Published for SISSA by Springer

Received: *February 5, 2024*
Accepted: *April 25, 2024*
Published: *May 31, 2024*

Space-time generalization of mutual information

Paolo Glorioso,[a,b] **Xiao-Liang Qi**[a] **and Zhenbin Yang**[a,c]

[a]*Stanford Institute for Theoretical Physics, Stanford University 450 Jane Stanford Way, Stanford, CA 94305, U.S.A.*
[b]*Department of Physics & Condensed Matter Theory Center, University of Maryland, College Park, MD 20742, U.S.A.*
[c]*Institute for Advanced Study, Tsinghua University, 30 Shuangqing Rd, Haidian District, Beijing, 100190, China*

E-mail: paolomorphic@gmail.com, xlqi@stanford.edu, zhenbiny@tsinghua.edu.cn

Abstract: The mutual information characterizes correlations between spatially separated regions of a system. Yet, in experiments we often measure dynamical correlations, which involve probing operators that are also separated in time. Here, we introduce a space-time generalization of mutual information which, by construction, satisfies several natural properties of the mutual information and at the same time characterizes correlations across subsystems that are separated in time. In particular, this quantity, that we call the *space-time mutual information*, bounds all dynamical correlations. We construct this quantity based on the idea of the quantum hypothesis testing. As a by-product, our definition provides a transparent interpretation in terms of an experimentally accessible setup. We draw connections with other notions in quantum information theory, such as quantum channel discrimination. Finally, we study the behavior of the space-time mutual information in several settings and contrast its long-time behavior in many-body localizing and thermalizing systems.

Keywords: Lattice Quantum Field Theory, Non-Equilibrium Field Theory, Quantum Dissipative Systems

ArXiv ePrint: 2401.02475

https://doi.org/10.1007/JHEP05(2024)338

JHEP05(2024)338

Contents

1 Introduction

A prominent experimental diagnostic for the properties of physical systems are correlation functions. These offer a window into the complex dependencies that exist among different subregions in space and time of a given system, and provide an organized and transparent way to characterize its fundamental properties. For example, the electric conductivity is determined by the current-current response function. Angle-resolved photo emission experiments measure the two-point function of the electron as a function of momentum and energy. Phases of matter are also characterized by correlation functions: for example, spontaneously broken symmetry phases, such as ferromagnetism and crystalline order, correspond to long-range correlations in space. In general, a two-point correlation function is defined for two operators each residing in a region of spacetime.

JHEP05(2024)338

One question that naturally arises is whether there exists a quantum information measure to quantify the amount of correlations between two space-time regions. For two spatial regions, A and B, defined at the same time, mutual information $I(A:B) = S_A + S_B - S_{AB}$ is a natural measure, where S_A represents the von Neumann entropy of subsystem A. $I(A:B) = 0$ if and only if the states of AB factorize as $\rho_{AB} = \rho_A \otimes \rho_B$, which implies that there is no connected correlation between the two regions. A non-zero $I(A:B)$ provides a quantitative measure of correlations that is independent of the specific operators that are correlated. For instance, if we consider two ferromagnetic states, one with long-range correlations of the spin-z component $\langle S_z(r) S_z(r') \rangle_c$ and another with an equal amount of correlation in S_x, then, keeping everything else fixed, both states will yield the same mutual information. Importantly, the mutual information provides an upper bound on the connected correlation functions between the two regions [1].

In physical experiments, dynamical correlation functions can be measured for regions located at different times, just like equal-time correlators. However, using mutual information to measure correlations between different times is not feasible. In quantum field theory, the mutual information can typically be defined for algebras associated with space-like separated regions A and B that are not adjacent, but it is likely undefined for time-like separated regions, particularly when region B is inside the future lightcone of region A. For finite dimensional quantum systems, the issue arises because the state ρ_{AB} is typically not defined for time-like separated regions. Our study is motivated by the idea that a physical system should be characterized by observables. Thus, we aim to generalize mutual information for a pair of space-time regions A and B that is valid even when they are not space-like separated, and hence when ρ_{AB} is not defined. This is the primary objective of our research.

In this paper, we introduce a novel quantity, called the space-time mutual information (STMI) denoted as $J(A:B)$, which generalizes the mutual information for two arbitrary space-time regions A and B. Our approach is based on the idea of hypothesis testing where, in a gedanken experiment, an experimentalist has access to a physical system at regions A and B defined at separate times. For example, in a qubit chain with qubits labelled by $x = 1, 2, \ldots, N$, region A could be qubit $x = 1$ at time $t = t_1$ and B could be the qubits $x = 2, 3$ at time $t = t_2$. The experimentalist is allowed to couple a general ancilla to regions A and B of the system. This can include ordinary measurements as well as more sophisticated quantum couplings with the ancilla, such as applying a quantum perturbation at A and measure its consequence in region B. The goal of the experimentalist is to distinguish between two situations. The first situation is where the experimentalist accesses regions A and B of the *same* system, and therefore can measure correlations between them. The second situation where the coupling at A and B occurs in two *independent* copies of the original system. By construction, in the second situation there is no correlation between the two subsystems. The difficulty of distinguishing the two situations measures the amount of correlation, which can be characterized by a relative entropy. The STMI is defined by optimizing this relative entropy over all possible system-ancilla coupling schemes. The advantage of our quantity is that it directly describes an experimentally accessible setup.

The remainder of the paper is organized as follows. In section 2 we introduce the intuition and the definition of the STMI, and explore some of its simple properties. We show that

JHEP05(2024)338

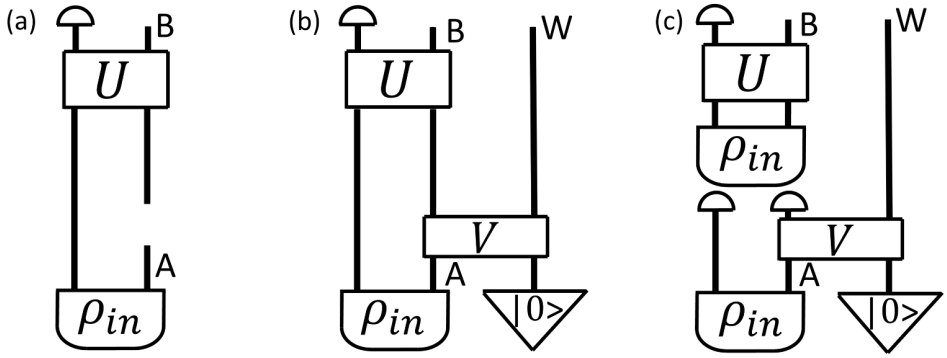

Figure 1. (a) System with initial state ρ_{in} undergoing evolution with unitary U. The upper cap stands for tracing over the corresponding region. (b) Coupling between subsystem A and ancilla W giving rise to the connected state ρ_{BW}. (c) Disconnected state $\rho_{B,0} \otimes \rho_W$, where $\rho_{B,0}$ is the unperturbed evolved state reduced to subsystem B.

the STMI reduces to the ordinary mutual information when regions A and B are spacelike separated. We investigate various properties of STMI such as its monotonicity under local operations at A and B separately. In section 3 we prove that the STMI is an upper-bound of connected *dynamical* correlation functions, a direct generalization of the corresponding inequality in ref. [1] for static correlations. In section 4 we show that, in certain cases, when three subregions are considered, our quantity satisfies a space-time generalization of the Markov property. In section 5 we draw a connection between the STMI and quantum channel discrimination and use this to prove additivity of the STMI in a special setting. Section 6 proposes a simplification of the definition of the STMI, which holds in a restricted regime and which we apply to provide semi-analytic results when studying the examples of section 7. Finally, we introduce a classical counterpart of the STMI in section 8 and discuss conclusions and outlook in section 9.

2 Definition of space-time mutual information

2.1 General intuition

Before presenting the rigorous definition of STMI, we would like to provide some intuition by presenting a simplified version of this quantity. Let us consider a system in an initial state ρ_{in} defined on a Hilbert space on region $A\bar{A}$ that evolves by time evolution U into an output state $\rho_{\text{out}} \equiv U\rho_{\text{in}}U^\dagger$, which can be partitioned into region B and $\bar{B}$. Here, A and B denote two subregions before and after the evolution, respectively, as in figure 1(a). We would like to introduce a generalization of mutual information that is applicable to general subregions A and B, in particular to when these are causally connected. In the latter case the standard definition of mutual information does not apply, as there is no joint state on AB. To overcome this issue we couple subsystem A to an ancilla W, which plays the role of an idler, allowing us to "carry to the future" the information encoded in A, as in figure 1(b). We shall denote the resulting state as ρ_{BW}. The operator V that couples ancilla and system, as well as the dimension of W, are for now arbitrary. Without loss of generality, as we will see momentarily,

JHEP05(2024)338

we can always assume V to be unitary, and the initial state of W to be pure. Now, recall that, when A and B are spatially separated and can be embedded in the same Hilbert space, the standard mutual information is the relative entropy between the connected state reduced on AB, ρ_{AB}, and the disconnected state $\rho_A \otimes \rho_B$. By analogy with this, in the case when A and B are causally connected, we consider the relative entropy between the connected state ρ_{BW} and an analog of the disconnected state $\rho_A \otimes \rho_B$. This disconnected state should be such that subregion B is unaffected by the presence of the perturbation V acting on A. The natural choice for such state is then given in figure 1(c), which we write as $\rho_{B,0} \otimes \rho_W$, where $\rho_{B,0}$ denotes the unperturbed evolved state in B, and $\rho_W = \mathrm{tr}_B(\rho_{BW})$ is the state of W after coupling with A, which is determined by ρ_{in} and coupling V (which is therefore independent from the time evolution U). Intuitively, the state $\rho_{B,0} \otimes \rho_W$ is the state that determines the disconnected term $\langle O_B \rangle \langle O_A \rangle$ in correlation functions, for any operators O_A, O_B.

Finally, we define the space-time mutual information $J_1(A : B)$ by maximizing the relative entropy over the ancilla-system coupling V:

$$J_1(A : B) = \sup_V S(\rho_{BW} | \rho_{B,0} \otimes \rho_W) \,. \tag{2.1}$$

We now see that it is sufficient to consider unitary coupling between system and ancilla. Indeed, if we take V to be a generic quantum channel, this is equivalent to having a unitary coupling to a bigger W followed by partial trace, which will only reduce the relative entropy, so for the purpose of taking the supremum it is sufficient to consider unitaries. Since we assume W is arbitrarily large, we can also take the initial state of W to be a pure state. If the initial state is a mixed state, we can purify it by enlarging W.

As we will show below, this definition already satisfies two important requirements: it reduces to the standard mutual information when A and B are spatially separated, and it bounds all space-time correlation functions of operators supported on A and B, with any normal (i.e., defined on the Schwinger-Keldysh time contour) time ordering.

The general definition of STMI is similar to eq. (2.1) except that it is defined for N copies of the initial state. In the next subsection we shall introduce the general definition based on a more rigorous setup of quantum hypothesis testing.

2.2 Definition

In this subsection we provide the rigorous reasoning behind our definition of STMI, and present the general definition. We begin by reviewing the hypothesis testing interpretation of relative entropy [2, 3]. To this aim, consider a black box that may contain either N copies of quantum state ρ, or N copies of quantum state σ. We are allowed to perform arbitrary measurements on this N-copied system to tell whether it is $\rho^{\otimes N}$ or $\sigma^{\otimes N}$. Now, if we make a hypothesis that the state is $\sigma^{\otimes N}$, and carry out some measurement, we can compute the probability of a measurement output *assuming* the state is $\sigma^{\otimes N}$. If our hypothesis is correct, this probability will reach 1 in large N, but if our hypothesis is incorrect, i.e. if the state is actually ρ, then the typical result has a smaller probability $P_N(\rho|\sigma)$, which is lower bounded by the relative entropy: $P_N(\rho|\sigma) \geq e^{-NS(\rho|\sigma)}$. Here, $S(\rho|\sigma) = \mathrm{tr}(\rho \log \rho - \rho \log \sigma)$. A smaller probability means that one can conclude the hypothesis is wrong with a higher confidence. Therefore, the relative entropy provides the fastest rate at which one can identify a wrong

JHEP05(2024)338

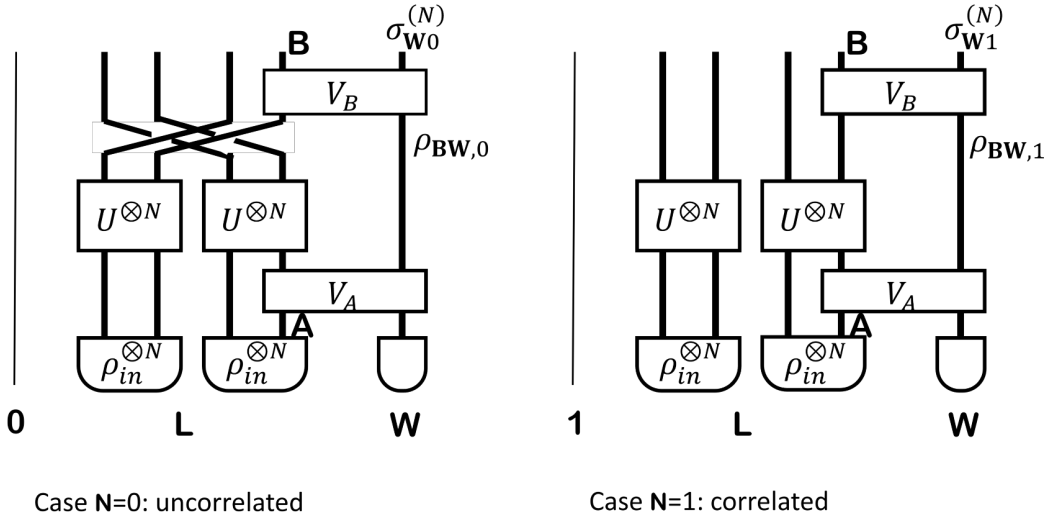

Figure 2. Definition of $J_N(A:B)$ in eq. (2.2) for the uncorrelated and correlated case.

hypothesis. If ρ and σ are close to each other, $S(\rho|\sigma)$ is small and it will be difficult to distinguish them. In contrast, if σ is not full rank, there are states which never appear in σ. For example, in a qubit system with states $|0\rangle, |1\rangle$, if $\sigma = |0\rangle\langle 0|$, then the probability of seeing $|1\rangle$ is zero. Thus if the probability of $|1\rangle$ is nonzero in ρ, if we perform a measurement in this basis we are certain that the state is ρ the moment we observe $|1\rangle$, which means $P_N = 0$ for a finite N. This corresponds to a diverging relative entropy.

For two spacelike separated regions A, B, the mutual information $I(A:B)$ is a relative entropy $I(A:B) = S(\rho_{AB}|\rho_A \otimes \rho_B)$. Thus mutual information determines the probability that ρ_{AB} is mistaken to be the uncorrelated state $\rho_A \otimes \rho_B$. Inspired by this intepretation, as a space-time generalization we consider the two situations illustrated in figure 2. In both cases, the experimentalist controls the ancilla W and the gates V_A, V_B which couple W with the physical system L. This coupling characterizes the most general experiments that can occur, which includes measurements of correlation functions in the original system, and also includes more general quantum processes. For example, we can swap a qubit in A with a qubit in W and later measure the swapped-out qubit together with B in an entangled basis. In situation 1, W is coupled with the same N copies of L at the space-time region A and B. In situation 0, the coupling at B occurs with a new set of N copies of systems, such that there won't be any correlation between B and W. The experimentalist does not know whether it is situation 0 or 1. If they make a conjecture that it is situation 0, and actually it is situation 1, the rate of finding out that the conjecture is wrong is determined by the relative entropy of the output state of W, denoted as $\sigma_{W1}^{(N)}$ and $\sigma_{W0}^{(N)}$ respectively. Thus it is natural to define the space-time mutual information $J(A:B)$ by optimizing this relative entropy over the choice of V_A, V_B:

$$J_N(A:B) = \frac{1}{N} \sup_{V_A,V_B} S\left(\sigma_{W,1}^{(N)} \middle| \sigma_{W,0}^{(N)}\right), \qquad J(A:B) = \lim_{N\to\infty} J_N(A:B) \tag{2.2}$$

More formally, this setup is an example of quantum algorithmic measurement (QUALM) defined in ref. [4]. The two situations in figure 2 are denoted as two "lab oracles," where

JHEP05(2024)338

the intrinsic dynamics U is given by nature, while the experimentalist has the option of choosing the coupling V_A, V_B between L and W. A QUALM refers to an algorithm, i.e. a choice of gates V_A, V_B for the purpose of achieving a particular task, similar to how a quantum algorithm is chosen to achieve a certain classical computation. In our case, the task is to distinguish the two lab oracles, with the optimal QUALM obtained by maximizing the relative entropy between the two output states.

The definition (2.2) can be simplified by realizing that the supremum over V_B can be achieved explicitly. Denote the state before applying V_B for B and W as $\rho_{B^N W,a}$, $a = 0, 1$. One can see that $\sigma^{(N)}_{W,a}$ is related to $\rho_{B^N W,a}$ by a quantum channel induced by V_B followed by a partial trace over B:

$$\sigma^{(N)}_{W,a} = \mathcal{C}\left(\rho_{B^N W,a}\right) \tag{2.3}$$

Due to the monotonicity of relative entropy under quantum channels, we have

$$S\left(\sigma^{(N)}_{W,1}|\sigma^{(N)}_{W,0}\right) \leq S\left(\rho_{B^N W,1}|\rho_{B^N W,0}\right) \tag{2.4}$$

More explicitly, we can define $W = W_A W_B$, and only W_A is acted upon by the coupling V_A. W_B has an initial state that is in direct product with W_A, and it has the same size as B. In this situation, $\rho_{B^N W,a} = \rho_{B^N W_A,a} \otimes \rho_{W_B}$. If we take V_B to be a swap between W_B and B, $\sigma^{(N)}_{W,a}$ is identical to $\rho_{B^N W_A,a}$, so that the relative entropy is the same. Therefore the swap operator achieves the optimization over V_B, and we can directly define the STMI using the state $\rho_{B^N W,a}$:

$$J_N(A:B) = \frac{1}{N} \sup_V S(\rho_{B^N W}|\rho^{\otimes N}_{B,0} \otimes \rho_W) \qquad J(A:B) = \lim_{N\to\infty} J_N(A:B), \tag{2.5}$$

where from now on we shall drop the subscript 1 from the connected state $\rho_{B^N W}$ for ease of notation, and we shall also drop the subscript A from W and V. In the above equation we have used the fact that $\rho_{B^N W,0}$ is by design in a factorized form $\rho_{B^N W,0} = \rho^{\otimes N}_{B,0} \otimes \rho_W$, where $\rho_W = \mathrm{tr}_{B^N}\left(\rho_{B^N W}\right)$ is determined by ρ_{in} and V.

As a side remark, one special choice of the ancilla and its coupling to subsystem A is to choose $W = W_1 W_2$, in which W_1 and W_2 each has the same Hilbert space dimension as that of A, and they are prepared in a maximally entangled state with each other. If we choose V to be a SWAP gate between A and W_1, ρ_{BW} is equivalent to the superdensity operator defined in ref. [5].

2.3 Properties of space-time mutual information

We now explore a few basic properties of $J(A:B)$.

Maximal size of W. In our definition of STMI, there is no restriction on the size of ancilla W. However, W does not need to be arbitrarily large. Without lost of generality, we can assume the initial state of $A\overline{A}$ and that of W to be pure, $|\Phi_{A\bar{A}}\rangle$ and $|\Phi_W\rangle$. If that is not the case, we can always enlarge the system and introduce its purification. For simplicity of notation let us consider the case with one copy. After applying unitary V we

JHEP05(2024)338

obtain a pure state $|\Psi_{A\overline{A}W}\rangle = V|\Phi_{A\overline{A}}\rangle \otimes |\Phi_W\rangle$. Expand this state in an arbitrary basis $|a_A\rangle$ of A, we can express

$$|\Psi_{A\overline{A}W}\rangle = \sum_{a,b=1}^{d_A} |b\rangle_A \otimes \langle a_A| \Phi_{A\overline{A}}\rangle \otimes V_{ba}|\Phi_W\rangle \tag{2.6}$$

This expression makes it explicit that the rank of the reduced state of W is always bounded by d_A^2 with d_A the Hilbert space dimension of A. Therefore it is always sufficient to take $d_W = d_A^2$. The discussion here easily generalizes to N copies, in which case it is sufficient to take $d_W = d_A^{2N}$. In other words, the size of W (number of qudits) can be taken as $2N$ copies of A. Additionally, by arbitrariness of V, we can choose $|\Phi_W\rangle$ to be the EPR state (or any other fixed reference state).

An alternative expression. We can decompose $J(A:B)$ into two terms:

$$J_N(A:B) = \sup_V \frac{1}{N}\left[-S\left(\rho_{B^N W}\right) - \mathrm{tr}\rho_{B^N W}\left(\log \rho_{B,0}^{\otimes N} + \log \rho_W\right)\right] \tag{2.7}$$

$$= \sup_V \frac{1}{N}\left[I^{(N)}\left(B:W\right) + S\left(\rho_{B^N}\middle|\rho_{B,0}^{\otimes N}\right)\right] \tag{2.8}$$

The first term is the ordinary mutual information in state $\rho_{B^N W}$, while the second term is the relative entropy between the state of B with and without the coupling with W. Here $\rho_B^{(N)} = \mathrm{tr}_W\left(\rho_{B^N W}\right)$. The second term is a consequence of the causal influence of A on B, which highlights the fact that even if we do not access the ancilla W, the coupling with W still has nontrivial effect on the state of B. Physically, these two terms represent two ways to distinguish the correlated state $\rho_{B^N W}$ and the uncorrelated state $\rho_{B,0}^{\otimes N} \otimes \rho_W$. The second term is sensitive to how much the coupling can change the state of B, while the first term implies that even if the state of B does not change, one can still measure the correlation between A and B by measuring that between B and W. For example, let us consider a simple case when the evolution operator is trivial, and A and B are the same spatial region. If initially A and $\overline{A}$ are in a maximally entangled EPR pair state, and W is in a maximally entangled EPR pair state of ancilla subregions W_1 and W_2, each with the same dimension as A, then we can take V to be a SWAP gate between A and W_1, after which $\rho_B = \rho_{B,0}$, so that the second term vanishes, but the first term is nonzero. Alternatively, if we prepare W_1 in a pure state and still apply a SWAP gate, the second term will be nonzero, while the first term is smaller. In general, the maximization over V is achieved by a compromise between these two terms.

Reduction to ordinary mutual information. If B and A are space-like separated so that we can define a quantum state ρ_{AB}, gates applied to A will never affect the state of B, so that $\rho_{B^N} = \rho_{B,0}^{\otimes N}$. In this case the second term in (2.8) vanishes. In addition, the mutual information $I^{(N)}\left(B:W\right)$ satisfies the monotonicity $I^{(N)}\left(B:W\right) \leq NI(A:B)$, where $I(A:B)$ is defined in the original system. Furthermore, the equal sign is achieved by choosing V to be a swap gate, so that we obtain $J_N(A:B) = I(A:B)$.

Monotonicity. The mutual information is monotonously non-increasing when a quantum channel is applied to A or B separately. The same applies to STMI. For region B, this

JHEP05(2024)338

follows directly from the monotonicity of relative entropy [6]. For a generic quantum channel $\mathcal{N}_B$ applied to B, we have

$$S\left(\mathcal{N}_B^{\otimes N}\left(\rho_{B^N W}\right)\middle|\mathcal{N}_B\left(\rho_{B,0}\right)^{\otimes N}\otimes\rho_W\right)\leq S\left(\rho_{B^N W}\middle|\rho_{B,0}^{\otimes N}\otimes\rho_W\right) \tag{2.9}$$

so that $J(A:B)$ is non-increasing.

For a quantum channel $\mathcal{N}_A$ applied to A, the proof is slightly more nontrivial. One can consider the dilation of $\mathcal{N}_A$, i.e. an isometry K from A to a bigger system $A\otimes\tilde{W}$. $\mathcal{N}_A$ is obtained by applying this isometry followed by tracing over $\tilde{W}$, i.e. $\mathcal{N}_A\left(\rho_A\right)=\mathrm{tr}_{\tilde{W}}\left(K\rho_A K^\dagger\right)$. Computing $J_N(A:B)$ after applying the channel $\mathcal{N}_A$ requires to apply V (which couples A and W) after applying K. Then we can merge $\tilde{W}$ with W and view $V\cdot K$ as a coupling between A and a bigger ancilla $\tilde{W}\otimes W$. Therefore for arbitrary V, we can apply the monotonicity of relative entropy by tracing over $\tilde{W}$ to obtain

$$S\left(\rho_{B^N W\tilde{W}}\middle|\left(\rho_{B,0}\right)^{\otimes N}\otimes\rho_{W\tilde{W}}\right)\geq S\left(\rho_{B^N W}\middle|\rho_{B,0}^{\otimes N}\otimes\rho_W\right). \tag{2.10}$$

Taking the supremum over V of the right-hand side yields the STMI $J_N(A:B)$ after applying quantum channel $\mathcal{N}_A$, while the left-hand side, upon optimization over V, yields the STMI without applying $\mathcal{N}_A$. This implies monotonicity of $J_N(A:B)$ with respect to the application of a quantum channel in either A or B.

As a special case of the above, $J_N(A:B)$ is non-increasing upon tracing over part of A or B. In other words, for disjoint regions A,C at the initial time, and disjoint regions B,D at the final time, we have $J(A:B)\leq J(AC:B)$, $J(A:B)\leq J(A:BD)$.

For the ordinary mutual information, when the inequality above is saturated, i.e. $I(AC:B)=I(A:B)$, the state ρ_{ABC} satisfies the Markov property and can be reconstructed from its marginal on AB. We obtained similar results for the STMI, which we will discuss in section 4.

Absence of upper bound. In eq. (2.8), the mutual information term is always finite, upper bounded by $2\log d_B$. The relative entropy term could diverge. The divergence occurs if $\rho_{B,0}$ is not full rank, and ρ_{B^N} has a nonzero probability to be in the null space of $\rho_{B,0}^{\otimes N}$. For example, if there is a conserved charge and the original state $\rho_{B,0}$ is supported in a charge range $[q_1,q_2]$, then as long as we can tune V to change the charge of B to be beyond this range, $J(A:B)$ diverges. If $\rho_{B,0}$ is full rank, with a nonzero minimal eigenvalue $p_{\min}$, one can prove that

$$S\left(\rho_{B^N}\middle|\rho_{B,0}^{\otimes N}\right)\leq\log\frac{1}{p_{\min}}-S\left(\rho_{B^N}\right)\leq\log\frac{1}{p_{\min}} \tag{2.11}$$

Thus we obtain $J(A:B)\leq 2\log d_B+\log\frac{1}{p_{\min}}$.

Relation to Choi state mutual information. A related quantity to the STMI is the mutual information of the Choi state corresponding to a given unitary evolution which, in terms of figure 3, is given by $I(B:W_2)$ [7]. Such quantity can be viewed as the mutual information term in (2.8) where we choose V to be the SWAP between A and W_2, with $W=W_1W_2$. We then infer that the Choi state mutual information cannot be larger than the STMI:

$$I(B:W_1)\leq J(A:B). \tag{2.12}$$

Additionally, note that in figure 3 $\bar{A}$ is in the infinite-temperature state. The space-time mutual information $J(A:B)$, on the other hand, is applicable for any initial state, including states where A and $\bar{A}$ are entangled.

JHEP05(2024)338

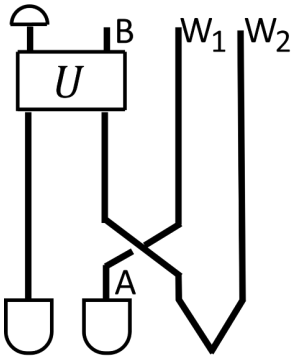

Figure 3. Mutual information of the Choi state corresponding to the unitary U.

Vanishing condition. The condition $J_N(A:B)=0$ requires $\rho_{B^N W}=\rho_{B,0}^{\otimes N}\otimes\rho_W$ for all V. This obviously requires all correlation functions in A and B to be disconnected. Conversely, assume that all (Keldysh-ordered) correlation functions between A and B factorize, i.e. (taking $N=1$ for simplicity)

$$\mathrm{Tr}\left(\mathcal{O}_B U\mathcal{O}_{A1}\rho_{\mathrm{in}}\mathcal{O}_{A2}U^\dagger\right)=\mathrm{Tr}\left(\mathcal{O}_B U\rho_{\mathrm{in}}U^\dagger\right)\mathrm{Tr}\left(O_{A1}\rho_{\mathrm{in}}\mathcal{O}_{A2}\right)\,. \tag{2.13}$$

Then, the BW correlation function in the system coupled with ancilla W is

$$\begin{aligned}&\mathrm{Tr}\left(\mathcal{O}_B\mathcal{O}_W UV\rho_{\mathrm{in}}V^\dagger U^\dagger\right)=\sum_{ij}\mathrm{Tr}\left(\mathcal{O}_B UV_i\rho_{\mathrm{in}}V_j^\dagger U^\dagger\right)\mathcal{O}_W^{ji}\\&=\sum_{ij}\mathrm{Tr}\left(\mathcal{O}_B U\rho_{\mathrm{in}}U^\dagger\right)\mathrm{Tr}\left(V_i\rho_{\mathrm{in}}V_j^\dagger\right)\mathcal{O}_W^{ij}=\mathrm{Tr}\left(\mathcal{O}_B U\rho_{\mathrm{in}}U^\dagger\right)\mathrm{Tr}\rho_{\mathrm{in}}\mathcal{O}_W\end{aligned} \tag{2.14}$$

also factorizes, where we decomposed the isometry $V=\sum_i V_i|i\rangle$, with $|i\rangle$ a basis on W and V_i an operator acting on A, and $\mathcal{O}_W^{ji}=\langle j|\mathcal{O}_W|i\rangle$. This in turn implies that ρ_{BW} necessarily factorizes, $\rho_{BW}=\rho_B\otimes\rho_W$, with $\rho_B=\rho_{B,0}$, thus implying $J_1(A:B)=0$. Similarly, when we couple N copies of A to W, we can also express the BW correlation as a sum over Keldysh-contour ordered correlators of A and B, which in turn proves that $J_N(A:B)=0$ if all AB correlation functions factorize.

Thermodynamics. If U is a chaotic time evolution $U=e^{-iHt}$ and B is smaller than half of the system size, for long enough t the subregion B will reach thermal equilibrium, so that the only dependence of ρ_B on V will by given by thermodynamic quantities. For simplicity let us consider a system with energy conservation and no other conservation law. In this case, the only possible change caused by the coupling V_A with the ancilla is the change of temperature. Thus we have $\rho_{B^N}=\rho_{\beta'}^{\otimes N}$ and $\rho_{B,0}^{\otimes N}=\rho_\beta^{\otimes N}$ are thermal states at temperatures β' and β, respectively. Here we have assumed that V_A does not cause a large energy fluctuation in the system, such that β' is well-defined. In this case, the mutual information $I(B^N:W)$ is negligible, and the main contribution to STMI comes from the relative entropy term:

$$J(A:B)\simeq S\left(\rho_{\beta'}|\rho_\beta\right)=\beta\left(\langle H\rangle_{\beta'}-\langle H\rangle_\beta\right)-\left(S_{\beta'}-S_\beta\right)=\beta\Delta F_\beta \tag{2.15}$$

This is the change of thermal free energy $F_\beta[\rho]=\langle H\rangle_\rho-\beta^{-1}S(\rho)$.

JHEP05(2024)338

More generically, V_A can cause a large energy fluctuation. For example, consider a single copy of the system coupled with W, and V_A can create a Schroedinger cat state in the system, so that the reduced density matrix of BW is

$$\rho_{BW} = \rho_{\beta'} \otimes |0\rangle\langle 0| + \rho_{\beta''} \otimes |1\rangle\langle 1| \tag{2.16}$$

where 0 and 1 are states of W. If this happens, there will be a nontrivial classical mutual information $I(B:W)$. Since energy is the only variable that W can correlate with, and quantum correlation cannot be preserved in B after a chaotic evolution, the mutual information is limited by the energy uncertainty. If V_A has the ability of changing the energy of B by ΔE at most, then the mutual information is upper-bounded by

$$I(B:W) \lesssim \log \frac{\Delta E}{\delta E} \tag{2.17}$$

with $\delta E = \sqrt{\langle H^2\rangle_B - \langle H\rangle_B^2}$ the intrinsic energy uncertainty. We expect this to be a small contribution. If we consider the limit that B is a finite (smaller than half) portion of the entire system, and consider the thermodynamic limit, then at most the mutual information term will be proportional to $\log|B|$, while the relative entropy term is proportional to $|B|$ as long as $\beta' \neq \beta$.

3 Bound on space-time correlation functions

An important property satisfied by the STMI is that, as we will now show, it bounds any two-point function between two possibly causally connected subregions. This generalizes the bound of the standard mutual information of spatial correlation functions [1].

Theorem 1. *The STMI bounds all two-point correlation functions between subsystems A and B. Explicitly, for any $N \geq 1$, we have the following bounds on symmetric and retarded correlation functions:*

$$J_N(A:B) \geq \frac{1}{8}\left(\frac{-i\mathrm{Tr}\rho_{\mathrm{in}}[\mathcal{O}_B(t),\mathcal{O}_A]}{||\mathcal{O}_A||_\infty||\mathcal{O}_B||_\infty}\right)^2 \tag{3.1}$$

$$J_N(A:B) \geq \frac{1}{8}\left(\frac{(\mathrm{Tr}\rho_{\mathrm{in}}\{\mathcal{O}_B(t),\mathcal{O}_A\})_c}{||\mathcal{O}_A||_\infty||\mathcal{O}_B||_\infty}\right)^2 \tag{3.2}$$

where $\mathcal{O}_B(t)$ is a (Heisenberg) operator supported in subregion B, and similarly for $\mathcal{O}_A$, $||\cdot||_\infty$ denotes the operator norm, $(\cdots)_c$ the connected component of a correlator, and $\mathcal{O}_A, \mathcal{O}_B$ are assumed to be Hermitian.

The numerators on the right-hand sides of the above inequalities correspond to retarded and connected symmetric two-point functions; these two correlators are sufficient to linearly generate any other Keldysh time ordering (such as Feynman, time-ordered correlators, etc.).

To prove the theorem, it is sufficient to consider the single-copy STMI $J_1(A:B)$. We take a specific choice of ancilla-system coupling, defined by

$$\rho_{\mathrm{in}} \otimes (|0\rangle\langle 0|)_W \to \sum_{ij} X_i \rho_{\mathrm{in}} X_j^\dagger \otimes (|i\rangle\langle j|)_W\,, \tag{3.3}$$

JHEP05(2024)338

where $i, j = 0, 1, 2$, and the operators X_i act on A and are defined as

$$X_0 = \sqrt{\tfrac{1}{2}}\mathbb{1}, \qquad X_1 = \sqrt{\tfrac{1}{2}}\frac{\mathcal{O}_A}{||\mathcal{O}_A||_\infty}, \qquad X_2 = \sqrt{\tfrac{1}{2}}\sqrt{\mathbb{1} - \frac{\mathcal{O}_A^2}{||\mathcal{O}_A||_\infty^2}}\,. \tag{3.4}$$

These operators satisfy $\sum_i X_i^\dagger X_i = \mathbb{1}$, thus guaranteeing that the coupling to the ancilla is an isometry, and can thus be extended to a unitary operator. The coupling (3.3) can be thought of as a control-$\mathcal{O}_A$ gate and was suitably chosen so that, as we will see, it will reproduce the two-point function we want to bound. Now, to prove the bound for the retarded two-point function (3.1), define the following operator acting on W:

$$Y_W = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{3.5}$$

We then find that the retarded two-point function can be viewed as the following expectation value over the state ρ_{BW}:

$$\mathrm{Tr}(\rho_{BW} Y_W \mathcal{O}_B) = \sum_{ij} \mathrm{Tr}\left[U X_i \rho_{\mathrm{in}} X_j U^\dagger \otimes (|i\rangle\langle j|)_W Y_W \mathcal{O}_B\right] = -\frac{i}{2}\mathrm{Tr}\rho_{\mathrm{in}}[\mathcal{O}_B(t), \mathcal{O}_A]\,. \tag{3.6}$$

Now consider the following sequence of inequalities

$$S(\rho_{BW}|\rho_{B0}\otimes\rho_W) \geq \frac{1}{2}||\rho_{BW} - \rho_{B0}\otimes\rho_W||_1^2 \geq \frac{1}{2}\left|\frac{\mathrm{Tr}(\rho_{BW} - \rho_{B0}\otimes\rho_W)Y_W\mathcal{O}_B}{||Y_W||_\infty||\mathcal{O}_B||_\infty}\right|^2, \tag{3.7}$$

where in the first step we applied the quantum Pinsker's inequality, and in the second step we used Hölder's inequality. Comparing (3.7) with (3.6), we arrive at (3.1). Note that the disconnected state does not contribute to the trace on the right-hand side of (3.7) due to the form of Y_W, which is equivalent to saying that the retarded two-point function does not have a disconnected component. The vanishing of this trace also implies that only the mutual information term in (2.8) contributes to the bound, i.e. we have a tighter bound given by

$$I(B:W) \geq \frac{1}{8}\left(\frac{-i\mathrm{Tr}\rho_{\mathrm{in}}[\mathcal{O}_B(t), \mathcal{O}_A]}{||\mathcal{O}_A||_\infty||\mathcal{O}_B||_\infty}\right)^2. \tag{3.8}$$

Similar steps lead to the bound on the symmetric two-point function (3.2), this time using, instead of Y_W,

$$X_W = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{3.9}$$

When A and B are causally disconnected, the symmetric two-point function reduces to the spatial one $\frac{1}{2}(\mathrm{Tr}\rho_{\mathrm{in}}\{\mathcal{O}_B(t), \mathcal{O}_A\})_c \to (\mathrm{Tr}\rho_{\mathrm{in}}\mathcal{O}_B\mathcal{O}_A)_c$ and we recover the standard bound of [1], with the correct numerical prefactor.

We emphasize that taking the supremum over V in the definition of J_1 is crucial for the above proof of the bounds (3.1), (3.2). Indeed, consider e.g. fixing $V = S_{AW_1} \otimes \mathbb{1}_{W_2}$, where S_{AW_1} is the swap between A and W_1. With this choice, ρ_{BW} constitutes an example of superdensity operator defined in [5]. In this case, as we show in appendix A, using similar steps as above, one can prove weaker bounds than (3.1), (3.2), with an overall dimensional suppression factor.

JHEP05(2024)338

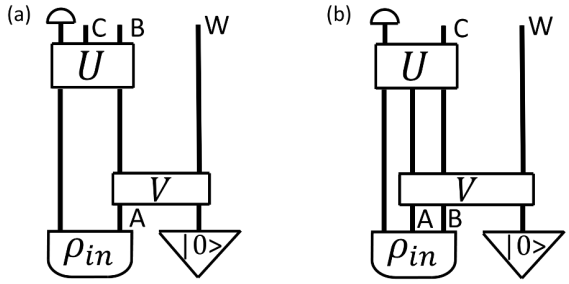

Figure 4. Illustration of the two situations involving three regions ABC, for the discussion of Markovian condition in section 4.

4 Markov property

In the case of ordinary mutual information, for three regions ABC one can define conditional mutual information

$$I(A:C|B) = I(A:BC) - I(A:B) \tag{4.1}$$

which is non-negative due to monotonicity of relative entropy under the action of quantum channels. $I(A:C|B)$ is the decrease of relative entropy $S(\rho_{ABC}|\rho_A \otimes \rho_{BC})$ under the partial trace over C. If $I(A:C|B) = 0$, the state ρ_{ABC} can be recovered from ρ_{AB} and ρ_{BC} using the Petz map [8, 9]:

$$\rho_{ABC} = \hat{T}\rho_{AB} \equiv \rho_{BC}^{\frac{1}{2}} \left(\rho_B^{-\frac{1}{2}} \rho_{AB} \rho_B^{-\frac{1}{2}} \otimes \mathrm{Id}_C \right) \rho_{BC}^{\frac{1}{2}} . \tag{4.2}$$

Since ρ_{ABC} is determined by ρ_{AB} and ρ_{BC}, correlation functions between A and BC can be determined by that between A and B. More explicitly, for any operator O_A supported on A and O_{BC} supported on BC, we can define

$$\tilde{O}_B = \rho_B^{-\frac{1}{2}} \rho_{BC}^{\frac{1}{2}} O_{BC} \rho_{BC}^{\frac{1}{2}} \rho_B^{-\frac{1}{2}} \tag{4.3}$$

such that

$$\mathrm{tr}\left(\rho_{ABC} O_A O_{BC}\right) = \mathrm{tr}\left(\rho_{AB} O_A \tilde{O}_B\right) . \tag{4.4}$$

Now we consider the situation with spacetime mutual information. Consider three regions A, B, C with B and C defined at the same future time, and A defined at an earlier time, as is shown in figure 4(a). Assume that

$$J_N(A:BC) = J_N(A:B) . \tag{4.5}$$

Naively, the unitary V_1 coupling A and W that optimize $J_N(A:B)$ may be different from the unitary V_2 that optimize $J_N(A:BC)$. However, they must actually be the same, which can be proven by contradiction. If $V_1 \neq V_2$, we can take $V = V_1$ and compute the relative entropy

$$\tilde{J}_N(A:BC) \equiv \frac{1}{N} S\left(\rho_{B^N C^N W} | \rho_{BC,0}^{\otimes N} \otimes \rho_W\right)\Big|_{V=V_1} \tag{4.6}$$

JHEP05(2024)338

If V_1 does not maximize this quantity, then we have

$$\tilde{J}_N(A:BC) < J_N(A:BC) = J_N(A:B) \tag{4.7}$$

Since $J_N(A:B)$ and $\tilde{J}_N(A:BC)$ are computed for the same system (with the same gate V_1), eq. (4.7) contradicts with monotonicity of relative entropy under partial trace over C. Therefore we have proven that for the same $V = V_1 = V_2$, the relative entropies

$$S\left(\rho_{B^N C^N W} | \rho_{BC,0}^{\otimes N} \otimes \rho_W\right) = S\left(\rho_{B^N W} | \rho_{B,0}^{\otimes N} \otimes \rho_W\right), \tag{4.8}$$

where both reach the maximum. Consequently, we can apply the Petz map and express

$$\rho_{BCW} = \hat{T}\rho_{BW} \equiv \rho_{BC0}^{\frac{1}{2}} \left(\mathrm{Id}_C \otimes \rho_{B0}^{-\frac{1}{2}} \rho_{BW} \rho_{B0}^{-\frac{1}{2}}\right) \rho_{BC0}^{\frac{1}{2}}. \tag{4.9}$$

Interestingly, the map acts trivially on W. For any operator O_{Bc} and O_W, there is a corresponding operator $\tilde{O}_B = \rho_{B0}^{-\frac{1}{2}} \rho_{BC0}^{\frac{1}{2}} O_{BC} \rho_{BC0}^{\frac{1}{2}} \rho_{B0}^{-\frac{1}{2}}$ such that $\mathrm{tr}\left(\rho_{BCW} O_{BC} O_W\right) = \mathrm{tr}\left(\rho_{BW} \tilde{O}_B O_W\right)$. This inturn implies that any correlation function between A and BC that one can measure indirectly through measuring correlation between BC and W can actually be converted into a measurement that only involves A and B. In other words, C does not directly correlate with A. The correlation between A and C are only generated through BC correlation and AB correlation. This is exactly in parallel with the Markovian condition in the ordinary spatial mutual information case.

The other situation is shown in figure 4(b), when regions A and B are at equal time and C is at a later time. In that case, if we have $J_N(AB:C) = J_N(B:C)$, it simply means that the optimal coupling between AB and W that maximizes the relative entropy does not involve A. It is sufficient to couple the ancilla to B in our optimization in order to find $J_N(AB:C)$, and A remains untouched. In this case, A is not a system that is traced out, and we cannot directly apply Petz map. A natural question is whether this condition also implies a Markov property on the joint ancilla-system state. We leave the exploration of this problem for future work.

5 Quantum channel discrimination and additivity

Quantum channel discrimination can be viewed as a natural extension of quantum hypothesis testing, where the aim is to discriminate channels instead of states. The formulation of this problem resembles that presented in section 2, although, as we now illustrate, it is applied to a slightly different context. Consider two quantum channels $\mathcal{N}_1$ and $\mathcal{N}_2$, both with input and output systems A and B, respectively. To discriminate whether the system measured has evolved through $\mathcal{N}_1$ or $\mathcal{N}_2$, one implements adaptive strategies consisting of alternating applications of the channel to discriminate and auxiliary channels. More explicitly, the application of $\mathcal{N}_1$ or $\mathcal{N}_2$ is alternated with channels $\mathcal{A}^{(i)}$ that map the output B, jointly with an ancilla W, to the input A of the subsequent application of the channel, together with the ancilla, as shown in figure 5(a). After N repetitions of this process, the final state is measured. From the formulation of the quantum hypothesis testing reviewed in section 2.2 we know that the probability of incorrectly concluding that the channel is $\mathcal{N}_2$ is $e^{-NS(\rho_1^{(\mathcal{A})} | \rho_2^{(\mathcal{A})})}$,

JHEP05(2024)338

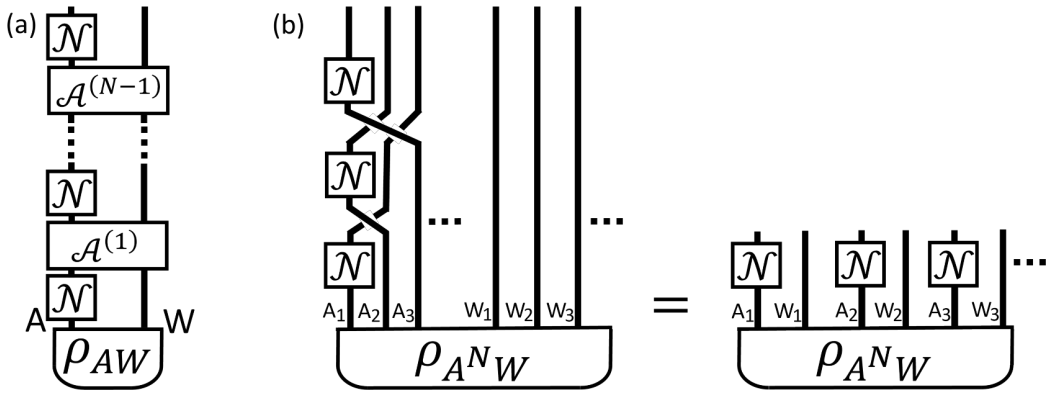

Figure 5. (a) Definition of $\mathcal{N}_N(\rho_{AW})$. (b) By specifying the $\mathcal{A}^{(i)}$ to suitable swaps, $\mathcal{N}_N(\rho_{AW})$ reduces to our (2.5).

where $\rho_1^{(\mathcal{A})}$ and $\rho_2^{(\mathcal{A})}$ are the output states that use channel $\mathcal{N}_1$ and $\mathcal{N}_2$, respectively, and the superscript $(\mathcal{A})$ indicates that we applied a given sequence of system-ancilla channels $\mathcal{A}^{(1)}, \mathcal{A}^{(2)}, \ldots$, corresponding to a particular strategy. The best strategy is then characterized by the optimization

$$\sup_{\rho, \mathcal{A}^{(i)}} S(\rho_1^{(\mathcal{A})} | \rho_2^{(\mathcal{A})}) \,, \tag{5.1}$$

where ρ is the input state in figure 5(a). Due to the joint convexity of the relative entropy and to the fact that $\rho_1^{(\mathcal{A})}$ and $\rho_2^{(\mathcal{A})}$ are linear in ρ, it is sufficient to restrict to pure states ρ.

Quantum channel discrimination has many applications, e.g.: quantum illumination, to enhance the detection of targets in the presence of thermal noise through entangled photon pairs [10]; quantum metrology, to estimate unknown parameters of quantum channels [11], and quantum reading, which involves the use of nonclassical transmitters to read data from classical digital memories [12].

5.1 Additivity of the STMI for initial pure states

Let us now come back to the setup of section 2.2. Assuming the initial state reduced on A is pure, i.e. $\rho_{\text{in}} = \rho_{\bar{A}} \otimes \rho_A$, with ρ_A pure, the optimization problem (2.5) reduces to a special case of quantum channel discrimination. To see the connection, note that both the system A^N and the ancilla W in eq. (2.5) are in a pure state, and optimizing over V corresponds to optimizing over the most general pure state of the joint system $A^N W$. One then writes the connected state $\rho_{B^N W}$ as N tensor copies of channel $\mathcal{N}$ acting on the state $\rho_{A^N W}$ on subregion A^N, where channel $\mathcal{N}$ is obtained by tracing the time evolution in (2.5) over $\bar{B}$. As shown in figure 5(b) this setup is then equivalent, for an appropriate choice of $\mathcal{A}^i$, to the setup of figure 5(a), where $\mathcal{A}^{(i)}$ simply swaps the i-th copy of A. For the disconnected state $\rho_{B,0}^{\otimes N} \otimes \rho_W$ we apply a similar reasoning, but instead of $\mathcal{N}$ we now have the replacer channel: $\mathcal{R}(\rho) = \rho_{B,0}$ for any state ρ of A. We then reduced (2.5) to an instance of quantum channel discrimination, and can write

$$J_N(A:B) = \frac{1}{N} \sup_{\rho_{A^N W}} S\left(\mathcal{N}^{\otimes N}(\rho_{A^N W}) | \mathcal{R}^{\otimes N}(\rho_{A^N W})\right) \tag{5.2}$$

JHEP05(2024)338

In particular, $J_1(A:B)$ identifies with the *channel relative entropy* between $\mathcal{N}$ and $\mathcal{R}$ [13]. We will see in section 5.2 how, in the general case, the STMI (2.5) can be formulated as a "constrained" quantum channel discrimination.

We now prove that the STMI $J_N(A:B)$, when ρ_A is factorized, is independent of N, i.e. it is additive. Additivity is a fundamental question in quantum channel discrimination and, fortunately, this property was recently proven when the alternative hypothesis (i.e., the second argument in the relative entropy) is a replacer channel [13], which precisely corresponds to our setup. Theorem 1 of [13] essentially states that, for $N \to \infty$, (5.1) is equal to the channel relative entropy $\sup_{\rho_{AW}} S(\mathcal{N}_1\rho_{AW}|\mathcal{N}_2\rho_{AW})$ when $\mathcal{N}_2$ is a replacer channel. Using the setup of the previous paragraph with $\mathcal{N}_1 = \mathcal{N}$ and $\mathcal{N}_2 = \mathcal{R}$, we then have

$$\sup_{\rho_{AW}} S\left(\mathcal{N}\rho_{AW}|\mathcal{R}\rho_{AW}\right) = \lim_N \frac{1}{N} \sup_{\rho,\mathcal{A}} S\left((\mathcal{N}\rho_{AW})^{\mathcal{A}}|(\mathcal{R}\rho_{AW})^{\mathcal{A}}\right) \geq J_N(A:B) \tag{5.3}$$

where the equality comes from the theorem mentioned above, and the inequality comes from that $J_N(A:B)$ corresponds to a special choice of $\mathcal{A}$, as shown in 5(b). Moreover, since $\sup_{\rho_{AW}} S(\mathcal{N}\rho_{AW}|\mathcal{R}\rho_{AW}) = J_1(A:B)$ due to (5.2) and $J_N \geq J_1$, we conclude additivity $J_N = J_1$.

We observe that, in quantum channel discrimination, there exist counterexamples to additivity when the alternative hypothesis is a more general channel [14].

We currently do not know if additivity holds for general initial states. In appendix B we show that the N-replicated optimization problem (2.5) admits a stationary point (i.e., V satisfies $\delta_V S(\rho_{B^NW}|\rho_{B^NW,0}) = 0$) of the form $V_N = V_1^{\otimes N}$, where V_1 is a stationary point for J_1. This is a necessary but not sufficient condition for additivity.

5.2 STMI as a constrained quantum channel discrimination

In the previous section we showed that when the initial state ρ_A is pure, the STMI can be viewed as a quantum channel discrimination. In this section, we will discuss the situation when A and $\bar{A}$ are entangled and in particular ρ_A is not pure. We shall see that it is most natural to think of this case as a constrained quantum channel discrimination. First, we can assume that the initial state of the system ρ_{in} is pure by extending $\bar{A}$ and tracing over the extension. We can then rewrite the STMI as

$$J_N(A:B) = \frac{1}{N} \sup_{\rho\in\mathcal{S}} S(\mathcal{N}^{\otimes N}(\rho)|\mathcal{R}^{\otimes N}(\rho)), \qquad \mathcal{S} = \{\rho : \text{Tr}_{A^NW}\rho = \rho_{\bar{A}}^{\otimes N}\}, \tag{5.4}$$

where $\rho_{\bar{A}} = \text{Tr}_A\rho_{\text{in}}$, and ρ is a state on $A^N\bar{A}^NW$. Also, $\mathcal{N}^{\otimes N}$ denotes N tensor copies of channel $\mathcal{N} : \bar{A}A \to B$, and $\mathcal{R}$ is the replacer channel introduced in section 5.1. Due to the joint convexity of the relative entropy and to the convexity of $\mathcal{S}$, it suffices to optimize over pure states.

One might ask if the constraint $\mathcal{S}$ in (5.4) can be implemented as a quantum channel in such a way that (5.4) can still be viewed as an unconstrained channel discrimination. More explicitly, we ask if there is a quantum channel $\mathcal{Q} : DW \to \bar{A}^NA^NW$, for some subsystem D, whose image $\mathcal{Q}(DW)$ is exactly $\mathcal{S}$ and acts trivially on W. The latter condition is necessary to maintain the structure of channel discrimination, i.e. W plays the role of an idler on which no channel is applied. The answer to this question is negative. Indeed, any state of $\mathcal{S}$ can be

JHEP05(2024)338

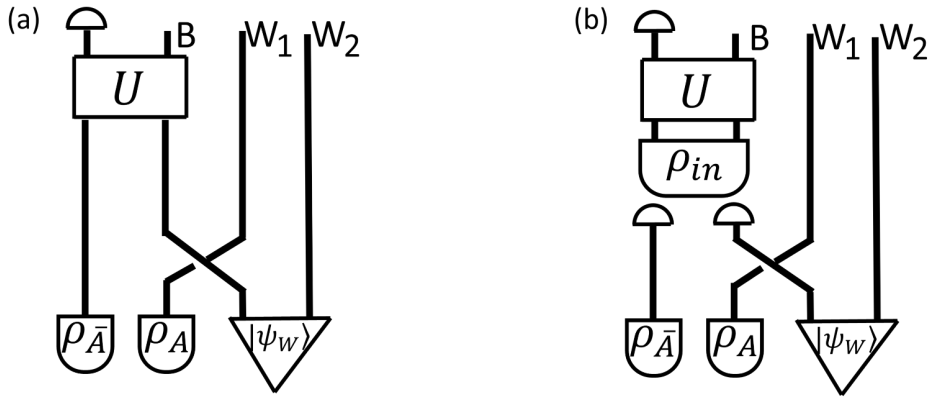

Figure 6. Representation of the ansatz (6.2) for the $N = 1$ correlated (a) and uncorrelated (b) states ρ_{BW} and $\rho_{B,0} \otimes \rho_W$.

written as $\rho = V\rho_{\text{in}}V^\dagger$, where $V : \bar{A}A \to \bar{A}AW$ is an isometry acting trivially on $\bar{A}$, ρ_{in} is a pure state, and for simplicity we are assuming $N = 1$. This means that $\text{Tr}_{\bar{A}}\rho$ has a fixed entanglement spectrum, independent of V. Assume that the quantum channel $\mathcal{Q}$ exists. Then, considering two arbitrary initial states σ_1 and σ_2 on DW, the states $\rho_1 = \text{Tr}_{\bar{A}}\mathcal{Q}(\sigma_1)$ and $\rho_2 = \text{Tr}_{\bar{A}}\mathcal{Q}(\sigma_1)$ should have the same entanglement spectrum. If their eigenstates are different, then $p\rho_1 + (1-p)\rho_2 = \text{Tr}_{\bar{A}}\mathcal{Q}(p\sigma_1 + (1-p)\sigma_2)$ will not have the same eigenvalues for $0 < p < 1$, and we thus infer that $\rho_1 = \rho_2$. This implies that $\text{Tr}_D\sigma_1 = \text{Tr}_A\rho_1 = \text{Tr}_A\rho_2 = \text{Tr}_D\sigma_2$, which contradicts the assumption that σ_1 and σ_2 are arbitrary states.

6 An ansatz for factorized initial states

The optimization problem (2.5) can be very non-trivial, especially for large Hilbert space dimension. In this section we propose an ansatz leading to a simplification that applies when the initial state is factorized, i.e.

$$\rho_{\text{in}} = \rho_{\bar{A}} \otimes \rho_A \,. \tag{6.1}$$

We shall restrict to a single replica $N = 1$. Keeping into account the bound $|W| \leq 2|A|$ from the discussion around eq. (2.6), it is sufficient to take $W = W_1 \otimes W_2$, with both W_1 and W_2 isomorphic to A. The ansatz consists of replacing V with a swap S between A and W_1, and optimize over a generic initial state $|\psi_W\rangle$ of W:

$$J_N(A:B) = \frac{1}{N}\sup_{\psi_W} S(\rho_{B^N W}|\rho_{B,0}^{\otimes N} \otimes \rho_W)\,, \tag{6.2}$$

as in figure 6.

We shall see below that the optimization reduces to a self-consistent equation for the reduced state of the ancilla W_2. This reduces the number of parameters to optimize from $2d_A^4 - d_A^2$ (an isometry going from $A \to AW_1W_2$) to d_A^2 (number of independent mixed states on W_2). In section 7 we show several numerical examples leveraging this ansatz, including MBL and thermalizing systems.

JHEP05(2024)338

As a support of this ansatz note that, when ρ_A is pure, following a similar reasoning as that around (2.6), we find that $|W| \leq |A|$. After applying the ancilla-system coupling V, we then have a pure state jointly defined on AW, and the optimization (2.5) reduces to optimizing over such state. This, in turn, is precisely the setup of figure 5(b), thus coinciding with our ansatz and proving its validity for pure ρ_A. As additional support for generic ρ_A, consider the relative entropy $S(\rho_B|\rho_{B0})$ as defined in eq. (2.8). One can easily see that, for any ancilla-system coupling V, the corresponding value of $S(\rho_B|\rho_{B0})$ can be reproduced by a suitable ansatz with ancilla state $|\psi_W\rangle$. Of course, to prove that the ansatz leads to a global maximum one would need to verify this statement for the entire expression in (2.8).

It is easy to see that (6.2) does not recover the optimum in (2.5) when A and $\bar{A}$ are entangled. In appendix C we show an explicit counterexample.

We now show how the above ansatz leads to a self-consistent equation for the state on the ancilla W_2. We shall restrict to the single-replica case $N = 1$ for simplicity. Below, $\mathcal{N} : A \to B$ will denote the quantum channel obtained from the evolution of the total system after tracing out $\bar{B}$ and choosing $\rho_{\bar{A}}$ as the initial state for $\bar{A}$. Using the ansatz, the relative entropy in (2.5) for $N = 1$ specifies to:

$$\begin{aligned} S(\rho_{BW}|\rho_{BW,0}) &= -S(\mathcal{N}(\rho_{AW})) + S(\rho_W) - \mathrm{Tr}(\mathcal{N}(\rho_{AW})(\log \mathcal{N}(\rho_{\mathrm{in}}) \otimes \mathrm{Id}_W) \\ &= -S(\mathcal{N}(|\psi\rangle\langle\psi|)) + S(\rho_W) - \mathrm{Tr}\rho_W \mathcal{N}^\dagger \log \mathcal{N}(\rho_{\mathrm{in}}) \end{aligned} \tag{6.3}$$

where ρ_{AW} denotes the joint state of system-ancilla after the swap has been applied, $\rho_W = \mathrm{Tr}_{W_2}|\psi_W\rangle\langle\psi_W|$. The reduced state of W_2 in $|\psi_W\rangle$ is equivalent to ρ_W by unitary conjugation. Since all quantities we discuss are invariant under applying a unitary operator to W_2, we can assume the reduced density operator of W_1 and W_2 to be the same ρ_W without loss of generality.

Writing the quantum channel $\mathcal{N}(\rho) = \sum_I K_I \rho K_I^\dagger$, where K_I are the Kraus operators satisfying $\sum_I K_I^\dagger K_I = \mathrm{Id}$, it is useful to view the indices I as labeling states $|I\rangle_E$ of an auxiliary system E, and to introduce the unitary evolution operator

$$\sum_I K_I \otimes |I\rangle_E : \mathcal{H}_A \to \mathcal{H}_B \otimes \mathcal{H}_E \tag{6.4}$$

Tracing over E leads to the channel $\mathcal{N}$, and tracing over B leads to the complement channel $\tilde{\mathcal{N}}$, with

$$\tilde{\mathcal{N}}(\rho_W) = \sum_{IJ} \Gamma_{IJ} |I\rangle_E\langle J|_E, \qquad \Gamma_{IJ} = \mathrm{Tr}(K_I \rho_W K_J^\dagger)\,. \tag{6.5}$$

In this notation, the first term in (6.3) reduces to the entropy of the ancilla's output state, so that we can express eq. (6.3) in terms of only the reduced state ρ_W:

$$S(\rho_{BW}|\rho_{BW,0}) = -S(\tilde{\mathcal{N}}(\rho_W)) + S(\rho_W) - \mathrm{Tr}\rho_W \mathcal{N}^\dagger \log \mathcal{N}(\rho_{\mathrm{in}})\,. \tag{6.6}$$

The optimization problem (2.5) then reduces to optimizing over ρ_W. The variation of (6.3) with respect to ρ_W can be written as

$$\delta S = \mathrm{Tr}\left(\delta\rho_W \left(\tilde{\mathcal{N}}^\dagger \log \tilde{\mathcal{N}}\,(\rho_W) - \log \rho_W - \mathcal{N}^\dagger \log \mathcal{N} \rho_{\mathrm{in}}\right)\right) \tag{6.7}$$

JHEP05(2024)338

thus leading to a self-consistent equation for ρ_W:

$$\rho_W = C \exp\left[\tilde{\mathcal{N}}^\dagger \log \tilde{\mathcal{N}}\left(\rho_W\right) - \mathcal{N}^\dagger \log \mathcal{N}\left(\rho_{\text{in}}\right)\right] \tag{6.8}$$

with C a normalization constant. One needs to remember to verify if the restricted (to $V = SWAP$) maximization is also a maximum (or at least a saddle point) for the unrestricted problem.

We now verify that the above ansatz is a stationary point of the optimization problem. Keeping into account factorization of the initial state (6.1), the variation of the relative entropy $S(\rho_{BW}|\rho_{BW,0})$ that is inside the sup argument in eq. (2.5) with respect to an infinitesimal change of the unitary $V \to (\text{Id} + iT)V$, with T an infinitesimal Hermitian operator acting on AW, is

$$\delta S = i\text{Tr}\left(T[\rho_{AW}, -\mathcal{N}^\dagger(\log \mathcal{N}(\rho_{AW})) + \text{Id}_A \otimes \log \bar{\rho}_W + \mathcal{N}^\dagger(\log \mathcal{N}(\rho_{\text{in}})) \otimes \text{Id}_W]\right), \tag{6.9}$$

where $\bar{\rho}_W = \text{Tr}_B \rho_{BW}$. We now plug in the ansatz by taking $W = W_1 \otimes W_2$, setting V to be the swap operator between A and W_1, and initializing the ancilla in a generic state $|\psi\rangle$. As a consequence, $\bar{\rho}_W = \rho_{\text{in}} \otimes \rho_W$, where ρ_W was defined around (6.3), $\rho_{AW} = (|\psi\rangle\langle\psi|)_{AW_2} \otimes (\rho_{\text{in}})_{W_1}$, and the above variation simplifies to

$$\delta S = i\text{Tr}\left(T\left[\text{Id}_{W_1} \otimes [|\psi\rangle\langle\psi|, -\mathcal{N}^\dagger(\log \mathcal{N}(|\psi\rangle\langle\psi|)) + \log \rho_W \otimes \text{Id}_{W_2} + \mathcal{N}^\dagger(\log \mathcal{N}(\rho_{\text{in}})) \otimes \text{Id}_{W_2}]\right]\right), \tag{6.10}$$

where we used that $[|\psi\rangle\langle\psi|, \text{Id} \otimes \rho_W] = [|\psi\rangle\langle\psi|, \rho_W \otimes \text{Id}]$. Assuming that $\mathcal{N}$ is unitary, the first term in the commutator vanishes. Plugging in (6.13), we find that latter is a stationary point of S with respect to the general variation (6.9). With generic $\mathcal{N}$ instead, we plug in (6.8) and find

$$\delta S = i\text{Tr}\, T(\text{Id}_{W_1} \otimes [|\psi\rangle\langle\psi|, -\mathcal{N}^\dagger(\log \mathcal{N}(|\psi\rangle\langle\psi|)) + \tilde{\mathcal{N}}^\dagger \log \tilde{\mathcal{N}}(\rho_W) \otimes \text{Id}]). \tag{6.11}$$

Noting that

$$\begin{aligned} |\psi\rangle\langle\psi|\mathcal{N}^\dagger(\mathcal{N}(|\psi\rangle\langle\psi|))^n &= (\Gamma^n)_{IJ}|\psi\rangle\langle\psi|K_J^\dagger K_I \\ \mathcal{N}^\dagger(\mathcal{N}(|\psi\rangle\langle\psi|))^n|\psi\rangle\langle\psi| &= (\Gamma^n)_{IJ}K_J^\dagger K_I|\psi\rangle\langle\psi|\,, \end{aligned} \tag{6.12}$$

we conclude that the commutator in (6.11) vanishes. We thus showed that for a generic quantum channel $\mathcal{N}$, (6.8) is a stationary point. It would be interesting to prove whether the ansatz is a local or global minimum in the general case.

Finally, we note that if the quantum channel $\mathcal{N} : A \to B$ is unitary, the ansatz allows us to analytically solve the optimization problem (2.5). In this case the ancilla's entropy is zero, $S(\tilde{\mathcal{N}}(\rho_W)) = 0$, and eq. (6.8) reduces to

$$\rho_W = \frac{\rho_{\text{in}}^{-1}}{\text{Tr}\rho_{\text{in}}^{-1}}\,. \tag{6.13}$$

Plugging the solution in the relative entropy (6.3) leads to the space-time mutual information

$$J_1(A:B) = \log \text{Tr}\rho_{\text{in}}^{-1}\,. \tag{6.14}$$

This indicates that, when the initial state is pure, the state of the ancilla can be chosen so that $J_1(A:B)$ diverges, even when the Hilbert spaces of A and B are finite-dimensional. As we commented around eq. (2.8), J_1 is unbounded, unlike the standard mutual information.

JHEP05(2024)338

7 Examples

In this section we shall study the behavior of the space-time mutual information in various quantum systems. We will first study its behavior in single-qubit systems subject to two different types of evolution and contrast the different behavior of the STMI in these two situations. We will then focus on many-body systems in two extreme cases: fully thermalizing and many-body localizing dynamics.

7.1 Single-qubit system

As a first example we consider the case where A is the entire system and is given by a single qubit. The time evolution is defined by a quantum channel $\mathcal{N}$ which maps it to an output state in B, also a single qubit. We will consider two types of quantum channel: depolarizing channel $\mathcal{N}_{\text{dpl}}$ and dephasing channel $\mathcal{N}_{\text{dph}}$, defined by

$$\mathcal{N}_{\text{dpl}}(\rho) = (1-p)\rho + \frac{p}{2}\text{Id}, \qquad \mathcal{N}_{\text{dph}}(\rho) = \left(1-\frac{p}{2}\right)\rho + \frac{p}{2}\sigma_3\rho\sigma_3\,, \tag{7.1}$$

with $0 \leq p \leq 1$. Here σ_i are the Pauli matrices. We will focus on the STMI (2.5) with a single replica of the system $N = 1$, i.e. $J_1(A:B)$.

To numerically optimize over V, we initialize the state of $A \otimes W_1 \otimes W_2$ to be $\rho_{AW} = \tilde{V}\rho_{\text{in}}\tilde{V}^\dagger$, where $\tilde{V} : A \to A \otimes W_1 \otimes W_2$ is a random isometry, and perform the following iterative updates on ρ_{AW}:

$$\rho_{AW} \to e^{-iM\eta}\rho_{AW}e^{iM\eta}, \qquad M = iV\frac{\delta S}{\delta V}\,, \tag{7.2}$$

where η is a positive number that sets the increment. The explicit expression of $V\frac{\delta S}{\delta V}$ can be obtained from (6.9).[1] Figure 7 shows the plots of $J_1(A:B)$ for the depolarizing and the dephasing channels as a function of p. Close to $p = 0$ the evolution approaches the identity, and we recover (6.14). As $p \to 1$, $\mathcal{N}_{\text{dpl}}$ becomes a fully depolarizing channel; all the information from the past is lost and the space-time mutual information approaches zero. On the other hand, as $p \to 1$ the dephasing channel $\mathcal{N}_{\text{dph}}$ still preserves classical information of the initial state, and therefore the space-time mutual information approaches a nonzero constant.

We shall now analytically evaluate the STMI $J_1(A:B)$ in a particular limit, using the ansatz of section 6. We focus first on the dephasing channel $\mathcal{N}_{\text{dph}}$. We write the initial state as

$$\rho_{\text{in}} = \frac{1}{2}\left(\text{Id} + \vec{a}\cdot\vec{\sigma}\right) \tag{7.3}$$

where $\vec{a} \in \mathbb{R}^3$ is a Bloch vector, with $|\vec{a}| = 1$, and we take $a_1 = \varepsilon$ and $a_3 = \sqrt{1-\varepsilon^2}$ with ε small, i.e., the state is almost an eigenstate of the evolution. In this case, the dephasing channel is close to the identity and we thus expect the space-time mutual information to diverge with ε. Adopting the ansatz (6.2), we take expression (6.6) as our starting point and we optimize it over the reduced state of the ancilla ρ_W, which we parameterize using (7.3)

[1] The infinitesimal change $\delta V = V' - V$ is related to T defined in (6.9) through $V' = (\text{Id} + iT)V$, so that $V\frac{\delta S}{\delta V} = -i\frac{\delta S}{\delta T}$.

JHEP05(2024)338

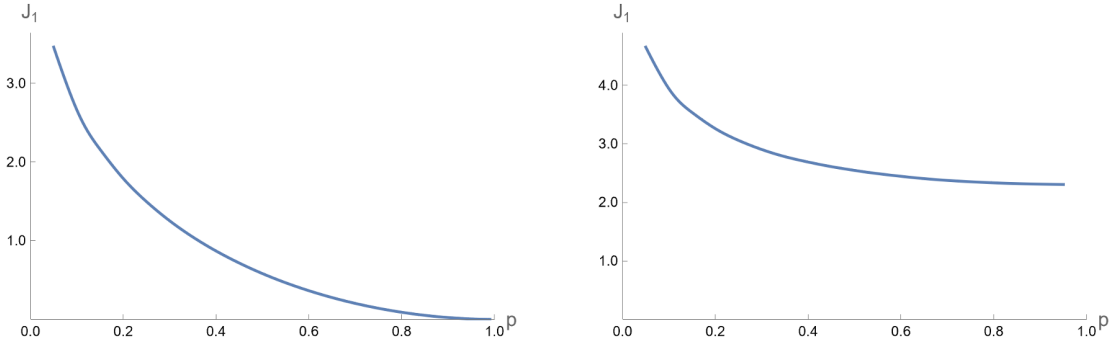

Figure 7. Plot of the STMI as a function of p for the depolarizing channel $\mathcal{N}_{\rm dpl}$ with initial state $\rho_{\rm in} = |0\rangle\langle 0|$ (left) and for the dephasing channel $\mathcal{N}_{\rm dph}$ with $\rho_{\rm in} = \sqrt{1-\varepsilon}|0\rangle\langle 0| + \sqrt{\varepsilon}|1\rangle\langle 1|$ and $\varepsilon \ll 1$ (right). For the dephasing channel, the STMI asymptotes to a finite value meaning that some information is preserved.

using the Bloch vector $\vec{b}$. We shall be interested in obtaining only the divergent part and it thus suffices to keep only the last term in (6.6):

$$S(\rho_{BW}|\rho_{BW,0}) \approx -\mathrm{Tr}\,\mathcal{N}_{\rm dph}(\rho_W)\log\mathcal{N}_{\rm dph}(\rho_{\rm in})\,. \tag{7.4}$$

From $\mathcal{N}_{\rm dph}(\rho_{\rm in}) = \frac{1}{2}(\mathrm{Id}+(1-p)a_1\sigma_1+a_3\sigma_3)$, and from the identity $\frac{1}{2}(\mathrm{Id}+\tanh\beta\vec{n}\cdot\vec{\sigma}) = \frac{e^{-\beta\vec{n}\cdot\vec{\sigma}}}{2\cosh\beta}$, with $\vec{n}^2 = 1$, we find

$$\log\mathcal{N}\rho_{\rm in} = 2\log\varepsilon|1\rangle\langle 1|\,, \tag{7.5}$$

which leads to

$$J_1(A:B) \approx -\sup_{\rho_W}\left(2\log\varepsilon\,\langle 1|\mathcal{N}_{\rm dph}(\rho_W)|1\rangle\right) = -2\log\varepsilon\,. \tag{7.6}$$

We then find that the space-time mutual information diverges as the initial state approaches an eigenstate of the evolution, and the divergent contribution is independent of p. We will see in the next subsection that a similar behavior happens for many-body localized systems.

For the depolarizing channel $\mathcal{N}_{\rm dpl}$ we consider a pure initial state, $\rho_{\rm in} = |0\rangle\langle 0|$. Using again the ansatz (6.6) we can exploit the symmetry of the evolution and choose the ancilla state

$$\rho_W = \frac{e^{\beta\sigma_3}}{2\cosh\beta} = \frac{1}{2}\mathrm{Id} + \frac{1}{2}\tanh\beta\sigma_3\,, \tag{7.7}$$

and we expect eq. (6.8) to be solved by $\rho_W \propto c_1\mathrm{Id} + c_2\sigma_3$. The complementary channel $\tilde{\mathcal{N}}$ given in terms of matrix Γ defined in (6.5) is[2]

$$\Gamma = \begin{pmatrix} \alpha_0^2 & \alpha_0\alpha_j n_j\tanh\beta \\ \alpha_0\alpha_i n_i\tanh\beta & \alpha_i\alpha_j(\delta_{ij} - i\tanh\beta\varepsilon_{ijk}n_k)\end{pmatrix}\,, \tag{7.8}$$

where $n_i = (0,0,1)$, ε_{ijk} is the totally antisymmetric tensor with $\varepsilon_{123} = 1$, and

$$\alpha_0 = \sqrt{1-\frac{3}{4}p}, \qquad \alpha_i = \sqrt{\frac{p}{4}}\,. \tag{7.9}$$

[2]We are using the Kraus decomposition of the depolarizing channel: $\mathcal{N}_{\rm dpl}(\rho) = \alpha_0^2\rho + \sum_i \alpha_i^2\sigma_i\rho\sigma_i$.

JHEP05(2024)338

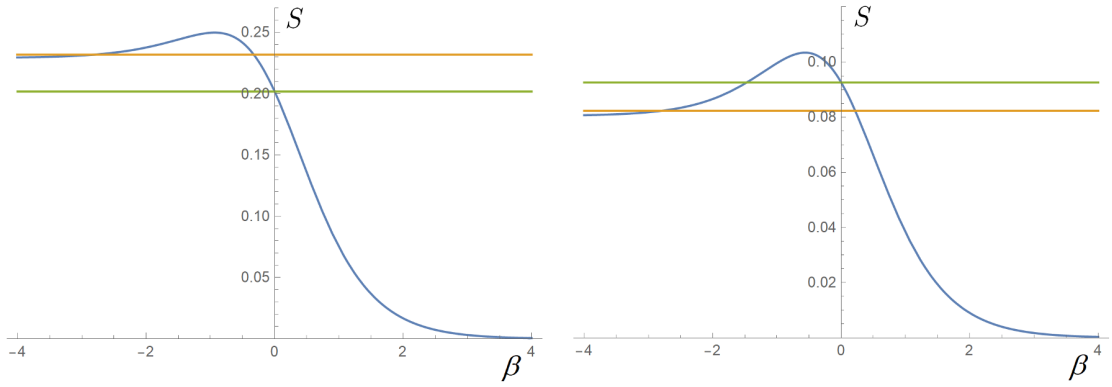

Figure 8. Plot of (6.3) as a function of β for a depolarizing channel with $p = 0.5$ (left) and with $p = 0.9$ (right). The STMI $J_1(A:B)$ is given by the maximum value of the blue curve. The green line represents the relative entropy using superdensity operator entanglement ($\beta = 0$), while the orange line represents the entropy obtained by substituting (6.13) in (6.3). Depending on p, one quantity can be better than the other.

Using these expressions, we evaluate $J_1(A:B)$ by extremizing (6.6) over β. A plot showing the relative entropy (6.3) as a function of β for various values of p is given in figure 8. The plot also illustrates a comparison between the STMI, the superdensity operator mutual information described at the end of section 2.2 (corresponding to $\beta = 0$), and the mutual information obtained by choosing the ancilla state (6.13).

We now find the exact value of the STMI for the depolarizing channel in the limit $p \to 1$. Using (6.6),

$$S(\rho_{BW}|\rho_{BW,0}) = -S(\tilde{\mathcal{N}}(\rho_W)) + S(\rho_W) - \mathrm{Tr}\mathcal{N}(\rho_W)\log\mathcal{N}(\rho_{\mathrm{in}})\,. \tag{7.10}$$

It is straightforward to evaluate each term in (7.10). As $p \to 1$, we obtain

$$S(\rho_{BW}|\rho_{BW,0}) = A(1-p)^2 + B(1-p)^3 + \cdots \tag{7.11}$$

where

$$\begin{aligned} A &= e^{-2\beta}(1+\coth\beta)(1+\beta+\beta\coth\beta)\tanh\beta \\ B &= -\beta\,\mathrm{csch}\beta\,\mathrm{sech}\beta\,. \end{aligned} \tag{7.12}$$

Numerically solving the optimization (2.5) for this relative entropy, we find that $\beta \approx -0.72 - (1-p)0.68$, and we thus see that $J_1(A:B) \to 0$ linearly as $1-p \to 0$. In the fully depolarizing case $p = 1$ the optimization problem is trivial as A and B are disconnected, thus any ancilla state ρ_W solves the optimization problem. However, what we find shows that the limit $p \to 1$ selects a unique ancilla state.

7.2 MBL and thermalization

We will now explore two extreme cases of many-body dynamics. The first example, a many-body localized (MBL) system, preserves an extensive amount of local operators, while the second example concerns a thermalizing system, and thus all local information is efficiently

JHEP05(2024)338

scrambled across the system. These two examples thus constitute contrasting cases where the STMI should display very different phenomenology.

We shall start with MBL and study the single-replica STMI $J_1(A:B)$, with A a single qubit at the initial time $t=0$ and B the same qubit at time t. As a model we consider a truncation of the MBL fixed-point Hamiltonian [15–17]:

$$H = \sum_i h_i \sigma_i^3 + \sum_{i<j} J_{ij}\sigma_i^3\sigma_j^3 + \sum_{i<j<k} J_{ijk}\sigma_i^3\sigma_j^3\sigma_k^3\,, \tag{7.13}$$

where the Pauli matrix on each site σ_i^3 is a local conserved operator. Here, $J_{ij} = e^{-|i-j|/\xi}\tilde{J}_{ij}$ and $J_{ijk} = e^{-|i-k|/\xi}\tilde{J}_{ijk}$, and $h_i, \tilde{J}_{ij}, \tilde{J}_{ijk} \in [-w, w]$ are drawn from a uniform distribution. Assuming the initial state is factorized between A and $\bar{A}$, tracing over $\bar{A}$ gives a channel $\mathcal{N}: A \to B$. We consider initial pure states on A of the form $\rho_{\text{in}} = |\chi\rangle\langle\chi|$, where

$$|\chi\rangle = \cos\alpha|0\rangle + \sin\alpha|1\rangle\,. \tag{7.14}$$

It is easy to see that states $|0\rangle, |1\rangle$ are preserved by the evolution. This means that, writing in matrix form using the $|0\rangle, |1\rangle$ basis, $\mathcal{N}$ must act as

$$\mathcal{N}\rho_{\text{in}} = \begin{pmatrix} \cos^2\alpha & f(t)\cos\alpha\sin\alpha \\ f^*(t)\cos\alpha\sin\alpha & \sin^2\alpha \end{pmatrix} \tag{7.15}$$

where $f(0) = 1$ and, as t grows, we expect generically that $f(t)$ vanishes, similarly to the fully dephasing channel $\mathcal{N}_{\text{dph}}$ discussed in the previous subsection. If $\alpha = 0$, the state does not evolve and remains pure, so that $J_1(A:B) = \infty$ for all times, consistently with the discussion around (6.14). For general α, and because ρ_{in} is pure, we can decompose the relative entropy $S(\rho_{BW}|\rho_{BW,0})$ as in (6.6):

$$S(\rho_{BW}|\rho_{BW,0}) = -S(\tilde{\mathcal{N}}(\rho_W)) + S(\rho_W) - \text{Tr}\left(\mathcal{N}\rho_W \log\mathcal{N}(\rho_{\text{in}})\right)\,. \tag{7.16}$$

For small α, the first two terms in (7.16) are finite, while the last one diverges. Applying the same manipulations as those around eq. (7.4), we find that

$$J_1(A:B) \approx -2\log\alpha\,, \tag{7.17}$$

showing that the closer the initial state (7.14) is to an eigenstate of the conserved operator σ^3, the larger the STMI will be at late time, consistently with the intuition that the STMI quantifies the information preserved by the system. Figure 9 shows the time-dependence of $J_1(A:B)$ and confirms this behavior.

For the thermalizing case, we consider a Floquet system whose evolution is generated by the unitary

$$\begin{aligned} U &= e^{-i\frac{\tau}{2}H_x}e^{-i\tau H_z}e^{-i\frac{\tau}{2}H_x} \\ H_x &= g\sum_{j=1}^{L}\sigma_j^x, \qquad H_z = \sum_{j=1}^{L-1}\sigma_j^z\sigma_{j+1}^z + h\sum_{j=1}^{L}\sigma_j^z\,. \end{aligned} \tag{7.18}$$

For $(g, h, \tau) = (0.9045, 0.8090, 0.8)$, this system is known to thermalize efficiently and have weak finite-size effects [18]. We consider the same STMI $J_1(A:B)$ as in the MBL case with

JHEP05(2024)338

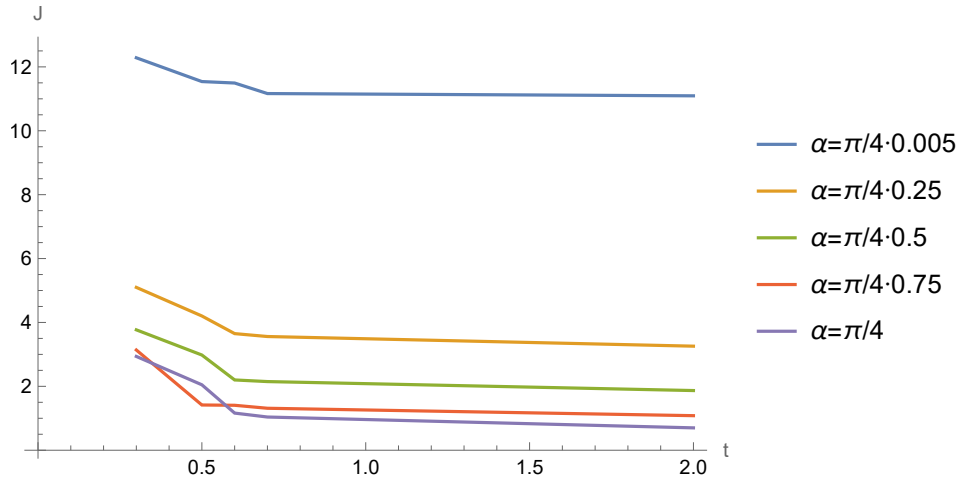

Figure 9. Plot of $J_1(A:B)$ for various values of α with evolution given by the MBL fixed point Hamiltonian (7.13). Each of these plots is obtained from a single disorder realization, with $w = 10$ and $\xi = 2$.

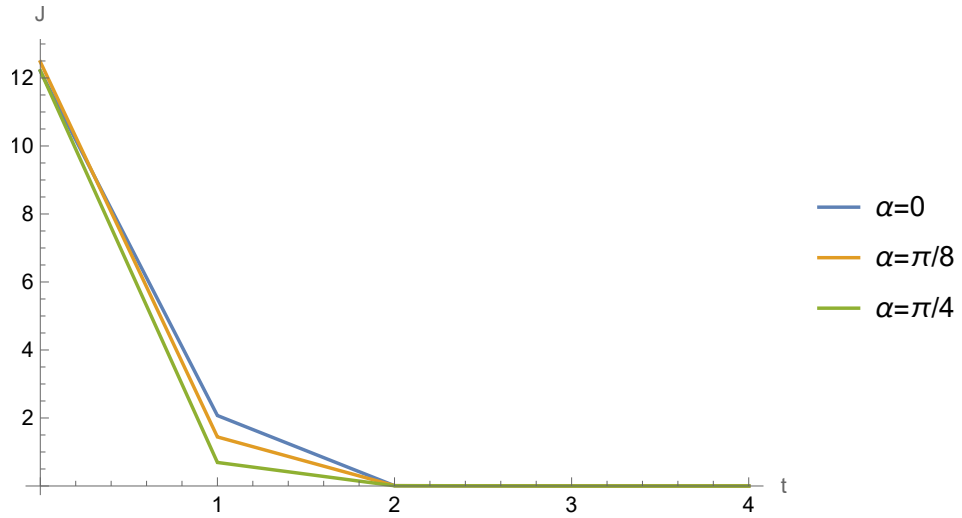

Figure 10. Plot of $J_1(A:B)$ for the Floquet evolution generated by (7.18). The initial state is $\rho_{\text{in}} = (1-\varepsilon)|\chi\rangle\langle\chi| + \frac{\varepsilon}{2}\,\text{Id}$, with $|\chi\rangle$ given in (7.14), and where $\varepsilon = 10^{-5}$ is used to regularize the initial value of the STMI.

coincident location of input and output qubits. Since there are no conserved operators in this case (including energy), we expect that $J_1(A:B)$ will quickly drop to zero for any initial state. Figure 10 shows $J_1(A:B)$ for various initial states, confirming our expectation. The timescale characteristic of this drop is given by the rate of decoherence caused by the effective channel describing the evolution of a single qubit.

8 Classical space-time mutual information

We now discuss how the STMI can be defined for classical systems, and make connection to information-theoretic quantities discussed in earlier literature. Consider a classical system

JHEP05(2024)338

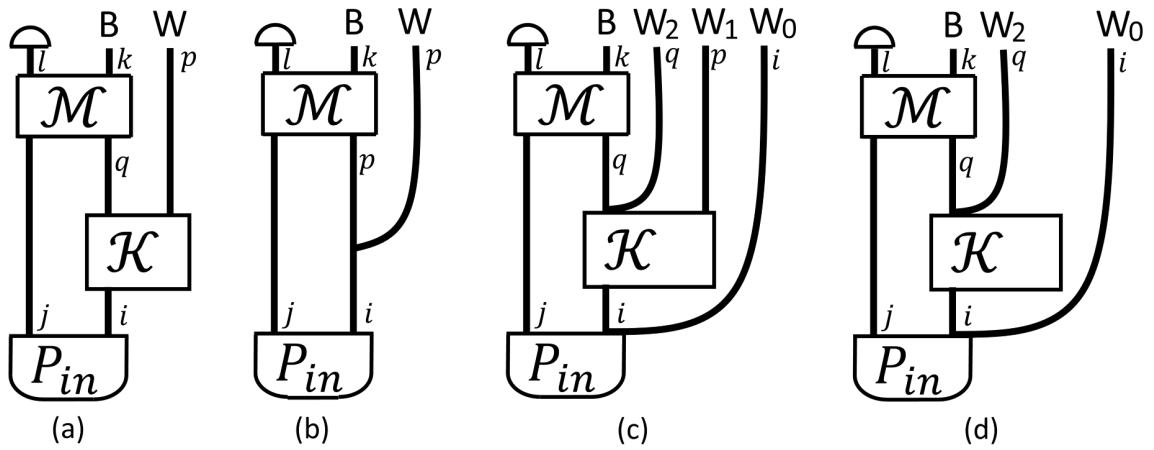

Figure 11. Representations of the probability P_{BW} for various ancilla-system couplings.

whose initial state is characterized by a probability over a state space S, which we denote by $P_{\text{in}}(i)$, where i labels a state in S. This probability is then mapped to an output probability through a stochastic map $\mathcal{M}: S \to S'$ acting as $P_{\text{in}}(i) \to \sum_i \mathcal{M}(j|i)P_{\text{in}}(i)$, where j labels states of S', and $\mathcal{M}(j|i)$ denotes the transfer matrix associated to the map $\mathcal{M}$ that satisfies normalization and positivity: $\sum_j \mathcal{M}(j|i) = 1$ and $\mathcal{M}(j|i) \geq 0$. The notion of STMI also applies to this classical setting with the important difference that the ancilla, as well as its coupling to the system, are themselves restricted to be classical. With $A \subseteq S$ and $B \subseteq S'$, we introduce a classical ancilla W and a stochastic map $\mathcal{K}$ that acts jointly on AW. From this we can define connected and disconnected ancilla-system probabilities, similarly to section 2. Taking a single copy of the system $N = 1$, these probabilities are:

$$\begin{aligned} P_{BW}(kp) &= \sum_{ljqri} \mathcal{M}(kl|qj)\mathcal{K}(qp|i)P_{\text{in}}(ij) \\ P_{BW,0}(kp) &= P_{B,0}(k)P_W(p)\,, \end{aligned} \tag{8.1}$$

where $P_{B,0}(k) = \sum_{ijl} \mathcal{M}(kl|ij)P_{\text{in}}(ij)$ is the unperturbed output state and $P_W(p) = \sum_k P_{BW}(kp)$. Here, k, l, p, i and j label states in $B, \bar{B}, W, A$ and $\bar{A}$ respectively. We represented the state for generic $N \geq 1$ in figure 11(a). The classical counterpart of the STMI is then

$$J_N(A:B) = \frac{1}{N} \sup_{\mathcal{K}} D(P_{B^N W,1}|P_{B^N W,0})\,, \tag{8.2}$$

where $D(\cdot|\cdot)$ denotes the Kullbeck-Leibler (KL) divergence, and we assume W to be sufficiently large.

One can show that, similarly to the quantum STMI, (8.2) bounds correlation functions as well as response functions. In fact, correlation functions can be already bounded by the input-output mutual information,[3] which can be obtained from the KL divergence in (8.2) by choosing $\mathcal{K}$ to be the copy channel: $\mathcal{K}(qp|i) = 1$ if $q = p = i$ and 0 otherwise, see figure 11(b).

[3]By intput-output mutual information, we simply mean the mutual information between A and B, where these have joint probability distribution $p(ki) = \sum_{lj} \mathcal{M}(kl|ij)P_{\text{in}}(ij)$ [19].

JHEP05(2024)338

To see this, we write the correlation function of two observables $\mathcal{O}_A$ and $\mathcal{O}_B$ as

$$\langle \mathcal{O}_B(t)\mathcal{O}_A\rangle = \sum_{ijkl} \mathcal{M}(kl|ij)P_{\text{in}}(ij)\mathcal{O}_B(k)\mathcal{O}_A(i)\,. \tag{8.3}$$

Using similar steps as in section 3, one easily obtains the bound $I(B:A) \geq \frac{1}{2}\left|\frac{\langle \mathcal{O}_B(t)\mathcal{O}_A\rangle_c}{||\mathcal{O}_B||_\infty||\mathcal{O}_A||_\infty}\right|^2$, where $||\mathcal{O}||_\infty = \sup_i |\mathcal{O}(i)|$. We then see that, in the classical case, correlations can be bound without adaptively optimizing over the system-ancilla coupling.

To bound response functions, on the other hand, the input-output mutual information is not enough: we need an adaptive ancilla-system coupling, like in the quantum case. Consider perturbing the evolution before applying channel $\mathcal{M}$ through a small perturbation from the identity, which itself can be viewed as a channel. Its transfer matrix can be written as $\mathcal{N}_A(k|i) = \delta_{ki} + \varepsilon N_A(k|i)$, with $\sum_k N_A(k|i) = 0$, $N_A(k|i) \geq 0$ for $k \neq i$ and $\varepsilon \geq 0$ small. The response function is then the leading order contribution in ε to the one-point function of $\mathcal{O}_B$:

$$\begin{aligned} G_R(\mathcal{O}_B, N_A) &= \lim_{\varepsilon\to 0} \varepsilon^{-1}\left(\langle \mathcal{O}_B(t)\rangle_{\mathcal{N}_A} - \langle \mathcal{O}_B(t)\rangle_{\mathcal{N}_A=\text{Id}}\right) \\ &= \sum_{jiklp} \mathcal{O}_B(j)\mathcal{M}(jl|kp)N_A(k|i)P_{\text{in}}(ip)\,. \end{aligned} \tag{8.4}$$

The input-output mutual information will not in general bound such two-point functions. For example, take $P_{\text{in}}(i) = \delta_{i0}$ to be a pure state, then $I(A:B) = 0$. One can easily see that there are choices of $\mathcal{O}_B$ and N_A with a nonzero response function so that the bound is violated.

We shall now prove the bound using $J_1(A:B)$, adopting a similar approach as in section 3. Introduce a 2-bit ancilla with states $p = 0, 1$, and the ancilla-system coupling

$$\mathcal{K}(j0|i) = \frac{1}{2}\left(\delta_{ij} + \frac{N_A(j|i)}{||N_A||_\infty}\right), \qquad \mathcal{K}(j1|i) = \frac{1}{2}\delta_{ij}\,, \tag{8.5}$$

where $||N_A||_\infty = \sup_j \sum_i |N_A(i|j)|$. Note that $\mathcal{K}$ satisfies positivity and normalization. Further introducing the observable $\mathcal{O}_W$ acting on the ancilla W, with $\mathcal{O}_W(p=0) = 1 = -\mathcal{O}_W(p=1)$, we find

$$\sum_{ipkj} \mathcal{O}_B(j)\mathcal{O}_W(p)\mathcal{M}(jl|k)\mathcal{K}(kp|i)P_{\text{in}}(i) = \frac{1}{2}\frac{G_R(\mathcal{O}_B, N_A)}{||N_A||_\infty}\,, \tag{8.6}$$

where for simplicity we left implicit the dependence on the indices in $\bar{B}$ and $\bar{A}$. Applying similar steps as in section 3 one then shows that the desired bound holds $J_1(A:B) \geq \frac{1}{8}\left(\frac{G_R(\mathcal{O}_B,N_A)}{||\mathcal{O}_B||_\infty||N_A||_\infty}\right)^2$, where the factor of $\frac{1}{8}$ also appeared in Theorem 1. for the quantum case.

Finally, we note that the expression of the classical STMI can be slightly simplified. Indeed, still taking to $N = 1$ for simplicity, we now show that to find the optimal J_1 one can restrict the state P_{BW} in (8.1) to the form

$$P_{BW}(kqi) = \sum_{lj} \mathcal{M}(kl|qj)\mathcal{K}(q|i)P_{\text{in}}(ij) \tag{8.7}$$

as illustrated in figure 11(d). The key is that copying to the ancillas W_0 and W_2 the state before and after applying $\mathcal{K}$, as in figure 11(c), does not require introducing any additional

JHEP05(2024)338

probing of the system, i.e. W_1 is not necessary. More precisely, the conditional mutual information $I(B : W_1|W_0W_2)$ can be readily shown to vanish. Due to this fact, and using the classical counterpart of (2.8), we have

$$D(P_{BW}|P_{BW,0}) = I(B : W_0W_2) + D(P_B|P_{B,0}) \,, \tag{8.8}$$

thus showing that marginalizing over W_1 does not affect $D(P_{BW}|P_{BW,0})$, and the optimal $\mathcal{K}$ in (8.2) can be achieved using (8.7). A consequence of this fact is that, if the initial state is factorized, i.e. $I(A : \bar{A}) = 0$, from figure 11(d) it is clear that maximization over $\mathcal{K}$ can be replaced by a maximization over the state of A with $\mathcal{K}$ fixed to the identity.

While we are not aware of discussions of the STMI in the literature, a restricted version of our implementation has been considered in the context of classical channel discrimination, where one optimizes over the input state [20, 21].

9 Conclusions

In this paper, we introduced the space-time mutual information (STMI), a quantity that generalizes mutual information to spatial subregions that can be separated in time. This was achieved by demanding that the STMI satisfies some of the natural properties possessed by the standard mutual information. The most stringent property leading to our definition (2.5) is that the STMI should bound space-time correlation functions between the two subregions.

We then investigated several properties that descend from our proposal, such as the Markov property and the relationship to quantum channel discrimination. We studied the behavior of the STMI in MBL and thermalizing many-body systems and found very distinct behaviors, thus, in a sense, providing a characterization of these two types of dynamics. Finally, we discussed a classical counterpart of the STMI.

Our framework can be extended in several directions. First, in this work we studied the time dependence of the STMI for two extreme cases in the context of many-body dynamics (MBL and Floquet thermalization). A natural next step is to look at more intermediate situations, e.g. thermalizing systems conserving a finite number of quantities such as energy or charge, or kinematically constrained models [22–24]. For subregions small enough compared to system size we expect the STMI to decay polynomially in time for these systems. When the subregions considered become large, we saw around eq. (2.15) that the STMI asymptotes to a finite value; it would be interesting to find how this asymptotic value is approached at late times. Another intriguing avenue for investigation involves examining the time dependence of the STMI as a diagnostic tool to differentiate between integrable and non-integrable systems, as explored in recent studies such as those highlighted in [25, 26]. Additionally, it will also be interesting to consider restrictions of the optimization over V which may characterize the type of information that the ancilla is able to extract from the system. For example, one can restrict V to be a one-way LOCC (local operations and classical communication) from A to W, which corresponds to an experimentalist who can only carry out classical measurements.

On an information-theoretic level, it is still an open question whether the STMI satisfies additivity. A positive answer to this would imply that it is sufficient to restrict to a single replica $N = 1$ in the definition of the STMI (2.5). We proved additivity in a restricted case

JHEP05(2024)338

where we could map our quantity to the channel relative entropy. While we could not find counter-examples to additivity in more general settings, it is still possible that additivity might not hold in full generality; we leave this question to future work.

We thank Aditya Cowsik, Patrick Hayden, Anton Kapustin, Vedika Khemani and Jinzhao Wang for illuminating discussions. This work is supported by the Department of Energy through Award DE-SC0019380 (PG and XLQ), the National Science Foundation under grant No.2111998 (XLQ), the It from Qubit Simons Collaboration (ZY), the Simons Fundation through Award No. 560571 (XLQ), the CAREER Award DMR-2045181 (PG), the Laboratory for Physical Sciences through the Condensed Matter Theory Center of the University of Maryland, College Park (PG), and the Simons Foundation through Award No. 620869 (PG). We are grateful to the long term workshop YITP-T-23-01 held at YITP, Kyoto University, where a part of this work was done. XLQ is grateful for the hospitality of Institute for Advanced Study, Tsinghua University (IASTU), where he visited while carrying out part of this work.

A Bound on correlations with superdensity operator

Here we obtain a bound on correlations similar to (3.1), (3.2) using the relative entropy $S\left(\rho_{BW}|\rho_{B,0}\otimes\rho_W\right)$, where $V=S_{AW_1}\otimes\mathbb{1}$ is the swap between A and W_1, and the ancilla W_1W_2 is prepared in an EPR state (see also figure 2). Using the norm bound,

$$S\left(\rho_{BW}|\rho_{B,0}\otimes\rho_W\right)\geq\frac{1}{2}\left\|\rho_{BW_1W_21}-\rho_{BW_1W_20}\right\|_1^2\geq\frac{\left(\langle\mathcal{O}_B\mathcal{O}_W\rangle_{\mathrm{c}}\right)^2}{2\left\|\mathcal{O}_B\right\|_\infty^2\left\|\mathcal{O}_W\right\|_\infty^2} \tag{A.1}$$

for any Hermitian operators $\mathcal{O}_B,\mathcal{O}_W$. The norm in the denominator is the operator norm, i.e. absolute value of the maximal eigenvalue, and where $\langle\mathcal{O}_B(t)\mathcal{O}_W\rangle_{\mathrm{c}}$ denotes any connected Schwinger-Keldysh time ordered two-point function. For example, consider a traceless Hermitian operator O_A, and define an operator on W by

$$\mathcal{O}_W=i\left(\left|I\right\rangle\left\langle I\right|O_A-O_A\left|I\right\rangle\left\langle I\right|\right) \tag{A.2}$$

with $|I\rangle$ the initial state of W, and O_A acts on W_1. When O_A is traceless, $\langle I|O_A|I\rangle=0$, and the norm of $\mathcal{O}_W$ satisfies

$$\left\|\mathcal{O}_W\right\|_\infty\leq2\left\|\mathcal{O}_A\right\|_\infty \tag{A.3}$$

On the other hand, inserting $\mathcal{O}_W$ into eq. (A.1) leads to

$$\langle\mathcal{O}_B\mathcal{O}_W\rangle_{\mathrm{c}}=d_A^{-2}i\,\mathrm{tr}\left(\left[\mathcal{O}_B(t),O_A\right]\rho_{\mathrm{in}}\right)\,. \tag{A.4}$$

Therefore we obtain an upper bound of response functions:

$$S\left(\rho_{BW}|\rho_{B,0}\otimes\rho_W\right)\geq\frac{\left(i\,\mathrm{tr}\left(\left[\mathcal{O}_B(t),O_A\right]\rho_{\mathrm{in}}\right)\right)^2}{8d_A^4\left\|\mathcal{O}_B\right\|_\infty^2\left\|O_A\right\|_\infty^2} \tag{A.5}$$

Note that the left-hand side can be replaced by the mutual information term $I(B:W)$ (the super-density operator mutual information [5]) because the disconnected part of the

JHEP05(2024)338

correlator vanishes, $\langle \mathcal{O}_W \rangle = 0$. The tighter bound is then

$$I(B:W) \geq \frac{\left(i \operatorname{tr}\left([\mathcal{O}_B(t), O_A]\,\rho_{\text{in}}\right)\right)^2}{8 d_A^4 \left\|\mathcal{O}_B\right\|_\infty^2 \left\|O_A\right\|_\infty^2} \,. \tag{A.6}$$

Note that (A.5) and (A.6) come with an additional factor of d_A^{-4} compared to (3.1) and (3.8), as stated in the main text, thus the bound we just proved is weaker. Similarly, one can obtain a bound for the symmetric two-point function (3.2) suppressed by the same factor.

In concluding this appendix, we note that we can bound the causal influence $CI(A:B)$ [27] by inserting a unitary U_A in A. This corresponds to

$$\mathcal{O}_W = U_A \left|I\right\rangle \left\langle I\right| U_A^\dagger \tag{A.7}$$

where in the above, U_A acts on W_1. Note that $\left\|\mathcal{O}_W\right\|_\infty = 1$. The inequality (A.1) becomes

$$S\left(\rho_{BW} | \rho_{B,0} \otimes \rho_W\right) \geq \frac{\left(\langle \mathcal{O}_B(t) \rangle \left(U_A\right) - \langle \mathcal{O}_B(t) \rangle \left(\mathbb{1}_A\right)\right)^2}{2 d_A^2 \left\|\mathcal{O}_B\right\|_\infty^2} \,. \tag{A.8}$$

B Additivity for general initial states

For generic N the initial state is $\rho_{\text{in}} = \rho_{\text{in,single copy}}^{\otimes N}$, and the state after coupling to the ancilla is $\rho_{(\bar{A}A)^N W} = V(\rho_{\text{in}} \otimes |\psi\rangle\langle\psi|)V^\dagger$. Restricting to unitary evolution for simplicity, and denoting the evolution of the replicated system by $U_N = U^{\otimes N}$, the connected and disconnected final states are

$$\rho_{B^N W} = \operatorname{Tr}_{\bar{B}^N} U_N \rho_{(\bar{A}A)^N W} U_N^\dagger \qquad \rho_{B^N W 0} = \left(\operatorname{Tr}_{\bar{B}^N} U_N \rho_{\text{in}} U_N^\dagger\right) \otimes \rho_W \,, \tag{B.1}$$

where $\rho_W = \operatorname{Tr}_{(\bar{A}A)^N} \rho_{(\bar{A}A)^N W}$. Consider an infinitesimal variation of V, $V \to (\mathrm{Id} + iT)V$, with T a Hermitian matrix acting on $A^N W$. This gives $\delta\rho_{(\bar{A}A)^N W} = i[T, \rho_{(\bar{A}A)^N W}]$, so that

$$\delta S(\rho_{B^N W} | \rho_{B^N W,0}) = i \operatorname{Tr}_{(\bar{A}A)^N W} T \left[\rho_{(\bar{A}A)^N W}, M_N\right] \,, \tag{B.2}$$

where

$$\begin{aligned} M_N = {} & U_N^\dagger \left(\mathrm{Id}_{\bar{B}^N} \otimes \log \rho_{B^N W}\right) U_N - U_N^\dagger \left(\mathrm{Id}_{\bar{B}^N} \otimes \log\left(\operatorname{Tr}_{\bar{B}^N} U_N \rho_{\text{in}} U_N^\dagger\right)\right) U_N \otimes \mathrm{Id}_W \\ & - \mathrm{Id}_{(\bar{A}A)^N} \otimes \log \rho_W \,. \end{aligned} \tag{B.3}$$

Restricting to factorized ancilla-system coupling $V = V_1^{\otimes N}$, we have $\rho_{(\bar{A}A)^N W} = (\rho_{\bar{A}AW})^{\otimes N}$, and thus $M_N = M_1 \otimes \mathrm{Id}^{\otimes(N-1)} + \mathrm{Id} \otimes M_1 \otimes \mathrm{Id}^{\otimes(N-2)} + \cdots$, where M_1 is the single-copy version of (B.3). We then find

$$\begin{aligned} \delta S(\rho_{B^N W} | \rho_{B^N W,0}) = i \operatorname{Tr}_{(\bar{A}A)^N W} T \Big(& [\rho_{\bar{A}AW}, M_1] \otimes (\rho_{\bar{A}AW})^{\otimes(N-1)} \\ & + \rho_{\bar{A}AW} \otimes [\rho_{\bar{A}AW}, M_1] \otimes (\rho_{\bar{A}AW})^{\otimes(N-2)} + \cdots \Big) = 0 \,, \end{aligned} \tag{B.4}$$

where in the last step we used that V_1 is a stationary point for $S(\rho_{BW}|\rho_{BW,0})$, as this is equivalent to $[\rho_{\bar{A}AW}, M_1]$. We thus showed that $V_N = V_1^{\otimes N}$ is a stationary point for the N-replica optimization (2.5). This however does not imply that such V_N is a global minimum, and thus we cannot conclude the additivity of J_N for general initial states. We leave this as an open question for future work.

JHEP05(2024)338

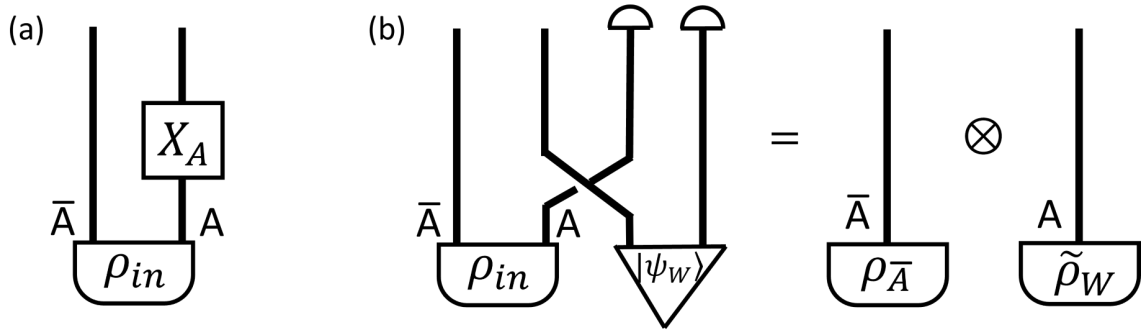

Figure 12. The setup in (a) yields a larger relative entropy compared to the one in (b) when ρ_{in} is given by (C.1), thus providing a counterexample to the ansatz of section 6 when A and $\bar{A}$ are entangled.

C Entangled initial state

We here discuss a counterexample to the ansatz proposed in section 6 when A and $\bar{A}$ are entangled. Suppose A and $\bar{A}$ are two qubits, with an entangled initial state

$$\rho_{\text{in}} = (1-\varepsilon)|\Gamma_{\bar{A}A}\rangle\langle\Gamma_{\bar{A}A}| + \frac{\varepsilon}{4}\text{Id}\,, \tag{C.1}$$

where $|\Gamma\rangle$ denotes an EPR state. Let us take the evolution of the system to be trivial for simplicity: $\rho_{B0} = \rho_{\text{in}}$, with B isomorphic to $\bar{A}A$. We first evaluate $J(A:B)$ by considering $V = X_A$ (i.e. the coupling to ancilla is the X Pauli matrix acting on A: coupling to the ancilla W is trivial and we take $W = \emptyset$), i.e.

$$\rho_B = \text{Tr}_W \rho_{BW} = X_A \rho_{\text{in}} X_A = (1-\varepsilon) X_A |\Gamma_{\bar{A}A}\rangle\langle\Gamma_{\bar{A}A}| X_A + \frac{\varepsilon}{4}\text{Id}\,, \tag{C.2}$$

see figure 12(a). We have

$$J(A:B) \geq S(\rho_B \rho_{B0}) = -\log\varepsilon + O(\varepsilon\log\varepsilon) \tag{C.3}$$

where in the last step we only kept the divergent part in ε, and where we used

$$\text{Tr}\rho_B \log\rho_B = \left(1-\frac{3}{4}\varepsilon\right)\log\left(1-\frac{3}{4}\varepsilon\right) + \frac{3}{4}\varepsilon\log\frac{\varepsilon}{4} \tag{C.4}$$

$$-\text{Tr}\rho_B \log\rho_{B0} = -\left(1-\frac{3}{4}\varepsilon\right)\log\frac{\varepsilon}{4} - \frac{\varepsilon}{4}\log\left(1-\frac{3}{4}\varepsilon\right) - 2\frac{\varepsilon}{4}\log\frac{\varepsilon}{4}\,. \tag{C.5}$$

Let us now estimate $J(A:B)$ using $V = \text{swap}$, and placing W in an arbitrary initial state. Then (see figure 12(b))

$$\rho_B = \rho_{\bar{A}} \otimes \tilde{\rho}_W\,, \quad \rho_{\bar{A}} = \frac{1}{2}\text{Id}, \quad \tilde{\rho}_W = \text{Tr}_{W_2}|\psi_W\rangle\langle\psi_W| \tag{C.6}$$

$$\log\rho_{B0} = \log\rho_{\text{in}} = \left(\log\left(1-\frac{3}{4}\varepsilon\right) - \log\frac{\varepsilon}{4}\right)|\Gamma_{\bar{A}A}\rangle\langle\Gamma_{\bar{A}A}| + \log\frac{\varepsilon}{4}\text{Id} \tag{C.7}$$

$$\text{Tr}\rho_B \log\rho_{B0} = \log\frac{\varepsilon}{4} + \left(\log\left(1-\frac{3}{4}\varepsilon\right) - \log\frac{\varepsilon}{4}\right)\langle\Gamma_{\bar{A}A}|\rho_B|\Gamma_{\bar{A}A}\rangle = \frac{3}{4}\log\varepsilon + O(\varepsilon) \tag{C.8}$$

JHEP05(2024)338

where we used

$$\langle \Gamma_{\bar{A}A}|\rho_B|\Gamma_{\bar{A}A}\rangle = \frac{1}{4}\sum_{ij}\langle ii|\mathrm{Id}_{\bar{A}}\otimes\tilde{\rho}_W|jj\rangle = \frac{1}{4}\,. \tag{C.9}$$

Then we have

$$J_{\mathrm{swap}}(A:B) = -\frac{3}{4}\log\varepsilon + \cdots \tag{C.10}$$

where the dots stands for terms that are bounded as $\varepsilon \to 0$ and include contributions from the mutual information $I_{\mathrm{swap}}(B:W)$. We see that the swap (C.10) leads to a smaller J than $V = X_A$ in (C.3).

References

[1] M.M. Wolf, F. Verstraete, M.B. Hastings and J.I. Cirac, *Area Laws in Quantum Systems: Mutual Information and Correlations*, *Phys. Rev. Lett.* **100** (2008) 070502 [arXiv:0704.3906] [INSPIRE].

[2] F. Hiai and D. Petz, *The proper formula for relative entropy and its asymptotics in quantum probability*, *Commun. Math. Phys.* **143** (1991) 99 [INSPIRE].

[3] M. Hayashi, *Asymptotic attainment for quantum relative entropy*, *J. Phys. A* **34** (2001) 3413 [quant-ph/9704040] [INSPIRE].

[4] D. Aharonov, J. Cotler and X.-L. Qi, *Quantum algorithmic measurement*, *Nature Commun.* **13** (2022) 887 [arXiv:2101.04634] [INSPIRE].

[5] J. Cotler, C.-M. Jian, X.-L. Qi and F. Wilczek, *Superdensity Operators for Spacetime Quantum Mechanics*, *JHEP* **09** (2018) 093 [arXiv:1711.03119] [INSPIRE].

[6] M.A. Nielsen and I.L. Chuang, *Quantum information and quantum computation*, Cambridge University Press (2000), p. 23 [ISBN:9780521632355].

[7] P. Hosur, X.-L. Qi, D.A. Roberts and B. Yoshida, *Chaos in quantum channels*, *JHEP* **02** (2016) 004 [arXiv:1511.04021] [INSPIRE].

[8] D. Petz, *Monotonicity of quantum relative entropy revisited*, *Rev. Math. Phys.* **15** (2003) 79 [quant-ph/0209053] [INSPIRE].

[9] P. Hayden, R. Jozsa, D. Petz and A. Winter, *Structure of States Which Satisfy Strong Subadditivity of Quantum Entropy with Equality*, *Commun. Math. Phys.* **246** (2004) 359 [INSPIRE].

[10] S. Lloyd, *Enhanced Sensitivity of Photodetection via Quantum Illumination*, *Science* **321** (2008) 1160627 [INSPIRE].

[11] S. Pirandola, R. Laurenza, C. Lupo and J.L. Pereira, *Fundamental limits to quantum channel discrimination*, *npj Quantum Inf.* **5** (2019) 50 [INSPIRE].

[12] S. Pirandola, *Quantum Reading of a Classical Digital Memory*, *Phys. Rev. Lett.* **106** (2011) 090504 [INSPIRE].

JHEP05(2024)338

[13] T. Cooney, M. Mosonyi and M.M. Wilde, *Strong Converse Exponents for a Quantum Channel Discrimination Problem and Quantum-Feedback-Assisted Communication*, *Commun. Math. Phys.* **344** (2016) 797.

[14] K. Fang, O. Fawzi, R. Renner and D. Sutter, *Chain rule for the quantum relative entropy*, *Phys. Rev. Lett.* **124** (2020) 100501 [arXiv:1909.05826].

[15] M. Serbyn, Z. Papić and D.A. Abanin, *Local Conservation Laws and the Structure of the Many-Body Localized States*, *Phys. Rev. Lett.* **111** (2013) 127201 [INSPIRE].

[16] D.A. Huse, R. Nandkishore and V. Oganesyan, *Phenomenology of fully many-body-localized systems*, *Phys. Rev. B* **90** (2014) 174202 [INSPIRE].

[17] A. Chandran, I.H. Kim, G. Vidal and D.A. Abanin, *Constructing local integrals of motion in the many-body localized phase*, *Phys. Rev. B* **91** (2015) 085425 [INSPIRE].

[18] L. Zhang, V. Khemani and D.A. Huse, *A floquet model for the many-body localization transition*, *Phys. Rev. B* **94** (2016) 224202 [arXiv:1609.00390].

[19] T. Cover and J. Thomas, *Elements of Information Theory*, Wiley (2012) [DOI:10.1002/0471200611].

[20] M. Hayashi, *Discrimination of two channels by adaptive methods and its application to quantum system*, arXiv:0804.0686 [DOI:10.1109/TIT.2009.2023726].

[21] A.W. Harrow, A. Hassidim, D.W. Leung and J. Watrous, *Adaptive versus nonadaptive strategies for quantum channel discrimination*, *Phys. Rev. As* **81** (2010) 032339 [arXiv:0909.0256].

[22] P. Glorioso, J. Guo, J.F. Rodriguez-Nieva and A. Lucas, *Breakdown of hydrodynamics below four dimensions in a fracton fluid*, *Nature Phys.* **18** (2022) 912 [arXiv:2105.13365] [INSPIRE].

[23] J. Richter and A. Pal, *Anomalous hydrodynamics in a class of scarred frustration-free Hamiltonians*, *Phys. Rev. Res.* **4** (2022) L012003 [arXiv:2107.13612] [INSPIRE].

[24] H. Singh, B.A. Ware, R. Vasseur and A.J. Friedman, *Subdiffusion and Many-Body Quantum Chaos with Kinetic Constraints*, *Phys. Rev. Lett.* **127** (2021) 230602 [arXiv:2108.02205] [INSPIRE].

[25] A. Lerose, M. Sonner and D.A. Abanin, *Scaling of temporal entanglement in proximity to integrability*, *Phys. Rev. B* **104** (2021) 035137 [arXiv:2104.07607] [INSPIRE].

[26] G. Giudice et al., *Temporal Entanglement, Quasiparticles, and the Role of Interactions*, *Phys. Rev. Lett.* **128** (2022) 220401 [arXiv:2112.14264] [INSPIRE].

[27] J. Cotler, X. Han, X.-L. Qi and Z. Yang, *Quantum Causal Influence*, *JHEP* **07** (2019) 042 [arXiv:1811.05485] [INSPIRE].

JHEP05(2024)338