

# Sharpening the second law of thermodynamics with the quantum Bayes theorem

Hrant Gharibyan and Max Tegmark

*Department of Physics & MIT Kavli Institute, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139*

(Received 3 October 2013; revised manuscript received 16 May 2014; published 19 September 2014)

We prove a generalization of the classic Groenewold-Lindblad entropy inequality, combining decoherence and the quantum Bayes theorem into a simple unified picture where decoherence increases entropy while observation decreases it. This provides a rigorous quantum-mechanical version of the second law of thermodynamics, governing how the entropy of a system (the entropy of its density matrix, partial-traced over the environment and conditioned on what is known) evolves under general decoherence and observation. The powerful tool of spectral majorization enables both simple alternative proofs of the classic Lindblad and Holevo inequalities without using strong subadditivity, and also novel inequalities for decoherence and observation that hold not only for von Neumann entropy, but also for arbitrary concave entropies.

DOI: [10.1103/PhysRevE.90.032125](https://doi.org/10.1103/PhysRevE.90.032125)

PACS number(s): 05.70.Ln, 03.65.Yz, 03.65.Ta, 03.67.Ac

## I. INTRODUCTION

More than a century after its formulation, the second law of thermodynamics remains at the forefront of physics research, with continuing progress on generalizing its applications and clarifying its foundations. For example, it is being extended to nonequilibrium statistical mechanics [1], quantum heat engines [2], biological self-replication [3], and cosmological inflation [4]. As to the quantum-mechanical version of the second law, Lloyd showed that it can be derived from imperfectly known quantum evolution [5], and one of the authors of the present paper showed how it can be generalized to observed open systems [4,6]. The goal of the present paper is to complete this generalization by providing the required mathematical proofs.

As emphasized by Von Neumann [7] and Feynman [8], state of a quantum system is completely described by a density matrix  $\rho$ , which encodes everything we need to know to make the best possible predictions for its future behavior. However, if you are interested in using physics to make predictions about your own future, knowing  $\rho$  for the entire universe (or multiverse) is neither sufficient nor necessary [4].

(1) *Is not sufficient*: You also need to know what branch of the global wavefunction you are in. In particular, you need to take into account what you know about your location, both in 3D space and in Hilbert space.

(2) *Is not necessary*: You only need to know the quantum state “nearby,” both in 3D space and in Hilbert space.

To predict what you will observe in a quantum physics laboratory, you need to take into account both which of the many existing physics laboratories you happen to be in and also which quantum state preparation has been performed. On the other hand, you do not need to take into account the current state of the Andromeda Galaxy or branches of the wavefunction that have permanently decohered from your own.

To quantify this, you can always decompose the total system (the entire cosmos) into three subsystems, as illustrated in Fig. 1: the degrees of freedom corresponding to your subjective perceptions (the subject), the degrees of freedom being studied (the object), and everything else (the environment). Computing the correct density matrix for your object of interest, therefore, involves two steps [4]:

(1) Condition on what you know (on all subject degrees of freedom).

(2) Marginalize over what you don’t care about (partial-trace over all environment degrees of freedom).

The first step, including quantum state preparation and observation, can be thought of as the quantum generalization of Bayes’s Theorem [4]. The second step produces decoherence, helping explain the emergence of a classical world [9–12]. As we will see below, decoherence always increases entropy, whereas observation on average decreases it. Although the latter was proved by Shannon for the case of classical physics, the corresponding quantum theorem has hitherto eluded proof except in special cases. We will sharpen the second law as follows:

(1) *Observation*: When an object is probed by the subject, its entropy on average decreases.

(2) *Decoherence*: When an object is probed by the environment, its entropy increases.

Below we will define “probed” and provide rigorous mathematical proofs of these entropy inequalities.

## II. EFFECTS OF DECOHERENCE AND OBSERVATION

Let us begin by briefly reviewing the unitary cosmology formalism of Refs. [4,6] necessary for our proofs. We define *probing* as a nontrivial interaction between the object and some other system that leaves the object unchanged in some basis, i.e., such that the unitary dynamics  $\mathbf{U}$  merely changes the system state in a way that depends on the object state  $|o_i\rangle$ . We will justify and generalize this definition in Sec. III.

*Object-environment*: If the object is probed by the environment the unitary dynamics of object-environment system is given by

$$\mathbf{U}|o_i\rangle|e_*\rangle = |o_i\rangle|\epsilon_i\rangle, \quad (1)$$

where  $|e_*\rangle$  and  $|\epsilon_i\rangle$  denote the initial and final states of the environment for the object state  $|o_i\rangle$ . The density matrix describing the object is the object-environment density matrix partial-traced over the environment. Equation 3 implies that the object-density matrix  $\rho$  evolves as [4,9]

$$\rho \mapsto \rho \circ \mathbf{E}, \quad (2)$$

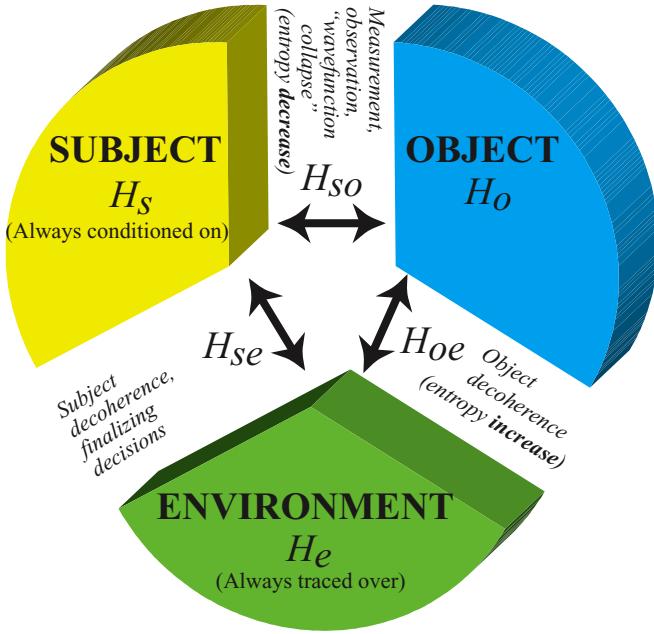


FIG. 1. (Color online) “EOS-decomposition” of the world. The subsystem Hamiltonians,  $H_S$ ,  $H_O$ ,  $H_E$ , and the interaction Hamiltonians,  $H_{SO}$ ,  $H_{OE}$ ,  $H_{SE}$ , can cause qualitatively different effects, providing a unified picture including both observation and decoherence.

where the matrix  $\mathbf{E}$  is defined by  $E_{ij} \equiv \langle \epsilon_j | \epsilon_i \rangle$  and the symbol  $\circ$  denotes what mathematicians know as the *Schur product*. Schur multiplying two matrices simply corresponds to multiplying their corresponding components, i.e.,  $(\rho \circ \mathbf{E})_{ij} = \rho_{ij} E_{ij}$ .

*Object-subject:* If the object is instead probed by an observer, the unitary dynamics of the object-subject system is given by

$$\mathbf{U}|o_i\rangle|s_*\rangle = |o_i\rangle|\sigma_i\rangle, \quad (3)$$

where  $|s_*\rangle$  and  $|\sigma_i\rangle$  denote the initial and final states of the subject for the object state  $|o_i\rangle$ . Let  $|s_k\rangle$  denote the basis states that the subject can perceive, which are robust to decoherence (discussed in Refs. [4] and [13]) and will correspond to “pointer states” [14] for the case of human observer. Since the subject will rapidly decohere, the observer will with probability  $p_k$  find that her state is  $|s_k\rangle$  and that the object density matrix is  $\rho^{(k)}$ , where [4]

$$\rho^{(k)} = \frac{\rho \circ (\mathbf{s}^k \mathbf{s}^{k\dagger})}{p_k}, \quad p_k \equiv \sum_i \rho_{ii} |s_i^k|^2. \quad (4)$$

Here the vector  $\mathbf{s}^k$  is the  $k$ th column of the matrix  $S_{ik} \equiv \langle s_k | \sigma_i \rangle$ , i.e.,  $s_i^k \equiv S_{ik}$  and  $\rho_{ij}^{(k)} = \rho_{ij} S_{ik} S_{jk}^* / p_k$ . Equation (4) can be thought of as the quantum-mechanical version of Bayes’s Theorem [4], showing how the observer’s state of knowledge about a system gets updated by conditioning on new observed information.

These effects of decoherence and observation are intimately related. Since the states  $|s_k\rangle$  form a complete basis, let us define  $F_{ij} = \langle \sigma_j | \sigma_i \rangle = \sum_k \langle \sigma_j | s_k \rangle \langle s_k | \sigma_i \rangle = \sum_k S_{jk}^* S_{ik}$ , i.e.,

$$\mathbf{F} = \mathbf{S}\mathbf{S}^\dagger. \quad (5)$$

Equations (4) and (5) imply that

$$\rho \circ \mathbf{F} = \sum_k p_k \rho^{(k)}. \quad (6)$$

This means that decoherence [Eq. (2)] can be interpreted as an observation that we do not know the outcome of: after being probed by its environment (setting  $\mathbf{F} = \mathbf{E}$ ), the object is in one of the states  $\rho^{(k)}$  (with probability  $p_k$ ), and we simply do not know which.

The most revealing probing is when  $\mathbf{S} = \mathbf{E} = \mathbf{I}$ , the identity matrix: then the system gains complete information about the object, so decoherence makes  $\rho$  diagonal (so-called von Neumann reduction [7]) and observation produces pure states  $\rho^{(k)} = |\tilde{o}_k\rangle\langle\tilde{o}_k|$ , for some orthonormal basis  $|\tilde{o}_k\rangle$ . Another interesting special case is when all elements  $S_{ij}$  are zero or unity. Then the decoherence Eq. (2) reduces to what is known as Lüders projection [15]

$$\rho \circ \mathbf{E} = \sum_k \mathbf{P}_k \rho \mathbf{P}_k, \quad (7)$$

and the observation Eq. (4) gives  $\rho^{(k)} \propto \mathbf{P}_k \rho \mathbf{P}_k$ , where  $\mathbf{P}_k \equiv \sum_i S_{ik} |o_i\rangle\langle o_i|$  are orthogonal projection operators (satisfying  $\mathbf{P}_i \mathbf{P}_j = \delta_{ij} \mathbf{P}_i$ ,  $\sum \mathbf{P}_i = \mathbf{I}$ ). The least revealing probing is the trivial case when  $\rho^{(k)}$  and  $\rho \circ \mathbf{E}$  equal  $\rho$  up to a unitary transformation, so that the subject or environment learns nothing about the object. We define “probing” to exclude this trivial case, which occurs for example when  $\mathbf{S}$  is of the form  $S_{jk} = e^{i(\theta_j + \phi_k)}$  or, for the decoherence case, when  $\rho$  is diagonal.

### III. ENTROPY INEQUALITIES ( $S$ THEOREMS)

We will now prove the main result of this paper: that observation on average decreases entropy, while decoherence increases entropy. Specifically, we will prove the theorem

$$\sum_k p_k S[\rho^{(k)}] \leq S(\rho) \leq S \left[ \sum_k p_k \rho^{(k)} \right], \quad (8)$$

relating the expected entropy after observation (left), the initial entropy (middle), and the entropy after decoherence (right). Both “ $\leq$ ” become “ $<$ ” when the probing is nontrivial. We will refer to these entropy inequalities as two  $S$  theorems. Our proof below holds for the general types of observation from Eq. (4) and decoherence from Eq. (2), and for a very general definition of entropy: for any quantity of the form

$$S(\rho) \equiv \text{tr } h(\rho), \quad (9)$$

where  $h$  is a concave function on the unit interval [ $h''(x) < 0$  for  $0 \leq x \leq 1$ ]. This includes the Shannon-von Neumann entropy [ $h(x) = -x \ln x$ ], the linear entropy [ $h(x) = 1 - x^2$ ], the rescaled exponentiated Renyi entropy [ $H(x) = \pm x^\alpha$ ] and the log-determinant [ $h(x) = \ln x$ ].

In Appendix B, we study interactions that are more general than probing, obtaining the results illustrated in Fig. 2. By providing counterexamples that violate both inequalities in Eq. (8), we prove that for a general positive operator valued measure (POVM) [16], neither of the two  $S$  theorems holds true. We also show that our proof of the left (observation) inequality of Eq. (8) can be generalized from probing to what

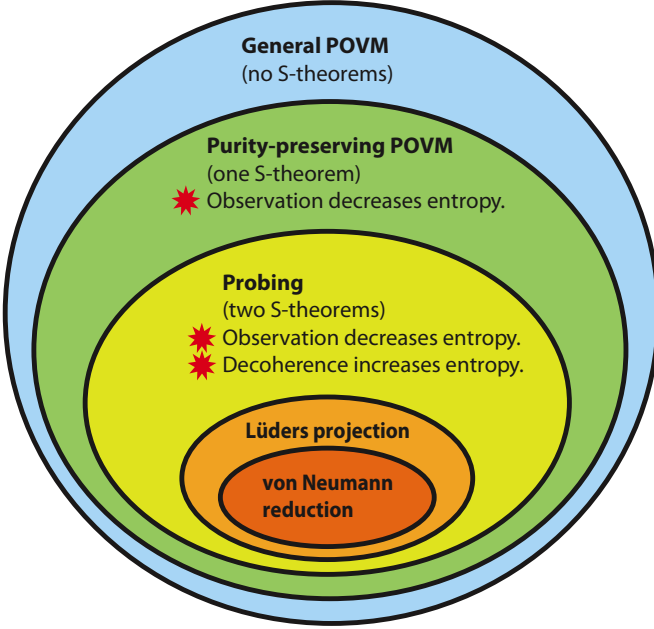


FIG. 2. (Color online) Our two entropy theorems generalize previous results from complete measurements (von Neumann reduction) and projective measurements (Lüders projection) to probing, the most general interaction that can be interpreted as an object measurement. The observation  $S$  theorem further generalizes to purity-preserving POVMs, whereas general POVMs violate both  $S$  theorems.

we term a purity-preserving POVM—“PPPOVM,” the special type of POVM that maps any pure state into pure states. We find that for PPPOVMs, observation on average decreases entropy, but decoherence does *not* always increase entropy. The interactions we defined as probing are simply the subset of PPPOVMs that leave the object unchanged in at least one basis. These are the most general interactions in Fig. 2 that can be interpreted as measurements. To see this, consider that simply replacing the object state by some fixed and *a priori* known pure state (as in counterexample 2 in Appendix B) is a PPPOVM, and it would be ridiculous to view this as a measurement of the object.

### A. Majorization and entropy

Our proof uses numerous inequalities involving the notion of *majorization* [17], which we will now briefly review. One writes

$$\lambda > \mu \quad (10)$$

and says that the vector  $\lambda$  with components  $\lambda_1, \dots, \lambda_n$  *majorizes* the vector  $\mu$  with components  $\mu_1, \dots, \mu_n$  if they have the same sum and

$$\sum_{i=1}^j \lambda_i \geq \sum_{i=1}^j \mu_i \quad \text{for } j = 1, \dots, n, \quad (11)$$

i.e., if the partial sums of the latter never beat the former:  $\lambda_1 \geq \mu_1$ ,  $\lambda_1 + \lambda_2 \geq \mu_1 + \mu_2$ , etc. It is not difficult to show (see, e.g., Ref. [17] or Appendix A of Ref. [4]) that if  $\lambda > \mu$ ,

then

$$\sum_i h(\lambda_i) \leq \sum_i h(\mu_i) \quad (12)$$

for any concave function  $h$ . This means that if the vectors are probability distributions (so that  $\lambda_i \geq 0$ ,  $\sum_i \lambda_i = 1$ ) with entropies defined as

$$S(\lambda) = \sum_i h(\lambda_i), \quad (13)$$

then the majorization  $\lambda > \mu$  implies the entropy inequality  $S(\lambda) \leq S(\mu)$ . Letting  $\lambda(\rho)$  denote the eigenvalues of the density matrix  $\rho$  sorted in decreasing order, and using Eq. (9), we thus have the powerful result

$$\lambda(\rho_1) > \lambda(\rho_2) \Rightarrow S(\rho_1) \leq S(\rho_2), \quad (14)$$

so if two density matrices  $\rho_1$  and  $\rho_2$  satisfy  $\lambda(\rho_1) > \lambda(\rho_2)$ , then they satisfy the entropy inequality  $S(\rho_1) \leq S(\rho_2)$ .

### B. The proof

The right part of the entropy theorem Eq. (8) (that decoherence increases entropy) was proven in Ref. [4]. By Eq. (6), it is equivalent to

$$S(\rho \circ \mathbf{E}) \geq S(\rho), \quad (15)$$

which follows from Eq. (14) and the majorization

$$\lambda(\rho \circ \mathbf{E}) < \lambda(\rho), \quad (16)$$

which is Corollary J.2.a in Ref. [17] [their Eq. (7)], which in turn follows from a 1985 theorem by Bapat and Sunder. To prove the other half of the entropy theorem Eq. (8) (that observation decreases entropy), we will need the following theorem, which is proven in Appendix A: For any Hermitean matrix  $\rho$  and any complete orthogonal set of projection operators  $\mathbf{P}_i$  (satisfying  $\sum_i \mathbf{P}_i = \mathbf{I}$ ,  $\mathbf{P}_i \mathbf{P}_j = \delta_{ij} \mathbf{P}_i$ ,  $\mathbf{P}_i^\dagger = \mathbf{P}_i$ ),

$$\sum_i \lambda(\mathbf{P}_i \rho \mathbf{P}_i) > \lambda(\rho) > \lambda\left(\sum_i \mathbf{P}_i \rho \mathbf{P}_i\right). \quad (17)$$

Applying Eq. (14) to the left part gives

$$\begin{aligned} S(\rho) &\geq S\left[\sum_k \lambda(\mathbf{P}_k \rho \mathbf{P}_k)\right] = S\left[\sum_k p_k \lambda\left(\frac{\mathbf{P}_k \rho \mathbf{P}_k}{p_k}\right)\right] \\ &= \sum_i h\left[\sum_k p_k \lambda_i\left(\frac{\mathbf{P}_k \rho \mathbf{P}_k}{p_k}\right)\right] \\ &\geq \sum_{ik} p_k h\left[\lambda_i\left(\frac{\mathbf{P}_k \rho \mathbf{P}_k}{p_k}\right)\right] = \sum_k p_k S\left(\frac{\mathbf{P}_k \rho \mathbf{P}_k}{p_k}\right), \end{aligned}$$

where we used Jensen’s inequality in the penultimate step. This, Eq. (17), and Eq. (14) thus imply that for any density matrix  $\rho$ ,

$$\sum_i p_i S\left(\frac{\mathbf{P}_i \rho \mathbf{P}_i}{p_i}\right) \leq S(\rho) \leq S\left(\sum_i \mathbf{P}_i \rho \mathbf{P}_i\right), \quad (18)$$

where  $p_i \equiv \text{tr } \mathbf{P}_i \rho \mathbf{P}_i$ . The right half of this double inequality is an alternative proof of Eq. (15) for the special case where decoherence is a Lüders projection as in Eq. (7). Using instead

TABLE I. Summary of the two basic quantum processes discussed in the text.

For observer to predict future, global $\rho$ ...	...is not sufficient	...is not necessary
Mathematical operation	Condition (on subject degrees of freedom)	Marginalize (over environment degrees of freedom)
Interaction	Object-subject	Object-environment
Process	Observation	Decoherence
Dynamics	$\rho_{ij} \mapsto \rho_{ij}^{(k)} = \rho_{ij} \frac{S_{ik} S_{jk}^*}{p_k}$ , $p_k \equiv \sum_i \rho_{ii}  S_{ij} ^2$	$\rho_{ij} \mapsto \rho_{ij} E_{ij}$
Entropy inequality	Decrease: $\sum_k p_k S(\rho^{(k)}) \leq S(\rho)$	Increase: $S(\rho) \leq S(\rho \circ E)$

the left half of Eq. (18), we obtain the following proof of the remaining (left) part of our entropy theorem Eq. (8):

$$\begin{aligned}
 \langle S \rangle &\equiv \sum_k p_k S[\rho^{(k)}] = \sum_k p_k S[\rho^{(k)} \otimes |s_k\rangle\langle s_k|] \\
 &= \sum_k p_k S\left(\frac{\mathbf{P}_k \mathbf{U} \rho^* \mathbf{U} \mathbf{P}_k}{p_k}\right) \leq S(\mathbf{U} \rho^* \mathbf{U}) \\
 &= S(\rho^*) = S(\rho \otimes |s_*\rangle\langle s_*|) = S(\rho). \quad (19)
 \end{aligned}$$

Here  $\rho^* \equiv \rho \otimes |s_*\rangle\langle s_*|$  is the initial state of the combined object-observer system, which the observation process evolves into  $\mathbf{U} \rho^* \mathbf{U}$ , which in turn decoheres into  $\sum_k \mathbf{P}_k \mathbf{U} \rho^* \mathbf{U} \mathbf{P}_k = \sum_k p_k \rho^{(k)} \otimes |s_k\rangle\langle s_k|$  as derived in Ref. [4], where  $\mathbf{P}_k \equiv \mathbf{I} \otimes |s_k\rangle\langle s_k|$  are projection operators acting on the combined object-observer system. The first and last equal signs in Eq. (19) hinge on the fact that tensor multiplying a density matrix by a pure state leaves its entropy unchanged, merely augmenting its spectrum by a number of vanishing eigenvalues. The inequality step uses Eq. (18), and the subsequent step uses the fact that unitary evolution leaves entropy unchanged.

#### IV. CONCLUSIONS

We have proved the entropy inequality Eq. (8), which states that decoherence increases entropy, whereas observation on average decreases it. Its left half is a direct generalization of the Groenewold-Lindblad inequality [18], which corresponds to the special case of the projective (Lüders) form of measurement and the special case of von Neumann entropy; our results also hold Renyi entropy, linear entropy, and indeed any entropy defined by a concave function. Both of these entropy inequalities hold for interactions *probing* the object, defined as the most general interactions leaving the object unchanged in some basis. Of the various classes of interactions we considered, probing constitutes the most general one that can be interpreted as a measurement of the object. We showed that our observation inequality, but not our decoherence inequality, holds also for more general interactions that are purity-reserving POVMs, generalizing Ozawa's result for quasicomplete measurements beyond von Neumann entropy [19]. None of our entropy inequalities hold for arbitrary POVMs.

To prove inequalities Eq. (8), we used the link between unitary quantum mechanics and spectral majorization. This link was first noticed by Uhlmann [20] and later proved extremely helpful in the study of pure state transformations and entanglement [21]. We showed that majorization techniques can also be used to provide simple alternative proofs of entropy inequalities for observation and decoherence. Also, it complements Holevo's inequality [22]  $\sum_k p_k S(\rho_k) \leq S(\rho)$ , where

$\rho = \sum_k p_k \rho_k$ , which follows immediately from Eq. (A2) and Jensen's inequality. It also generalizes Shannon's classical version thereof, which states that observation on average reduces the entropy by an amount equal to the mutual information. The quantum entropy reduction cannot always be that large (which would give negative entropy after observing a qubit with  $\mathbf{S} = \mathbf{I}$  and two bits of mutual information), but we have proved that it is never negative.

Our results complete the formalism of Refs. [4,6] for handling observed open systems. We have seen that quantum statistical mechanics still works flawlessly as long as we avoid sloppy talk of *the* density matrix and *the* entropy: we each have our own personal density matrix encoding everything we know about our object of interest, and we have simple formulas (Table I) for how it changes under observation and decoherence, decreasing and increasing entropy.

#### ACKNOWLEDGMENTS

The authors thank Seth Lloyd, Robert Penna, Benjamin Schumacher, Harold Shapiro, and Wojciech Zurek for helpful discussions. This work was supported by NSF Grant No. AST-1105835 and the MIT UROP program.

#### APPENDIX A: PROOF OF EQ. (17)

Let us first review three useful facts that we will use in our proof. As proven in Ref. [23], any Hermitean matrix  $\mathbf{H}$  written in block form can be decomposed as

$$\mathbf{H} = \begin{bmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{C}^\dagger & \mathbf{B} \end{bmatrix} = \mathbf{U} \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{U}^\dagger + \mathbf{V} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{bmatrix} \mathbf{V}^\dagger \quad (A1)$$

for some unitary matrices  $\mathbf{U}$  and  $\mathbf{V}$ . Second, for any two Hermitean matrices  $\mathbf{A}$  and  $\mathbf{B}$ ,

$$\lambda(\mathbf{A} + \mathbf{B}) < \lambda(\mathbf{A}) + \lambda(\mathbf{B}). \quad (A2)$$

This theorem was suggested and proved by Fan in 1949 and the proof is provided in Ref. [24] as Theorem 10.21. Finally, because the spectrum of a matrix is invariant under unitary transformations, we have

$$\lambda(\mathbf{U} \mathbf{H} \mathbf{U}^\dagger) = \lambda(\mathbf{H}) \quad (A3)$$

for any Hermitean matrix  $\mathbf{H}$  and any unitary matrix  $\mathbf{U}$ .

Combining these three facts, we obtain

$$\begin{aligned}
 \lambda \begin{pmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{C}^\dagger & \mathbf{B} \end{pmatrix} &= \lambda \left[ \mathbf{U} \begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{U}^\dagger + \mathbf{V} \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{pmatrix} \mathbf{V}^\dagger \right] \\
 &< \lambda \left[ \mathbf{U} \begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{U}^\dagger \right] + \lambda \left[ \mathbf{V} \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{pmatrix} \mathbf{V}^\dagger \right] \\
 &= \lambda \begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} + \lambda \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{pmatrix}, \quad (A4)
 \end{aligned}$$

where the three logical steps use Eqs. (A1), (A2), and (A3), respectively.

Now consider a complete orthogonal set of Hermitean projection operators  $\mathbf{P}_i$ ,  $i = 1, \dots, n$ , satisfying the standard relations  $\sum_{i=1}^n \mathbf{P}_i = \mathbf{I}$  and  $\mathbf{P}_i \mathbf{P}_j = \delta_{ij} \mathbf{P}_i$ . Since  $\mathbf{P}_i^2 = \mathbf{P}_i$ , all eigenvalues are either 0 or 1. Since all matrices  $\mathbf{P}_i$  commute, there is basis where they are all diagonal, and where each matrix vanishes except for a block of ones somewhere along the diagonal. In this basis,  $\mathbf{P}_i \mathbf{H} \mathbf{P}_i$  is simply  $\mathbf{H}$  with all elements set to zero except for a corresponding square block. For example, for  $n = 2$  we can write

$$\begin{aligned} \mathbf{P}_1 \begin{pmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{C}^\dagger & \mathbf{B} \end{pmatrix} \mathbf{P}_1 &= \begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \\ \mathbf{P}_2 \begin{pmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{C}^\dagger & \mathbf{B} \end{pmatrix} \mathbf{P}_2 &= \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{pmatrix}. \end{aligned} \quad (\text{A5})$$

This means that we can rewrite the inequality Eq. (A4) as

$$\lambda(\mathbf{H}) < \sum_{i=1}^2 \lambda(\mathbf{P}_i \mathbf{H} \mathbf{P}_i) \quad (\text{A6})$$

for any Hermitean matrix  $\mathbf{H}$ . The sum of any of our two projection operators is a new projection operator, so by iterating Eq. (A6), we can trivially generalize it to the case of arbitrary  $n$ :

$$\lambda(\mathbf{H}) < \sum_{i=1}^n \lambda(\mathbf{P}_i \mathbf{H} \mathbf{P}_i). \quad (\text{A7})$$

This concludes the proof of the left part of the double inequality Eq. (17). The right part follows directly from the Eq. (16) when making the choice

$$\mathbf{E}_{ij} \equiv \sum_{k=1}^n (\mathbf{P}_k)_{ii} (\mathbf{P}_k)_{jj} \quad (\text{A8})$$

in the basis where all projectors are diagonal. In other words,  $\mathbf{E}$  is chosen to be the matrix with ones in all square blocks picked out by the projectors and zeros everywhere else.

## APPENDIX B: POVMs and measurements

In this Appendix, we derive the extent to which our entropy inequalities can be extended to interactions more general than probing. A POVM [16] is a projective measurement on the larger system that has the object as a subsystem. The additional quantum system is typically referred to as an ancilla (A). Specifically, a POVM is a mapping

$$\rho \rightarrow \{p_k, \rho^{(k)}\}, \quad (\text{B1})$$

where the resulting object state is  $\rho^{(k)}$  with probability of outcome  $p_k$ .

### 1. General POVM

$\{p_k, \rho^{(k)}\}$  are given by

$$\rho^{(k)} = \frac{\chi_k(\rho)}{\text{tr } \chi_k(\rho)}, \quad p_k = \text{tr } \chi_k(\rho), \quad (\text{B2})$$

where

$$\chi_k(\rho) = \text{tr}_A [\mathbf{P}_{\text{OA}}^{(k)} (\mathbf{U} \rho \otimes \rho_A \mathbf{U}^\dagger)]. \quad (\text{B3})$$

Here  $A$  denotes the ancilla system in some initial state  $\rho_A$  and  $\{\mathbf{P}_{\text{OA}}^{(k)}\}$  is an orthonormal set of projectors acting on the object-ancilla system. Without loss of generality, we can take the ancilla state to be pure ( $\rho_A = |a^*\rangle\langle a^*|$ ), because for a mixed ancilla state  $\rho_A = \sum_i w_i |a_i\rangle\langle a_i|$ , the purified state  $|a^*\rangle \equiv \sum_i w_i^{1/2} |a_i\rangle \otimes |a_i\rangle$  defines the same POVM.

For a general POVM, neither of the two inequalities in the Eq. (8) holds true, so we have no  $S$  theorems for general POVMs. We will now prove this by providing POVM counterexamples that violate both the left (observation) and right (decoherence) inequalities in Eq. (8).

(i) *Counterexample 1:* The object and ancilla are single qubits in a state  $|0\rangle_O |0\rangle_A$  and the POVM is defined by  $\mathbf{U} = \mathbf{I}$  and  $\mathbf{P}_{\text{AO}}^{(k)} = |\Psi_k\rangle\langle\Psi_k|$ , where

$$|\Psi_1\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}, \quad |\Psi_2\rangle = \frac{|00\rangle - |11\rangle}{\sqrt{2}} \quad (\text{B4})$$

$$|\Psi_3\rangle = \frac{|01\rangle + |10\rangle}{\sqrt{2}}, \quad |\Psi_4\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}. \quad (\text{B5})$$

This gives

$$\rho^{(1)} = \rho^{(2)} = \rho^{(3)} = \rho^{(4)} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}, \quad (\text{B6})$$

$$p_1 = p_2 = \frac{1}{2}, \quad p_3 = p_4 = 0, \quad (\text{B7})$$

so that the initial von Neumann entropy  $S(\rho) = 0$  increases to a final entropy of 1 bit, violating the left (observation) part of Eq. (8).

(ii) *Counterexample 2:* The object and ancilla are single qubits in initial states

$$\rho = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \quad \text{and} \quad \rho_A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad (\text{B8})$$

and  $\mathbf{U}$  is the unitary two-qubit operation that exchanges the states of these qubits (a swap gate):

$$\mathbf{U} \rho \otimes |0\rangle\langle 0| \mathbf{U}^\dagger = |0\rangle\langle 0| \otimes \rho. \quad (\text{B9})$$

The projections  $\mathbf{P}_{\text{OA}}^{(k)}$  are

$$\mathbf{P}_{\text{OA}}^{(1)} = \mathbf{I} \otimes |0\rangle\langle 0|, \quad (\text{B10})$$

$$\mathbf{P}_{\text{OA}}^{(2)} = \mathbf{I} \otimes |1\rangle\langle 1|, \quad (\text{B11})$$

giving

$$\rho^{(1)} = \rho^{(2)} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad (\text{B12})$$

$$p_1 = p_2 = \frac{1}{2},$$

so that the initial entropy  $S(\rho) = 1$  bit drops to zero, violating the right (decoherence) side of the inequality Eq. (8).

### 2. Purity-preserving POVM

We define a PPOVM as a POVM that keeps a pure state pure, i.e., if  $\rho$  is a pure state, then  $\rho^{(k)}$  is pure for all  $k$ . A

POVM is purity preserving if and only if

$$\mathbf{P}_{\text{OA}}^{(k)} = \mathbf{I} \otimes |a_k\rangle\langle a_k|, \quad (\text{B13})$$

where  $|a_k\rangle$  is an orthonormal basis of the ancilla system. To show this, we need to prove both the “if” and “only if” parts.

(i) Part 1 (“if”): For a pure object state  $\rho = |\psi\rangle\langle\psi|$ , the pure total state

$$|\Psi\rangle^{\text{OA}} \equiv \mathbf{U}|\psi\rangle|a^*\rangle \quad (\text{B14})$$

can always be decomposed as

$$|\Psi\rangle^{\text{OA}} = \sum_k \lambda_k |\psi_k\rangle|a_k\rangle, \quad (\text{B15})$$

where  $|\psi_k\rangle$  are a normalized object states and  $\lambda_k$  are complex numbers. If  $\mathbf{P}_{\text{OA}}^{(k)} = \mathbf{I} \otimes |a_k\rangle\langle a_k|$ , then this decomposition gives

$$\chi_k(\rho) = |\lambda_k|^2 |\psi_k\rangle\langle\psi_k|. \quad (\text{B16})$$

Since the state  $\rho_k = |\psi_k\rangle\langle\psi_k|$  is pure for all  $k$ , we have shown that if  $\mathbf{P}_{\text{OA}}^{(k)} = \mathbf{I} \otimes |a_k\rangle\langle a_k|$ , then the POVM is purity preserving.

(ii) Part 2 (“only if”): If the POVM is purity preserving, then we can write  $\rho = |\psi\rangle\langle\psi|$  and  $\rho_k = |\psi_k\rangle\langle\psi_k|$ , so we have

$$\chi_k(\rho) = \text{tr}_A(\mathbf{P}_{\text{OA}}^{(k)}|\Psi\rangle^{\text{OA}}\langle\Psi|^{\text{OA}}) = p_k |\psi_k\rangle\langle\psi_k|. \quad (\text{B17})$$

This means that after normalization, the state  $\mathbf{P}_{\text{OA}}^{(k)}|\Psi\rangle^{\text{OA}}$  must be pure and separable for all  $k$ :

$$\mathbf{P}_{\text{OA}}^{(k)}|\Psi\rangle^{\text{OA}} = \lambda_k |\psi_k\rangle|a_k\rangle, \quad (\text{B18})$$

where  $|\lambda_k|^2 = p_k$  and  $|a_k\rangle$  is some state of the ancilla system. We can now express  $\mathbf{P}_{\text{OA}}^{(k)}$  as

$$\mathbf{P}_{\text{OA}}^{(k)} = \mathbf{I} \otimes |a_k\rangle\langle a_k|, \quad (\text{B19})$$

and because these terms form an orthonormal set of projectors, we conclude that  $|a_k\rangle$  form an orthonormal basis for ancilla. We have thus shown that if the POVM is purity preserving, then  $\mathbf{P}_{\text{OA}}^{(k)} = \mathbf{I} \otimes |a_k\rangle\langle a_k|$  for some orthonormal basis  $|a_k\rangle$ .

In Ref. [19], Ozawa referred to PPPOVMs as “quasicomplete.” For the special case of von Neumann entropy for one-dimensional position measurements, he showed that the left inequality in Eq. (8) is true if and only if a measurement is quasicomplete. As we will now show, our proof of the left part of Eq. (8) is readily generalized to PPPOVMs for our much more general definition of entropy, and without use of strong subadditivity. Here the ancilla plays the role of the subject and projections on the object-ancilla system is given by Eq. (B13), which are equivalent to the projectors  $\mathbf{P}_k = \mathbf{I} \otimes |s_k\rangle\langle s_k|$  used in Eqs. (17) and (19). Because the key Eqs. (17) and (19) hold for arbitrary  $\mathbf{U}$ , the proof for the left (observation) part of Eq. (8) can be immediately generalized to PPPOVMs. In contrary, the right part of Eq. (8) does not hold for PPPOVMs, which can be seen using the counterexample 2 from the previous section. So only one of the  $S$  theorems holds for PPPOVMs: observation always decreases average entropy, but decoherence does not always increase it.

### 3. Probing

Probing is a special case of a PPPOVM, where the role of the ancilla is played by either the subject or the environment, and the unitary dynamics  $\mathbf{U}$  of the object-ancilla system leaves the object unchanged in some basis, merely changing the ancilla state in a way that depends on the object state  $|o_i\rangle$ :

$$\mathbf{U}|o_i\rangle|a^*\rangle = |o_i\rangle|a_i\rangle. \quad (\text{B20})$$

In this paper, we have shown that both inequalities in Eq. (8) hold true for probing, so that observation always decreases average entropy and decoherence always increases it.

- 
- [1] G. Bunin, L. D’Alessio, Y. Kafri, and A. Polkovnikov, *Nat. Phys.* **7**, 913 (2011).
  - [2] P. Skrzypczyk, A. J. Short, and S. Popescu (2013), [arXiv:1307.1558](https://arxiv.org/abs/1307.1558).
  - [3] J. L. England, *J. Chem. Phys.* **139**, 121923 (2013).
  - [4] M. Tegmark, *Phys. Rev. D* **85**, 123517 (2012).
  - [5] S. Lloyd, Ph.D. thesis, Rockefeller University, 1988.
  - [6] M. Tegmark, *Phys. Rev. E* **61**, 4194 (2000).
  - [7] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, NJ, 1996), translated from German.
  - [8] R. Feynman, *Statistical Mechanics: A Set of Lectures*, Advanced Books Classics Series (Westview Press, Boulder, CO, 1998).
  - [9] H. Zeh, *Found. Phys.* **1**, 69 (1970).
  - [10] D. Giulini *et al.*, *Decoherence and the Appearance of a Classical World in Quantum Theory* (Springer, Berlin, 1996).
  - [11] W. H. Zurek, *Nat. Phys.* **5**, 181 (2009).
  - [12] M. Schlosshauer, *Decoherence and the Quantum-to-Classical Transition*, The Frontiers Collection (Baker & Taylor, Charlotte, NC, 2007).
  - [13] M. Tegmark (2014), [arXiv:1401.1219](https://arxiv.org/abs/1401.1219).
  - [14] W. H. Zurek, S. Habib, and J. P. Paz, *Phys. Rev. Lett.* **70**, 1187 (1993).
  - [15] G. Lüders, *Annalen der Physik* **443**, 322 (1950).
  - [16] J. Preskill, *Lecture Notes for Physics 229: Quantum Information and Computation* (California Institute of Technology, 1998).
  - [17] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: Theory of Majorization and its Applications* (Springer Science + Business Media, LLC, New York, 2011).
  - [18] G. Lindblad, *Commun. Math. Phys.* **28**, 245 (1972).
  - [19] M. Ozawa, *J. Math. Phys.* **27**, 759 (1986).
  - [20] A. Uhlmann, *Math. Nat. R.* **20**, 633 (1971).
  - [21] M. A. Nielsen, *Phys. Rev. Lett.* **83**, 436 (1999).
  - [22] A. S. Holevo, *Probl. Peredachi Inf.* **9**, (1973).
  - [23] J.-C. Bourin and E.-Y. Lee, *Bull. London Math. Soc.* **44**, 1085 (2012).
  - [24] F. Zhang, *Matrix Theory: Basic Results and Techniques* (Springer, Berlin, 2011).