Entanglement and thermodynamics in general probabilistic theories

Giulio Chiribella and Carlo Maria Scandolo
Center for Quantum Information, Institute for Interdisciplinary Information Sciences, Tsinghua University, Beijing 100084, China

Entanglement is one of the most striking features of quantum mechanics, and yet it is not specifically quantum. It shows up in a number of alternative theories, such as the theory of nonlocal boxes, where it provides a powerful information-theoretic resource. Here we explore the ways in which this resource can be manipulated in a general physical theory, searching for operational links between transformations of entanglement and thermodynamic transformations. We first address the question whether an entangled state can be transformed into another by means of local operations and classical communication (LOCC). Under two operational requirements, we prove an analogue of the Lo-Popescu theorem, stating that every LOCC protocol acting on a pure bipartite state can be simulated by a protocol that uses a single round of classical communication. We then introduce an ordering of single-system states according to their degree of mixedness, providing an operational version of the majorization criterion. Based on this notion, we show that the degree of entanglement of a pure bipartite state coincides with the degree of mixedness of its marginals in every physical theory where all processes arise from pure states and reversible interactions at the fundamental level.

I. INTRODUCTION

The discovery of quantum entanglement [1, 2] introduced the revolutionary idea that a composite system can be in pure state while its components are in mixed states. In Schrödinger’s words: “maximal knowledge of a total system does not necessarily imply maximal knowledge of all its parts” [2]. This new possibility, in radical contrast with the paradigm of classical physics, is at the root of quantum non-locality [3–6] in all its counterintuitive manifestations [7–13]. With the advent of quantum information, it quickly became clear that entanglement was not only a source of foundational puzzles, but also a resource [14]. Harnessing this resource has been the key to the invention of seminal protocols such as quantum teleportation [15], dense coding [16], and secure quantum key distribution [17, 18], whose implications deeply impacted physics and computer science [19, 20].

The key to understand entanglement as a resource is to consider distributed scenarios where spatially-separated parties perform local operations (LO) in their laboratories and exchange classical communication (CC) from one laboratory to another [21, 22]. The protocols that can be implemented in this scenario, known as LOCC protocols, provide a means to characterize entangled states and to compare their degree of entanglement. Precisely, a state is i) entangled if it cannot be generated by an LOCC protocol, and ii) more entangled than another if there exists an LOCC protocol that transforms the former into the latter.

Comparing the degree of entanglement of two quantum states is generally a hard problem [24–30]. Nevertheless, the solution is simple for pure bipartite states, where the majorization criterion [31] provides a necessary and sufficient condition for LOCC convertibility. Essentially, the criterion identifies the degree of entanglement of a bipartite system with the degree of mixedness of its parts: the more entangled a pure bipartite state, the more mixed its marginals. Mixed states are compared here according to their spectra, with a state being more mixed than another if the spectrum of the latter majorizes the spectrum of the former [32–35].

The majorization criterion shows that for pure bipartite states the notion of entanglement as a resource beautifully matches Schrödinger’s idea of entanglement as non-maximal knowledge about the parts of a composite system. Moreover, it establishes an intriguing duality between entanglement and thermodynamics [36–40], where the reduction of entanglement caused by LOCC protocols becomes dual to the increase of mixedness (and therefore entropy [41]) caused by thermodynamic transformations. This duality has far-reaching consequences, such as the existence of a unique measure of entanglement in the asymptotic limit [36–37, 42]. In addition, it has provided guidance for the development of entanglement theory beyond the simple case of pure bipartite states [39].

The duality between entanglement and thermodynamics is a profound and fundamental fact. As such, one might expect it to follow directly from basic principles. However, it is not a priori clear what these principle are: up to now, the relation between entanglement and thermodynamics has been addressed in a way that is heavily dependent on the Hilbert space framework, using technical results that lack an operational interpretation (such as, e.g. the singular value decomposition). It is then natural to search for a derivation of the entanglement-thermodynamics duality that uses only high-level quantum features, such as the impossibility of instantaneous signalling or the no-cloning theorem. In the same vein, one can ask whether the duality holds for physical theo-
The current page contains a detailed discussion on general probabilistic theories (GPTs) and their implications on thermodynamics. The text explores the consequences of the No Entropy Sinks requirement and the role of entanglement and local exchangeability in these theories. It also touches upon the relation of these concepts with quantum mechanics and information-theoretic tasks.

In the GPT framework, processes are described in terms of states and transformations, which can be composed in sequence, in parallel, or both. The framework is particularly useful for analyzing entangled states and their purification, which is a key concept in quantum mechanics.

The paper is organized as follows: In section II, the framework adopted in the paper, including the operational version of the Lo-Popescu theorem, is introduced. Section III focuses on the entanglement of pure bipartite states and establishes a duality between entanglement and purity. Section IV provides a generalized Lo-Popescu theorem, which is used in section V to discuss operational aspects.

Section VI proves the duality between entanglement and thermodynamics, and section VII addresses the issue of Symmetric Purification. Finally, section VIII draws conclusions from the established results.

The existence of symmetric purifications is a key feature of the quantum resource theory of purity. This theory is further generalized in section IX to non-relativistic theories, providing a natural framework for studying information-theoretic tasks.

In summary, the paper aims to connect the concepts of entanglement and thermodynamics within the GPT framework, highlighting the importance of symmetric purification and the No Entropy Sinks requirement. These concepts are crucial for understanding the potential of GPTs in various information-theoretic tasks.
a single outcome. We will often refer to deterministic transformations as channels, following the terminology of quantum information. Among all possible channels, reversible channels are particularly important: a channel \( \mathcal{U} \) from A to B is called reversible if there exists a channel \( \mathcal{U}^{-1} \) from B to A such that \( \mathcal{U}^{-1}\mathcal{U} = \mathcal{I}_A \) and \( \mathcal{U}\mathcal{U}^{-1} = \mathcal{I}_B \), where \( \mathcal{I}_X \) is the identity channel on system X. If there exists a reversible channel transforming A into B we say that A and B are operationally equivalent, denoted by \( A \simeq B \). The composition of systems is required to be symmetric \([18, 41]\), meaning that \( A \otimes B \simeq B \otimes A \). The reversible channel that implements the equivalence is the swap channel, \( \text{SWAP} \), which satisfies the condition

\[
\begin{array}{cccc}
\text{B} & \text{E} & \text{B}' \\
\text{A} & \text{A}'
\end{array} = \begin{array}{cccc}
\text{B} & \text{E} & \text{B}' \\
\text{A} & \text{SWAP} & \text{A}' & \text{B}' & \text{SWAP} & \text{A}'
\end{array},
\]  

(1)

for every pair of transformations \( \mathcal{A} \) and \( \mathcal{B} \) and for generic systems \( A, A', B, B' \).

In this paper we restrict our attention only to causal theories \([77]\), where the choice of future measurement settings does not influence the outcome probability of present experiments. Mathematically, causality is equivalent to the fact that for every system A there is only one deterministic effect, which we denote here by \( \text{Tr}_A \), in analogy with the partial trace in quantum mechanics. In a causal theory, one can define the norm of a state \( \rho \) as

\[
\|\rho\| := \text{Tr} \rho.
\]

The set of normalized states of A will be denoted by

\[
\text{St}_1 (A) := \{ \rho \in \text{St}(A) \mid \|\rho\| = 1 \}.
\]

In a causal theory, every state is proportional to a normalized state \([47]\). In quantum mechanics, \( \text{St}_1 (A) \) is the set of normalized density matrices of system A, while \( \text{St}(A) \) is the set of all sub-normalized density matrices.

In a causal theory channels admit a simple characterization, which will be useful later in the paper:

**Proposition 1** Let \( \mathcal{C} \in \text{Transf}(A, B) \). \( \mathcal{C} \) is a channel if and only if \( \text{Tr}_B \mathcal{C} = \text{Tr}_A \).

The proof can be found in lemma 5 of Ref. \([77]\).

### A. Pure states and transformations

In every probabilistic theory one can define pure states, and, more generally, pure transformations. Both concepts are based on the notion of coarse-graining, i.e. the operation of joining two or more outcomes of a test. More precisely, a test \( \{ \mathcal{C}_i \}_{i \in \mathcal{X}} \) is a coarse-graining of the test \( \{ \mathcal{D}_j \}_{j \in \mathcal{Y}} \) if there is a partition \( \{ \mathcal{Y}_i \}_{i \in \mathcal{X}} \) of \( \mathcal{Y} \) such that \( \mathcal{C}_i = \sum_{j \in \mathcal{Y}_i} \mathcal{D}_j \) for every \( i \in \mathcal{X} \). In this case, we say that \( \{ \mathcal{D}_j \}_{j \in \mathcal{Y}} \) is a refinement of \( \{ \mathcal{C}_i \}_{i \in \mathcal{X}} \). The refinement of a given transformation is defined via the refinement of a test: if \( \{ \mathcal{D}_j \}_{j \in \mathcal{Y}} \) is a refinement of \( \{ \mathcal{C}_i \}_{i \in \mathcal{X}} \), then the transformations \( \{ \mathcal{D}_j \}_{j \in \mathcal{Y}} \) are a refinement of the transformation \( \mathcal{C}_i \).

A transformation is called pure if it has only trivial refinements:

**Definition 1** The transformation \( \mathcal{C} \in \text{Transf}(A, B) \) is pure if for every refinement \( \{ \mathcal{D}_j \}_{j \in \mathcal{Y}} \) one has \( \mathcal{D}_j = p_j \mathcal{C} \), where \( \{ p_j \} \) is a probability distribution.

Pure transformations are those for which the experimenter has maximal information about the evolution of the system. In the special case of states (transformations with no input), the above definition coincides with the usual definition of pure state. We denote the set of pure states of system A as \( \text{PurSt}(A) \). As usual, non-pure states are called mixed.

Pure states will play a key role in this paper. An elementary property of pure states is that they are preserved by reversible transformations.

**Proposition 2** Let \( \mathcal{U} \in \text{Transf}(A, B) \) be a reversible channel. Then a state \( \psi \in \text{St}(A) \) is pure if and only if the state \( \mathcal{U} \psi \in \text{St}(B) \) is pure.

The proof is standard and is reported in Appendix A for convenience of the reader.

### III. Entanglement

#### A. The resource theory of entanglement

The resource theory of quantum entanglement \([14]\) is based on the notion of LOCC protocols, that is, protocols in which distant parties are allowed to communicate classically to one another and to perform local operations in their laboratories \([21, 22]\). This is an operational notion and, as such, can be easily exported to arbitrary theories.

In this paper we will consider protocols involving two parties, Alice and Bob. A generic LOCC protocol consists of a sequence of tests, performed by Alice and Bob, with the property that the choice of the test at a given step can depend on the outcomes produced at the previous steps. For example, consider a two-way protocol where

1. Alice performs a test \( \{ \mathcal{A}_1 \} \) and communicates the outcome to Bob
2. Bob performs a test \( \{ \mathcal{B}^{(1)}_2 \} \) and communicates the outcome to Alice
3. Alice performs a test \( \{ \mathcal{A}^{(1,2)}_3 \} \).

Every instance of the protocol will correspond to a sequence of outcomes, which can be represented by a circuit of the form

\[
\begin{array}{cccc}
A_0 & A_1 & A_2 \\
\text{A}_1 & \text{A}_1 & \text{A}_1 \\
\text{B}_1 & \text{B}_1 & \text{B}_1
\end{array}
\]
where the dashed lines represent classical communication. By coarse-graining over all possible outcomes, one obtains a channel, given by

\[ \mathcal{L} = \sum_{i_1, i_2, i_3} \left[ A_{i_3, i_2} A_{i_1} \otimes B_{i_2, i_3} \right]. \]

Entangled states are those states that cannot be generated using an LOCC protocol. Equivalently, they can be characterized as those states that are not separable, i.e., they are not of the form

\[ \rho = \sum_i p_i \alpha^{(i)} \otimes \beta^{(i)}, \]

where \( \{p_i\} \) is a probability distribution allowed by the theory, \( \alpha^{(i)} \) is a state of A, and \( \beta^{(i)} \) is a state of B.

Like in quantum theory, LOCC protocols can be used to compare entangled states.

**Definition 2** Given two states \( \rho \in \text{St}(A \otimes B) \) and \( \rho' \in \text{St}(A' \otimes B') \), we say that \( \rho \) is more entangled than \( \rho' \), denoted by \( \rho \succeq_{\text{ent}} \rho' \), if there exists an LOCC protocol that transforms \( \rho \) into \( \rho' \), i.e., if \( \rho' = \mathcal{L} \rho \) for some LOCC channel \( \mathcal{L} \).

Mathematically, the relation \( \succeq_{\text{ent}} \) is a preorder, i.e., it is reflexive and transitive. Moreover, it is stable under tensor products, namely \( \rho \otimes \sigma \succeq_{\text{ent}} \rho' \otimes \sigma' \) whenever \( \rho \succeq_{\text{ent}} \rho' \) and \( \sigma \succeq_{\text{ent}} \sigma' \). In other words, the relation \( \succeq_{\text{ent}} \) turns the set of all bipartite states into a preordered monoid, a mathematical structure arising in all resource theories \[64\]. The resource theory of entanglement fits completely into the framework of Ref. \[64\], with the LOCC channels as free operations. The states which can be prepared by LOCC (i.e., the separable states) are free by definition, and all the other states represent resources. If a state can be converted into another by free operations, then it constitutes a more valuable resource. If \( \rho \succeq_{\text{ent}} \rho' \) and \( \rho' \succeq_{\text{ent}} \rho \), then we say that \( \rho \) and \( \rho' \) are equally entangled, denoted by \( \rho \sim_{\text{ent}} \rho' \). Note that \( \rho \sim_{\text{ent}} \rho' \) does not imply that \( \rho \) and \( \rho' \) are equal; for example, every two separable states are equally entangled. Similarly, every two states that differ by a local reversible channel are equally entangled: this is the case, e.g., of two pure states \( \Psi \neq \Psi' \) such that

\[ \begin{array}{ccc}
\Psi'\quad & A & \quad \Psi \\
\downarrow & \quad \Uparrow & \quad \downarrow \\
\Psi & B & \quad \Psi
\end{array} \]

for some reversible channel \( \Uparrow \) on \( A \) [recall, indeed, that reversible transformations send pure states into pure states, cf. lemma \[2\]].

IV. AN OPERATIONAL LO-POPESCU THEOREM

Given two bipartite states, it is natural to ask whether one is more entangled than the other. A priori, answering the question requires one to check all possible LOCC protocols, which in general is a hard task. However, the situation is much simpler when the initial state is pure. Now we prove that in this case every LOCC protocol can be replaced without loss of generality by a protocol involving only one round of classical communication—i.e., a one-way LOCC protocol. Our argument provides an operational version of the Lo-Popescu theorem \[23\], which lays the foundation for the quantum theory of pure-state entanglement.

A. Two operational requirements

Our derivation of the operational Lo-Popescu theorem is based on two physical requirements. The first is

**Axiom 1 (Purity Preservation \[45, 48, 72\])** The sequential and parallel composition of two pure transformations yield a pure transformation.

As a special case, Purity Preservation implies that the product of two pure states is a pure state:

\[ \alpha \in \text{PurSt}(A), \beta \in \text{PurSt}(B) \implies \alpha \otimes \beta \in \text{PurSt}(A \otimes B). \]

Note that this conclusion could also be obtained from the Local Tomography axiom \[47\]. Nevertheless, there exist examples of theories that satisfy Purity Preservation and violate Local Tomography. An example is quantum theory on real vector spaces \[23, 74\]. In general, we regard Purity Preservation as a fundamental requirement, even more fundamental than Local Tomography. Considering the theory as an algorithm to make deductions about physical processes, Purity Preservation ensures that, when presented with maximal information about two processes, the algorithm outputs maximal information about their composition.

Our second requirement imposes a symmetry on pure bipartite states.

**Axiom 2 (Local Exchangeability)** For every pure bipartite state \( \Psi \in \text{PurSt}(A \otimes B) \), there exist two channels \( \mathcal{C} \in \text{Transf}(A, B) \) and \( \mathcal{D} \in \text{Transf}(B, A) \) such that

\[ \begin{array}{ccc}
\Psi & \mathcal{C} & \mathcal{D} \\
\downarrow & \quad \Uparrow & \quad \downarrow \\
\Psi & \mathcal{D} & \mathcal{C}
\end{array} = \begin{array}{ccc}
\Psi & \quad \Uparrow \quad \mathcal{C} & \quad \mathcal{D} \\
\downarrow & \quad \Uparrow & \quad \downarrow \\
\Psi & \quad \mathcal{D} & \quad \mathcal{C}
\end{array}, \]

where \( \text{SWAP} \) is the swap operation [cf. Eq. \([1]\)].

This requirement is trivially satisfied by classical probability theory, where all pure states are of the product form. Less trivially, it is satisfied by quantum theory, on both complex and real Hilbert spaces.

**Example 1** Suppose that \( A \) and \( B \) are quantum systems, and let \( \mathcal{H}_A \) and \( \mathcal{H}_B \) be the corresponding Hilbert spaces. By the Schmidt decomposition, every pure state in the tensor product Hilbert space can be written as

\[ |\Psi\rangle = \sum_{i=1}^r \sqrt{p_i} |\alpha_i\rangle \otimes |\beta_i\rangle, \]

where the Schmidt coefficients \( p_i \) satisfy \( \sum_i p_i = 1 \).
where $\{|\alpha_i\rangle\rangle_{i=1}^r \subset \mathcal{H}_A$ and $\{|\beta_i\rangle\rangle_{i=1}^r \subset \mathcal{H}_B$ are orthonormal vectors. This implies the relation

$$\text{SWAP} |\Psi\rangle = (C \otimes D) |\Psi\rangle,$$

where $C$ and $D$ are the partial isometries $C := \sum_{i=1}^r |\alpha_i\rangle\langle \alpha_i|$ and $D := \sum_{i=1}^r |\beta_i\rangle\langle \beta_i|$. From the partial isometries $C$ and $D$ it is immediate to construct the desired channels $C$ and $D$, which can be defined as

$$C(\rho) := C \rho C^\dagger + \sqrt{1 - C^\dagger C} \rho \sqrt{1 - C^\dagger C}$$

$$D(\sigma) := D \sigma D^\dagger + \sqrt{1 - D^\dagger D} \sigma \sqrt{1 - D^\dagger D},$$

where $\rho$ and $\sigma$ are generic input states of systems $A$ and $B$, respectively. With this definition, one has

$$(C \otimes D)(|\Psi\rangle\langle \Psi|) = \text{SWAP} |\Psi\rangle\langle \Psi| \text{SWAP},$$

which is the Hilbert space version of the local exchangeability condition of Eq. (2).

Note that Local Exchangeability does not violate the no-signalling principle. In order to exchange their systems, Alice and Bob need to know in advance what pure state they share, because the two channels $C$ and $D$ depend on the state $\Psi$.

Local Exchangeability is also satisfied by all the extreme bipartite non-local correlations characterized in the literature [59–62].

**Example 2** Consider a scenario where two space-like separated parties, Alice and Bob, perform measurements on a pair of systems, $A$ and $B$, respectively. Let $x$ (resp. $y$) be the index labeling Alice’s (resp. Bob’s) measurement setting and let $a$ (resp. $b$) the index labeling the outcome of the measurement done by Alice (resp. Bob). In the theory known as box world [64, 72] all no-signaling probability distributions $p_{ab|xy}$ are physically realizable and represent states of the composite system $A \otimes B$. Such probability distributions form a convex polytope [61] and the extreme points are the pure states of the theory.

For $x, y, a, b \in \{0, 1\}$ the systems $A$ and $B$ are operationally equivalent, and we will denote by $\mathbb{I}$ the reversible transformation that converts $A$ into $B$. The extreme non-local correlations have been characterized in [59] and are known to be equal to the standard PR-box correlation [60]

$$p_{ab|xy} = \begin{cases} \frac{1}{2} & a + b \equiv xy \mod 2 \\ 0 & \text{otherwise} \end{cases}$$

(3)

up to exchange of 0 with 1 in the local settings of Alice and Bob and in the outcomes of their measurements. In the circuit picture, these operations are described by local reversible transformations: denoting by $\Phi$ the standard PR-box state, one has that every other pure state $\Psi \in \text{PurSt} (A \otimes B)$ is of the form

$$\Psi_{AB} = \Phi_{BA} U_{AB} \Gamma_{AB} ,$$

where $\mathcal{U}$ and $\mathcal{V}$ are reversible transformations.

To see that the Local Exchangeability property holds, note that swapping systems $A$ and $B$ is equivalent to swapping $x$ with $y$ and $a$ with $b$. Now, the standard PR-box correlation of Eq. (3) is invariant under the swappings $x \leftrightarrow y$, and $a \leftrightarrow b$, meaning that one has

$$\Phi_{AB} \rightarrow \Phi_{BA} ,$$

Then it is clear that every pure state of $A \otimes B$ can be swapped by local operations: defining $\mathcal{C} := \mathcal{V} \mathcal{U}^{-1}$ and $\mathcal{D} := \mathcal{U} \mathcal{V}^{-1}$, one immediately obtains

$$\Psi_{AB} \rightarrow \Psi_{BA} ,$$

where $\mathcal{C} := \mathcal{V} \mathcal{I} \mathcal{U}^{-1}$ and $\mathcal{D} := \mathcal{U} \mathcal{I}^{-1} \mathcal{V}^{-1}$. This proves the Local Exchangeability property for all pure bipartite states in the 2-setting 2-outcome scenario.

The situation is analogous in the case of 2 settings and arbitrary number of outcomes. If $x, y \in \{0, 1\}$, and $a$ and $b$ take $d_{A}$ values and $c$ and $d_{B}$ values, all extreme non-local correlations are characterized in Ref. [61]. Up to local reversible transformations, they are labeled by a parameter $k \in \{2, \ldots, \min\{d_{A}, d_{B}\}\}$ and they are such that

$$p_{ab|xy} = \begin{cases} \frac{1}{k} & b - a \equiv xy \mod k \\ 0 & \text{otherwise} \end{cases}$$

(5)

Thanks to the local equivalence, it is enough to prove the validity of Local Exchangeability for correlations in the standard form of Eq. (5). We distinguish between the two cases $xy = 0$ and $xy = 1$. For $xy = 0$, swapping $x$ with $y$ and $a$ with $b$ has no effect on $p_{ab|xy}$. For $xy = 1$, by swapping $x$ with $y$ and $a$ with $b$, one obtains the probability distribution

$$p_{ab|xy}' = \begin{cases} \frac{1}{k} & a - b \equiv 1 \mod k \\ 0 & \text{otherwise} \end{cases}$$

This probability distribution can be obtained from the original one by relabeling the outputs as $a' := k - a$ and $b' := k - b$. Such relabeling corresponds to local reversible operations on $A$ and $B$. In other words, Local Exchangeability holds.
Finally, the last category of extreme non-local correlations characterized in the literature corresponds to the case of arbitrary number of settings and to 2-outcome measurements. In this case, the extreme correlations are characterized explicitly in Ref. [22]. Up to local reversible transformations, the pure states are invariant under swap. Hence, the same argument used in Eq. (4) shows that Local Exchangeability holds.

B. Inverting the direction of classical communication

Purity Preservation and Local Exchangeability have an important consequence. For one-way protocols acting on a pure input state, the direction of classical communication is irrelevant: every one-way LOCC protocol with communication from Alice to Bob can be replaced by a one-way LOCC protocol with communication from Bob to Alice, as shown by the following.

Lemma 1 (Inverting CC) Let $\Psi$ be a pure state of $A \otimes B$ and let $\rho'$ be a (possibly mixed) state of $A' \otimes B'$. Under the validity of axioms [1] and [2] the following are equivalent:

1. $\Psi$ can be transformed into $\rho'$ by a one-way LOCC protocol with communication from Alice to Bob

2. $\Psi$ can be transformed into $\rho'$ by a one-way LOCC protocol with communication from Bob to Alice.

Proof. Suppose that $\Psi$ can be transformed into $\rho'$ by one-way LOCC protocol with communication from Alice to Bob, namely

$$\rho'_{A'B'} = \sum_{i \in X} \Psi_{A'} \otimes B' \otimes A_i \otimes B_i$$

where $\{A_i\}_{i \in X}$ is a test, and, for every outcome $i \in X$, $B_i$ is a channel. Note that one can assume every transformation $A_i$ to be pure without loss of generality.

For every fixed $i \in X$, one has

$$\rho'_{A'B'} = \sum_{i \in X} \Psi_{A'} \otimes B' \otimes A_i \otimes B_i$$

having defined $A_i := D(i) C(i) A_i D(i)$. By construction $\{B_i\}_{i \in X}$ is a test, because it can be realized by performing the test $\{A_i\}$ after the channel $D$ and subsequently applying the channel $C(i) B_i D(i) C(i)$, depending on the outcome $i$. On the other hand, $A(i)$ is a channel for every $i \in X$. Hence, we constructed a one-way LOCC protocol with communication from Bob to Alice. Combining Eqs. (6) and (7) we obtain

$$\rho'_{A'B'} = \sum_{i \in X} \Psi_{A'} \otimes B' \otimes A_i$$

meaning that $\Psi$ can be transformed into $\rho'$ by a one-way protocol with communication from Bob to Alice. Clearly, the same argument can be applied to prove the converse direction.

Note that the target state $\rho'$ need not be pure: the fact that the direction of classical communication can be exchanged is based only on the purity of the input state $\Psi$.

C. Reduction to one-way protocols

We are now ready to derive our operational Lo-Popescu theorem. The theorem states that the action of an arbi-
trary LOCC protocol on a pure state can be simulated by a one-way LOCC protocol:

**Theorem 1 (Operational Lo-Popescu theorem)**
Let \( \Psi \) be a pure state of \( A \otimes B \) and \( \rho' \) be a (possibly mixed) state of \( A' \otimes B' \). Under the validity of axioms \[1\] and \[2\], the following are equivalent:

1. \( \Psi \) can be transformed into \( \rho' \) by an LOCC protocol.
2. \( \Psi \) can be transformed into \( \rho' \) by a one-way LOCC protocol.

**Proof.** The non-trivial implication is \( 1 \implies 2 \). Suppose that \( \Psi \) can be transformed into \( \rho' \) by an LOCC protocol with \( N \) rounds of classical communication. Without loss of generality, we assume that Alice starts the protocol and that all transformations occurring in the first \( N - 1 \) rounds are pure.

Let \( s = (i_1, i_2, \ldots, i_{N-1}) \) be the sequence of all classical outcomes obtained by Alice and Bob up to step \( N - 1 \), \( p_s \) be the probability of the sequence \( s \), and \( \Psi_s \) be the pure state after step \( N - 1 \) conditional on the occurrence of \( s \). For concreteness, suppose that the outcome \( i_{N-1} \) has been generated on Alice’s side. Then, the rest of the protocol consists in a test \( \{B(s)\} \), performed on Bob’s side, followed by a channel \( A(s, i_{N-1}) \) performed on Alice’s side. By definition, one has

\[
\rho'_{A' B'} = \sum_s p_s \times \sum_{i_N} \Psi_s \otimes B_{N-1} \overset{i_{N-1}}{\rightarrow} B(s)_{i_{N-1}} \otimes B'_{i_N} \rightarrow A(s, i_{N-1})_{i_{N-1}} \otimes A'_{i_N}.
\]

Now, using lemma \[1\] one can invert the direction of the classical communication in the last round, obtaining

\[
\rho'_{A' B'} = \sum_s p_s \times \sum_{i_N} \Psi_s \otimes B_{N-1} \overset{i_{N-1}}{\rightarrow} \tilde{A}_{i_{N-1}}(s)_{i_{N-1}} \otimes \tilde{B}_{i_N}(s, i_{N-1}) \rightarrow A'_{i_N} \otimes B'_{i_N},
\]

for a suitable test \( \{\tilde{A}_{i_{N-1}}(s)\} \) and suitable channels \( \tilde{B}(s, i_{N-1}) \).

Now, since both the \((N - 1)\)-th and the \(N\)-th tests are performed by Alice, these two test can be merged into a single test, thus reducing the original LOCC protocol to an LOCC protocol with \( N - 1 \) rounds. Iterating this argument for \( N - 1 \) times one finally obtains a one-way protocol. \( \blacksquare \)

In quantum theory, the Lo-Popescu theorem provides the foundation for the resource theory of pure state entanglement. Having the operational version of this result will be crucial for our study of the relation between entanglement and thermodynamics.

### V. PURITY

#### A. A resource theory of dynamical control

Consider the scenario where a closed system \( A \) undergoes a reversible dynamics governed by some parameters which can be controlled by an experimenter. For example, system \( A \) could be a charged particle moving in an electric field, whose intensity and direction can be engineered in order to obtain a desired trajectory. In general, the experimenter may not have full control and the actual values of the control parameters may fluctuate randomly.

As a result, the evolution of the system will be described by a Random Reversible (RaRe) channel, that is a channel \( \mathcal{R} \) of the form

\[
\mathcal{A}_\mathcal{R} \mathcal{A} = \sum_{i \in X} p_i \mathcal{A} \mathcal{U}^{(i)} \mathcal{A}^\dagger,
\]

where \( \{\mathcal{U}^{(i)} | i \in X\} \) is a set of reversible transformations and \( \{p_i | i \in X\} \) is their probability distribution. Assuming that the system remains closed during the whole evolution, RaRe channels are the most general transformations the experimenter can implement. Letting system \( A \) evolve jointly with another system \( B \) is not an allowed evolution here, because it would generally require an interaction, which by definition is not possible for closed systems.

An important question in all problems of control is whether a given input state can be driven to a target state using the allowed dynamics. With respect to this task, an input state is more valuable than another if the set of target states that can be reached from the former contains the set of target states that can be reached from the latter. In our model, this idea leads to the following definition.

**Definition 3 (More controllable states)** Given two states \( \rho \) and \( \rho' \) of system \( A \), we say that \( \rho \) is more controllable than \( \rho' \), denoted by \( \rho \succeq \rho' \), if \( \rho' \) can be obtained from \( \rho \) via a RaRe channel.

This definition fits into the general framework of resource theories \[2\], with the RaRe channels playing the role of free operations. Note that at this level of generality there are no free states: since the ability of the experimenter is limited to controlling the evolution, every state is regarded as a resource. Physically, this is in agreement with the fact that the input state in a control problem is not chosen by the experimenter—for example, it can be a thermal state or the ground state of an unperturbed Hamiltonian.

As it is always the case in resource theories, the relation \( \succeq \) is clearly reflexive and transitive, i.e. it is a preorder. Moreover, since the tensor product of two RaRe channels is a RaRe channel, the relation \( \succeq \) is stable under tensor products, namely \( \rho \otimes \sigma \succeq \rho' \otimes \sigma' \) whenever \( \rho \succeq \rho' \) and \( \sigma \succeq \sigma' \).
B. From dynamical control to purity

There is a close relation between the controllability of a state and its purity. For example, if a state is more controllable than a pure state, then it must be pure.

**Proposition 3** If $\psi \in \text{St}(A)$ is a pure state and $\rho$ is more controllable than $\psi$, then $\rho$ must be pure. Specifically, $\rho = U\psi$ for some reversible channel $U$.

**Proof.** Since $\psi$ is pure, the condition $\sum_i p_i U^{(i)}\rho = \psi$ implies that $U^{(i)}\rho = \psi$ for every $i$, meaning that $\rho = \psi^{(i)}\psi$, where $\psi^{(i)}$ is the inverse of $U^{(i)}$. Proposition 2 then guarantees that $\rho$ is pure. ■

In other words, pure states can be reached only from pure states. A natural question is whether every state can be reached from some pure state. The answer is positive in quantum theory and in a large class of GPTs, but there are nevertheless counterexamples, which prevent an easy identification of the theory of dynamical control with a “theory of purity”. Three different examples are illustrated in Fig. 1.

---

**Example 3** Consider first a system with a state space like in Fig. 1a. In this case, there are only two reversible transformations, namely the identity and the reflection around the vertical symmetry axis. As a consequence, there is no way to obtain the mixed states on the two vertical sides by applying a RaRe channel to a pure state. These states represent a valuable resource, even though they are not pure. Hence, in this case the resource theory of dynamical control cannot be thought as a resource theory of purity.

As a second example, consider instead a system whose state space is a half-disk, like in Fig. 1b. Also in this case there are only two reversible transformations (the identity and the reflection around the vertical axis). However, now every mixed state can be generated from some pure state through a RaRe channel. The state space can be foliated into horizontal segments generated by pure states under the action of RaRe channels. As a result, the pure states are the most useful resources and one can interpret the relation $\succ$ as a way to compare the degree of purity of different states. Nevertheless, pure states on different segments are inequivalent resources: in this “resource theory of purity” there are different, inequivalent classes of pure states.

Finally, consider a system with a square state space, like in Fig. 1c and suppose that all the symmetry transformations in the dihedral group $D_4$ are allowed reversible transformations. In this case, all the pure states are equivalent under reversible transformations and every mixed state can be obtained by applying a RaRe channel to a fixed pure state. Here, the resource theory of dynamical control becomes a full-fledged resource theory of purity.

The above examples show that not every GPT supports a sensible resource theory of purity. Motivated by
the examples, we put forward the following definition:

**Definition 4** A theory of purity is a resource theory of dynamical control where every state \( \rho \) can be compared with at least one pure state. The theory is called canonical if every pure state is more controllable than every other state.

In this paper we will focus on canonical theories of purity. Canonical theories can be characterized as those where all pure states are comparable to one another:

**Proposition 4** The following are equivalent:

1. the theory is a canonical theory of purity
2. for every system \( A \), the group of reversible channels acts transitively on the set of pure states.
3. for every system \( A \), there exists at least one state that is more controllable than every state.

**Proof.** 1 \( \implies \) 2. Suppose that the theory is canonical. By definition, this means that every pure state \( \psi \in \text{PurSt}(A) \) is more controllable than every other state. In particular, \( \psi \) must be more controllable than any other pure state \( \varphi \), i.e. \( \varphi = \sum_i p_i U^{(i)} \psi \), for some probabilities \( \{p_i\} \) and some reversible channels \( \{U^{(i)}\} \). Since \( \varphi \) is pure, this relation implies \( \varphi = U^{(i)} \psi \) for every \( i \), which in turn implies that \( \psi \) is pure and \( \psi = V^{(i)} \varphi \), with \( V^{(i)} \) being the inverse of \( U^{(i)} \). Hence, \( \psi \) and \( \varphi \) are connected by a reversible channel. We conclude that the group of reversible transformations acts transitively on the set of pure states.

2 \( \implies \) 3. Every state \( \rho \) can be expressed as a convex combination of the form \( \rho = \sum_i p_i \varphi_i \), where \( \{p_i\} \) is a probability distribution allowed by the theory and \( \varphi_i \) are pure states. Now, suppose that \( \psi \) is a pure state. For every \( i \), by picking a reversible channel \( U^{(i)} \) such that \( U^{(i)} \psi = \varphi_i \), one obtains the relation \( \rho = \sum_i p_i U^{(i)} \psi \), meaning that \( \psi \) is more controllable than \( \rho \). Since \( \rho \) is generic, we conclude that \( \psi \) is more controllable than every state.

3 \( \implies \) 1. Suppose that there exists a state \( \rho \) that is more controllable than every state. In particular, \( \rho \) must be more controllable than every pure state \( \psi \). By proposition 3, \( \rho \) must be pure and there exists a reversible transformation \( U \) such that \( \rho = U^{-1} \psi \). Hence, the pure state \( \psi = U \rho \) is more controllable than every state. Since \( \psi \) is generic, this proves that every pure state is more controllable than every other state, i.e. the theory is canonical.

Starting from Hardy’s work [13], the transitivity of the reversible channels on pure states has featured in a number of axiomatizations of quantum theory, either directly as an axiom [26, 28] or indirectly as a consequence of an axiom [18]. Proposition 4 provides a new angle on this axiom, as a necessary and sufficient condition for a well-behaved theory of purity, and, ultimately, for a well-behaved thermodynamics.

C. Maximally mixed states

In a canonical theory of purity, we say that \( \rho \) is purer than \( \rho' \) iff \( \rho \geq \rho' \) and we adopt the notation \( \rho \geq_{\text{pur}} \rho' \). In this case, we also say that \( \rho' \) is more mixed than \( \rho \), denoted by \( \rho' \geq_{\text{mix}} \rho \). When \( \rho \geq_{\text{mix}} \rho' \) and \( \rho' \geq_{\text{mix}} \rho \) we say that \( \rho \) and \( \rho' \) are equally mixed, denoted by \( \rho \geq_{\text{mix}} \rho' \). Clearly, every two states that differ by a reversible channel are equally mixed. We say that a state \( \chi \in \text{St}(A) \) is maximally mixed iff it satisfies the property

\[
\forall \rho \in \text{St}(A) : \rho \geq_{\text{mix}} \chi \implies \rho = \chi.
\]

Maximally mixed states can be characterized as the states that are invariant under all reversible channels:

**Proposition 5** A state \( \chi \in \text{St}(A) \) is maximally mixed if and only if it is invariant, i.e. if and only if \( \chi = U \chi \) for every reversible channel \( U : A \to A \).

The proof is straightforward. Note that maximally mixed states do not exist in every theory: for example, for infinite-dimensional quantum systems there is no maximally mixed density operator, i.e. no trace-class operator that is invariant under the action of the full unitary group. For finite-dimensional canonical theories, however, the maximally mixed state exists and is unique under the standard assumption of compactness of the state space [47, 72]. In this case, the state \( \chi \) is not only a maximal element, but also the maximum of the relation \( \geq_{\text{mix}} \), that is,

\[
\chi \geq_{\text{mix}} \rho \iff \forall \rho \in \text{St}(A).
\]

This is in analogy with the quantum case, where the maximally mixed state is given by the density matrix \( \chi = I/d \), where \( I \) is the identity operator on the system’s Hilbert space and \( d \) is the Hilbert space dimension. Another example of finite-dimensional canonical theory is provided by the square bit:

**Example 4** Consider a system whose state space is a square, as in Fig. 1c and pick a generic (mixed) state \( \rho \). The states that are more mixed than \( \rho \) are obtained by applying all possible reversible transformations to \( \rho \) (i.e. all the elements of the dihedral group \( D_4 \)) and taking the convex hull of the orbit. The set of all states that are more mixed than \( \rho \) form an octagon depicted in blue in Fig. 2. All the vertexes of the octagon are equally mixed. The centre of the square is the maximally state \( \chi \), the unique invariant state of the system.

VI. ENTANGLEMENT-THERMODYNAMICS DUALITY

In quantum theory, it is well known that the ordering of pure bipartite states according to the degree of entanglement is equivalent to the ordering of their marginals.
Figure 2: Mixedness relation for the state space of a square bit: vertices of the octagon represent the states that can be reached from a given state $\rho$ via reversible transformations. Their convex hull is the set of states that are more mixed than $\rho$. Note that it contains the invariant state $\chi$, which can be characterized as the maximally mixed state. According to the degree of mixedness [31–34, 63]. In this section, we will prove the validity of this equivalence based only on first principles. The key axiom to prove this equivalence is the purification principle.

A. Purification

In order to establish the entanglement-thermodynamics duality, we consider theories that satisfy Purification [47, 48], which expresses a strengthened version of the principle of conservation of information [72]. Purification characterizes all the physical theories admitting a description where all processes are pure and reversible at a fundamental level [72, 79]. In a causal theory, one can define marginal states.

**Definition 5** The marginal state of a bipartite state $\rho_{AB}$ on system $A$ is the state $\rho_A := \text{Tr}_B \rho_{AB}$ obtained by applying the deterministic effect on $B$.

Purification ensures that every state arises as a marginal state of a pure state of a larger system:

**Axiom 3 (Purification)** Every state has a purification, unique up to reversible channels on the purifying system, namely for every two pure states $\Psi, \Psi' \in \text{PurSt}(A \otimes B)$ such that

$$\Psi' \begin{array}{c} A \\ B \end{array} = \begin{array}{c} A \\ \text{Tr} \end{array} \Psi \begin{array}{c} B \\ \text{Tr} \end{array},$$

one has

$$\Psi' \begin{array}{c} A \\ B \end{array} = \Psi \begin{array}{c} A \\ B \text{U} \end{array}, \quad (11)$$

for some reversible transformation $U : B \rightarrow B$.

Purification has a lot of important consequences [47]. First of all, purification implies that every two pure states are connected to one another by a reversible transformation:

**Proposition 6** For every system $B$ and every pair of pure states $\psi, \psi' \in \text{PurSt}(B)$ there is a reversible channel $U : B \rightarrow B$ such that $\psi' = U \psi$.

The existence of a reversible transformation connecting $\psi$ and $\psi'$ is a consequence of the uniqueness of purification [Eq. (11)] in the special case where $A$ is the trivial system [47] (no wire corresponding to system $A$). Since all pure states are equivalent under reversible transformations, every theory with purification gives rise to a canonical theory of purity, in the sense of subsection V.B.

Another important consequence of Purification is the steering property, stating that every ensemble decomposition of a given state can be induced by a measurement on its purifying system:

**Proposition 7 (Steering)** Let $\rho$ be a state of system $A$ and let $\Psi \in \text{PurSt}(A \otimes B)$ be a purification of $\rho$. For every ensemble of states $\{\rho_i\}_{i \in X}$ such that $\sum_i \rho_i = \rho$, there exists a measurement $\{b_i\}_{i \in X}$ on the purifying system $B$ and such that

$$\rho_i \begin{array}{c} A \\ B \end{array} = \Psi \begin{array}{c} B \end{array} b_i \forall i \in X.$$  

See theorem 6 of Ref. [47] for the proof. The steering property will turn out to be essential in establishing the duality between entanglement and thermodynamics.

B. One-way protocols transforming pure states into pure states

The operational Lo-Popescu theorem guarantees that every LOCC protocol acting on a pure bipartite input state can be simulated by a one-way protocol. We now consider one-way protocols where both the input and the output states are pure. In this case, Purification guarantees that every such protocol can be simulated by a one-way protocol where all conditional operations are reversible.

**Lemma 2** Let $\Psi$ and $\Psi'$ be two pure states of $A \otimes B$. Under the validity of Purification and Purity Preservation, every one-way protocol transforming $\Psi$ into $\Psi'$ can be simulated by a one-way protocol where all conditional operations are reversible.

**Proof.** Suppose that $\Psi$ can be transformed into $\Psi'$ via a one-way protocol where Alice performs a test $\{A_i\}_{i \in X}$
and Bob performs a channel $\mathcal{B}^{(i)}$ conditional on outcome $i$. By definition, we have

$$
\sum_i \left( \Psi \xrightarrow{\mathcal{A}_i} \right)_{A \rightarrow B} \mathcal{B}^{(i)}_{B \rightarrow B} = \Psi' \xrightarrow{\mathcal{A}}_{A \rightarrow B}.
$$

Since $\Psi'$ is pure, this implies that there exists a probability distribution $\{p_i\}$ such that

$$
\Psi \xrightarrow{\mathcal{A}_i} \mathcal{B}^{(i)}_{B \rightarrow B} = p_i \Psi' \xrightarrow{\mathcal{A}}_{A \rightarrow B}.
$$

(12)

for every outcome $i$. Now, without loss of generality each transformation $\mathcal{A}_i$ can be assumed to be pure (if not, one can always decompose it into pure transformations).

Then, Purity Preservation guarantees that the normalized state $\Psi_i$ defined by

$$
\Psi_i \xrightarrow{\mathcal{A}_i} \mathcal{B}^{(i)}_{B \rightarrow B} := p_i^{-1} \left( \Psi \xrightarrow{\mathcal{A}_i} \mathcal{B}^{(i)}_{B \rightarrow B} \right)
$$

is pure. With this definition, Eq. (12) becomes

$$
\Psi_i \xrightarrow{\mathcal{A}_i} \mathcal{B}^{(i)}_{B \rightarrow B} = \Psi' \xrightarrow{\mathcal{A}}_{A \rightarrow B}.
$$

(13)

Tracing out system $B$ on both sides one obtains

$$
\Psi' \xrightarrow{\text{Tr}} = \Psi_i \xrightarrow{\mathcal{A}_i} \mathcal{B}^{(i)}_{B \rightarrow B} \xrightarrow{\text{Tr}} = \Psi' \xrightarrow{\mathcal{A}}_{A \rightarrow B}.
$$

the second equality coming from the normalization of the channel $\mathcal{B}^{(i)}$ (cf. proposition 1). Hence, the pure states $\Psi_i$ and $\Psi'$ have the same marginal on $A$. By the uniqueness of Purification, they must differ by a reversible channel $\mathcal{U}^{(i)}$ on the purifying system $B$, namely

$$
\Psi_i \xrightarrow{\mathcal{B}^{(i)}} \xrightarrow{\mathcal{U}^{(i)}} \mathcal{B}^{(i)}_{B \rightarrow B} = \Psi' \xrightarrow{\mathcal{A}}_{A \rightarrow B}.
$$

(14)

In conclusion, we obtained

$$
\sum_i \left( \Psi \xrightarrow{\mathcal{A}_i} \mathcal{B}^{(i)}_{B \rightarrow B} \right) = p_i \left( \Psi' \xrightarrow{\mathcal{A}}_{A \rightarrow B} \right).
$$

where we used Eqs. (12), (14), and (13). In other words, the initial protocol can be simulated by a protocol where Alice performs the test $\{A_i\}$ and Bob performs the reversible transformation $\mathcal{U}^{(i)}$ conditionally to the outcome $i$. ■

The reduction to one-way protocols with reversible operations is the key to connect the resource theory of entanglement with the resource theory of dynamical control, defined earlier in subsection VA.

C. The more entangled a pure state, the more mixed its marginals

**Lemma 3** Let $\Psi$ and $\Psi'$ be two pure states of system $A \otimes B$ and let $\rho, \rho'$ and $\sigma, \sigma'$ be their marginals on system $A$ and $B$, respectively. Under the validity of Purification, Purity Preservation, and Local Exchangeability, if $\Psi$ is more entangled than $\Psi'$, then $\rho$ (resp. $\sigma$) is more mixed than $\rho'$ (resp. $\sigma'$).

**Proof.** By the operational Lo-Popescu theorem, we know that there exists a one-way protocol transforming $\Psi$ into $\Psi'$. Moreover, thanks to Purification, the conditional operations in the protocol can be chosen to be reversible (lemma 2). Let us choose a protocol with classical communication from Alice to Bob, in which Alice performs the test $\{A_i\}_{i \in X}$ and Bob performs the reversible transformation $U^{(i)}$ conditional on outcome $i$. Since $\Psi'$ is pure, we must have

$$
\Psi \xrightarrow{\mathcal{A}_i} \mathcal{B}^{(i)}_{B \rightarrow B} \xrightarrow{\mathcal{U}^{(i)}} = p_i \Psi' \xrightarrow{\mathcal{A}}_{A \rightarrow B} \forall i \in X,
$$

where $\{p_i\}$ is a suitable probability distribution. Denoting by $U^{(i)}$ the inverse of $U^{(i)}$ and applying it on both sides of the equation, we obtain

$$
\Psi \xrightarrow{\mathcal{A}_i} \mathcal{B}^{(i)}_{B \rightarrow B} \xrightarrow{U^{(i)}} = p_i \Psi' \xrightarrow{\mathcal{A}}_{A \rightarrow B} \xrightarrow{U^{(i)}}.
$$

Summing over all outcomes the equality becomes

$$
\sum_i \left( \Psi \xrightarrow{\mathcal{A}_i} \mathcal{B}^{(i)}_{B \rightarrow B} \right) \xrightarrow{U^{(i)}} = p_i \left( \Psi' \xrightarrow{\mathcal{A}}_{A \rightarrow B} \right) \xrightarrow{U^{(i)}}.
$$

(15)
with $A := \sum_i A_i$ and $R := \sum_j p_i \gamma^{(i)}$. Finally, we obtain

$$\sigma_{AB} = \Psi_{AB} \mathrel{\operatorname{Tr}} B = \Psi_{A} \mathrel{\operatorname{Tr}} B = \Psi'_{A} \mathrel{\operatorname{Tr}} B = \sigma'_{AB} \mathrel{\operatorname{Tr}} B,$$

where we used the normalization of channel $A$ in the second equality and Eq. [15] in the third. Since $R$ is a RaRe channel by construction, we proved that $\sigma$ is more mixed than $\sigma'$. The fact that $\rho$ is more mixed than $\rho'$ can be proven by the same argument, starting from a one-way protocol with classical communication from Bob to Alice. \qed

The relation between degree of entanglement of a pure state and degree of mixedness of its marginals holds not only for bipartite states, but also for multipartite states. Indeed, suppose that $\Psi$ and $\Psi'$ are two pure states of system $A_1 \otimes A_2 \otimes \cdots \otimes A_N$ and that $\Psi$ is more entangled than $\Psi'$, in the sense that there exists a (multipartite) LOCC protocol converting $\Psi$ into $\Psi'$. For every subset $S \subset \{1, \ldots, N\}$ one can define $A := \bigotimes_{n \in S} A_n$ and $B := \bigotimes_{n \in S^c} A_n$ and apply lemma 3. As a result, one obtains that the marginals of $\Psi$ are more mixed than the marginals of $\Psi'$ on every subsystem.

**D. The more mixed a state, the more entangled its purification**

We now prove the converse direction of the entanglement-thermodynamics duality: if a state is more mixed than another, then its purification is more entangled. Remarkably, the proof of this fact requires only the validity of Purification.

**Lemma 4** Let $\rho$ and $\rho'$ be two states of system $A$ and let $\Psi, \Psi' \in \text{PurSt}(A \otimes B)$ be two purifications of $\rho, \rho'$, respectively. Under the validity of Purification, if $\rho$ is more mixed than $\rho'$, then $\Psi$ is more entangled than $\Psi'$.

**Proof.** By hypothesis, one has

$$\rho_{AB} = \Psi_{AB} \mathrel{\operatorname{Tr}} B = \Psi'_{AB} \mathrel{\operatorname{Tr}} B = \rho'_{AB} \mathrel{\operatorname{Tr}} B,$$

for some RaRe channel $R := \sum_i p_i \gamma^{(i)}$. Let us define the bipartite state as

$$\Theta_{AB} := \Psi'_{AB} \mathrel{\operatorname{Tr}} B = \sum_i p_i \Psi'_{AB} \mathrel{\operatorname{Tr}} B = \sum_i p_i \Psi'_{AB} \mathrel{\operatorname{Tr}} B = \rho_{AB} \mathrel{\operatorname{Tr}} B.$$



By construction, $\Theta$ is an extension of $\rho$: indeed, one has

$$\Theta_{AB} := \Psi'_{AB} \mathrel{\operatorname{Tr}} B = \rho_{AB} \mathrel{\operatorname{Tr}} B = \rho_{AB} \mathrel{\operatorname{Tr}} B.$$

Let us take a purification of $\Theta$, say $\Gamma \in \text{PurSt}(A \otimes B \otimes C)$. Clearly, $\Gamma$ is a purification of $\rho$, since one has

$$\Gamma_{ABC} = \Theta_{AB} \mathrel{\operatorname{Tr}} B = \rho_{AB} \mathrel{\operatorname{Tr}} B = \rho_{AB} \mathrel{\operatorname{Tr}} B.$$

Then, the uniqueness of purification implies that $\Gamma$ must be of the form

$$\Gamma_{ABC} = \Psi_{AB} \mathrel{\operatorname{Tr}} B = \Psi_{AB} \mathrel{\operatorname{Tr}} B = \Gamma_{AB} \otimes \left( \mathcal{U}_{BC} \mathrel{\operatorname{Tr}} C \right) = \Theta_{AB} \mathrel{\operatorname{Tr}} B = \rho_{AB} \mathrel{\operatorname{Tr}} B = \rho_{AB} \mathrel{\operatorname{Tr}} B,$$

for some reversible transformation $\mathcal{U}$ and some pure state $\gamma$. In other words, $\Psi$ can be transformed into $\Theta$ by local operations on Bob’s side.

Now, Eq. [16] implies that the states $\{p_i (\mathcal{U}^{(i)} \otimes I_B) \Psi \}_{i \in X}$ are an ensemble decomposition of $\Theta$. Hence, the steering property (proposition 7) implies that there exists a measurement $\{c_i\}_{i \in X}$ such that

$$p_i \left( \Psi_{AB} \mathrel{\operatorname{Tr}} B \right) = \Gamma_{AB} \mathrel{\operatorname{Tr}} B = \rho_{AB} \mathrel{\operatorname{Tr}} B = \rho_{AB} \mathrel{\operatorname{Tr}} B = \rho_{AB} \mathrel{\operatorname{Tr}} B = \rho_{AB} \mathrel{\operatorname{Tr}} B.$$

Combining Eqs. [17] and [18], we obtain the desired result.

$$\Psi_{AB} \mathrel{\operatorname{Tr}} B = \sum_{i \in X} \Psi_{AB} \mathrel{\operatorname{Tr}} B = \sum_{i \in X} \Psi_{AB} \mathrel{\operatorname{Tr}} B = \sum_{i \in X} \Psi_{AB} \mathrel{\operatorname{Tr}} B.$$
where \( \{B_i\}_{i \in \mathbb{X}} \) is the test defined by

\[
\begin{array}{ccc}
  & B & B_i & B \\
\gamma & C & U & C & \overleftarrow{C_i}
\end{array}
\]

In conclusion, if the marginal state of \( \Psi \) is more mixed than the marginal state of \( \Psi' \), then \( \Psi \) can be converted into \( \Psi' \) by a (one-way) LOCC protocol. ■

E. The duality

Combining lemmas 3 and 4 we identify the degree of entanglement of a pure bipartite state with the degree of mixedness of its marginals:

**Theorem 2 (Entanglement-thermodynamics duality)** Let \( \Psi \) and \( \Psi' \) be two pure states of system \( A \otimes B \) and let \( \rho, \rho' \) and \( \sigma, \sigma' \) be their marginals on system \( A \) and \( B \), respectively. Under the validity of Purification, Purity Preservation, and Local Exchangeability, the following statements are equivalent:

1. \( \Psi \) is more entangled than \( \Psi' \)
2. \( \rho \) is more mixed than \( \rho' \)
3. \( \sigma \) is more mixed than \( \sigma' \).

**Proof.** The implications 1 \( \Rightarrow \) 2 and 1 \( \Rightarrow \) 3 follow from lemma 3 and require the validity of all the three axioms. The implications 2 \( \Rightarrow \) 1 and 3 \( \Rightarrow \) 1 follow from lemma 4 and require only the validity of Purification. ■

The duality can be illustrated by the commutative diagrams

\[
\begin{array}{ccc}
  & \Psi & \text{LOCC} & \Psi' \\
\text{Tr}_B & \rho \underset{\text{RaRe}}{\rightarrow} & \rho' \\
\end{array}
\]

and is implemented operationally by discarding one of the component systems. Another illustration of the duality is via the diagram

\[
\begin{array}{ccc}
  & \Psi & \text{LOCC} & \Psi' \\
\text{purification} & \rho \underset{\text{RaRe}}{\rightarrow} & \rho' \\
\end{array}
\]

Here the map implementing the duality is (a choice of) the purification. Such a map cannot be realized as a physical operation [47]. Instead, it corresponds to the theoretical operation of modelling mixed states as marginals of pure states.

VII. CONSEQUENCES OF THE DUALITY

In this section we discuss the simplest consequences of the entanglement-thermodynamics duality, including the relation between maximally mixed and maximally entangled states, as well as a link between information erasure and generation of entanglement. From now on, the axioms used to derive the duality will be treated as standing assumptions and will not be written explicitly in the statement of the results.

A. Maximally entangled states

As a consequence of the duality, there exists a correspondence between maximally mixed and “maximally entangled” states, the latter being defined as follows

**Definition 6** A pure state \( \Phi \) of system \( A \otimes B \) is maximally entangled if no other pure state of \( A \otimes B \) is more entangled than \( \Phi \), except for the states that are equivalent to \( \Phi \) under local reversible transformations—i.e. if for every \( \Psi \in \text{PurSt} \ (A \otimes B) \) one has

\[
\Psi \succeq_{\text{ent}} \Phi \quad \implies \quad \Psi = (\mathcal{U} \otimes \mathcal{V}) \Phi
\]

for some reversible transformations \( \mathcal{U} : A \rightarrow A \) and \( \mathcal{V} : B \rightarrow B \).

Theorem 2 directly implies the following.

**Corollary 1** The purification of a maximally mixed state is maximally entangled.

**Proof.** Suppose that \( \Psi \in \text{PurSt} \ (A \otimes B) \) is more entangled than \( \Phi \), where \( \Phi \) is a purification of the maximally mixed state of system \( A \) (assuming such a state exists for system \( A \)). By theorem 2 the marginal of \( \Psi \) on system \( A \), denoted by \( \rho \), must satisfy \( \rho \succeq_{\text{mix}} \chi \). Since \( \chi \) is maximally mixed, this implies \( \rho = \chi \). The uniqueness of purification then implies the condition \( \Psi = (\mathcal{I}_A \otimes \mathcal{V}_B) \Phi \) for some reversible transformation \( \mathcal{V}_B \) on \( B \). ■

Under standard assumptions [17, 17], the maximally mixed state is not only a maximal element of the mixedness relation, but also the maximum—i.e. it is more mixed than every state [cf. Eq. (19)]. Under the same standard assumptions, it is immediate to obtain that the purification of a maximally mixed state is more entangled than every state, namely

\[
\Phi \succeq_{\text{ent}} \Sigma \quad \forall \Sigma \in \text{St} \ (A \otimes B) . \tag{19}
\]

The relation follows directly from theorem 2 when \( \Sigma \) is a pure state, and in the general case can be proved by convexity, using the fact that the set of LOCC channels is closed under convex combinations.
B. Duality for states on different systems

Theorem 2 concerns the convertibility of states of the same system. To generalize it to arbitrary systems, it is enough to observe that the tensor product with local pure states does not change the degree of entanglement: for arbitrary pure states $\Psi$, $\alpha'$, and $\beta'$ of systems $A \otimes B$, $A'$, and $B'$ one has

$$\Psi \simeq_{\text{ent}} \alpha' \otimes \Psi \otimes \beta',$$  \hspace{1cm} (20)

relative to the bipartition $(A' \otimes A) \otimes (B \otimes B')$. As a consequence, one has the equivalence

$$\Psi \succeq_{\text{ent}} \Psi' \iff \alpha' \otimes \Psi \otimes \beta' \succeq_{\text{ent}} \alpha \otimes \Psi' \otimes \beta \quad \text{for arbitrary states } \alpha, \alpha', \beta, \beta' \text{ of } A, A', B, B', \text{ respectively.}$$

This fact leads directly to the generalization of the duality to states of different systems:

Corollary 2 Let $\Psi$ and $\Psi'$ be two pure states of systems $A \otimes B$ and $A' \otimes B'$, respectively, and let $\rho$, $\rho'$, $\sigma$ and $\sigma'$ be their marginals on system $A, A', B$ and $B'$, respectively. Under the validity of Purification, Purity Preservation, and Local Exchangeability, the following statements are equivalent:

1. $\Psi$ is more entangled than $\Psi'$.
2. $\rho \otimes \alpha'$ is more mixed than $\alpha \otimes \rho'$ for every pair of states $\alpha \in \text{PurSt}(A)$ and $\alpha' \in \text{PurSt}(A')$.
3. $\sigma \otimes \beta'$ is more mixed than $\beta \otimes \sigma'$ for every pair of states $\beta \in \text{PurSt}(B)$ and $\beta' \in \text{PurSt}(B')$.

The duality is now implemented by the operation of erasure assisted by a catalyst, as illustrated by the commutative diagrams

$$\Psi \xrightarrow{\text{LOCC}} \Psi' \quad \iff \quad \rho \otimes \alpha' \xrightarrow{\text{LOCC}} \alpha \otimes \rho' \quad \iff \quad \sigma \otimes \beta' \xrightarrow{\text{LOCC}} \beta \otimes \sigma'.$$

At this point, a cautionary remark is in order. Inspired by Eq. (20) and (21) one may be tempted compare the degree of mixedness of states of different systems, by postulating the relation

$$\rho \simeq_{\text{mix}} \rho \otimes \alpha',$$  \hspace{1cm} (21)

for arbitrary states $\rho$ and arbitrary pure states $\alpha'$. The appeal of this choice is that the duality would maintain the simple form

$$\Psi \succeq_{\text{ent}} \Psi' \iff \rho \succeq_{\text{mix}} \rho'.$$  \hspace{1cm} (22)

even for states of different systems. However, Eq. (21) would trivialize the resource theory of purity: as a special case, it would imply the relation $1 \simeq_{\text{mix}} \alpha'$ for a generic pure state of a generic system, meaning that pure states can be freely generated from nothing and can be freely erased. Since in a canonical theory of purity the pure states are the most resourceful ones, having pure states for free would mean having every state for free.

C. Information erasure and entanglement generation

The entanglement-thermodynamics duality establishes a link between the two tasks of erasing information and generating entanglement. By “erasing information” we mean resetting a mixed state to a pure state of the same system (cf. [66]). Clearly, erasure is a costly operation in the resource theory of purity: there is no way to reset a mixed state to a pure state using only RaRe channels (cf. proposition 3). The corresponding operation in the resource theory of entanglement is the generation of entangled states from product states. By the duality, the impossibility of erasing information by RaRe channels and the impossibility of generating entanglement by LOCC are one and the same thing.

Let us consider now the task of erasure assisted by a catalyst, represented by a system C whose state remains unaffected by the erasure operation. In this case, the operation of erasure transforms the product state $\rho \otimes \gamma \in \text{St}(A \otimes C)$ into the state $\alpha_0 \otimes \gamma$ for some pure state $\alpha_0 \in \text{PurSt}(A)$. By duality, it is immediate to see that catalyst-assisted erasure is equivalent to catalyst-assisted entanglement generation:

Corollary 3 Let $\Psi$ and $\Gamma$ be two pure states of systems $A \otimes B$ and $C \otimes D$, respectively, and let $\rho$ and $\gamma$ be their marginals on systems $A$ and $C$, respectively. Then, the following are equivalent:

1. $\rho$ can be erased by a RaRe channel using $\gamma$ as a catalyst.
2. $\Psi$ can be generated by a LOCC channel using $\Gamma$ as a catalyst.

Now, it is well known that that in quantum mechanics no LOCC protocol can generate entanglement from nothing, even in the presence of a catalyst. Equivalently, no random-unitary channel can erase information, even in the presence of a catalyst. Quite surprisingly, however, these impossibilities do not seem to follow from the axioms used to derive the duality: in principle, there may exist theories that satisfy Purification, Purity-Preservation, and Local Exchangeability and nevertheless allow one to catalytically erase information and generate entanglement for free. In such theories, the catalysts would behave like “entropy sinks”, which can absorb mixed states without becoming more mixed, or like “entanglement reservoirs”, from which entanglement can be borrowed indefinitely. This scenario can be excluded at the level of first principles, by postulating the following

Axiom 4 (No Entropy Sinks) Random reversible dynamics cannot erase information, even with the assistance of a catalyst.

Axiom 4 is a necessary requirement for a quantitative theory of entanglement. Indeed, suppose that one is trying to define a measure of entanglement $M$ that i) assigns
a non-zero value to every entangled state and ii) it is additive (or super-additive) on product of entangled states, namely

\[ M(\Psi \otimes \Gamma) \geq M(\Psi) + M(\Gamma). \]

Clearly, such measure can exist only in theories satisfying Axiom 4. A similar consideration holds for measures of mixedness, which ultimately provide the foundation of the notion of entropy and of a quantitative theory of thermodynamics. Axiom 4 will also lead to an elegant reformulation of our results, which sheds light to the physical picture underlying the entanglement-thermodynamics duality.

**VIII. SYMMETRIC PURIFICATION**

Here we show that the axioms of Purification, Local Exchangeability, and, No Entropy Sinks lead to a new property, namely the existence of a symmetric purification for every state. By a symmetric purification of the state \( \rho \in \text{St}(A) \) we mean a purification \( \Psi \in \text{PurSt}(A \otimes A) \), where marginals on both systems are equal to \( \rho \). By adding the requirement that the purifications of a state should be defined uniquely up to local reversible transformations, we obtain the following

**Axiom 5 (Symmetric Purification)** Every state has a symmetric purification. All the purifications of a given state are equivalent up to local reversible transformations.

Clearly, symmetric purifications are locally exchangeable: indeed, if \( \Psi \) is a symmetric purification one has

\[
\begin{align*}
\begin{array}{c}
\Psi \\
\text{SWAP} \\
\text{Tr}
\end{array}
\end{align*}
\begin{array}{c}
A \\
A \\
A \\
A
\end{array}
= 
\begin{align*}
\begin{array}{c}
\Psi \\
\text{SWAP} \\
\text{Tr}
\end{array}
\begin{array}{c}
A \\
A \\
A \\
A
\end{array}
\end{align*}
\]

and, by the uniqueness of purification,

\[
\begin{align*}
\begin{array}{c}
\Psi \\
\text{SWAP} \\
\text{Tr}
\end{array}
\begin{array}{c}
A \\
A \\
A \\
A
\end{array}
= 
\begin{align*}
\begin{array}{c}
\Psi \\
\text{SWAP} \\
\text{Tr}
\end{array}
\begin{array}{c}
A \\
A \\
A \\
A
\end{array}
\end{align*}
\]

for some reversible channel \( \mathcal{U} \). Since all purifications of \( \rho \) are equivalent to \( \Psi \) under local operations and since \( \Psi \) is locally swappable, we conclude that every purification of \( \rho \) is locally exchangeable [by the same argument used in Eq. (4)]. Therefore Symmetric Purification is a stronger assumption than Local Exchangeability.

However, in the present setting we need not assume Symmetric Purification, because it is a consequence of the other axioms, as shown in the the following.

**Proposition 8** Every theory satisfying Purification, Purity Preservation, Local Exchangeability, and No Entropy Sinks satisfy Symmetric Purification.

**Proof.** Let \( \rho \) be a pure state of system \( A \), and let \( \Psi \in \text{PurSt}(A \otimes B) \) be one of its purifications. By Local Exchangeability there exist two channels \( \mathcal{C} \) and \( \mathcal{D} \) such that

\[
\begin{align*}
\begin{array}{c}
\Psi \\
\\text{SWAP} \\
\text{Tr}
\end{array}
\begin{array}{c}
A \\
B \\
A \\
B
\end{array}
= 
\begin{align*}
\begin{array}{c}
\Psi \\
\\text{SWAP} \\
\text{Tr}
\end{array}
\begin{array}{c}
A \\
B \\
A \\
B
\end{array}
\end{align*}
\]

Now, in a theory satisfying Purification, every channel can be realized through a reversible transformation acting on the system and on an environment, initially in a pure state and finally discarded [47]. In particular, channel \( \mathcal{C} \) can be realized as

\[
\begin{align*}
\begin{array}{c}
A \\
B \\
C \\
B
\end{array}
= 
\begin{align*}
\begin{array}{c}
\eta \\
E \\
F \\
E'
\end{array}
\begin{array}{c}
A \\
B \\
A \\
B
\end{array}
\end{align*}
\]

where \( E \) and \( E' \) are suitable systems, \( \mathcal{U} \) is a reversible transformation, and \( \eta \) is a pure state. Similarly, channel \( \mathcal{D} \) can be realized as

\[
\begin{align*}
\begin{array}{c}
A \\
B \\
D \\
A
\end{array}
= 
\begin{align*}
\begin{array}{c}
\phi \\
F \\
F' \\
F'
\end{array}
\begin{array}{c}
A \\
B \\
A \\
B
\end{array}
\end{align*}
\]

Inserting the realizations of \( \mathcal{C} \) and \( \mathcal{D} \) in the local exchangeability condition, we obtain

\[
\begin{align*}
\begin{array}{c}
\Psi \\
\\text{SWAP} \\
\text{Tr}
\end{array}
\begin{array}{c}
A \\
B \\
C \\
B
\end{array}
= 
\begin{align*}
\begin{array}{c}
\Psi \\
\\text{SWAP} \\
\text{Tr}
\end{array}
\begin{array}{c}
A \\
B \\
C \\
B
\end{array}
\end{align*}
\]

and, by the uniqueness of purification and Purity Preservation,

\[
\begin{align*}
\begin{array}{c}
\psi \\
\\text{Swap} \\
\text{Tr}
\end{array}
\begin{array}{c}
A \\
B \\
C \\
B
\end{array}
= 
\begin{align*}
\begin{array}{c}
\Gamma \\
\psi \\
\psi
\end{array}
\begin{array}{c}
A \\
B \\
C \\
B
\end{array}
\end{align*}
\]

for some pure state \( \Gamma \). The above equation shows that the state \( \Gamma \) can be generated by LOCC using \( \Psi \) as a catalyst. By the No Entropy Sinks requirement, we have that \( \Gamma \) must be a product state, i.e. \( \Gamma = \eta' \otimes \phi' \) for two pure states \( \eta' \) and \( \phi' \). Hence, the local exchangeability condition becomes

\[
\begin{align*}
\begin{array}{c}
\Psi \\
\\text{SWAP} \\
\text{Tr}
\end{array}
\begin{array}{c}
A \\
B \\
C \\
B
\end{array}
= 
\begin{align*}
\begin{array}{c}
\Psi \\
\\text{SWAP} \\
\text{Tr}
\end{array}
\begin{array}{c}
A \\
B \\
C \\
B
\end{array}
\end{align*}
\]

for some pure state \( \Gamma \). The above equation shows that the state \( \Gamma \) can be generated by LOCC using \( \Psi \) as a catalyst. By the No Entropy Sinks requirement, we have that \( \Gamma \) must be a product state, i.e. \( \Gamma = \eta' \otimes \phi' \) for two pure states \( \eta' \) and \( \phi' \). Hence, the local exchangeability condition becomes

\[
\begin{align*}
\begin{array}{c}
\Psi \\
\\text{SWAP} \\
\text{Tr}
\end{array}
\begin{array}{c}
A \\
B \\
C \\
B
\end{array}
= 
\begin{align*}
\begin{array}{c}
\Psi \\
\\text{SWAP} \\
\text{Tr}
\end{array}
\begin{array}{c}
A \\
B \\
C \\
B
\end{array}
\end{align*}
\]

for some pure state \( \Gamma \). The above equation shows that the state \( \Gamma \) can be generated by LOCC using \( \Psi \) as a catalyst. By the No Entropy Sinks requirement, we have that \( \Gamma \) must be a product state, i.e. \( \Gamma = \eta' \otimes \phi' \) for two pure states \( \eta' \) and \( \phi' \). Hence, the local exchangeability condition becomes

\[
\begin{align*}
\begin{array}{c}
\Psi \\
\\text{SWAP} \\
\text{Tr}
\end{array}
\begin{array}{c}
A \\
B \\
C \\
B
\end{array}
= 
\begin{align*}
\begin{array}{c}
\Psi \\
\\text{SWAP} \\
\text{Tr}
\end{array}
\begin{array}{c}
A \\
B \\
C \\
B
\end{array}
\end{align*}
\]
or, equivalently,

\[
\begin{array}{c}
\eta_E \\
\Psi \quad B \\
\varphi \quad F \\
\end{array} = 
\begin{array}{c}
\eta'_{E'} \\
\Psi' \quad B \\
\varphi' \quad F' \\
\end{array} = 
\begin{array}{c}
\Psi \quad B \\
\varphi \quad F \\
SVAP \\
\end{array} 
\begin{array}{c}
\eta'_{E'} \\
\Psi' \quad B \\
\varphi' \quad F' \\
\end{array}
\]

Discarding system \(E\) one obtains

\[
\begin{array}{c}
\Psi \quad B \\
\varphi \quad F \\
\end{array} = 
\begin{array}{c}
\Sigma \\
\end{array}
\]

for some suitable state \(\Sigma\). Since the l.h.s. is a pure state, \(\Sigma\) must be a pure state. Now discard the second system \(A\) and system \(F'\). We have, recalling Eq. (23),

\[
\begin{array}{c}
\Psi \quad B \\
\varphi \quad F \\
\end{array} = 
\begin{array}{c}
\Sigma \\
\end{array}
\]

whence, by proposition [1] we conclude that

\[
\begin{array}{c}
\rho \quad A \\
\end{array} = 
\begin{array}{c}
\Psi \quad B \\
\varphi \quad F \\
\end{array} = 
\begin{array}{c}
\Sigma \\
\end{array}
\]

Therefore \(\Sigma\) is a symmetric purification of \(\rho\). By the uniqueness of the purification, all symmetric purifications of \(\rho\) differ by local reversible transformations. Since \(\rho\) is arbitrary, we conclude that every state has a symmetric purification, unique up to local reversible transformations. ■

**IX. CONCLUSIONS**

In this paper we explored the links between entanglement and thermodynamics in general probabilistic theories. Specifically, we focussed on the relation between LOCC transformations of pure bipartite states and random reversible transformations of mixed single-system states. We associated random-reversible transformations to a resource theory of dynamical controllability, which can be used to order single-system states. The resource theory of controllability can be interpreted as a resource theory of purity when the set of reversible transformations acts transitively on the set of pure states of the system. Using these notions, one can formulate a general duality between bipartite entanglement and an single-system purity, whereby a pure bipartite state is more entangled than another if and only if the marginal states of the latter are purer than the marginal states of the former.

After formulating the duality, we set out to prove it. Our proof follows from a set of operational axioms, including Causality, Purity Preservation, and Purification—plus a new axiom, which we named Local Exchangeability. As the first step in our proof, we established a Lo-Popescu theorem for general probabilistic theories, showing that every LOCC protocol with a pure input state can be replaced by a one-way protocol producing the same output state. Interestingly, this result does not need Purification.

Finally, we used the duality to explore the task of erasing information. We found out that our four axioms are not enough to guarantee that information cannot be erased for free: *a priori*, a theory satisfying the axioms could have entropy sinks, which can absorb mixed states without becoming more mixed themselves. To exclude this possibility, we formulated the absence of entropy sinks as an axiom. Surprisingly, in the context of our four axioms, the No Entropy Sinks requirement turns out to be equivalent to a stronger version of Purification, in which every state admits a purification where the purified and purifying system are identical. This fact strikes us as an additional piece of evidence that Purification plays an essential role in the axiomatic definition of a sensible theory of thermodynamics.

**ACKNOWLEDGMENTS**

This work is supported by the National Basic Research Program of China (973) 2011CBA00300 (2011CBA00301), by the National Natural Science Foundation of China through Grants 11450110096, 11350110207, 61033001, and 61061130540, by the Foundational Questions Institute through the large grant “The fundamental principles of information dynamics”, and by the 1000 Youth Fellowship Program of China.

**Appendix A: Proof of proposition [2]**

Suppose that \(U\psi\) can be written as a coarse-graining as follows

\[
U\psi = \sum_i \rho_i. \quad (A1)
\]

To prove that the state \(U\psi\) is pure, now we show that the refinement \(\{\rho_i\}\) is trivial. Indeed, applying \(U^{-1}\) to both sides of Eq. (A1), we obtain

\[
\psi = \sum_i U^{-1}\rho_i.
\]

Since \(\psi\) is pure, this implies that \(U^{-1}\rho_i = p_i \psi\) for some probability distribution \(\{p_i\}\). Hence, by applying \(U\) on both sides, we obtain \(\rho_i = p_i U\psi\). This concludes the proof that \(\{\rho_i\}\) is a trivial refinement of \(U\psi\) and, therefore, that \(U\psi\) is pure. The converse can be proven...
in the same way by applying the reversible channel $U^{-1}$ to $U\psi$. ■
