Deconstruction and conditional erasure of quantum correlations

Mario Berta
Department of Computing, Imperial College London, London SW7 2AZ, United Kingdom

Fernando G. S. L. Brandão
Institute for Quantum Information and Matter, California Institute of Technology, Pasadena, California 91125, USA

Christian Majenz
Institute for Language, Logic, and Computation, University of Amsterdam, and QuSoft, 1098XG Amsterdam, Netherlands and Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen Ø, Denmark

Mark M. Wilde
Hearne Institute for Theoretical Physics, Department of Physics and Astronomy, Center for Computation and Technology, Louisiana State University, Baton Rouge, Louisiana 70803, USA

(Received 8 August 2018; published 15 October 2018)

We define the deconstruction cost of a tripartite quantum state on systems ABE as the minimum rate of noise needed to apply to the AE systems, such that there is negligible disturbance to the marginal state on the BE systems, while the system A of the resulting state is locally recoverable from the E system alone. We refer to such actions as deconstruction operations and protocols implementing them as state deconstruction protocols. State deconstruction generalizes Landauer erasure of a single-party quantum state as well the erasure of correlations of a two-party quantum state. We find that the deconstruction cost of a tripartite quantum state on systems ABE is equal to its conditional quantum mutual information (CQMI) $I(A;B|E)$, thus giving the CQMI an operational interpretation in terms of a state deconstruction protocol. We also define a related task called conditional erasure, in which the goal is to apply noise to systems AE in order to decouple system A from systems BE, while causing negligible disturbance to the marginal state of systems BE. We find that the optimal rate of noise for conditional erasure is also equal to the CQMI $I(A;B|E)$. State deconstruction and conditional erasure lead to operational interpretations of the quantum discord and squashed entanglement, which are quantum correlation measures based on the CQMI. We find that the quantum discord is equal to the cost of simulating einselection, the process by which a quantum system interacts with an environment, resulting in selective loss of information in the system. The squashed entanglement is equal to half the minimum rate of noise needed for deconstruction and/or conditional erasure if Alice has available the best possible system E to help in the deconstruction and/or conditional erasure task.

DOI: 10.1103/PhysRevA.98.042320

I. INTRODUCTION

The Landauer erasure principle represents a deep link between information theory and thermodynamics [1]. An informal summary of the principle is that the work cost of erasing the contents of a computer memory is proportional to the amount of information stored there. This insight has now sparked a whole literature, a consequence of which has been a deepening of the connection between information theory and thermodynamics (see, e.g., Ref. [2] for a review).

One generalization of Landauer's insight goes beyond the single-system setup mentioned above. In Ref. [3], Groisman *et al.* considered a setting in which two parties share a quantum state ρ_{AB}. Their goal was to determine the work cost of erasing the correlations present in the state, by acting locally on one system, such that the resulting state has a tensor-product form $\sigma_A \otimes \omega_B$, where σ_A and ω_B are quantum states. Groisman *et al.* solved the problem in the framework of quantum Shannon theory [4], whereby they allowed the two parties to have many copies of the state ρ_{AB} and quantified the minimum rate of noise that needs to be applied to the A systems such that the resulting state is a tensor product between the A systems and the B systems. They found that the optimal rate of noise is equal to the quantum mutual information of the state ρ_{AB}, defined as

$$I(A;B)_\rho \equiv H(A)_\rho + H(B)_\rho - H(AB)_\rho, \tag{1}$$

where the quantum entropy of a state σ_G on system G is defined as $H(G)_\sigma \equiv -\operatorname{Tr}\{\sigma_G \log_2 \sigma_G\}$. An important consequence of their theorem is that we can assign a physical meaning to, or operational interpretation of, the quantum mutual information as the minimum rate of noise needed to completely erase the correlations present in a two-party quantum state. Thus, we can say that quantum mutual information is equal to the work cost of correlation destruction.

On the other hand, quantum mutual information has also been interpreted in a communication-theoretic task (now

 ©2018 American Physical Society

called *coherent state merging* [5]) as the optimal rate of entanglement creation when transferring the system A of ρ_{AB} to a party possessing system B [6], while using quantum communication at a fixed rate. These dual interpretations of quantum mutual information in terms of destruction and creation perhaps come at no surprise if one is familiar with the unitarity of quantum mechanics and the purification principle. Information can never truly be destroyed in quantum mechanics, which means that the apparent destruction of correlations between two parties implies the creation of correlations elsewhere, i.e., with another party who possesses a purification of the state ρ_{AB}. In fact, this insight is the main idea underlying the decoupling principle [7,8], which is a method for proving the above theorem [6] and others similar to it.

In this paper, we are interested in further generalizations of the erasure of correlations to a three-system scenario, i.e., for a tripartite quantum state ρ_{ABE} (see also our companion paper [9]). The tasks we are interested in accomplishing are more delicate than the destruction of correlations mentioned above.

The first task we consider is a *state deconstruction* protocol, whose aim is to deconstruct (literally, "to break into constituent components") the correlations in a three-party quantum state. To make the setting precise, consider a state ρ_{ABE}, and suppose that Alice possesses system A, Bob system B, and Eve system E. We would like a deconstruction protocol to result in a state for which Eve is the mediator of correlations between Alice and Bob, while the original correlations shared between Eve and Bob are negligibly disturbed. The setup begins with Alice and Eve in the same laboratory and Bob in a different laboratory, and we also operate in the framework of quantum Shannon theory, allowing them to share n copies of the state ρ_{ABE}, where n can be a large number. Following Groisman *et al.* [3], we allow for a local unitary randomizing channel acting on the AE systems and an ancilla. The rate of noise is equal to the logarithm of the number of unitaries in such a channel divided by the number n of copies of the state ρ_{ABE}. We define the *deconstruction cost* of a tripartite state ρ_{ABE} to be the minimum rate of noise needed to apply to the AE systems and an ancilla, such that the resulting state satisfies the following:

(1) The resulting system of Alice is *locally recoverable* from Eve's system alone, and

(2) the correlations between Eve and Bob are *negligibly disturbed*.

See Sec. IV A for a more detailed definition and Fig. 2 for a depiction of a state deconstruction protocol along with the conditions of local recoverability and negligible disturbance.

The second task we consider is *conditional erasure*. Such a task is very similar to state deconstruction: We allow for a local channel to act on the AE systems and an ancilla. However, we define the conditional erasure cost to be the minimum rate of noise such that the resulting system of Alice is decoupled from the BE systems and the marginal state of the BE systems is negligibly disturbed. A protocol that accomplishes conditional erasure also accomplishes state deconstruction: This is because a decoupled system is locally recoverable.

The negligible disturbance condition is critical in both state deconstruction and conditional erasure: It could be the case that Eve and Bob would want to use their systems for some later quantum information processing task, so that keeping the correlations intact is essential for the systems to be useful later on. For example, Eve's and Bob's systems might contain some entanglement which could be useful for a subsequent distributed quantum computation. This condition also highlights an essential difference between semiclassical and fully quantum protocols: In the case that the system E is classical, the negligible disturbance condition is not necessary because one could always observe the value in Eve's system without causing any disturbance to it. However, in the quantum case, the uncertainty principle forbids us from taking a similar action, so that it is necessary for fully quantum protocols to proceed with a greater sleight of hand.

State deconstruction and conditional erasure are far more delicate than decoupling, the latter sometimes described as having the "relatively indiscriminate goal of destruction" [6]. That is, a naive application of the decoupling method is too blunt of a tool to apply in these protocols. Applying it naively would result in the annihilation of correlations such that if correlations between systems B and E were present beforehand, they would be destroyed and thus no longer useful for a future quantum information processing task.

II. MAIN RESULT

The main result of this paper is that both the deconstruction cost and the conditional erasure cost of a tripartite state ρ_{ABE} are equal to its conditional quantum mutual information (CQMI), defined as

$$I(A;B|E)_\rho \equiv I(AE;B)_\rho - I(E;B)_\rho. \tag{2}$$

(See Theorems 4 and 7.) Thus, our result assigns a physical meaning to the CQMI, in terms of erasure or thermodynamical tasks that generalize Landauer's original scenario as well as the erasure of correlations scenario from Ref. [3]. The deconstruction and conditional erasure tasks are intimately related to properties of the CQMI itself, which has previously been related to local recoverability [10–12] as well as the condition of negligible disturbance [13].

The state deconstruction and conditional erasure tasks are also closely related to the protocol of quantum state redistribution [14,15], which, prior to our contribution, was the only protocol giving an operational meaning for the CQMI. A quantum state redistribution protocol begins with many independent copies of a four-party pure state ψ_{ABER}, with a sender possessing the A and E systems, a receiver possessing the R systems, and the sender and receiver sharing noiseless entanglement before communication begins. The main result of Refs. [14,15] is that the optimal rate of quantum communication needed to redistribute the A systems from the sender to the receiver is equal to $\frac{1}{2}I(A;B|R)_\psi$. In the present paper, the state redistribution protocol is one of the main tools that we use for establishing that the deconstruction and conditional erasure costs are each equal to the CQMI.

The other main tool that we use is a quantity known as the *fidelity of recovery* of a tripartite state ρ_{ABE} [16]:

$$F(A;B|E)_\rho \equiv \sup_{\mathcal{R}_{E\to AE}} F(\rho_{ABE}, \mathcal{R}_{E\to AE}(\rho_{BE})), \tag{3}$$

where the quantum fidelity between states ω and τ is defined as $F(\omega,\tau) \equiv \|\sqrt{\omega}\sqrt{\tau}\|_1^2$ [17] and the supremum is with respect to all recovery channels $\mathcal{R}_{E\to AE}$.

Our main results then lead to operational interpretations of quantum correlation measures based on CQMI, including quantum discord [18,19] and squashed entanglement [20]. We find that the quantum discord is equal to the optimal rate of simulating einselection [21], the process by which a system interacts with an environment in such a way as to cause selective loss of information in the system. In particular, given a bipartite state ρ_{AB} and measurement Λ_A, we find that the discord is equal to the minimum rate of noise needed to apply to the A system of ρ_{AB}, such that the resulting state is locally recoverable after performing a measurement on the A system and its postmeasurement state is indistinguishable from the postmeasurement state after Λ_A acts on ρ_{AB}. We find that the squashed entanglement of a state ρ_{AB} is equal to half the minimum rate of noise needed in a deconstruction operation which has the best possible quantum side information in system E to help in the deconstruction task.

An outline of the rest of the paper is as follows. In Sec. III, we provide more background on quantum information basics and the conditional quantum mutual information, and we review the state redistribution protocol in more detail. Section IV A defines a state deconstruction protocol and the deconstruction cost of a tripartite state ρ_{ABE}, and Sec. IV B discusses a slightly different model for state deconstruction. In Sec. V, we prove that the deconstruction cost is bounded from below by the CQMI. After that, Sec. VI proves the other inequality by showing how a state redistribution protocol leads to one for state deconstruction. In Sec. VII, we define the conditional erasure task and show how a conditional erasure protocol is equivalent to a quantum state redistribution protocol, in the sense that the existence of one implies the existence of the other. We then establish the CQMI as the optimal conditional erasure cost. Section VIII details how quantum discord is equal to the optimal rate of einselection simulation, and the following section gives the aforementioned operational interpretation of squashed entanglement. We finally conclude in Sec. X with a summary and some open questions.

III. BACKGROUND

A. Basics of quantum information

We review some basic aspects of quantum information before proceeding with the main development (see, e.g., Ref. [4] for a review). Let $\mathcal{L}(\mathcal{H})$ denote the algebra of bounded linear operators acting on a Hilbert space $\mathcal{H}$ (we consider finite-dimensional Hilbert spaces throughout this paper). Let $\mathcal{L}_+(\mathcal{H})$ denote the subset of positive semidefinite operators. An operator ρ is in the set $\mathcal{D}(\mathcal{H})$ of density operators (or states) if $\rho \in \mathcal{L}_+(\mathcal{H})$ and $\mathrm{Tr}\{\rho\} = 1$. Throughout this paper, we let π denote the maximally mixed state on a given Hilbert space $\mathcal{H}$, so that $\pi \equiv I_{\mathcal{H}}/\dim(\mathcal{H})$. The tensor product of two Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ is denoted by $\mathcal{H}_A \otimes \mathcal{H}_B$ or $\mathcal{H}_{AB}$. Given a multipartite density operator $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, we unambiguously write $\rho_A = \mathrm{Tr}_B\{\rho_{AB}\}$ for the reduced density operator on system A. We use ρ_{AB}, σ_{AB}, τ_{AB}, ω_{AB}, etc. to denote general density operators in $\mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, while ψ_{AB}, φ_{AB}, ϕ_{AB}, etc. denote rank-1 density operators (pure states) in $\mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ (with it implicit, clear from the context, and the above convention implying that ψ_A, φ_A, ϕ_A may be mixed if ψ_{AB}, φ_{AB}, ϕ_{AB} are pure). A purification $|\phi^\rho\rangle_{RA} \in \mathcal{H}_R \otimes \mathcal{H}_A$ of a state $\rho_A \in \mathcal{D}(\mathcal{H}_A)$ is such that $\rho_A = \mathrm{Tr}_R\{|\phi^\rho\rangle\langle\phi^\rho|_{RA}\}$. An isometry $U : \mathcal{H} \to \mathcal{H}'$ is a linear map such that $U^\dagger U = I_{\mathcal{H}}$. Often, an identity operator is implicit if we do not write it explicitly (and should be clear from the context).

A linear map $\mathcal{N}_{A\to B} : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_B)$ is positive if $\mathcal{N}_{A\to B}(\sigma_A) \in \mathcal{L}(\mathcal{H}_B)_+$ whenever $\sigma_A \in \mathcal{L}(\mathcal{H}_A)_+$. Let id_A denote the identity map acting on a system A. A linear map $\mathcal{N}_{A\to B}$ is completely positive if the map $\mathrm{id}_R \otimes \mathcal{N}_{A\to B}$ is positive for a reference system R of arbitrary size. A linear map $\mathcal{N}_{A\to B}$ is trace preserving if $\mathrm{Tr}\{\mathcal{N}_{A\to B}(\tau_A)\} = \mathrm{Tr}\{\tau_A\}$ for all input operators $\tau_A \in \mathcal{L}(\mathcal{H}_A)$. A quantum channel is a linear map which is completely positive and trace preserving (CPTP). A quantum channel $\mathcal{U} : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_B)$ is an isometric channel if it has the action $\mathcal{U}(X_A) = U X_A U^\dagger$, where $X_A \in \mathcal{L}(\mathcal{H}_A)$ and $U : \mathcal{H}_A \to \mathcal{H}_B$ is an isometry.

The trace distance between two quantum states $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ is equal to $\|\rho - \sigma\|_1$. It has a direct operational interpretation in terms of the distinguishability of these states. That is, if ρ or σ are prepared with equal probability and the task is to distinguish them via some quantum measurement, then the optimal success probability in doing so is equal to $(1 + \|\rho - \sigma\|_1/2)/2$. The trace distance and fidelity are related by the Fuchs–van de Graaf inequalities [22]:

$$1 - \sqrt{F(\rho,\sigma)} \leqslant \tfrac{1}{2}\|\rho - \sigma\|_1 \leqslant \sqrt{1 - F(\rho,\sigma)}. \tag{4}$$

The rightmost quantity above is known to be a distance measure, satisfying the triangle inequality, as proposed and shown in Ref. [23]. This quantity was generalized to subnormalized states and given the name "purified distance" in Ref. [24].

Let $\{|i\rangle_A\}$ denote the standard, orthonormal basis for a Hilbert space $\mathcal{H}_A$, and let $\{|i\rangle_B\}$ be defined similarly for $\mathcal{H}_B$. If the dimensions of these spaces are equal [$\dim(\mathcal{H}_A) = \dim(\mathcal{H}_B) = d$], then we define the maximally entangled state $|\Phi\rangle_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B$ as

$$|\Phi\rangle_{AB} \equiv \frac{1}{\sqrt{d}} \sum_{i=1}^{d} |i\rangle_A \otimes |i\rangle_B. \tag{5}$$

The generalized Pauli shift operator X is defined by $X_A|i\rangle_A = |i \oplus 1\rangle_A$, where addition is modulo d. The generalized Pauli phase operator Z is defined by $Z_A|k\rangle_A = \exp(2\pi i k/d)|k\rangle_A$. The Heisenberg-Weyl group is defined as $\{X_A^j Z_A^k\}_{j,k\in\{1,\ldots,d\}}$, and satisfies

$$\frac{1}{d}\,\mathrm{Tr}\left\{X_A^j Z_A^k\right\} = \delta_{d,j}\delta_{d,k}. \tag{6}$$

The generalized Bell basis is defined as $\{|\Phi^{j,k}\rangle_{AB}\}_{j,k\in\{1,\ldots,d\}}$, where

$$|\Phi^{j,k}\rangle_{AB} = \left(X_A^j Z_A^k \otimes I_B\right)|\Phi\rangle_{AB}. \tag{7}$$

It is an orthonormal basis as a consequence of (6).

B. Conditional quantum mutual information

Here we briefly provide more background on the conditional quantum mutual information (CQMI). The CQMI is understood informally as quantifying the correlations between systems A and B from the perspective of a party possessing system E [14,15]. The CQMI is symmetric with respect to the exchange of the A and B systems of a state ρ_{ABE}: $I(A;B|E)_\rho = I(B;A|E)_\rho$. One of the powerful properties of the CQMI is that it obeys a chain rule of the following form for a state $\sigma_{A_1\cdots A_n BE}$:

$$I(A_1\ldots A_n;B|E)_\sigma = \sum_{i=1}^{n} I\left(A_i;B|EA_1^{i-1}\right)_\sigma, \tag{8}$$

where $A_1^{i-1} \equiv A_1\ldots A_{i-1}$, so that we can think of the correlations between $A_1\ldots A_n$ and B, as observed by E, being built up one system at a time. The CQMI is always nonnegative $I(A;B|E)_\rho \geqslant 0$, an entropy inequality known as strong subadditivity [25,26]. A first relation of CQMI to recoverability was established in Ref. [10], in which it was shown that $I(A;B|E)_\rho = 0$ if and only if there exists a recovery quantum channel $\mathcal{R}_{E\to AE}$ such that the global state ρ_{ABE} can be reconstructed by acting on one share E of the marginal state ρ_{BE}:

$$\rho_{ABE} = \mathcal{R}_{E\to AE}(\rho_{BE}). \tag{9}$$

More recently, it was shown that these results are robust [11,12]: The CQMI is approximately equal to zero [i.e., $I(A;B|E)_\rho \approx 0$] if and only if the global state is approximately recoverable by acting on one share E of the marginal ρ_{BE} [i.e., $\rho_{ABE} \approx \mathcal{R}_{E\to AE}(\rho_{BE})$]. In more detail, Ref. [11] established the inequality

$$I(A;B|E)_\rho \geqslant -\log_2 F(A;B|E)_\rho, \tag{10}$$

and Refs. [11,12] established a converse relation. Using some recent tools [27] and the Fuchs–van de Graaf inequalities in (4), the following refinement of the converse holds [4, Theorem 11.10.5]: If $F(A;B|E)_\rho \geqslant 1-\varepsilon$ for $\varepsilon \in (0,1)$, then

$$I(A;B|E)_\rho \leqslant 2\sqrt{\varepsilon}\log|B| + (1+\sqrt{\varepsilon})h_2(\sqrt{\varepsilon}/[1+\sqrt{\varepsilon}]), \tag{11}$$

where the binary entropy $h_2(x)$ is defined for $x \in (0,1)$ as

$$h_2(x) \equiv -x\log_2 x - (1-x)\log_2(1-x), \tag{12}$$

with the property that $\lim_{x\to 0} h_2(x) = 0$. From the above, we see that the CQMI is a witness to quantum Markovianity: If it is small, then we can understand the correlations between A and B as being mediated by system E via the recovery channel $\mathcal{R}_{E\to AE}$.

C. Quantum state redistribution

This section provides some background on quantum state redistribution [14,15]. A quantum state redistribution protocol begins with a sender, a receiver, and a reference party sharing many independent copies of a four-system pure state ψ_{ABER}. The sender has the AE systems, the receiver the R systems, and the reference the B systems. The goal is to use entanglement and noiseless quantum communication to redistribute the systems such that the sender ends up with the E systems, the receiver the AR systems, and the reference the B systems. As a side benefit, the protocol can also generate entanglement shared between the sender and receiver at the end.

More formally, let $n \in \mathbb{N}$, $M \in \mathbb{N}$, and $\varepsilon \in [0,1]$. An (n, M, ε) state redistribution protocol consists of an encoding channel $\mathcal{E}_{A^nE^nA'\to\bar{A}_0A_0\hat{E}^n}$ and a decoding channel $\mathcal{D}_{\bar{A}_0R'R^n\to\hat{A}^n\hat{R}^nR_0}$, such that the following state

$$\xi_{\hat{A}^nB^n\hat{E}^n\hat{R}^nA_0R_0} \equiv \mathcal{D}_{\bar{A}_0R'R^n\to\hat{A}^n\hat{R}^nR_0}(\varphi_{\bar{A}_0A_0\hat{E}^nB^nR^nR'}), \tag{13}$$

where

$$\varphi_{\bar{A}_0A_0\hat{E}^nB^nR^nR'} \equiv \mathcal{E}_{A^nE^nA'\to\bar{A}_0A_0\hat{E}^n}\left(\psi_{ABER}^{\otimes n}\otimes\Phi_{A'R'}\right), \tag{14}$$

has fidelity larger than $1-\varepsilon$ with the following pure state:

$$\psi_{\hat{A}B\hat{E}\hat{R}}^{\otimes n}\otimes\Phi_{A_0R_0}, \tag{15}$$

where $\Phi_{A'R'}$ and $\Phi_{A_0R_0}$ denote maximally entangled states of Schmidt ranks $|A'|$ and $|A_0|$, respectively. That is, an (n, M, ε) state redistribution protocol satisfies

$$F\left(\xi_{\hat{A}^nB^n\hat{E}^n\hat{R}^nA_0R_0}, \psi_{\hat{A}B\hat{E}\hat{R}}^{\otimes n}\otimes\Phi_{A_0R_0}\right) \geqslant 1-\varepsilon. \tag{16}$$

The parameter M is the dimension of the quantum system $\bar{A}_0$ that is communicated from sender to receiver:

$$M \equiv |\bar{A}_0|. \tag{17}$$

Definition 1 (Achievable rate). A rate R is *achievable* for state redistribution of ψ_{ABER} if for all $\varepsilon \in (0,1)$, $\delta > 0$, and sufficiently large n, there exists an $(n, 2^{n[R+\delta]}, \varepsilon)$ state redistribution protocol.

Definition 2 (Quantum communication cost). The *quantum communication cost* $\mathcal{Q}(\psi_{ABER})$ of state redistribution of ψ_{ABER} is equal to the infimum of all rates which are achievable for redistribution of ψ_{ABER}.

The following theorem from Refs. [14,15] gives a precise characterization of the quantum communication cost:

Theorem 1 ([14,15]). The quantum communication cost of state redistribution is equal to half the conditional quantum mutual information:

$$\mathcal{Q}(\psi_{ABER}) = \tfrac{1}{2}I(A;B|R)_\psi. \tag{18}$$

The achievability part of the above theorem was simplified in Ref. [28], which is the formulation of state redistribution that we will use to characterize deconstruction cost.

Remark 1. The results of Refs. [28–30] establish that the encoding channel and decoding channel for state redistribution can be chosen as unitaries, a key fact that we will use in what follows. Let $U^{\mathcal{E}}_{A^nE^nA'\to\bar{A}_0A_0\hat{E}^n}$ denote the unitary encoder and $U^{\mathcal{D}}_{\bar{A}_0R'R^n\to\hat{A}^n\hat{R}^nR_0}$ the unitary decoder for these protocols, and note that the state $\xi_{\hat{A}^nB^n\hat{E}^n\hat{R}^nA_0R_0}$ in (13) can be taken as a pure state as a consequence. See Fig. 1 for a depiction of such a state redistribution protocol.

We can also quantify the entanglement cost of a quantum state redistribution protocol. In such a case, for $L \in \mathbb{N}$, we define an (n, M, L, ε) quantum state redistribution protocol specified exactly as given above, except we set

$$L \equiv |A'|/|A_0|. \tag{19}$$

With this convention, there is an entanglement cost if $L \geqslant 1$ and there is an entanglement gain if $L \leqslant 1$. A rate pair

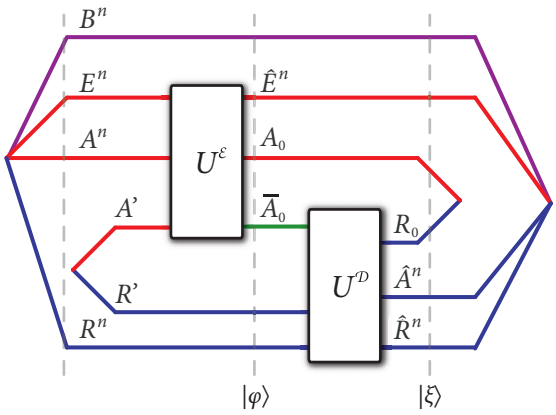


FIG. 1. Quantum state redistribution with a unitary encoding and decoding. By using shared entanglement in systems A' and R' and noiseless quantum communication of the system $\bar{A}_0$, a sender can transfer her quantum systems A^n to a receiver, such that the resulting state of systems $\hat{A}^n B^n \hat{R}^n \hat{E}^n$ has arbitrarily high fidelity with the initial state of systems $A^n B^n R^n E^n$. At the same time, the protocol generates entanglement in the registers A_0 and R_0.

(R, E) is *achievable* for state redistribution of ψ_{ABER} if for all $\varepsilon \in (0, 1)$, $\delta > 0$, and sufficiently large n, there exists an $(n, 2^{n[R+\delta]}, 2^{n[E+\delta]}, \varepsilon)$ state redistribution protocol. The achievable rate region of state redistribution of ψ_{ABER} is equal to the union of all rate pairs which are achievable for redistribution of ψ_{ABER}.

References [14,15] proved that the rate pair

$$(I(A;B|R)_\psi/2, [I(A;E)_\psi - I(A;R)_\psi]/2) \tag{20}$$

is achievable and that the optimal rate region is equal to

$$R \geqslant \tfrac{1}{2} I(A;B|R)_\psi, \tag{21}$$

$$R + E \geqslant H(A|R)_\psi. \tag{22}$$

Thus, the rate pair in (20) corresponds to an optimal corner point of the region in (21) and (22). The protocol from Ref. [28] consumes entanglement at a rate equal to $I(A;E)_\psi/2$ and generates entanglement at a rate equal to $I(A;R)_\psi/2$.

IV. STATE DECONSTRUCTION PROTOCOL

Here we provide an operational definition for the *deconstruction cost* of a tripartite state ρ_{ABE}. We frame the problem in the formalism of quantum Shannon theory [4], which, as we will show, ultimately leads to the CQMI being equal to the deconstruction cost after taking a limit. In what follows, we consider two seemingly different models, called the local unitary randomizing model and the Landauer-Bennett erasure model. In Sec. IV C, we show that these two models are in fact equivalent to each other, in the sense that a protocol from one model can simulate a protocol from the other, with the same resource consumption and performance.

A. Local unitary randomizing model

We begin by defining a state deconstruction protocol in the local unitary randomizing model. Let $n \in \mathbb{N}$, $M \in \mathbb{N}$, and $\varepsilon \in [0, 1]$. An (n, M, ε) state deconstruction protocol consists of an ensemble of M unitaries $\{p_i, U^i_{A^n A' E^n}\}_{i=1}^M$ that lead to the following local unitary randomizing channel:

$$\mathcal{N}_{A^n A' E^n}(\tau_{A^n A' E^n}) \equiv \sum_i p_i U^i_{A^n A' E^n} \tau_{A^n A' E^n} \left(U^i_{A^n A' E^n}\right)^\dagger, \tag{23}$$

for a density operator $\tau_{A^n A' E^n}$, with system A' an auxiliary system. We also refer to such an action as an *ε-deconstruction operation* and are interested in its action on the state $\rho_{ABE}^{\otimes n} \otimes \theta_{A'}$, where $\theta_{A'}$ is an auxiliary density operator that plays the role of a catalyst in the sense of Ref. [31] to help in the deconstruction task. The state resulting from a deconstruction operation acting on $\rho_{ABE}^{\otimes n} \otimes \theta_{A'}$ is as follows:

$$\omega_{A^n A' B^n E^n} \equiv \mathcal{N}_{A^n A' E^n}\left(\rho_{ABE}^{\otimes n} \otimes \theta_{A'}\right). \tag{24}$$

We demand for such a deconstruction operation to satisfy the property of *negligible disturbance* and for the state resulting from the operation to be *locally recoverable*. In particular, the negligible disturbance condition means that the deconstruction operation $\mathcal{N}_{A^n A' E^n}$ causes little disturbance to the residual state of the $B^n E^n$ systems, in the sense that

$$F\left(\omega_{B^n E^n}, \rho_{BE}^{\otimes n}\right) \geqslant 1 - \varepsilon. \tag{25}$$

The condition of local recoverability means that the resulting state $\omega_{A^n A' B^n E^n}$ is such that the $A^n A'$ systems are locally recoverable by acting on the E^n systems alone. That is, there exists a recovery channel $\mathcal{R}_{E^n \to A^n A' E^n}$ such that

$$F(\omega_{A^n A' B^n E^n}, \mathcal{R}_{E^n \to A^n A' E^n}(\omega_{B^n E^n})) \geqslant 1 - \varepsilon. \tag{26}$$

Equivalently, we demand for the following fidelity of recovery to be large:

$$F(A^n A'; B^n | E^n)_\omega \geqslant 1 - \varepsilon. \tag{27}$$

Figure 2 depicts a state deconstruction protocol in the local unitary randomizing model.

Definition 3 (Achievable rate). A rate R is *achievable* for state deconstruction of ρ_{ABE} if for all $\varepsilon \in (0, 1)$, $\delta > 0$, and sufficiently large n, there exists an $(n, 2^{n[R+\delta]}, \varepsilon)$ state deconstruction protocol.

Definition 4 (Deconstruction cost). The *deconstruction cost* $\mathcal{D}(A;B|E)_\rho$ of a state ρ_{ABE} is equal to the infimum of all rates which are achievable for state deconstruction of ρ_{ABE}.

Remark 2. [12, Proposition 35] and [4, Theorem 11.10.5] implies that the deconstruction cost of ρ_{ABE} is equal to the minimum rate of noise needed to deconstruct the correlations in $\rho_{ABE}^{\otimes n}$ in such a way that the resulting state has vanishing normalized CQMI. Specifically, the state $\omega_{A^n A' B^n E^n}$ resulting from an (n, M, ε) state deconstruction protocol is such that

$$\frac{1}{n} I(A^n A'; B^n | \hat{E}^n)_\omega \leqslant 2\sqrt{\varepsilon} \log |B| + \frac{1}{n}(1 + \sqrt{\varepsilon}) h_2(\sqrt{\varepsilon}/[1 + \sqrt{\varepsilon}]). \tag{28}$$

Remark 3. Operational tasks related to state deconstruction were previously explored in Ref. [32], where a class of "Markovianizing operations" were defined and subsequently

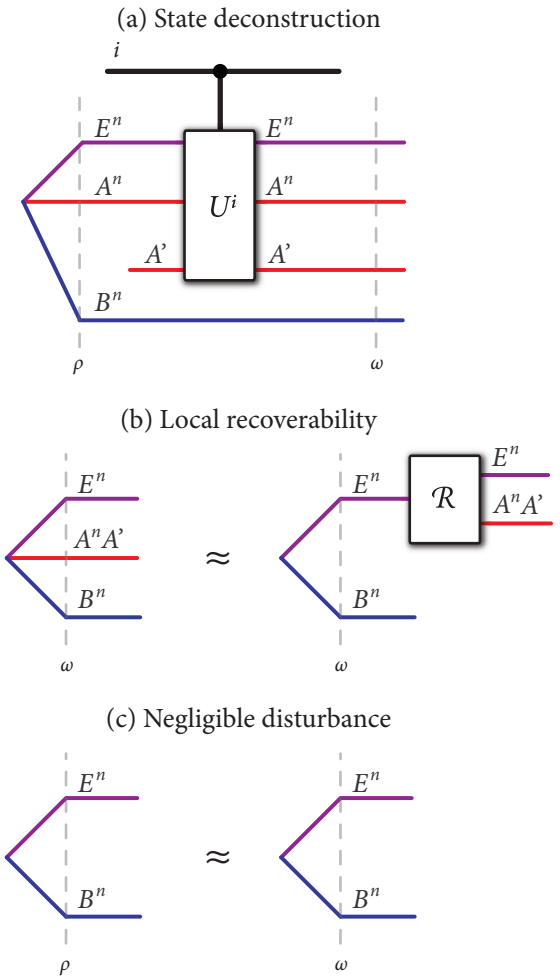


FIG. 2. Depiction of (a) a state deconstruction protocol in the local unitary randomizing model along with the conditions of (b) local recoverability and (c) negligible disturbance.

broadened in Refs. [33,34]. Deconstruction operations are different in that we allow for a catalyst, a unitary interaction between the A^nE^n systems and the catalyst, and we demand for the condition of negligible disturbance to hold. Whereas our converse (Theorem 2) holds for the model of Ref. [32] as well, the CQMI cannot be achieved: The fact that Ref. [32] does not allow for an interaction with the E systems leads to a strictly larger optimal rate function based on the Koashi-Imoto decomposition [35] (at least for pure states). This proves that the CQMI cannot be achieved without having access to the E systems. The result of Ref. [32] is motivated from questions in distributed computation [36] but has the disadvantage that the Koashi-Imoto decomposition is not continuous in the state.

Remark 4. In Appendix B, we give a strictly classical example that demonstrates how the conditional mutual information cannot be achieved without having access to the E systems.

B. Landauer-Bennett erasure model

We can think of deconstruction operations in an alternative way, akin to the Landauer-Bennett model of erasure [1,37] and discussed in Ref. [3, Remark II.4], in which we interact the systems of interest unitarily (reversibly) with a catalyst and subsequently perform a partial trace over some subsystem. The deconstruction cost in this case is then related to the size of the system that we trace out. In this alternative model, we define a deconstruction operation $\mathcal{N}_{A^nE^n\to A'_1\hat{E}^n}$ as

$$\omega_{A'_1B^n\hat{E}^n} \equiv \mathcal{N}_{A^nE^n\to A'_1\hat{E}^n}\left(\rho_{ABE}^{\otimes n}\right) \tag{29}$$

$$\equiv \mathrm{Tr}_{A'_2}\left\{\mathcal{U}_{A^nE^nA'\to A'_1A'_2\hat{E}^n}\left(\rho_{ABE}^{\otimes n}\otimes\theta_{A'}\right)\right\}, \tag{30}$$

with $\theta_{A'}$ an arbitrary ancilla state and $\mathcal{U}_{A^nE^nA'\to A'_1A'_2\hat{E}^n}$ a unitary quantum channel. An (n, M, ε) deconstruction protocol in this case has n defined again as the number of copies of ρ_{ABE} and ε defined via (25) and $F(A'_1; B^n|\hat{E}^n)_\omega \geqslant 1-\varepsilon$. However, in this Landauer-Bennett erasure model, we take M defined as

$$M \equiv |A'_2|^2. \tag{31}$$

In this model, we take the convention of squaring the dimension of the removed system $|A'_2|^2$ when calculating M, because we are interested in measuring the amount of noise needed to remove the A'_2 system (i.e., the amount of noise needed to physically implement a partial trace). One way to do so is to apply a randomizing channel of the following form, which realizes a partial trace:

$$\frac{1}{|A'_2|^2}\sum_{i=1}^{|A'_2|^2} V^i_{A'_2}\mathcal{U}_{A^nE^nA'\to A'_1A'_2\hat{E}^n}\left(\rho_{ABE}^{\otimes n}\otimes\theta_{A'}\right)\left(V^i_{A'_2}\right)^\dagger$$

$$= \pi_{A'_2}\otimes\mathrm{Tr}_{A'_2}\left\{\mathcal{U}_{A^nE^nA'\to A'_1A'_2\hat{E}^n}\left(\rho_{ABE}^{\otimes n}\otimes\theta_{A'}\right)\right\}, \tag{32}$$

where $\{V^i_{A'_2}\}_{i=1}^{|A'_2|^2}$ is a unitary one design and $\pi_{A'_2}\equiv I_{A'_2}/|A'_2|$ is the maximally mixed state. It is known that $|A'_2|^2$ unitaries are necessary and sufficient for physically implementing a partial trace in the above sense [38].

We can then define achievable rates and the deconstruction cost for this alternative model just as in Definitions 3 and 4. This model might seem as if it is slightly different from the local unitary randomizing one, but we show in the next section that they are equivalent and thus lead to the same deconstruction cost.

C. Equivalence of the two models

In this section, we show that the local unitary randomizing model and the Landauer-Bennett erasure models are equivalent, in the sense that they can simulate one another with the same performance and resource consumption. This equivalence was shown for a special case in Ref. [31], and here we generalize the argument to the settings considered in this paper. As a consequence of our simulation argument, there is no need to consider two different notions of deconstruction cost, since the simulation argument implies that the costs are in fact the same.

First, we show that the local unitary randomizing model can simulate the Landauer-Bennett erasure model. To this end, suppose that we are given a catalyst state $\theta_{A'}$ and an interaction unitary $U_{A^nE^nA'\to A'_1A'_2\hat{E}^n}$, such that the Landauer-Bennett erasure deconstruction operation is as given in (30). We can simulate such an operation by choosing an ensemble

of unitaries to be as follows:

$$\left\{1/|A_2'|^2,\, W^i_{A^nE^nA'\to A_1'A_2'\hat{E}^n}\right\}_{i=1}^{|A_2'|^2}, \tag{33}$$

where

$$W^i_{A^nE^nA'\to A_1'A_2'\hat{E}^n} \equiv V^i_{A_2'} U_{A^nE^nA'\to A_1'A_2'\hat{E}^n} \tag{34}$$

and $\{V^i_{A_2'}\}_{i=1}^{|A_2'|^2}$ is a set of Heisenberg-Weyl unitaries that realize a partial trace. The result is that a local unitary randomizing channel in (23) formed from the ensemble in (33) can realize the deconstruction operation in (30):

$$\begin{aligned}&\frac{1}{|A_2'|^2}\sum_i W^i\left(\rho^{\otimes n}_{ABE}\otimes\theta_{A'}\right)W^{i\dagger}\\ &\quad=\pi_{A_2'}\otimes \mathrm{Tr}_{A_2'}\left\{\mathcal{U}_{A^nE^nA'\to A_1'A_2'\hat{E}^n}\left(\rho^{\otimes n}_{ABE}\otimes\theta_{A'}\right)\right\}\\ &\quad\equiv \pi_{A_2'}\otimes\omega_{A_1'B^n\hat{E}^n}.\end{aligned} \tag{35}$$

Both the negligible disturbance and the local recoverability conditions hold with the same quality as in the original protocol. This is clear for the negligible disturbance condition, and to see it for the local recoverability condition, we can invoke a special case of the multiplicativity of fidelity of recovery with respect to tensor-product states [39]:

$$F(A_1'A_2';B^n|\hat{E}^n)_{\pi\otimes\omega} = F(A_1';B^n|\hat{E}^n)_{\omega}. \tag{36}$$

Showing the other simulation requires a bit more effort. To this end, consider an arbitrary ensemble of unitaries $\{p_i, U^i_{A^nA'E^n}\}_{i=1}^M$ and an ancilla $\theta_{A'}$. We need to show how it is possible to simulate the effect of a local unitary randomizing channel of the form in (23) built from this ensemble, by bringing in an ancilla state, performing a global unitary, and ending with a partial trace. We take the ancilla to be the following state:

$$\pi_{S_A}\otimes\pi_{T_A}\otimes\sum_{i=1}^M p_i|i\rangle\langle i|_{\hat{M}_A}\otimes\theta_{A'}, \tag{37}$$

where S_A and T_A are quantum systems each having dimension equal to $\sqrt{M}$. (Note that if $\sqrt{M}$ is not an integer, then we can "zero-pad" the probability distribution $\{p_i\}$ such that its cardinality becomes a power of 2—this has the negligible effect of incrementing by one the number of bits needed to describe the indices i corresponding to the entries of the probability distribution $\{p_i\}$ and at the same time ensures that $\sqrt{M}$ is an integer.) It is helpful to recall the following equality:

$$\pi_{S_A}\otimes\pi_{T_A} = \frac{1}{M}\sum_{j,k}|\Phi^{j,k}\rangle\langle\Phi^{j,k}|_{S_AT_A}, \tag{38}$$

where $\{|\Phi^{j,k}\rangle_{S_AT_A}\}$ denotes the Bell basis reviewed in Sec. III A. We take the unitary interaction between the ancilla systems $S_AT_A\hat{M}_AA'$ and the data systems A^nE^n to be a serial concatenation of the following two controlled unitaries:

$$\sum_i |i\rangle\langle i|_{\hat{M}_A}\otimes U^i_{A^nA'E^n}, \tag{39}$$

$$\sum_{j,k}|\Phi^{j,k}\rangle\langle\Phi^{j,k}|_{S_AT_A}\otimes X^{(j-1)\cdot d+k}_{\hat{M}_A}. \tag{40}$$

The state resulting from applying these two controlled unitaries sequentially [(39) and then (40)] to the systems $S_AT_A\hat{M}_AA^nE^nF$ is as follows:

$$\begin{aligned}&\frac{1}{M}\sum_{j,k}|\Phi^{j,k}\rangle\langle\Phi^{j,k}|_{S_AT_A}\\ &\quad\otimes\sum_i p_i X^{(j-1)\cdot d+k}_{\hat{M}_A}|i\rangle\langle i|_{\hat{M}_A}\left[X^{(j-1)\cdot d+k}_{\hat{M}_A}\right]^\dagger\\ &\quad\otimes U^i_{A^nA'E^n}\left(\rho^{\otimes n}_{ABE}\otimes\theta_{A'}\right)\left(U^i_{A^nA'E^n}\right)^\dagger.\end{aligned} \tag{41}$$

After tracing over the S_A register, which requires $\log M$ bits of noise according to our convention in (31), the state becomes as follows:

$$\begin{aligned}&\pi_{T_A}\otimes\pi_{\hat{M}_A}\otimes\sum_i p_i U^i_{A^nA'E^n}\left(\rho^{\otimes n}_{ABE}\otimes\theta_{A'}\right)\left(U^i_{A^nA'E^n}\right)^\dagger\\ &\quad\equiv\pi_{T_A}\otimes\pi_{\hat{M}_A}\otimes\omega_{A^nA'B^nE^n}.\end{aligned} \tag{42}$$

One can verify this explicitly, or see that it follows intuitively from a cascade: Tracing over system S_A has the effect of "forgetting" j and k, which has the effect of randomizing the classical system $\hat{M}_A$ with a uniform mixture of the shift operators $X^{(j-1)\cdot d+k}_{\hat{M}_A}$, which in turn has the effect of "forgetting" i, which then applies the local unitary randomizing channel to the systems $A^nA'E^n$. Both the negligible disturbance and the local recoverability conditions hold with the same quality as in the original protocol. This is clear for the negligible disturbance condition, and to see it for the local recoverability condition, we can invoke a special case of the multiplicativity of fidelity of recovery with respect to tensor-product states [39]:

$$F(T_A\hat{M}_AA^nA';B^n|E^n)_{\pi\otimes\pi\otimes\omega} = F(A^nA';B^n|E^n)_{\omega}. \tag{43}$$

V. DECONSTRUCTION COST IS LOWER BOUNDED BY CQMI

In this section, we prove that the deconstruction cost of a tripartite state ρ_{ABE} is lower bounded by its conditional quantum mutual information $I(A;B|E)_\rho$. We prove such a converse theorem in the Landauer-Bennett erasure model. By the simulation argument given in Sec. IV C, this theorem also serves as a converse bound for deconstruction cost in the local unitary randomizing model. For the interested reader, Appendix A offers two alternative converse proofs for optimality of the deconstruction cost in the local unitary randomizing model. One of them has a flavor similar to the converse proof given below, and the other is similar to those from prior works [3,33,34].

Theorem 2. The conditional quantum mutual information $I(A;B|E)_\rho$ of a tripartite state ρ_{ABE} is a lower bound on its deconstruction cost $\mathcal{D}(A;B|E)_\rho$:

$$I(A;B|E)_\rho \leqslant \mathcal{D}(A;B|E)_\rho. \tag{44}$$

Proof. To prove this theorem, we employ entropy inequalities and properties of CQMI. Consider a general (n, M, ε) Landauer-Bennett state deconstruction protocol as outlined in

Sec. IV B. Then the following chain of inequalities holds

$$\begin{aligned} & nI(A;B|E)_\rho \\ & = I(A^n;B^n|E^n)_{\rho^{\otimes n}} \\ & = H(B^n|E^n)_{\rho^{\otimes n}} - H(B^n|A^nE^n)_{\rho^{\otimes n}} \\ & = H(B^n|E^n)_{\rho^{\otimes n}} - H(B^n|A^nA'E^n)_{\rho^{\otimes n}\otimes\theta} \\ & \leqslant H(B^n|\hat{E}^n)_\omega + f(n,\varepsilon) - H(B^n|A_1'A_2'\hat{E}^n)_{\mathcal{U}(\rho^{\otimes n}\otimes\theta)} \\ & \leqslant H(B^n|\hat{E}^n)_\omega + f(n,\varepsilon) - H(B^n|A_1'\hat{E}^n)_\omega + 2\log_2|A_2'| \\ & = 2\log_2|A_2'| + I(A_1';B^n|\hat{E}^n)_\omega + f(n,\varepsilon) \\ & \leqslant 2\log_2|A_2'| + g(n,\varepsilon) + f(n,\varepsilon). \end{aligned} \tag{45}$$

The first equality follows because the CQMI is additive with respect to tensor-product states. The second equality follows from the definition of CQMI. The third equality follows because the conditional entropy is invariant with respect to tensoring in a product state to be part of the conditioning system. The first inequality follows because the conditional entropy is invariant with respect to a local unitary acting on the conditioning system:

$$H(B^n|A^nA'E^n)_{\rho^{\otimes n}\otimes\theta} = H(B^n|A_1'A_2'\hat{E}^n)_{\mathcal{U}(\rho^{\otimes n}\otimes\theta)}. \tag{46}$$

Also, we have applied the negligible disturbance condition from (25), the Fuchs–van de Graaf inequalities in (4), and the continuity of conditional entropy [27,40], with

$$f(n,\varepsilon) = 2\sqrt{\varepsilon}n\log|B| + (1+\sqrt{\varepsilon})h_2(\sqrt{\varepsilon}/[1+\sqrt{\varepsilon}]). \tag{47}$$

The second inequality follows from a rewriting and applying a dimension bound for CQMI (see, e.g., Ref. [4, Exercise 11.7.9]):

$$\begin{aligned} & H(B^n|A_1'\hat{E}^n)_{\mathcal{U}(\rho^{\otimes n}\otimes\theta)} - H(B^n|A_1'A_2'\hat{E}^n)_{\mathcal{U}(\rho^{\otimes n}\otimes\theta)} \\ & = I(B^n;A_2'|A_1'\hat{E}^n)_{\mathcal{U}(\rho^{\otimes n}\otimes\theta)} \leqslant 2\log_2|A_2'|. \end{aligned} \tag{48}$$

The last equality follows from the definition of CQMI. The final inequality follows by applying the local recoverability condition $F(A_1';B^n|\hat{E}^n)_\omega \geqslant 1-\varepsilon$ and because locally recoverable states have small CQMI as reviewed in (11). In particular, we can take

$$g(n,\varepsilon) \equiv 2n\sqrt{\varepsilon}\log|B| + (1+\sqrt{\varepsilon})h_2(\sqrt{\varepsilon}/[1+\sqrt{\varepsilon}]). \tag{49}$$

Thus, recalling our convention that $M = |A_2'|^2$, we conclude that the following bound holds for any (n, M, ε) state deconstruction protocol:

$$I(A;B|E)_\rho \leqslant \tfrac{1}{n}\log_2 M + \tfrac{1}{n}[g(n,\varepsilon) + f(n,\varepsilon)]. \tag{50}$$

By taking the limit as $n\to\infty$, then $\varepsilon\to 0$, and applying definitions, we can conclude the inequality $I(A;B|E)_\rho \leqslant \mathcal{D}(A;B|E)_\rho$. ∎

VI. FROM STATE REDISTRIBUTION TO STATE DECONSTRUCTION

To show that the deconstruction cost is achievable [i.e., that $\mathcal{D}(A;B|E)_\rho \leqslant I(A;B|E)_\rho$], we employ the quantum state redistribution protocol, reviewed in Sec. III C. We begin by proving that a state redistribution protocol implies the existence of a state deconstruction protocol.

Theorem 3. An (n, M, ε) protocol for state redistribution of a four-system pure state ψ_{ABER}, as specified in Sec. III C, realizes an $(n, M^2, 4\varepsilon)$ protocol for state deconstruction of $\rho_{ABE} = \mathrm{Tr}_R\{\psi_{ABER}\}$, as specified in Sec. IV.

Proof. Let ψ_{ABER} be a purification of ρ_{ABE}. Given is an (n, M, ε) state redistribution protocol, which by Remark 1 means that there is a unitary encoder $U^{\mathcal{E}}_{A^nE^nA'\to\bar{A}_0A_0\hat{E}^n}$ and a unitary decoder $U^{\mathcal{D}}_{\bar{A}_0R'R^n\to\hat{A}^n\hat{R}^nR_0}$ satisfying (16). We will show the existence of an $(n, M^2, 4\varepsilon)$ protocol for state deconstruction of ρ_{ABE} in the Landauer-Bennett erasure model. By the monotonicity of fidelity with respect to partial trace over the systems $\hat{A}^n\hat{R}^nR_0$ [4, Lemma 9.2.1], Eq. (16) implies that

$$F\left(\xi_{A_0B^n\hat{E}^n}, \pi_{A_0}\otimes\rho_{B\hat{E}}^{\otimes n}\right) \geqslant 1-\varepsilon. \tag{51}$$

In our protocol for state deconstruction, we take the deconstruction operation to be

(1) tensoring in the maximally mixed state $\pi_{A'}$,
(2) application of the unitary $U^{\mathcal{E}}_{A^nE^nA'\to\bar{A}_0A_0\hat{E}^n}$, and
(3) a partial trace over the $\bar{A}_0$ system.

Let

$$\begin{aligned} \omega_{A_0B^n\hat{E}^n} &\equiv \mathrm{Tr}_{\bar{A}_0}\left\{U^{\mathcal{E}}\left(\rho_{AB\hat{E}}^{\otimes n}\otimes\pi_{A'}\right)U^{\mathcal{E}\dagger}\right\} && (52)\\ &= \xi_{A_0B^n\hat{E}^n}, && (53)\end{aligned}$$

where $U^{\mathcal{E}} \equiv U^{\mathcal{E}}_{A^nE^nA'\to\bar{A}_0A_0\hat{E}^n}$.

Now we show that the protocol satisfies the requirements of negligible disturbance and local recoverability, as outlined in Sec. IV B. The condition of negligible disturbance follows directly from (51), after a partial trace over system A_0, because

$$\xi_{B^n\hat{E}^n} = \mathrm{Tr}_{A_0}\left\{\omega_{A_0B^n\hat{E}^n}\right\}. \tag{54}$$

The condition of local recoverability follows rather directly as well from (51). If the system A_0 is lost, then the remaining state is $\xi_{B^n\hat{E}^n}$. We can then take the recovery channel to merely tensor in a maximally mixed state π_{A_0}, and (51) guarantees that the resulting state is close to the original one. Indeed, by employing the fact that $\sqrt{1-F(\rho,\sigma)}$ is a distance measure [23] and thus obeys the triangle inequality, we find that

$$\begin{aligned} & \sqrt{1-F(\xi_{A_0B^n\hat{E}^n}, \pi_{A_0}\otimes\xi_{B^n\hat{E}^n})} \\ & \leqslant \sqrt{1-F\left(\xi_{A_0B^n\hat{E}^n}, \pi_{A_0}\otimes\rho_{B\hat{E}}^{\otimes n}\right)} \\ & \quad + \sqrt{1-F\left(\pi_{A_0}\otimes\rho_{B\hat{E}}^{\otimes n}, \pi_{A_0}\otimes\xi_{B^n\hat{E}^n}\right)} \leqslant 2\sqrt{\varepsilon}, \end{aligned} \tag{55}$$

where the second inequality follows from (51) and the fact that

$$F\left(\pi_{A_0}\otimes\rho_{B\hat{E}}^{\otimes n}, \pi_{A_0}\otimes\xi_{B^n\hat{E}^n}\right) = F\left(\rho_{B\hat{E}}^{\otimes n}, \xi_{B^n\hat{E}^n}\right) \geqslant 1-\varepsilon. \tag{56}$$

Then we find that

$$F\left(\xi_{A_0B^n\hat{E}^n}, \pi_{A_0}\otimes\xi_{B^n\hat{E}^n}\right) \geqslant 1-4\varepsilon, \tag{57}$$

concluding the proof. ∎

The following is then a direct corollary of Theorem 3, the definitions of state redistribution and state deconstruction in Secs. III C and IV, respectively, and Theorem 1:

Corollary 1. The deconstruction cost $\mathcal{D}(A;B|E)_\rho$ of a tripartite state ρ_{ABE} is bounded from above by its CQMI

$I(A;B|E)_\rho$:

$$\mathcal{D}(A;B|E)_\rho \leqslant I(A;B|E)_\rho. \tag{58}$$

As a consequence of Theorem 2 and Corollary 1, we can conclude one of our main results, as stated at the beginning of Sec. II.

Theorem 4. The deconstruction cost $\mathcal{D}(A;B|E)_\rho$ of a tripartite state ρ_{ABE} is equal to its CQMI $I(A;B|E)_\rho$:

$$\mathcal{D}(A;B|E)_\rho = I(A;B|E)_\rho. \tag{59}$$

A. Special case of classical side information

The state deconstruction protocol can be simplified in the case that the system E is classical. If this is the case, then the tripartite state ρ_{ABE} has the form $\rho_{ABE} = \sum_e p_E(e)\rho^e_{AB} \otimes |e\rangle\langle e|_E$, where $p_E(e)$ is a probability distribution, $\{\rho^e_{AB}\}$ is a set of states, $\{|e\rangle_E\}$ is an orthonormal basis, and the symbol e is chosen from an alphabet $\mathcal{E}$. In this case, we have

$$\rho^{\otimes n}_{ABE} = \sum_{e^n} p_{E^n}(e^n)\rho^{e^n}_{A^nB^n} \otimes |e^n\rangle\langle e^n|_{E^n}, \tag{60}$$

$$p_{E^n}(e^n) \equiv \prod_{j=1}^{n} p_E(e_j), \tag{61}$$

$$\rho^{e^n}_{A^nB^n} = \rho^{e_1}_{A_1B_1} \otimes \cdots \otimes \rho^{e_n}_{A_nB_n}, \tag{62}$$

$$|e^n\rangle_{E^n} = |e_1\rangle_{E_1} \otimes \cdots \otimes |e_n\rangle_{E_n}. \tag{63}$$

The protocol proceeds by performing a typical subspace measurement of the systems E^n [4], keeping only the classical sequences which are typical (i.e., those with empirical distribution close to the distribution p_E). All such sequences can be partitioned into $|\mathcal{E}|$ blocks, each consisting of the same symbol $e \in \mathcal{E}$ and with length $\approx np_E(e)$. For each block, we then employ the erasure of correlations protocol from Ref. [3], which implies that $\approx np_E(e)I(A;B)_{\rho^e}$ bits of noise are used to erase the correlations in a given block. Thus, the total rate of noise needed in this case is equal to $\sum_e p_E(e)I(A;B)_{\rho^e} = I(A;B|E)_\rho$. The above protocol falls into the class of deconstruction operations because it causes zero disturbance to the marginal state on systems B^nE^n. Furthermore, the state afterward is locally recoverable. The result of the erasure of correlations protocol is to produce a state close to one of the form $\sum_{e^n} p_{E^n}(e^n)\omega^{e^n}_{A^n} \otimes \omega^{e^n}_{B^n} \otimes |e^n\rangle\langle e^n|_{E^n}$, for which the recovery procedure is clear: If system A^n gets lost, look in system E^n for the classical sequence e^n and then prepare the state $\omega^{e^n}_{A^n}$ in the A systems.

One further observation is that the protocol given above does not require access to a catalyst in this special case. It is largely open to determine whether a catalyst is actually needed in the fully quantum case (i.e., when the E system does not admit a classical description).

VII. CONDITIONAL ERASURE

We now turn to conditional erasure and begin by providing an operational definition of a conditional erasure protocol, doing so in the Landauer-Bennett erasure model from Sec. IV B. There are some similarities between state deconstruction and conditional erasure, but in our development for conditional erasure, we also quantify the rate of noise being consumed or generated by a given protocol. To this end, we distinguish and quantify two types of noise, which we call active noise and passive noise.

Active noise is synonymous with a partial trace in the Landauer-Bennett erasure model from Sec. IV B. The amount of active noise being applied in the operation in (30) is equal to $M = |A'_2|^2$ and the rate of active noise is equal to $[\log_2 M]/n$. We use the term *active noise* to describe this kind of noise because one needs to apply a physical procedure, consisting of local randomizing unitaries, in order to implement an active noise operation and realize a partial trace.

Passive noise is synonymous with a catalyst that is brought in to help accomplish an erasure task. Here, we consider passive noise as a resource and quantify it as follows: The amount of passive noise is equal to the dimension d of the catalyst and the rate of passive noise is equal to $[\log_2 d]/n$. We use the term *passive noise* to describe this kind of noise because one only needs to bring in a maximally mixed state as a resource: There is no need to apply local randomizing unitaries to create passive noise. It is also clear that active noise can create passive noise but not vice versa.

With these notions in mind, we can now define a conditional erasure protocol. Let $n \in \mathbb{N}$, $M, L \in \mathbb{N}$, and $\varepsilon \in [0,1]$. An (n, M, L, ε) conditional erasure protocol consists of a unitary quantum channel $\mathcal{U}_{A^nE^nA' \to A'_1A'_2\hat{E}^n}$ and an auxiliary catalyst state $\pi_{A'}$, which is maximally mixed. The state at the end of the protocol is $\omega_{A'_1B^n\hat{E}^n}$, as given in (29). The parameter M is equal to $|A'_2|^2$ as before. We require that a conditional erasure protocol satisfies the property of negligible disturbance, as specified in (25). We also require that the resulting state $\omega_{A'_1B^n\hat{E}^n}$ is such that the A'_1 system is decoupled from the BE systems, in the sense that

$$F\left(\omega_{A'_1B^n\hat{E}^n}, \pi_{A'_1} \otimes \omega_{B^n\hat{E}^n}\right) \geqslant 1 - \varepsilon, \tag{64}$$

where $\pi_{A'_1}$ is a maximally mixed state. We take the parameter

$$L = (|A'|/|A'_1|)^2, \tag{65}$$

or equivalently, $\log_2 L = 2[\log_2 |A'| - \log_2 |A'_1|]$. The parameter L thus quantifies the gain or consumption of passive noise in a conditional erasure protocol. If passive noise is gained in a conditional erasure protocol, then it can be used as a resource for a future erasure task.

We can see by inspecting (64) that conditional erasure achieves the task of state deconstruction, with the local recovery channel taken to be a preparation of the state $\pi_{A'_1}$ after the system A'_1 of $\omega_{A'_1B^n\hat{E}^n}$ is lost.

A. Conditional erasure is equivalent to state redistribution

In this section, we show that the task of conditional erasure is equivalent to state redistribution, in the sense that the existence of a conditional erasure protocol implies the existence of a state redistribution protocol and vice versa. We begin with the following implication:

Theorem 5. An (n, M, L, ε) protocol for state redistribution of a four-system pure state ψ_{ABER}, as specified in Sec. III C, realizes an $(n, M^2, L^2, 4\varepsilon)$ conditional erasure protocol of $\rho_{ABE} = \mathrm{Tr}_R\{\psi_{ABER}\}$, as specified in Sec. VII.

Proof. A proof of this theorem directly follows along the lines given in the proof of Theorem 3. Following the proof there, we arrive at (57), which is equivalent to the desired condition in (64). The parameter L for the state redistribution protocol is equal to $|A'|/|A_0|$, which becomes L^2 in the conditional erasure protocol per our convention in (65). ■

We now state the other implication:

Theorem 6. An (n, M, L, ε) protocol for conditional erasure of a four-system pure state ψ_{ABER}, as specified in Sec. VII, realizes an $(n, \lceil\sqrt{M}\rceil, \lceil\sqrt{L}\rceil, 4\varepsilon)$ state redistribution protocol of $\rho_{ABE} = \mathrm{Tr}_R\{\psi_{ABER}\}$, as specified in Sec. III C.

Proof. This follows simply by applying Uhlmann's theorem for fidelity [17] to a conditional erasure protocol in order to realize a decoder for state redistribution. To this end, suppose we are given a unitary quantum channel $\mathcal{U}_{A^nE^nA'\to A'_1A'_2\hat{E}^n}$ and an auxiliary catalyst state $\pi_{A'}$, as part of a conditional erasure protocol. Suppose further that they satisfy the negligible disturbance condition in (25) and the decoupled condition in (64). Combining these via the triangle inequality for $\sqrt{1-F(\rho,\sigma)}$ [similar to how we did previously in (55)], we find that the following condition holds:

$$F\left(\omega_{A'_1B^n\hat{E}^n}, \pi_{A'_1}\otimes\psi^{\otimes n}_{BE}\right) \geqslant 1-4\varepsilon. \tag{66}$$

A purification of the state $\omega_{A'_1B^n\hat{E}^n}$ is the following state:

$$\varsigma_{A'_1A'_2B^n\hat{E}^nR^nR'} \equiv \mathcal{U}_{A^nE^nA'\to A'_1A'_2\hat{E}^n}\left(\psi^{\otimes n}_{ABER}\otimes\Phi_{A'R'}\right). \tag{67}$$

That is, we obtain the state $\omega_{A'_1B^n\hat{E}^n}$ by tracing over the A'_2R^nR' systems of the above state. A purification of the state $\pi_{A'_1}\otimes\psi^{\otimes n}_{BE}$ is the following state:

$$\Phi_{A'_1R'_1}\otimes\psi^{\otimes n}_{ABER}. \tag{68}$$

Thus, Uhlmann's theorem for fidelity applied to (66) implies the existence of an isometric channel $\mathcal{V}_{A'_2R^nR'\to R'_1A^nR^n}$ such that

$$F\left(\mathcal{V}(\varsigma), \Phi_{A'_1R'_1}\otimes\psi^{\otimes n}_{ABER}\right) \geqslant 1-4\varepsilon, \tag{69}$$

where we have used the shorthand $\mathcal{V}(\varsigma)\equiv\mathcal{V}_{A'_2R^nR'\to R'_1A^nR^n}(\varsigma_{A'_1A'_2B^n\hat{E}^nR^nR'})$. Thus, the channel $\mathcal{V}_{A'_2R^nR'\to R'_1A^nR^n}$ can function as a decoder for a quantum state redistribution (QSR) protocol.

Summarizing, a purification $\Phi_{A'R'}$ of the catalyst state $\pi_{A'}$ functions as a maximally entangled resource in QSR, the unitary channel $\mathcal{U}_{A^nE^nA'\to A'_1A'_2\hat{E}^n}$ functions as an encoder in QSR, the system A'_2 is sent over a noiseless quantum channel in QSR, the isometric channel $\mathcal{V}_{A'_2R^nR'\to R'_1A^nR^n}$ functions as a decoder in QSR, and a purification $\Phi_{A'_1R'_1}$ of the state $\pi_{A'_1}$ functions as a maximally entangled resource shared between sender and receiver at the end of the QSR protocol. This completes the proof. ■

B. Optimal rate region for conditional erasure

We now define the achievable rate region for conditional erasure, which consists of achievable rate pairs (R_A, R_P), where R_A is equal to the rate of active noise and R_P is equal to the rate of passive noise. A rate pair (R_A, R_P) is *achievable* for conditional erasure of ψ_{ABER} if for all $\varepsilon\in(0,1)$, $\delta>0$, and sufficiently large n, there exists an $(n, 2^{n[R_A+\delta]}, 2^{n[R_P+\delta]}, \varepsilon)$ conditional erasure protocol. The achievable rate region of conditional erasure of ψ_{ABER} is equal to the union of all rate pairs which are achievable for conditional erasure of ψ_{ABER}.

Because of the equivalence between conditional erasure and state redistribution, given in the previous section, and the results about quantum state redistribution recalled in (21) and (22), we can immediately conclude the following theorem:

Theorem 7. The rate pair

$$(I(A;B|R)_\psi,\ I(A;E)_\psi - I(A;R)_\psi) \tag{70}$$

is achievable for conditional erasure of ψ_{ABER}, and the optimal rate region is equal to

$$R_A \geqslant I(A;B|R)_\psi, \tag{71}$$

$$R_A + R_P \geqslant 2H(A|R)_\psi. \tag{72}$$

Remark 5. The above theorem indicates that sometimes a catalyst is not actually needed to complete the conditional erasure task. In particular, if the inequality $I(A;E)_\psi \leqslant I(A;R)_\psi$ holds, then the protocol generates passive noise and hence only a vanishing, sublinear rate of passive noise is in fact needed to accomplish the conditional erasure task. Indeed, we could double block the protocol into N blocks, each consisting of n copies of ψ_{ABER}. For the first block of the protocol, we could supply $\approx nI(A;E)_\psi$ bits of passive noise and then the protocol would generate $\approx nI(A;R)_\psi$ bits of passive noise. Since the condition $I(A;E)_\psi \leqslant I(A;R)_\psi$ is assumed to hold, we could reinvest $\approx nI(A;E)_\psi$ bits of passive noise for the second block of the protocol while generating $\approx nI(A;R)_\psi$ bits of passive noise. For each block, we have an excess of $\approx n[I(A;R)_\psi - I(A;E)_\psi]$ bits of passive noise available. Repeating this procedure until the Nth block, we find that the rate of passive noise consumed is equal to $\approx nI(A;E)_\psi/nN$, since it was only consumed in the first block, and this rate vanishes in the limit as $n, N\to\infty$.

VIII. QUANTUM DISCORD AS EINSELECTION COST

Environment-induced superselection (abbreviated *einselection*) is a process in which an interaction between a system of interest and a large environment causes selective loss of information from the system [21]. The interaction with the environment has the effect of monitoring particular observables of the system, such that only eigenstates of these observables can persist in the system, being unaffected by the interaction. The quantum discord was originally proposed as a measure of the decrease of correlations after einselection is complete [18,19] and can be generalized to include arbitrary measurements (positive operator-valued measures, POVMs) rather than just measurements corresponding to system observables (see, e.g., Refs. [41,42] for reviews of discord and related measures).

To define the quantum discord, we begin with a bipartite state ρ_{AB} and a POVM $\Lambda\equiv\{\Lambda^x_A\}$, with $\Lambda^x_A\geqslant 0$ for all x and $\sum_x\Lambda^x_A = I_A$. The (unoptimized) quantum discord is a measure of the loss of correlation between A and B under the measurement Λ:

$$D(\overline{A};B)_{\rho,\Lambda} \equiv I(A;B)_\rho - I(X;B)_\zeta, \tag{73}$$

where

$$\zeta_{XB} \equiv \sum_x |x\rangle\langle x|_X \otimes \mathrm{Tr}_A\left\{\Lambda_A^x \rho_{AB}\right\}. \tag{74}$$

Here we continue with the main theme of this paper, namely, erasure of correlations, and define an operational task that we call an *einselection-simulation protocol*, which is a simulation of the einselection process via local randomizing unitaries. The starting point for such a protocol is a bipartite state ρ_{AB} and a POVM $\Lambda \equiv \{\Lambda_A^x\}$, and the objective is to determine the minimum rate of noise needed to apply to the A system of ρ_{AB}, such that the resulting state σ_{AB} is approximately einselected. By this, we mean that

(1) there is a measurement corresponding to σ_{AB}, such the state σ_{AB} is locally recoverable after performing this measurement on system A of σ_{AB}, and

(2) the corresponding postmeasurement state is indistinguishable from the postmeasurement state in (74).

By Ref. [16, Proposition 21], the state σ_{AB} having negligible discord is equivalent to the condition of local recoverability of σ_{AB} after a measurement is performed on system A.

More formally, for $n, M \in \mathbb{N}$, and $\varepsilon \in [0, 1]$, we define an (n, M, ε) einselection-simulation protocol for a state ρ_{AB} and a POVM $\Lambda_A \equiv \{\Lambda_A^x\}$ to consist of an ensemble $\{p_i, U^i_{A^nA'}\}_{i=1}^M$ of einselection-simulating unitaries, a catalyst state $\theta_{A'}$, and a measurement channel $\mathcal{M}_{A^nA'\to X^n}$ such that the state $\sigma_{A^nA'B^n}$ resulting from local unitary randomization

$$\sigma_{A^nA'B^n} \equiv \sum_{i=1}^M p_i U^i_{A^nA'}\left(\rho_{AB}^{\otimes n} \otimes \theta_{A'}\right)\left(U^i_{A^nA'}\right)^\dagger \tag{75}$$

and the measurement channel $\mathcal{M}_{A^nA'\to X^n}$ satisfy the following two requirements:

(1) The state $\sigma_{A^nA'B^n}$ is locally recoverable from the classical system X^n after the measurement channel $\mathcal{M}_{A^nA'\to X^n}$ is applied, in the sense that there exists a preparation channel $\mathcal{P}_{X^n\to A^nA'}$ such that

$$F(\sigma_{A^nA'B^n}, (\mathcal{P}\circ\mathcal{M})(\sigma_{A^nA'B^n})) \geqslant 1-\varepsilon, \tag{76}$$

where $\mathcal{P} \equiv \mathcal{P}_{X^n\to A^nA'}$ and $\mathcal{M} \equiv \mathcal{M}_{A^nA'\to X^n}$. In this sense, we say that $\sigma_{A^nA'B^n}$ has been approximately einselected. In Ref. [16], this was described as the state $\sigma_{A^nA'B^n}$ being negligibly disturbed by the action of an entanglement-breaking channel.

(2) The postmeasurement state $\mathcal{M}_{A^nA'\to X^n}$ is indistinguishable from many copies of the postmeasurement state in (74), in the sense that

$$F\left(\mathcal{M}_{A^nA'\to X^n}(\sigma_{A^nA'B^n}), \zeta_{XB}^{\otimes n}\right) \geqslant 1-\varepsilon. \tag{77}$$

This latter condition ensures that the einselection-simulating unitaries perform a *faithful* simulation of the einselection process: They do not destroy the correlations remaining between X and B after the measurement Λ_A occurs (i.e., they only destroy the correlations in ρ_{AB} lost in the application of the measurement Λ_A).

Definition 5 (Achievable rate). A rate R of einselection simulation for a state ρ_{AB} and a POVM Λ_A is achievable if for all $\varepsilon \in (0, 1)$, $\delta > 0$, and sufficiently large n, there exists an $(n, 2^{n[R+\delta]}, \varepsilon)$ einselection-simulation protocol.

Definition 6 (Einselection cost). The einselection cost $\mathcal{E}(\rho_{AB}, \Lambda_A)$ of a state ρ_{AB} and a POVM Λ_A is equal to the infimum of all achievable rates for einselection simulation of ρ_{AB} and Λ_A.

Our main result in this section is the following physical meaning for the quantum discord:

Theorem 8. The einselection cost $\mathcal{E}(\rho_{AB}, \Lambda_A)$ of a state ρ_{AB} and a POVM Λ_A is equal to its quantum discord $D(\overline{A}; B)_{\rho,\Lambda}$:

$$\mathcal{E}(\rho_{AB}, \Lambda_A) = D(\overline{A}; B)_{\rho,\Lambda}, \tag{78}$$

where $D(\overline{A}; B)_{\rho,\Lambda}$ is defined in (73).

Proof. A proof of the above theorem requires two parts: the achievability part and the converse. We begin with the converse and note that it bears some similarities to a converse given in Appendix A and the proof of [16, Proposition 21]. Consider an arbitrary (n, M, ε) einselection simulation protocol for ρ_{AB} and Λ_A, which consists of $\{p_i, U^i_{A^nA'}\}_{i=1}^M$, $\theta_{A'}$, $\mathcal{M}_{A^nA'\to X^n}$, and $\mathcal{P}_{X^n\to A^nA'}$ as defined above. Let $\sigma_{\hat{M}A^nA'B^n}$ denote the following state,

$$\sigma_{\hat{M}A^nA'B^n} \equiv \sum_{i=1}^M p_i |i\rangle\langle i|_{\hat{M}} \otimes U^i_{A^nA'}\left(\rho_{AB}^{\otimes n} \otimes \theta_{A'}\right)\left(U^i_{A^nA'}\right)^\dagger, \tag{79}$$

and let $\kappa_{\hat{M}X^nB^n}$ denote the following state after the measurement channel $\mathcal{M}_{A^nA'\to X^n}$ acts

$$\kappa_{\hat{M}X^nB^n} \equiv \mathcal{M}_{A^nA'\to X^n}(\sigma_{\hat{M}A^nA'B^n}), \tag{80}$$

For such a protocol, the following chain of inequalities holds:

$$\begin{aligned} nD(\overline{A}; B)_{\rho,\Lambda} &= n[H(B|X)_\zeta - H(B|A)_\rho] && (81)\\ &= H(B^n|X^n)_{\zeta^{\otimes n}} - H(B^n|A^n)_{\rho^{\otimes n}} && (82)\\ &\leqslant H(B^n|X^n)_\kappa + f(n,\varepsilon) - H(B^n|A^n)_{\rho^{\otimes n}}. && (83)\end{aligned}$$

The first equality follows from a simple manipulation of the definition in (73), noting that $H(B)_\rho = H(B)_\zeta$. The second equality follows from additivity of the conditional entropies with respect to tensor-product states. The inequality follows from (77) (faithfulness of the einselection simulation), the Fuchs–van de Graaf inequalities in (4), and Ref. [27, Lemma 2], with

$$f(n,\varepsilon) \equiv 2n\sqrt{\varepsilon}\log_2|B| + (1+\sqrt{\varepsilon})h_2(\sqrt{\varepsilon}/[1+\sqrt{\varepsilon}]). \tag{84}$$

We now focus on bounding the two entropic terms $H(B^n|X^n)_\kappa$ and $-H(B^n|A^n)_{\rho^{\otimes n}}$ separately. Consider that

$$\begin{aligned} H(B^n|X^n)_\kappa &\leqslant H(B^n|A^nA')_{\mathcal{P}(\kappa)} && (85)\\ &\leqslant H(B^n|A^nA')_\sigma + f(n,\varepsilon). && (86)\end{aligned}$$

The first inequality follows because the conditional entropy does not decrease under the action of a channel on the conditioning system, in this case the channel being the preparation channel $\mathcal{P}_{X^n\to A^nA'}$. The second inequality follows from the local recoverability condition in (76), the Fuchs–van de Graaf inequalities in (4) and Ref. [27, Lemma 2]. We now bound the

term $-H(B^n|A^n)_{\rho^{\otimes n}}$ from above:

$$-H(B^n|A^n)_{\rho^{\otimes n}} = -H(B^n|\hat{M}A^nA')_{\sigma_{\hat{M}}\otimes\rho^{\otimes n}\otimes\theta} \quad (87)$$

$$= -H(B^n|\hat{M}A^nA')_{\sigma} \quad (88)$$

$$\leqslant -H(B^n|A^nA')_{\sigma} + \log_2|\hat{M}|. \quad (89)$$

The first equality follows because the conditional entropy is invariant with respect to tensoring in the product states $\sigma_{\hat{M}}\otimes\theta_{A'}$ to be part of the conditioning system, with $\sigma_{\hat{M}} = \sum_{i=1}^{M} p_i|i\rangle\langle i|_{\hat{M}}$. The second equality follows because the conditional entropy is invariant with respect to the following controlled unitary acting on the systems $\hat{M}A^nA'$ of $\sigma_{\hat{M}}\otimes\rho_{AB}^{\otimes n}\otimes\theta_{A'}$:

$$\sum_{i=1}^{M}|i\rangle\langle i|_{\hat{M}}\otimes U^i_{A^nA'}. \quad (90)$$

The inequality follows from a rewriting and a dimension bound for CQMI [4, Exercise 11.7.9] when one of the conditioned systems is classical (in this case system $\hat{M}$):

$$H(B^n|A^nA')_{\sigma} - H(B^n|\hat{M}A^nA')_{\sigma}$$
$$= I(B^n;\hat{M}|A^nA')_{\sigma} \leqslant \log_2|\hat{M}|. \quad (91)$$

Putting everything together, we find the following lower bound on the rate of an arbitrary (n, M, ε) einselection simulation protocol:

$$D(\overline{A};B)_{\rho,\Lambda} \leqslant \tfrac{1}{n}\log_2|\hat{M}| + \tfrac{1}{n}[f(n,\varepsilon)+g(n,\varepsilon)]. \quad (92)$$

Taking the limit as $n\to\infty$ and then as $\varepsilon\to 0$, we can conclude that the quantum discord is a lower bound on the einselection cost:

$$D(\overline{A};B)_{\rho,\Lambda} \leqslant \mathcal{E}(\rho_{AB},\Lambda_A). \quad (93)$$

We now turn to the achievability part, which makes use of a state deconstruction protocol. Let $V_{AE_0\to XE}$ denote a unitary extension of a measurement channel corresponding to the POVM Λ_A. In particular, we can define $V_{AE_0\to XE}$ as follows by its action on a state vector $|\psi\rangle_A$:

$$V_{AE_0\to XE}|\psi\rangle_A|0\rangle_{E_0} \equiv \sum_x \left(\sqrt{\Lambda_A^x}|\psi\rangle_A\right)_{\tilde{E}}|x\rangle_X|x\rangle_{\tilde{E}}, \quad (94)$$

where we set $E\equiv\tilde{E}\tilde{E}$ and $\{|x\rangle\}_x$ is an orthonormal basis. We define the isometric channel $\mathcal{V}_{A\to XE}(\psi_A)\equiv V(\psi_A\otimes|0\rangle\langle 0|_{E_0})V^\dagger$ and note that tracing over system E gives back the original measurement channel:

$$\mathrm{Tr}_E\{\mathcal{V}_{A\to XE}(\psi_A)\} = \sum_x \mathrm{Tr}\left\{\Lambda_A^x\psi_A\right\}|x\rangle\langle x|_X. \quad (95)$$

We now show how an (n, M, ε) state deconstruction protocol for the state $\rho_{XEB}\equiv\mathcal{V}_{A\to XE}(\rho_{AB})$ leads to an $(n, M, 9\varepsilon)$ einselection-simulation protocol for ρ_{AB} and Λ_A. We consider a state deconstruction protocol in the Landauer-Bennett erasure model. To this end, let $\mathcal{U}_{E^nX^nE'\to E_1'E_2'X^n}$ be a unitary channel, and let $\theta_{E'}$ denote an ancilla state. Let $\omega_{E_1'E_2'X^nB^n}$ denote the following state resulting from a deconstruction operation:

$$\omega_{E_1'X^nB^n}\equiv\mathrm{Tr}_{E_2'}\left\{\mathcal{U}_{E^nX^nE'\to E_1'E_2'X^n}\left(\rho_{XEB}^{\otimes n}\otimes\theta_{E'}\right)\right\}. \quad (96)$$

The following two properties, discussed in Sec. IV B, hold for a state deconstruction protocol:

(1) There exists a recovery channel $\mathcal{R}_{X^n\to X^nE_1'}$ such that system E_1' is locally recoverable from X^n:

$$F(\omega_{E_1'X^nB^n},\mathcal{R}_{X^n\to X^nE_1'}(\omega_{X^nB^n}))\geqslant 1-\varepsilon. \quad (97)$$

(2) The deconstruction protocol causes neliglible disturbance to the marginal state on systems X^nB^n:

$$F\left(\omega_{X^nB^n},\rho_{XB}^{\otimes n}\right)\geqslant 1-\varepsilon. \quad (98)$$

We now specify the components of the einselection-simulation protocol. It consists of the following ensemble of unitaries:

$$\left\{1/M, V^{\dagger\otimes n}U^\dagger W_{E_2'}^i U V^{\otimes n}\right\}_{i=1}^M, \quad (99)$$

where U is the unitary operator corresponding to the unitary channel $\mathcal{U}_{E^nX^nE'\to E_1'E_2'X^n}$ and $\{W_{E_2'}^i\}_{i=1}^M$ is a Heisenberg-Weyl set of unitaries for system E_2'. The ancilla state for einselection simulation is $\theta_{E'}\otimes|0\rangle\langle 0|_{E_0}$, such that the resulting approximately einselected state $\sigma_{A^nA'B^n}$ is as follows,

$$\sigma_{A^nA'B^n}\equiv V^{\dagger\otimes n}U^\dagger\left(\omega_{E_1'X^nB^n}\otimes\pi_{E_2'}\right)UV^{\otimes n}, \quad (100)$$

where we are setting system $A'\equiv E'E_0$, since the systems $E'E_0$ will now serve as the ancilla system A' for an einselection-simulation protocol. We define the measurement channel $\mathcal{M}_{A^nA'\to X^n}$ as follows:

$$\mathcal{M}_{A^nA'\to X^n}(\tau_{A^nA'})\equiv\overline{\Delta}_{X^n}\circ$$
$$\mathrm{Tr}_{E_1'E_2'}\left\{\mathcal{U}_{E^nX^nE'\to E_1'E_2'X^n}\left(\mathcal{V}_{A\to XE}^{\otimes n}(\tau_{A^nA'})\right)\right\}, \quad (101)$$

where $\overline{\Delta}_{X^n}$ denotes a completely dephasing channel, defined as

$$\overline{\Delta}_{X^n}(\xi_{X^n})\equiv\sum_{x^n}|x^n\rangle\langle x^n|_{X^n}\xi_{X^n}|x^n\rangle\langle x^n|_{X^n}, \quad (102)$$

$$|x^n\rangle\equiv|x_1\rangle_{X_1}\otimes\cdots\otimes|x_n\rangle_{X_n}. \quad (103)$$

We take the preparation channel $\mathcal{P}_{X^n\to A^nA'}$ to be

$$\mathcal{P}_{X^n\to A^nA'}(\xi_{X^n})\equiv V^{\dagger\otimes n}U^\dagger\left(\mathcal{R}_{X^n\to X^nE_1'}(\xi_{X^n})\otimes\pi_{E_2'}\right)UV^{\otimes n}, \quad (104)$$

which consists of applying the recovery channel $\mathcal{R}_{X^n\to X^nE_1'}$, appending the maximally mixed state $\pi_{E_2'}$, inverting the deconstruction unitary U, and inverting the unitary dilation $V^{\otimes n}$ of the measurement channel for $\Lambda_A^{\otimes n}$.

We now demonstrate that the two conditions for einselection simulation hold. We begin by establishing the faithfulness condition in (77). From definitions, we have that

$$\mathcal{M}_{A^nA'\to X^n}(\sigma_{AA'B^n})=\overline{\Delta}_{X^n}(\omega_{X^nB^n}). \quad (105)$$

Furthermore, the negligible disturbance condition in (98) and the monotonicity of fidelity with respect to quantum channels imply that

$$F\left(\overline{\Delta}_{X^n}(\omega_{X^nB^n}),\overline{\Delta}_{X^n}\left(\rho_{XB}^{\otimes n}\right)\right)\geqslant 1-\varepsilon. \quad (106)$$

But this is equivalent to

$$F\left(\mathcal{M}_{A^nA'\to X^n}(\sigma_{AA'B^n}),\rho_{XB}^{\otimes n}\right)\geqslant 1-\varepsilon, \quad (107)$$

by applying (105) and the fact that the classical-quantum state $\rho_{XB}^{\otimes n}$ is invariant under the action of the dephasing channel $\overline{\Delta}_{X^n}$, so this establishes the faithfulness condition in (77).

We now establish the local recoverability condition in (76). Consider that (107), (98), the triangle inequality for the metric $\sqrt{1-F}$, and a rewriting imply that

$$F(\mathcal{M}_{A^nA'\to X^n}(\sigma_{AA'B^n}), \omega_{X^nB^n}) \geq 1-4\varepsilon. \tag{108}$$

The monotonicity of fidelity with respect to quantum channels applied to (108) then implies that

$$F((\mathcal{P}\circ\mathcal{M})(\sigma_{AA'B^n}), \mathcal{P}_{X^n\to A^nA'}(\omega_{X^nB^n})) \geq 1-4\varepsilon. \tag{109}$$

Invariance of the fidelity in (97) with respect to tensoring in $\pi_{E_2'}$, applying the unitary $U^\dagger$ followed by $V^{\dagger\otimes n}$, and applying definitions implies that

$$F(\sigma_{A^nA'B^n}, \mathcal{P}_{X^n\to A^nA'}(\omega_{X^nB^n})) \geq 1-\varepsilon. \tag{110}$$

We can then apply the triangle inequality to (109) and (110) with respect to the metric $\sqrt{1-F}$ and rewrite to find that

$$F(\sigma_{A^nA'B^n}, (\mathcal{P}\circ\mathcal{M})(\sigma_{AA'B^n})) \geq 1-9\varepsilon. \tag{111}$$

This then establishes the local recoverability condition in (76). Thus, we have demonstrated that an (n, M, ε) state deconstruction protocol leads to an $(n, M, 9\varepsilon)$ einselection-simulation protocol.

What remains is to show that the discord is an achievable rate for einselection simulation. In our protocol for state deconstruction (the particular setup considered here), an achievable rate is

$$\frac{1}{n}\log_2|M| \approx I(E;B|X)_{\mathcal{V}(\rho)}, \tag{112}$$

which implies via the simulation argument given above that $I(E;B|X)_{\mathcal{V}(\rho)}$ is an achievable rate for einselection simulation. It is known from Refs. [43,44] that

$$D(\overline{A};B)_{\rho,\Lambda} = I(E;B|X)_{\mathcal{V}(\rho)}, \tag{113}$$

where $\mathcal{V}(\rho) = \mathcal{V}_{A\to XE}(\rho_{AB})$, and so we establish the inequality

$$D(\overline{A};B)_{\rho,\Lambda} \geq \mathcal{E}(\rho_{AB}, \Lambda_A), \tag{114}$$

completing the proof when combined with (93). ■

Remark 6. The operational interpretation for quantum discord given here builds upon the previous interpretation from Ref. [45, Sec. 6(c)], given in terms of quantum state redistribution (see Refs. [46,47] for other operational, information-theoretic interpretations of discord). In Ref. [45, Sec. 6(c)], it was established via the relation in (113) that the discord is equal to twice the rate of quantum communication needed in a state redistribution protocol to transmit the environment system E of $\mathcal{V}_{A\to XE}(\rho_{AB})$ to an inaccessible environmental system R, which purifies the state ρ_{AB}. The interpretation written there is that discord "characterizes the amount of quantum information lost in the measurement process." On the one hand, we now see that the einselection-simulation protocol discussed above perhaps gives a more natural operational interpretation of quantum discord, in the original spirit of the discussions from Refs. [18,19]. On the other hand, we see that at the core of the achievability proof above is the state redistribution protocol and the method from Ref. [45, Sec. 6(c)], given that we showed in Sec. VI how state redistribution can simulate state deconstruction.

IX. SQUASHED ENTANGLEMENT

Our main result in Theorem 4 also provides an operational interpretation of the squashed entanglement [20], which is an entanglement measure satisfying many desirable properties (see Refs. [48] and references therein). A communication-theoretic interpretation for squashed entanglement was given in Ref. [49], and our interpretation here largely follows the interpretation of Ref. [49]. There, it was argued that squashed entanglement of ρ_{AB} is equal to the fastest rate at which Alice could send her systems to a third party possessing the best possible quantum side information to help in decoding.

Recall that the squashed entanglement of a bipartite state ρ_{AB} is defined as

$$E_{\mathrm{sq}}(A;B)_\rho \equiv \tfrac{1}{2}\inf_{\zeta_{ABE}}\{I(A;B|E)_\zeta : \rho_{AB} = \mathrm{Tr}_E\{\zeta_{ABE}\}\}.$$

Because of its connection with CQMI, we thus see that the squashed entanglement is equal to half the minimum rate of noise needed in a deconstruction operation if Alice has available the best possible third correlated system E to help in the deconstruction task. That is, suppose that the state that Alice and Bob begin with is ρ_{AB}. If there is no third system available, then the deconstruction task reduces to decorrelating and the optimal rate of noise for deconstructing is equal to the mutual information $I(A;B)_\rho$. However, if Alice is provided with a third system E, such that the global state is ζ_{ABE} with $\rho_{AB} = \mathrm{Tr}_E\{\zeta_{ABE}\}$, then the rate of noise needed to achieve deconstruction is equal to $I(A;B|E)$ and could potentially be reduced, such that fewer local randomizing unitaries are needed in a deconstruction operation. By inspecting the formula for squashed entanglement, we see that $E_{\mathrm{sq}}(A;B)_\rho$ is equal to half the minimum rate of noise needed in a deconstruction operation if optimal quantum side information in E is available. Also, loosely speaking, we see that the more entangled a state is (as measured by E_{sq}), the more difficult it is to deconstruct it with respect to any possible third system E.

Applying the insights of Ref. [48], we see that squashed entanglement is equal to half the minimum rate of noise needed to produce a state on Alice, Bob, and Eve's systems, such that Alice's system of the resulting state is locally recoverable from Eve's system. By Ref. [48], the resulting state is thus highly extendible and furthermore arbitrarily close to a separable state in 1-LOCC (one-way local operations and classical communication) distance in the many-copy limit. In more detail, let $\omega_{A'B^nE^n}$ denote the state resulting from applying a state deconstruction protocol to $\rho_{ABE}^{\otimes n}$, where ρ_{ABE} is an extension of ρ_{AB}. Using the argument from Ref. [48] (repeated in Ref. [16]), along with the fact that $\sqrt{1-F(\rho,\sigma)}$ is a distance measure [23] and thus obeys the triangle inequality, we find that $F(A';B^n|E^n)_\omega \geq 1-\varepsilon$ implies that

$$\sup_{\gamma_{A'B^n}\in\mathcal{E}_k(A':B^n)} F(\omega_{A'B^n}, \gamma_{A'B^n}) \geq 1-k^2\varepsilon, \tag{115}$$

where $\mathcal{E}_k(A':B^n)$ denotes the set of k-extendible states, defined as the set of all states $\gamma_{A'B^n}$ such that there exists a k extension $\gamma_{A_1'\cdots A_k'B^n}$, with $\gamma_{A_1'\cdots A_k'B^n}$ invariant with respect

to permutations of the systems $A'_1 \dots A'_k$ and $\gamma_{A'B^n} = \mathrm{Tr}_{A'_2 \cdots A'_k B^n}\{\gamma_{A'_1 \cdots A'_k B^n}\}$ [50,51]. Since we can take ε to be an exponentially decreasing function of n [15], we can take k growing to infinity, say, proportional to n^2, such that $\sup_{\gamma_{A'B^n} \in \mathcal{E}_k(A':B^n)} F(\omega_{A'B^n}, \gamma_{A'B^n}) \to 1$ as $k, n \to \infty$. Thus, the squashed entanglement can be interpreted in terms of k extendibility as stated above.

To get the statement about 1-LOCC distance to separable states, we need only apply a result from Ref. [52], which states that the 1-LOCC distance between k-extendible states and separable states can be bounded from above by a term $\propto \sqrt{(\log_2 |A'|)/k}$. In our case, $\log_2 |A'|$ is linear in n, and with $k \propto n^2$, the 1-LOCC distance between k-extendible states and separable states vanishes in the large n limit.

X. DISCUSSION

We have provided an operational interpretation of the conditional quantum mutual information (CQMI) $I(A;B|E)_\rho$ of a tripartite state ρ_{ABE} as the minimal rate of noise needed to apply in a deconstruction operation, such that it has negligible disturbance of the marginal state ρ_{BE} while producing a state that is locally recoverable from system E alone. Equivalently, we find that CQMI is equal to the minimal rate of noise needed to result in a state that has vanishing normalized CQMI. The method for showing achievability of CQMI in such a state deconstruction task relies upon the quantum state redistribution protocol [14,15]. We showed how the state deconstruction protocol simplifies significantly if the system E is classical. We also considered the task of conditional erasure, in which the goal is to apply a noisy operation to the AE systems such that the BE systems are negligibly disturbed and the resulting A system is decoupled from the BE systems. We find again that the minimal rate of noise for conditional erasure is equal to the CQMI $I(A;B|E)$. We also provided operational interpretations of quantum correlation measures which have CQMI at their core, including quantum discord [18,19] and squashed entanglement [20]. We should also mention that our operational interpretation of CQMI seems natural in the context of the recent contribution of Ref. [53], which discussed scrambling of information due to bipartite unitary interactions.

Going forward from here, we suspect that it should be possible to generalize our results to multipartite CQMI quantities [54,55]. We also think there are major obstacles to be overcome before we can determine a satisfying one-shot generalization of these results, just as there are obstacles in doing so for quantum state redistribution [29,30,56]. We would also like to know whether the CQMI is generally achievable for the task of state deconstruction if no catalyst is available. Remark 5 discusses how a catalyst is sometimes not actually needed for state deconstruction or conditional erasure, but we would like to know whether this might generally be the case.

ACKNOWLEDGMENTS

We are indebted to Siddhartha Das, Gilad Gour, Matt Hastings, Mio Murao, Marco Piani, Kaushik Seshadreesan, Eyuri Wakakuwa, and Andreas Winter for valuable discussions. We also acknowledge the catalyzing role of the open problems session at Beyond IID 2016, which ultimately led to the solution of the i.i.d result presented here. M.B. acknowledges funding by the SNSF through a fellowship, funding by the Institute for Quantum Information and Matter (IQIM), an NSF Physics Frontiers Center (NFS Grant No. PHY-1125565) with support of the Gordon and Betty Moore Foundation (GBMF-12500028), and funding support from the ARO grant for Research on Quantum Algorithms at the IQIM (No. W911NF-12-1-0521). C.M. acknowledges financial support from the European Research Council (ERC Grant Agreement No. 337603), the Danish Council for Independent Research (Sapere Aude) and VILLUM FONDEN via the QMATH Centre of Excellence (Grant No. 10059). M.M.W. acknowledges support from the NSF under Award No. CCF-1350397.

APPENDIX A: ALTERNATIVE CONVERSE PROOFS

In this appendix, we detail two alternative converse proofs for Theorem 2, which are tailored to the local unitary randomizing model. One of the proofs bears similarities to the current proof of Theorem 2 and the other is similar to those appearing in prior work [3,33,34]. We begin with the former proof.

Let $\{p_i, U^i_{A^n A' E^n}\}_{i=1}^M$ denote any ensemble of unitaries and $\theta_{A'}$ a corresponding ancilla state realizing an (n, M, ε) state deconstruction protocol, such that the deconstruction operation $\mathcal{N}_{A^n A' E^n}$ is as given in (23) and satisfies the conditions of negligible disturbance in (25) and local recoverability in (26).

Let $\sigma_{\hat{M} A^n B^n E^n A'}$ denote the following state:

$$\sigma_{\hat{M} A^n B^n E^n A'} \equiv \sum_{i=1}^M p_i |i\rangle\langle i|_{\hat{M}} \otimes U^i_{A^n A' E^n} \left(\rho^{\otimes n}_{ABE} \otimes \theta_{A'}\right) \left(U^i_{A^n A' E^n}\right)^\dagger. \quad \text{(A1)}$$

Then consider that

$$\begin{aligned} n I(A;B|E)_\rho &= I(A^n;B^n|E^n)_{\rho^{\otimes n}} && \text{(A2)} \\ &= H(B^n|E^n)_{\rho^{\otimes n}} - H(B^n|A^n E^n)_{\rho^{\otimes n}} && \text{(A3)} \\ &\leqslant H(B^n|E^n)_\sigma + f(n,\varepsilon) - H(B^n|A^n E^n)_{\rho^{\otimes n}}. && \text{(A4)} \end{aligned}$$

The inequality follows for a similar reason as given for the first inequality in (45), with $f(n,\varepsilon)$ chosen as in (47). We now focus on bounding the term $-H(B^n|A^n E^n)_{\rho^{\otimes n}}$:

$$\begin{aligned} -H(B^n|A^n E^n)_{\rho^{\otimes n}} &= -H(B^n|\hat{M} A^n E^n A')_{\sigma_{\hat{M}} \otimes \rho^{\otimes n} \otimes \theta} && \text{(A5)} \\ &= -H(B^n|\hat{M} A^n E^n A')_\sigma && \text{(A6)} \\ &\leqslant -H(B^n|A^n E^n A')_\sigma + \log_2 |\hat{M}|. && \text{(A7)} \end{aligned}$$

The first equality follows because we are tensoring in the product states $\sigma_{\hat{M}} = \sum_{i=1}^M p_i |i\rangle\langle i|_{\hat{M}}$ and $\theta_{A'}$ for the conditioning system of the conditional entropy, which leave it invariant. The second equality follows because the conditional entropy is invariant under the application of the following controlled unitary to the systems $\hat{M} A^n E^n A'$ of $\sigma_{\hat{M}} \otimes \rho^{\otimes n}_{ABE} \otimes \theta_{A'}$:

$$\sum_i |i\rangle\langle i|_{\hat{M}} \otimes U^i_{A^n A' E^n}. \quad \text{(A8)}$$

The inequality follows from a rewriting and a dimension bound for CQMI [4, Exercise 11.7.9]:

$$\begin{aligned} & H(B^n|A^nE^nA')_\sigma - H(B^n|\hat{M}A^nE^nA')_\sigma \\ & = I(B^n;\hat{M}|A^nE^nA')_\sigma \leqslant \log_2|\hat{M}|. \end{aligned} \tag{A9}$$

Combining these inequalities, we find that

$$\begin{aligned} nI(A;B|E)_\rho &\leqslant H(B^n|E^n)_\sigma - H(B^n|A^nE^nA')_\sigma \\ &\quad + f(n,\varepsilon) + \log_2|\hat{M}| && \text{(A10)} \\ &= I(A^nA';B^n|E^n)_\sigma + f(n,\varepsilon) + \log_2|\hat{M}| && \text{(A11)} \\ &\leqslant g(n,\varepsilon) + f(n,\varepsilon) + \log_2|\hat{M}|. && \text{(A12)} \end{aligned}$$

The inequality follows by applying the local recoverability condition in (26) and because locally recoverable states have small CQMI as reviewed in (11), where we choose $g(n,\varepsilon)$ as in (49). We can then rewrite this as

$$I(A;B|E)_\rho \leqslant \frac{1}{n}\log_2|\hat{M}| + \frac{1}{n}[g(n,\varepsilon) + f(n,\varepsilon)]. \tag{A13}$$

By taking the limit as $n\to\infty$, then $\varepsilon\to 0$, and applying definitions, we can conclude the inequality $I(A;B|E)_\rho \leqslant \mathcal{D}(A;B|E)_\rho$.

We now detail the other proof, which is similar to those given in Refs. [3,33,34]. We define the following pure state:

$$\begin{aligned} & |\varphi\rangle_{M_1M_2A^nA'E^nB^nR^nR'} \\ & \equiv \sum_i \sqrt{p_i}|i\rangle_{M_1}|i\rangle_{M_2}U^i_{A^nA'E^n}|\psi\rangle^{\otimes n}_{ABER}\otimes|\theta\rangle_{A'R'}, \end{aligned} \tag{A14}$$

where $|\psi\rangle_{ABER}$ purifies ρ_{ABE} and $|\theta\rangle_{A'R'}$ purifies the ancilla $\theta_{A'}$. The state $|\varphi\rangle$ above is a purification of $\omega_{A^nA'B^nE^n}$ in (24) and is helpful in our analysis. Consider that

$$\begin{aligned} \log_2 M &\geqslant H(\{p_i\}) = H(M_1)_\varphi && \text{(A15)} \\ &= H(M_2A^nA'E^nB^nR^nR')_\varphi && \text{(A16)} \\ &\geqslant H(A^nA'E^nB^nR^nR')_\varphi && \text{(A17)} \\ &\geqslant H(A^nA'E^nB^n)_\varphi - H(R^nR')_\varphi && \text{(A18)} \\ &= H(A^nA'E^nB^n)_\omega - H(R^nR')_{\psi^{\otimes n}\otimes\theta} && \text{(A19)} \\ &= H(A^nA'E^nB^n)_\omega - H(A^nB^nE^nA')_{\psi^{\otimes n}\otimes\theta} && \text{(A20)} \\ &= H(A^nA'E^nB^n)_\omega - nH(ABE)_\psi - H(A')_\theta. && \text{(A21)} \end{aligned}$$

The first inequality follows because the logarithm of the cardinality of the probability distribution $\{p_i\}$ is an upper bound on its entropy $H(\{p_i\})$. The first equality follows because the reduced state of φ on system M_1 is classical with probability distribution $\{p_i\}$. The second equality follows because the entropies of the marginals of a bipartite pure state are equal (the bipartite cut here being between system M_1 and systems $M_2A^nA'E^nB^nR^nR'$). The second inequality follows the entropy cannot decrease when adding a classical system (in this case, the M_2 system of the reduced state on systems $M_2A^nA'E^nB^nR^nR'$ is classical, being decohered after a partial trace over system M_1). The third inequality is a consequence of the Araki-Lieb triangle inequality [57], which states that $H(KL)_\tau \geqslant H(K)_\tau - H(L)_\tau$ for a bipartite state τ_{KL}. The third equality follows because $\varphi_{A^nA'E^nB^n} = \omega_{A^nA'E^nB^n}$ and $\varphi_{R^nR'} = \psi_R^{\otimes n}\otimes\theta_{R'}$. The fourth equality follows because the state $\psi^{\otimes n}_{ABER}\otimes\theta_{A'R'}$ is pure, so that $H(R^nR')_{\psi^{\otimes n}\otimes\theta} = H(A^nB^nE^nA')_{\psi^{\otimes n}\otimes\theta}$. The last equality follows because entropy is additive with respect to tensor-product states. Focusing on the term $H(A^nA'E^nB^n)_\varphi$, we continue with

$$\begin{aligned} & H(A^nA'E^nB^n)_\omega \\ & \geqslant H(A^nA'E^n)_\omega + H(B^n|E^n)_\omega - g(n,\varepsilon) && \text{(A22)} \\ & \geqslant H(A^nA'E^n)_\omega + H(B^n|E^n)_{\psi^{\otimes n}} - g(n,\varepsilon) - f(n,\varepsilon) && \text{(A23)} \\ & = H(A^nA'\hat{E}^n)_\omega + nH(B|E)_\psi - g(n,\varepsilon) - f(n,\varepsilon). && \text{(A24)} \end{aligned}$$

The first inequality follows by applying the local recoverability condition in (26), the Fuchs–van de Graaf inequalities in (4), and because locally recoverable states have small CQMI as reviewed in (11). In particular, we can take

$$g(n,\varepsilon) \equiv 2n\sqrt{\varepsilon}\log|B| + (1+\sqrt{\varepsilon})h_2(\sqrt{\varepsilon}/[1+\sqrt{\varepsilon}]), \tag{A25}$$

and find that

$$I(A^nA';B^n|E^n)_\omega \leqslant g(n,\varepsilon), \tag{A26}$$

which when rewritten is equivalent to the first inequality. The second inequality follows from the negligible disturbance condition from (25), the Fuchs–van de Graaf inequalities in (4), and the continuity of conditional quantum entropy [27,40], with

$$f(n,\varepsilon) = 2\sqrt{\varepsilon}n\log|B| + (1+\sqrt{\varepsilon})h_2(\sqrt{\varepsilon}/[1+\sqrt{\varepsilon}]). \tag{A27}$$

The equality holds because entropy is additive with respect to tensor-product states. Now focusing on the term $H(A^nA'E^n)_\omega$, we continue with

$$\begin{aligned} H(A^nA'E^n)_\omega &\geqslant \sum_i p_i H(A^nA'E^n)_{U^i(\psi^{\otimes n}\otimes\theta)U^{i\dagger}} && \text{(A28)} \\ &= \sum_i p_i H(A^nE^nA')_{\psi^{\otimes n}\otimes\theta} && \text{(A29)} \\ &= nH(AE)_\psi + H(A')_\theta. && \text{(A30)} \end{aligned}$$

The first inequality follows from the concavity of quantum entropy. The first equality follows from unitary invariance of entropy, and the last again from additivity of entropy for tensor-product states. Putting everything together, we find that the following bound holds for any (n, M, ε) state deconstruction protocol:

$$\frac{1}{n}\log_2 M + \frac{1}{n}[g(n,\varepsilon) + f(n,\varepsilon)] \geqslant I(A;B|E)_\rho. \tag{A31}$$

By taking the limit as $n\to\infty$, then $\varepsilon\to 0$, and applying definitions, we can conclude the inequality $I(A;B|E)_\rho \leqslant \mathcal{D}(A;B|E)_\rho$.

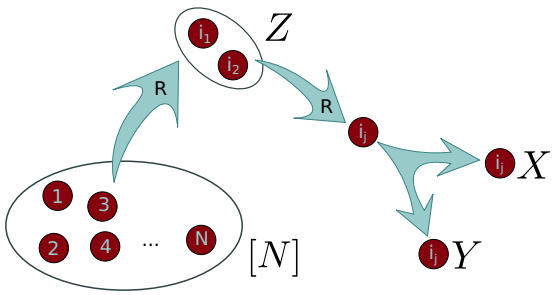


FIG. 3. As an example justifying why conditional correlations cannot be erased using a noise rate equal to the CQMI without conditioning on the pivot system, consider random variables X, Y, and Z constructed in the following way. Z is constructed by picking a random pair of numbers $i_1 \neq i_2$ from $[N] = \{1, 2, \ldots, N\}$, and $X = Y = i_b$ for a uniformly random bit b that is independent of Z. Clearly $I(X : Y|Z) = 1$. The classical analog of a random unitary channel is a random permutation channel. When conditioning on $Z = (i, j)$, one can just randomize the pair (i, j) using one bit of noise. Conversely, for a given pair (i_1, i_2), the mutual information $I(X : Y|Z = (i_1, i_2))$ is erased if and only if this pair is randomized by the random permutation channel. When operating on X alone, this implies that all pairs have to be randomized, which essentially implies that all numbers $1, \ldots, N$ have to be randomized jointly, requiring $\log_2 N$ bits of noise.

APPENDIX B: REQUIREMENT OF THE ACCESS TO THE CONDITIONING SYSTEM: CLASSICAL EXAMPLE

The following example shows that, even for the classical analog of state deconstruction, access to the conditioning system is necessary. Otherwise, the Markovianizing cost, as it is called in Ref. [32], can be arbitrarily large compared to the conditional mutual information. The example is informally explained in Fig. 3.

Let X and Y be random variables, each taking values from the alphabet $[n] = \{0, 1, \ldots, n-1\}$, and let Z be a random variable taking values in $[n] \times [n]$, such that the joint probability distribution for (X, Y, Z) is as follows:

$$p_{XYZ}(i, j, (k, l)) = \begin{cases} \frac{1}{n(n-1)} & \text{if } k < l,\ i = j \in \{k, l\} \\ 0 & \text{else} \end{cases}. \tag{B1}$$

For $k < l$, conditioned on $Z = (k, l)$, X and Y are maximally correlated fair coins, so that

$$I(X : Y|Z) = 1. \tag{B2}$$

One classical analog of tracing out a subsystem is applying a function $f : [n] \to [m]$, where $n/m \in \mathbb{N}$ such that $|f^{-1}(\{j\})| = n/m$ for all $j \in [m]$.

Let $f : [n] \to [m]$ be such a function and set $X' = f(X)$. Let $X', Y, Z \sim p'$ so that

$$p'_{X'YZ}(i, j, (k, l)) = \begin{cases} \frac{1}{n(n-1)} & \text{if } k < l,\ i = f(j),\ j \in \{k, l\} \\ 0 & \text{else} \end{cases}. \tag{B3}$$

Let us look at the fidelity of recovery $F(X'; Y|Z)$. It is easy to see that we can restrict to classical recovery channels: Let $\mathcal{R}_{Z\to ZX'}$ be an arbitrary (quantum) recovery channel, denote the measurement of the computational basis on system X by Λ_X, etc., and let $\rho_{X'YZ}$ be the diagonal quantum state representing p'. The fidelity does not decrease under the application of $\Lambda_{X'} \otimes \Lambda_Z$ and any classical state is invariant; therefore, $\mathcal{R}'_{Z\to ZX'} = (\Lambda_{X'} \otimes \Lambda_Z) \circ \mathcal{R}_{Z\to ZX'}$ is at least as good a recovery channel as $\mathcal{R}_{Z\to ZX'}$. Now, as $\Lambda_Z(\rho_{YZ}) = \rho_{YZ}$, we can also precompose a measurement without changing the fidelity; i.e., the desired classical recovery channel that is as good as $\mathcal{R}_{Z\to ZX'}$ is $\mathcal{R}'_{Z\to ZX'} = (\Lambda_{X'} \otimes \Lambda_Z) \circ \mathcal{R}_{Z\to ZX'} \circ \Lambda_Z$.

Let us then take an arbitrary classical recovery channel given by a conditional probability distribution $q_{X'Z'|Z}$. The resulting recovered distribution is

$$\hat{p}_{X'YZ}(i, j, \{k, l\}) = \sum_{\substack{k'<l' \\ k', l' \in [n]}} p_{YZ}(j, \{k', l'\}) q_{X'Z'|Z}(i, \{k, l\}|\{k', l'\}) \tag{B4}$$

$$= \frac{1}{n(n-1)} \sum_{\substack{k'<l' \\ k', l' \in [n]}} (\delta_{k'j} + \delta_{l'j}) q_{X'Z'|Z}(i, \{k, l\}|\{k', l'\}) \tag{B5}$$

$$= \frac{1}{n(n-1)} \sum_{\substack{l' \in [n] \\ l' \neq j}} q_{X'Z'|Z}(i, \{k, l\}|\{j, l'\}). \tag{B6}$$

Now we look at the fidelity with the original distribution, i.e.,

$$\sqrt{F(p'_{X'YZ}, \hat{p}_{X'YZ})} = \sum_{\substack{k<l \\ k, l \in [n]}} \sum_{i \in [m]} \sum_{j \in [n]} \sqrt{p'_{X'YZ}(i, j, \{k, l\}) \hat{p}_{X'YZ}(i, j, \{k, l\})} \tag{B7}$$

$$= \frac{1}{n(n-1)} \sum_{\substack{k \neq l \\ k, l \in [n]}} \sqrt{\sum_{\substack{l' \in [n] \\ l' \neq k}} q_{X'Z'|Z}(f(k), \{k, l\}|\{k, l'\})}. \tag{B8}$$

It is obvious that the optimal recovery channel has $q_{X'Z'|Z}(i, \{k, l\}|\{k', l'\}) = 0$ whenever k, l, k', l' are all different or $f(k) \neq i \neq f(l)$. Let us therefore assume this is the case. Let $\lambda_{kl} = \sum_{\substack{l' \in [n] \\ l' \neq k}} q_{X'Z'|Z}(f(k), \{k, l\}|\{k, l'\})$. Then we have

$$\sum_{\substack{k<l \\ f(k)=f(l)}} \lambda_{kl} + \sum_{\substack{k \neq l \\ f(k)\neq f(l)}} \lambda_{kl} = n(n-1)/2 \tag{B9}$$

due to the normalization of the conditional distribution q. Suppose first (B9) and $\lambda_{kl} \geqslant 0$ are the only restrictions on the possible λ_{kl}. Then the optimal choice is

$$\lambda_{kl} = \frac{(n-1)}{2(n-1) - (n/m - 1)},$$

i.e., constant λ_{kl}. We can now bound the fidelity of recovery

$$\sqrt{F(X';Y|Z)_{p'}} = \max_q \sqrt{F(p',\hat{p})} \tag{B10}$$

$$\leqslant \max_{\lambda_{kl}\geqslant 0} \frac{1}{n(n-1)} \sum_{\substack{k\neq l \\ k,\,l\in[n]}} \sqrt{\lambda_{kl}} \tag{B11}$$

$$= \sqrt{\frac{(n-1)}{2(n-1)-(n/m-1)}}. \tag{B12}$$

Here the maxima are taken over conditional probability distributions and the positive λ_{kl} that sum to $n(n-1)/2$, respectively. The inequality is due to the fact that by relaxing the conditions on λ_{kl} we maximize over a larger set. For $F(X';Y|Z)_{p'} \geqslant 1-\varepsilon$ this implies

$$\log(n/m) \geqslant \log(n-1) + \log\left(\frac{1-2\varepsilon}{1-\varepsilon}\right). \tag{B13}$$

In words, the required noise can be arbitrarily large compared to the conditional mutual information. A similar analysis can be done for many i.i.d. copies of X, Y, and Z.

[1] R. Landauer, Irreversibility and heat generation in the computing process, IBM J. Res. Dev. **5**, 183 (1961).

[2] J. Goold, M. Huber, A. Riera, L. del Rio, and P. Skrzypczyk, The role of quantum information in thermodynamics—a topical review, J. Phys. A: Math. Theor. **49**, 143001 (2016).

[3] B. Groisman, S. Popescu, and A. Winter, Quantum, classical, and total amount of correlations in a quantum state, Phys. Rev. A **72**, 032317 (2005).

[4] M. M. Wilde, From classical to quantum Shannon theory, arXiv:1106.1445v7.

[5] J. Oppenheim, State redistribution as merging: Introducing the coherent relay, arXiv:0805.1065.

[6] A. Abeyesinghe, I. Devetak, P. Hayden, and A. Winter, The mother of all protocols: Restructuring quantum information's family tree, Proc. R. Soc. London, Ser. A **465**, 2537 (2009).

[7] B. Schumacher and M. D. Westmoreland, Approximate quantum error correction, Quant. Info. Proc. **1**, 5 (2002).

[8] P. Hayden, M. Horodecki, A. Winter, and J. Yard, A decoupling approach to the quantum capacity, Open Syst. Info. Dyn. **15**, 7 (2008).

[9] M. Berta, F. G. S. L. Brandão, C. Majenz, and M. M. Wilde, Conditional Decoupling of Quantum Information, Phys. Rev. Lett. **121**, 040504 (2018).

[10] P. Hayden, R. Jozsa, D. Petz, and A. Winter, Structure of states which satisfy strong subadditivity of quantum entropy with equality, Commun. Math. Phys. **246**, 359 (2004).

[11] O. Fawzi and R. Renner, Quantum conditional mutual information and approximate Markov chains, Commun. Math. Phys. **340**, 575 (2015).

[12] M. Berta, K. Seshadreesan, and M. M. Wilde, Rényi generalizations of the conditional quantum mutual information, J. Math. Phys. **56**, 022205 (2015).

[13] M. M. Wilde, N. Datta, M.-H. Hsieh, and A. Winter, Quantum rate-distortion coding with auxiliary resources, IEEE Trans. Inf. Theory **59**, 6755 (2013).

[14] I. Devetak and J. Yard, Exact Cost of Redistributing Multipartite Quantum States, Phys. Rev. Lett. **100**, 230501 (2008).

[15] J. Yard and I. Devetak, Optimal quantum source coding with quantum side information at the encoder and decoder, IEEE Trans. Inf. Theory **55**, 5339 (2009).

[16] K. P. Seshadreesan and M. M. Wilde, Fidelity of recovery, squashed entanglement, and measurement recoverability, Phys. Rev. A **92**, 042321 (2015).

[17] A. Uhlmann, The "transition probability" in the state space of a *-algebra, Rep. Math. Phys. **9**, 273 (1976).

[18] W. H. Zurek, Einselection and decoherence from an information theory perspective, Ann. Phys. **9**, 855 (2000).

[19] H. Ollivier and W. H. Zurek, Quantum Discord: A Measure of the Quantumness of Correlations, Phys. Rev. Lett. **88**, 017901 (2001).

[20] M. Christandl and A. Winter, Squashed entanglement: An additive entanglement measure, J. Math. Phys. **45**, 829 (2004).

[21] W. H. Zurek, Decoherence, einselection, and the quantum origins of the classical, Rev. Mod. Phys. **75**, 715 (2003).

[22] C. A. Fuchs and J. van de Graaf, Cryptographic distinguishability measures for quantum mechanical states, IEEE Trans. Inf. Theory **45**, 1216 (1998).

[23] A. Gilchrist, N. K. Langford, and M. A. Nielsen, Distance measures to compare real and ideal quantum processes, Phys. Rev. A **71**, 062310 (2005).

[24] M. Tomamichel, R. Colbeck, and R. Renner, A fully quantum asymptotic equipartition property, IEEE Trans. Inf. Theory **55**, 5840 (2009).

[25] E. H. Lieb and M. B. Ruskai, Proof of the strong subadditivity of quantum-mechanical entropy, J. Math. Phys. **14**, 1938 (1973).

[26] E. H. Lieb and M. B. Ruskai, A Fundamental Property of Quantum-Mechanical Entropy, Phys. Rev. Lett. **30**, 434 (1973).

[27] A. Winter, Tight uniform continuity bounds for quantum entropies: Conditional entropy, relative entropy distance and energy constraints, Commun. Math. Phys. **347**, 291 (2016).

[28] M.-Y. Ye, Y.-K. Bai, and Z. D. Wang, Quantum state redistribution based on a generalized decoupling, Phys. Rev. A **78**, 030302 (2008).

[29] M. Berta, M. Christandl, and D. Touchette, Smooth entropy bounds on one-shot quantum state redistribution, IEEE Trans. Inf. Theory **62**, 1425 (2016).

[30] N. Datta, M.-H. Hsieh, and J. Oppenheim, An upper bound on the second order asymptotic expansion for the quantum communication cost of state redistribution, J. Math. Phys. **57**, 052203 (2016).

[31] C. Majenz, M. Berta, F. Dupuis, R. Renner, and M. Christandl, Catalytic Decoupling of Quantum Information, Phys. Rev. Lett. **118**, 080503 (2017).

[32] E. Wakakuwa, A. Soeda, and M. Murao, Markovianizing cost of tripartite quantum states, IEEE Trans. Inf. Theory **63**, 1280 (2017).

[33] E. Wakakuwa, A. Soeda, and M. Murao, The cost of randomness for converting a tripartite quantum state to be approximately recoverable, IEEE Trans. Inf. Theory **63**, 5360 (2017).

[34] M. Berta, F. G. S. L. Brandao, and M. M. Wilde (unpublished).
[35] M. Koashi and N. Imoto, Operations that do not disturb partially known quantum states, Phys. Rev. A **66**, 022318 (2002).
[36] E. Wakakuwa, A. Soeda, and M. Murao, A coding theorem for bipartite unitaries in distributed quantum computation, IEEE Trans. Inf. Theory **63**, 5372 (2017).
[37] C. H. Bennett, Logical reversibility of computation, IBM J. Res. Dev. **17**, 525 (1973).
[38] A. Ambainis, M. Mosca, A. Tapp, and R. de Wolf, Private quantum channels, in *Proceedings of the 41st Annual Symposium on Foundations of Computer Science* (IEEE Computer Society, Washington, DC, 2000), pp. 547–553.
[39] M. Berta and M. Tomamichel, The fidelity of recovery is multiplicative, IEEE Trans. Inf. Theory **62**, 1758 (2016).
[40] R. Alicki and M. Fannes, Continuity of quantum conditional information, J. Phys. A: Math. Gen. **37**, L55 (2004).
[41] K. Modi, A. Brodutch, H. Cable, T. Paterek, and V. Vedral, The classical-quantum boundary for correlations: Discord and related measures, Rev. Mod. Phys. **84**, 1655 (2012).
[42] G. Adesso, T. R. Bromley, and M. Cianciaruso, Measures and applications of quantum correlations, J. Phys. A: Math. Theor. **49**, 473001 (2016).
[43] M. Piani, Problem with geometric discord, Phys. Rev. A **86**, 034101 (2012).
[44] K. P. Seshadreesan, M. Berta, and M. M. Wilde, Rényi squashed entanglement, discord, and relative entropy differences, J. Phys. A: Math. Theor. **48**, 395303 (2015).
[45] M. M. Wilde, Multipartite quantum correlations and local recoverability, Proc. R. Soc. London, Ser. A **471**, 20140941 (2014).
[46] V. Madhok and A. Datta, Interpreting quantum discord through quantum state merging, Phys. Rev. A **83**, 032323 (2011).
[47] D. Cavalcanti, L. Aolita, S. Boixo, K. Modi, M. Piani, and A. Winter, Operational interpretations of quantum discord, Phys. Rev. A **83**, 032324 (2011).
[48] K. Li and A. Winter, Squashed entanglement, k-extendibility, quantum Markov chains, and recovery maps, Found. Phys. **48**, 910 (2018).
[49] J. Oppenheim, A paradigm for entanglement theory based on quantum communication, arXiv:0801.0458.
[50] R. F. Werner, An application of Bell's inequalities to a quantum state extension problem, Lett. Math. Phys. **17**, 359 (1989).
[51] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri, Complete family of separability criteria, Phys. Rev. A **69**, 022308 (2004).
[52] F. G. S. L. Brandao and A. W. Harrow, Quantum de Finetti theorems under local measurements with applications, Commun. Math. Phys. **353**, 469 (2017).
[53] D. Ding, P. Hayden, and M. Walter, Conditional mutual information of bipartite unitaries and scrambling, J. High Energy Phys. 12 (2016) 145.
[54] D. Yang, K. Horodecki, M. Horodecki, P. Horodecki, J. Oppenheim, and W. Song, Squashed entanglement for multipartite states and entanglement measures based on the mixed convex roof, IEEE Trans. Inf. Theory **55**, 3375 (2009).
[55] D. Avis, P. Hayden, and I. Savov, Distributed compression and multiparty squashed entanglement, J. Phys. A: Math. Theor. **41**, 115301 (2008).
[56] A. Anshu, V. K. Devabathini, and R. Jain, Near Optimal Bounds on Quantum Communication Complexity of Single-Shot Quantum State Redistribution, Phys. Rev. Lett. **119**, 120506 (2017).
[57] H. Araki and E. H. Lieb, Entropy inequalities, Commun. Math. Phys. **18**, 160 (1970).