{N}(\sigma_D)^m \otimes \mathbb{1}_{B'}) U \right] \sigma_D^{\frac{1}{2}}. \tag{2.69}$$

⁷This was proven for Type I von Neumann algebras but was recently generalized to the Type III algebras relevant to quantum field theory that we will model [40]. This technicality will not play an important role for us.

In some of our discussion below, an alternative analytic continuation will also be useful:

$$\mathcal{P}_{\sigma_D, \mathcal{N}} \circ \mathcal{N}(\rho_D) = \lim_{\substack{n_1 \rightarrow -\frac{1}{2}, \\ n_2 \rightarrow -\frac{1}{2}}} \sigma_D^{\frac{1}{2}} \text{Tr}_B \left[\rho_B U^\dagger (\mathcal{N}(\sigma_D)^{n_1} \mathcal{N}(\rho_D) \mathcal{N}(\sigma_D)^{n_2} \otimes \mathbb{1}_{B'}) U \right] \sigma_D^{\frac{1}{2}}. \quad (2.70)$$

When we take B' to be empty, then \mathcal{N} is unitary, and \mathcal{N}^\dagger is the inverse evolution, for which the above equation gives ρ_D . In the opposite limit, with B' being the full system, we have $\mathcal{P}_{\sigma_D, \mathcal{N}} \circ \mathcal{N}(\rho_D) = \sigma_D$ for any ρ_D .

In the case where ρ_D is a pure state, the fidelity of the Petz map can be written as an overlap

$$\begin{aligned} F(\rho_D, \mathcal{P}_{\sigma_D, \mathcal{N}} \circ \mathcal{N}(\rho_D)) &= \text{Tr}_D (\rho_D \mathcal{P}_{\sigma_D, \mathcal{N}} \circ \mathcal{N}(\rho_D)) \\ &= \lim_{m \rightarrow -\frac{1}{2}} \text{Tr} \left[U \left(\sigma_D^{\frac{1}{2}} \rho_D \sigma_D^{\frac{1}{2}} \otimes \rho_B \right) U^\dagger \mathcal{N}(\sigma_D)^m \mathcal{N}(\rho_D) \mathcal{N}(\sigma_D)^m \right]. \end{aligned} \quad (2.71)$$

or a similar overlap using (2.70). Eqs. (2.69), (2.70), and (2.71) can all be evaluated by using the equilibrium approximation. If the Petz map works perfectly, without error, the fidelity will be one. At worst, the Petz map should output a random answer in which case the fidelity would be exponentially small in the entropy.

3 Equilibrium approximation for quantum-informational quantities

In addition to the Renyi and entanglement entropies, the equilibrium approximation can in principle be used to calculate any quantum informational quantities which can be defined using replicas. In this section, we consider a few quantities that will be used in the later discussion of the paper: Renyi and logarithmic negativities, relative entropy, reflected entropy, and the Petz recovery map.

3.1 Renyi and logarithmic negativities

We now describe the calculation of Renyi negativities between A_1 and A_2 by adding an auxiliary system B to A , and applying the equilibrium approximation to an equilibrated pure state $|\Psi\rangle$ describing the full system $S = A \cup B$. There is also an alternative way to calculate the negativities between A_1 and A_2 by directly applying the equilibrium approximation to the mixed state ρ_A , analogous to the discussion for the Renyi entropies in section 2.3.2.

Suppose the system is partitioned to $A \cup B$ with $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ and the subsystem A is in turn partitioned to $A = A_1 \cup A_2$ with $\mathcal{H}_A = \mathcal{H}_1 \otimes \mathcal{H}_2$. We are interested in

$$\mathcal{Z}_n^{(\text{PT})} := \text{Tr} \left(\rho_A^{T_2} \right)^n = \text{Tr}_A \left(\text{Tr}_B \left(U \rho_0 U^\dagger \right)^{T_2} \right)^n \quad (3.1)$$

and the corresponding logarithmic negativity $\mathcal{E}(A_1, A_2) = \lim_{n \rightarrow \frac{1}{2}} \log \mathcal{Z}_{2n}^{(\text{PT})}$.

Using the replica space $(\mathcal{H} \otimes \mathcal{H})^n$ we can write (3.1) as

$$\mathcal{Z}_n^{(\text{PT})} = \langle \eta_{A_1} \otimes \eta_{A_2}^{-1} \otimes e_B | \left(U \otimes U^\dagger \right)^n | \rho_0, e \rangle, \quad (3.2)$$

where $\eta = (n, n-1, \dots, 1)$ is the cyclic permutation and its inverse η^{-1} is the anti-cyclic permutation.

We can insert the identity (2.34) twice in (3.2) to arrive at

$$\mathcal{Z}_n^{(\text{PT})} = \langle \eta_{A_1} \otimes \eta_{A_2}^{-1} \otimes e_B | P_\alpha | \rho_0, e \rangle + \langle \eta_{A_1} \otimes \eta_{A_2}^{-1} \otimes e_B | Q (U \otimes U^\dagger)^n Q | \rho_0, e \rangle. \quad (3.3)$$

In the equilibrium approximation, we drop the second term in the above equation, which leads to a time-independent expression,

$$\begin{aligned} \mathcal{Z}_n^{(\text{PT})} &\approx [\mathcal{Z}_n^{(\text{PT})}]_{\text{eq approx}} := \frac{1}{Z_2^n} \sum_{\sigma, \tau} g^{\tau\sigma} \langle \eta_{A_1} \otimes \eta_{A_2}^{-1} \otimes e_B | \mathcal{I}_\alpha, \tau \rangle \langle \mathcal{I}_\alpha, \sigma | \rho_0, e \rangle \\ &= \frac{1}{Z_1^n} \sum_{\sigma, \tau} g^{\tau\sigma} \langle \eta_{A_1} \otimes \eta_{A_2}^{-1} \otimes e_B | \mathcal{I}_\alpha, \tau \rangle \\ &\approx \frac{1}{Z_1^n} \sum_{\tau} \langle \eta_{A_1} \otimes \eta_{A_2}^{-1} \otimes e_B | \mathcal{I}_\alpha, \tau \rangle. \end{aligned} \quad (3.4)$$

In the second line above we used (2.40) and in the third line, we again used the fact that in the large Z_1 limit $g_{\sigma\tau}$ can be approximated by the identity matrix. This final expression is independent of the initial state and only depends on the equilibrium density operator.

Each term in the final expression in (3.4) can be given a diagrammatic representation, as shown in figure 3. We can insert the identity to write

$$\begin{aligned} \langle \eta_{A_1} \otimes \eta_{A_2}^{-1} \otimes e_B | \mathcal{I}_\alpha, \tau \rangle &= \sum_{i_1, i'_1, \dots, i_n, i'_n} \langle \eta_{A_1} \otimes \eta_{A_2}^{-1} \otimes e_B | i_1 \bar{i}'_1 \dots i_n \bar{i}'_n \rangle \langle i_1 \bar{i}'_1 \dots i_n \bar{i}'_n | \mathcal{I}_\alpha, \tau \rangle, \\ |i_m\rangle &= |i_{m_a}\rangle_{A_1} |i_{m_{\bar{a}}}\rangle_{A_2} |i_{m_b}\rangle_B, \quad |\bar{i}'_m\rangle = |\bar{i}'_{m_a}\rangle_{A_1} |\bar{i}'_{m_{\bar{a}}}\rangle_{A_2} |\bar{i}'_{m_b}\rangle_B. \end{aligned} \quad (3.5)$$

The lower half of each diagram represents $\langle \eta_{A_1} \otimes \eta_{A_2}^{-1} \otimes e_B | i_1 \bar{i}'_1 \dots i_n \bar{i}'_n \rangle$ by connecting i_{m_a} with $i'_{\eta(m)_a}$ using dashed lines, $i_{m_{\bar{a}}}$ with $i'_{\eta^{-1}(m)_{\bar{a}}}$ using dotted lines, and i_{m_b} with i'_{m_b} using solid lines, as shown in figure 3(a). The upper half of the diagram represents $\langle i_1 \bar{i}'_1 \dots i_n \bar{i}'_n | \mathcal{I}_\alpha, \tau \rangle$, by connecting i_m with $i'_{\tau(m)}$, as shown for two examples in figure 3(b) and (c). In the resulting diagram, roughly each solid loop gives a power of d_B , each dashed loop gives a power of d_{A_1} , and each dotted loop gives a power of d_{A_2} , where d_P is the effective Hilbert space dimension of subsystem P . The number of solid, dashed, and dotted loops in a diagram is respectively equal to $k(\tau)$, $k(\eta^{-1}\tau)$, and $k(\eta\tau)$. We therefore find

$$\langle \eta_{A_1} \otimes \eta_{A_2}^{-1} \otimes e_B | \mathcal{I}_\alpha, \tau \rangle \sim d_{A_1}^{k(\eta^{-1}\tau)} d_{A_2}^{k(\eta\tau)} d_B^{k(\tau)}. \quad (3.6)$$

The sum of these powers satisfies the inequality

$$k(\eta^{-1}\tau) + k(\eta\tau) + k(\tau) \leq \begin{cases} 3n/2 + 2 & n \text{ even} \\ (3n+3)/2 & n \text{ odd} \end{cases}. \quad (3.7)$$

The permutations for which this inequality is saturated are described in appendix A.

We note

$$\langle \eta_{A_1} \otimes \eta_{A_2}^{-1} \otimes e_B | \mathcal{I}_\alpha, \tau \rangle = \langle (\mu_1\eta)_{A_1} (\mu_2\eta^{-1})_{A_2} (\mu_3)_B | \mathcal{I}_\alpha, (\mu_1\tau)_{A_1} (\mu_2\tau)_{A_2} (\mu_3\tau)_B \rangle \quad (3.8)$$

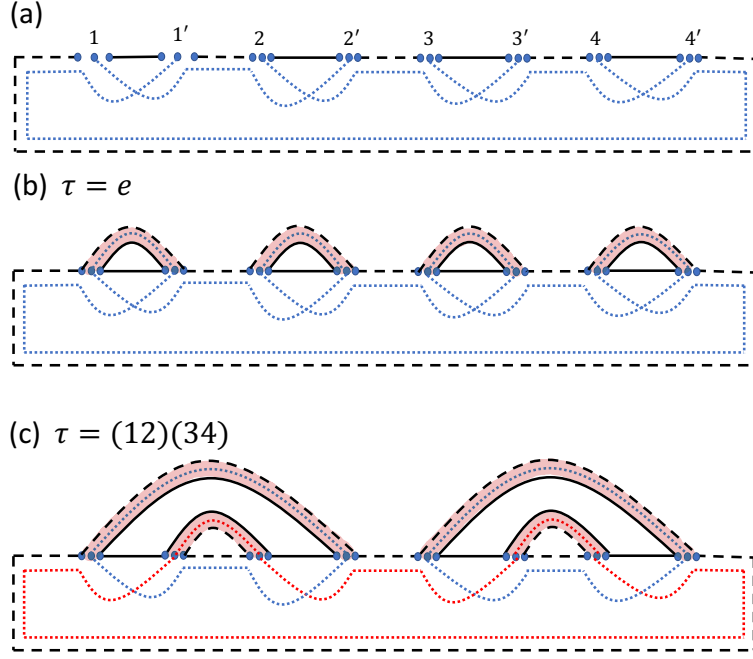


Figure 3. (a) shows the common lines of all diagrams for different terms in (3.4), and (b) and (c) show examples for two choices of τ , for the case $n = 4$. The dotted loops are shown in two different colors in (c) to make the distinction between different loops clear.

for arbitrary $\mu_{1,2,3} \in \mathcal{S}_n$, and

$$\langle \eta_{A_1} \otimes \eta_{A_2}^{-1} \otimes e_B | \mathcal{I}_\alpha, \tau \rangle = \langle (\eta\mu)_{A_1} (\eta^{-1}\mu)_{A_2} \mu_B | \mathcal{I}_\alpha, (\tau\mu)_{A_1} (\tau\mu)_{A_2} (\tau\mu)_B \rangle, \quad \mu \in \mathcal{S}_n. \quad (3.9)$$

The expression (3.4) passes various consistency checks. Firstly, we have

$$\mathcal{Z}_n^{(\text{PT})}(A_1, A_2) = \mathcal{Z}_n^{(\text{PT})}(A_2, A_1), \quad (3.10)$$

which can be shown as follows. Since η and η^{-1} have the same cycle structure, there exists some element $\sigma \in \mathcal{S}_n$ such that $\eta^{-1} = \sigma\eta\sigma^{-1}$ and $\eta = \sigma\eta^{-1}\sigma^{-1}$. We then have from (3.8)–(3.9)

$$\langle \eta_{A_1} \otimes \eta_{A_2}^{-1} \otimes e_B | \mathcal{I}_\alpha, \tau \rangle = \langle \eta_{A_1}^{-1} \otimes \eta_{A_2} \otimes e_B | \mathcal{I}_\alpha, \sigma^{-1}\tau\sigma \rangle \quad (3.11)$$

and thus (3.10) results upon summing over τ .

Next, consider $n = 2$, where $\mathcal{Z}_2^{(\text{PT})}$ should give the same answer as $\mathcal{Z}_2^{(A)}$, the second moment of the density matrix that has not been partially transposed. For $n = 2$, $\eta = \eta^{-1}$ because both are swap operators, so this check immediately passes.

Finally, consider the case where B is not present, i.e. ρ_A is a pure state

$$\mathcal{Z}_n^{(\text{PT})} \simeq \frac{1}{Z_1^n} \sum_{\tau \in \mathcal{S}_n} \langle \eta_{A_1} \otimes \eta_{A_2}^{-1} | \mathcal{I}_\alpha, \tau \rangle. \quad (3.12)$$

From (3.9) with $\mu = \eta$ we have

$$\mathcal{Z}_n^{(\text{PT})} \simeq \frac{1}{Z_1^n} \sum_{\tau \in \mathcal{S}_n} \langle \eta_{A_1}^2 \otimes e_{A_2} | \mathcal{I}_\alpha, \tau \rangle. \quad (3.13)$$

For even $n = 2m$, we have $\eta^2 = (2m-1, 2m-3, \dots, 3, 1)(2m, 2m-2, \dots, 4, 2)$. We can separate the sum over $\tau \in \mathcal{S}_{2m}$ into that over the elements of the subgroup $\mathcal{S}_m \times \mathcal{S}_m$ which permutes separately odd and even numbers, and the rest. We then find

$$\mathcal{Z}_{2m}^{(\text{PT})} \simeq (\mathcal{Z}_{m, A_1})^2 + \sum_{\tau \in \mathcal{S}_{2m} - (\mathcal{S}_m \times \mathcal{S}_m)} \langle \eta_{A_1}^2 \otimes e_{A_2} | \mathcal{I}_\alpha, \tau \rangle \quad (3.14)$$

In the replica limit, the first term gives the $n = 1/2$ Renyi entropy as expected for the pure state limit of negativity. The second term is suppressed by additional factors of Z_1 as they correspond to non-planar diagrams.

The size of the second term in (3.3), which we ignored in (3.4), can be estimated as Δ_N , where

$$\Delta_N^2 := [(\mathcal{Z}_n^{(\text{PT})})^2]_{\text{eq approx}} - ([\mathcal{Z}_n^{(\text{PT})}]_{\text{eq approx}})^2 \quad (3.15)$$

In appendix B, we show that

$$\frac{\Delta_N}{[\mathcal{Z}_n^{(\text{PT})}]_{\text{eq approx}}} \sim Z_1^{-1/2}. \quad (3.16)$$

Hence, the fluctuations around the equilibrium value of $\mathcal{Z}_n^{(\text{PT})}$ are suppressed.

We can obtain the equilibrium approximation for the logarithmic negativity using analytic continuation in cases where $\mathcal{Z}_n^{(\text{PT})}$ can be written as an analytic function of n , or, more reliably, by using the resolvent (2.27).

3.2 Relative entropy

Consider a system evolved from two possible initial states specified respectively by density operators ρ_0, σ_0 . We will assume that the support of ρ_0 lies inside that of σ_0 so that the relative entropy

$$D(\rho_0 || \sigma_0) = \text{Tr} \rho_0 \log \rho_0 - \text{Tr} \rho_0 \log \sigma_0 \quad (3.17)$$

is finite. Suppose $\rho = U \rho_0 U^\dagger$ and $\sigma = U \sigma_0 U^\dagger$ can be approximated at the macroscopic level by the same equilibrium density operator $\rho^{(\text{eq})}$. We are interested in calculating the relative entropy

$$D(\rho_A || \sigma_A) = \text{Tr}_A \rho_A \log \rho_A - \text{Tr}_A \rho_A \log \sigma_A \quad (3.18)$$

$$= \lim_{n \rightarrow 1} \frac{1}{n-1} \left(\log \text{Tr}_A [\rho_A^n] - \log \text{Tr}_A [\rho_A \sigma_A^{n-1}] \right) \quad (3.19)$$

between the reduced density operators ρ_A, σ_A of some subsystem A . The calculation of the first term was already reviewed in section 2.3, see e.g. (2.58). Here we discuss how to use the equilibrium approximation to compute the second term, which depends on two distinct density matrices.

As before, we rewrite the second term in (3.19) as a transition amplitude in the replica Hilbert space

$$\mathcal{D}_{n,A} = \text{Tr}_A \left[\rho_A \sigma_A^{n-1} \right] = \langle \eta_A \otimes e_{\bar{A}} | (U \otimes U^\dagger)^{\otimes n} | \rho_0 \otimes \sigma_0^{\otimes(n-1)}, e \rangle, \quad (3.20)$$

where the state $|\rho \otimes \sigma^{\otimes(n-1)}, \tau\rangle$ for a permutation τ is defined as

$$\langle i_1 \bar{i}'_1 i_2 \bar{i}'_2 \cdots i_n \bar{i}'_n | \rho \otimes \sigma^{\otimes(n-1)}, \tau \rangle = \rho_{i_1 i'_{\tau(1)}} \sigma_{i_2 i'_{\tau(2)}} \cdots \sigma_{i_n i'_{\tau(n)}} \quad (3.21)$$

which is inhomogeneous in the replicas. Applying the equilibrium approximation, we find

$$\mathcal{D}_{n,A} \approx \frac{1}{Z_2^n} \sum_{\tau \in \mathcal{S}_n} \langle \eta_A \otimes e_{\bar{A}} | \mathcal{I}_\alpha, \tau \rangle \langle \mathcal{I}_\alpha, \tau | \rho_0 \otimes \sigma_0^{\otimes(n-1)}, e \rangle \quad (3.22)$$

$$\begin{aligned} &= \frac{1}{Z_2^n} \sum_{\tau \in \mathcal{S}_n} \langle \eta_A \otimes e_{\bar{A}} | \mathcal{I}_\alpha, \tau \rangle \\ &\quad \times \text{Tr} \left[\mathcal{I} \rho_0 (\mathcal{I}_\alpha \sigma_0)^{m_1-1} \right] \text{Tr} [(\mathcal{I}_\alpha \sigma_0)^{m_2}] \cdots \text{Tr} [(\mathcal{I} \sigma_0)^{m_{k(\tau)}}] \end{aligned} \quad (3.23)$$

$$= \frac{1}{Z_1^n} \sum_{\tau \in \mathcal{S}_n} \langle \eta_A \otimes e_{\bar{A}} | \mathcal{I}_\alpha, \tau \rangle \frac{\text{Tr}(\rho_0 \sigma_0^{m_1-1})}{\text{Tr}(\sigma_0^{m_1})} \prod_{i=1}^{k(\tau)} \text{Tr}(\sigma_0^{m_i}) \quad (3.24)$$

where m_i is the number of elements in i -th cycle of τ , the $i = 1$ cycle is taken to be the one containing the first copy of the Hilbert space, and we have used self-consistency conditions (derived similarly as (2.57))

$$\text{Tr} \left[\mathcal{I} \rho_0 (\mathcal{I} \sigma_0)^{n-1} \right] \approx \frac{Z_2^n}{Z_1^n} \text{Tr}(\rho_0 \sigma_0^{n-1}), \quad \text{Tr} [(\mathcal{I} \sigma_0)^n] \approx \frac{Z_2^n}{Z_1^n} \text{Tr}(\sigma_0^n). \quad (3.25)$$

We now make some general comments on the structure of (3.24). In (3.24), we can divide τ 's into those with $m_1 = 1$, and those with $m_1 > 1$. Denoting the two sets respectively as $\mathcal{S}_{n,1}$ and $\mathcal{S}_{n,2}$, we have

$$\mathcal{D}_{n,A} = \sum_{\tau \in \mathcal{S}_{n,1}} \mathcal{Z}_{n,A}(\tau) + \sum_{\tau \in \mathcal{S}_{n,2}} \frac{\text{Tr}(\rho_0 \sigma_0^{m_1-1})}{\text{Tr}(\sigma_0^{m_1})} \mathcal{Z}_{n,A}(\tau) \quad (3.26)$$

where

$$\mathcal{Z}_{n,A}(\tau) = \frac{1}{Z_1^n} \langle \eta_A \otimes e_{\bar{A}} | \mathcal{I}_\alpha, \tau \rangle \prod_{i=1}^{k(\tau)} \text{Tr}(\sigma_0^{m_i}) \quad (3.27)$$

is the contribution of τ to the Renyi partition function for A with initial state σ_0 .

From the discussion around (2.62), we then conclude that when A is small $\tau = e$ dominates,⁸ giving

$$\mathcal{D}_{n,A} = \mathcal{Z}_{n,A}^{(\text{eq})}, \quad \text{for } S_{n,A}^{(\text{eq})} \ll S_{n,\bar{A}}^{(\text{eq})} + S_n(\sigma_0). \quad (3.28)$$

For A to be sufficiently large we expect the first term is dominated by η_{n-1} which cyclicly permutes $2, \cdots, n$ and gives a contribution $\mathcal{Z}_{n,A}(\eta_{n-1})$, while the dominant permutation for

⁸Note $\text{Tr}(\rho_0 \sigma_0^{m-1}) < \text{Tr}(\sigma_0^m)$.

the second term is η , giving a contribution $\text{Tr}(\rho_0 \sigma_0^{n-1}) \mathcal{Z}_{n,A}^{(\text{eq})}$. From our general discussion $\mathcal{Z}_{n,A}(\eta_{n-1})$ is smaller than $\mathcal{Z}_{n,A}^{(\text{eq})}$ by at least a factor Z_1^{-1} . Thus for the case $\text{Tr}(\rho_0 \sigma_0^{n-1})$ is not too small (i.e. much larger than Z_1^{-1}) we then have

$$\mathcal{D}_{n,A} \approx \text{Tr}(\rho_0 \sigma_0^{n-1}) \mathcal{Z}_{n,A}^{(\text{eq})}, \quad \text{for } S_{n,A}^{(\text{eq})} \gg S_{n,A}^{(\text{eq})} + S_n(\sigma_0). \quad (3.29)$$

Now combining the above discussion with (2.62), and assuming that we can analytically continue to $n = 1$, we find (3.19) can be written as

$$D(\rho_A || \sigma_A) \simeq \begin{cases} 0, & S_A^{(\text{eq})} \ll S_A^{(\text{eq})} + S(\rho_0) \\ S_A^{(\text{eq})} - S_A^{(\text{eq})} - S(\rho_0), & S_A^{(\text{eq})} + S(\rho_0) \ll S_A^{(\text{eq})} \ll S_A^{(\text{eq})} + S(\sigma_0) \\ D(\rho_0 || \sigma_0), & S_A^{(\text{eq})} \gg S_A^{(\text{eq})} + S(\sigma_0) \end{cases}. \quad (3.30)$$

where we have assumed $S(\sigma_0) \gg S(\rho_0)$. The above expressions are intuitively reasonable. When subregion A is sufficiently small, the two density matrices are entirely indistinguishable, a manifestation of thermalization in an isolated quantum system. Once we move beyond the first regime of (3.30), the state σ_0 becomes important. In particular, the relative entropy rises from 0 to $D(\rho_0 || \sigma_0)$ where it plateaus; as we gain information, the density matrices become more and more distinguishable. The relative entropy never increases beyond $D(\rho_0 || \sigma_0)$ due to the monotonicity of relative entropy under quantum channels.

Less conservatively, if we trust the analytic continuations of dominant permutations, we find the following sharper version of (3.30):

$$D(\rho_A || \sigma_A) \simeq \begin{cases} 0, & S_A^{(\text{eq})} < S_A^{(\text{eq})} + S(\rho_0) \\ S_A^{(\text{eq})} - S_A^{(\text{eq})} - S(\rho_0), & S_A^{(\text{eq})} + S(\rho_0) < S_A^{(\text{eq})} < S_A^{(\text{eq})} + S(\rho_0) + D(\rho_0 || \sigma_0) \\ D(\rho_0 || \sigma_0), & S_A^{(\text{eq})} > S_A^{(\text{eq})} + S(\rho_0) + D(\rho_0 || \sigma_0) \end{cases}. \quad (3.31)$$

We also found numerical evidence for this equation in small spin chains and it is consistent with an infinite temperature result we derive exactly in section 6. It would be interesting to test this equation more generally.

3.3 The fidelity of the Petz map

We now turn to the calculation of the fidelity (2.71) of the Petz map using the equilibrium approximation. Recall that in (2.71), we assume ρ_D is pure.

We can write (2.71) as a transition amplitude in the replica space as

$$F(\rho_D, \mathcal{P}_{\sigma_D, \mathcal{N}} \circ \mathcal{N}(\rho_D)) = \lim_{m \rightarrow -\frac{1}{2}} \langle \eta_R \otimes e_{B'} | (U \otimes U^\dagger)^{\otimes (2m+2)} | \chi, e \rangle := \lim_{m \rightarrow -\frac{1}{2}} F_m, \quad (3.32)$$

where χ has the form

$$\chi = \chi_1 \otimes \chi_2^{\otimes m} \otimes \chi_3 \otimes \chi_2^{\otimes m}, \quad \chi_1 = \sigma_D^{\frac{1}{2}} \rho_D \sigma_D^{\frac{1}{2}} \otimes \rho_B, \quad \chi_2 = \sigma_D \otimes \rho_B, \quad \chi_3 = \rho_D \otimes \rho_B. \quad (3.33)$$

Applying the equilibrium approximation (assuming the equilibrated states of $\chi_{1,2,3}$ can all be described by the same macroscopic state \mathcal{I}_α)

$$F_m = \frac{1}{Z_2^{2m+2}} \sum_{\tau \in \mathcal{S}_{2m+2}} \langle \eta_R \otimes e_{B'} | \mathcal{I}_\alpha, \tau \rangle \langle \mathcal{I}_\alpha, \tau | \chi, e \rangle := \sum_{\tau \in \mathcal{S}_{2m+2}} F_m(\tau) \quad (3.34)$$

where

$$\langle \mathcal{I}_\alpha, \tau | \chi, e \rangle = \prod_{i=1}^{k(\tau)} \text{Tr} \left[(\mathcal{I}_\alpha \chi_1)^{a_i} (\mathcal{I}_\alpha \chi_2)^{b_i} (\mathcal{I}_\alpha \chi_3)^{c_i} (\mathcal{I}_\alpha \chi_2)^{d_i} \right] \quad (3.35)$$

$$= \frac{Z_2^{2m+2}}{Z_1^{2m+2}} \prod_{i=1}^{k(\tau)} \text{Tr} \left[\chi_1^{a_i} \chi_2^{b_i} \chi_3^{c_i} \chi_2^{d_i} \right] \quad (3.36)$$

where a_i, b_i, c_i, d_i denote the number of appearances of $\chi_{1,2,3}$ in (3.33) in the i -th cycle of τ .⁹ Clearly $a_1 = 1$ and $a_{i>1} = 0$. In the second line we again have used the consistency conditions as in (2.57).

The contribution from the identity permutation can be written as

$$F_m(e) = Z_{2m+2,R}^{(\text{eq})} \text{Tr}(\sigma_D \rho_D) = Z_{2m+2,R}^{(\text{eq})} F(\rho_D, \sigma_D), \quad (3.37)$$

which we expect to dominate for R much smaller than B' . Analytically continuing the above expression to $m = -\frac{1}{2}$ we then find

$$F(\rho_D, \mathcal{P}_{\sigma_D, \mathcal{N}} \circ \mathcal{N}(\rho_D)) = F(\rho_D, \sigma_D), \quad (3.38)$$

which is consistent with statement below (2.69) that in the limit where R is small, the Petz map simply gives σ_D for any ρ_D .

For R to be much larger than B' , we expect the contribution from $\tau = \eta$ dominates giving

$$F_m(\eta) = Z_{2m+2,B'}^{(\text{eq})} \text{Tr}(\rho_D \sigma_D^{m+\frac{1}{2}} \rho_D \sigma_D^{m+\frac{1}{2}}) \text{Tr} \rho_B^{2m+2}. \quad (3.39)$$

Analytically continuing to $m = -\frac{1}{2}$ we have

$$F(\rho_D, \mathcal{P}_{\sigma_D, \mathcal{N}} \circ \mathcal{N}(\rho_D)) = 1. \quad (3.40)$$

which is consistent with that in the limit B' is empty, the Petz map becomes the identity map.

There is a crossover between the above two extremes where other permutations could become important. This crossover behaviour will be dependent on the choice of \mathcal{I}_α , and we explain how to calculate it for particular choices of \mathcal{I}_α in section 6.

Note that starting from (2.70), we could alternatively use the equilibrium approximation to get expressions involving permutations in $\mathcal{S}_{n_1+n_2+2}$ instead of \mathcal{S}_{2m+2} by a similar series of steps, which give the same results for $F(\rho_D, \mathcal{P}_{\sigma_D, \mathcal{N}} \circ \mathcal{N}(\rho_D))$ on analytic continuation to $n_1 \rightarrow -\frac{1}{2}$, $n_2 \rightarrow -\frac{1}{2}$.

⁹Note that in general, the factors of $\chi_1^{a_i}$, $\chi_2^{b_i}$, $\chi_3^{c_i}$, and $\chi_2^{d_i}$ can also appear in other orders.

3.4 Reflected entropy

Another interesting quantity that may be computed using the replica trick is the reflected entropy [41]. Consider $A = A_1 \cup A_2$ in a mixed state ρ_A and its canonical purification $\rho_A \rightarrow |\sqrt{\rho_A}\rangle_{AA^*} \in \mathcal{H}_A \otimes \mathcal{H}_{A^*}$ with $A^* = A_1^* \cup A_2^*$. The reflected entropy is then defined to be the von Neumann entropy of $A_1 A_1^*$ in this pure state.

To calculate the reflected entropy using replicas, we consider

$$\mathcal{R}_{n,m} = -\frac{1}{n-1} \log \text{Tr}_{A_1 A_1^*} \left(\text{Tr}_{A_2 A_2^*} |\rho_A^m\rangle \langle \rho_A^m| \right)^n := -\frac{1}{n-1} \log \mathcal{Y}_{n,m} \quad (3.41)$$

where $|\rho_A^m\rangle \in \mathcal{H}_A \otimes \mathcal{H}_{A^*}$ is defined as

$$\langle i\bar{i}' | \rho_A^m \rangle = (\rho_A^m)_{ii'} \quad (3.42)$$

where $|i\rangle$ is a basis for \mathcal{H}_A . The reflected entropy is then given by

$$S_{\mathcal{R}} := \lim_{n \rightarrow 1} S_{\mathcal{R}}^{(n)} = \lim_{n \rightarrow 1} \lim_{m \rightarrow \frac{1}{2}} \mathcal{R}_{n,m} \quad (3.43)$$

where $S_{\mathcal{R}}^{(n)}$ are the Renyi reflected entropies.

We take A to be embedded in a larger system $S = A \cup B$ with

$$\rho_A = \text{Tr}_B(U \rho_0 U^\dagger) \quad (3.44)$$

where ρ_0 is a pure state. We can then write $\mathcal{Y}_{n,m}$ as a transition amplitude in $\mathcal{H}_S^{\otimes 2mn}$

$$\mathcal{Y}_{n,m} = \left\langle (\sigma_1)_{A_1} \otimes (\sigma_2)_{A_2} \otimes e_B | (U \otimes U^\dagger)^{2mn} | \rho_0, e \right\rangle \quad (3.45)$$

where σ_1 and σ_2 denote the following permutations in \mathcal{S}_{2mn}

$$\sigma_1 = \prod_{k=1}^n (k, k+n, \dots, k+n(m-1), k+1+nm, \dots, k+1+n(2m-1)), \quad (3.46)$$

$$\sigma_2 = \prod_{k=1}^n (k, k+n, \dots, k+n(2m-1)). \quad (3.47)$$

In the equilibrium approximation we then have

$$\mathcal{Y}_{n,m} \approx \frac{1}{Z_1^{2nm}} \sum_{\tau \in \mathcal{S}_{2nm}} \langle (\sigma_1)_{A_1} \otimes (\sigma_2)_{A_2} \otimes e_B | \mathcal{I}_\alpha, \tau \rangle. \quad (3.48)$$

When B is sufficiently larger than A , the identity element will dominate the sum, leading to

$$\mathcal{Y}_{n,m} = \frac{(Z_{2m,A})^n}{(Z_{1,A})^{2nm}}. \quad (3.49)$$

Plugging this into (3.43), we immediately find all Renyi reflected entropies to be zero.

When A is sufficiently large, there will be correlations between A_1 and A_2 . When A_1 is much larger than A_2 , we can be confident that $\tau = \sigma_1^{-1}$ will dominate the sum. Note that

$$\sigma_2 \sigma_1^{-1} = (1, 2, \dots, n)(n(m+1), n(m+1) - 1, \dots, nm + 1), \quad (3.50)$$

which has two cycles of length n and $n(2m - 2)$ trivial cycles. This leads to

$$\mathcal{Y}_{n,m} = \frac{(Z_{n,A_2})^2 (Z_{2m,B})^n}{(Z_{1,A_2})^{2n} (Z_{1,B})^{2nm}}. \quad (3.51)$$

Taking $m \rightarrow \frac{1}{2}$, we find that the Renyi reflected entropy is given by twice the equilibrium Renyi entropy of A_2 .

$$S_{\mathcal{R}}^{(n)} = \frac{2}{1-n} \log \frac{Z_{n,A_2}}{(Z_{1,A_2})^n} = 2S_{n,A_2}^{(\text{eq})}. \quad (3.52)$$

In the same way, when $\tau = \sigma_{\bar{g}}^{-1}$ dominates,

$$S_{\mathcal{R}}^{(n)} = 2S_{n,A_1}^{(\text{eq})}. \quad (3.53)$$

It is natural to ask whether there are additional important permutations interpolating between these limits. Indeed, such permutations are identified in refs. [42, 43] where the random tensor network (infinite temperature) result is studied in detail using the reflected entropy spectrum. However, in the $n \rightarrow 1$ limit, the naive analytic continuations of the previously discussed permutations always give the correct answer to leading order. Here, we simply assume that this remains to be the case at finite temperature when $n \rightarrow 1$ and leave a rigorous justification of this assumption to future work. To find the transition points between the phases, we maximize the partition functions at $n \sim 1$, giving

$$S_{\mathcal{R}} = \begin{cases} 0, & S_A^{(\text{eq})} < S_B^{(\text{eq})} \\ 2S_{A_1}^{(\text{eq})}, & S_{A_1}^{(\text{eq})} < S_{A_2}^{(\text{eq})}, S_A^{(\text{eq})} > S_B^{(\text{eq})} \\ 2S_{A_2}^{(\text{eq})}, & S_{A_2}^{(\text{eq})} < S_{A_1}^{(\text{eq})}, S_A^{(\text{eq})} > S_B^{(\text{eq})} \end{cases}. \quad (3.54)$$

The interplay between the two replica numbers is delicate for the analytic continuation, and the order of limits in (3.43) is an additional assumption in the computation. For a more reliable analysis, one must evaluate the reflected entropy spectrum.

4 Entanglement phase diagram of an equilibrated mixed state

4.1 General setup

We would like to explore the entanglement structure of a system A in a mixed state ρ_A , which is in a macroscopic equilibrium but can be far in trace distance from the usual equilibrium density operators describing thermal ensembles. For this purpose, we consider various bi-partite quantum entanglement measures between a subsystem A_1 and its complement A_2 in A , including the Renyi and logarithmic negativities, as well as the mutual

information and Renyi mutual informations. As reviewed earlier, the logarithmic negativity and the mutual information are of particular interest: the former because it is sensitive only to quantum entanglement correlations and gives a lower bound on the PPT entanglement cost, $E_c^{(\text{ppt,exact})} \geq \mathcal{E}$, and the latter because it gives an upper bound on the PPT distillable entanglement, $\frac{1}{2}I \geq E_d$. ρ_A can be characterized by an infinite number of parameters: $z_n = \text{Tr} \rho_A^n$, $n = 2, 3, \dots$. In principle, the entanglement structure between A_1, A_2 can depend on the relations among these infinite number of parameters in a complicated way. In other words, if we use an “entanglement phase diagram” to characterize different entanglement structures, the diagram is in principle drawn on an infinite dimensional space. For ρ_A in macroscopic equilibrium, we expect that the story should be much simpler in the thermodynamic limit, not depending on microscopic details of ρ_A . Our goal is to extract the universal behavior of the entanglement structure in this regime.

A general setup for exploring entanglement correlations in a mixed state in A is to imagine that A is embedded in a larger system $S = A \cup B$, with the total system S in a pure state $|\Psi\rangle$ in macroscopic equilibrium, and $\rho_A = \text{Tr}_B |\Psi\rangle \langle \Psi|$. In many questions of interest, such a B naturally exists. For example, consider the evaporation of a black hole formed from gravitational collapse of a star in a pure state, a system of central interest for the black hole information paradox. If we take A to be the collection of the Hawking radiation and ask about the entanglement correlations between different parts of the radiation, the corresponding B is the black hole emitting the Hawking radiation. If quantum gravity is compatible with the usual rules of quantum mechanics, the full system of the black hole plus the Hawking radiation would be in a pure state in macroscopic equilibrium. Alternatively, we may simply view B as an auxiliary system used to purify ρ_A .

In this setup, with $A_1 \cup A_2 \cup B$ in an equilibrated pure state $|\Psi\rangle$, the equilibrium approximation reviewed in section 2.3 can be generalized to calculate the Renyi and logarithmic negativities between A_1 and A_2 , as we discussed in section 3.1.

We will show that the entanglement structure, i.e. the qualitative behaviors of the negativities and mutual informations between A_1 and A_2 , can be characterized by the equilibrium density operator $\rho^{(\text{eq})}$ (2.30), and two parameters describing the relative sizes of A_1, A_2 and of A, B . Suppose the system A has a volume¹⁰ V_A . We denote the von Neumann entropy of ρ_A as S_A , with the entropy density given by

$$s_A = \frac{S_A}{V_A}. \tag{4.1}$$

We are interested in the thermodynamic limit $V_A \rightarrow \infty$ with s_A finite. We will consider $\mathcal{E}(A_1, A_2)$ and $I(A_1, A_2)$ at leading order in the thermodynamic limit. In this limit, (2.39) and (3.4) can both be approximated by terms from a subset of permutations τ , which give the dominant contribution. These sets of permutations can change as we vary two parameters

$$\lambda := \frac{S_{A_1}^{(\text{eq})}}{S_A^{(\text{eq})}}, \quad c := \frac{S_A^{(\text{eq})}}{S_A^{(\text{eq})} + S_B^{(\text{eq})}}, \tag{4.2}$$

¹⁰For a lattice system V_A correspond to the number of sites. For a system with no spatial extent, such as the SYK model, V_A corresponds to the number of fermions in A .

where $S_{A_1, A, B}^{(\text{eq})}$ are respectively the von Neumann entropies for A_1, A, B in the state $\rho^{(\text{eq})}$. These parameters can be seen as a way of measuring the relative sizes of the subsystems in the general case where the system $S = A \cup B$ is inhomogeneous; when the full system is homogeneous, λ and c are simply the volume fractions of various subsystems, $\lambda = V_{A_1}/V_A$ and $c = V_A/(V_A + V_B)$. The change in the dominant contribution on varying c and λ leads to qualitative changes in the behaviors of $\mathcal{Z}_n^{(\text{PT})}$ and $\mathcal{Z}_{n,R}$, and correspondingly of $\mathcal{E}(A_1, A_2)$ and $I(A_1, A_2)$. We refer to such changes as entanglement phase transitions.

4.2 Entanglement structure at infinite temperature

We now proceed to apply the techniques developed in section 3.1 to calculate the negativities between A_1 and A_2 in various situations. We first consider the case where the system has a finite-dimensional Hilbert space and is sufficiently excited that it can be treated as being at infinite temperature. Here, we find a universal entanglement phase structure which is independent of any details of A or B . The structure also coincides with that obtained from the Haar average of a random state [9].

The dimensions of the Hilbert spaces for A, A_1, A_2 and B will be denoted respectively as d_A, d_1, d_2, d_B . Assuming the system is homogeneous, we then have $\log d_A = V_A \log q$ where q is the dimension of the Hilbert space at a single site. The parameters λ and c (4.2) can also be written in terms of dimensions of various Hilbert spaces as

$$S_A^{(\text{eq})} = \log d_A, \quad \lambda := \frac{\log d_1}{\log d_A}, \quad c := \frac{\log d_A}{\log d_A + \log d_B}, \quad d_A = d_1 d_2. \quad (4.3)$$

We will also denote $S_0 := \log d_A + \log d_B$. Taking $\mathcal{I}_\alpha = \mathbf{1}$ in (3.4) and (2.39), we get the following approximations for the quantities $\mathcal{Z}_n^{(\text{PT})}(A_1, A_2)$ and $\mathcal{Z}_n^{(A)}$:

$$\mathcal{Z}_n^{(\text{PT})}(A_1, A_2) \approx \frac{1}{e^{nS_0}} \sum_{\tau} e^{S_0 \mathcal{A}(\tau)}, \quad \mathcal{Z}_n^{(A)} \approx \frac{1}{e^{nS_0}} \sum_{\tau} e^{S_0 \mathcal{B}(\tau)}, \quad (4.4)$$

where

$$\mathcal{A}(\tau) = c\lambda k(\eta^{-1}\tau) + c(1-\lambda)k(\eta\tau) + (1-c)k(\tau), \quad (4.5)$$

$$\mathcal{B}(\tau) = c k(\eta^{-1}\tau) + (1-c)k(\tau). \quad (4.6)$$

For any choice of c, λ , there is some set of permutations for which $\mathcal{A}(\tau)$ is maximized, and some set for which $\mathcal{B}(\tau)$ is maximized. Let τ_m, τ'_m respectively refer to any elements of these sets. Then, in the thermodynamic limit, we have

$$\log \mathcal{Z}_n^{(\text{PT})} \approx S_0 \mathcal{A}(\tau_m) - nS_0, \quad (4.7)$$

and

$$R_n(A_1, A_2) \approx b_n S_0 (\mathcal{A}(\tau_m) - \mathcal{B}(\tau'_m)). \quad (4.8)$$

The set of permutations that maximize $\mathcal{B}(\tau)$ was reviewed earlier in section 2.3, with $\tau'_m = e$ for $c < 1/2$, and $\tau'_m = \eta$ for $c > 1/2$.

To maximize $\mathcal{A}(\tau)$, it is convenient to write (4.5) in a few different ways,

$$\mathcal{A}(\tau) = c\lambda \left(k(\eta^{-1}\tau) + k(\tau) \right) + c(1-\lambda) \left(k(\eta\tau) + k(\tau) \right) + (1-2c)k(\tau) \quad (4.9)$$

$$= (2c\lambda - 1)k(\eta^{-1}\tau) + c(1-\lambda) \left(k(\eta\tau) + k(\eta^{-1}\tau) \right) + (1-c) \left(k(\tau) + k(\eta^{-1}\tau) \right) \quad (4.10)$$

$$= c\lambda \left(k(\eta^{-1}\tau) + k(\eta\tau) \right) + (2c(1-\lambda) - 1)k(\eta\tau) + (1-c) \left(k(\tau) + k(\eta\tau) \right) \quad (4.11)$$

$$= \left(\frac{1}{2} - c\lambda \right) \left(k(\eta\tau) + k(\tau) \right) + \left(\frac{1}{2} - c(1-\lambda) \right) \left(k(\eta^{-1}\tau) + k(\tau) \right) \\ + \left(c - \frac{1}{2} \right) \left(k(\eta\tau) + k(\eta^{-1}\tau) \right). \quad (4.12)$$

Different expressions above are convenient for different ranges of parameters c and λ , from which we find three entanglement phases, which are shown in figure 4:

1. Phase of no entanglement.¹¹

For $c < \frac{1}{2}$, i.e. $S_A^{(\text{eq})} < S_B^{(\text{eq})}$, all coefficients in (4.9) are positive and $\mathcal{A}(\tau)$ is maximized by having $k(\eta^{-1}\tau) + k(\tau)$, $k(\eta\tau) + k(\tau)$ and $k(\tau)$ all reach their maximum values simultaneously. This happens for $\tau = e$, which gives

$$R_n(A_1, A_2) = 0, \quad n \geq 3, \quad \lim_{n \rightarrow 2} R_n(A_1, A_2) = 0, \\ \log \mathcal{Z}_n^{(\text{PT})} = S_A^{(\text{eq})} (1 - n), \quad n \geq 2. \quad (4.13)$$

By analytically continuing either R_{2m} or $\log \mathcal{Z}_{2m}^{(\text{PT})}$, we find that the logarithmic negativity is given by

$$\mathcal{E}(A_1, A_2) = \lim_{m \rightarrow \frac{1}{2}} R_{2m} = \lim_{m \rightarrow \frac{1}{2}} \log \mathcal{Z}_{2m}^{(\text{PT})} = 0. \quad (4.14)$$

Furthermore, we find that all Renyi mutual informations vanish

$$I_n(A_1, A_2) = 0, \quad n \geq 2, \quad I = \lim_{n \rightarrow 1} I_n(A_1, A_2) = 0. \quad (4.15)$$

Since the equilibrium approximation calculates only the leading order contribution of order $O(\log d_A)$, in this phase there is no extensive entanglement. Both the negativities and mutual information may have nontrivial higher order sub-extensive contributions.

It is quite intuitive that in this case there is no entanglement between any subsystems. In the language of purification, all degrees of freedom of A are maximally entangled with those in B , and from the monogamy of entanglement, there is no entanglement within A .

¹¹“No entanglement” here should be understood as no “volume” entanglement, i.e. there is no contribution at the order of $O(\log d_A)$.

2. Maximally entangled phase.

For $c > \frac{1}{2}$, $\lambda > \frac{1}{2c}$, i.e. $S_B^{(\text{eq})} < S_A^{(\text{eq})}$, $S_{A_2}^{(\text{eq})} < \frac{1}{2}(S_A^{(\text{eq})} - S_B^{(\text{eq})})$, all coefficients in (4.10) are positive, and $\mathcal{A}(\tau)$ is maximized by having $k(\eta^{-1}\tau)$, $k(\eta\tau) + k(\eta^{-1}\tau)$ and $k(\tau) + k(\eta^{-1}\tau)$ reach maximum values simultaneously. This happens for $\tau = \eta$, which gives

$$R_n(A_1, A_2) = \log d_2 = S_{A_2}^{(\text{eq})} \quad n \geq 3, \quad \lim_{n \rightarrow 2} R_n(A_1, A_2) = \log d_2 = S_{A_2}^{(\text{eq})},$$

$$\log \mathcal{Z}_n^{(\text{PT})} = \begin{cases} -(n-2)S_{A_2}^{(\text{eq})} - (n-1)S_B^{(\text{eq})} & n \geq 2, n \text{ even} \\ -(n-1)(S_{A_2}^{(\text{eq})} + S_B^{(\text{eq})}) & n \geq 3, n \text{ odd} \end{cases} \quad (4.16)$$

By analytic continuation, we get

$$\mathcal{E}(A_1, A_2) = \lim_{m \rightarrow \frac{1}{2}} R_{2m} = \lim_{m \rightarrow \frac{1}{2}} \log \mathcal{Z}_{2m}^{(\text{PT})} = \log d_2 = S_{A_2}^{(\text{eq})}. \quad (4.17)$$

For this parameter range, $V_A > V_B$, $V_{A_2} < \frac{1}{2}(V_A - V_B) < V_{A_1}$, (4.17) is the maximal value of $\mathcal{E}(A_1, A_2)$ can have, and implies that A_2 is maximally entangled with A_1 . We also have $V_{A_2} + V_B < \frac{1}{2}(V_A + V_B) < V_{A_1}$, i.e. the effective number of degrees of freedom in A_2 and B together is smaller than that in A_1 . Thus both A_2 and B should be maximally entangled with A_1 and there should be no entanglement between A_2 and B .

Similarly for $c > \frac{1}{2}$, $\lambda < 1 - \frac{1}{2c}$, i.e. $S_B^{(\text{eq})} < S_A^{(\text{eq})}$, $S_{A_1}^{(\text{eq})} < \frac{1}{2}(S_A^{(\text{eq})} - S_B^{(\text{eq})})$, from (4.11), the dominant permutation for $\mathcal{A}(\tau)$ is η^{-1} and A_1 is maximally entangled with A_2 , with R_n and $\log \mathcal{Z}_n^{(\text{PT})}$ obtained by exchanging A_1 and A_2 in (4.16) and (4.17).

From the equilibrium approximation for the mutual information, we find that for both range parameters

$$I_n(A_1, A_2) = 2 \min(S_{A_1}^{(\text{eq})}, S_{A_2}^{(\text{eq})}) \quad n \geq 2,$$

$$I = \lim_{n \rightarrow 1} I_n(A_1, A_2) = 2 \min(S_{A_1}^{(\text{eq})}, S_{A_2}^{(\text{eq})}). \quad (4.18)$$

which is also consistent with the maximally entangled interpretation given.

3. Entanglement saturation phase.

For $c > \frac{1}{2}$, $1 - \frac{1}{2c} < \lambda < \frac{1}{2c}$, i.e. $S_B^{(\text{eq})} < S_A^{(\text{eq})}$, $\frac{1}{2}(S_A^{(\text{eq})} - S_B^{(\text{eq})}) < S_{A_1}^{(\text{eq})} < \frac{1}{2}(S_A^{(\text{eq})} + S_B^{(\text{eq})})$, from (4.12), $\mathcal{A}(\tau)$ is maximized by permutations which maximize simultaneously $k(\tau) + k(\eta^{-1}\tau)$, $k(\tau) + k(\eta\tau)$, and $k(\eta^{-1}\tau) + k(\eta\tau)$. For even n , $\mathcal{A}(\tau)$ is maximized by the set $\{\tau\}^*$ of non-crossing permutations which consist only of 2-cycles.¹² For odd n , the dominant permutations for $\mathcal{A}(\tau)$ are the set $\{\tau\}_{\text{odd}}^*$ of non-crossing permutations which consist of $\frac{n-1}{2}$ 2-cycles and one 1-cycle. See appendix A for more details. We then find that¹³

$$R_n(A_1, A_2) = \frac{1}{2}(S_A^{(\text{eq})} - S_B^{(\text{eq})}), \quad n \geq 2, \quad (4.19)$$

$$\log \mathcal{Z}_n^{(\text{PT})} = \begin{cases} -(\frac{n}{2} - 1)S_A^{(\text{eq})} - \frac{n}{2}S_B^{(\text{eq})}, & n \geq 2, n \text{ even} \\ -\frac{n-1}{2}(S_A^{(\text{eq})} + S_B^{(\text{eq})}), & n \geq 3, n \text{ odd} \end{cases}. \quad (4.20)$$

¹²A k -cycle refers to a cycle with k elements.

¹³Note that in the discussion here, we consider only the leading order $O(V)$ contribution to R_n , ignoring the $O(1)$ term coming from the degeneracy of different permutations.

From analytic continuation, the logarithmic negativity is given by

$$\mathcal{E}(A_1, A_2) = \lim_{m \rightarrow \frac{1}{2}} R_{2m} = \lim_{m \rightarrow \frac{1}{2}} \log \mathcal{Z}_{2m}^{(\text{PT})} = \frac{1}{2}(S_A^{(\text{eq})} - S_B^{(\text{eq})}). \quad (4.21)$$

The mutual information is given by

$$I_n(A_1, A_2) = S_A^{(\text{eq})} - S_B^{(\text{eq})}, \quad n \geq 2, \quad I = \lim_{n \rightarrow 1} I_n(A_1, A_2) = S_A^{(\text{eq})} - S_B^{(\text{eq})}. \quad (4.22)$$

Both the negativity and mutual information depend only on the difference $S_A^{(\text{eq})} - S_B^{(\text{eq})}$, and do not change as we vary the size of A_1 (as long as we stay in the aforementioned parameter range). For this reason, this is referred to as the entanglement saturation (ES) phase. Also notice that R_n and \mathcal{E} are half of the corresponding values for the mutual information as in a pure state, even though here ρ is mixed.

In terms of purification, in this range parameters, we have $V_A > V_B$, $\frac{1}{2}(V_A - V_B) < V_{A_1}, V_{A_2} < \frac{1}{2}(V_A + V_B)$, and the sum of any two V_{A_1}, V_{A_2}, V_B is larger than the third. Thus, any two of the systems are entangled with each other. There is a simple intuitive interpretation of (4.21) and (4.22) in terms of mutual bi-partite entanglement: since $V_B < V_A$, $\log d_B = S_B^{(\text{eq})}$ degrees of freedom in A are entanglement with B , and the remaining $\log d_A - \log d_B = S_A^{(\text{eq})} - S_B^{(\text{eq})}$ are entangled between A_1 and A_2 , resulting in (4.21) and (4.22).

We should emphasize, however, that this “mechanical” way of assigning entanglement likely does not reflect the genuine entanglement structure of the system in this phase. A state of mostly bi-partite entanglement has been proposed in [44] which satisfies the relation $S_R(A_1, A_2) - I(A_1, A_2) = 0$ at leading order [45], which in our current context means the absence of a volume law contribution. From the discussion of section 3.4, we find at leading order in the infinite temperature case

$$S_R(A_1, A_2) - I(A_1, A_2) = \begin{cases} 0 & \text{NE phase} \\ 0 & \text{ME phase} \\ S_B^{(\text{eq})} - |S_{A_1}^{(\text{eq})} - S_{A_2}^{(\text{eq})}| & \text{ES phase} \end{cases}. \quad (4.23)$$

This gives some hint that the ES phase likely has significant multi-partite entanglement.

In the above discussion, we obtained the behavior of \mathcal{E} by analytic continuation in n , but it can be checked that the same results follow from calculating the full negativity spectrum using the resolvent. See appendix C.2.

4.3 Entanglement phase structure at a finite temperature: general discussion

The infinite temperature case only applies to a system with a finite dimensional Hilbert space at sufficiently high energies or without energy conservation. When the energy of a system is not large enough, or if a system has an infinite dimensional Hilbert space such as in a field theory, energy constraints have to be imposed.

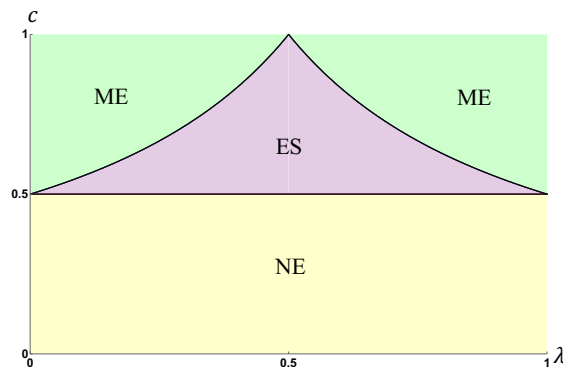


Figure 4. Entanglement phase diagram for the infinite temperature equilibrated pure state.

Now there are many more possibilities for $\rho^{(\text{eq})}$, which depend on the ensembles we choose and properties of B . We may view the different possibilities as giving different universality classes. In this subsection, we summarize the general structure and the main results, leaving explicit discussions of specific choices of $\rho^{(\text{eq})}$ to subsequent subsections.

At infinite temperature, the Renyi negativities R_n , logarithmic negativity \mathcal{E} , Renyi mutual information I_n , and mutual information I all have the same n -independent phase structure. At finite temperature, in general each of these infinite number of quantities can give rise to a different phase diagram as a function of the parameters c and λ defined in (4.2). This reveals intricate patterns of entanglement structure,¹⁴ but also makes it all but impossible to draw a phase diagram which reflects the behavior of all these quantities. Of particular interest is the behavior of logarithmic negativity \mathcal{E} , as it directly reflects quantum entanglement correlations, and as reviewed in section 2, can be used to bound the PPT entanglement cost. Furthermore, its relation with mutual information I can shed light on the irreversibility of the preparation of a mixed state by different types of operations. A significant technical complication at finite temperature is that the extraction of the logarithmic negativity using analytic continuation becomes a priori unreliable for various parameter ranges due to non-uniform dependence on n of R_n . The resolvent is needed to calculate \mathcal{E} , which fortunately can be done for some choices of $\rho^{(\text{eq})}$ and can be used to illustrate the general structure. Here is a summary of the main features we observe for finite temperature ensembles:

1. As we will discuss immediately below, the mutual information $I(A_1, A_2)$ has exactly the same phase structure as that at infinite temperature. Thus, the mutual information appears to be insensitive to finite temperature effects.

The behavior of Renyi mutual informations in general become n -dependent, with n -dependent phase boundaries. See figure 8 for an example.

2. For negativities, each n can have its own version of finite-temperature generalizations of the NE, ME, and ES phases, with phase boundaries between them, which in general are n -dependent. See figure 6 for an example.

¹⁴Recall that other than the logarithmic negativities, all other quantities are sensitive to classical correlations.

3. In the phase diagram of the logarithmic negativity, there are finite-temperature generalizations of the NE, ME and ES phases. Between these phases, there can be new phases that have different \mathcal{E} behavior, and are separated from other phases by non-analytic dependences on the c, λ parameters. Such phases cannot be deduced by analytic continuation of $\mathcal{Z}_n^{(\text{PT})}$ or R_n . See figure 11 for an example.
4. There are regimes where \mathcal{E} has a nonzero volume-law contribution, but the mutual information I is sub-extensive. This is a surprising result, as it is generally believed that mutual information contains both quantum and classical correlations. That it is sub-extensive intuitively implies there cannot be any volume-like quantum entanglement correlations. Our results indicate this intuition cannot be correct. We will elaborate further on this relation between logarithmic negativity and mutual information in section 5.
5. For various choices of $\rho^{(\text{eq})}$, the Renyi and logarithmic negativities can often be expressed in terms of equilibrium Renyi entropies, but this appears to not always be the case.
6. Due to the n -dependence of phase boundaries of R_n and I_n , for a given λ, c , there can be intricate patterns of entanglement. For example, it is possible that for all n , the negativities show the behavior of the ME phase, while for all n the mutual information show the behavior of the ES phase. It can also happen that for different n , R_n show different phases, i.e. some n can be in the ME phase while other n are in the ES or NE phases. See figure 5 for some examples.

To conclude this general discussion, let us now consider the mutual information. From our general discussion in section 2.3, we have

$$S_{n,A} = \min \left(S_{n,A}^{(\text{eq})}, S_{n,A}^{(\text{eq})} \right) \tag{4.24}$$

for $n \geq 2$, as well as

$$S_A = \min \left(S_A^{(\text{eq})}, S_A^{(\text{eq})} \right). \tag{4.25}$$

Given that the von Neumann entropy of a thermal density operator is extensive, i.e.

$$S_{A_1}^{(\text{eq})} + S_{A_2}^{(\text{eq})} = S_A^{(\text{eq})}, \quad S_{A_1 B}^{(\text{eq})} = S_{A_1}^{(\text{eq})} + S_B^{(\text{eq})}, \tag{4.26}$$

we then have

$$I(A_1, A_2) = \begin{cases} 0 & c < \frac{1}{2} \text{ i.e. } S_A^{(\text{eq})} = S_B^{(\text{eq})} \\ S_A^{(\text{eq})} - S_B^{(\text{eq})} & c > \frac{1}{2}, 1 - \frac{1}{2c} < \lambda < \frac{1}{2c} \\ 2S_{A_2}^{(\text{eq})} & \lambda > \frac{1}{2c} \text{ i.e. } S_{A_1}^{(\text{eq})} > S_{A_2}^{(\text{eq})} + S_B^{(\text{eq})} \\ 2S_{A_1}^{(\text{eq})} & \lambda < 1 - \frac{1}{2c} \text{ i.e. } S_{A_2}^{(\text{eq})} > S_{A_1}^{(\text{eq})} + S_B^{(\text{eq})} \end{cases} \tag{4.27}$$

where c and λ were defined in (4.2). Thus, we find exactly the same phase structure as that at infinite temperature and the same phase diagram figure 4. For the Renyi mutual

information, we have a similar phase structure, but the transitions are in general at n -dependent values of c, λ . More explicitly, suppose $\mathcal{I}_\alpha = \mathcal{I}_A \otimes \mathcal{I}_B$ can be factorized between A and B , then from (4.24),

$$I_n(A_1, A_2) = \begin{cases} S_{n,A_1}^{(\text{eq})} + S_{n,A_2}^{(\text{eq})} - S_{n,A}^{(\text{eq})} & S_{n,A}^{(\text{eq})} < S_{n,B}^{(\text{eq})} \\ S_{n,A_1}^{(\text{eq})} + S_{n,A_2}^{(\text{eq})} - S_{n,B}^{(\text{eq})} & S_{n,A}^{(\text{eq})} > S_{n,B}^{(\text{eq})}, S_{n,B}^{(\text{eq})} > |S_{n,A_1}^{(\text{eq})} - S_{n,A_2}^{(\text{eq})}| \\ 2S_{n,A_2}^{(\text{eq})} & S_{n,A_1}^{(\text{eq})} > S_{n,A_2}^{(\text{eq})} + S_{n,B}^{(\text{eq})} \\ 2S_{n,A_1}^{(\text{eq})} & S_{n,A_2}^{(\text{eq})} > S_{n,A_1}^{(\text{eq})} + S_{n,B}^{(\text{eq})} \end{cases}, \quad (4.28)$$

Note that the Renyi entropies from subsystems may not be additive, i.e. $S_{n,A_1}^{(\text{eq})} + S_{n,A_2}^{(\text{eq})}$ does not have to be equal to $S_{n,A}^{(\text{eq})}$.

The behavior of the Renyi and logarithmic negativities is technically much more complicated at finite temperature, and we turn to this next.

4.4 Canonical ensemble

In this subsection, we consider more explicitly the entanglement structure in the canonical ensemble.

4.4.1 General setup

In the canonical ensemble, we consider an effective identity operator of the form

$$\mathcal{I} = e^{-\beta_A H_A} \otimes e^{-\beta_B H_B}, \quad (4.29)$$

which can be interpreted in several ways. One possible scenario is that after some brief interactions, subsystems A and B are separated and thus can achieve separate equilibria at different temperatures. When $\beta_A = \beta_B = \beta$ it also applies to a system AB with a local Hamiltonian $H = H_A + H_B + H_{AB}$ for which $e^{-\beta H} \approx e^{-\beta H_A} \otimes e^{-\beta H_B}$. We can drop the H_{AB} term as it should give contributions proportional to the areas of the subsystems in the exponent, and thus can be neglected in the thermodynamic limit.

We assume that the system is homogeneous within A , so that $Z_{m,R} = \text{Tr}[e^{-m\beta_R H_R}]$ for different subsystems R can be written as

$$Z_{m,A_1} = e^{V_{A_1} f_A(m\beta_A)}, \quad Z_{m,A_2} = e^{V_{A_2} f_A(m\beta_A)}, \quad Z_{m,B} = e^{V_B f_B(m\beta_B)}. \quad (4.30)$$

The corresponding equilibrium Renyi entropy density can be written as

$$s_{n,R}^{(\text{eq})}(\beta) := \frac{n f_R(\beta) - f_R(n\beta)}{n-1}, \quad S_{n,R}^{(\text{eq})}(\beta) = V_R s_{n,R}^{(\text{eq})}(\beta). \quad (4.31)$$

We allow n to take any non-negative real value. The $n \rightarrow 1$ limit of (4.31) gives the density of the equilibrium von Neumann entropy (i.e. the thermodynamic entropy density),

$$s_R^{(\text{eq})}(\beta) := f_R(\beta) - \beta f'_R(\beta), \quad S_R^{(\text{eq})}(\beta) = V_R s_R^{(\text{eq})}(\beta). \quad (4.32)$$

The free energy densities $f_A(\beta)$ and $f_B(\beta)$ both satisfy the following properties:

1. Since $\langle (H - \langle H \rangle_\beta)^2 \rangle_\beta = N f''(\beta)$ is non-negative,

$$f''(\beta) \geq 0. \quad (4.33)$$

2. From the non-negativity of the thermodynamic entropy,

$$f(\beta) - \beta f'(\beta) \geq 0. \quad (4.34)$$

3. From the non-negativity of (4.31),

$$n f(\beta) - f(n\beta) \geq 0, \quad n \geq 2. \quad (4.35)$$

4. From the fact that $s_n^{(\text{eq})}$ decreases with the Renyi index n , we have

$$\frac{d}{dn} \left(\frac{n f(\beta) - f(n\beta)}{n-1} \right) \leq 0 \Rightarrow f(\beta) - \beta f'(n\beta) \geq f(n\beta) - n\beta f'(n\beta) \quad (4.36)$$

Note that (4.36), together with (4.34) at inverse temperature $n\beta$, implies that

$$f(\beta) - \beta f'(n\beta) \geq 0. \quad (4.37)$$

From (3.4), we find

$$\mathcal{Z}_n^{(\text{PT})}(A_1, A_2) = \frac{1}{e^{nF}} \sum_{\tau} e^{\mathcal{A}(\tau)}, \quad \mathcal{Z}_n^{(A)} = \frac{1}{e^{nF}} \sum_{\tau} e^{\mathcal{B}(\tau)} \quad (4.38)$$

where

$$\mathcal{A}(\tau) = V_{A_1} \sum_{i=1}^{k(\eta^{-1}\tau)} f_A(a_i\beta_A) + V_{A_2} \sum_{j=1}^{k(\eta\tau)} f_A(b_j\beta_A) + V_B \sum_{k=1}^{k(\tau)} f_B(c_k\beta_B), \quad (4.39)$$

$$\mathcal{B}(\tau) = V_A \sum_{i=1}^{k(\eta^{-1}\tau)} f_A(a_i\beta_A) + V_B \sum_{k=1}^{k(\tau)} f_B(c_k\beta_B), \quad (4.40)$$

$$F = V_A f_A(\beta_A) + V_B f_B(\beta_B) \quad (4.41)$$

and $\{a_i\}, \{b_j\}, \{c_k\}$ respectively are the numbers of elements of the cycles in $\eta^{-1}\tau, \eta\tau$ and τ . Now if τ_m denotes one of the permutations that maximizes $\mathcal{A}(\tau)$ and τ'_m denotes one of the permutations that maximize $\mathcal{B}(\tau)$ for a particular choice of V_{A_1}, V_{A_2}, V_B , then

$$\log \mathcal{Z}_n^{(\text{PT})}(A_1, A_2) \approx \mathcal{A}(\tau_m) - nF \quad (4.42)$$

and

$$R_n(A_1, A_2) = b_n (\mathcal{A}(\tau_m) - \mathcal{B}(\tau'_m)) = b_n \left(\log \mathcal{Z}_n^{(\text{PT})} - \log \mathcal{Z}_n^{(A)} \right). \quad (4.43)$$

Using the equilibrium Renyi entropies (4.31) we can write the contribution from a permutation τ to (4.38) as

$$\log \mathcal{Z}_n^{(\text{PT})}(\tau) = \sum_{i=1}^{k(\eta\tau^{-1})} -(a_i - 1)S_{a_i, A_1}^{(\text{eq})} + \sum_{i=1}^{k(\eta\tau)} -(b_i - 1)S_{b_i, A_2}^{(\text{eq})} + \sum_{i=1}^{k(\tau)} -(c_i - 1)S_{c_i, B}^{(\text{eq})} \quad (4.44)$$

$$\log \mathcal{Z}_{n,A}(\tau) = \sum_{i=1}^{k(\eta\tau^{-1})} -(a_i - 1)S_{a_i, A}^{(\text{eq})} + \sum_{i=1}^{k(\tau)} -(c_i - 1)S_{c_i, B}^{(\text{eq})} \quad (4.45)$$

For the Renyi partition function (4.45), the dominant permutations were already reviewed in section 2.3 and the end of last section, given by either $\tau = e$ or $\tau = \eta$, with

$$\log \mathcal{Z}_{n,A}(e) = -(n-1)S_{n,A}^{(\text{eq})}(\beta_A), \quad \log \mathcal{Z}_{n,A}(\eta) = -(n-1)S_{n,B}^{(\text{eq})}(\beta_B). \quad (4.46)$$

To find the negativities we need to maximize (4.44) among different permutations τ , which is a difficult task without knowing the explicit forms of functions f_A, f_B and relative magnitude of β_A, β_B . We will give detailed discussion for some specific examples and here summarize some general features. In all our investigations we find that for a given n , only one of the following set of permutations could dominate: $\tau = e$, $\tau = \eta, \eta^{-1}$, $\tau = \tau_{ES}$, which may be considered as giving finite temperature generalizations of the NE, ME and ES phases. The set of τ_{ES} denotes,

$$\tau_{ES} = \begin{cases} (12)(34) \cdots (n-1, n) \text{ and } (23)(45) \cdots (n1), & \text{even } n \\ \{(23)(45) \cdots (n-1, n), (34)(56) \cdots (n1), \dots, (12) \cdots (n-2, n-1)\}, & \text{odd } n \end{cases}, \quad (4.47)$$

which is a subset of the permutations $\tau^*, \tau_{\text{odd}}^*$ for the ES phase at the infinite temperature.¹⁵ The corresponding expressions of (4.44) for these permutations are given by

$$\log \mathcal{Z}_n^{(\text{PT})}(e) = -(n-1)S_{n,A}^{(\text{eq})}(\beta_A) \quad (4.48)$$

$$\log \mathcal{Z}_n^{(\text{PT})}(\eta) = \begin{cases} -(n-1)S_{n,B}^{(\text{eq})}(\beta_B) - (n-2)S_{\frac{n}{2}, A_2}^{(\text{eq})}(\beta_A) & n \geq 2, n \text{ even} \\ -(n-1)(S_{n, A_2}^{(\text{eq})}(\beta_A) + S_{n, B}^{(\text{eq})}(\beta_B)) & n \geq 3, n \text{ odd} \end{cases} \quad (4.49)$$

$$\log \mathcal{Z}_n^{(\text{PT})}(\eta^{-1}) = \begin{cases} -(n-1)S_{n,B}^{(\text{eq})}(\beta_B) - (n-2)S_{\frac{n}{2}, A_1}^{(\text{eq})}(\beta_A) & n \geq 2, n \text{ even} \\ -(n-1)(S_{n, A_1}^{(\text{eq})}(\beta_A) + S_{n, B}^{(\text{eq})}(\beta_B)) & n \geq 3, n \text{ odd} \end{cases} \quad (4.50)$$

$$\log \mathcal{Z}_n^{(\text{PT})}(\tau_{ES}) = \begin{cases} -\frac{n-2}{2}S_{\frac{n}{2}, A}^{(\text{eq})}(\beta_A) - \frac{n}{2}S_{2, B}^{(\text{eq})}(\beta_B) & n \geq 2, n \text{ even} \\ -\frac{n-1}{2} \left(S_{\frac{n+1}{2}, A}^{(\text{eq})}(\beta_A) + S_{2, B}^{(\text{eq})}(\beta_B) \right) & n \geq 3, n \text{ odd} \end{cases}. \quad (4.51)$$

The regions where the different permutations dominate can be found by comparing their magnitudes. For example, for even n , τ_{ES} dominates over e when

$$(n-1)S_{n,A}^{(\text{eq})} - \frac{n-2}{2}S_{\frac{n}{2}, A}^{(\text{eq})} \geq \frac{n}{2}S_{2, B}^{(\text{eq})} \quad (4.52)$$

¹⁵In some special finite temperature cases, all permutations in the τ^* or τ_{odd}^* give degenerate contributions, but more generally the degeneracy is broken. We note that these permutations are also relevant to the moments of the partial transpose for tensor network states with non-maximally entangled bonds [46].

and the equal sign defines the transition line for the change of dominance. These transition lines are generally n -dependent.

If we can analytically continue the expressions (4.48)–(4.51) for even n to $n = 1$ we would obtain the logarithmic negativity for the corresponding finite temperature generalizations of NE, ME, and ES phases with

$$\mathcal{E}_{NE} = 0, \tag{4.53}$$

$$\mathcal{E}_{ME} = S_{\frac{1}{2},A_2}^{(\text{eq})}(\beta_A) \quad \text{or} \quad \mathcal{E}_{ME} = S_{\frac{1}{2},A_1}^{(\text{eq})}(\beta_A), \tag{4.54}$$

$$\mathcal{E}_{ES} = \frac{1}{2} \left(S_{\frac{1}{2},A}^{(\text{eq})}(\beta_A) - S_{2,B}^{(\text{eq})}(\beta_B) \right). \tag{4.55}$$

Comparing with the behavior of mutual information (4.27), notice that in (4.54), $S_{\frac{1}{2},A_2}^{(\text{eq})}(\beta)$ appears and is no longer related to the third line of (4.27) by a simple factor of $\frac{1}{2}$ as it was in the infinite temperature case. Similarly, in (4.55), $S_{\frac{1}{2},A}^{(\text{eq})}$ and $S_{2,B}^{(\text{eq})}$ appear, in contrast to the second line of (4.27). Analytically continuing (4.52) to $n = 1$ gives

$$S_{\frac{1}{2},A}^{(\text{eq})} = S_{2,B}^{(\text{eq})} \tag{4.56}$$

which may be interpreted as the transition from NE to ES phase for \mathcal{E} . As we emphasized before, the correctness of such analytic continuations is not warranted in particular in regions of parameters where the dependence on n is not uniform. In section 4.4.3 we discuss a case where the resolvent can be calculated numerically, and we checked in the NE and ES regimes that the expressions (4.53), (4.55) and (4.56) do apply.

It is instructive to compare (4.56) with the transition from NE to ES for the mutual information I which from (4.27) is given by $c = \frac{1}{2}$, i.e.

$$S_A^{(\text{eq})} = S_B^{(\text{eq})}. \tag{4.57}$$

Given that S_n monotonically decreases with n , we have $S_{\frac{1}{2},A}^{(\text{eq})} > S_A^{(\text{eq})}$ and $S_{2,B}^{(\text{eq})} < S_B^{(\text{eq})}$, so the transition (4.56) must happen for $S_A^{(\text{eq})} < S_B^{(\text{eq})}$, i.e. somewhere with $c < \frac{1}{2}$. Therefore, (4.56) implies a region in the phase diagram where the logarithmic negativity is extensive, but the mutual information is sub-extensive.

Below, we will consider the following representatives of two classes of examples:

1. In discrete systems, $f(\beta)$ has a well-defined $\beta \rightarrow 0$ limit. An example is

$$f(\beta) = \log 2 + \log \cosh(\beta J) \tag{4.58}$$

which is observed in a variety of spin chain models.

2. In continuum systems, $f(\beta)$ does not have a well-defined $\beta \rightarrow 0$ limit. For instance, one form that appears in quantum field theories is

$$f(\beta) = a\beta^{-\alpha}, \quad \alpha \geq 0. \tag{4.59}$$

4.4.2 A and B are uniform

As an explicit illustration, we consider the case where the Hamiltonian is translationally invariant and $\beta_A = \beta_B$ for which we have $f_A(\beta) = f_B(\beta) = f(\beta)$. It will be useful to consider the explicit examples in both (4.58) and (4.59). In this case, it is not clear how to perform a resolvent calculation, so we are not able to find the phase diagram for \mathcal{E} and will only present the structure for R_n .

Since the system is homogeneous, for $\log \mathcal{Z}_n^{(A)}$, $\tau = e$ dominates for $c < \frac{1}{2}$ and $\tau = \eta$ dominates for $c > \frac{1}{2}$. In this case, c_n and λ_n coincide with c and λ for all n , so the Renyi mutual information (4.28) has the same phase structure as the mutual information (4.27).

To understand the structure of $\log \mathcal{Z}_n^{(\text{PT})}$, it is convenient to introduce $V = V_A + V_B$ and rewrite $\mathcal{A}(\tau)$ in four ways,

$$\frac{1}{V} \mathcal{A}(\tau) = c\lambda \left(G(\beta, \tau) + G(\beta, \eta^{-1}\tau) \right) + c(1-\lambda) \left(G(\beta, \tau) + G(\beta, \eta\tau) \right) + (1-2c)G(\beta, \tau) \quad (4.60)$$

$$= (2c\lambda - 1)G(\beta, \eta^{-1}\tau) + c(1-\lambda) \left(G(\beta, \eta\tau) + G(\beta, \eta^{-1}\tau) \right) + (1-c) \left(G(\beta, \tau) + G(\beta, \eta^{-1}\tau) \right) \quad (4.61)$$

$$= c\lambda \left(G(\beta, \eta^{-1}\tau) + G(\beta, \eta\tau) \right) + (2c(1-\lambda) - 1)G(\beta, \eta\tau) + (1-c) \left(G(\beta, \tau) + G(\beta, \eta\tau) \right) \quad (4.62)$$

$$= \left(\frac{1}{2} - c\lambda \right) \left(G(\beta, \tau) + G(\beta, \eta\tau) \right) + \left(\frac{1}{2} - c(1-\lambda) \right) \left(G(\beta, \tau) + G(\beta, \eta^{-1}\tau) \right) + \left(c - \frac{1}{2} \right) \left(G(\beta, \eta\tau) + G(\beta, \eta^{-1}\tau) \right) \quad (4.63)$$

where we have introduced

$$G(\beta, \tau) := \sum_{i=1}^{k(\tau)} f(c_i \beta) \quad (4.64)$$

with $\{c_i\}$ the lengths of the cycles in τ . We show in appendix D that $G(\beta, \tau)$ and $G(\beta, \eta^{-1}\tau)$ are respectively maximized by $\tau = e$ and $\tau = \eta$, and $G(\beta, \tau) + G(\beta, \eta^{-1}\tau)$ is maximized by both $\tau = e$ and $\tau = \eta$. Similarly, $G(\beta, \eta\tau)$ is maximized by $\tau = \eta^{-1}$, and $G(\beta, \tau) + G(\beta, \eta\tau)$ is maximized by both $\tau = e$ and $\tau = \eta^{-1}$. Furthermore, we observe numerically that $G(\beta, \eta\tau) + G(\beta, \eta^{-1}\tau)$ is maximized for even n by four permutations: $\tau = \eta, \eta^{-1}$ and $\tau_{ES} = \{(12)(34) \dots (n-1 n), (23)(45) \dots (n1)\}$, and for odd n by two permutations: $\tau = \eta, \eta^{-1}$.

From the above, we then find:

1. For $c < \frac{1}{2}$, from (4.60) $\tau = e$ is the dominant permutation for $\mathcal{A}(\tau)$ for all $n \geq 3$, with $\log \mathcal{Z}_n^{(\text{PT})}$ given by (4.48). Since $\tau = e$ is also the dominant permutation for $\mathcal{B}(\tau)$, we have

$$R_n(A_1, A_2) = 0, \quad n \geq 3, \quad \lim_{n \rightarrow 2} R_n(A_1, A_2) = 0, \quad (4.65)$$

This is the natural generalization of the NE phase for the R_n .

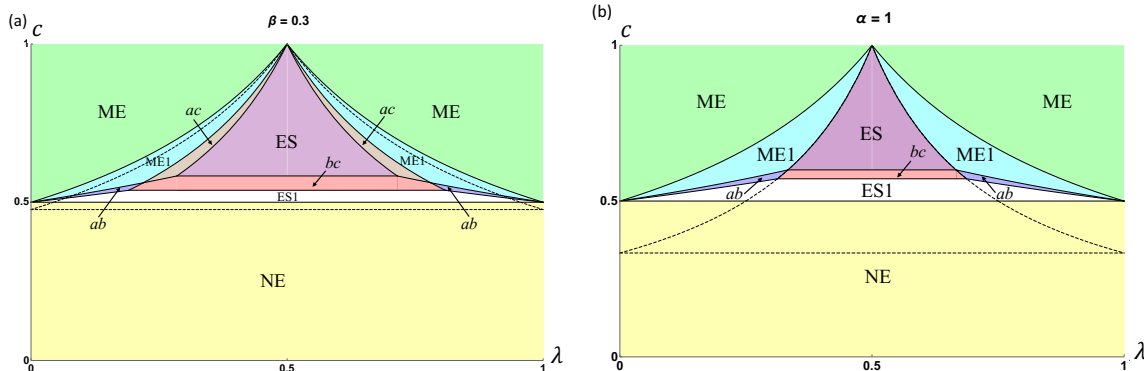


Figure 5. Phase diagrams for the canonical ensemble in a homogeneous system. (a) shows the case with $f(\beta)$ as in (4.58), and (b) shows the case with $f(\beta)$ in (4.59). Note that in (a), in addition to the ab , bc and ac regions, there is also a small abc region near the intersections of these regions which is not shown explicitly in the figure. The dashed lines show the prediction for the phase transition in the logarithmic negativity from the naive analytic continuation prescription, where we set $\mathcal{E}_{ES} = \mathcal{E}_{NE}$ or $\mathcal{E}_{ES} = \mathcal{E}_{ME}$ in (4.53)–(4.55).

- For $c > \frac{1}{2}$, $\lambda > \frac{1}{2c}$, from (4.61) the dominant permutation for $\mathcal{A}(\tau)$ is η , with $\log \mathcal{Z}_n^{(\text{PT})}$ given by (4.49) and

$$R_n(A_1, A_2) = \begin{cases} V_{A_2} \frac{nf(\beta) - f(n\beta)}{n-1} = S_{n, A_2}^{(\text{eq})} & n \geq 3, n \text{ odd} \\ V_{A_2} \frac{nf(\beta) - 2f(\frac{n}{2}\beta)}{n-2} = S_{\frac{n}{2}, A_2}^{(\text{eq})} & n \geq 4, n \text{ even} \end{cases} \quad (4.66)$$

$$\lim_{n \rightarrow 2} R_n(A_1, A_2) = S_{A_2}^{(\text{eq})}(\beta), \quad (4.67)$$

From analytic continuation, \mathcal{E} is given by (4.54). This may naturally be interpreted as the finite temperature generalization of the maximally entangled (ME) phase. Notice in (4.66) for even n , $S_{\frac{n}{2}, A_2}^{(\text{eq})}(\beta)$ appears instead of $S_{n, A_2}^{(\text{eq})}(\beta)$ as in the third line of (4.28) for Renyi mutual information.

For $c > \frac{1}{2}$, $\lambda < 1 - \frac{1}{2c}$, the dominant permutation is η^{-1} , and we get similar expressions with A_2 replaced by A_1 .

- For $c > 1/2$, $1 - \frac{1}{2c} < \lambda < \frac{1}{2c}$, the dominant permutations now appear to depend on values of c and λ , and are model-dependent. We present the phase diagrams for examples (4.58) and (4.59) based on the behavior of even n in figure 5 (a) and (b) respectively. In addition to the NE and ME phases discussed in items 1 and 2 above, there are also the following regions in the phase diagrams:

- ME1 region in figure 5 where $\tau_m = \eta$ or $\tau_m = \eta^{-1}$ dominates for all n . The behavior of Renyi and logarithmic negativities are the same as (4.66) and (4.54). But note that the mutual information in this region are given by the second line of (4.27) rather than the third line or fourth line. So the mutual information and Renyi mutual information in this region differ from those of ME in item 2, and they should be considered as describing distinct phases. We refer to this region

as the ME1 phase. It is curious that in this region, the negativities behave like those in the maximally entangled phase, while the mutual informations behave like those for the entanglement saturation phase. In contrast to the relation (2.25) for pure states, we have for even n

$$R_n(A_1, A_2) > \frac{1}{2} I_{n/2}(A_1, A_2). \quad (4.68)$$

- (b) ES1 region in figure 5 where $\tau_m = e$ dominates for all $n \geq 3$. But since for $c > \frac{1}{2}$ the dominant permutation for $\mathcal{B}(\tau)$ is $\tau = \eta$, the values of R_n are different from those in (4.65),

$$R_n(A_1, A_2) = \begin{cases} \frac{n-1}{n-2}(S_{n,A}^{(\text{eq})} - S_{n,B}^{(\text{eq})}) & n \text{ even} \\ S_{n,A}^{(\text{eq})} - S_{n,B}^{(\text{eq})} & n \text{ odd} \end{cases}, \quad n \geq 3 \quad (4.69)$$

$$\log \mathcal{Z}_n^{(\text{PT})} = \begin{cases} -(n-1)S_{n,A}^{(\text{eq})} & n \geq 3 \\ -S_{2,B}^{(\text{eq})} & n = 2 \end{cases}. \quad (4.70)$$

We cannot obtain $\mathcal{E}(A_1, A_2)$ by analytic continuation from the above expressions, as the analytical continuation of (4.70) for even $n \geq 4$ to $n = 2$ already gives the wrong expression, not to mention to $n = 1$. The mutual information of this region is given by the second line of (4.27), which is also nonzero. So this should describe a distinct phase from the NE phase of item 1. Since (4.69)–(4.70) and the Renyi mutual informations in this phase only depend on A , and not on A_1, A_2 , we will refer to it as the ES1 phase. For $n = 2$, there is no ES1 region.

- (c) ES region in figure 5 where $\tau_m = \tau_{ES}$ for all n with $\log \mathcal{Z}_n^{(\text{PT})}$ given by (4.51). We find

$$R_n(A_1, A_2) = \begin{cases} \frac{1}{2} S_{\frac{n-1}{2}, A}^{(\text{eq})}(\beta) - S_{n,B}^{(\text{eq})}(\beta) + \frac{1}{2} S_{2,B}^{(\text{eq})}(\beta) & n \geq 3, n \text{ odd} \\ \frac{1}{2} (S_{\frac{n}{2}, A}^{(\text{eq})}(\beta) - S_{\frac{n}{2}, B}^{(\text{eq})}(2\beta)) & n \geq 4, n \text{ even} \end{cases} \quad (4.71)$$

$$\lim_{n \rightarrow 2} R_n(A_1, A_2) = \frac{1}{2} S_A^{(\text{eq})}(\beta) - \frac{1}{2} S_B^{(\text{eq})}(2\beta) \quad (4.72)$$

and the logarithmic negativity is given by (4.55). Notice all quantities depend only on properties of A , not on A_1 or A_2 . We can thus identify this as the finite temperature generalization of the ES phase. For even n we also have (4.68).

- (d) In the remaining regions of figure 5, the set of dominant permutations now depends on n . For example, in the region denoted by ac there exists a critical value n_c for $n \leq n_c$, the dominant permutations are that of item (a), while for $n > n_c$, the dominant permutations are those of item (c). Similarly with bc and ab regions. There is in general also an abc region where all three options appear at a given c, λ for different n . These regions also give potentially new phases. Clearly in these “mixed” regions, there is no way to perform analytic continuation in n to obtain $\mathcal{E}(A_1, A_2)$.

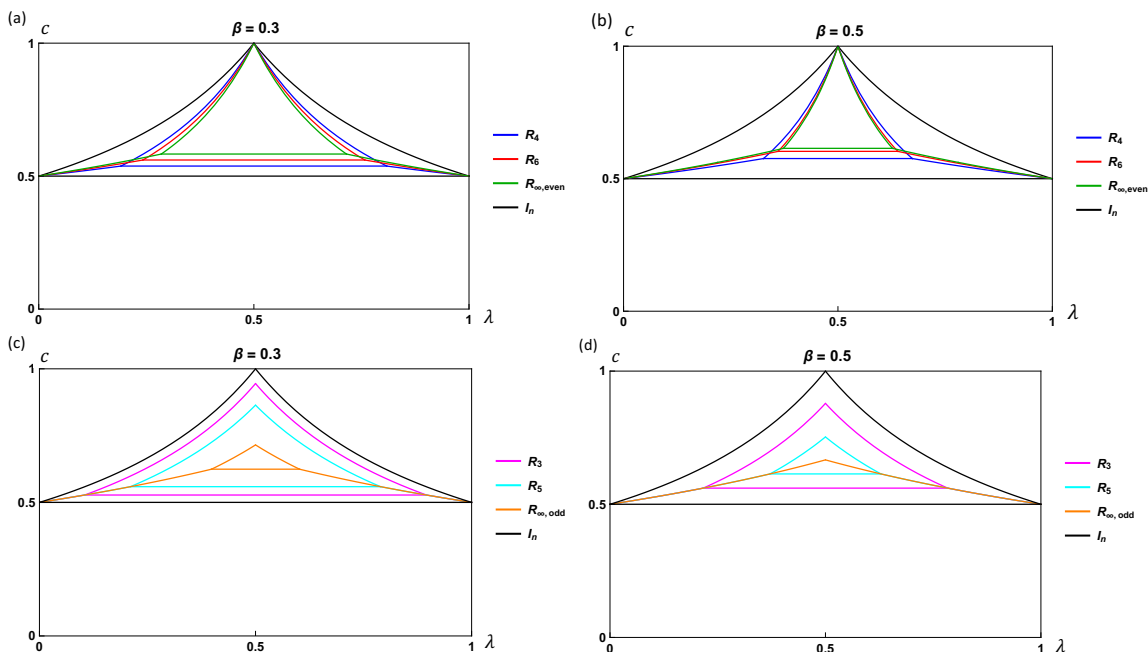


Figure 6. Transition lines for R_n at two different values of β for $f(\beta)$ in (4.58).

In the above, we have assumed that for a region if a given set of permutations dominate for all n and there is no apparent contradiction, we can analytically continue n to obtain the logarithmic negativity. But we should caution that it is not warranted such a procedure always gives the correct answer.

4. As the temperature decreases, the ES phase region shrinks.

The n -dependent transition lines for R_n are shown for a few examples of even and odd n in figures 6 and 7. Figure 6 shows the case of $f(\beta)$ as in (4.58), at two different values of β . Figure 7 shows the case of $f(\beta)$ in (4.59), where the transition lines turn out to be independent of β , at two different values of the parameter α . Note that there are qualitative differences between the diagrams for even and odd n . In particular, for sufficiently large β or α , the odd n cases have no region where τ_{ES} dominates.

4.4.3 A at infinite temperature

As another illustration, here we consider the case where A is at infinite temperature, while B is at a finite temperature β . We find numerically for $n = 4, 6$ that the transition the dominant permutation contributing to $\mathcal{Z}_n^{(PT)}$ is always one out of e, τ^* , and η, η^{-1} , and assume that this is true for all n . The corresponding expressions for $\mathcal{Z}_n^{(PT)}$ and analytically continued expressions for \mathcal{E} are given by (4.48)–(4.51) and (4.53)–(4.55). The phase transitions for both the mutual information and the Renyi negativities are n -dependent in this case, and are shown in figure 8.

These phase diagrams have a similar structure for both discrete and continuum examples (4.58) and (4.59). The phase diagrams for both R_n and I_n are n -dependent in this

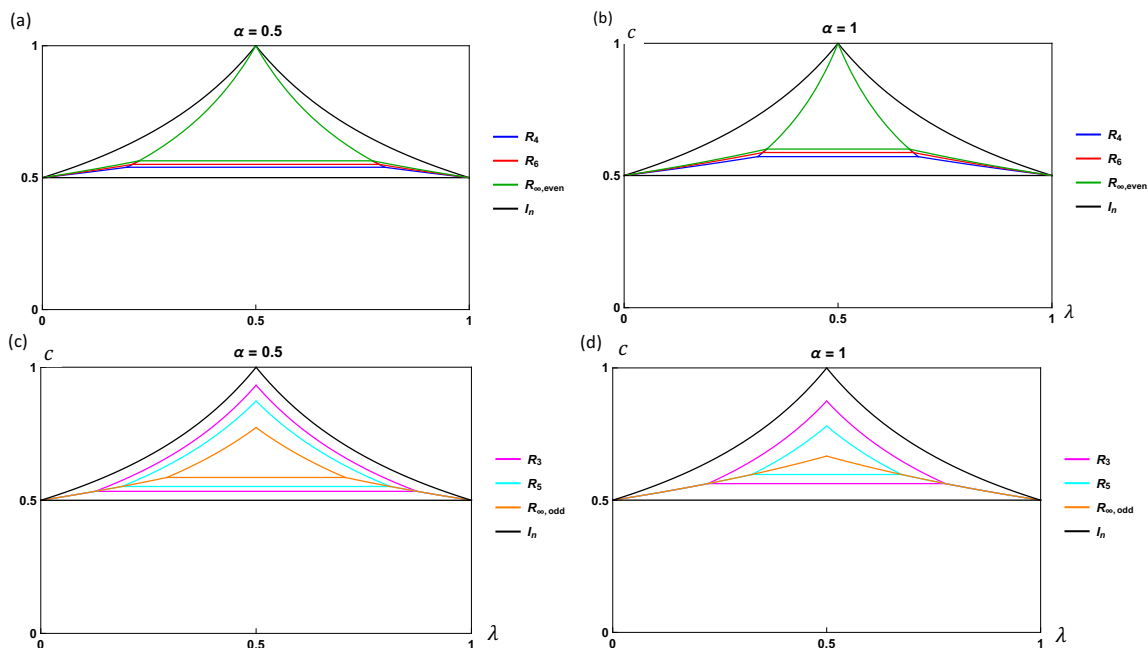


Figure 7. Transition lines for R_n at two different values of α for $f(\beta)$ in (4.59).

case. However, the transition from dominance of $\tau = e$ to $\tau = \tau_{ES}$ for all R_n happens along the same line, indicated as the purple dashed line in the figure. This coincides with the transition in I_2 from NE to ES, and takes place at a smaller value of c than the corresponding transition in I . As explained in appendix C.3, the transition in \mathcal{E} obtained from the semicircle approximation in the resolvent also lies along the same line. More explicitly, in the semi-circle approximation, the logarithmic negativity is given by

$$\mathcal{E}(A_1, A_2) = \log \left(\frac{2}{\pi} \sin^{-1} \left[\frac{1}{2} \sqrt{\frac{e^{S_{2,B}^{(eq)}}}{d_A}} \right] + \frac{\left(\sqrt{\frac{e^{S_{2,B}^{(eq)}}}{d_A}} + 8 \sqrt{\frac{d_A}{e^{S_{2,B}^{(eq)}}}} \right) \sqrt{1 - \frac{1}{4} \frac{e^{S_{2,B}^{(eq)}}}{d_A}}}{3\pi} \right), \quad (4.73)$$

$$\simeq \begin{cases} 0, & \log d_A < S_{2,B}^{(eq)} \\ \frac{1}{2} \left(\log d_A - S_{2,B}^{(eq)} \right) + \log \frac{8}{3\pi}, & \log d_A > S_{2,B}^{(eq)} \end{cases} \quad (4.74)$$

Comparison between the mutual information and negativity is shown in figure 9. Here, we explicitly see a regime before the Page time when the negativity is extensive and hence much larger than the mutual information.

4.5 Microcanonical ensemble

4.5.1 General description

In cases where the initial state is supported on a narrow interval $I_{E,\Delta} = [E - \Delta, E + \Delta]$ of energies, we can take the effective identity operator for the equilibrium approximation to be

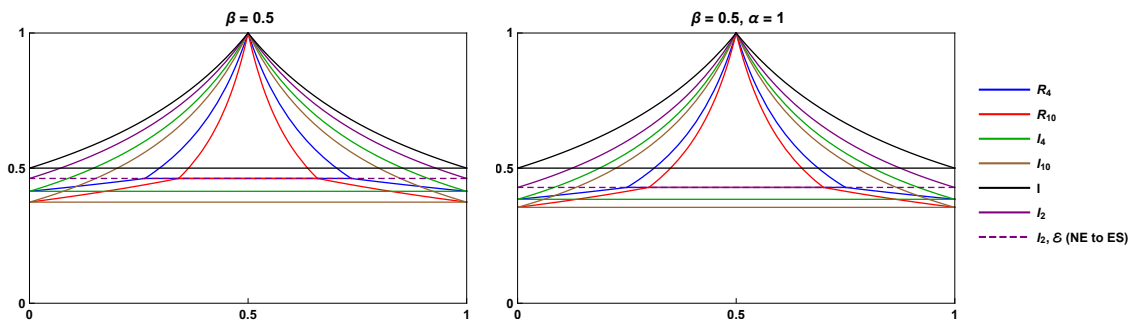


Figure 8. Phase transition lines for R_n and I_n for a few different n for f_B as in (4.58) and (4.59) are shown on the left and right respectively.

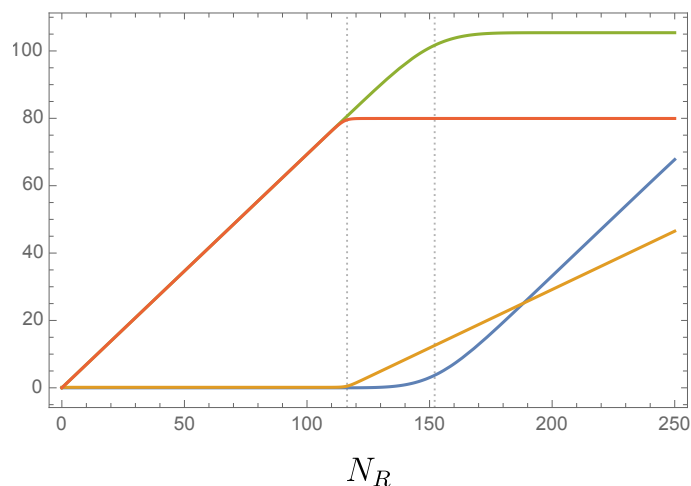


Figure 9. The mutual information (blue), negativity (orange), von Neumann entropy of R (green), and second Rényi entropy of R (red) are shown when B is in the canonical ensemble and A at infinite temperature. While the mutual information within A becomes volume-law when the von Neumann entropy of A saturates, the negativity becomes volume-law when the second Rényi entropy saturates, a significantly earlier time. We also note that the finite temperature corrections around the transition point are significantly larger for the mutual information than they are for negativity, which has a sharper transition. We take $V_B = 100$ and $\beta = 1/2$ for a Cardy-like density of states for B , $\rho(E) = e^{V_B \sqrt{\frac{E}{V_B}}}$.

the projector onto the eigenstates in this microcanonical interval. Again assuming that the interaction terms between different subsystems A_1, A_2 and B give small contributions to the energy compared to the terms involving only degrees of freedom within the subsystems, we can approximate this effective identity operator as

$$\mathcal{I}_E = \sum_{E_n \in I_{E,\Delta}} |n\rangle \langle n| \approx \sum_{E_p^{A_1} + E_q^{A_2} + E_r^B \in I_{E,\Delta}} (|p\rangle \langle p|)_{A_1} \otimes (|q\rangle \langle q|)_{A_2} \otimes (|r\rangle \langle r|)_B. \quad (4.75)$$

where $|p\rangle, |q\rangle$ and $|r\rangle$ are respectively eigenstates of H_{A_1}, H_{A_2} and H_B with energies $E_p^{A_1}, E_q^{A_2}$ and E_r^B . Below we will denote $N_E := \text{Tr}[\mathcal{I}_E]$. The contributions to the equilibrium

approximation for $\mathcal{Z}_n^{(\text{PT})}$ from general permutations τ with this choice of \mathcal{I}_α in a general homogenous system are complicated to evaluate, and in particular do not have simple expressions in terms of the cycle numbers and lengths of various permutations like the canonical ensemble expressions (4.38). Some examples are shown in appendix E. We consider two simplifying cases below, with B or A_2 at infinite temperature. These examples are useful as they allow a calculation of the logarithmic negativity through the resolvent, as explained in appendices C.4 and C.5, revealing new features in their phase diagrams which do not have an analog at infinite temperature and cannot be deduced from the phase diagrams for the Renyi negativities. In both examples, we find that the logarithmic negativity becomes extensive in a regime where the mutual information is sub-extensive.

For concreteness, in both cases below, it will sometimes be useful to consider a specific example of the entropy density as a function of energy density in the part of the system with energy conservation,

$$s(\epsilon) = g\sqrt{\epsilon}. \quad (4.76)$$

where g is a constant.

4.5.2 B at infinite temperature

Let us first assume that the total energy in A is conserved and it equilibrates to the microcanonical ensemble, while B equilibrates to infinite temperature. The effective identity operator is

$$\mathcal{I}_\alpha = \sum_{E_{a_1}^{A_1} + E_{a_2}^{A_2} \in I_{E,\Delta}} |a_1\rangle \langle a_1| \otimes |a_2\rangle \langle a_2| \otimes \mathbf{1}_B \quad (4.77)$$

As we vary the volumes of the different subsystems, we fix the average energy density in A to a value $\frac{E}{V_A} = \epsilon$, and the infinite temperature entropy density in B to a value s_0 .

Note first that since \mathcal{I}_α factorizes as $\mathcal{I}_A \otimes \mathcal{I}_B$ in this example, the Renyi mutual information is again given by the general expression (4.28). Note that $S_{n,A}^{(\text{eq})}$ and $S_{n,B}^{(\text{eq})}$ for all n in this example are equal to $S_A^{(\text{eq})}$ and $S_B^{(\text{eq})} = \log d_B$ respectively, so the transition from the first to the second line of (4.28) occurs at $S_A^{(\text{eq})} = \log d_B$ for all n . But note that in this example, the first line of (4.28) is not zero for $n \neq 1$: it is positive for $n < 1$, and negative for $n > 1$.

The contributions to $\mathcal{Z}_n^{(\text{PT})}$ for even $n \geq 2$ from different permutations are

$$\log \mathcal{Z}_n^{(\text{PT})}(e) = -(n-1)S_A^{(\text{eq})} \quad (4.78)$$

$$\log \mathcal{Z}_n^{(\text{PT})}(\eta) = -(n-1)\log d_B - (n-2)S_{\frac{n}{2},A_2}^{(\text{eq})} \quad (4.79)$$

$$\log \mathcal{Z}_n^{(\text{PT})}(\eta^{-1}) = -(n-1)\log d_B - (n-2)S_{\frac{n}{2},A_1}^{(\text{eq})} \quad (4.80)$$

$$\log \mathcal{Z}_n^{(\text{PT})}(\tau_{ES}) = -\frac{n-2}{2}S_{\frac{n}{2},A_1}^{(\text{eq})} - \frac{n-2}{2}S_{\frac{n}{2},A_2}^{(\text{eq})} - \frac{n}{2}\log d_B \quad (4.81)$$

The derivation of these expressions is explained in appendix C.4. On noticing that $S_{n,A}^{(\text{eq})} = S_A^{(\text{eq})}$ and $S_{n,B}^{(\text{eq})} = \log d_B$ for all n in this example, and that the equilibrium Renyi entropies of different subsystems are additive in the canonical ensemble, we can see that these expressions in terms of the equilibrium entropies turn out to be the same as the contributions

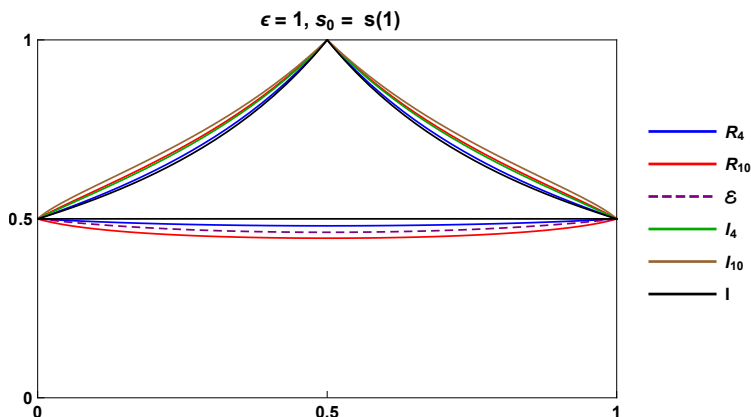


Figure 10. Phase transition lines for R_n and I_n for $n = 4, 10$. The transition from NE to ES for \mathcal{E} obtained from the semi-circle approximation is shown with the dashed purple line. The transition from the first to the second line of (4.28) for all Renyi mutual informations coincides with the black $n = 1$ transition line in this case.

from the leading permutations in the canonical ensemble case, (4.48)–(4.51). From analytic continuation, the above expressions correspond to the following values of the logarithmic negativity:

$$\mathcal{E}_{NE} = 0, \tag{4.82}$$

$$\mathcal{E}_{ME} = S_{\frac{1}{2}, A_2}^{(\text{eq})} \quad \text{or} \quad \mathcal{E}_{ME} = S_{\frac{1}{2}, A_1}^{(\text{eq})}, \tag{4.83}$$

$$\mathcal{E}_{ES} = \frac{1}{2} \left(S_{\frac{1}{2}, A_1}^{(\text{eq})} + S_{\frac{1}{2}, A_2}^{(\text{eq})} - \log d_B \right). \tag{4.84}$$

Depending on the permutations τ_m and τ'_m that dominate in $\mathcal{Z}_n^{(\text{PT})}$ and $\mathcal{Z}_n^{(A)}$ respectively, the Renyi negativities for even n can take three possible forms,

$$R_n = \begin{cases} 0 & \tau_m = e, \tau'_m = e \\ \min(S_{\frac{n}{2}, A_2}^{\text{eq}}, S_{\frac{n}{2}, A_1}^{\text{eq}}) & \tau_m = \eta^{-1}, \eta, \tau'_m = \eta \\ \frac{1}{2}(S_{n/2, A_1}^{\text{eq}} + S_{n/2, A_2}^{\text{eq}} - \log d_B) & \tau_m = \tau_{ES}, \tau'_m = \eta \end{cases} \tag{4.85}$$

Note that this example does not have an ES1 phase like (4.69) where $\tau_m = e, \tau'_m = \eta$ for $n > 2$. The n -dependent phase transition lines for R_n and I_n are shown in figure 10, taking the entropy density in A to be of the form (4.76). Analytic continuation of the condition $\mathcal{Z}_n^{(\text{PT})}(\tau_{ES}) = \mathcal{Z}_n^{(\text{PT})}(e)$ suggests that the transition from the NE to the ES phase for the logarithmic negativity takes place when

$$S_{1/2, A_1}^{(\text{eq})} + S_{1/2, A_2}^{(\text{eq})} = \log d_B. \tag{4.86}$$

This analytically continued phase transition line is indicated with the dashed purple line in figure 10. When this transition takes place, the mutual information is still sub-extensive from (4.27). As discussed in appendix C.4, a computation of the resolvent in the “semicircle

approximation” for this case indicates that the naive analytic continuation of the phase transition (4.86), and the analytically continued value (4.84) of the logarithmic negativity in the ES phase, are correct. However, our computation of the full resolvent in a different setup in the following subsection suggests that the semicircle approximation may in principle be insensitive to new phases that cannot be deduced by analytic continuation. Irrespective of whether such new phases also exist in this example, the conclusion that the negativity is extensive in a regime where the mutual information is sub-extensive holds.

4.5.3 A_2 at infinite temperature

Consider an inhomogeneous example, where the system A_2 is taken to be at infinite temperature while A_1B is a homogenous system that satisfies energy conservation and equilibrates to the microcanonical ensemble. More explicitly, we take the effective identity operator to be

$$\mathcal{I}_\alpha = \mathbf{1}_{A_2} \otimes \sum_{E_p^{A_1} + E_r^B \in I_{E,\Delta}} (|p\rangle\langle p|)_{A_1} \otimes (|r\rangle\langle r|)_B. \quad (4.87)$$

As we vary the volumes of the different subsystems to find the phase diagram, we keep the average energy density in A_1B fixed to some value $\frac{E}{V_{A_1} + V_B} = \epsilon$. We denote the infinite temperature equilibrium entropy density in A_2 by s_0 .

Using (2.49), we find that the Renyi mutual informations in this setup are given by

$$I_n(A_1, A_2) = \begin{cases} 0 & S_{n,A_1}^{(\text{eq})} + \log d_{A_2} < S_{n,B}^{(\text{eq})} \\ S_{n,A_1}^{(\text{eq})} + \log d_{A_2} - S_{n,B}^{(\text{eq})}, & S_{n,A_1}^{(\text{eq})} + \log d_{A_2} > S_{n,B}^{(\text{eq})}, \quad S_{n,A_1}^{(\text{eq})} < \log d_{A_2} + S_{n,B}^{(\text{eq})}, \\ & \log d_{A_2} < S_{n,A_1B}^{(\text{eq})} \\ 2 \log d_{A_2} & \log d_{A_2} > S_{n,A_1B}^{(\text{eq})} \\ S_{n,A_1}^{(\text{eq})} + S_{n,A_1B}^{(\text{eq})} - S_{n,B}^{(\text{eq})} & S_{n,A_1}^{(\text{eq})} > \log d_{A_2} + S_{n,B}^{(\text{eq})} \end{cases} \quad (4.88)$$

Note that this differs slightly from (4.28), as the Renyi entropies for $n \neq 1$ in A_1B are not additive in this example. For $I(A_1, A_2)$, we recover the general expression (4.27). The phase transitions for different I_n are n -dependent in this example.

We find that the dominant permutation in $\mathcal{Z}_n^{(\text{PT})}$ for even n is given by one of the choices $\tau = e$, $\tau = \eta, \eta^{-1}$ or $\tau = \tau_{ES}$,¹⁶ and the corresponding values are

$$\log Z_n^{(\text{PT})}(e) = (1 - n) (\log d_{A_2} + S_{n,A_1}^{(\text{eq})}), \quad (4.89)$$

$$\log Z_n^{(\text{PT})}(\eta^{-1}) = S_{\frac{1}{2},A_1}^{(\text{eq})} - (n - 1)(S_B^{(\text{eq})} + S_{A_1}^{(\text{eq})}), \quad (4.90)$$

$$\log Z_n^{(\text{PT})}(\eta) = (2 - n) \log d_{A_2} + (1 - n) S_{B,n}^{(\text{eq})}, \quad (4.91)$$

$$\log Z_n^{(\text{PT})}(\tau_{ES}) = \left(1 - \frac{n}{2}\right) \log d_{A_2} + S_{\frac{n}{n+2},A_1}^{(\text{eq})} - \frac{n}{2}(S_{A_1}^{(\text{eq})} + S_B^{(\text{eq})}). \quad (4.92)$$

¹⁶We checked this numerically for $n = 4$ and $n = 6$ for the entropy density function in (4.76).

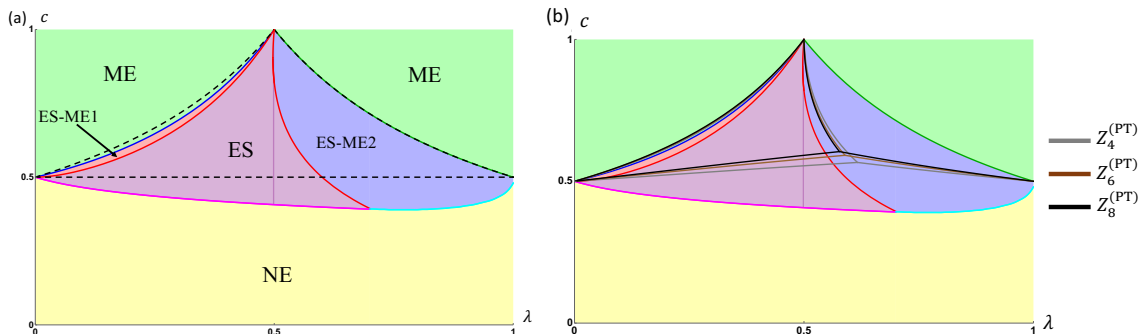


Figure 11. (a) shows the phases of \mathcal{E} with \mathcal{I}_α as in (4.87), $s(\epsilon) = \sqrt{\epsilon}$ in A_1B , and $\frac{E}{V} = 0.5$. The phase transitions of the mutual information are shown with dashed black lines. In (b), we show the same phase transition lines of \mathcal{E} alongside the phase transition lines of $\mathcal{Z}_n^{(\text{PT})}$ for $n = 4, 6, 8$. Note that the precise shape of all phase transition lines will depend on the choices of $s(\epsilon)$ and $\frac{E}{V}$.

From analytic continuation of the above expressions, we get the following forms of the logarithmic negativity:

$$\mathcal{E}_{NE} = 0, \quad (4.93)$$

$$\mathcal{E}_{ME} = \log d_{A_2} \quad \text{or} \quad \mathcal{E}_{ME} = S_{\frac{1}{2}, A_1}^{(\text{eq})}, \quad (4.94)$$

$$\mathcal{E}_{ES} = \frac{1}{2} \log d_{A_2} + S_{\frac{1}{3}, A_1}^{(\text{eq})} - \frac{1}{2} S_{A_1}^{(\text{eq})} - \frac{1}{2} S_B^{(\text{eq})}. \quad (4.95)$$

We also find the following different forms of the Renyi negativities for even n depending on the permutations τ_m and τ'_m that dominate in $\mathcal{Z}_n^{(\text{PT})}$ and $\mathcal{Z}_n^{(A)}$ respectively:

$$R_n(A_1, A_2) = \begin{cases} 0 & \tau_m = e, \tau'_m = e \\ \frac{n-1}{n-2} (S_{n,A}^{(\text{eq})} - S_{n,A}) & \tau_m = e, \tau'_m = \eta \\ \frac{1}{2} \log d_{A_2} + \frac{n}{2(n-2)} (S_{A_1}^{(\text{eq})} + S_B^{(\text{eq})}) & \tau_m = \eta^{-1}, \tau'_m = \eta \\ -\frac{1}{n-2} S_{\frac{n}{n+2}, A_1}^{(\text{eq})} - \frac{n-1}{n-2} S_{n,B}^{(\text{eq})} & \tau_m = \eta, \tau'_m = \eta \\ S_{A_2}^{(\text{eq})} & \tau_m = \eta, \tau'_m = \eta \\ \frac{(n-1)S_{1/n, A_1}^{(\text{eq})} - S_{1/2, A_1}^{(\text{eq})}}{n-2} & \tau_m = \tau_{ES}, \tau'_m = \eta \end{cases} \quad (4.96)$$

The phase transition lines for R_n and I_n for $n = 4, 6$ and 8 are shown in figure 11(a), where we have assumed that the entropy density in A_1B is given by (4.76). Qualitatively, this has a similar structure to the phase diagrams seen in previous cases, with analogs of each of the regions NE, ES1, ES, ME1, ME, ab , bc , ac and abc that we saw previously. For this example, we are further able to use the resolvent to find the full phase diagram of the logarithmic negativity. We present the details of this calculation in appendix C.5. The resulting phase diagram is shown in figure 11(b). Each of the five different shaded regions, labelled NE, ES, ME, ES-ME1 and ES-ME2, represents a distinct phase of the logarithmic negativity. The phase transition lines for the mutual information are also shown in the

same figure with dashed lines, but the corresponding regions are not shaded. In the top row of figure 12, both \mathcal{E} and I are shown as functions of the parameter $c = \frac{S_A^{(\text{eq})}}{S_A^{(\text{eq})} + S_B^{(\text{eq})}}$ along different vertical lines marked in the phase diagram of figure 11(b), which have different fixed values of the ratio $\lambda = \frac{S_{A_1}^{(\text{eq})}}{S_A^{(\text{eq})}}$. The various phases shown in the diagram should be interpreted as follows:

1. In the phases labelled by NE, ES, and ME, \mathcal{E} is given by the analytically continued expressions (4.93)–(4.95).
2. In the phases labelled by ES-ME1 and ES-ME2, \mathcal{E} cannot be obtained by analytic continuation of $Z_n^{(\text{PT})}$ or \mathcal{E} . The expressions for the negativity in these phases are given respectively by

$$\mathcal{E}_{ES-ME1} = S_{A_2}^{(\text{eq})} + V_{A_1}(s(\theta_3) - s(\epsilon)) - V_B s(\epsilon). \quad (4.97)$$

and

$$\mathcal{E}_{ES-ME2} = V_{A_1}(2s(\theta_2) - s(\epsilon)) - V_B s(\epsilon). \quad (4.98)$$

where θ_2 and θ_3 are solutions to the equations

$$\begin{aligned} V_{A_2} s_0 + V_B s \left(\frac{\epsilon(V_{A_1} + V_B) - V_{A_1} \theta_2}{V_B} \right) &= V_{A_1} s(\theta_2), \\ V_{A_1} s(\theta_3) + V_B s \left(\frac{\epsilon(V_{A_1} + V_B) - V_{A_1} \theta_3}{V_B} \right) &= V_{A_2} s_0. \end{aligned} \quad (4.99)$$

3. Only one of the phase transition lines shown in the figure — the pink line dividing the NE and ES phases — can be correctly deduced by naive analytic continuation. Putting $n = 1$ in the condition $Z_n^{(\text{PT})}(\tau_{ES}) \geq Z_n^{(\text{PT})}(\epsilon)$, we would deduce that the transition line between the NE and ES phases for \mathcal{E} is given by

$$\frac{1}{2} S_{A_2}^{\text{eq}} + S_{1/3, A_1}^{\text{eq}} - \frac{1}{2} S_{A_1}^{\text{eq}} - \frac{1}{2} S_B^{\text{eq}} = 0 \quad (4.100)$$

Eq. (4.100) agrees with the pink line in figure 11(b), but it cannot be used to determine the point where the pink line ends and the cyan line begins, as the naive analytic continuation does not “know about” the existence of the ES-ME2 phase. The semi-circle approximation for the resolvent in this example also correctly predicts the pink line for the transition from NE to ES, but is again insensitive to the ES-ME2 phase.

4. In the transition from NE to ES or ES-ME2, there is a discontinuity in the first derivative of \mathcal{E} with respect to c . The remaining phase transitions correspond to a discontinuity in the second derivative. We show $\frac{\partial \mathcal{E}}{\partial c}$ along different vertical lines in the phase diagram in the bottom row of figure 12. This is unlike the infinite temperature case, where all phase transitions involve discontinuities in the first derivative of \mathcal{E} .
5. In the region between the dashed black line and the pink and cyan lines, the mutual information is sub-extensive while the negativity is extensive. Comparing (4.100)

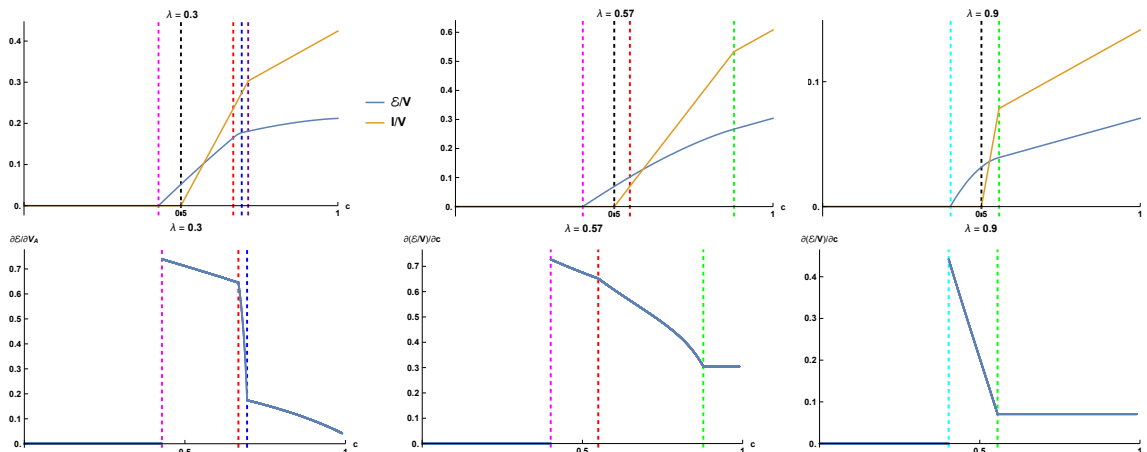


Figure 12. Top row: the blue and orange curves respectively show \mathcal{E} and I as a function of c along three vertical lines in the phase diagram of figure 11(b). Bottom row: $\frac{\partial \mathcal{E}}{\partial c}$ is shown along the same lines. The relevant phase transitions lines from figure 11(b) are also indicated with dashed vertical lines.

with the NE to ES transition line for the mutual information in (4.57), we see quite generally that since $S_{1/3, A_1}^{(eq)} > S_{A_1}^{(eq)}$, such a regime should exist for any choice of $s(\epsilon)$, as long as the naive transition line for the logarithmic negativity in (4.100) is correct.

Note also that the NE-ES transition line for $I_{1/2}$ lies slightly below, but quite close to, that of the mutual information. In particular, it also lies above the lines where the logarithmic negativity starts to become extensive.

4.6 Numerical examples: negativity spectrum in chaotic spin chains

Throughout this work, we have used an analytic approximation that holds in the thermodynamic limit to calculate various information-theoretic quantities. In this subsection, we present results for the spectrum of the partially transposed density matrix in the canonical ensemble. Here we avoid all approximations and carry out numerical simulations in chaotic spin chains, but are necessarily limited to small sizes.

We consider a chaotic 1D spin chain with Hamiltonian

$$H = - \sum_{i=1}^N (Z_i Z_{i+1} + h_x X_i + h_z Z_i). \tag{4.101}$$

This Hamiltonian has been studied frequently as it undergoes a well-understood integrable to chaotic transition [47]. We find, numerically, that this model has the following partition function

$$g(\beta) = \frac{\log Z(\beta)}{N} = \log (2 \cosh (f(h_x, h_z)\beta)) \tag{4.102}$$

to extremely precise accuracy. Here, there is only a single fitting parameter, f , that depends on the coupling constants chosen. This can be fixed by the low-temperature behavior of

the partition function

$$\frac{\log Z(\beta \rightarrow \infty)}{N} = -\frac{\beta E_0}{N} = \beta f(h_x, h_z). \quad (4.103)$$

Thus, we have

$$g(\beta) = \log (2 \cosh (\varepsilon_0 \beta)), \quad (4.104)$$

where ε_0 is the ground state energy density. We set $h_x = -1.05$ and $h_z = 0.5$ in our calculations below.

If we want to model the canonical ensemble, it is convenient to study the “canonical thermal pure quantum states” of ref. [48]

$$|\beta\rangle \propto e^{-\beta H/2} |R\rangle, \quad (4.105)$$

where $|R\rangle$ is a random state vector. Being random, it is approximately an equal weighting of all energy eigenstates

$$|\beta\rangle \simeq \frac{1}{\sqrt{Z(\beta)}} \sum_{E_n} e^{-\beta E_n/2} |E_n\rangle. \quad (4.106)$$

The exponential of the Hamiltonian is a dense matrix, so a priori, we will be very limited in the system sizes. This can be improved by approximating the exponential as

$$e^{-\beta H/2} \simeq e^{-\beta(H_A \otimes \mathbb{1}_B)/2} e^{-\beta(\mathbb{1}_A \otimes H_B)/2}, \quad (4.107)$$

where, as in the equilibrium approximation, we have considered the coupling H_{AB} to be small. A Haar random vector can be written as [49]

$$|R\rangle = \sum_{i\alpha} X_{i\alpha} |i\rangle_A \otimes |\alpha\rangle_B, \quad (4.108)$$

where the $X_{i\alpha}$ ’s are independent complex Gaussian random variables and $|i\rangle, |\alpha\rangle$ are orthonormal bases on the sub-Hilbert spaces. The (unnormalized) reduced density matrix on subsystem A may then be expressed as

$$\rho_A(\beta) = e^{-\beta H_A/2} X X^* e^{-\beta H_A/2}, \quad (4.109)$$

where X is now the $d_A \times d_B$ rectangular matrix with matrix elements $X_{i\alpha}$. This is a great numerical simplification because we avoid ever forming dense, $(d_A d_B) \times (d_A d_B)$ matrices.

In figure 13, we show the exact negativity spectra at both infinite and finite temperature. Certain aspects of the infinite temperature spectra are robust such as the topological transition from the single connected component to two connected components. However, other details are not robust. For example, when all systems are at finite temperature, the spectrum never resembles a semi-circle.

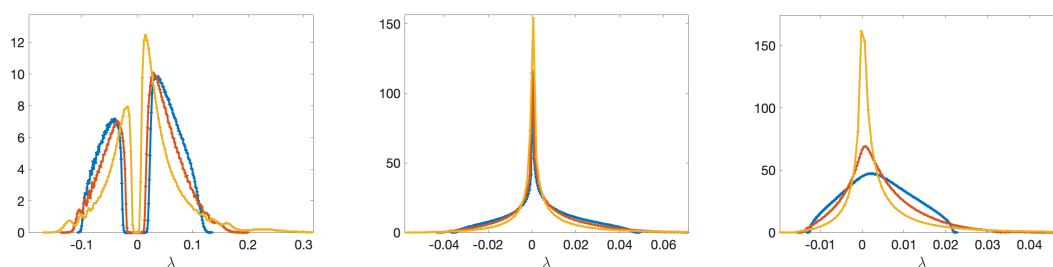


Figure 13. The probability density functions for the negativity spectra at $|\beta\epsilon_0| = 0$ (blue), $\frac{1}{4}$ (orange), and $\frac{1}{2}$ (yellow). From left to right, we have $N_B = 2, 4,$ and 6 . The topological transition is robust to finite temperature. We always take $N_{A_1} = 2$ and $N_{A_2} = 6$.

5 Implications for mixed-state entanglement and black hole radiation

We now discuss the operational implications of the above results for \mathcal{E} and I . As reviewed in section 2.2, the mutual information gives an upper bound the distillable entanglement

$$E_d(A_1, A_2) \leq \frac{1}{2}I(A_1, A_2), \tag{5.1}$$

while the logarithmic negativity gives a lower bound on the exact PPT entanglement cost

$$E_c^{(\text{ppt,exact})}(A_1, A_2) \geq \mathcal{E}(A_1, A_2). \tag{5.2}$$

We observed that at a finite temperature, for various choices of $\rho^{(\text{eq})}$, there is a region in the entanglement phase diagram where \mathcal{E} is volume-like while I is sub-volume, and hence $\mathcal{E} \gg \frac{1}{2}I$. Our results suggest that such behavior is likely to be generic for all finite temperature chaotic systems. In this part of the phase diagram, we then have

$$E_c^{(\text{exact})} \gg E_d^{(\text{exact})}. \tag{5.3}$$

If we further assume that $E_c^{(\text{ppt})}$ is extensive when $E_c^{(\text{ppt,exact})}$ is extensive, then a stronger statement can be made,

$$E_c \gg E_d. \tag{5.4}$$

This implies that the preparation of the state of A from EPR pairs by LOCC operations is highly irreversible, and it has significant bound entanglement.

It is interesting to contrast the results here with those from random states in [50], which can be seen as thermalized states at infinite temperature. In such states, a regime was found where E_c is extensive while I (and hence E_d) is sub-extensive. However, this effect took place for a range of subsystem dimensions which all correspond to $c = \frac{1}{2}$ in the thermodynamic limit, as opposed to an $O(1)$ range of c between some c_0 and $\frac{1}{2}$, as we find in the finite temperature examples here. The conclusion about E_c in [50] was drawn using entanglement of formation as opposed to logarithmic negativity. On calculating \mathcal{E} in the infinite temperature case in the same regime at $c = \frac{1}{2}$, we do not find that it becomes extensive while I is still sub-extensive.

We can ask whether the preparation of the state ρ_A by PPT operations, as opposed to LOCC, is also irreversible in this regime. This question cannot be conclusively answered from our results. With the assumption that $E_c^{(\text{ppt})}$ is extensive when $E_c^{(\text{ppt,exact})}$ is extensive, there are two logical possibilities, both with interesting physical consequences:

1. If $E_d^{(\text{ppt})}$ is comparable to \mathcal{E} , i.e. volume-like, as in some known systems [7], then although LOCC distillable entanglement is small, PPT distillable entanglement is large, implying that finite temperature greatly enhances distillable entanglement if we use PPT operations.
2. If $E_d^{(\text{ppt})}$ is comparable to I , i.e. sub-volume-like, then both the PPT and LOCC distillable entanglements are small, and the state is also highly PPT-irreversible.

The results of the previous section apply to generic chaotic many-body systems, and hence also have implications for an evaporating black hole, which is believed to be highly chaotic.

Consider a black hole formed from the collapse of a star initially in a pure state. The black hole subsequently evaporates by emitting Hawking radiation. The evaporation process is very slow in microscopic scales, and we can treat the remaining black hole as well as the radiation as being in macroscopic equilibrium.¹⁷ The full system is still in a pure state, and the reduced density operator of each subsystem can be far from the density operators of thermal ensembles. The results of the previous section can be used to predict the entanglement structure *within* the radiation subsystem, or entanglement between a part of radiation and the remaining black hole, at both infinite and finite temperatures. For definiteness, we will take A to be the radiation and B to be the black hole.

In [51], from averaging over random pure states, Page found that the entanglement entropy of the radiation undergoes a transition from increasing to decreasing behavior at some time scale $t = t_p$ where $\log d_A = \log d_B$. Page’s calculation was in the infinite temperature case, where from our discussion in section 4.2, we know that t_p is also the time scale at which entanglement within the radiation, as quantified by either I or \mathcal{E} , starts to become extensive. In other words, before the Page time t_p there is no entanglement within the radiation.

The natural finite-temperature generalization of the Page time is when $S_A^{(\text{eq})} = S_B^{(\text{eq})}$, corresponding to $c = \frac{1}{2}$, which was already used in [52]. Our results from the equilibrium approximation and resolvent calculations confirm that this is indeed the time at which the entanglement entropy transitions from increasing to decreasing behavior. Interestingly, our results for logarithmic negativity give a new prediction for the quantum-informational properties of the radiation at finite temperature: there are significant entanglement correlations within the radiation long before the Page time. This suggests the existence of another time scale t_b when quantum entanglement within the radiation starts becoming extensive. For time scales $t \in (t_b, t_p)$, the entanglement correlations within the radiation cannot be distilled using LOCC, i.e. they are an example of “bound entanglement.” It

¹⁷The two systems can be in separate macroscopic equilibria, because in general the radiation is separated from the black hole after it is emitted.

is an interesting open question whether this entanglement could be distillable using more general operations such as PPT operations.

Our results from the previous section can also be used to derive replica wormhole contributions to the calculation of the negativity between parts of the radiation in Euclidean gravity. We show these replica wormholes explicitly for the model of [4] in appendix C.3.1. (See figure 26.) Due to specific features of the density of states in JT gravity, the difference between t_b and t_p is suppressed in the large quantity S_0 in this model. However, this difference should be macroscopically large for higher-dimensional black holes.

To make predictions for the mixed-state entanglement between the black hole and some part of the radiation, we can take A_1 and B to be parts of the radiation and A_2 to be the black hole.

To explore further the nature of entanglement among various parts of the full black hole system, it is instructive to consider the Hayden-Preskill thought experiment at a finite temperature: we throw a diary into the black hole, and see when the information of the diary is recoverable from the radiation. We turn to this problem next. We will find that both time scales t_p and t_b can be given operational interpretations in the context of this question.

6 Implications for information transfer from black hole to radiation

In this section, we investigate how information is transferred from the black hole to the radiation during the evaporation process using the Hayden-Preskill thought experiment [10]. Suppose we throw a secret diary D into a black hole B at an early stage in the evaporation process. As the black hole evaporates, we collect all of the radiation and refer to it as R , and refer to the remaining black hole as B' . Suppose we know the initial state of B , and have access to a universal quantum computer with which we can act on the radiation R . We then ask how much of the radiation we need in order to learn the initial state of the diary, ρ_D .

We can make this question more explicit in the following way. The relevant decompositions of the Hilbert space at the initial and final times respectively are $\mathcal{H} = \mathcal{H}_D \otimes \mathcal{H}_B$ and $\mathcal{H} = \mathcal{H}_{B'} \otimes \mathcal{H}_R$. We can consider the following quantum channel from \mathcal{H}_D to \mathcal{H}_R ,

$$\mathcal{N}(\rho_D) = \text{Tr}_{B'} \left[U (\rho_D \otimes \rho_B) U^\dagger \right], \tag{6.1}$$

where U is the time-evolution operator for the black hole and radiation, together with the diary, and $\rho_B = |\psi_0\rangle \langle \psi_0|_B$ is some fixed initial pure state of the black hole. Asking whether we can learn the state of the diary from the radiation is equivalent to asking whether there exists a universal recovery channel \mathcal{R} such that $\mathcal{R} \circ \mathcal{N}(\rho_D) = \rho_D$ for all ρ_D .

Note that while we will mostly discuss the above question in terms of evaporating black holes in this section, it can also be seen as a more general question about thermalized states in quantum many-body systems. Suppose we put some information in a small subsystem D , and then let D evolve together with the rest of the system, B . Then given some subsystem R in the thermalized state, can we learn the initial state of D ?

We will take three different approaches to this question, each of which reveals different aspects of the transfer of information from the black hole to the radiation. We will use the equilibrium approximation in each of these approaches, which allows us to understand the

transfer of information both at infinite and finite temperature. Note that all statements at finite temperature below are based on the canonical ensemble universality class, where we take $\mathcal{I}_\alpha = e^{\beta_1 H_{B'}} \otimes e^{\beta_2 H_R}$. We will assume that the Hilbert space of the diary has a finite dimension d_D . We will in general consider the case where the diary can be large, so that d_D can be $O(e^{1/G_N})$.

1. We first introduce a reference system Q with the same Hilbert space dimension as D , and consider an initial state in which D is maximally entangled with Q . The time-evolution on \mathcal{H}_Q is trivial, while $\mathcal{H}_B \otimes \mathcal{H}_D$ is evolved with U . We then look at the mutual information $I(Q, R)$ under time-evolution, which can be seen as a way of quantifying the extent to which the radiation contains information about the diary. The behavior of this quantity motivates us to define two natural time-scales:

- t_{p_1} , the time at which $S_R^{(\text{eq})} = S_{B'}^{(\text{eq})} - \log d_D$. $I(Q, R)$ starts increasing from its initial value of zero at this time.
- t_{p_2} , at which $S_R^{(\text{eq})} = S_{B'}^{(\text{eq})} + \log d_D$. $I(Q, R)$ reaches its maximal value of $2 \log d_D$ at this time.

The standard Page time t_p , at which we have $S_R^{(\text{eq})} = S_{B'}^{(\text{eq})}$, lies between these two time scales.

When $I(Q, R)$ reaches its maximal value at t_{p_2} , this can be interpreted by saying that all the information that was initially in the diary is now present in the radiation. This statement can be understood operationally in terms of the quantum channel \mathcal{N} in (6.1). From the results of [53], $I(Q, R) = 2 \log d_D$ implies the existence of a universal recovery channel \mathcal{R} for \mathcal{N} , i.e. $\mathcal{R} \circ \mathcal{N}(\rho_D) = \rho_D$ for all ρ_D . This is consistent with the fact that from a gravitational perspective, t_{p_2} is the latest time at which an island can form [2] (see figure 14).

2. For some choice of reference state σ_D , and any state ρ_D , we compute the difference in the relative entropies before and after applying the channel \mathcal{N} ,

$$D(\mathcal{N}(\rho_D) || \mathcal{N}(\sigma_D)) - D(\rho_D || \sigma_D), \tag{6.2}$$

in order to find a lower bound for the fidelity of some recovery channel \mathcal{R} using (2.65). The relevant time scales for this quantity turn out in general to depend on the choice of σ_D and ρ_D , and are given by

- $t_p(\rho_D)$, the time at which $S_R^{(\text{eq})} = S_{B'}^{(\text{eq})} + S(\rho_D)$. The lower bound on the fidelity of some recovery channel \mathcal{R} first starts to increase from an exponentially small value at this time.
- $t_{p_2}(\sigma_D, \rho_D)$, at which $S_R^{(\text{eq})} = S_{B'}^{(\text{eq})} + D(\rho_D || \sigma_D) + S(\rho_D) = S_{B'}^{(\text{eq})} - \text{Tr}[\rho_D \log \sigma_D]$. The lower bound on the fidelity reaches its maximal value of 1 at this time-scale.

Note that if we take ρ_D to be pure, then $t_p(\rho_D) = t_p$, the standard Page time. If we take $\sigma_D = \mathbf{1}/d_D$, then $t_{p_2}(\sigma_D, \rho_D)$ becomes equal to t_{p_2} defined in the previous point

independently of ρ_D , implying that universal recovery is possible after t_{p_2} . Recall that in the previous point, universal recovery at t_{p_2} was deduced from the complementary perspective that the mutual information $I(Q, R)$ becomes maximal at this time.

The lower bound on the fidelity provides an operational way of seeing the gradual transfer of information from the black hole to the radiation between times t_p and t_{p_2} , as the fidelity increases from its minimal value to one in this range of times. However, this quantity does not seem to have a regime which reflects the growth of $I(Q, R)$ from time t_{p_1} to t_p which we observed in the previous point.

3. We explicitly calculate the fidelity of the Petz recovery map \mathcal{P} for \mathcal{N} , taking the initial state ρ_D to be pure to simplify calculations. This reveals the following new time scale, which the lower bound in terms of the relative entropy from the previous point is not sensitive to:

- t_b , the time at which $S_{\frac{1}{2}, R}^{(\text{eq})} = S_{2, B'}^{(\text{eq})}$. For a large diary, the fidelity of the Petz map starts growing from its initial value of $F(\rho_D || \sigma_D)$ at this time. Recall from our discussion in section 4.4.1, that this is also the time scale at which the logarithmic negativity between two parts of the radiation starts to become extensive in the canonical ensemble.¹⁸

The time at which the fidelity saturates to unity is t_{p_2} , consistent with the conclusions of the two other approaches in the previous points.

Note that at infinite temperature and for small diaries such that $\log d_D \ll S_{B'}^{(\text{eq})}, S_R^{(\text{eq})}$, all the different time scales above coincide, i.e. $t_{p_1} = t_b = t_p = t_{p_2}$. More generally, allowing for finite temperature and large diaries, we have $t_{p_1} \leq t_p \leq t_{p_2}$ and $t_b \leq t_p \leq t_{p_2}$.

6.1 From the perspective of mutual information

We first discuss the transfer of information from the perspective of the mutual information of the radiation with a reference system Q . This approach was considered at infinite temperature in the original discussion of [10] by using random states, and has been recently generalized to generic chaotic systems at infinite temperature using properties of operator growth [54]. We now provide a generalization to finite temperature using the equilibrium approximation.

We take the initial state of the full system to be

$$|\Psi_0\rangle = \frac{1}{\sqrt{d_D}} \sum_{n=0}^{d_D-1} |n\rangle_Q \otimes |n\rangle_D \otimes |\psi_0\rangle_B. \tag{6.3}$$

We will take the time evolution in Q to be trivial, so that the n -th Renyi entropy of the time-evolved state in subsystem A is given by

$$\mathcal{Z}_{n,A} = \langle \eta_A \otimes e_{\bar{A}} | (\mathbf{1} \otimes \mathbf{1})_Q^n \otimes (U \otimes U^\dagger)_{DB}^n | \rho_0, e \rangle, \quad \rho_0 = |\Psi_0\rangle \langle \Psi_0|. \tag{6.4}$$

¹⁸At the end of this section, we briefly consider an example with the microcanonical ensemble, taking \mathcal{I}_α as in (4.87). For this setup, we find that the fidelity of a large diary starts increasing at a time $t_r < t_p$, which can be either earlier or later than t_b .

From the fact that the time-evolution in Q is trivial, we immediately have at all times

$$S_{n,Q} = \log d_D. \quad (6.5)$$

With the initial state (6.3), (6.4) for $A = R$ becomes

$$\mathcal{Z}_{n,R} = \frac{1}{d_D^n} \langle e_{B'} \otimes \eta_R | (U \otimes U^\dagger)^n | (\tilde{\rho}_0)_B \otimes \frac{\mathbf{1}_D}{d_D}, e \rangle \quad (6.6)$$

where $\tilde{\rho}_0 = |\psi_0\rangle\langle\psi_0|$. Note that (6.6) can also be seen as the time-evolution of the entanglement entropy of a mixed state with initial entropy $\log d_D$. We can then use (2.26) to find

$$S_{n,R} = \begin{cases} S_{n,R}^{(\text{eq})} & S_{n,R}^{(\text{eq})} < S_{n,B'}^{(\text{eq})} + \log d_D \\ S_{n,B'}^{(\text{eq})} + \log d_D & S_{n,R}^{(\text{eq})} > S_{n,B'}^{(\text{eq})} + \log d_D \end{cases} \quad (6.7)$$

Since the state on the full system is pure,

$$\mathcal{Z}_{n,QR} = \mathcal{Z}_{n,B'} = \langle \eta_{B'} \otimes e_R | (U \otimes U^\dagger)^n | (\tilde{\rho}_0)_B \otimes \frac{\mathbf{1}_D}{d_D}, e \rangle. \quad (6.8)$$

Again using (2.26), we have

$$S_{n,QR} = S_{n,B'} = \begin{cases} S_{n,R}^{(\text{eq})} + \log d_D & S_{n,R}^{(\text{eq})} + \log d_D < S_{n,B'}^{(\text{eq})} \\ S_{n,B'}^{(\text{eq})} & S_{n,R}^{(\text{eq})} + \log d_D > S_{n,B'}^{(\text{eq})} \end{cases}. \quad (6.9)$$

The n -th Renyi mutual information between Q and R is then given by

$$I_n(Q, R) = \begin{cases} 0 & S_{n,R}^{(\text{eq})} < S_{n,B'}^{(\text{eq})} - \log d_D \\ \log d_D + S_{n,R}^{(\text{eq})} - S_{n,B'}^{(\text{eq})} & S_{n,B'}^{(\text{eq})} - \log d_D < S_{n,R}^{(\text{eq})} < S_{n,B'}^{(\text{eq})} + \log d_D \\ 2 \log d_D & S_{n,R}^{(\text{eq})} > S_{n,B'}^{(\text{eq})} + \log d_D \end{cases} \quad (6.10)$$

and by analytic continuation,

$$I(Q, R) = \begin{cases} 0 & S_R^{(\text{eq})} < S_{B'}^{(\text{eq})} - \log d_D \\ \log d_D + S_R^{(\text{eq})} - S_{B'}^{(\text{eq})} & S_{B'}^{(\text{eq})} - \log d_D < S_R^{(\text{eq})} < S_{B'}^{(\text{eq})} + \log d_D \\ 2 \log d_D & S_R^{(\text{eq})} > S_{B'}^{(\text{eq})} + \log d_D \end{cases}. \quad (6.11)$$

Note that at all times, $I(Q, B') = 2 \log d_D - I(Q, R)$. From [53], when $I(Q, R) = 2 \log d_D$ and $I(Q, B') = 0$, there exists a recovery channel \mathcal{R} satisfying $\mathcal{R} \otimes \mathcal{N}(\rho_D) = \rho_D$ for all ρ_D under the channel \mathcal{N} defined in (6.1). Hence, the universal recovery channel exists after a time scale t_{p_2} , at which $S_R^{(\text{eq})} = S_{B'}^{(\text{eq})} + \log d_D$. This result from the equilibrium approximation is consistent with the expectations from [11] and [2].

Note from (6.11) that while the mutual information between the radiation and the auxiliary system becomes maximal at t_{p_2} , it begins to grow from zero at an earlier time t_{p_1} , at which $S_R^{(\text{eq})} = S_B^{(\text{eq})} - \log d_D$.

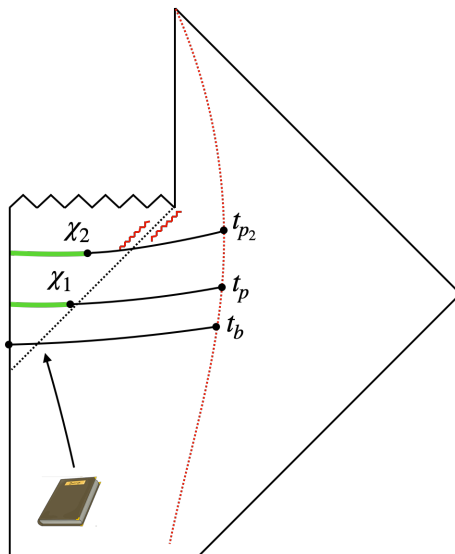


Figure 14. A Penrose diagram is shown for an evaporating black hole with a large diary thrown into it. We ask when the diary can be reconstructed from the radiation (the fields to the right of the dotted red line). At early times, the entanglement wedge of the radiation is trivial, so the diary cannot be recovered. If the diary was initialized in a pure state, an island (green) forms at t_p where the nontrivial extremal surface, χ_1 , becomes smaller than the entropy of the radiation. If the diary was initialized in the maximally mixed state, the island will not form until t_{p_2} because the generalized entropy includes the entropy of the diary. While the diary cannot be reconstructed with $O(1)$ fidelity until the island forms, the fidelity exponentially increases from its minimal value starting at the earlier time t_b .

6.2 Recovery channel: bound from relative entropy

We now consider the evolution of the fidelity of recovery channels during the evaporation process. In this subsection, we use (2.65) to put a lower bound on the fidelity of some recovery channel \mathcal{R} , by calculating the change of the relative entropy under \mathcal{N} using the equilibrium approximation. In next subsection, we calculate the fidelity of the Petz map explicitly. Note that in both this subsection and the next, we do not include a reference system Q , and the input state for \mathcal{N} is simply ρ_D .

6.2.1 Infinite temperature

For simplicity, let us first consider the infinite temperature case, and take σ_D to be $\mathbf{1}_D/d_D$ and ρ_D to be a pure state. Then for the first term in the expression (3.19) for $D(\mathcal{N}(\rho_D)||\mathcal{N}(\sigma_D))$, in the infinite temperature case we can sum over all permutations to find (see for instance [1])

$$\text{Tr} [\mathcal{N}(\rho_D)^n] = \frac{1}{(d_R d_{B'})^n} \sum_{\tau \in S_n} d_R^{k(\eta^{-1}\tau)} d_{B'}^{k(\tau)} = d_R^{1-n} {}_2F_1 \left(1-n, -n, 2, \frac{d_R}{d_{B'}} \right) \quad (6.12)$$

where we consider the regime where d_R and $d_{B'}$ are both large, but with no restrictions on the relative sizes. The resulting von Neumann entropy coincides with Page's formula [51]

$$S(\mathcal{N}(\rho_D)) = \lim_{n \rightarrow 1} \frac{1}{1-n} \log \text{Tr} [\mathcal{N}(\rho_D)^n] = \begin{cases} \log d_R - \frac{d_R}{2d_{B'}}, & d_R < d_{B'} \\ \log d_{B'} - \frac{d_{B'}}{2d_R}, & d_R > d_{B'} \end{cases}. \quad (6.13)$$

Using (3.24), the second term in (3.19) is given by a similar sum

$$\text{Tr} [\mathcal{N}(\rho_D) \mathcal{N}(\sigma_D)^{n-1}] = \frac{1}{(d_D d_R d_{B'})^n} \sum_{\tau \in S_n} d_R^{k(\eta^{-1}\tau)} (d_D d_{B'})^{k(\tau)}. \quad (6.14)$$

This can be interpreted as the entropy of the radiation if it were coupled to an additional bath with a Hilbert space dimension identical to that of the diary. The $n \rightarrow 1$ limit is given by

$$\lim_{n \rightarrow 1} \frac{1}{1-n} \log \text{Tr} [\mathcal{N}(\rho_D) \mathcal{N}(\sigma_D)^{n-1}] = \begin{cases} \log d_R - \frac{d_R}{2d_{B'} d_D}, & d_R < d_{B'} d_D \\ \log d_{B'} d_D - \frac{d_{B'} d_D}{2d_R}, & d_R > d_{B'} d_D \end{cases}. \quad (6.15)$$

The relative entropy is therefore

$$D(\mathcal{N}(\rho_D) || \mathcal{N}(\sigma_D)) = \begin{cases} \frac{d_R(d_D-1)}{2d_{B'} d_D}, & d_R < d_{B'} \\ \log \frac{d_R}{d_{B'}} + \frac{d_{B'}}{2d_R} - \frac{d_R}{2d_{B'} d_D}, & d_{B'} < d_R < d_{B'} d_D \\ \log d_D + \frac{d_{B'}(1-d_D)}{2d_R}, & d_R > d_{B'} d_D \end{cases}. \quad (6.16)$$

Plugging back into (2.65), we obtain a lower bound on the fidelity:

$$F(\rho, [\mathcal{R}_{\sigma, \mathcal{N}} \circ \mathcal{N}](\rho)) \geq \begin{cases} \frac{1}{d_D} \exp\left(\frac{d_R(d_D-1)}{2d_{B'} d_D}\right), & d_R < d_{B'} \\ \frac{d_R}{d_{B'} d_D} \exp\left(\frac{d_{B'}}{2d_R} - \frac{d_R}{2d_{B'} d_D}\right), & d_{B'} < d_R < d_{B'} d_D \\ \exp\left(\frac{d_{B'}(1-d_D)}{2d_R}\right), & d_R > d_{B'} d_D \end{cases}. \quad (6.17)$$

The transitions between the different lines of (6.17) occur at the infinite-temperature versions of the times t_p and t_{p_2} respectively. The fidelity is exponentially small prior to t_p . After t_p , the fidelity is still small but exponentially increases until it reaches a value close to one at t_{p_2} . While the formula for $\exp(D(\mathcal{N}(\rho_D) || \mathcal{N}(\sigma_D)) - D(\rho_D || \sigma_D))$ on the r.h.s. of (6.17) is only exact in the limit of large Hilbert space dimensions, we find that it is remarkably accurate even for small system sizes in figure 15.

6.2.2 Finite temperature

Let us now consider finite temperature, and allow for general choices of the initial state of the diary ρ_D and the reference state σ_D . Now, in general, we can no longer perform the sum over all permutations appearing in the equilibrium approximation. We can instead

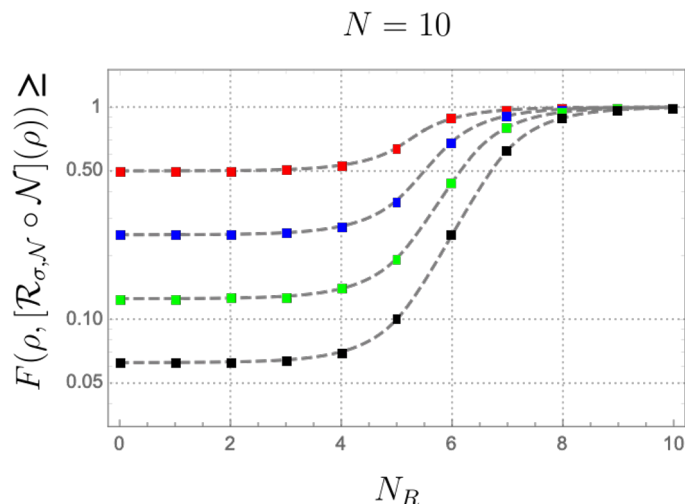


Figure 15. The lower bound on the fidelity of the Petz map to recover the diary thrown into the black hole from the radiation as a function of the number of qubits N_R in the radiation, taking the total number of qubits in the system to be $N = 10$. The diary is composed of one (red), two (blue), three (green), or four (black) qubits. The dashed lines are the analytic predictions calculated at large N , (6.17), although the agreement is already very precise. In the numerical data, we take a fiducial pure state tensored with the diary state and apply a random unitary matrix, computing the relative entropy before and after the random channel. We take 10^3 disorder realizations. Error bars are plotted, though barely visible due to small fluctuations.

use the formula (3.31) based on the dominant permutations in different regimes, to find¹⁹

$$D(\mathcal{N}(\rho_D) || \mathcal{N}(\sigma_D)) \simeq \begin{cases} 0, & S_R^{(\text{eq})} < S_{B'}^{(\text{eq})} + S(\rho_D) \\ S_R^{(\text{eq})} - S_{B'}^{(\text{eq})} - S(\rho_D), & S_{B'}^{(\text{eq})} + S(\rho_D) < S_R^{(\text{eq})} \\ & < S_{B'}^{(\text{eq})} + D(\rho_D || \sigma_D) + S(\rho_D) \\ D(\rho_D || \sigma_D), & S_R^{(\text{eq})} > S_{B'}^{(\text{eq})} + D(\rho_D || \sigma_D) \end{cases} . \quad (6.18)$$

As an aside, note that we can interpret this result in terms of quantum hypothesis testing, from the standard operational interpretation of the relative entropy. Say Alice were to ask Bob a yes or no question and Bob responded to Alice by writing his answer in his diary (i.e. encoding one of two states ρ_D and σ_D in the diary), then throwing it into the black hole, knowing that Alice would collect the radiation from the black hole in order to decode his message using a quantum measurement. If this is repeated over and over, the error rate of Alice misidentifying Bob’s answer is given by $e^{-D(\mathcal{N}(\rho_D) || \mathcal{N}(\sigma_D))}$.

¹⁹Note that this formula breaks down and becomes infinite in the large R regime if the initial diary states ρ_D and σ_D are close to orthogonal. This is guaranteed not to be the case if we choose σ_D to be full rank.

Eq. (6.18) implies the following bound on the fidelity of recovery on using (2.65):

$$F(\rho, [\mathcal{R}_{\sigma, \mathcal{N}} \circ \mathcal{N}](\rho)) \geq \begin{cases} e^{-D(\rho_D || \sigma_D)}, & S_R^{(\text{eq})} < S_{B'}^{(\text{eq})} + S(\rho_D) \\ e^{S_R^{(\text{eq})} - S_{B'}^{(\text{eq})} - S(\rho_D) - D(\rho_D || \sigma_D)}, & S_{B'}^{(\text{eq})} + S(\rho_D) < S_R^{(\text{eq})} \\ & < S_{B'}^{(\text{eq})} + D(\rho_D || \sigma_D) + S(\rho_D) \\ 1, & S_R^{(\text{eq})} > S_{B'}^{(\text{eq})} + D(\rho_D || \sigma_D) + S(\rho_D) \end{cases}. \quad (6.19)$$

We now see that the lower bound on fidelity first starts increasing at a time $t_p(\rho_D)$, at which $S_R^{(\text{eq})} = S_{B'}^{(\text{eq})} + S(\rho_D)$, and reaches its maximal value at $t_{p_2}(\sigma_D, \rho_D)$, at which $S_R^{(\text{eq})} = S_{B'}^{(\text{eq})} + D(\rho_D || \sigma_D) + S(\rho_D) = S_{B'}^{(\text{eq})} - \text{Tr}[\rho_D \log \sigma_D]$.

6.3 Recovery channel: the Petz map and its fidelity

We progress to explicitly evaluating the fidelity of the Petz recovery map. The quantum channel that we seek to reverse is

$$\begin{aligned} \mathcal{N} : \mathcal{L}(\mathcal{H}_D) &\rightarrow \mathcal{L}(\mathcal{H}_R) \\ \rho_D &\mapsto \text{Tr}_{B'} \left[U^\dagger (\rho_D \otimes \rho_B) U \right], \end{aligned} \quad (6.20)$$

As a reminder, the Petz map is given by

$$\begin{aligned} \mathcal{P}_{\sigma, \mathcal{N}} : \mathcal{L}(\mathcal{H}_R) &\rightarrow \mathcal{L}(\mathcal{H}_D) \\ X &\mapsto \sigma^{\frac{1}{2}} \mathcal{N}^\dagger \left(\mathcal{N}(\sigma)^{-\frac{1}{2}} X \mathcal{N}(\sigma)^{-\frac{1}{2}} \right) \sigma^{\frac{1}{2}}. \end{aligned} \quad (6.21)$$

6.3.1 Infinite temperature

We first study Petz recovery at infinite temperature, again taking ρ_D to be pure and σ_D to be maximally mixed. Using the equilibrium approximation for this quantity in (3.34) with $\mathcal{I}_\alpha = \mathbf{1}$, we have $F(\rho, [\mathcal{P}_{\sigma, \mathcal{N}} \circ \mathcal{N}](\rho)) = \lim_{m \rightarrow \frac{1}{2}} F_m$, where

$$F_m = \frac{1}{d_D^{2m+3} (d_R d_{B'})^{2m+2}} \sum_{\tau \in \mathcal{S}_{2m+2}} d_R^{k(\eta^{-1}\tau)} d_{B'}^{k(\tau)} d_D^{k(\tau) + \zeta(\tau)} \quad (6.22)$$

where $\zeta(\tau)$ is zero if the first and $(m+2)^{\text{th}}$ elements are in different cycles in the permutation τ , and 1 if they are in the same cycle.

At early times, $d_{B'} \gg d_R$ and the identity element will dominate. The identity has $\zeta(e) = 0$, so that

$$F_m = \frac{1}{d_D d_R^{2m+1}} + O(d_{B'}^{-1}), \quad F(\rho, [\mathcal{P}_{\sigma, \mathcal{N}} \circ \mathcal{N}](\rho)) = \frac{1}{d_D} = F(\rho_D, \sigma_D). \quad (6.23)$$

At late times, when $d_R \gg d_{B'} d_D$ the cyclic permutation will dominate, which has $\zeta(\eta) = 1$, so that

$$F_m = \frac{1}{(d_{B'} d_D)^{2m+1}} + O(d_R^{-1}), \quad F(\rho, [\mathcal{P}_{\sigma, \mathcal{N}} \circ \mathcal{N}](\rho)) = 1. \quad (6.24)$$

To understand how the recovery process improves from (6.23) to (6.24) as the size of the radiation grows, we can perform the full sum in (6.22), as we explain later in this subsection. Before turning to this detailed calculation, we can try to understand relevant time scales by looking at the leading corrections at both early and late times:

1. To find the leading corrections to (6.23), note that the permutations with the next largest value of $k(\tau)$ after the identity are those with a single transposition, e.g. (12). If it were not for the $d_D^{\zeta(\tau)}$ term, we would need to sum over all such permutations. However, only the permutation that transposes the first and $(m+2)^{th}$ elements has $\zeta(\tau) = 1$. Therefore, this permutation gives the leading correction, and on including it we find

$$F_m = \frac{1}{d_D d_R^{2m+1}} + \frac{1}{d_D d_{B'} d_R^{2m}} + O(d_D^{-2} d_{B'}^{-1}). \quad (6.25)$$

so that we find the leading and next-to-leading order contributions to the fidelity at early times to be

$$F(\rho, [\mathcal{P}_{\sigma, \mathcal{N}} \circ \mathcal{N}](\rho)) = \frac{1}{d_D} + \frac{d_R}{d_D d_{B'}}, \quad (6.26)$$

which grows as more radiation is collected, reflecting the improved recovery. This correction starts to give a contribution comparable to the leading term at time t_p , which indicates that the fidelity starts to grow from its initial value at t_p . Note also that this correction has the same scaling as the correction in the lower bound (6.17), but is twice as large.

When the diary is small, the terms with $\zeta(\tau) = 0$ can also be important. There are $\binom{2m+2}{2} - 1$ leading terms, so that

$$F_m = \frac{1}{d_D d_R^{2m+1}} + \frac{1}{d_D d_{B'} d_R^{2m}} + \frac{m(3+2m)}{d_D^2 d_{B'} d_R^{2m}} + O(d_{B'}^{-2} d_D^{-2}). \quad (6.27)$$

The replica limit $m \rightarrow -\frac{1}{2}$ gives

$$F(\rho, [\mathcal{P}_{\sigma, \mathcal{N}} \circ \mathcal{N}](\rho)) = \frac{1}{d_D} \left(1 + \frac{d_R}{d_{B'}} - \frac{d_R}{d_D d_{B'}} \right), \quad (6.28)$$

2. To understand the time scale at which the late-time value (6.24) is reached, let us now evaluate the leading correction to (6.24). There are now $m(m+1)$ permutations satisfying both $k(\eta^{-1}\tau) = 2m+1$ and $\zeta(\tau) = 1$, leading to

$$F_m = \frac{1}{(d_{B'} d_D)^{2m+1}} + \frac{m(m+1)}{d_R (d_{B'} d_D)^{2m}} + O(d_R^{-1} d_D^{-2m-1}). \quad (6.29)$$

The leading and next-to-leading contributions to the fidelity at late times are therefore

$$F(\rho, [\mathcal{P}_{\sigma, \mathcal{N}} \circ \mathcal{N}](\rho)) = 1 - \frac{d_{B'} d_D}{4 d_R} + O(d_R^{-1} d_D^0). \quad (6.30)$$

Note that the correction had to be negative because the fidelity is bounded above by unity. The correction becomes comparable to the leading term at times earlier than t_{p_2} . Note also that this correction has the same scaling as, but is half the size of, the lower bound correction in (6.17).

For small diaries, there are $(m + 1)^2$ other terms which are also important in the correction to the late-time value, and we find

$$F(\rho, [\mathcal{P}_{\sigma, \mathcal{N}} \circ \mathcal{N}](\rho)) = 1 - \frac{d_{B'} d_D}{4d_R} + \frac{d_{B'}}{4d_R} + O(d_R^{-2} d_D^{-1}). \quad (6.31)$$

Hence, for the infinite temperature case, (6.26) and (6.30) together give the same predictions for the time scales at which the fidelity of \mathcal{P} first increases and the time scale at which it saturates, as the predictions based on the lower bound for some recovery channel in (6.17).

Finally, we can also examine the crossover regime of (6.22), where the fidelity becomes $O(1)$. At the time t_{p_2} , $d_R = d_{B'} d_D := d$, so the sum simplifies to

$$F_m \simeq \frac{1}{d^{4m+4}} \sum_{\tau \in \mathcal{S}_{2m+2}} d^{k(\eta^{-1}\tau) + k(\tau)} d_D^{\zeta(\tau) - 1}. \quad (6.32)$$

The exponent is maximized at $2m+4$ when τ is a non-crossing permutation that has the first and $(m+2)^{th}$ factors in the same cycle. Out of the C_{2m+2} total non-crossing permutations (where C_n is the n -th Catalan number), only C_{m+1}^2 have $\zeta(\tau) = 1$, a statement we will soon prove. Therefore,

$$F_m \simeq \frac{C_{m+1}^2}{d^{2m+1}}. \quad (6.33)$$

Taking the $m \rightarrow -\frac{1}{2}$ limit, we find the fidelity at t_{p_2} to have an $O(1)$ value independent of the dimensions,

$$F(\rho, [\mathcal{P}_{\sigma, \mathcal{N}} \circ \mathcal{N}](\rho)) = \frac{64}{9\pi^2} \simeq 0.72. \quad (6.34)$$

This fidelity is markedly larger than the lower bound from (6.17), which is $e^{-\frac{1}{2}} \simeq 0.61$. Again, for small diaries, we include the subleading non-crossing permutation with $\zeta(\tau) = 0$ to find

$$F(\rho, [\mathcal{P}_{\sigma, \mathcal{N}} \circ \mathcal{N}](\rho)) = \frac{64}{9\pi^2} + \frac{1 - \frac{64}{9\pi^2}}{d_D}. \quad (6.35)$$

We now turn to the full calculation of the fidelity in the planar limit, i.e., we carry out the full sum in (6.22) analytically. Recall that the planar limit corresponds to non-crossing permutations τ where $k(\eta^{-1}\tau) + k(\tau) = 2m + 3$. We claim that the subset of these permutations with $\zeta(\tau) = 1$ and $k(\eta^{-1}\tau) = p$ is given by a product of Narayana numbers

$$\begin{aligned} NC_{2m+2, \zeta=1, p} &= \sum_{p_1 + p_2 = p} N_{m+1, p_1} N_{m+1, p_2} = \sum_{p_1=1}^p N_{m+1, p_1} N_{m+1, p-p_1} \\ &= \frac{\binom{m+1}{p-2} \binom{m+1}{p-1} {}_4F_3(-m-1, -m, 1-p, 2-p; 2, m-p+3, m-p+4; 1)}{m+1}. \end{aligned} \quad (6.36)$$

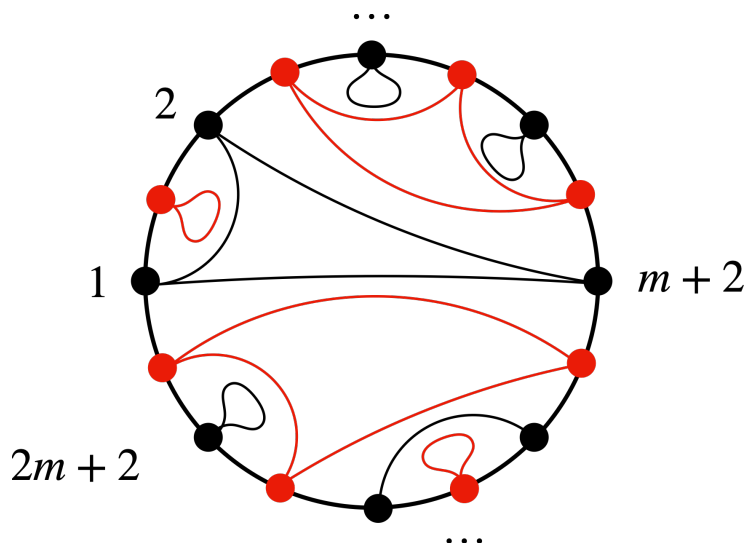


Figure 16. A circular lattice of $2m + 2$ elements is shown as black dots. The dual lattice is the red dots. We show an example of a non-crossing permutation on the red lattice that factorizes across the first and $m + 2$ nd elements of the original lattice. Clearly, the Kreweras complement (shown in black lines) connects the first and $m + 2$ nd elements such that $\zeta(\tau) = 1$.

We note that the previous statement regarding the square of Catalan numbers above (6.33) is the special case where we sum the above equation from $p = 2$ to $p = 2m + 2$.

Eq. (6.36) can be proven by considering the elements of the permutation group as a circular lattice of $2m + 2$ points then moving to the dual lattice (see figure 16). In the dual lattice, we count the number of non-crossing permutations that factorize into non-crossing permutations of $m + 1$ and $m + 1$ elements. There is a unique non-crossing permutation of the original lattice corresponding to each one of these permutations, commonly known as the “Kreweras complement.” This is the maximally extended permutation in the dual graph that does not cross the permutation. It is clear that for each factorized permutation in the dual lattice, the Kreweras complement in the original lattice has $\zeta(\tau) = 1$. Indeed, the number of non-crossing permutations in the dual graph with $k(\tau) = p$ is the same as the number of non-crossing permutations in the original graph with $k(\eta^{-1}\tau) = p$. Using this insight and the prior knowledge of counting the number of non-crossing permutations with a given $k(\tau)$, we arrive at (6.36).

We can now find the sum over all permutations with $\zeta(\tau) = 1$ in (6.22), which we call \tilde{F}_m :

$$\begin{aligned} \tilde{F}_m &= \frac{1}{(d_R d_{B'} d_D)^{2m+2}} \sum_{p=2}^{2m+2} NC_{2m+2, \zeta=1, p} d_R^p (d_D d_{B'})^{2m+3-p} \\ &= \frac{1}{d_{B'} d_D} \begin{cases} \left(d_R^{-m} {}_2F_1 \left(-m, -m-1, 2, \frac{d_R}{d_D d_{B'}} \right) \right)^2, & d_R < d_D d_{B'} \\ \left((d_D d_{B'})^{-m} {}_2F_1 \left(-m, -m-1, 2, \frac{d_D d_{B'}}{d_R} \right) \right)^2, & d_R > d_D d_{B'} \end{cases}. \end{aligned} \quad (6.37)$$

To enumerate the terms with $\zeta(\tau) = 0$ and $k(\eta^{-1}\tau) = p$, we note that this must simply be the remaining non-crossing permutations of which there are $N_{2m+2, p} - NC_{2m+2, \zeta=1, p}$. We

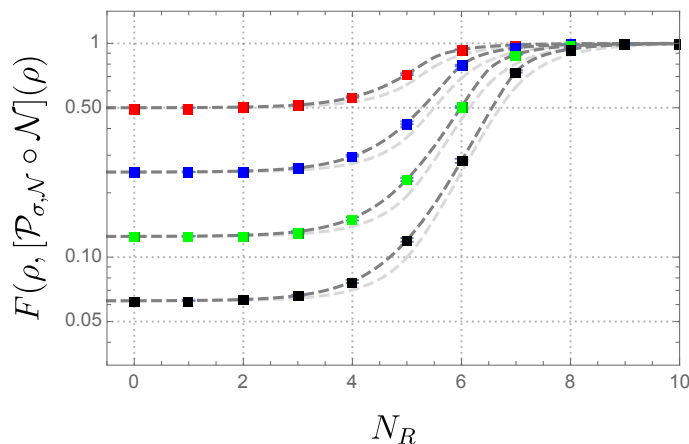


Figure 17. The fidelity of the Petz map for the Hayden-Preskill protocol at infinite temperature, as a function of the number of qubits N_R in the radiation. The total number of qubits in the black hole and the radiation is 10 in all cases, and the different curves correspond to different sizes of the diary, ranging from 1 to 4 qubits from top to bottom. The gray dashed curve is the analytic result given by (6.40). The light gray curves are the lower bounds set by the change in relative entropy. In the numerical data, we take a fiducial pure state tensored with the diary state and apply a random unitary matrix and partial trace, followed by the Petz recovery map. We numerically evaluate the between this recovered state and the initial state. Again, we show the average over 10^3 disorder realizations and the corresponding error bars.

refer to this second contribution to F_m as \bar{F}_m , and immediately find

$$\bar{F}_m = \begin{cases} \frac{1}{d_D d_R^{2m+1}} {}_2F_1\left(-2m-1, -2m-2, 2, \frac{d_R}{d_D d_{B'}}\right) - \frac{\tilde{F}_m}{d_D}, & d_R < d_D d_{B'} \\ \frac{1}{d_D (d_D d_{B'})^{2m+1}} {}_2F_1\left(-2m-1, -2m-2, 2, \frac{d_D d_{B'}}{d_R}\right) - \frac{\tilde{F}_m}{d_D}, & d_R > d_D d_{B'} \end{cases} \quad (6.38)$$

This expression simplifies in the $m \rightarrow -\frac{1}{2}$ limit to

$$\bar{F}_{-\frac{1}{2}} = \frac{1}{d_D} - \frac{\tilde{F}_{-\frac{1}{2}}}{d_D} \quad (6.39)$$

Therefore, the fidelity is

$$F(\rho, [\mathcal{P}_{\sigma_r, N} \circ \mathcal{N}](\rho)) = \begin{cases} \left(1 - \frac{1}{d_D}\right) \frac{d_R}{d_{B'} d_D} {}_2F_1\left(\frac{1}{2}, -\frac{1}{2}, 2, \frac{d_R}{d_D d_{B'}}\right)^2 + \frac{1}{d_D} & d_R < d_D d_{B'} \\ \left(1 - \frac{1}{d_D}\right) {}_2F_1\left(\frac{1}{2}, -\frac{1}{2}, 2, \frac{d_D d_{B'}}{d_R}\right)^2 + \frac{1}{d_D} & d_R > d_D d_{B'} \end{cases} \quad (6.40)$$

This expression is plotted in figure 17, where it matches perfectly with numerical tests for finite system sizes.

6.3.2 Finite temperature

At finite temperature, we have far less control over the contributions from subleading permutations, and hence we are able to find the value of the fidelity corresponding to the sum over all permutations in the equilibrium approximation only in some special cases.

First, note that from the general discussion in section 3.3 based on leading permutations, we expect again that at early times when the identity permutation dominates, the fidelity is given by

$$F(\rho, [\mathcal{P}_{\sigma, \mathcal{N}} \circ \mathcal{N}](\rho)) = F(\rho_D, \sigma_D). \quad (6.41)$$

where in the last expression we take σ_D to be an equilibrium density matrix on D , and ρ_D to be some pure state. Similarly, at sufficiently late times, the cyclic permutation should dominate, so that we have

$$F(\rho, [\mathcal{P}_{\sigma, \mathcal{N}} \circ \mathcal{N}](\rho)) = 1. \quad (6.42)$$

At intermediate times, the fidelity will interpolate between these values. To probe this regime and find the relevant time scales, let us first look at the leading corrections to the dominant permutations. The first correction to early-time contribution from the identity permutation comes from the permutation that swaps the first and $(m+2)^{th}$ copies,

$$F_m = \frac{Z_{2m+2,R}}{Z_{1,R}^{2m+2}} F(\rho_D, \sigma_D) + \frac{Z_{2,B'} Z_{m+1,R}^2}{Z_{1,B'}^2 Z_{1,R}^{2m+2}} \text{Tr} \left[\sigma_D^{\frac{1}{2}} \rho_D \sigma_D^{\frac{1}{2}} \rho_D \right]. \quad (6.43)$$

In the replica limit, this can be rewritten in terms of Renyi entropies as

$$F(\rho, [\mathcal{P}_{\sigma, \mathcal{N}} \circ \mathcal{N}](\rho)) = F(\rho_D, \sigma_D) + e^{\frac{S_{\frac{1}{2},R}^{(eq)}}{2} - S_{2,B'}^{(eq)}} \text{Tr} \left[\sigma_D^{\frac{1}{2}} \rho_D \sigma_D^{\frac{1}{2}} \rho_D \right]. \quad (6.44)$$

We can assume $\text{Tr} \left[\sigma_D^{\frac{1}{2}} \rho_D \sigma_D^{\frac{1}{2}} \rho_D \right]$ is of roughly the same magnitude as $F(\rho_D, \sigma_D)$ (note that if we take σ_D to be that maximally mixed state and ρ_D to be pure, then both are equal to $\frac{1}{d_D}$). This suggests that the fidelity begins to grow significantly from its initial value at the time scale t_b , at which $S_{\frac{1}{2},R}^{(eq)} = S_{2,B'}^{(eq)}$.

However, in order to reliably calculate the fidelity past t_b , we must sum over the contributions from all permutations corresponding to planar diagrams, a calculation we explain in appendix G.1 using a generating functional method. For this calculation, we assume that σ_D is maximally mixed, ρ_D is pure, the radiation is at infinite temperature, and the black hole is at inverse temperature β , with the density of states as in AdS_3 , $\rho(E) = e^{cV\sqrt{E/V}}$. Now if we define

$$x \equiv \frac{\log d_R}{V}, \quad y \equiv \frac{\log d_D}{V}, \quad (6.45)$$

where V is the volume of B' and d_R is the dimension of the radiation, then the various relevant time scales are given by:

$$t_b : \quad x = \frac{3c^2}{8\beta}, \quad t_p : \quad x = \frac{c^2}{2\beta}, \quad t_{p_2} : \quad x = \frac{c^2}{2\beta} + y. \quad (6.46)$$

We find that for small diaries (i.e. y is $O(V^{-1})$), the fidelity grows rapidly at t_p and the time scale t_b does not turn out to be relevant,

$$F(\rho, [\mathcal{P}_{\sigma, \mathcal{N}} \circ \mathcal{N}](\rho)) \approx \begin{cases} \frac{1}{d_D} & x < \frac{c^2}{2\beta} \\ 1 & x > \frac{c^2}{2\beta} \end{cases} \quad (6.47)$$

with the transition occurring in a region of size $O((V\beta)^{-\frac{1}{2}})$ in x . On the other hand, for a sufficiently large diary such that there is a regime where $\frac{3c^2}{8\beta} < x < \frac{5c^2}{16\beta} + y$, we have

$$F(\rho, [\mathcal{P}_{\sigma, \mathcal{N}} \circ \mathcal{N}](\rho)) = \begin{cases} \frac{1}{d_D} & x < \frac{3c^2}{8\beta} \\ e^{V(-\frac{3e^2}{8\beta} + x - y)} = e^{\log d_R - S_{2, B'} - \log d_D} & \frac{3c^2}{8\beta} < x < \frac{5c^2}{16\beta} + y \\ e^{V(c\sqrt{-\frac{c^2}{4\beta^2} + \frac{x-y}{\beta}} - x + y)} & \frac{5c^2}{16\beta} + y < x < \frac{c^2}{2\beta} + y \\ 1 & x > \frac{c^2}{2\beta} + y \end{cases} \quad (6.48)$$

So for a sufficiently large diary, the fidelity starts increasing exponentially from its initial value of $\frac{1}{d_D}$ at time t_b , and the initial increase is precisely as predicted by (6.44). Note that the lower bound on the fidelity from the change in relative entropy from (6.19) is not sensitive to the time scale t_b , and only starts growing at the Page time t_p .

When y is $O(1)$ but the diary is not sufficiently large such that there is a regime where $\frac{3c^2}{8\beta} < x < \frac{5c^2}{16\beta} + y$, the fidelity starts increasing from $\frac{1}{d_D}$ at a time t_r between t_b and t_p defined by

$$t_r : \quad x = \frac{c^2}{2\beta} + 2y - c\sqrt{\frac{y}{\beta}} \quad (6.49)$$

so that we have

$$F(\rho, [\mathcal{P}_{\sigma, \mathcal{N}} \circ \mathcal{N}](\rho)) = \begin{cases} \frac{1}{d_D} & x < \frac{c^2}{2\beta} + 2y - c\sqrt{\frac{y}{\beta}} \\ e^{V(c\sqrt{-\frac{c^2}{4\beta^2} + \frac{x-y}{\beta}} - x + y)} & \frac{c^2}{2\beta} + 2y - c\sqrt{\frac{y}{\beta}} < x < \frac{c^2}{2\beta} + y \\ 1 & x > \frac{c^2}{2\beta} + y \end{cases} \quad (6.50)$$

t_r is increasingly earlier for larger $\log d_D$, and eventually becomes equal to t_b .

Eqs. (6.47), (6.48) and (6.50) are derived with several approximations for the thermodynamic limit. We plot a more exact expression resulting from the calculations of appendix G.1 for finite but large volumes in figure 18. Note in particular that for the case where d_D is large, we see the initial exponential growth of fidelity $e^{\frac{S_{\frac{1}{2}, R}^{(\text{eq})} - S_{2, B'}^{(\text{eq})} - \log d_D}{2}}$ predicted by (6.44). For a small diary, the curves obtained from increasing system sizes with fixed $\log d_D$ gradually approach an increasingly sharp transition, as expected from (6.47).

The discussion in this section so far has entirely been for the canonical ensemble case. In appendix G.2, we also consider the example of \mathcal{I}_α as in the microcanonical example of (4.87), corresponding to a case where the radiation is divided into two parts R_1 and R_2 such that there is energy conservation between R_1 and B' while R_2 is at infinite temperature. For a small diary, we again find that the fidelity grows from $\frac{1}{d_D}$ to 1 rapidly at t_p . For a large diary, we find that the fidelity reaches 1 at t_{p2} , and that there is a time scale $t_r < t_p$ when the fidelity starts to grow above $\frac{1}{d_D}$. But for this example, t_r does not seem to be related to the time t_b at which $\mathcal{E}(R_1, R_2)$ starts to grow in any regime.

7 Conclusions and discussions

In this paper, we generalized the equilibrium approximation introduced in [1] to various quantum-informational measures beyond the entanglement entropy, such as Renyi and

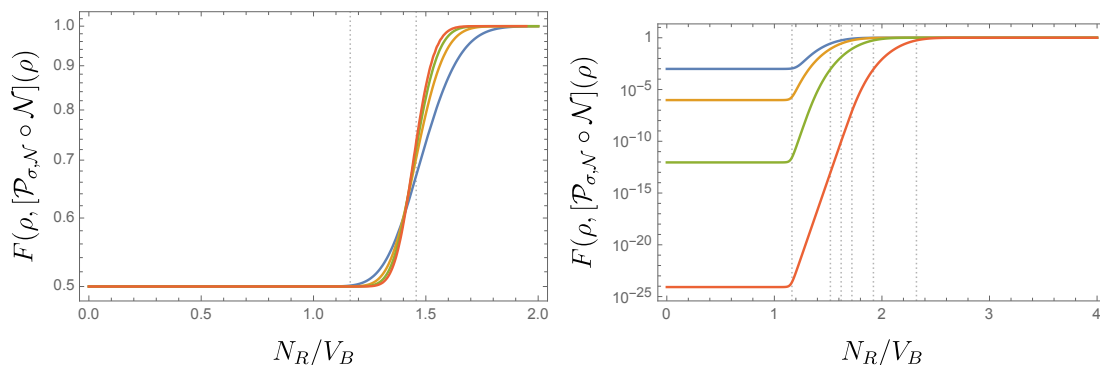


Figure 18. Left: the fidelity of the Petz recovery map is shown for a diary consisting of a single qubit. The different plots are for different $V = 100, 200, 300,$ and 400 (from blue to red). The transition approaches the second vertical line, which corresponds to t_p in the thermodynamic limit. The first vertical line is t_b . Right: the fidelity of the Petz recovery map for large diaries of sizes 10, 20, 40, and 80 qubits (top to bottom). The first two vertical lines denote t_b and t_p which are separated even in the thermodynamic limit. We also display the t_{p_2} 's for the respective diary sizes, always with $V = 100$. We take R to be at infinite temperature and B' to be at finite temperature with Cardy-like density of states $\rho(E) = e^{V\sqrt{\frac{E}{V}}}$ in both plots, taking $\beta = 1/2$. Notably, the fidelity increases significantly between t_b and t_p and does not saturate to one until after t_{p_2} .

logarithmic negativities, relative entropy, Petz map fidelity, and reflected entropy. We then studied in detail the entanglement correlations within equilibrated pure states at a *finite* temperature. We found rich entanglement structures exhibited by the phase diagrams of the logarithmic negativity, mutual information, Renyi negativities, and Renyi mutual informations. In addition to their implications for general quantum chaotic systems, our results also give predictions for the entanglement structure of an evaporating black hole. The verification of these predictions from gravity calculations would help in further probing the quantum nature of black holes as well as informational aspects of quantum gravity. In particular, we predicted the existence of a new time scale t_b before the Page time, when entanglement correlations within the Hawking radiation start to emerge.

We also studied in detail the process of information recovery from a subsystem, using the Petz map. This issue is central to understanding how and when quantum information is transferred from a black hole to its radiation. We showed that even at finite temperature, the fidelity of the recovery map as well as an upper bound on it provided by relative entropy can be expressed in terms of natural quantum-informational measures in thermal density matrices. Intriguingly, we found evidence that t_b is the relevant time scale at which the fidelity of recovery starts to increase for a large diary, which suggests that correlations that are detected by negativity but not detected by mutual information play an important role in information recovery.

There are many future directions to explore. Our formulations of the equilibrium approximation for relative entropy and reflected entropy should find other applications. For example, it would be interesting to explore in detail the universal behavior of reflected entropy for finite temperature systems, which we only briefly touched upon. Such studies

should also provide further understanding of multipartite entanglement in a quantum many-body system, especially among different parts of the radiation and the black hole.

It would be highly desirable to understand better the nature of the entanglement in the region of the phase diagram where $\mathcal{E} \gg I$. The question is of particular interest for black hole physics. In this regard, it would be interesting to understand whether it is possible to obtain a Lorentzian derivation of the non-zero negativity before the Page time, and see whether there is any semi-classical or geometric description of bound entanglement. Another interesting question is about how new phases which might not correspond to analytic continuation, similar to the ones we found in figure 11, would be manifested in a gravity calculation.

From a purely quantum information-theoretic perspective, it would be interesting to investigate if there are deeper direct connections between logarithmic negativity and recoverability of quantum information. In our calculations for the Hayden-Preskill protocol, the known lower bound on the fidelity of recovery from the relative entropy turned out not to be sensitive to the t_b time scale at which both the Petz map fidelity and the logarithmic negativity start to grow. Our results thus hint that there could be strictly stronger lower bounds on fidelity of recovery if one considered notions related to the partial transpose.

Acknowledgments

We thank Chris Akers, Tom Faulkner, Simon Lin, and Pratik Rath for sharing their manuscripts with us. This work is supported by the Office of High Energy Physics of U.S. Department of Energy under grant Contract Numbers DE-SC0012567 and DE-SC0020360 (MIT contract # 578218).

A Properties of the permutation group

We can define a notion of distance between elements σ, τ of the permutation group \mathcal{S}_n with the Cayley distance,

$$d(\sigma, \tau) = n - k(\sigma\tau^{-1}). \tag{A.1}$$

This definition of the distance is consistent with the triangle inequality,

$$d(\sigma, \tau) + d(\tau, \rho) \geq d(\sigma, \rho). \tag{A.2}$$

Eqs. (A.1) and (A.2) can be used to show that

$$k(\eta\tau^{-1}) + k(\tau) \leq n + 1 \tag{A.3}$$

$$k(\eta^{-1}\tau^{-1}) + k(\tau) \leq n + 1 \tag{A.4}$$

$$k(\eta^{-1}\tau) + k(\eta\tau) \leq \begin{cases} n + 2 & n \text{ even} \\ n + 1 & n \text{ odd} \end{cases} \tag{A.5}$$

where we use the fact $k(\eta) = 1$ for all n , $k(\eta^2) = 2$ for even n , and $k(\eta^2) = 1$ for odd n . We then also have

$$\begin{aligned}
 & k(\eta^{-1}\tau) + k(\eta\tau) + k(\tau) \\
 &= \frac{k(\eta\tau^{-1}) + k(\tau)}{2} + \frac{k(\eta^{-1}\tau^{-1}) + k(\tau)}{2} + \frac{k(\eta\tau^{-1}) + k(\eta\tau)}{2} \leq \begin{cases} \frac{3n}{2} + 2 & n \text{ even} \\ \frac{3n+3}{2} & n \text{ odd} \end{cases} \quad (\text{A.6})
 \end{aligned}$$

where the upper bound follows from the inequalities for the individual terms in the second line, and we have used (for instance) that $k(\eta\tau) = k(\tau^{-1}\eta^{-1}) = k(\tau^{-1}(\eta^{-1}\tau^{-1})\tau) = k(\eta^{-1}\tau^{-1})$. We will now find the set of permutations that saturates (A.6), which is the same as the set that saturates the inequalities for each of the three terms on the l.h.s. of the second line.

First note that the permutations that saturate the inequality for both the first and the second term are in one-to-one correspondence with the non-crossing partitions of n elements. Given a non-crossing partition $\{a_1^1, \dots, a_{m_1}^1\}, \{a_1^2, \dots, a_{m_2}^2\}, \dots, \{a_1^k, \dots, a_{m_k}^k\}$, where the elements of each set are listed in ascending order, the corresponding permutations are given by

$$\tau_1 = (a_1^1 a_2^1 \dots a_{m_1}^1)(a_1^2 \dots a_{m_2}^2) \dots (a_1^k \dots a_{m_k}^k) : \quad k(\eta^{-1}\tau_1^{-1}) + k(\tau_1) = n + 1 \quad (\text{A.7})$$

$$\tau_2 = (a_{m_1}^1 \dots a_2^1 a_1^1) (a_{m_2}^2 \dots a_1^2) \dots (a_{m_k}^k \dots a_1^k) : \quad k(\eta\tau_2^{-1}) + k(\tau_2) = n + 1 \quad (\text{A.8})$$

The permutations that maximize the sum of the first two terms in the second line of (A.6) are those that lie in the intersection of the above values of τ_1 and τ_2 . The difference in the ordering of the elements within the cycles in τ_1 and τ_2 does not change the permutation if and only if the permutation consists entirely of one-cycles and two-cycles. Hence, the permutations that simultaneously saturate the inequalities for the first and second terms are in one-to-one correspondence with the non-crossing partitions of n elements that have blocks only of size 1 and 2.

Next, note that for any permutation τ that saturates the inequalities for both the first and the second terms, we have

$$k(\eta^{-1}\tau^{-1}) + k(\eta\tau^{-1}) = 2n + 2 - 2k(\tau) \quad (\text{A.9})$$

To maximize the r.h.s., we need to minimize $k(\tau)$ among permutations that only have one-cycles and two-cycles.

For even n , $k(\tau)$ is minimized for permutations that only have two cycles, where it is given by $n/2$. Hence, the permutations saturating (A.6) for even n are the ones corresponding to non-crossing partitions of n elements that have blocks only of size 2. We refer to this set of permutations as $\{\tau\}^*$. For any such permutation, $k(\eta\tau^{-1}) = k(\eta\tau) = \frac{n}{2} + 1$. The number of such permutations is equal to $C_{n/2}$, where $C_{n/2}$ is the Catalan number.

Similarly, for odd n , the permutations saturating (A.6) are the ones corresponding to non-crossing partitions of n elements that have $\frac{n-1}{2}$ blocks of size 2 and 1 block of size 1. We refer to this set of permutations as $\{\tau\}_{\text{odd}}^*$. For any such permutation, $k(\tau) = k(\eta\tau^{-1}) = k(\eta\tau) = \frac{n+1}{2}$. The number of such permutations is equal to $n C_{\frac{n-1}{2}}$.

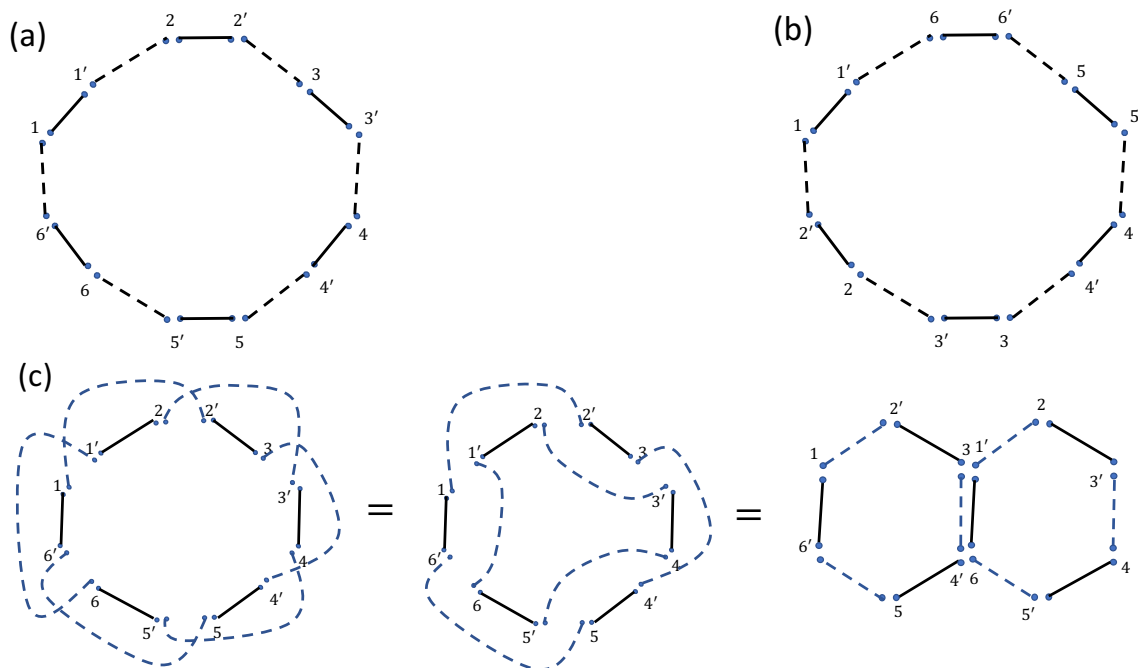


Figure 19. Boundary conditions for the diagrams corresponding to the three terms in the second line of (A.6).

An alternative way to understand the inequality (A.6) and the permutations that saturate it is in terms of 'tHooft-like double-line diagrams. The first line of (A.6) is equal to the total number of loops in diagrams of the kind shown in figure 3, which involve sets of three lines, so we cannot directly modify 'tHooft's method to that expression. It is useful to separately consider the three terms in the second line of (A.6), each of which can be equated with the total number of loops in a diagram with a different boundary condition. The boundary conditions for the first two terms are as shown in figure 19(a) and (b), and the ones for the third term are explained in 19(c), where it will be convenient to use the third form to apply 'tHooft's method.

By adding an extra loop around each of the diagrams, they can be seen as double-line diagrams, which can then be mapped to polygons by replacing the double lines with single lines. An example of the mapping to polygons for a particular τ that saturates (A.6), with the boundary conditions of figures 19(a) and (c), is shown in figure 20. The total number of loops in the double-line diagram is equal to the total number of faces of the corresponding polygon. If the polygon can be flattened on a surface of minimum genus h , then

$$F = E - V + 2 - 2h. \tag{A.10}$$

Now for any given τ , in the polygons corresponding to the boundary conditions in figure 19 (a) and (b), we always have $E = 3n$ and $V = 2n$, so if for instance $F_{(a)}$ is the number of faces in the diagram for τ with the boundary conditions of figure 19(a) and $h_{(a)}$ is the

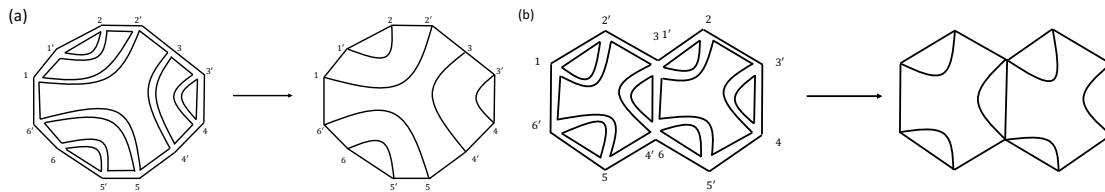


Figure 20. For $n = 6$ and $\tau = (12)(34)(56)$, we show the planar double-line diagrams and the corresponding polygons that we get for the first and third terms in the second line of (A.6), using the boundary conditions from (19) (a) and (c).

corresponding genus, we have

$$\begin{aligned} k(\tau) + k(\tau^{-1}\eta) &= F_{(a)} - 1 = n + 1 - 2h_{(a)} \leq n + 1, \\ k(\tau) + k(\tau^{-1}\eta^{-1}) &= F_{(b)} - 1 = n + 1 - 2h_{(b)} \leq n + 1 \end{aligned} \quad (\text{A.11})$$

where we subtract 1 from the total number of faces to remove the contribution from the extra line we added around the diagrams. In the polygons corresponding to the boundary conditions in the last figure in 19 (c), we have $E = 3n - 1$ and $V = 2n - 2$, so

$$k(\eta\tau^{-1}) + k(\tau\eta) = F_{(c)} - 1 = n + 2 - 2h_{(c)} \leq n + 2 \quad (\text{A.12})$$

Hence, the set of permutations that saturate (A.6) are those that give planar diagrams for each of the figures 19 (a), (b) and (c).

B Estimate of fluctuations in the equilibrium approximation for the negativity

In this appendix, we show that the fluctuations in $\mathcal{Z}_n^{(\text{PT})}$ around the equilibrium value, as measured self-consistently by the quantity Δ_N defined in (3.16), are suppressed relative to the equilibrium value. For this, we need to find the equilibrium approximation for the quantity $(\mathcal{Z}_n^{(\text{PT})})^2$, and compare it to $([\mathcal{Z}_n^{(\text{PT})}]_{\text{eq approx}})^2$. Let d_X be the effective Hilbert space dimension of subsystem X , and $\alpha_2 := \eta \otimes \eta$. Then

$$\begin{aligned} ((\mathcal{Z}_n^{(\text{PT})})^2)_{\text{eq approx}} &= \sum_{\tau_2 \in \mathcal{S}_{2n}} \langle e_B \otimes \eta_{A_1} \otimes \eta_{A_2}^{-1} | \langle e_B \otimes \eta_{A_1} \otimes \eta_{A_2}^{-1} | \mathcal{I}_\alpha, \tau_2 \rangle \\ &\sim \sum_{\tau_2 \in \mathcal{S}_{2n}} d_B^{k(\tau_2)} d_{A_1}^{k(\tau_2 \alpha_2^{-1})} d_{A_2}^{k(\tau_2 \alpha_2)} \end{aligned} \quad (\text{B.1})$$

Note that

$$([\mathcal{Z}_n^{(\text{PT})}]_{\text{eq approx}})^2 = \sum_{\tau, \sigma \in \mathcal{S}_n} \langle e_B \otimes \eta_{A_1} \otimes \eta_{A_2}^{-1} | \langle e_B \otimes \eta_{A_1} \otimes \eta_{A_2}^{-1} | \mathcal{I}_\alpha, \tau \otimes \sigma \rangle \quad (\text{B.2})$$

So

$$\begin{aligned} \Delta_N^2 &= \sum_{\tau_2 \in \mathcal{S}_{2n}, \tau_2 \neq \tau \otimes \sigma} \langle e_B \otimes \eta_{A_1} \otimes \eta_{A_2}^{-1} | \langle e_B \otimes \eta_{A_1} \otimes \eta_{A_2}^{-1} | \mathcal{I}_\alpha, \tau_2 \rangle \\ &\sim \sum_{\tau_2 \in \mathcal{S}_{2n}, \tau_2 \neq \tau \otimes \sigma} d_B^{k(\tau_2)} d_{A_1}^{k(\tau_2 \alpha_2^{-1})} d_{A_2}^{k(\tau_2 \alpha_2)} \end{aligned} \quad (\text{B.3})$$

Let us first consider the entanglement saturation phase, where $d_{A_1} \sim d_{A_2} \sim d_B \sim d$. Then

$$[(\mathcal{Z}_n^{(\text{PT})})^2]_{\text{eq approx}} \sim \sum_{\tau_2 \in \mathcal{S}_{2n}} d^{k(\tau_2) + k(\tau_2 \alpha_2^{-1}) + k(\tau_2 \alpha_2)} \quad (\text{B.4})$$

Using the Cayley distance and the triangle inequality for three permutations $\sigma_2, \tau_2, \rho_2 \in \mathcal{S}_{2n}$, we have

$$k(\sigma_2 \tau_2^{-1}) + k(\tau_2 \rho_2^{-1}) \leq 2n + k(\sigma_2 \rho_2^{-1}) \quad (\text{B.5})$$

which gives the following upper bound for the exponent in (B.4):

$$\begin{aligned} & k(\tau_2) + k(\tau_2 \alpha_2^{-1}) + k(\tau_2 \alpha_2) \quad (\text{B.6}) \\ &= \frac{k(\tau_2^{-1}) + k(\tau_2 \alpha_2^{-1})}{2} + \frac{k(\alpha_2 \tau_2^{-1}) + k(\tau_2 \alpha_2)}{2} + \frac{k(\tau_2^{-1}) + k(\tau_2 \alpha_2)}{2} \leq \begin{cases} 3n + 4 & n \text{ even} \\ 3n + 3 & n \text{ odd} \end{cases} \end{aligned}$$

We can then check that if $\tau, \sigma \in \mathcal{S}_n$ both saturate (A.6), then $\tau \otimes \sigma \in \mathcal{S}_{2n}$ saturates (B.6). We will now show using 'tHooft-like line diagrams that for any $\tau_2 \neq \tau \otimes \sigma$, (B.6) is not saturated. The three terms on the second line of (B.6) can be equated with the total number of loops respectively in the diagrams with boundary conditions shown in figure 21 (a), (b) and (c). The leading contribution comes from the permutations that give planar diagrams for each of the three boundary conditions. In particular, any non-factorized permutation gives a non-planar diagram for each of (a), (b) and (c), corresponding to genres that can be labelled $h_1, h_2, h_3 \geq 1$. With manipulations of the diagrams to the ones used in appendix A, we have

$$\begin{aligned} k(\tau_2^{-1}) + k(\tau_2 \alpha_2^{-1}) &= 2n + 2 - 2h_1, \\ k(\tau_2^{-1}) + k(\tau_2 \alpha_2) &= 2n + 2 - 2h_2, \\ k(\alpha_2 \tau_2^{-1}) + k(\tau_2 \alpha_2) &= 2n + 4 - 2h_3. \end{aligned} \quad (\text{B.7})$$

Hence, all terms contributing to Δ_N^2 are of the form

$$d^{(2n+2-2h_1)/2 + (2n+2-2h_2)/2 + (2n+4-2h_3)/2} \leq d^{3n/2+2-3} \quad (\text{B.8})$$

We therefore find that in the entanglement saturation phase, the fluctuations are suppressed as

$$\frac{\Delta_N}{[\mathcal{Z}_n^{(\text{PT})}]_{\text{eq approx}}} \sim d^{-3/2} \sim Z_1^{-1/2} \quad (\text{B.9})$$

where Z_1 is the effect dimension of the full Hilbert space.

Next, consider the maximally entangled phase. In this case WLOG, $d_{A_1} \gg d_{A_2}, d_B$, so the dominant contribution to $[\mathcal{Z}_n^{(\text{PT})}]_{\text{eq approx}}$ comes from $\tau = \eta$, and the dominant contribution to $[(\mathcal{Z}_n^{(\text{PT})})^2]_{\text{eq approx}}$ comes from $\tau_2 = \alpha = \eta \otimes \eta$. From any $\tau_2 \neq \eta \otimes \eta$, and in particular any non-factorized τ_2 , the contribution is suppressed relative to the leading contribution by a factor of at least $d_{A_1}^{-1} \sim D^{-1}$, so

$$\frac{\Delta_N}{[\mathcal{Z}_n^{(\text{PT})}]_{\text{eq approx}}} \sim D^{-1/2} \quad (\text{B.10})$$

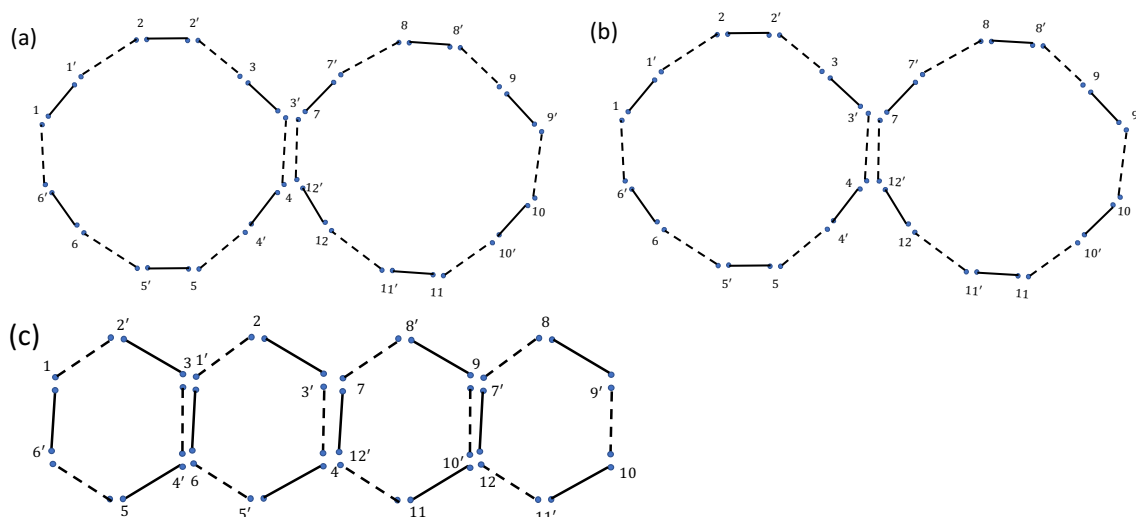


Figure 21. Boundary conditions for the diagrams corresponding to the three terms in the second line of (B.6) are shown in order in (a), (b) and (c).

Finally, consider the PPT phase. In this case $d_B \gg d_{A_1}, d_B$, so the dominant contribution to $[\mathcal{Z}_n^{(\text{PT})}]_{\text{eq approx}}$ comes from $\tau = e$, and the dominant contribution to $[(\mathcal{Z}_n^{(\text{PT})})^2]_{\text{eq approx}}$ comes from $\tau_2 = \alpha = e \otimes e$. Again we have

$$\frac{\Delta_N}{[\mathcal{Z}_n^{(\text{PT})}]_{\text{eq approx}}} \sim D^{-1/2}. \quad (\text{B.11})$$

C Resolvent calculations of logarithmic negativity in various cases

C.1 General diagrammatic approach

In order to find the logarithmic negativity, it is useful to first find the equilibrium approximation for the resolvent

$$R_N(\lambda) = \frac{1}{\lambda} \sum_{n=0}^{\infty} \frac{1}{\lambda^n} \text{Tr}[(\rho_A^{T_2})^n]. \quad (\text{C.1})$$

It is useful to see R_N as the trace of a matrix

$$R_{pq} = \frac{1}{\lambda} \sum_{n=0}^{\infty} \frac{1}{\lambda^n} (\rho_A^{T_2})^n_{pq}, \quad |p\rangle = |p_1\rangle_{A_1} |p_2\rangle_{A_2}, \quad |q\rangle = |q_1\rangle_{A_1} |q_2\rangle_{A_2} \quad (\text{C.2})$$

We can apply the equilibrium approximation to each $(\rho_A^{T_2})^n_{pq}$. The common lower half of the diagrams for all permutations in this case can be deduced from figure 3 (a) for $\text{Tr}[(\rho_A^{T_2})^n]$ by erasing the dashed line connecting i'_{a_n} and i_{a_1} , and the dotted line connecting $i'_{\bar{a}_1}$ and $i_{\bar{a}_n}$, and instead taking the inner product of $|i_{a_1}\rangle, |i'_{\bar{a}_1}\rangle, |i_{a_n}\rangle, |i'_{\bar{a}_n}\rangle$ with $|p_1\rangle, |p_2\rangle, |q_1\rangle, |q_2\rangle$ respectively. The resulting lines are shown in figure 22, which also explains how the factors of $\frac{1}{\lambda^{n+1}}$ and $\frac{1}{Z_1^n}$ that are common to all terms with index n in (C.1) (the second factor comes from the equilibrium approximation) are incorporated into these lines. In the

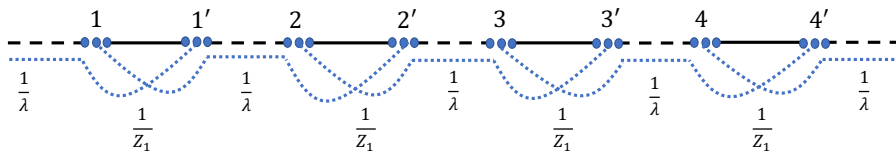


Figure 22. “Boundary” lines for the equilibrium approximation for R_{pq} with $n = 4$.

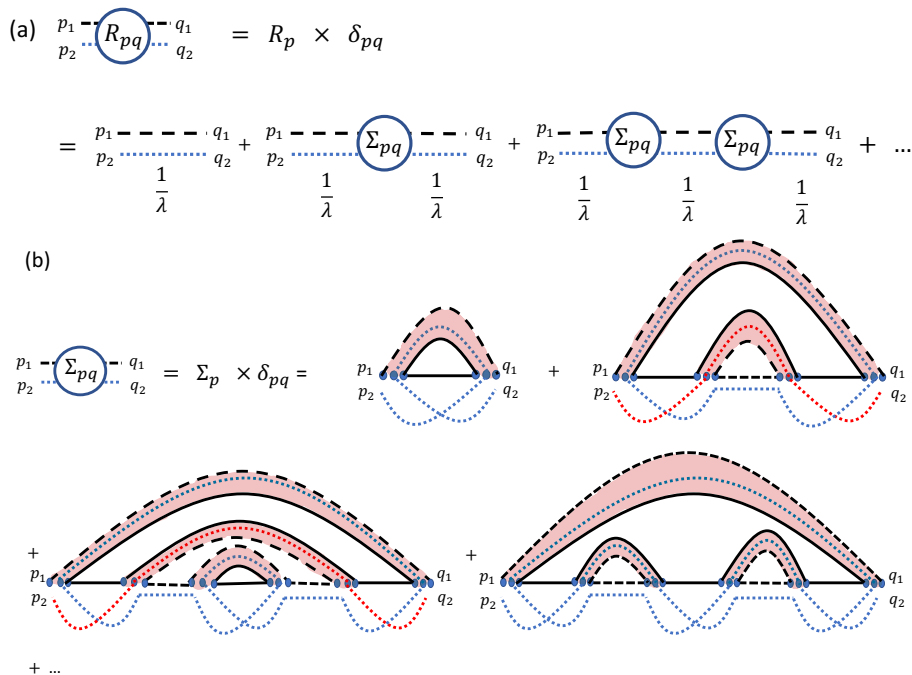


Figure 23. Diagrammatic representation of R_{pq} and the first few diagrams contributing to Σ_{pq} in the case where A_1 is larger than A_2 .

limit where the effective dimension of A_1 is much larger than that of A_2 , it is sufficient to consider contributions from planar diagrams for all n to R_{pq} . We can write R_{pq} in terms of a self-energy Σ_{pq} as shown in figure 23(a). Σ_{pq} is a sum of diagrams without any disconnected parts connected by $\frac{1}{\lambda}\delta_{pq}$, to which the first few contributions are shown in figure 23(b). We take $|p_1\rangle, |q_1\rangle$ and $|p_2\rangle, |q_2\rangle$ to be elements of the energy eigenbasis in A_1 and A_2 respectively, so that we approximately have that

$$\langle p_1 | \langle p_2 | \mathcal{I}_\alpha | q_1 \rangle | q_2 \rangle \propto \delta_{p_1 q_1} \delta_{p_2 q_2} \quad (\text{C.3})$$

for both the canonical and microcanonical ensembles, and hence from the diagrams contributing to Σ_{pq} we can see that both Σ_{pq} and R_{pq} are proportional to δ_{pq} .

We can immediately see by summing the geometric series on the r.h.s. of figure 23(a) that

$$R_N = \sum_p R_p, \quad R_p = \frac{1}{\lambda - \Sigma_p}. \quad (\text{C.4})$$

We can systematically include all planar contributions to Σ_p using the Schwinger-Dyson equation shown diagrammatically in figure 24. For general choices of \mathcal{I}_α , this Schwinger-

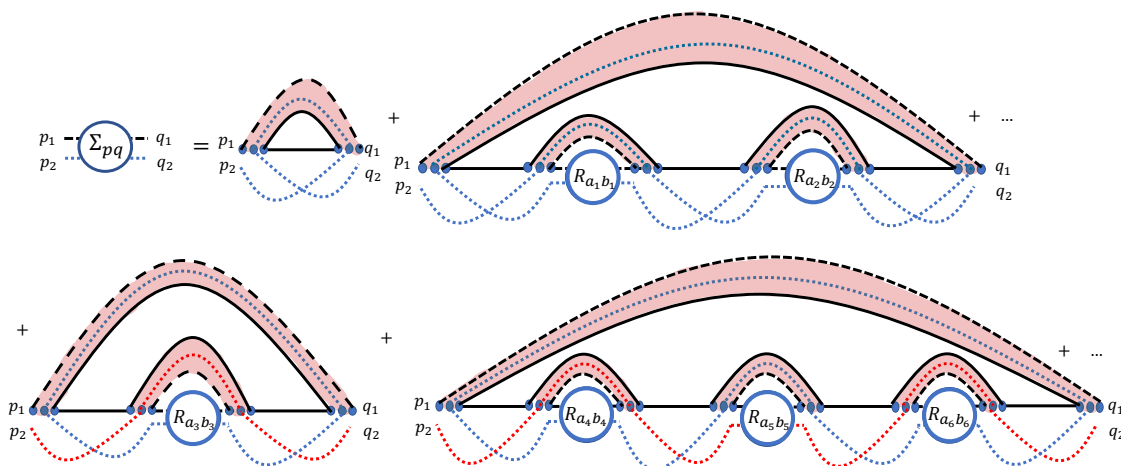


Figure 24. Rewriting of the r.h.s. of figure 23(b) as a Schwinger-Dyson equation. All a_i, b_i are independently summed over (unless they have delta functions among each other or with p, q according to the diagram).

Dyson equation in general leads to a complicated set of equations relating R_p for all different p to each other. Below we consider a few choices of \mathcal{I}_α for which the Schwinger-Dyson equation simplifies.

C.2 Infinite temperature

We first consider the case where the effective identity operator is simply the identity,

$$\mathcal{I}_\alpha = \mathbf{1}_{A_1} \otimes \mathbf{1}_{A_2} \otimes \mathbf{1}_B. \tag{C.5}$$

Here the Schwinger-Dyson equation for R_p becomes independent of the index p in A , and we have

$$R_N^\infty = d_A R_1, \quad R_1 = \frac{1}{\lambda - \Sigma_1}. \tag{C.6}$$

Each line on the r.h.s. of figure 24 now simplifies to a geometric series, and hence

$$\begin{aligned} \Sigma_1 &= \left(\frac{1}{d_A} + \alpha R_1 \right) \left(1 + R_1^2 \beta + (R_1^2 \beta)^2 + \dots \right) = \left(\frac{1}{d_A} + \alpha R_1 \right) \frac{1}{1 - R_1^2 \beta}, \\ \alpha &= \frac{1}{d_A d_B}, \quad \beta = \frac{1}{(d_{A_2} d_B)^2} \end{aligned} \tag{C.7}$$

Substituting this into (C.6), we get a cubic equation for R_1 ,

$$\beta \lambda R_1^3 + (\alpha - \beta) R_1^2 + \left(\frac{1}{d_A} - \lambda \right) R_1 + 1 = 0. \tag{C.8}$$

It is difficult to obtain a general analytic expression for $\mathcal{Z}^{(\text{PT})} := \sum_i |\lambda_i| = e^\mathcal{E}$ by finding the analytic form of $D_N(\lambda)$ from the solutions to (C.8). However, the density and the logarithmic negativity can be found numerically from (C.8), and turn out to agree with

the analytic continuation in (4.17), (4.14), (4.21), as discussed in [9]. That is, we have

$$\mathcal{Z}_\infty^{(\text{PT})}(d_{A_1}, d_{A_2}, d_B) = \begin{cases} 1 & \log d_A < \log d_B \\ d_{A_1} & \log(d_{A_1} d_B) < \log d_{A_2} \\ d_{A_2} & \log(d_{A_2} d_B) < \log d_{A_1} \\ \left(\frac{d_A}{d_B}\right)^{1/2} & \log d_A > \log d_B, \quad |\log d_{A_1} - \log d_{A_2}| < \log d_B \end{cases} \quad (\text{C.9})$$

Now consider the regime where $\frac{d_{A_1}}{d_{A_2} d_B} \ll 1$. This corresponds to being outside the ME phase. Then since $\alpha \gg \beta$ in this regime, R_1 is $O(1)$, and λ is $O(1/d_A)$, (C.8) simplifies to a quadratic equation for R_1 ,

$$\alpha R_1^2 + \left(\frac{1}{d_A} - \lambda\right) R_1 + 1 = 0. \quad (\text{C.10})$$

The same quadratic equation can be obtained diagrammatically by ignoring all contributions to the Schwinger-Dyson equation in figure 24 except the first term in each line on the r.h.s. This corresponds to including contributions from permutations with only one-cycles and two-cycles to $\mathcal{Z}_n^{(\text{PT})}$ for all n . We can now solve (C.10) to get a simple semicircle form of $D_N(\lambda)$ from $R_N(\lambda)$,

$$D_\infty^{\text{sc}}(\lambda, d_A, d_B) = \frac{d_A^2 d_B}{2\pi} \sqrt{\frac{4}{d_A d_B} - \left(\lambda - \frac{1}{d_A}\right)^2}, \quad (\text{C.11})$$

which can be integrated analytically to get $\mathcal{Z}^{(\text{PT})}$,

$$\mathcal{Z}_\infty^{(\text{PT}), \text{sc}}(d_A, d_B) = \begin{cases} 1 & \log d_A < \log d_B \\ \left(\frac{d_A}{d_B}\right)^{1/2} & \log d_A > \log d_B \end{cases}. \quad (\text{C.12})$$

As expected, this approximation misses the ME phase of (C.9).

C.3 Canonical ensemble with A at infinite temperature

Next, let us consider the case where A is at infinite temperature and B is at finite temperature in the canonical ensemble, that is,

$$\mathcal{I}_\alpha = \mathbf{1}_{A_1} \otimes \mathbf{1}_{A_2} \otimes e^{-\beta H_B}. \quad (\text{C.13})$$

Then the r.h.s. of figure 24 implies that R_p, Σ_p again become independent of the index p , but in this case the r.h.s. does not simplify to a geometric series, so that we have

$$R_N = d_A R_1, \quad R_1 = \frac{1}{\lambda - \Sigma_1}, \quad (\text{C.14})$$

where

$$\Sigma_1 = \frac{1}{d_A} + \sum_{n=1}^{\infty} \frac{Z_{2n+1,B}}{(d_A Z_{1,B})^{2n+1}} (d_{A_1} R_1)^{2n} + \sum_{n=1}^{\infty} \frac{Z_{2n,B}}{(d_A Z_{1,B})^{2n}} d_{A_2} (d_{A_1} R_1)^{2n-1} \quad (\text{C.15})$$

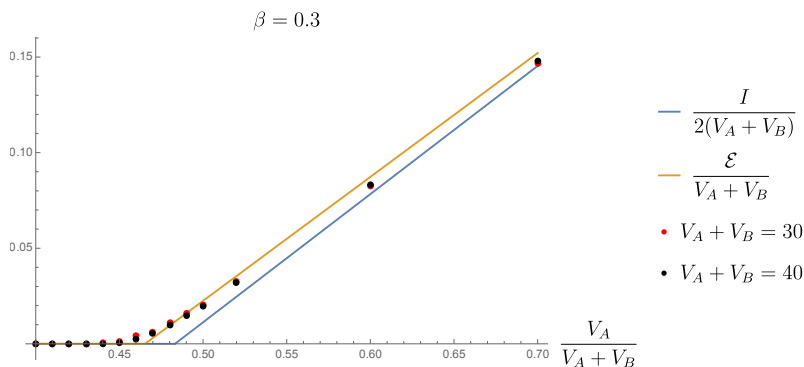


Figure 25. Left: for \mathcal{I}_α given by (C.13), the logarithmic negativity is computed numerically using the resolvent (C.17) and compared to the naive analytic continuation in (4.53), (4.55) and (4.56), with excellent agreement.

where $Z_{n,B} = \text{Tr}[e^{-n\beta H_B}]$. The resulting equation for R_N can be written as

$$\lambda R_N = d_A + d_{A_2} \int dE \rho(E) \sum_{k=1}^{\infty} \left[\left(\frac{R_N e^{-\beta E}}{d_A d_{A_2} Z_{1,B}} \right)^{2k-1} + d_{A_2} \left(\frac{R_N e^{-\beta E}}{d_A d_{A_2} Z_{1,B}} \right)^{2k} \right]. \quad (\text{C.16})$$

Completing the geometric sums, we find

$$\lambda R_N = d_A + \int dE \rho(E) \frac{d_{A_2}^2 R_N (d_A Z_{1,B} e^{\beta E} + R_N)}{d_A^2 d_{A_2}^2 Z_{1,B}^2 e^{2\beta E} - R_N^2}. \quad (\text{C.17})$$

where $\rho(E) = e^{Vs(E/V)}$ is the density of states for B . On specifying the density of states, this equation can be solved numerically for $R_N(\lambda)$ and used to obtain \mathcal{E} . The solution with a gaussian entropy density is shown in figure 25 as a function of c at $\lambda = \frac{1}{2}$. The result agrees with the expressions for \mathcal{E} in the NE and ES phases in (4.53) and (4.55) and the naive phase transition line between them in (4.56).

Similar to the infinite temperature case, we can consider the regime where the effective dimensions d_{A_1}, d_{A_2}, d_B of the different subsystems are such that $\frac{d_{A_1}}{d_{A_2} d_B} \ll 1$. Assuming that the same set of permutations contributes in this regime at finite temperature as the one that contributed at infinite temperature, we can then include only the first diagram from each line on the r.h.s. of figure 24. We then have

$$\Sigma_1 = \frac{1}{d_A} + \frac{Z_{2,B}}{d_A (Z_{1,B})^2} R_1 \quad (\text{C.18})$$

which gives a quadratic equation

$$\alpha' R_1^2 + \left(\frac{1}{d_A} - \lambda \right) R_1 + 1 = 0, \quad \alpha' = \frac{Z_{2,B}}{(Z_{1,B})^2 d_A} \quad (\text{C.19})$$

Similar to (C.10), this gives a semicircle form of the density of eigenvalues, which for this case gives

$$\mathcal{Z}_{\infty, \text{sc}}^{(\text{PT})} = \begin{cases} 1 & \log d_A < S_{2,B}^{(\text{eq})} \\ e^{\frac{1}{2}(\log d_A - S_{2,B}^{(\text{eq})})} & \log d_A > S_{2,B}^{(\text{eq})} \end{cases}. \quad (\text{C.20})$$

This gives the value of the logarithmic negativity we would expect in the ES phase from naive analytic continuation, and also the transition between the NE and ES phases we would expect from naive analytic continuation.

C.3.1 Toy model for black hole evaporation in JT gravity

As discussed in [1], the rules for calculating the entanglement entropy of the radiation in the holographic model of black hole evaporation in [4] can be derived from the equilibrium approximation, taking A to be the radiation R , B to be the black hole, and

$$\mathcal{I}_\alpha = \mathbf{1}_R \otimes f(H_B)e^{-\beta H_B} \tag{C.21}$$

which is a special case of (C.13), with an additional factor $f(H_B)$ which captures the effect of end-of-the-world (EOW) branes, and with the density of states as in JT gravity, so that

$$\mathcal{Z}_{m,B} = e^{S_0} z_m(\beta) \tag{C.22}$$

where S_0 is the large ground state entropy and $z_m(\beta)$ is $O(1)$. The rules for calculating $\mathcal{Z}_n^{(\text{PT})}$ in this model can similarly be derived from the equilibrium approximation. The negativity within the radiation in this model was discussed in refs. [55, 56].

Like in other cases, the dominant permutation in $\mathcal{Z}_n^{(\text{PT})}$ for this choice of \mathcal{I}_α is always one out of e , η or η^{-1} , and τ_{ES} . We can read off the “boundary” expressions for $\mathcal{Z}_n^{(\text{PT})}(\tau)$ in this holographic model, in terms of $d_{R_1}, d_{R_2}, Z_{n,B}$, from the diagrams in figure 3 (where $Z_{n,B} = \text{Tr}[(e^{-\beta H_B} f(H_B))^n]$). Using the relation between bulk and boundary partition functions in holography, we then find geometries like the ones shown in figure 26 for evaluating $\mathcal{Z}_n^{(\text{PT})}$ in the ES and ME phases. Both involve “replica wormholes,” which are connected Euclidean gravity path integrals between multiple asymptotic boundaries. Note that more generally, (3.4) can be used to systematically derive replica wormhole prescriptions for calculating $\mathcal{Z}_n^{(\text{PT})}$ in other gravity setups, including the model of [5].

Note that (C.22) implies in particular that

$$\begin{aligned} S_{2,B}^{(\text{eq})} &= S_0 + s_2(\beta), & S_B^{(\text{eq})} &= S_0 + s_1(\beta), \\ s_1(\beta) &> s_2(\beta) &\sim O(1) \end{aligned} \tag{C.23}$$

From analytic continuation of the expressions for $\mathcal{Z}_n^{(\text{PT})}$ for even n to $n \rightarrow 1$ in this model, the expressions for the logarithmic negativity in (4.53)–(4.55) become

$$\mathcal{E}_{NE}(R_1, R_2) = 0 \tag{C.24}$$

$$\mathcal{E}_{ME}(R_1, R_2) = \log d_{R_1} \quad \text{or} \quad \mathcal{E}_{ME} = \log d_{R_2} \tag{C.25}$$

$$\mathcal{E}_{ES}(R_1, R_2) = \frac{1}{2} \left(\log d_R - S_0 - s_2(\beta) + \log \frac{8}{3\pi} \right) \tag{C.26}$$

Note that we have explicitly included a term $\log \frac{8}{3\pi}$ coming from the degeneracy of the permutations in the ES phase in this case (which we have ignored for other choices of \mathcal{I}_α in the paper), as it contributes at the same order as $\log \frac{z_2(\beta)}{z_1(\beta)^2}$. Recall that the Page time t_p corresponds to $c = \frac{1}{2}$. The transition from the NE phase, where the logarithmic negativity

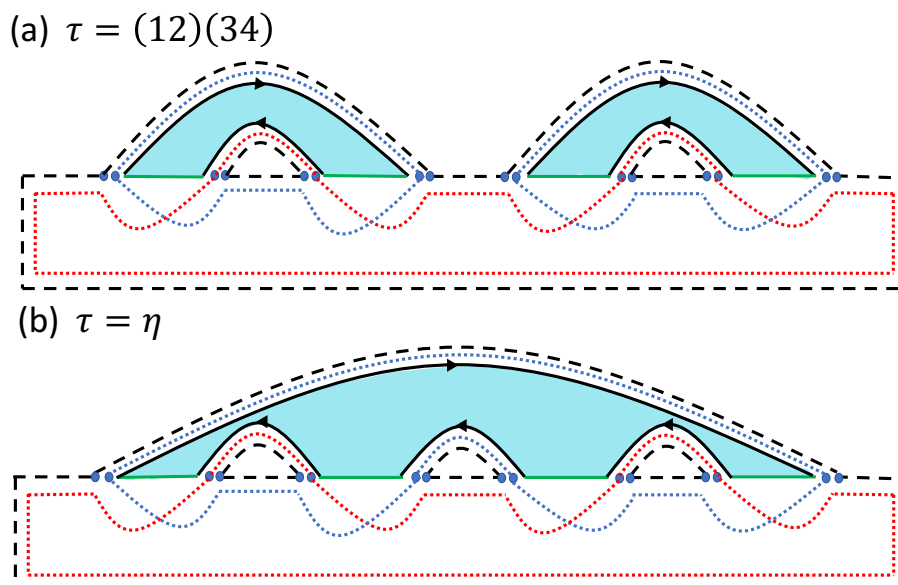


Figure 26. Euclidean gravity path integrals with replica wormholes for calculating $\mathcal{Z}_n^{(\text{PT})}$ in the model of [4], for $n = 4$. The dominant contribution in the ES phase, shown in (a), involves connected partition functions $Z_{2,B}$ between two asymptotic boundaries, while the ME phase contribution in (b) involves $Z_{n,B}$. The black lines with arrows are asymptotic boundaries in JT gravity (each with length β , although the lengths appear to differ in the figure), and the green lines are EOW branes. The dashed and dotted loops simply give factors of d_{A_1} and d_{A_2} respectively.

within the radiation is sub-extensive, to the ES phase, where it becomes extensive, takes place at

$$c = c_0 \equiv \frac{1}{2} \left(1 - \frac{s_1(\beta) - s_2(\beta) + \log 8\pi/3}{S_0} \right) \tag{C.27}$$

For this model, the difference between $c = \frac{1}{2}$ and $c = c_0$ is suppressed due to the special structure of the density of states in JT gravity compared to other black hole models.

It is also interesting to consider the negativity between the black hole and a subsystem of the radiation, R_1 . In this case, the permutation that dominates in $\mathcal{Z}_n^{(\text{PT})}$ at early times when $e^{S_0} \gg d_{R_1} d_{R_2}$ is $\tau = \eta^{-1}$, leading on analytic continuation to

$$\mathcal{E}(R_1, B) = \log d_{R_1}. \tag{C.28}$$

irrespective of the relative sizes of R_1 and R_2 . This means that each Hawking quantum is maximally entangled with the black hole at early times, as expected. After the Page time, if $d_{R_1} \gg d_{R_2} e^{S_0}$, $\tau = \eta$ dominates, and by analytic continuation the negativity is then

$$\mathcal{E}(R_1, B) = S_0 + \log \frac{z_{\frac{1}{2}}(\beta)^2}{z_1(\beta)} = S_{\frac{1}{2}, B}^{(\text{eq})}. \tag{C.29}$$

Most interestingly, we can also consider the regime after the Page time where $d_{R_2} \gg d_{R_1} e^{S_0}$ i.e. how small amounts of the radiation are entangled with the black hole. The usual story of black hole evaporation says that these quanta are maximally entangled with the black

hole, more specifically, their pair behind the horizon. This entanglement played a starring role in the firewall paradox [57]. But the dominant term in $\mathcal{Z}_n^{(\text{PT})}$ is the identity, so that on analytic continuation,

$$\mathcal{E}(R_1, B) = 0. \quad (\text{C.30})$$

This is the statement that Hawking quanta after the Page time are not actually entangled with the black hole, but only entangled with the early radiation, avoiding paradoxes with entanglement monogamy.

Note that we can numerically compute the resolvent and spectrum in this model using the general equation (C.17), now with

$$\rho(E) = e^{S_0} \frac{\sqrt{2E}}{2\pi^2} \sinh(2\pi\sqrt{2E}) \quad (\text{C.31})$$

This problem is studied systematically in ref. [56]. For the special case of the “microcanonical ensemble” in [4], the equation for the resolvent reduces to the infinite temperature case in (C.8), with the dimension of the black hole given by $e^{\mathbf{S}}$, where \mathbf{S} is the microcanonical entropy of the black hole.

C.4 Microcanonical ensemble with B at infinite temperature

We now consider \mathcal{I}_α given by (4.77), which we repeat here:²⁰

$$\mathcal{I}_\alpha = \sum_{E_{a_1}^{A_1} + E_{a_2}^{A_2} \in I_{E, \Delta}} |a_1\rangle \langle a_1| \otimes |a_2\rangle \langle a_2| \otimes \mathbf{1}_B. \quad (\text{C.32})$$

(i) Contributions to $\mathcal{Z}_n^{(\text{PT})}$ from different permutations. Let us first understand the contributions to $\mathcal{Z}_n^{(\text{PT})}$ from different permutations in this case. Recall that the contribution from τ can be expanded as

$$\begin{aligned} \langle \eta_{A_1} \otimes \eta_{A_2}^{-1} \otimes e_B | \mathcal{I}_\alpha, \tau \rangle &= \sum_{i_1, i'_1, \dots, i_n, i'_n} \langle \eta_{A_1} \otimes \eta_{A_2}^{-1} \otimes e_B | i_1 \bar{i}'_1 \dots i_n \bar{i}'_n \rangle \\ &\times \langle i_1 | \mathcal{I}_\alpha | i'_{\tau(1)} \rangle \langle i_2 | \mathcal{I}_\alpha | i'_{\tau(2)} \rangle \dots \langle i_n | \mathcal{I}_\alpha | i'_{\tau(n)} \rangle \end{aligned} \quad (\text{C.33})$$

where

$$|i_m\rangle = |i_{m_a}\rangle_{A_1} |i_{m_{\bar{a}}}\rangle_{A_2} |i_{m_b}\rangle_B, \quad |\bar{i}'_m\rangle = |\bar{i}'_{m_a}\rangle_{A_1} |\bar{i}'_{m_{\bar{a}}}\rangle_{A_2} |\bar{i}'_{m_b}\rangle_B. \quad (\text{C.34})$$

It is useful to take the states $|i_{m_a}\rangle_{A_1}$, $|i_{m_{\bar{a}}}\rangle_{A_2}$, $|i_{m_b}\rangle_B$ to be energy eigenstates in A_1, A_2, B . Then note that for (4.77), each factor in the second line of (C.33) simultaneously gives delta functions between the indices on A_1, A_2 and B in the bra and the ket, and enforces the condition that the energies in A_1 and A_2 add up to E . For instance,

$$\langle i_{a_1} i_{\bar{a}_1} i_{b_1} | \mathcal{I}_\alpha | i'_{a_1} i'_{\bar{a}_1} i'_{b_1} \rangle = \delta_{i_{a_1} i'_{a_1}} \delta_{i_{\bar{a}_1} i'_{\bar{a}_1}} \delta_{i_{b_1} i'_{b_1}} \delta_{E_{a_1}^{A_1} + E_{\bar{a}_1}^{A_2}, E} \quad (\text{C.35})$$

In figure 27, we explain the resulting contributions from different τ with the help of diagrams, showing how the energies of different index loops are related.

²⁰Similar resolvent calculations are done in ref. [58].

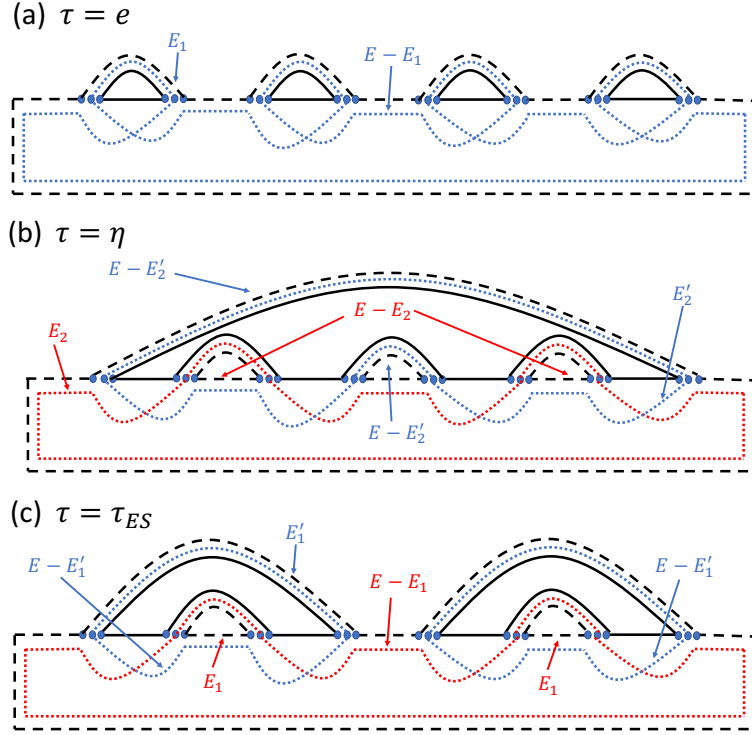


Figure 27. Diagrams for different possible dominant contributions to $\mathcal{Z}_n^{(\text{PT})}$ in the microcanonical ensemble with B at infinite temperature. Recall that dashed, dotted and solid lines respectively represent A_1 , A_2 , B . E_s, E'_s are energies in A_s for $s = 1, 2$, and each of these energies is independently summed over in the expression for $\mathcal{Z}_n^{(\text{PT})}(\tau)$.

Then if the density of states in subsystem P at energy E_P is $d_{E_P}^P$, we find

$$\mathcal{Z}_n^{(\text{PT})}(e) = \frac{1}{N_E^n} d_B^n \sum_{E_1} d_{E_1}^{A_1} d_{E-E_1}^{A_2}, \quad (\text{C.36})$$

$$\mathcal{Z}_n^{(\text{PT})}(\tau_{ES}) = \frac{1}{N_E^n} d_B^{n/2} \left(\sum_{E_1} (d_{E_1}^{A_1})^{n/2} d_{E-E_1}^{A_2} \right) \left(\sum_{E'_1} d_{E'_1}^{A_1} (d_{E-E'_1}^{A_2})^{n/2} \right), \quad (\text{C.37})$$

$$\mathcal{Z}_n^{(\text{PT})}(\eta) = \frac{1}{N_E^n} d_B \left(\sum_{E_2} d_{E_2}^{A_2} (d_{E-E_2}^{A_1})^{n/2} \right) \left(\sum_{E'_2} d_{E'_2}^{A_2} (d_{E-E'_2}^{A_1})^{n/2} \right) \quad (\text{C.38})$$

where

$$N_E = \text{Tr}[\mathcal{I}_\alpha] = d_B \sum_{E_1} d_{E_1}^{A_1} d_{E-E_1}^{A_2}. \quad (\text{C.39})$$

On using the saddle point approximation for each of these terms, we find that the saddle point values of the energies in the resulting expressions can be identified as the saddle points that appear in expressions for Renyi entropies of equilibrium density matrices for various subsystems and indices. As a result, we can further simplify (C.36)–(C.38) to get the expressions in terms of equilibrium Renyi entropies in (4.78)–(4.81).

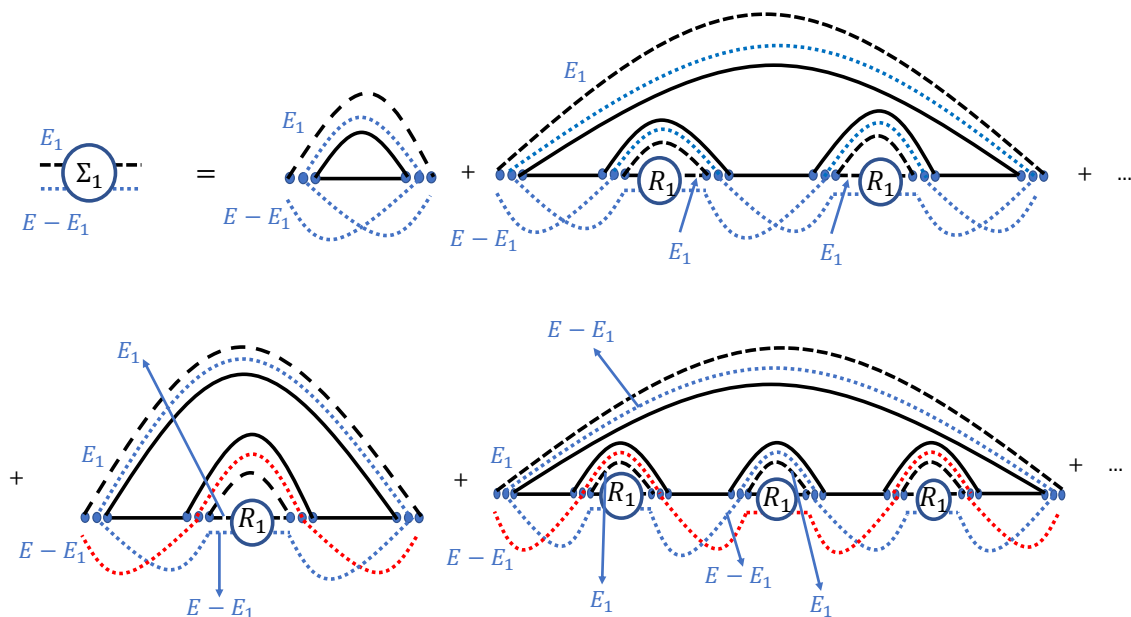


Figure 28. Schwinger-Dyson equation for different contributions to $\Sigma_1(E_1)$.

(ii) Resolvent and logarithmic negativity. Let us now understand what the Schwinger-Dyson equations in figure 23(a) and 24 give for this choice of \mathcal{I}_α . Recall that we take $|p_s\rangle, |q_s\rangle$ to be energy eigenstates in A_s . Then we can see that R_p depends only on the energies E_1 and E_2 of $|p_1\rangle_{A_1}$ and $|p_2\rangle_{A_2}$. So we can express the resolvent as a sum

$$R_N(\lambda) = \sum_{E_1, E_2} d_{E_1}^{A_1} d_{E_2}^{A_2} \tilde{R}(E_1, E_2), \quad \tilde{R}(E_1, E_2) = \begin{cases} R_1(E_1) & E_1 + E_2 = E \\ R_2(E_1, E_2) & E_1 + E_2 \neq E \end{cases} \quad (\text{C.40})$$

The set of diagrams contributing to $\tilde{R}(E_1, E_2)$ differs depending on whether or not the external energies E_1 and E_2 add up to E , giving the two cases R_1 and R_2 in (C.40). For example, out of the diagrams explicitly shown on the r.h.s. of figure 23(b), in the case where $E_1 + E_2 = E$, all diagrams on the r.h.s. of give a non-zero contribution, whereas in the case where $E_1 + E_2 \neq E$, only the second diagram gives a non-zero contribution, as the remaining diagrams enforce $E_1 + E_2 = E$. Then we have

$$R_1(E_1) = \frac{1}{\lambda - \Sigma_1(E_1)}, \quad R_2(E_1, E_2) = \frac{1}{\lambda - \Sigma_2(E_1, E_2)} \quad (\text{C.41})$$

where Σ_1, Σ_2 are expressed in terms of R_1, R_2 through the Schwinger-Dyson equations in figure 28 and 29.

Contribution from permutations with only one-cycles and two-cycles

To evaluate Σ_1 and Σ_2 , we can first consider the contributions from all diagrams which have only one-cycles and two-cycles, based again on the expectation that this should capture the regime where the effective dimensions are such that $\frac{d_{A_1}}{d_B d_{A_2}} \ll 1$. We can restrict to such contributions by keeping only the first term in each line on the r.h.s. of figure 28, and only the first term on the r.h.s. of figure 29.

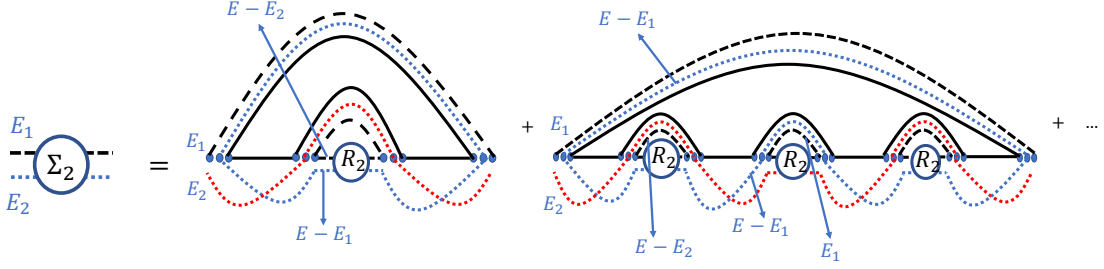


Figure 29. Schwinger-Dyson equation for different contributions to $\Sigma_2(E_1, E_2)$.

We then find

$$\Sigma_1(E_1) = \frac{d_B}{N_E} + \frac{d_{E-E_1}^{A_2} d_{E_1}^{A_1} d_B}{N_E^2} R_1(E_1) \quad (\text{C.42})$$

which together with (C.41) gives a quadratic equation for R_1 ,

$$\alpha_E R_1^2 + \left(\frac{d_B}{N_E} - \lambda \right) R_1 + 1 = 0, \quad \alpha_E = \frac{d_{E_1}^{A_1} d_{E-E_1}^{A_2} d_B}{N_E^2} \quad (\text{C.43})$$

which has the solution

$$R_1(E_1, E - E_1) = \frac{\lambda - \frac{d_B}{N_E}}{2\alpha_E} - \frac{1}{2\alpha_E} \sqrt{\left(\lambda - \frac{d_B}{N_E} \right)^2 - 4\alpha_E}. \quad (\text{C.44})$$

From the first diagram on the r.h.s. of (29), we find

$$\Sigma_2(E_1, E_2) = \frac{d_B d_{E-E_2}^{A_1} d_{E-E_1}^{A_2}}{N_E^2} R_2(E - E_2, E - E_1) \quad (\text{C.45})$$

Using the analogous equation for $\Sigma_2(E - E_2, E - E_1)$ together with (C.41), we get a quadratic equation for $R_2(E_1, E_2)$,

$$\lambda \alpha_1 R_2^2 - (\lambda^2 + \alpha_1 - \alpha_2) R_2 + \lambda = 0, \quad \alpha_1 = \frac{d_{E_1}^{A_1} d_{E_2}^{A_2} d_B}{N_E^2}, \quad \alpha_2 = \frac{d_{E-E_2}^{A_1} d_{E-E_1}^{A_2} d_B}{N_E^2} \quad (\text{C.46})$$

which has the solution

$$R_2(E_1, E_2) = \frac{\lambda^2 + \alpha_1 - \alpha_2 - \sqrt{(\lambda - a_1)(\lambda - a_2)(\lambda + a_2)(\lambda + a_1)}}{2\lambda\alpha_1} \quad (\text{C.47})$$

where

$$a_1 = \sqrt{\alpha_1} + \sqrt{\alpha_2}, \quad a_2 = |\sqrt{\alpha_1} - \sqrt{\alpha_2}|. \quad (\text{C.48})$$

By substituting R_1 and R_2 back into (C.40) and finding the imaginary part, we can find the density of eigenvalues of $\rho_A^{T_2}$:

$$D_N(\lambda) = \sum_{E_1} \frac{1}{p_{E_1}} D_\infty^{\text{sc}}(\tilde{\lambda}, d_{E_1}^{A_1} d_{E-E_1}^{A_2}, d_B) + \sum_{E_1 \neq E_2} D_{E_1, E_2}(\lambda) \quad (\text{C.49})$$

where the first term is expressed in terms of the infinite temperature value of $D_N(\lambda)$ in the semi-circle regime from (C.11), and we have introduced

$$p_{E_1} = \frac{d_{E_1}^{A_1} d_{E-E_1}^{A_2} d_B}{N_E}, \quad \tilde{\lambda} = \frac{\lambda}{p_{E_1}}. \quad (\text{C.50})$$

The second term of (C.49) is found by noting that the second term in (C.47) is imaginary between $-a_1$ and $-a_2$, and between a_2 and a_1 ,

$$D_{E_1, E_2}(\lambda) = \frac{1}{2}(d_{E_1}^{A_1} d_{E_2}^{A_2} - d_{E-E_2}^{A_1} d_{E-E_1}^{A_2})\delta(\lambda) + \frac{N_E^2}{2\pi|\lambda|d_B} \sqrt{a_1^2 - \lambda^2} \sqrt{\lambda^2 - a_2^2} \theta(a_1^2 - \lambda^2) \theta(\lambda^2 - a_2^2) \quad (\text{C.51})$$

So the exponential of the logarithmic negativity, $\sum_i |\lambda_i|$, is given by

$$\mathcal{Z}^{(\text{PT})} = \sum_{E_1} p_{E_1} \mathcal{Z}_{\infty, \text{sc}}^{(\text{PT})}(d_{E_1}^{A_1} d_{E-E_1}^{A_2}, d_B) + \sum_{E_1, E_2} \frac{N_E^2}{\pi d_B} \int_{a_2}^{a_1} d\lambda \sqrt{a_1^2 - \lambda^2} \sqrt{\lambda^2 - a_2^2} \quad (\text{C.52})$$

where $\mathcal{Z}_{\infty, \text{sc}}^{(\text{PT})}$ is given by (C.20).

The integral in (C.52) evaluates to

$$\int_{a_2}^{a_1} d\lambda \sqrt{a_1^2 - \lambda^2} \sqrt{\lambda^2 - a_2^2} = \frac{2}{3} a_1^3 E \left(1 - \frac{a_2^2}{a_1^2}\right) + \frac{2}{3} a_1 a_2^2 E \left(1 - \frac{a_2^2}{a_1^2}\right) - \frac{4}{3} a_1 a_2^2 K \left(1 - \frac{a_2^2}{a_1^2}\right) \quad (\text{C.53})$$

where K and E respectively refer to complete elliptic integrals of the first and second kind. Now since $a_2/a_1 \approx 1 + e^{-cV} + \dots$ for some positive coefficient c , to find the leading contribution in volume we can approximate both of the elliptic functions appearing above with the leading orders in their Taylor expansions as their argument goes to 0, which are both given by $\pi/2$. We can then approximate

$$\int_{a_2}^{a_1} d\lambda \sqrt{a_1^2 - \lambda^2} \sqrt{\lambda^2 - a_2^2} \approx \frac{\pi}{3} a_1 (a_1^2 - a_2^2) = \frac{4\pi}{3} (\sqrt{\alpha_1} + \sqrt{\alpha_2}) \sqrt{\alpha_1 \alpha_2} \quad (\text{C.54})$$

So we find

$$\begin{aligned} \mathcal{Z}^{(\text{PT})} &= \frac{d_B}{N_E} \sum_{E_1, d_{E_1}^{A_1} d_{E-E_1}^{A_2} < d_B} d_{E_1}^{A_1} d_{E-E_1}^{A_2} + \frac{\sqrt{d_B}}{N_E} \sum_{E_1, d_{E_1}^{A_1} d_{E-E_1}^{A_2} > d_B} (d_{E_1}^{A_1})^{3/2} (d_{E-E_1}^{A_2})^{3/2} \\ &+ \frac{8}{3} \frac{\sqrt{d_B}}{N_E} \sum_{E_1} d_{E_1}^{A_1} (d_{E-E_1}^{A_2})^{\frac{1}{2}} \times \sum_{E_2} d_{E_2}^{A_2} (d_{E-E_2}^{A_1})^{\frac{1}{2}} \end{aligned} \quad (\text{C.55})$$

The saddle point energy density for the first two terms in the last line is the same, and is equal to $\epsilon = E/V_A$. When $c = \frac{S_A^{(\text{eq})}}{S_A^{(\text{eq})} + S_B^{(\text{eq})}} < \frac{1}{2}$, this saddle point lies in the regime of the first term, which then contributes 1. The contribution from the second term is given by its boundary value in this regime, which is always less than the saddle-point value of the same expression, given by $e^{\frac{1}{2}(S_{A_1}^{(\text{eq})} + S_{A_2}^{(\text{eq})} - S_B^{(\text{eq})})}$. For $c > \frac{1}{2}$, the second term contributes

$e^{\frac{1}{2}(S_{A_1}^{(\text{eq})} + S_{A_2}^{(\text{eq})} - S_B^{(\text{eq})})}$ and dominates over the first. The last term can always be approximated with its saddle point value, which is given by $e^{\frac{1}{2}(S_{\frac{1}{2},A_1}^{(\text{eq})} + S_{\frac{1}{2},A_2}^{(\text{eq})} - S_B^{(\text{eq})})}$ (where we have ignored an overall factor that gives an $O(1)$ contribution to negativity). For $S_{\frac{1}{2},A_1}^{(\text{eq})} + S_{\frac{1}{2},A_2}^{(\text{eq})} > S_B^{(\text{eq})}$, this contribution dominates over the other two terms, giving

$$\mathcal{E} = \frac{1}{2} \left(S_{\frac{1}{2},A_1}^{(\text{eq})} + S_{\frac{1}{2},A_2}^{(\text{eq})} - S_B^{(\text{eq})} \right) \quad (\text{C.56})$$

Both the ES phase value of the negativity (C.56) and the transition line $S_{\frac{1}{2},A_1}^{(\text{eq})} + S_{\frac{1}{2},A_2}^{(\text{eq})} = S_B^{(\text{eq})}$ in this approximation thus agree with the naive analytic continuation in (4.84) and (4.86).

Contribution from all planar permutations

Let us now sum over all contributions on the r.h.s. of figure 28 and 29. While in this case we are not able to find the logarithmic negativity in all regimes from this calculation either analytically or numerically, we can use the expression for the full sum to systematically understand the regime in which the above result from all permutations with only one- and two- cycles should be valid.

On the r.h.s. of figure 28(b), each of the two lines can be obtained by summing over a geometric series, so that we get

$$\Sigma_1 = \left(\frac{d_B}{N_E} + \alpha_E R_1 \right) \frac{1}{1 - \beta_E R_1^2}, \quad \beta_E = \frac{(d_{E_1}^{A_1})^2}{N_E^2} \quad (\text{C.57})$$

Here we have suppressed the arguments of both Σ_1 and R_1 , which are the same in all expressions. This gives a cubic equation for $R_1(E_1)$:

$$\beta_E \lambda R_1^3 + (\alpha_E - \beta_E) R_1^2 + \left(\frac{d_B}{N_E} - \lambda \right) R_1 + 1 = 0. \quad (\text{C.58})$$

We see that this reduces to the contribution from only one-cycles and two-cycles in (C.43) when $\alpha_E \gg \beta_E$ for all E_1 , i.e. when

$$d_{E-E_1}^{A_2} d_B \gg d_{E_1}^{A_1}, \quad \text{for all } E_1. \quad (\text{C.59})$$

Comparing this to the cubic equation for the infinite temperature case (C.8), we find that the first term in (C.40) is given by

$$\sum_{E_1} d_{E_1}^{A_1} d_{E-E_1}^{A_2} R_1(E_1, E - E_1) = \sum_{E_1} \frac{1}{p_{E_1}} R_N^\infty(d_{E_1}^{A_1}, d_{E-E_1}^{A_2}, d_B) \quad (\text{C.60})$$

where R_N^∞ is the infinite temperature resolvent. The contribution to $\mathcal{Z}^{(\text{PT})}$ from this term is therefore

$$\mathcal{Z}_1^{(\text{PT})} = \sum_{E_1} p_{E_1} \mathcal{Z}_\infty^{(\text{PT})}(d_{E_1}^{A_1}, d_{E-E_1}^{A_2}, d_B). \quad (\text{C.61})$$

Now for evaluating Σ_2 , the r.h.s. of (29) also gives a geometric series (suppressing the λ argument in all Σ_2, R_2 below):

$$\begin{aligned}\Sigma_2(E_1, E_2) &= \frac{d_{E-E_2}^{A_1} d_{E-E_1}^{A_2} d_B}{N_E^2} R_2(E-E_2, E-E_1) (1 + \tilde{\Delta} + \tilde{\Delta}^2 + \dots), \\ \tilde{\Delta} &= \frac{d_{E_1}^{A_1} d_{E-E_2}^{A_1}}{N_E^2} R(E_1, E_2) R(E-E_2, E-E_1)\end{aligned}\tag{C.62}$$

To simplify the notation, let us fix some E_1 and E_2 , and denote

$$\Sigma_2 := \Sigma_2(E_1, E_2), \quad \Sigma'_2 := \Sigma_2(E-E_2, E-E_1), \quad R_2 := \Sigma_2(E_1, E_2), \quad R'_2 := R_2(E-E_2, E-E_1)\tag{C.63}$$

Then from (C.62) and (C.41), we have

$$\lambda - \frac{1}{R_2} = \Sigma_2 = \frac{d_{E-E_2}^{A_1} d_{E-E_1}^{A_2} d_B}{N_E^2} R'_2 \frac{1}{1 - \frac{d_{E_1}^{A_1} d_{E-E_2}^{A_1}}{N_E^2} R_2 R'_2}\tag{C.64}$$

$$\lambda - \frac{1}{R'_2} = \Sigma'_2 = \frac{d_{E_1}^{A_1} d_{E_2}^{A_2} d_B}{N_E^2} R_2 \frac{1}{1 - \frac{d_{E-E_2}^{A_1} d_{E_1}^{A_1}}{N_E^2} R_2 R'_2}\tag{C.65}$$

From (C.64) and (C.65), we get the following cubic equation for R_2 :

$$\alpha_3 \lambda^2 R_2^3 + \lambda(\alpha_4 + \alpha_1 - 2\alpha_3) R_2^2 - (\lambda^2 + \alpha_4 - \alpha_3 - \alpha_2 + \alpha_1) R_2 + \lambda = 0.\tag{C.66}$$

where

$$\alpha_3 = \frac{(d_{E_1}^{A_1})^2 d_{E_2}^{A_2}}{d_{E-E_1}^{A_2} N_E^2}, \quad \alpha_4 = \frac{d_{E_1}^{A_1} d_{E-E_2}^{A_1}}{N_E^2},\tag{C.67}$$

and α_1, α_2 were defined in (C.46). This reduces to the contribution from one-cycles and two-cycles in (C.46) if $\alpha_1 \gg \alpha_3, \alpha_1 \gg \alpha_4$, i.e. if we have (C.59) and the additional condition

$$d_{E_1}^{A_2} d_B \gg d_{E-E_1}^{A_2} \quad \text{for all } E_1.\tag{C.68}$$

The solutions to this equation are

$$R_2 = \frac{2a - b - c}{3a\lambda} + \frac{e^{-qi2\pi/3}}{6a\lambda^2} 2^{1/3} \frac{Z}{(X + \sqrt{Y})^{1/3}} + \frac{e^{iq2\pi/3}}{6a\lambda^2} 2^{2/3} (X + \sqrt{Y})^{1/3}\tag{C.69}$$

with $q = 0, 1, -1$, and

$$\begin{aligned}X &= \lambda^3 (-9a(a+b+c)\lambda^2 - (2a-b-c)(a^2 - 2(b+c)^2 - a(b+c-9d))) \\ Z &= \lambda^2 (6a\lambda^2 + 2(a^2 + (b+c)^2 - a(b+c+3d))) \\ Y &= X^2 - \frac{Z^3}{2}\end{aligned}\tag{C.70}$$

In the limit $\lambda \rightarrow \infty$,

$$\begin{aligned}Z &\rightarrow 6a\lambda^4 \left(1 + O\left(\frac{1}{\lambda^2}\right)\right), \\ (X + \sqrt{Y})^{1/3} &= e^{i\pi/6} 2^{1/3} \sqrt{3} a^{1/2} \lambda^2 \left(1 + i \frac{a(a+b+c)}{2\sqrt{3}a^{3/2}\lambda}\right)\end{aligned}\tag{C.71}$$

and hence

$$R_2 = \frac{2a - b - c}{3a\lambda} + \frac{e^{-i(\frac{2\pi q}{3} + \frac{\pi}{6})}}{\sqrt{3}\sqrt{a}} \left(1 - i \frac{a + b + c}{2\sqrt{3}a^{3/2}\lambda}\right) + \frac{e^{i(\frac{2\pi q}{3} + \frac{\pi}{6})}}{\sqrt{3}\sqrt{a}} \left(1 + i \frac{a + b + c}{2\sqrt{3}a^{3/2}\lambda}\right) + O\left(\frac{1}{\lambda^2}\right) \quad (\text{C.72})$$

For the solutions with $q = 0, 1$, the leading order in the $\lambda \rightarrow \infty$ limit is $O(\lambda^0)$, which implies that the leading behaviour of R coming from these solutions is also $O(\lambda^0)$. Since R should approach $\frac{d_A}{\lambda}$, these solutions are not allowed. For $q = -1$, we find in this limit

$$R_2(\lambda) = \frac{1}{\lambda} + O\left(\frac{1}{\lambda^2}\right) \quad (\text{C.73})$$

$R_1(E_1; \lambda)$ also approaches $\frac{1}{\lambda}$. We therefore find

$$R_N(\lambda) = \sum_{E_1, E_2} \frac{d_{E_1}^{A_1} d_{E_2}^{A_2}}{\lambda} + O\left(\frac{1}{\lambda^2}\right) = \frac{d_A}{\lambda} + O\left(\frac{1}{\lambda^2}\right) \quad (\text{C.74})$$

which is the expected behaviour.

When λ is sufficiently large, $Y < 0$, and R_2 solution is real. For $Y > 0$, we find

$$\text{Im}R_2 = \frac{1}{4\sqrt{3}a\lambda^2} \left| \frac{Z}{(|X + \sqrt{Y}|)^{1/3}} 2^{1/3} - 2^{2/3} (|X + \sqrt{Y}|)^{1/3} \right| \quad (\text{C.75})$$

This expression can in principle be used to find $D(\lambda)$ and $\mathcal{Z}^{(\text{PT})}$ for this case, but finding the support of $\text{Im}R_2$ and performing the integral to find the contribution to $\mathcal{Z}^{(\text{PT})}$ is difficult to do analytically.

C.5 Microcanonical ensemble with A_2 at infinite temperature

(i) Contributions to $\mathcal{Z}_n^{(\text{PT})}$ from different permutations. Unlike in the earlier example (4.77), when \mathcal{I}_α is given by (4.87), the contributions to $\mathcal{Z}_n^{(\text{PT})}$ from different τ can be expressed in a simple way in terms of the cycle numbers of τ , $\eta\tau$ and $\eta\tau^{-1}$. In this case, in any diagram, if we take the energy in any index loop of A_1 to be E_1 , the energy conservation constraint between A_1 and B immediately sets all energies in index loops of A_1 to be E_1 , and the energies in index loops of B to $E - E_1$. Some examples are shown in figure 30.

We therefore find in this case that

$$\mathcal{Z}_n^{(\text{PT})} = \sum_{\tau} \mathcal{Z}_n^{(\text{PT})}(\tau), \quad \mathcal{Z}_n^{(\text{PT})}(\tau) = \frac{1}{N_E^n} (d_{A_2})^{k(\eta\tau)} \sum_{E_1} (d_{E_1}^{A_1})^{k(\eta\tau^{-1})} (d_{E-E_1}^B)^{k(\tau)} \quad (\text{C.76})$$

where $N_E = d_{A_2} \sum_{E_1} d_{E_1}^{A_1} d_{E-E_1}^B$. Then on using the saddle-point approximation in these expressions and comparing to the saddle points appearing in various equilibrium Renyi entropies, we get the expressions for different possible dominant permutations in (4.89)–(4.92).

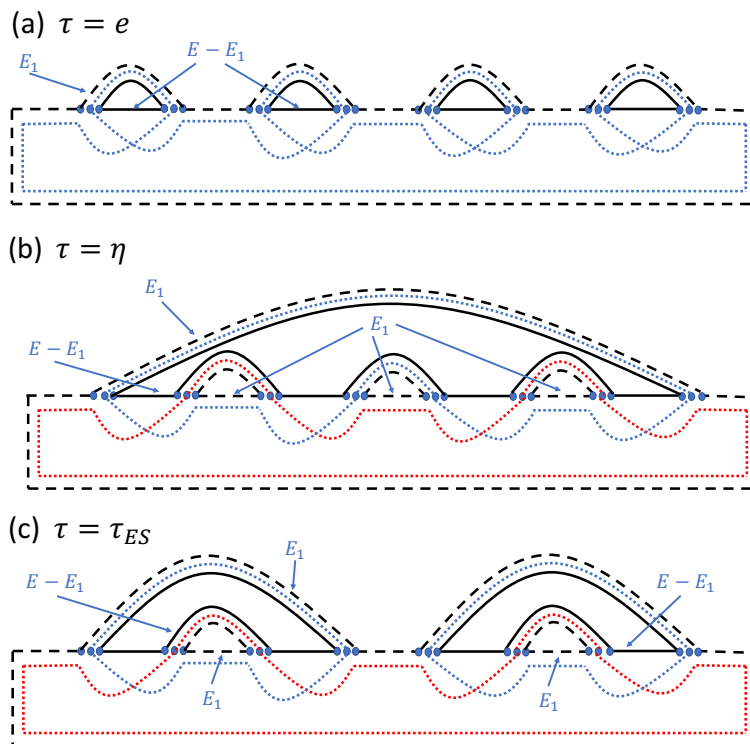


Figure 30. Diagrams for different possible dominant contributions to $\mathcal{Z}_n^{(\text{PT})}$ in the microcanonical ensemble with A_2 at infinite temperature. Recall that dashed, dotted and solid lines respectively represent A_1 , A_2 , B . In this example, all energies in A_1 get set to some E_1 which is summed over, and each of the energies in B gets set to $E - E_1$.

(ii) Logarithmic negativity via resolvent. In this example, due to the simplicity of the expressions for $\mathcal{Z}_n^{(\text{PT})}$, it is more useful to write down an expression for the resolvent by directly summing over $\mathcal{Z}_n^{(\text{PT})}/\lambda^n$ as in the definition (2.27), instead of using the Schwinger-Dyson equation like in previous examples. Recall that in the infinite temperature case

$$\mathcal{Z}_{n,\infty}^{(\text{PT})} = \frac{1}{(d_A d_B)^n} \sum_{\tau \in \mathcal{S}_n} (d_{A_2})^{k(\eta\tau)} \sum_{E_1} (d_{A_1})^{k(\eta\tau^{-1})} (d_B)^{k(\tau)} \quad (\text{C.77})$$

Then from (C.76), (C.77) and (2.27), we find

$$R_N(\lambda) = \sum_{E_1} \frac{1}{p_{E_1}} R_N^\infty(\tilde{\lambda}, d_{E_1}^{A_1}, d_{A_2}, d_{E-E_1}^B), \quad (\text{C.78})$$

where

$$p_{E_1} = \frac{d_{E_1}^{A_1} d_{A_2} d_{E-E_1}^B}{N_E}, \quad \tilde{\lambda} = \lambda/p_{E_1}. \quad (\text{C.79})$$

Using (2.28), we then find that

$$D_N(\lambda) = \sum_{E_1} \frac{1}{p_{E_1}} D_N^\infty(\tilde{\lambda}, d_{E_1}^{A_1}, d_{A_2}, d_{E-E_1}^B) \quad (\text{C.80})$$

where D_N^∞ is the spectral density of the partially transposed density matrix at infinite temperature, and hence

$$\mathcal{Z}^{(\text{PT})} := \sum_i |\lambda_i| = \sum_{E_1} p_{E_1} \mathcal{Z}_\infty^{(\text{PT})}(d_{A_1}^{E_1}, d_{A_2}, d_B^{E-E_1}) \quad (\text{C.81})$$

where $\mathcal{Z}_\infty^{(\text{PT})}$ is given by (C.9). We therefore have

$$\begin{aligned} \mathcal{Z}^{(\text{PT})} \approx & \sum_{E_1, d_{E_1}^{A_1} d_{A_2} < d_{E-E_1}^B} p_{E_1} + d_{A_2} \times \sum_{E_1, d_{A_2} d_{E-E_1}^B < d_{E_1}^{A_1}} p_{E_1} + \sum_{E_1, d_{E_1}^{A_1} d_{E-E_1}^B < d_{A_2}} p_{E_1} d_{E_1}^{A_1} \\ & + d_{A_2}^{1/2} \times \sum_{E_1, d_{E_1}^{A_1} d_{A_2} > d_{E-E_1}^B, \frac{1}{d_{E-E_1}^B} < \frac{d_{E_1}^{A_1}}{d_{A_2}} < d_{E-E_1}^B} p_{E_1} (d_{E_1}^{A_1})^{1/2} (d_{E-E_1}^B)^{-1/2} \end{aligned} \quad (\text{C.82})$$

In the limit where V is large, each term in (C.82) can be evaluated at the energy that maximizes its exponent, and $\mathcal{Z}^{(\text{PT})}$ is approximately equal to the maximum among the four terms. For each of the terms, there are some values of the volumes V_{A_1}, V_{A_2}, V_B such that the saddle-point energy lies within the range of E_1 specified for that term, and other values where it does not. In the latter case, we must use the boundary value rather than the saddle-point value for that term.

Before further analyzing (C.82), it will be useful to introduce some definitions. Recall that as we vary the sizes of various subsystems, we keep the average energy density $\frac{E}{V_{A_1} + V_B}$ fixed to ϵ , and the infinite temperature entropy density in B fixed to s_0 . If $\epsilon_1 = \frac{E_1}{V_{A_1}}$ is an energy density in A , we define the corresponding energy density in B due to energy conservation as $\bar{\epsilon}_1$,

$$\bar{\epsilon}_1 = \frac{\epsilon(V_{A_1} + V_B) - \epsilon_1 V_{A_1}}{V_B}. \quad (\text{C.83})$$

Also, we define $\epsilon_1^{(\alpha)}$ as the solution to the equation

$$s'(\epsilon_1^{(\alpha)}) = \alpha s'(\bar{\epsilon}_1^{(\alpha)}). \quad (\text{C.84})$$

For example, $\epsilon_1^{(1)}$ is the saddle point for evaluating $N_E = \text{Tr}[\mathcal{I}_\alpha]$, and is equal to ϵ for any form of $s(\epsilon)$. With $s(\epsilon)$ as in (4.76),

$$\epsilon_1^{(\alpha)} = \frac{\epsilon(V_{A_1} + V_B)}{V_{A_1} + V_B \alpha^2}. \quad (\text{C.85})$$

Also, note that the equilibrium entropies of the full system and different subsystems are given by

$$\begin{aligned} S^{(\text{eq})} &= V_{A_2} s_0 + (V_{A_1} + V_B) s(\epsilon), & S_{A_1}^{(\text{eq})} &= V_{A_1} s(\epsilon), & S_B^{(\text{eq})} &= V_B s(\epsilon), \\ S_{A_2}^{(\text{eq})} &= S_{n, A_2}^{(\text{eq})} = V_{A_2} s_0, \\ S_{\alpha, A_1}^{(\text{eq})} &= V_{A_1} \frac{\alpha s(\epsilon) - s(\epsilon_1^{(\alpha)})}{\alpha - 1} + V_B \frac{\alpha s(\epsilon) - \alpha s(\bar{\epsilon}_1^{(\alpha)})}{\alpha - 1}, & \alpha &\neq 1 & (\text{C.86}) \\ S_{\alpha, B}^{(\text{eq})} &= V_B \frac{\alpha s(\epsilon) - s(\epsilon_1^{(1/\alpha)})}{\alpha - 1} + V_{A_1} \frac{\alpha s(\epsilon) - \alpha s(\epsilon_1^{(1/\alpha)})}{\alpha - 1}, & \alpha &\neq 1. \end{aligned}$$

Considering when the dominant contributions to different terms in (C.82) are given by their saddle-point or boundary values, we then find that

$$\mathcal{E}(A_1, A_2) = \log \mathcal{Z}^{(\text{PT})} \approx \left[\max_{i=1, \dots, 4} f_i - S^{(\text{eq})} \right] \quad (\text{C.87})$$

where

$$f_1 = \begin{cases} S^{(\text{eq})} & S_A^{(\text{eq})} < S_B^{(\text{eq})} \\ S_{A_2}^{(\text{eq})} + V_{A_1} s(\theta_1) + V_B s(\bar{\theta}_1) & S_A^{(\text{eq})} > S_B^{(\text{eq})} \end{cases} \quad (\text{C.88})$$

$$f_2 = \begin{cases} S_{A_2}^{(\text{eq})} + S^{(\text{eq})} & S_{A_2}^{(\text{eq})} + S_B^{(\text{eq})} < S_{A_1}^{(\text{eq})} \\ 2S_{A_2}^{(\text{eq})} + V_{A_1} s(\theta_2) + V_{A_2} s(\bar{\theta}_2) & S_{A_2}^{(\text{eq})} + S_B^{(\text{eq})} > S_{A_1}^{(\text{eq})} \end{cases} \quad (\text{C.89})$$

$$f_3 = \begin{cases} S_{\frac{1}{2}, A_1}^{(\text{eq})} + S^{(\text{eq})} & V_{A_1} s(\epsilon_1^{(1/2)}) + V_B s\left(\overline{\epsilon_1^{(1/2)}}\right) < V_{A_2} s_0 \\ S_{A_2}^{(\text{eq})} + 2V_{A_1} s(\theta_3) + V_B s(\bar{\theta}_3) & V_{A_1} s(\epsilon_1^{(1/2)}) + V_B s\left(\overline{\epsilon_1^{(1/2)}}\right) > V_{A_2} s_0 \end{cases} \quad (\text{C.90})$$

$$f_4(c, \lambda) = \begin{cases} \frac{3}{2} S_{A_2}^{(\text{eq})} + S_{A_1, \frac{1}{3}}^{(\text{eq})} & V_{A_1} s(\epsilon_1^{(1/3)}) + V_{A_2} s_0 > V_B s\left(\overline{\epsilon_1^{(1/3)}}\right), \text{ and} \\ + \frac{1}{2} (S_{A_1}^{(\text{eq})} + S_B^{(\text{eq})}) & V_{A_2} s_0 < V_{A_1} s(\epsilon_1^{(1/3)}) + V_B s\left(\overline{\epsilon_1^{(1/3)}}\right), \text{ and} \\ & V_{A_1} s(\epsilon_1^{(1/3)}) < V_{A_2} s_0 + V_B s\left(\overline{\epsilon_1^{(1/3)}}\right) \\ \max_{i=1, 2, 3} \left[\frac{3}{2} S_{A_2}^{(\text{eq})} + \frac{3}{2} V_{A_1} s(\theta_i) \right. & \\ \left. + \frac{1}{2} V_B s(\bar{\theta}_i) \right] & \text{otherwise} \end{cases} \quad (\text{C.91})$$

In the above expressions, θ_i are the values of the energy density lying at the boundaries of the regimes corresponding to the different terms in (C.82), and are defined implicitly as solutions to the equations

$$\begin{aligned} V_{A_1} s(\theta_1) + V_{A_2} s_0 &= V_B s(\bar{\theta}_1), \\ V_{A_2} s_0 + V_B s(\bar{\theta}_2) &= V_{A_1} s(\theta_2), \\ V_{A_1} s(\theta_3) + V_B s(\bar{\theta}_3) &= V_{A_2} s_0, \end{aligned} \quad (\text{C.92})$$

which we can solve numerically for the entropy density in (4.76). For certain values of the subsystem volumes, we will find that solutions for some of the θ_i do not exist; this corresponds to cases where the range of energies corresponding to some of the terms in (C.82) do not exist.

Let us try to understand the phase diagram for the logarithmic negativity in terms of the parameters $c = \frac{S_A^{(\text{eq})}}{S_A^{(\text{eq})} + S_B^{(\text{eq})}}$ and $\lambda = \frac{S_{A_1}^{(\text{eq})}}{S_A^{(\text{eq})}}$. We first find the ranges of c, λ for which the different terms f_1 through f_4 are given by their saddle-point values (the first lines of equations (C.88)–(C.91)) in figure 31(a).

The parameter ranges where f_2 and f_3 take their saddle-point values are shown respectively by the green and blue regions in figure 31(a). In these regions, we find that f_2 and f_3 respectively are also larger than the other f_i . Hence, in these regions the logarithmic

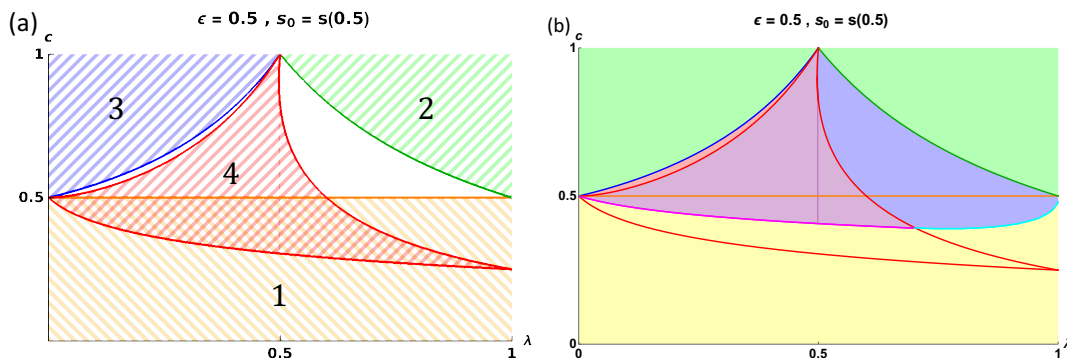


Figure 31. Regions 1-4 in (a) show the parameter ranges for which f_1 through f_4 are given by the saddle-point values in the first lines of (C.88)–(C.91). The resulting phase diagram for the logarithmic negativity is shown in (b). Different shaded regions correspond to distinct phases. All lines from (a), including those that do not correspond to phase transition lines of the logarithmic negativity, are also shown in (b) to make it easier to compare regions in the two figures.

negativity is given by

$$\mathcal{E} = \begin{cases} S_{1/2, A_1}^{(\text{eq})} & \lambda < \frac{1}{2} \\ S_{A_2}^{(\text{eq})} & \lambda > \frac{1}{2} \end{cases} \quad (\text{C.93})$$

There is some intersection between the orange and red shaded regions in figure 31(a) where f_1 and f_4 respectively take their saddle point values. On comparing the first lines of (C.88) and (C.91), we find that the latter dominates for

$$\frac{1}{2} S_{A_2}^{\text{eq}} + S_{1/3, A_1}^{\text{eq}} - \frac{1}{2} S_{A_1}^{\text{eq}} - \frac{1}{2} S_B^{\text{eq}} > 0. \quad (\text{C.94})$$

The condition (C.94) is shown with the pink line in figure 31(b). In the purple shaded region of figure 31(b), we find that f_4 is also larger than f_2 and f_3 , and hence in this region the logarithmic negativity is given by

$$\mathcal{E} = \frac{1}{2} S_{A_2}^{\text{eq}} + S_{1/3, A_1}^{\text{eq}} - \frac{1}{2} S_{A_1}^{\text{eq}} - \frac{1}{2} S_B^{\text{eq}}. \quad (\text{C.95})$$

In the unshaded regions of figure 31 (a), none of the terms of (C.82) take their saddle-point values. In the left unshaded region of figure 31 (a), which is the same as the region shaded in red in figure 31 (b), we find that the dominant contribution is given by the second lines of f_3 and f_4 , which coincide with each other. In this region, the logarithmic negativity is given by

$$\mathcal{E} = S_{A_2}^{(\text{eq})} + V_{A_1} (s(\theta_3) - s(\epsilon)) - V_B s(\epsilon). \quad (\text{C.96})$$

In the blue shaded region of figure 31(b), which includes both the right unshaded region of figure 31(a) and part of the orange shaded region there, the dominant contribution is given by the second lines of f_2 and f_4 , which coincide. The logarithmic negativity here is given by

$$\mathcal{E} = V_{A_1} (2s(\theta_2) - s(\epsilon)) - V_B s(\epsilon). \quad (\text{C.97})$$

Finally, in the yellow shaded region of figure 31(b), we have

$$\mathcal{E} = 0. \quad (\text{C.98})$$

Each of the expressions (C.93), (C.95), (C.96), (C.97), (C.98) corresponds to a distinct phase, and there is a discontinuity in either the first or the second derivative of \mathcal{E} across each of the phase transition lines, as shown in figure 12.

D Maxima of functions of permutations at finite temperature

In this appendix, we use general properties of the function $f(\beta)$ in any system to show that $G(\beta, \tau)$ (as defined in (4.64)) is maximized by $\tau = e$, while $G(\beta, \tau) + G(\beta, \eta^{-1}\tau)$ is maximized by both $\tau = e$ and $\tau = \eta$.

Recall

$$G(\beta, \tau) = \sum_{i=1}^{k(\tau)} f(c_i \beta) \quad (\text{D.1})$$

and

$$G(\beta, \tau) + G(\beta, \eta^{-1}\tau) = \sum_{i=1}^{k(\tau)} f(c_i \beta) + \sum_{j=1}^{k(\eta^{-1}\tau)} f(a_j \beta) \quad (\text{D.2})$$

Both (D.1) and (D.2) are examples of sums of the form

$$P(S, m, \{c_i\}) := \sum_{i=1}^m f(c_i \beta), \quad \sum_{i=1}^m c_i = S, \quad 1 \leq c_i \leq n \quad (\text{D.3})$$

where the m numbers c_i are all positive integers, and S is either n or $2n$. In fact, note that the above constraints together further restrict the range of the c_i , to

$$1 \leq c_i \leq S - (m - 1). \quad (\text{D.4})$$

To simplify the problem of finding τ that maximize (D.1) and (D.2), we first ask: with the constraint in (D.4), what are the values of m, c_i for which the function $P(S, m, \{c_i\})$ is maximized?

We first fix m , and see what set of m numbers c_i satisfying (D.4) maximizes $P(S, m, \{c_i\})$. Note that since each c_i lies between 1 and $S - (m - 1)$, it can be written as

$$c_i = \lambda + (1 - \lambda)(S - (m - 1)) \quad (\text{D.5})$$

where $\lambda = \frac{S - m + 1 - c_i}{S - m}$ lies between 0 and 1. Then using (4.33) for each term in $P(m, S, \{c_i\})$, we find

$$\begin{aligned} P(S, m, \{c_i\}) &\leq \sum_{i=1}^m \left(\frac{S - m + 1 - c_i}{S - m} f(\beta) + \frac{c_i - 1}{S - m} f((S - m + 1)\beta) \right) \\ &\leq Q(m, S) := (m - 1)f(\beta) + f((S - m + 1)\beta) \end{aligned} \quad (\text{D.6})$$

Note further that $Q(m, S)$ is an increasing function of m for both choices of S , since (4.37) implies in particular that

$$f(\beta) - \beta f'((2n - m + 1)\beta) \geq 0, \quad f(\beta) - \beta f'((n - m + 1)\beta) \geq 0. \quad (\text{D.7})$$

Hence, the following statements hold:

1.

$$G(\tau) = \sum_{i=1}^{k(\tau)} f(c_i\beta) \leq nf(\beta) \quad (\text{D.8})$$

This is because on comparing the definition of $G(\tau)$ to that of $P(S, m, \{c_i\})$, we see that $S = n$, and the maximum value of m that can appear in $G(\tau)$ is n . But the r.h.s. of (D.8) is the value of the l.h.s. for $\tau = e$, so $G(\tau) = \sum_{i=1}^{k(\tau)} f(c_i\beta)$ is maximized by $\tau = e$.

2.

$$G(\tau) + G(\eta\tau^{-1}) = \sum_{i=1}^{k(\tau)} f(c_i\beta) + \sum_{j=1}^{k(\eta\tau^{-1})} f(a_j\beta) \leq nf(\beta) + f(n\beta) \quad (\text{D.9})$$

This follows since the maximum number of terms in this full expression is $n+1$ since $k(\tau) + k(\eta\tau^{-1}) \leq n+1$, and here $S = 2n$. The r.h.s. is the value of the l.h.s. for $\tau = e$ and $\tau = \eta$, so $G(\tau) + G(\eta\tau^{-1})$ is maximized by both.

E $\mathcal{Z}_n^{(\text{PT})}$ in the microcanonical ensemble with energy conservation in AB

Let us now consider the equilibrium approximation with \mathcal{I}_α as in (4.75). Unlike in the simpler example in the previous subsection, we are not able to compute the logarithmic negativity through the resolvent in this example. However, we can find the phase diagram for $\mathcal{Z}_n^{(\text{PT})}$ and the Renyi negativities, which have the same qualitative features as those for the canonical example and the simple example of the previous subsection.

Let us assume that like in other universality classes, the dominant contribution is always given by one out of e , η , η^{-1} , and τ_{ES} . Let us try to understand the contributions that we get from each of these permutations. In each case, we can use diagrams similar to figure 27, now taking into account energy conservation between all three subsystems A_1 , A_2 , and B . We then find:

1.

$$\mathcal{Z}_n^{(\text{PT})}(e) = \frac{1}{N_E^n} \sum_{E_1, E_2} d_{E_1}^{A_1} d_{E_2}^{A_2} (d_{E-E_1-E_2}^B)^n \quad (\text{E.1})$$

Expressing each density of states in terms of the corresponding entropy density (e.g. $d_{E_1}^{A_1} = \exp(V_{A_1} s(\frac{E_1}{V_{A_1}}))$), we find that the saddle-point equations for the energy densities $\epsilon_1 = \frac{E_1}{V_{A_1}}$ and $\epsilon_2 = \frac{E_2}{V_{A_2}}$ are

$$s'(\epsilon_1) = ns' \left(\frac{V\epsilon - V_{A_1}\epsilon_1 - V_{A_2}\epsilon_2}{V_B} \right) \quad (\text{E.2})$$

$$s'(\epsilon_2) = ns' \left(\frac{V\epsilon - V_{A_1}\epsilon_1 - V_{A_2}\epsilon_2}{V_B} \right) \quad (\text{E.3})$$

where $\epsilon = \frac{E}{V}$. Since the r.h.s. of (E.2) and (E.3) is the same, we can assume that $\epsilon_1 = \epsilon_2 = \epsilon_e$, and solve

$$s'(\epsilon_e) = ns' \left(\frac{V\epsilon - V_A\epsilon_e}{V_B} \right) \quad (\text{E.4})$$

We therefore find

$$\mathcal{Z}_n^{(\text{PT})}(e) \approx e^{V_A s(\epsilon_e) + n V_B s\left(\frac{V\epsilon - V_A \epsilon e}{V_B}\right) - n V s(\epsilon)}. \quad (\text{E.5})$$

Taking the $n \rightarrow 1$ limit of (E.4) and (E.5) gives

$$\mathcal{E} = 0. \quad (\text{E.6})$$

2.

$$\mathcal{Z}_n^{(\text{PT})}(\eta) = \frac{1}{N_E^n} \sum_{E_B, E_2, E'_2} d_{E_B}^B d_{E_2}^{A_2} d_{E'_2}^{A_2} (d_{E-E_2-E_B}^{A_1})^{\frac{n}{2}} (d_{E-E'_2-E_B}^{A_1})^{\frac{n}{2}} \quad (\text{E.7})$$

The saddle-point equations for the energy densities $\epsilon_B = \frac{E_B}{V_B}$, $\epsilon_2 = \frac{E_2}{V_{A_2}}$, and $\epsilon'_2 = \frac{E'_2}{V_{A_2}}$ are

$$s'(\epsilon_B) - \frac{n}{2} s' \left(\frac{\epsilon V - V_{A_2} \epsilon_2 - V_B \epsilon_B}{V_{A_1}} \right) - \frac{n}{2} s' \left(\frac{\epsilon V - V_{A_2} \epsilon'_2 - V_B \epsilon_B}{V_{A_1}} \right) = 0 \quad (\text{E.8})$$

$$s'(\epsilon_2) - \frac{n}{2} s' \left(\frac{\epsilon V - V_{A_2} \epsilon_2 - V_B \epsilon_B}{V_{A_1}} \right) = 0 \quad (\text{E.9})$$

$$s'(\epsilon'_2) - \frac{n}{2} s' \left(\frac{\epsilon V - V_{A_2} \epsilon'_2 - V_B \epsilon_B}{V_{A_1}} \right) = 0 \quad (\text{E.10})$$

Since the system of equations is unchanged on exchanging ϵ_2 with ϵ'_2 everywhere, the solutions for both variables will be the same. Hence, we can simplify the equations by setting $\epsilon_2 = \epsilon'_2 = \epsilon_2^\eta$. Then we must solve the following two equations for ϵ_2^η and ϵ_B^η :

$$s'(\epsilon_B^\eta) - n s' \left(\frac{V\epsilon - V_{A_2} \epsilon_2^\eta - V_B \epsilon_B^\eta}{V_{A_1}} \right) = 0 \quad (\text{E.11})$$

$$s'(\epsilon_2^\eta) - \frac{n}{2} s' \left(\frac{V\epsilon - V_{A_2} \epsilon_2^\eta - V_B \epsilon_B^\eta}{V_{A_1}} \right) = 0 \quad (\text{E.12})$$

to find

$$\mathcal{Z}_n^{(\text{PT})}(\eta) \approx \exp(V_B s(\epsilon_B^\eta) + 2V_{A_2} s(\epsilon_2^\eta) + n V_{A_1} s\left(\frac{V - \epsilon V_{A_2} \epsilon_2^\eta - V_B \epsilon_B^\eta}{V_{A_1}}\right) - n V s(\epsilon)) \quad (\text{E.13})$$

For general forms of $s(\epsilon)$, there is no relation between (E.13) for general n and the Renyi entropies of $\rho^{(\text{eq})}$. But for $n = 1$, (E.11), (E.12) and (E.13) simplify such that we have

$$\mathcal{E} = S_{\frac{1}{2}, A_2}^{(\text{eq})} \quad (\text{E.14})$$

Similarly, taking the $n \rightarrow 1$ limit of $\mathcal{Z}_n^{(\text{PT})}(\eta^{-1})$, we find

$$\mathcal{E} = S_{\frac{1}{2}, A_1}^{(\text{eq})}. \quad (\text{E.15})$$

3. Now consider the contribution from τ_{ES} :

$$\begin{aligned} \mathcal{Z}_n^{(\text{PT})}(\tau_{ES}) &= \frac{1}{N_E^n} \sum_{\substack{E_1, E'_1, E_B^1, \\ E_B^2, \dots, E_B^{n/2}}} \left(\prod_{j=1}^{n/2} d_{E_B^j}^B \right) d_{E-E_1-E_B^1}^{A_2} \left(\prod_{j=1}^{n/2} d_{E-E'_1-E_B^j}^{A_2} \right) \\ &\quad \times d_{E'_1}^{A_1} d_{E_1}^{A_1} \left(\prod_{j=2}^{n/2} d_{E_1+E_B^1-E_B^j}^{A_1} \right) \end{aligned} \quad (\text{E.16})$$

The $n + 2$ saddle point equations are

$$\begin{aligned} s'(\epsilon_B^1) - s' \left(\frac{V\epsilon - V_{A_1}\epsilon_1 - V_B\epsilon_B^1}{V_{A_2}} \right) - s' \left(\frac{V\epsilon - V_{A_1}\epsilon'_1 - V_B\epsilon_B^1}{V_{A_2}} \right) \\ + \sum_{j=2}^{n/2} s' \left(\epsilon_1 + \frac{V_B}{V_{A_1}} (\epsilon_B^1 - \epsilon_B^j) \right) = 0 \end{aligned} \quad (\text{E.17})$$

$$s'(\epsilon_1) - s' \left(\frac{\epsilon - V_{A_1}\epsilon_1 - V_B\epsilon_B^1}{V_{A_2}} \right) + \sum_{j=2}^{n/2} s' \left(\epsilon_1 + \frac{V_B}{V_{A_2}} (\epsilon_B^1 - \epsilon_B^j) \right) = 0 \quad (\text{E.18})$$

$$s'(\epsilon'_1) - \sum_{j=1}^{n/2} s' \left(\frac{\epsilon V - V_{A_1}\epsilon'_1 - V_B\epsilon_B^j}{V_{A_2}} \right) = 0 \quad (\text{E.19})$$

$$s'(\epsilon_B^j) - s' \left(\frac{V\epsilon - V_{A_1}\epsilon'_1 - V_B\epsilon_B^j}{V_{A_2}} \right) - s' \left(\epsilon_1 + \frac{V_B}{V_{A_1}} (\epsilon_B^1 - \epsilon_B^j) \right) = 0, \quad j=2, \dots, n \quad (\text{E.20})$$

Note that this set of equations is invariant under permuting the ϵ_B^j from $j = 2$ to n . We can therefore assume that these variables have the same solution, so that we have a simpler set of four equations for any n :

$$\begin{aligned} s'(\epsilon_B^1) - s' \left(\frac{V\epsilon - V_{A_1}\epsilon_1 - V_B\epsilon_B^1}{V_{A_2}} \right) - s' \left(\frac{V\epsilon - V_{A_1}\epsilon'_1 - V_B\epsilon_B^1}{V_{A_2}} \right) \\ + \left(\frac{n}{2} - 1 \right) s' \left(\epsilon_1 + \frac{V_B}{V_{A_1}} (\epsilon_B^1 - \epsilon_B^j) \right) = 0 \end{aligned} \quad (\text{E.21})$$

$$s'(\epsilon_1) - s' \left(\frac{V\epsilon - V_{A_1}\epsilon_1 - V_B\epsilon_B^1}{V_{A_2}} \right) + \left(\frac{n}{2} - 1 \right) s' \left(\epsilon_1 + \frac{V_B}{V_{A_1}} (\epsilon_B^1 - \epsilon_B^j) \right) = 0 \quad (\text{E.22})$$

$$s'(\epsilon'_1) - s' \left(\frac{V\epsilon - V_{A_1}\epsilon'_1 - V_B\epsilon_B^1}{V_{A_2}} \right) - \left(\frac{n}{2} - 1 \right) s' \left(\frac{V\epsilon - V_{A_1}\epsilon'_1 - V_B\epsilon_B^j}{V_{A_2}} \right) = 0 \quad (\text{E.23})$$

$$\begin{aligned} s'(\epsilon_B^j) - s' \left(\frac{V\epsilon - V_{A_1}\epsilon'_1 - V_B\epsilon_B^j}{V_{A_2}} \right) - s' \left(\epsilon_1 + \frac{V_B}{V_{A_1}} (\epsilon_B^1 - \epsilon_B^j) \right) = 0, \\ j = 2, \dots, n \end{aligned} \quad (\text{E.24})$$

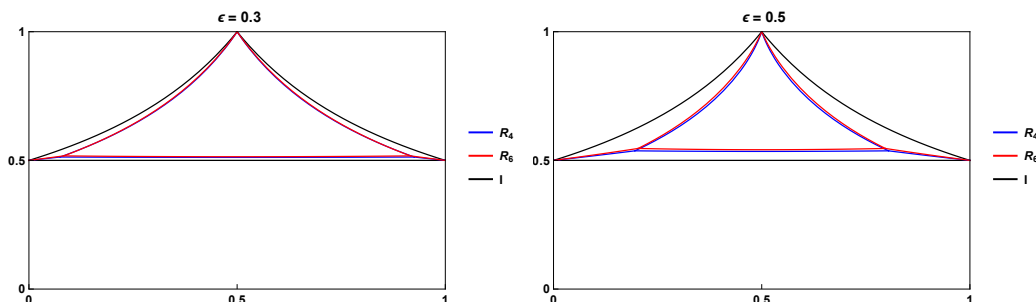


Figure 32. Phase transition lines for the $n = 4, 6$ Renyi negativities in the homogeneous micro-canonical ensemble case, with $s(\epsilon)$ given by (E.26).

In terms of the solution $\{\bar{\epsilon}_1, \bar{\epsilon}'_1, \bar{\epsilon}_B^j, \bar{\epsilon}_B^1\}$ to these equations, we can write

$$\begin{aligned} \mathcal{Z}_n^{(\text{PT})}(\tau_{\text{ES}}) \approx & \exp \left(V_B \left(\frac{n}{2} - 1 \right) s(\bar{\epsilon}_B^j) + V_B s(\bar{\epsilon}_B^1) + V_{A_2} s \left(\frac{V\epsilon - V_{A_1} \bar{\epsilon}_1 - V_B \bar{\epsilon}_B^1}{V_{A_2}} \right) \right) \quad (\text{E.25}) \\ & + V_{A_2} \left(\frac{n}{2} - 1 \right) s \left(\frac{V\epsilon - V_{A_1} \bar{\epsilon}'_1 - V_B \bar{\epsilon}_B^j}{V_{A_2}} \right) + V_{A_2} s \left(\frac{V\epsilon - V_{A_1} \bar{\epsilon}'_1 - V_B \bar{\epsilon}_B^1}{V_{A_2}} \right) \\ & + V_{A_1} s(\bar{\epsilon}'_1) + V_{A_1} s(\bar{\epsilon}_1) + V_{A_1} \left(\frac{n}{2} - 1 \right) s \left(\bar{\epsilon}_1 + \frac{V_B}{V_{A_1}} (\bar{\epsilon}_B^1 - \bar{\epsilon}_B^j) \right) - V n s(\epsilon) \end{aligned}$$

In this case, it is not clear how to take the $n \rightarrow 1$ limit, so we do not get simple expressions for \mathcal{E} like in the other phases.

Eqs. (E.21)–(E.24) are difficult to solve for general forms of $s(\epsilon)$, so that it is difficult to find even the phase diagram for $\mathcal{Z}_n^{(\text{PT})}$ for general $s(\epsilon)$. As one simple example, let us consider the Gaussian form of the entropy density

$$s(\epsilon) = \log 2 - \frac{1}{2} \epsilon^2 \quad (\text{E.26})$$

This form is an approximation for the entropy density of discrete systems, such as those with $g(\beta)$ given by (4.58), close to $\epsilon = 0$. The phase diagrams for $\mathcal{Z}_n^{(\text{PT})}$ with this form of $s(\epsilon)$ for the $n = 4$ and $n = 6$ are shown in figure 32.

F Finite temperature Page curve in inhomogeneous systems

In this appendix, we show that the expression (2.50) for the von Neumann entropy S_A in an equilibrated pure state holds in two examples where AB is inhomogeneous. In such examples, the Page transition in different Renyi entropies $S_{n,A}$ takes place at different values of $c = \frac{S_A^{(\text{eq})}}{S_A^{(\text{eq})} + S_B^{(\text{eq})}}$. Hence, we cannot a priori deduce the form of the von Neumann entropy by analytic continuation, and instead we must use the resolvent for the reduced density matrix ρ_A , defined in (2.13).

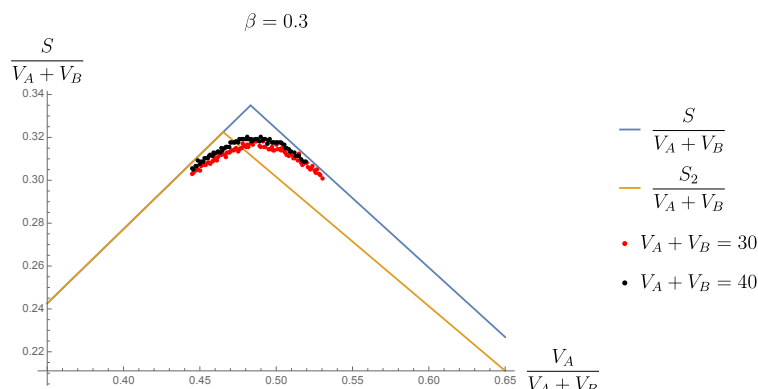


Figure 33. Numerical evaluation of the von Neumann entropy of A with \mathcal{I}_α as in (F.1) using the resolvent (F.3). We can see that the Page transition occurs at the point predicted by (2.50) and not at, for instance, the transition for the second Renyi entropy from (F.2). For simplicity, we have used a Gaussian density of states $\rho(E) \propto \exp\left[-\frac{E^2}{2N_B}\right]$.

F.1 Canonical ensemble with A at infinite temperature and \bar{A} at finite temperature

Consider the effective identity operator

$$\mathcal{I}_\alpha = \mathbf{1}_A \otimes e^{-\beta H_B}, \tag{F.1}$$

which we used previously to compute the negativity below (C.13). For this choice of \mathcal{I}_α ,

$$S_{n,A} = \min(\log d_A, S_{n,B}^{(\text{eq})}) \tag{F.2}$$

Since $S_{n,B}^{(\text{eq})}$ is n -dependent for $\beta > 0$, it is clear that the transition from increasing to decreasing behavior as a function of the volume fraction $\frac{V_A}{V}$ depends on n .

To find the von Neumann entropy, we can use a Schwinger-Keldysh equation to evaluate the resolvent. The calculation is exactly analogous to the derivation of (C.17) using the Schwinger-Keldysh equation, and we find an integral expression

$$\lambda \mathcal{R} = d_A + \int dE \rho(E) \sum_{k=1}^{\infty} \left(\frac{\mathcal{R} e^{-\beta E}}{d_A Z_{1,B}} \right)^k = d_A + \int dE \rho(E) \frac{\mathcal{R}}{d_A Z_{1,B} e^{\beta E} - \mathcal{R}}. \tag{F.3}$$

The von Neumann entropy can be evaluated numerically using this expression, confirming (2.50), as shown in figure 33.

F.2 Microcanonical ensemble with different entropy densities in A and B

If we take the effective identity operator to be

$$\mathcal{I}_\alpha = \sum_{p,q, E_p^A + E_q^{\bar{A}} = E} |p\rangle \langle p|_A \otimes |q\rangle \langle q|_B, \tag{F.4}$$

then the equilibrium approximation for $\mathcal{Z}_n^{(A)}$ in (2.39) gives

$$\mathcal{Z}_n^{(A)} \approx \sum_{\tau} \frac{(d_{\mathcal{E}}^A)^{k(\tau\eta^{-1})} (d_{E-\mathcal{E}}^B)^{k(\tau)}}{(N_E)^n}, \quad N_E = \sum_{\mathcal{E}} d_{\mathcal{E}}^A d_{E-\mathcal{E}}^B \quad (\text{F.5})$$

Let us consider the resolvent

$$\begin{aligned} R(\lambda) &= \frac{1}{\lambda} \sum_{n=0}^{\infty} \lambda^{-n} \mathcal{Z}_n^{(A)} = \sum_{\mathcal{E}} \frac{1}{\lambda} \sum_{n=0}^{\infty} \lambda^{-n} \frac{(d_{\mathcal{E}}^A d_{E-\mathcal{E}}^B)^n}{N_E^n} \sum_{\tau} \frac{(d_{\mathcal{E}}^A)^{k(\tau\eta^{-1})} (d_{E-\mathcal{E}}^B)^{k(\tau)}}{(d_{\mathcal{E}}^A d_{E-\mathcal{E}}^B)^n} \\ &= \sum_{\mathcal{E}} \frac{N_E}{d_{\mathcal{E}}^A d_{E-\mathcal{E}}^B} R_{\infty}(\tilde{\lambda}; d_{\mathcal{E}}^A, d_{E-\mathcal{E}}^B) = \sum_{\mathcal{E}} \frac{1}{p_{\mathcal{E}}} R_{\infty}(\tilde{\lambda}; d_{\mathcal{E}}^A, d_{E-\mathcal{E}}^B) \end{aligned} \quad (\text{F.6})$$

where

$$R_{\infty}(\lambda; d_A, d_B) = \frac{d_A - d_B}{2\lambda} + \frac{d_A d_B}{2} - \frac{d_A d_B}{2\lambda} \sqrt{(\lambda - \lambda_+) (\lambda - \lambda_-)}, \quad (\text{F.7})$$

is the resolvent for ρ_A in the infinite temperature case, and

$$\tilde{\lambda} = \frac{\lambda N_E}{d_{\mathcal{E}}^A d_{E-\mathcal{E}}^B}, \quad p_{\mathcal{E}} = \frac{d_{\mathcal{E}}^A d_{E-\mathcal{E}}^B}{N_E}. \quad (\text{F.8})$$

Eq. (F.6) can be used to express the density of eigenvalues $D(\lambda)$ of ρ_A , and its von Neumann and Renyi entropies, all in terms of their infinite temperature values,

$$D(\lambda) = \sum_{\mathcal{E}} \frac{1}{p_{\mathcal{E}}} D_{\infty}(\tilde{\lambda}; d_{\mathcal{E}}^A, d_{E-\mathcal{E}}^B) \quad (\text{F.9})$$

$$S_A = - \int D \lambda \lambda \log \lambda = \sum_{\mathcal{E}} p_{\mathcal{E}} S_A^{\infty}(\mathcal{E}) - \sum_{\mathcal{E}} p_{\mathcal{E}} \log p_{\mathcal{E}}, \quad (\text{F.10})$$

$$\mathcal{Z}_{n,A} = \int d\lambda D(\lambda) \lambda^n = \sum_{\mathcal{E}} (p_{\mathcal{E}})^n \mathcal{Z}_{n,A}^{\infty}(d_{\mathcal{E}}^A; d_{E-\mathcal{E}}^B). \quad (\text{F.11})$$

Let us first consider $S_n^{(A)}$ for $n \neq 1$. Since we know that

$$\mathcal{Z}_{n,A}^{\infty}(d_A, d_B) = \begin{cases} d_A^{1-n} & d_A < d_B \\ d_B^{1-n} & d_B < d_A \end{cases} \quad (\text{F.12})$$

we have

$$\mathcal{Z}_{n,A} = \sum_{\mathcal{E}, d_{\mathcal{E}}^A < d_{E-\mathcal{E}}^B} \frac{(d_{E-\mathcal{E}}^B)^n d_{\mathcal{E}}^A}{N_E^n} + \sum_{\mathcal{E}, d_{\mathcal{E}}^A > d_{E-\mathcal{E}}^B} \frac{d_{E-\mathcal{E}}^B (d_{\mathcal{E}}^A)^n}{N_E^n} \quad (\text{F.13})$$

Assuming that the saddle points for each of the above two integrals lie within their respective ranges, we then have

$$\mathcal{Z}_{n,A} \approx e^{-(n-1)S_{n,A}^{(\text{eq})}} + e^{-(n-1)S_{n,B}^{(\text{eq})}} \Rightarrow S_{n,A} = \min(S_{n,A}^{(\text{eq})}, S_{n,B}^{(\text{eq})}) \quad (\text{F.14})$$

If we take the entropy densities in A and B to have distinct forms, for instance,

$$d_E^A = e^{V_A s_A(\frac{E}{V_A})}, \quad d_E^B = e^{V_B s_B(\frac{E}{V_B})}, \quad s_A(\epsilon) = g \epsilon^{1/2}, \quad s_B(\epsilon) = g \epsilon^{3/4}, \quad (\text{F.15})$$

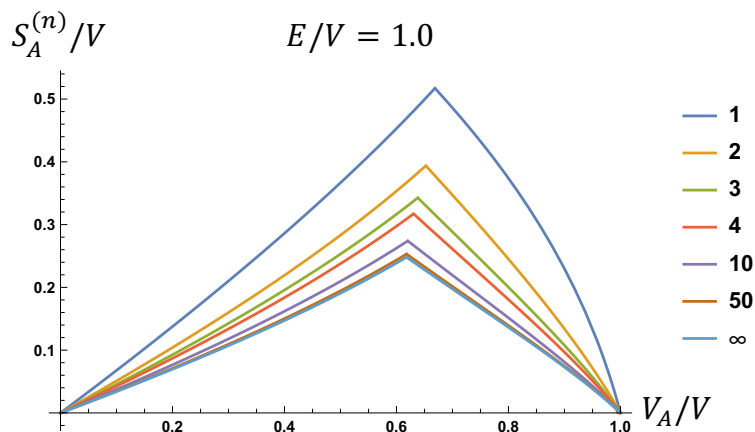


Figure 34. Renyi and von Neumann entropies as a function of volume fraction for the microcanonical ensemble with inhomogeneous entropy densities in A and B .

then as shown in figure 34, (F.14) implies that S_n for different n transition from increasing to decreasing behaviour as a function of the volume fraction $\frac{V_A}{V}$ at different points.

Let us now consider the von Neumann entropy in (F.10). We have

$$S_A = \sum_{\mathcal{E}} \frac{d_{\mathcal{E}}^A d_{E-\mathcal{E}}^B}{N_E} \min(\log d_{\mathcal{E}}^A, \log d_{E-\mathcal{E}}^B) - \sum_{\mathcal{E}} \frac{d_{\mathcal{E}}^A d_{E-\mathcal{E}}^B}{N_E} \log \left(\frac{d_{\mathcal{E}}^A d_{E-\mathcal{E}}^B}{N_E} \right) \quad (\text{F.16})$$

Let us evaluate each of the terms on the right-hand side in the saddle-point approximation. First note that

$$N_E = \sum_{\mathcal{E}} d_{\mathcal{E}}^A d_{E-\mathcal{E}}^B = V_A \int d\epsilon e^{V_A s_A(\epsilon) + V_B s_B\left(\frac{E-\epsilon V_A}{V_B}\right)} \approx \sqrt{\frac{2\pi}{B}} e^{S_A^{(\text{eq})} + S_B^{(\text{eq})}} \quad (\text{F.17})$$

where

$$B = -\frac{1}{V_A} s_A''(\bar{\epsilon}_A) - \frac{1}{V_B} s_B''(\bar{\epsilon}_B). \quad (\text{F.18})$$

Note that the von Neumann entropies $S_A^{(\text{eq})}$ and $S_B^{(\text{eq})}$ appear here as they are respectively the values of $V_A s_A(\epsilon)$ and $V_A s_A\left(\frac{E-V_A\epsilon}{V_A}\right)$ at the saddle-point value $\bar{\epsilon}$ of ϵ , which is the solution to

$$s'_A(\bar{\epsilon}) = s'_B\left(\frac{E - V_A\bar{\epsilon}}{V_B}\right). \quad (\text{F.19})$$

We can use the saddle-point approximation for the second term in (F.16) to see that it gives a contribution at $O(\log V)$. The saddle-point approximation for the first term is

$$S_A \approx V_A \int d\epsilon \frac{e^{-\frac{1}{2} B V_A^2 (\epsilon - \bar{\epsilon})^2}}{\sqrt{2\pi/B}} \min(S_A^{(\text{eq})} + V_A \beta (\epsilon - \bar{\epsilon}) + \dots, S_B^{(\text{eq})} - V_A \beta (\epsilon - \bar{\epsilon}) + \dots) \quad (\text{F.20})$$

Assume without loss of generality that $S_A^{(\text{eq})} < S_B^{(\text{eq})}$. We can write

$$\begin{aligned}
 S_A &\approx S_A^{(\text{eq})} \sqrt{\frac{B}{2\pi}} V_A \int_{2V_A\beta(\epsilon-\bar{\epsilon}) \leq S_B^{(\text{eq})} - S_A^{(\text{eq})}} d\epsilon e^{-\frac{1}{2}B V_A^2 (\epsilon-\bar{\epsilon})^2} \\
 &+ \sqrt{\frac{B}{2\pi}} V_A \int_{2V_A\beta(\epsilon-\bar{\epsilon}) \leq S_B^{(\text{eq})} - S_A^{(\text{eq})}} d\epsilon e^{-\frac{1}{2}B V_A^2 (\epsilon-\bar{\epsilon})^2} V_A\beta(\epsilon-\bar{\epsilon}) \\
 &+ \sqrt{\frac{B}{2\pi}} V_A \int_{2V_A\beta(\epsilon-\bar{\epsilon}) > S_B^{(\text{eq})} - S_A^{(\text{eq})}} d\epsilon e^{-\frac{1}{2}B V_A^2 (\epsilon-\bar{\epsilon})^2} (S_B^{(\text{eq})} - V_A\beta(\epsilon-\bar{\epsilon}))
 \end{aligned} \tag{F.21}$$

Let us consider the above expression in different regimes of the difference $S_B^{(\text{eq})} - S_A^{(\text{eq})}$.

1. If $S_B^{(\text{eq})} - S_A^{(\text{eq})} \gg \frac{1}{\sqrt{B}}$, then the first term in (F.21) is approximately S_A , while the remaining terms are approximately zero.
2. If $S_B^{(\text{eq})} = S_A^{(\text{eq})}$, then (F.21) becomes

$$\begin{aligned}
 S_A &= S_A^{(\text{eq})} \sqrt{\frac{B}{2\pi}} V_A \int_{-\infty}^{\infty} d\epsilon e^{-\frac{1}{2}B V_A^2 (\epsilon-\bar{\epsilon})^2} - \sqrt{\frac{B}{2\pi}} V_A \int_{-\infty}^{\infty} d\epsilon e^{-\frac{1}{2}B V_A^2 (\epsilon-\bar{\epsilon})^2} V_A\beta|\epsilon-\bar{\epsilon}| \\
 &= S_A^{(\text{eq})} - \sqrt{\frac{2\beta^2}{\pi B}}
 \end{aligned} \tag{F.22}$$

3. If $S_B^{(\text{eq})} - S_A^{(\text{eq})} \sim \frac{1}{\sqrt{B}}$, then the three terms in (F.21) should be evaluated more carefully, and we get corrections to the value of $S_A^{(\text{eq})}$ of order \sqrt{V} . This gives a smoothed out transition between the increasing and decreasing parts of the Page curve as a function of V_A . However, since $S_A^{(\text{eq})}$ and $S_B^{(\text{eq})}$ scale roughly linearly with V_A and V_B , the smoothing out of the transition is over a range of values of V_A of order \sqrt{V} . So as a function of the parameter $c := \frac{V_A}{V}$, the transition between the increasing and decreasing parts of the Page curve is still sharp in the thermodynamic limit.

The observations about enhanced corrections of order \sqrt{V} close to the Page transition in points 2 and 3 were made earlier in [59] and [60]. Similar effects were observed in the canonical ensemble in the model of [4]. Since we are interested in the sharp transition as a function of c in this discussion, we can ignore the effects of such corrections, and we then find that at leading order (F.20) implies

$$S_A = \min(S_A^{(\text{eq})}, S_B^{(\text{eq})}), \tag{F.23}$$

confirming (2.50) for this example. As shown in figure 34, when we take the inhomogeneous entropy densities in (F.15), the Page transition of the von Neumann entropy coming from (F.23) takes place at a distinct value of $\frac{V_A}{V}$ from the transitions of each of Renyi entropies from (F.14).

G Evolution of Petz map fidelity at finite temperature

G.1 Canonical ensemble

To evaluate the fidelity of the Petz map reconstruction $F(\rho, [\mathcal{P}_{\sigma, \mathcal{N}} \circ \mathcal{N}](\rho))$ systematically at finite temperature for the canonical ensemble, we use a method similar to section 3 of [4]. We take the reference state σ_D to be maximally mixed and ρ_D to be pure. For simplicity, we take R to be at infinite temperature, so that the effective identity operator is

$$\mathcal{I}_\alpha = e^{-\beta H_{B'}} \otimes \mathbf{1}_R. \quad (\text{G.1})$$

It is useful to define

$$F_{n_1, n_2} \equiv F_{n_1, n_2}^c + F_{n_1, n_2}^d \quad (\text{G.2})$$

for two non-negative integers n_1, n_2 , where

$$F_{n_1, n_2}^c \equiv \frac{1}{Z_1^{n_1+n_2+2}} \frac{1}{d_D^{n_1+n_2+2}} \sum_{\tau \in P_c} \langle \eta_R \otimes e_{B'} | \mathcal{I}_\alpha, \tau \rangle d_D^{k(\tau)} \quad (\text{G.3})$$

$$F_{n_1, n_2}^d \equiv \frac{1}{Z_1^{n_1+n_2+2}} \frac{1}{d_D^{n_1+n_2+3}} \sum_{\tau \in P_d} \langle \eta_R \otimes e_{B'} | \mathcal{I}_\alpha, \tau \rangle d_D^{k(\tau)} \quad (\text{G.4})$$

where P_c refers to the set of permutations in $S_{n_1+n_2+2}$ such that the first and n_1+2 -th element are in the same cycle, and P_d refers to the rest. The equilibrium approximation for the fidelity is then given by²¹

$$F(\rho, [\mathcal{P}_{\sigma, \mathcal{N}} \circ \mathcal{N}](\rho)) = F^c + F^d, \quad F^{c,d} = \lim_{n_1 \rightarrow -\frac{1}{2}, n_2 \rightarrow -\frac{1}{2}} F_{n_1, n_2}^{c,d} \quad (\text{G.5})$$

Note that

$$F_{n_1, n_2}^c + d_D F_{n_1, n_2}^d = \frac{1}{(Z_1 d_D)^{n_1+n_2+2}} \sum_{\tau \in S_{n_1+n_2+2}} \langle \eta_R \otimes e_{B'} \otimes e_{D'} | (\mathcal{I}_\alpha)_{B'R} \otimes \mathbf{1}_{D'}, \tau \rangle \quad (\text{G.6})$$

$$= \mathcal{Z}_{n_1+n_2+2, R} \quad (\text{G.7})$$

where we get the last expression by identifying the second-to-last expression to be the equilibrium approximation for $\mathcal{Z}_{n_1+n_2+2, R}$ for an equilibrated pure state $|\Psi\rangle$ in a Hilbert space $\mathcal{H}_{B'R} \otimes \mathcal{H}_{D'}$, with the effective identity operator $(\mathcal{I}_\alpha)_{B'R} \otimes \mathbf{1}_{D'}$, where D' is an auxiliary system with Hilbert space dimension d_D . This implies that

$$F^c + d_D F^d = \mathcal{Z}_{1, R} = 1, \quad (\text{G.8})$$

so that it is sufficient to calculate F^c and express the fidelity as

$$F(\rho, [\mathcal{P}_{\sigma, \mathcal{N}} \circ \mathcal{N}](\rho)) = \left(1 - \frac{1}{d_D}\right) F^c + \frac{1}{d_D}. \quad (\text{G.9})$$

In order to systematically obtain F^c using F_{n_1, n_2}^c , we can define the generating functional

$$F^c(\lambda_1, \lambda_2) = \sum_{n_1, n_2=0}^{\infty} \frac{F_{n_1, n_2}^c}{\lambda_1^{n_1} \lambda_2^{n_2}}. \quad (\text{G.10})$$

²¹Note that here we use the alternative analytic continuation from (2.70).

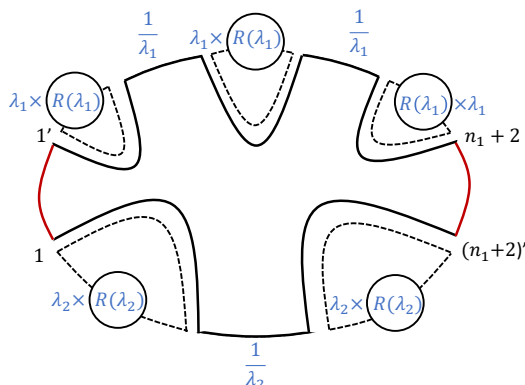


Figure 35. Contribution from a fixed $m_1 = 2$ and $m_2 = 1$ to all n_1 and n_2 in the expression for $F^c(\lambda_1, \lambda_2)$ in (G.10). The solid loop gives a factor of $\frac{d_D Z_{m_1+m_2+2, B'}}{(d_D d_R Z_{1, B'})^{m_1+m_2+2}}$, and we get $m_1, m_1 + 1, m_2$ and $m_2 + 1$ factors respectively of $\frac{1}{\lambda_1}, \lambda_1 R(\lambda_1), \frac{1}{\lambda_2}$ and $\lambda_2 R(\lambda_2)$, leading to (G.12).

Then for non-negative integers n_1, n_2 ,

$$F_{n_1, n_2}^c = \oint_{\infty} \frac{d\lambda_1}{2\pi i} \frac{d\lambda_2}{2\pi i} \lambda_1^{n_1-1} \lambda_2^{n_2-1} F^c(\lambda_1, \lambda_2) \quad (\text{G.11})$$

where the contour is taken to be around the point at infinity. Now in all permutations τ contributing to (G.10), the first and $(n_1 + 2)^{\text{th}}$ element are in a common cycle. Suppose this cycle also includes m_1 elements out of the n_1 elements between the first and $(n_1 + 2)^{\text{th}}$ element, and m_2 elements after the $(n_1 + 2)^{\text{th}}$ element. We can then consider the total contribution from a fixed m_1, m_2 to all n_1, n_2 , as shown in figure 35. Note that in the process we have introduced factors of $R(\lambda)$, the resolvent for $\rho_R \equiv \text{Tr}_{B'D'}[|\Psi\rangle\langle\Psi|]$, with $|\Psi\rangle$ as defined above (G.8).

We then find

$$F^c(\lambda_1, \lambda_2) = \sum_{m_1, m_2=0}^{\infty} \frac{d_D Z_{m_1+m_2+2, B'}}{(Z_{1, B'} d_R d_D)^{m_1+m_2+2}} R(\lambda_1)^{m_1+1} R(\lambda_2)^{m_2+1} \lambda_1 \lambda_2. \quad (\text{G.12})$$

and using (G.11),

$$F_{n_1, n_2}^c = d_D \int_0^{\infty} dE \rho(E) \left(\oint_{\infty} \frac{d\lambda}{2\pi i} \lambda^{n_1} R(\lambda) \frac{e^{-\beta E}}{Z_{1, B'} d_D d_R - \frac{e^{-\beta E} R(\lambda)}{Z_{1, B'} d_D}} \frac{1}{d_R - \frac{e^{-\beta E} R(\lambda)}{Z_{1, B'} d_D}} \right) \times \left(\oint_{\infty} \frac{d\lambda}{2\pi i} \lambda^{n_2} R(\lambda) \frac{e^{-\beta E}}{Z_{1, B'} d_D d_R - \frac{e^{-\beta E} R(\lambda)}{Z_{1, B'} d_D}} \frac{1}{d_R - \frac{e^{-\beta E} R(\lambda)}{Z_{1, B'} d_D}} \right) \quad (\text{G.13})$$

Note that $\lambda^{-\frac{1}{2}}$ has a branch point at infinity, so the above expression cannot be analytically continued as written. However, we can first deform the contour for integer n_1, n_2 to surround the branch cut on the real axis coming from $R(\lambda)$, and then analytically continue

to $n_1 = n_2 = -\frac{1}{2}$, to get

$$F^c = d_D \int_0^\infty dE \rho(E) \left(\oint_{\mathcal{C}} \frac{d\lambda}{2\pi i} \lambda^{-\frac{1}{2}} R(\lambda) \frac{e^{-\beta E}}{Z_{1,B'} d_D} \frac{1}{d_R - \frac{e^{-\beta E} R(\lambda)}{Z_{1,B'} d_D}} \right)^2 \quad (\text{G.14})$$

where \mathcal{C} is the contour that surrounds the branch cut on the real axis. In order to further simplify this expression, we use the following approximation for the resolvent:

$$R(\lambda) \approx \frac{d_R}{\lambda - \lambda_0} \quad (\text{G.15})$$

where λ_0 is the smallest eigenvalue of ρ_R , which turns out to be

$$\lambda_0 = \frac{1}{d_R} g, \quad g = \frac{1}{Z_{1,B'}} \int_{E_0}^\infty dE \rho(E) e^{-\beta E} \quad (\text{G.16})$$

with E_0 defined through

$$\int_0^{E_0} dE \rho(E) = \frac{d_R}{d_D}. \quad (\text{G.17})$$

In deriving (G.15)–(G.17), we use similar methods to [4], which we review for this context at the end of this subsection. Putting (G.15) into (G.14), we find

$$F^c \approx \frac{1}{Z_{1,B'}} \int_0^\infty dE \rho(E) \frac{e^{-2\beta E}}{d_D Z_{1,B'} \lambda_0 + e^{-\beta E}}. \quad (\text{G.18})$$

Now to further analyse (G.18), let us take the density of states to be as in AdS_3 , $\rho(E) = e^{cV\sqrt{E/V}}$, where V is the volume of B' . Then $Z_{1,B'}$ can be evaluated using the saddle point approximation and has the behavior

$$Z_{1,B'} = c' V^{\frac{1}{2}} e^{-\beta E_c + cV\sqrt{\epsilon_c}} = c' V^{\frac{1}{2}} e^{\frac{c^2 V}{4\beta}}, \quad E_c = V\epsilon_c, \quad \sqrt{\epsilon_c} = \frac{c}{2\beta}, \quad S = V \frac{c^2}{2\beta} \quad (\text{G.19})$$

where S is the thermal entropy and c' is some $O(1)$ constant. The second equilibrium Renyi entropy can also be evaluated from the saddle point

$$e^{-S_{2,B'}} = \frac{Z_{1,B'}(2\beta)}{Z_{1,B'}^2(\beta)} \sim V^{-\frac{1}{2}} e^{-\frac{3c^2 V}{8\beta}}, \quad S_{2,B'} = \frac{3c^2 V}{8\beta}. \quad (\text{G.20})$$

Now if we define

$$x \equiv \frac{\log d_R}{V}, \quad y \equiv \frac{\log d_D}{V}, \quad (\text{G.21})$$

then the various relevant time scales are given by:

$$t_b: \quad x = \frac{3c^2}{8\beta}, \quad t_p: \quad x = \frac{c^2}{2\beta}, \quad t_{p_2}: \quad x = \frac{c^2}{2\beta} + y. \quad (\text{G.22})$$

We will always take $x \sim O(1)$ while take $V \rightarrow \infty$, but y can be $O(V^{-1})$ or $O(1)$ depending on whether the diary is “small” or “large.”

Now let us examine the behavior of g defined in (G.16). Note that for $\epsilon_0 = E_0/V$,

$$V \int_0^{\epsilon_0} d\epsilon e^{cV\sqrt{\epsilon}} = \frac{2}{cV} \sqrt{\epsilon_0} e^{cV\sqrt{\epsilon_0}} = \frac{d_R}{d_D} \quad (\text{G.23})$$

$$\Rightarrow \sqrt{\epsilon_0} \approx \frac{x-y}{c} + O(V^{-1} \log V), \quad . \quad (\text{G.24})$$

so that g has the form

$$g = \frac{1}{2} \left[1 + \operatorname{erf} \left(\sqrt{V\beta} (\sqrt{\epsilon_c} - \sqrt{\epsilon_0}) \right) \right] + \sqrt{\frac{\beta}{\pi c^2 V}} e^{-V\beta(\sqrt{\epsilon_c} - \sqrt{\epsilon_0})^2} \quad (\text{G.25})$$

$$= \begin{cases} 1 + O(e^{-V}) & x - y < \frac{c^2}{2\beta} - O(V^{-\frac{1}{2}}) \\ \frac{\sqrt{\frac{\epsilon_0}{\epsilon_c}} e^{-V\beta(\sqrt{\epsilon_c} - \sqrt{\epsilon_0})^2}}{2\sqrt{\pi V\beta}(\sqrt{\epsilon_0} - \sqrt{\epsilon_c})} & x - y > \frac{c^2}{2\beta} + O(V^{-\frac{1}{2}}) \end{cases} . \quad (\text{G.26})$$

During the evaporation process, $\log d_R$ increases while V decreases. We would like to find the evolution of F^c as a function of x to see how the information recovery improves during the evaporation process.

Note first that from (G.26), for $x - y > \frac{c^2}{2\beta} + O(V^{-\frac{1}{2}})$, the first term in the denominator of the integrand of (G.18) can always be ignored relative to the second, implying that the $F^c \approx 1$ and hence $F \approx 1$ at t_{p_2} and later times. For $t < t_{p_2}$, $\lambda_0 \approx \frac{1}{d_R}$, and we can simplify (G.18) in this regime to

$$F^c \approx \sqrt{V} e^{-V\frac{c^2}{4\beta}} \int_0^\infty d\epsilon e^{cV\sqrt{\epsilon}} \frac{e^{-2\beta V\epsilon}}{e^{V(\frac{c^2}{4\beta} - x + y)} + e^{-\beta V\epsilon}} . \quad (\text{G.27})$$

The second term in the denominator dominates for

$$\epsilon < \epsilon_u \equiv -\frac{c^2}{4\beta^2} + \frac{x-y}{\beta} \quad (\text{G.28})$$

and the first term dominates otherwise. (Note that the regime (G.28) exists only for $x - y \geq \frac{c^2}{4\beta}$.) So we can write (ignoring sub-exponential prefactors)

$$F^c \sim e^{-Vc^2/(4\beta)} \int_0^{\epsilon_u} e^{cV\sqrt{\epsilon}} e^{-\beta V\epsilon} + e^{-Vc^2/(4\beta)} e^{\beta V\epsilon_u} \int_{\epsilon_u}^\infty e^{cV\sqrt{\epsilon}} e^{-2\beta V\epsilon} \quad (\text{G.29})$$

Recall that the saddle point for the first integral is at $\epsilon_c = \frac{c^2}{4\beta^2}$, and the saddle point for the second integral is at $\epsilon_c^{(2)} = \frac{c^2}{16\beta^2}$. We consider three regimes:

1. $x - y < \frac{c^2}{4\beta}$: we only have the second term in (G.29), which can be approximated with its saddle point, so

$$F^c \sim e^{V(-\frac{3c^2}{8\beta} + x - y)}, \quad x - y < \frac{c^2}{4\beta} \quad (\text{G.30})$$

2. $\frac{c^2}{4\beta} < x - y < \frac{5c^2}{16\beta}$: here the first term should be approximated with the boundary value at ϵ_u , and the second term with the saddle point at $\epsilon_c^{(2)}$. So we find

$$F^c \sim e^{V(c\sqrt{-\frac{c^2}{4\beta^2} + \frac{x-y}{\beta}} - x + y)} + e^{V(-\frac{3c^2}{8\beta} + x - y)} \quad (\text{G.31})$$

The second term is always greater than the first in this regime, so we have

$$F^c \approx e^{V(-\frac{3c^2}{8\beta} + x - y)}, \quad \frac{c^2}{4\beta} < x - y < \frac{5c^2}{16\beta}. \quad (\text{G.32})$$

3. $\frac{5c^2}{16\beta} < x - y < \frac{c^2}{2\beta}$: in this regime, both terms are approximated with their boundary values, which turn out to be the same, so we have

$$F^c \approx 2e^{V(c\sqrt{-\frac{c^2}{4\beta^2} + \frac{x-y}{\beta}} - x + y)}, \quad \frac{5c^2}{16\beta} < x - y < \frac{c^2}{2\beta} \quad (\text{G.33})$$

For a small diary, we can neglect y in (G.30)–(G.33), and we therefore find that $(1 - \frac{1}{d_D})F^c$ is exponentially suppressed in volume before t_p , while $\frac{1}{d_D}$ is $O(1)$. We therefore have

$$F(\rho, [\mathcal{P}_{\sigma, \mathcal{N}} \circ \mathcal{N}](\rho)) \approx \begin{cases} \frac{1}{d_D} & x < \frac{c^2}{2\beta} \\ 1 & x > \frac{c^2}{2\beta} \end{cases} \quad (\text{G.34})$$

So the fidelity for the recovery of a small diary improves rapidly from $\frac{1}{d_D}$ to 1 at the Page time.

For a diary sufficiently large such that there is a regime where $\frac{3c^2}{8\beta} < x < \frac{5c^2}{16\beta} + y$, we have

$$F(\rho, [\mathcal{P}_{\sigma, \mathcal{N}} \circ \mathcal{N}](\rho)) = \begin{cases} \frac{1}{d_D} & x < \frac{3c^2}{8\beta} \\ e^{V(-\frac{3c^2}{8\beta} + x - y)} = e^{\log d_R - S_{2, B'} - \log d_D} & \frac{3c^2}{8\beta} < x < \frac{5c^2}{16\beta} + y \\ e^{V(c\sqrt{-\frac{c^2}{4\beta^2} + \frac{x-y}{\beta}} - x + y)} & \frac{5c^2}{16\beta} + y < x < \frac{c^2}{2\beta} + y \\ 1 & x > \frac{c^2}{2\beta} + y \end{cases} \quad (\text{G.35})$$

So in this case, the fidelity starts increasing from its initial value of $\frac{1}{d_D}$ at time t_b (independently of $\log d_D$ as long as there is a non-trivial regime corresponding to the second line), and the initial improvement is exponential in $\log d_R - S_{2, B'}$, precisely as predicted by the leading correction in (6.44).

For a diary that is $O(1)$ but not sufficiently large such that there is a regime where $\frac{3c^2}{8\beta} < x < \frac{5c^2}{16\beta} + y$, we have

$$F(\rho, [\mathcal{P}_{\sigma, \mathcal{N}} \circ \mathcal{N}](\rho)) = \begin{cases} \frac{1}{d_D} & x < \frac{c^2}{2\beta} + 2y - c\sqrt{\frac{y}{\beta}} \\ e^{V(c\sqrt{-\frac{c^2}{4\beta^2} + \frac{x-y}{\beta}} - x + y)} & \frac{c^2}{2\beta} + 2y - c\sqrt{\frac{y}{\beta}} < x < \frac{c^2}{2\beta} + y \\ 1 & x > \frac{c^2}{2\beta} + y \end{cases} \quad (\text{G.36})$$

So in this case, the fidelity starts to increase from $\frac{1}{d_D}$ at a time

$$t_r : \quad x = \frac{c^2}{2\beta} + 2y - c\sqrt{\frac{y}{\beta}} \quad (\text{G.37})$$

which is macroscopically earlier than t_p when y is $O(1)$, and becomes increasingly earlier as the diary becomes larger.

Further, note that in (G.29), we can evaluate both integrals exactly using error functions. The first integral is approximately

$$F^c \approx \frac{\sqrt{\pi}c \left(1 - \operatorname{erf} \left(\frac{\sqrt{V}(c - \sqrt{4\beta(x-y) - c^2})}{2\sqrt{\beta}} \right) \right)}{2\beta^{3/2}}. \quad (\text{G.38})$$

Near t_{p_2} , this term dominates over the second integral. This shows that the transition regime is $O((V\beta)^{-\frac{1}{2}})$ in $x - y$, as this is the regime where the error function becomes $O(1)$.

We can also use (G.17) and (G.16) to get an exact expression for λ_0 in terms of d_R, d_D, V , which can then be used to evaluate the integral (G.18) for large values of V numerically. This leads to the evolution of the fidelity shown in figure 18.

Let us now justify the approximation we used for $R(\lambda)$ in (G.15). We use similar steps to section 2 of [4]. The Schwinger-Dyson equation for $R(\lambda)$ is

$$\lambda R = d_R + \sum_{n=1}^{\infty} d_D Z_{n,B'} \frac{R^n}{d_R^n Z_{1,B'}^n d_D^n} \quad (\text{G.39})$$

$$\Rightarrow \lambda R = d_R + d_D R \int_0^{\infty} dE \rho(E) \frac{e^{-\beta E}}{d_R d_D Z_{1,B'} - e^{-\beta E} R} \quad (\text{G.40})$$

It is useful to first find the smallest eigenvalue λ_0 of ρ_R . For $\lambda < \lambda_0$, we can see from the definition of R that R is real and negative, and we can write

$$\lambda \approx \frac{d_R}{R} - \frac{d_D}{R} \int_0^{E'_0} dE \rho(E) + d_D \int_{E'_0}^{\infty} dE \rho(E) \frac{e^{-\beta E}}{d_R d_D Z_{1,B'}} \quad (\text{G.41})$$

where E'_0 for a given set of $d_R, d_D, V_{B'}$ and λ is defined implicitly by

$$-e^{-\beta E'_0} = \frac{d_R d_D Z_{1,B'}}{R(\lambda)}. \quad (\text{G.42})$$

In the two terms in (G.41), we have assumed that different terms dominate in the denominator of the integrand of (G.40). Since $\lambda = \lambda_0$ is a branch point of $R(\lambda)$, we have $\frac{d\lambda}{dR} = 0$ at this point, which implies

$$\int_0^{E_0} dE \rho(E) \approx \frac{d_R}{d_D} \quad (\text{G.43})$$

where we use E_0 to refer to the value of E'_0 at λ_0 . We can treat (G.17) as the definition of E_0 . In terms of E_0 , λ_0 is given by

$$\lambda_0 = \frac{1}{d_R Z_{1,B'}} \int_{E_0}^{\infty} dE \rho(E) e^{-\beta E} \quad (\text{G.44})$$

Let us now return to (G.40) to obtain an approximation for the resolvent. We can divide the integral into two parts at E_0 :

$$\begin{aligned} \lambda R &= d_R + d_D R \int_0^{E_0} dE \rho(E) \frac{e^{-\beta E}}{d_R d_D Z_{1,B'} - e^{-\beta E} R} \\ &+ d_D R \int_{E_0}^{\infty} dE \rho(E) e^{-\beta E} \frac{e^{-\beta E}}{d_R d_D Z_{1,B'} - e^{-\beta E} R} \end{aligned} \quad (\text{G.45})$$

Note that $e^{-\beta E_0} R(\lambda_0) = -d_R d_D Z_{1,B'}$, and $e^{-\beta E} |R(\lambda)| < e^{-\beta E_0} |R(\lambda_0)|$ for $E > E_0$ and $\lambda > \lambda_0$, using the definition of R . Hence, the denominator in the second term of the integrand can be approximated as $d_R d_D Z_{1,B'}$, and we have

$$\begin{aligned} \lambda R &\approx d_R + d_D R \int_0^{E_0} dE \rho(E) \frac{e^{-\beta E}}{d_R d_D Z_{1,B'} - e^{-\beta E} R} \\ &\quad + \frac{1}{d_R Z_{1,B'}} R \int_{E_0}^{\infty} dE \rho(E) e^{-\beta E} \\ \Rightarrow \lambda R &\approx d_R + d_D R \int_0^{E_0} dE \rho(E) \frac{e^{-\beta E}}{d_R d_D Z_{1,B'} - e^{-\beta E} R} + R \lambda_0 \\ \Rightarrow R &= \frac{d_R}{\lambda - \lambda_0} + \frac{d_D}{\lambda - \lambda_0} R \int_0^{E_0} dE \rho(E) \frac{e^{-\beta E}}{d_R d_D Z_{1,B'} - e^{-\beta E} R} \end{aligned} \tag{G.46}$$

Treating the second term as a perturbation, we find to first order

$$R \approx \frac{d_R}{\lambda - \lambda_0} + \frac{d_D}{\lambda - \lambda_0} \int_0^{E_0} dE \rho(E) \frac{e^{-\beta E}}{Z_{1,B'} d_D} \frac{1}{\lambda - \lambda_0 - e^{-\beta E} / (Z_{1,B'} d_D)} \tag{G.47}$$

Let us now understand whether our approximation in treating the second term as a perturbation is self-consistent. Suppose we require that for some small ϵ ,

$$\lambda - \lambda_0 - \frac{e^{-\beta \epsilon}}{Z_{1,B'} d_D} > 0. \tag{G.48}$$

The integral in the second term can be divided into two parts,

$$\begin{aligned} I &= \frac{d_D}{\lambda - \lambda_0} \int_0^{\epsilon} dE \rho(E) \frac{e^{-\beta E}}{Z_{1,B'} d_D} \frac{1}{\lambda - \lambda_0 - e^{-\beta E} / (Z_{1,B'} d_D)} \\ &\quad + d_D \int_{\epsilon}^{E_0} dE \rho(E) \frac{e^{-\beta E} / (Z_{1,B'} d_D)}{\lambda - \lambda_0} \frac{1}{\lambda - \lambda_0 - e^{-\beta E} / (Z_{1,B'} d_D)} \end{aligned} \tag{G.49}$$

If ϵ is small, the first term is small. The integrand in the second term is positive and $< \rho(E)$ due to (G.48), so from (G.17), the second term is $< d_R$, and hence smaller than the leading term in (G.47). So the approximation (G.47) is valid as long as (G.48) is satisfied for small ϵ , and to a first approximation, we can ignore the second term in (G.47) and write $R(\lambda)$ as in (G.15).

G.2 Microcanonical ensemble

Let us now consider the question of Petz map fidelity taking the effective identity operator to be as in (4.87), with $A_1 = R_1$ and $A_2 = R_2$ corresponding to two parts of the radiation, and B to the black hole. We again take ρ_D to be pure and σ_D to be maximally mixed. R_2 is taken to be at infinite temperature while there is energy conservation between R_1 and B . We take the average energy density ϵ in $R_1 B$ to be

$$\epsilon = \frac{E}{V_{R_1} + V_B} \tag{G.50}$$

and view the evaporation process as a process where $S_R^{(\text{eq})} = V_{R_1}s(\epsilon) + \log d_{R_2}$ increases while $S_B^{(\text{eq})} = V_B s(\epsilon)$ decreases. t_p is the time when

$$t_p : \quad V_{R_1}s(\epsilon) + \log d_{R_2} = V_B s(\epsilon) \quad (\text{G.51})$$

and t_{p_2} is defined as the time when

$$t_{p_2} : \quad V_{R_1}s(\epsilon) + \log d_{R_2} = V_B s(\epsilon) + \log d_D. \quad (\text{G.52})$$

Recall that in this setup, the time t_b at which the logarithmic negativity starts to grow is given for sufficiently small $\lambda \equiv \frac{\log d_{R_2}}{V_{R_1}s(\epsilon) + \log d_{R_2}}$ by

$$t_b : \quad \log d_{R_2} + 2S_{\frac{1}{3}, R_1}^{(\text{eq})} - V_{R_1}s(\epsilon) = V_B s(\epsilon). \quad (\text{G.53})$$

The equilibrium approximation for the Petz map fidelity is given by

$$F(\rho, [\mathcal{P}_{\sigma, \mathcal{N}} \circ \mathcal{N}](\rho)) = \lim_{m \rightarrow \frac{1}{2}} F_m^E, \quad (\text{G.54})$$

where

$$\begin{aligned} F_m^E &\equiv \sum_{E_1} \frac{1}{d_D^{2m+3} (N_E d_{R_2})^{2m+2}} \sum_{\tau \in \mathcal{S}_{2m+2}} (d_{E_1}^{R_1} d_{R_2})^{k(\eta^{-1}\tau)} (d_{E-E_1}^{B'})^{k(\tau)} d_D^{k(\tau) + \zeta(\tau)} \\ &= \sum_{E_1} \left(\frac{d_{E_1}^{R_1} d_{E-E_1}^{B'}}{N_E} \right)^{2m+2} F_m^\infty(d_D, d_{E_1}^{R_1} d_{R_2}, d_{E-E_1}^{B'}) \end{aligned} \quad (\text{G.55})$$

with

$$N_E \equiv \sum_{E_1} d_{E_1}^{R_1} d_{E-E_1}^{B'} \approx c_1 e^{(V_1 + V_B)s(\epsilon)} \quad (\text{G.56})$$

where c_1 and all c_i we will introduce below are $O(1)$ constants, and

$$F_m^\infty(d_D, d_R, d_{B'}) \equiv \frac{1}{d_D^{2m+3} (d_R d_B)^{2m+2}} \sum_{\tau \in \mathcal{S}_{2m+2}} d_R^{k(\eta^{-1}\tau)} d_{B'}^{k(\tau)} d_D^{k(\tau) + \zeta(\tau)} \quad (\text{G.57})$$

can be identified to be the infinite-temperature value of F_m in (6.22). Then using (G.54) and (6.40),

$$F(\rho, [\mathcal{P}_{\sigma, \mathcal{N}} \circ \mathcal{N}](\rho)) = \frac{1}{d_D} + \left(1 - \frac{1}{d_D}\right) (F_1 + F_2) \quad (\text{G.58})$$

where

$$F_1 = \frac{1}{d_D} \frac{d_{R_2}}{N_E} \sum_{E_1 < V_{R_1} \epsilon_u} (d_{E_1}^{R_1})^2 {}_2F_1 \left(\frac{1}{2}, -\frac{1}{2}, 2, \frac{d_{R_2} d_{E_1}^{R_1}}{d_{E-E_1}^{B'} d_D} \right)^2, \quad (\text{G.59})$$

$$F_2 = \sum_{E_1 > V_{R_1} \epsilon_u} \frac{d_{E_1}^{R_1} d_{E-E_1}^{B'}}{N_E} {}_2F_1 \left(\frac{1}{2}, -\frac{1}{2}, 2, \frac{d_{E-E_1}^{B'} d_D}{d_{R_2} d_{E_1}^{R_1}} \right)^2 \quad (\text{G.60})$$

and ϵ_u is defined as the solution to

$$V_{R_1} s(\epsilon_u) + \log d_{R_2} = \log d_D + V_B s(\bar{\epsilon}_u), \quad \bar{\epsilon}_u = \frac{E - V_{R_1} \epsilon_u}{V_B}. \quad (\text{G.61})$$

We note that F_1 is always dominated by the value at upper limit of the sum in (G.59),

$$F_1 = c_2 e^{\Lambda_1}, \quad \Lambda_1 = 2V_{R_1} s(\epsilon_u) - (V_{R_1} + V_B) s(\epsilon) + \log d_{R_2} - \log d_D = \Lambda_2, \quad (\text{G.62})$$

$$\Lambda_2 = V_{R_1} (s(\epsilon_u) - s(\epsilon)) + V_B (s(\bar{\epsilon}_u) - s(\epsilon)) \quad (\text{G.63})$$

and F_2 is dominated by either its saddle point or its lower limit, i.e.

$$F_2 = \begin{cases} 1 & \epsilon_u < \epsilon \\ c_3 e^{\Lambda_2} & \epsilon_u > \epsilon \end{cases}. \quad (\text{G.64})$$

Note that by definition $\Lambda_2 \leq 0$, and thus F_1 is always exponentially suppressed except at $\epsilon_u = \epsilon$.

From (G.61) when we decrease the values of V_1 and $\log d_{R_2}$ (i.e. going to earlier times), the value of ϵ_u should increase. We can assume that ϵ_u is a smooth and monotonically decreasing function of t . At t_{p_2} we have $\epsilon_u = \epsilon$ while for $t < t_{p_2}$ ($t > t_{p_2}$) we have $\epsilon_u > \epsilon$ ($\epsilon_u < \epsilon$). From (G.62)–(G.64), we then conclude that for any d_D , $F_1 + F_2$ is exponentially small in volume for $t < t_{p_2}$ and becomes 1 at t_{p_2} over a very short range of time (volume suppressed).

Then for a small diary, precisely as in (G.34) for the canonical example, we have

$$F(\rho, [\mathcal{P}_{\sigma, \mathcal{N}} \circ \mathcal{N}](\rho)) = \begin{cases} \frac{1}{d_D} & t < t_p \\ 1 & t > t_p \end{cases}. \quad (\text{G.65})$$

For a large diary, we have

$$F(\rho, [\mathcal{P}_{\sigma, \mathcal{N}} \circ \mathcal{N}](\rho)) = \begin{cases} \frac{1}{d_D} & t < t_r \\ \frac{1}{d_D} e^{V_{R_1} (2s(\epsilon_u) - s(\epsilon)) + \log d_{R_2} - V_B s(\epsilon)} & t_r < t < t_{p_2} \\ = e^{V_{R_1} (s(\epsilon_u) - s(\epsilon)) + V_B (s(\bar{\epsilon}_u) - s(\epsilon))} & t_r < t < t_{p_2} \\ 1 & t > t_{p_2} \end{cases}. \quad (\text{G.66})$$

The growth of the fidelity above $1/d_D$ for a large diary starts at the time scale

$$t_r : \quad V_{R_1} (2s(\epsilon_u) - s(\epsilon)) + V_{R_2} s_0 = V_B s(\epsilon) \quad (\text{G.67})$$

Comparing this to t_p defined in (G.51) and noting that $\epsilon_u > \epsilon$ for $t < t_{p_2}$, we can see that $t_r < t_p$. Note that unlike in the canonical ensemble example in the previous subsection, t_r does not converge to t_b defined by (G.53) for sufficiently large diaries. For a given λ , depending on $\log d_D$, t_r can be either earlier or later than t_b .

Open Access. This article is distributed under the terms of the Creative Commons Attribution License ([CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited. SCOAP³ supports the goals of the International Year of Basic Sciences for Sustainable Development.

References

- [1] H. Liu and S. Vardhan, *Entanglement entropies of equilibrated pure states in quantum many-body systems and gravity*, *PRX Quantum* **2** (2021) 010344 [[arXiv:2008.01089](#)] [[INSPIRE](#)].
- [2] G. Penington, *Entanglement wedge reconstruction and the information paradox*, *JHEP* **09** (2020) 002 [[arXiv:1905.08255](#)] [[INSPIRE](#)].
- [3] A. Almheiri, N. Engelhardt, D. Marolf and H. Maxfield, *The entropy of bulk quantum fields and the entanglement wedge of an evaporating black hole*, *JHEP* **12** (2019) 063 [[arXiv:1905.08762](#)] [[INSPIRE](#)].
- [4] G. Penington, S.H. Shenker, D. Stanford and Z. Yang, *Replica wormholes and the black hole interior*, *JHEP* **03** (2022) 205 [[arXiv:1911.11977](#)] [[INSPIRE](#)].
- [5] A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian and A. Tajdini, *Replica wormholes and the entropy of Hawking radiation*, *JHEP* **05** (2020) 013 [[arXiv:1911.12333](#)] [[INSPIRE](#)].
- [6] S. Vardhan, J. Kudler-Flam, H. Shapourian and H. Liu, *Bound entanglement in thermalized states and black hole radiation*, *Phys. Rev. Lett.* **129** (2022) 061602 [[arXiv:2110.02959](#)] [[INSPIRE](#)].
- [7] K. Audenaert, M.B. Plenio and J. Eisert, *Entanglement cost under positive-partial-transpose-preserving operations*, *Phys. Rev. Lett.* **90** (2003) 027901.
- [8] M. Christandl and A. Winter, “Squashed entanglement”: an additive entanglement measure, *J. Math. Phys.* **45** (2004) 829.
- [9] H. Shapourian, S. Liu, J. Kudler-Flam and A. Vishwanath, *Entanglement negativity spectrum of random mixed states: a diagrammatic approach*, *PRX Quantum* **2** (2021) 030347 [[arXiv:2011.01277](#)] [[INSPIRE](#)].
- [10] P. Hayden and J. Preskill, *Black holes as mirrors: quantum information in random subsystems*, *JHEP* **09** (2007) 120 [[arXiv:0708.4025](#)] [[INSPIRE](#)].
- [11] P. Hayden and G. Penington, *Learning the alpha-bits of black holes*, *JHEP* **12** (2019) 007 [[arXiv:1807.06041](#)] [[INSPIRE](#)].
- [12] R.F. Werner, *Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model*, *Phys. Rev. A* **40** (1989) 4277 [[INSPIRE](#)].
- [13] L. Gurvits, *Classical complexity and quantum entanglement*, *J. Comput. Syst. Sci.* **69** (2004) 448.
- [14] M. Horodecki, P. Horodecki and R. Horodecki, *On the necessary and sufficient conditions for separability of mixed quantum states*, *Phys. Lett. A* **223** (1996) 1 [[quant-ph/9605038](#)] [[INSPIRE](#)].
- [15] K. Życzkowski, P. Horodecki, A. Sanpera and M. Lewenstein, *On the volume of the set of mixed entangled states*, *Phys. Rev. A* **58** (1998) 883 [[quant-ph/9804024](#)] [[INSPIRE](#)].
- [16] A. Peres, *Separability criterion for density matrices*, *Phys. Rev. Lett.* **77** (1996) 1413 [[quant-ph/9604005](#)] [[INSPIRE](#)].
- [17] J. Eisert and M.B. Plenio, *A comparison of entanglement measures*, *J. Mod. Opt.* **46** (1999) 145 [[quant-ph/9807034](#)] [[INSPIRE](#)].

- [18] R. Simon, *Peres-Horodecki separability criterion for continuous variable systems*, *Phys. Rev. Lett.* **84** (2000) 2726 [[quant-ph/9909044](#)] [[INSPIRE](#)].
- [19] G. Vidal and R.F. Werner, *Computable measure of entanglement*, *Phys. Rev. A* **65** (2002) 032314 [[quant-ph/0102117](#)] [[INSPIRE](#)].
- [20] M.B. Plenio, *Logarithmic negativity: a full entanglement monotone that is not convex*, *Phys. Rev. Lett.* **95** (2005) 090503 [[quant-ph/0505071](#)] [[INSPIRE](#)].
- [21] C.H. Bennett, D.P. DiVincenzo, J.A. Smolin and W.K. Wootters, *Mixed state entanglement and quantum error correction*, *Phys. Rev. A* **54** (1996) 3824 [[quant-ph/9604024](#)] [[INSPIRE](#)].
- [22] P.M. Hayden, M. Horodecki and B.M. Terhal, *The asymptotic entanglement cost of preparing a quantum state*, *J. Phys. A* **34** (2001) 6891.
- [23] C.H. Bennett, H.J. Bernstein, S. Popescu and B. Schumacher, *Concentrating partial entanglement by local operations*, *Phys. Rev. A* **53** (1996) 2046 [[quant-ph/9511030](#)] [[INSPIRE](#)].
- [24] M. Horodecki, P. Horodecki and R. Horodecki, *Mixed state entanglement and distillation: is there a ‘bound’ entanglement in nature?*, *Phys. Rev. Lett.* **80** (1998) 5239 [[quant-ph/9801069](#)] [[INSPIRE](#)].
- [25] T. Eggeling, K. Vollbrecht, R. Werner and M. Wolf, *Distillability via protocols respecting the positivity of partial transpose*, *Phys. Rev. Lett.* **87** (2001) 257902.
- [26] X. Wang and R. Duan, *Irreversibility of asymptotic entanglement manipulation under quantum operations completely preserving positivity of partial transpose*, *Phys. Rev. Lett.* **119** (2017) 180506.
- [27] A. Elben et al., *Mixed-state entanglement from local randomized measurements*, *Phys. Rev. Lett.* **125** (2020) 200501 [[arXiv:2007.06305](#)] [[INSPIRE](#)].
- [28] C.-M. Chung, V. Alba, L. Bonnes, P. Chen and A.M. Läuchli, *Entanglement negativity via the replica trick: a quantum Monte Carlo approach*, *Phys. Rev. B* **90** (2014) 064401.
- [29] J. Gray, L. Banchi, A. Bayat and S. Bose, *Machine-learning-assisted many-body entanglement measurement*, *Phys. Rev. Lett.* **121** (2018) 150503.
- [30] K.-H. Wu, T.-C. Lu, C.-M. Chung, Y.-J. Kao and T. Grover, *Entanglement Renyi negativity across a finite temperature transition: a Monte Carlo study*, *Phys. Rev. Lett.* **125** (2020) 140603 [[arXiv:1912.03313](#)] [[INSPIRE](#)].
- [31] T.-C. Lu, T.H. Hsieh and T. Grover, *Detecting topological order at finite temperature using entanglement negativity*, *Phys. Rev. Lett.* **125** (2020) 116801 [[arXiv:1912.04293](#)] [[INSPIRE](#)].
- [32] E. Wybo, M. Knap and F. Pollmann, *Entanglement dynamics of a many-body localized system coupled to a bath*, *Phys. Rev. B* **102** (2020) 064304 [[arXiv:2004.13072](#)] [[INSPIRE](#)].
- [33] A. Elben et al., *Mixed-state entanglement from local randomized measurements*, *Phys. Rev. Lett.* **125** (2020) 200501 [[arXiv:2007.06305](#)] [[INSPIRE](#)].
- [34] T.-C. Lu and T. Grover, *Entanglement transitions as a probe of quasiparticles and quantum thermalization*, *Phys. Rev. B* **102** (2020) 235110 [[arXiv:2008.11727](#)] [[INSPIRE](#)].
- [35] E. Wybo, M. Knap and F. Pollmann, *Dynamics of negativity of a Wannier-Stark many-body localized system coupled to a bath*, [arXiv:2110.07384](#).
- [36] D. Petz, *Sufficient subalgebras and the relative entropy of states of a von Neumann algebra*, *Commun. Math. Phys.* **105** (1986) 123 [[INSPIRE](#)].

- [37] D. Petz, *Sufficiency of channels over von Neumann algebras*, *Quart. J. Math. Oxford Ser.* **39** (1988) 97 [INSPIRE].
- [38] D. Petz, *Monotonicity of quantum relative entropy revisited*, *Rev. Math. Phys.* **15** (2003) 79 [quant-ph/0209053] [INSPIRE].
- [39] M. Junge, R. Renner, D. Sutter, M.M. Wilde and A. Winter, *Universal recovery maps and approximate sufficiency of quantum relative entropy*, *Ann. Henri Poincaré* **19** (2018) 2955 [arXiv:1509.07127] [INSPIRE].
- [40] T. Faulkner, S. Hollands, B. Swingle and Y. Wang, *Approximate recovery and relative entropy I: general von Neumann subalgebras*, *Commun. Math. Phys.* **389** (2022) 349 [arXiv:2006.08002] [INSPIRE].
- [41] S. Dutta and T. Faulkner, *A canonical purification for the entanglement wedge cross-section*, *JHEP* **03** (2021) 178 [arXiv:1905.00577] [INSPIRE].
- [42] C. Akers, T. Faulkner, S. Lin and P. Rath, *Reflected entropy in random tensor networks*, to appear.
- [43] C. Akers, T. Faulkner, S. Lin and P. Rath, *The Page curve for reflected entropy*, *JHEP* **06** (2022) 089 [arXiv:2201.11730] [INSPIRE].
- [44] S.X. Cui, P. Hayden, T. He, M. Headrick, B. Stoica and M. Walter, *Bit threads and holographic monogamy*, *Commun. Math. Phys.* **376** (2019) 609 [arXiv:1808.05234] [INSPIRE].
- [45] C. Akers and P. Rath, *Entanglement wedge cross sections require tripartite entanglement*, *JHEP* **04** (2020) 208 [arXiv:1911.07852] [INSPIRE].
- [46] X. Dong, X.-L. Qi and M. Walter, *Holographic entanglement negativity and replica symmetry breaking*, *JHEP* **06** (2021) 024 [arXiv:2101.11029] [INSPIRE].
- [47] M.C. Bañuls, J.I. Cirac and M.B. Hastings, *Strong and weak thermalization of infinite nonintegrable quantum systems*, *Phys. Rev. Lett.* **106** (2011) 050405.
- [48] S. Sugiura and A. Shimizu, *Canonical thermal pure quantum state*, *Phys. Rev. Lett.* **111** (2013) 010401 [arXiv:1302.3138] [INSPIRE].
- [49] B. Collins and I. Nechita, *Gaussianization and eigenvalue statistics for random quantum channels (III)*, arXiv:0910.1768.
- [50] P. Hayden, D.W. Leung and A. Winter, *Aspects of generic entanglement*, *Commun. Math. Phys.* **265** (2006) 95.
- [51] D.N. Page, *Average entropy of a subsystem*, *Phys. Rev. Lett.* **71** (1993) 1291 [gr-qc/9305007] [INSPIRE].
- [52] D.N. Page, *Time dependence of Hawking radiation entropy*, *JCAP* **09** (2013) 028 [arXiv:1301.4995] [INSPIRE].
- [53] M.A. Nielsen and D. Poulin, *Algebraic and information-theoretic conditions for operator quantum error correction*, *Phys. Rev. A* **75** (2007) 064304.
- [54] H. Liu and S. Vardhan, *A dynamical mechanism for the Page curve from quantum chaos*, *JHEP* **03** (2021) 088 [arXiv:2002.05734] [INSPIRE].
- [55] J. Kudler-Flam, V. Narovlansky and S. Ryu, *Negativity spectra in random tensor networks and holography*, *JHEP* **02** (2022) 076 [arXiv:2109.02649] [INSPIRE].

- [56] X. Dong, S. McBride and W.W. Weng, *Replica wormholes and holographic entanglement negativity*, *JHEP* **06** (2022) 094 [[arXiv:2110.11947](#)] [[INSPIRE](#)].
- [57] A. Almheiri, D. Marolf, J. Polchinski and J. Sully, *Black holes: complementarity or firewalls?*, *JHEP* **02** (2013) 062 [[arXiv:1207.3123](#)] [[INSPIRE](#)].
- [58] K. Hejazi and H. Shapourian, *Symmetry-protected entanglement in random mixed states*, *Phys. Rev. A* **106** (2022) 052428 [[arXiv:2112.00032](#)] [[INSPIRE](#)].
- [59] C. Murthy and M. Srednicki, *Structure of chaotic eigenstates and their entanglement entropy*, *Phys. Rev. E* **100** (2019) 022131 [[arXiv:1906.04295](#)] [[INSPIRE](#)].
- [60] X. Dong and H. Wang, *Enhanced corrections near holographic entanglement transitions: a chaotic case study*, *JHEP* **11** (2020) 007 [[arXiv:2006.10051](#)] [[INSPIRE](#)].