

# Complexity-Constrained Quantum Thermodynamics

Anthony Munson<sup>1,2</sup>, Naga Bhavya Teja Kothakonda,<sup>3,4</sup> Jonas Haferkamp,<sup>5</sup>  
Nicole Yunger Halpern<sup>1,2,6</sup>, Jens Eisert,<sup>4</sup> and Philippe Faist<sup>4,\*</sup>

<sup>1</sup>*Joint Center for Quantum Information and Computer Science, NIST and University of Maryland, College Park, Maryland 20742, USA*

<sup>2</sup>*Department of Physics, University of Maryland, College Park, Maryland 20742, USA*

<sup>3</sup>*Grup d'Informació Quàntica, Departament de Física, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain*

<sup>4</sup>*Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany*

<sup>5</sup>*School of Engineering and Applied Sciences, Harvard University, Massachusetts 02134, USA*

<sup>6</sup>*Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742, USA*



(Received 1 April 2024; accepted 3 February 2025; published 10 March 2025)

Quantum complexity measures the difficulty of realizing a quantum process, such as preparing a state or implementing a unitary. We present an approach to quantifying the thermodynamic resources required to implement a process if the process's complexity is restricted. We focus on the prototypical task of information erasure, or Landauer erasure, wherein an  $n$ -qubit memory is reset to the all-zero state. We show that the minimum thermodynamic work required to reset an arbitrary state in our model, via a complexity-constrained process, is quantified by the state's *complexity entropy*. The complexity entropy therefore quantifies a trade-off between the work cost and complexity cost of resetting a state. If the qubits have a nontrivial (but product) Hamiltonian, the optimal work cost is determined by the *complexity relative entropy*. The complexity entropy quantifies the amount of randomness a system appears to have to a computationally limited observer. Similarly, the complexity relative entropy quantifies such an observer's ability to distinguish two states. We prove elementary properties of the complexity (relative) entropy. In a random circuit—a simple model for quantum chaotic dynamics—the complexity entropy transitions from zero to its maximal value around the time corresponding to the observer's computational-power limit. Also, we identify information-theoretic applications of the complexity entropy. The complexity entropy quantifies the resources required for data compression if the compression algorithm must use a restricted number of gates. We further introduce a *complexity conditional entropy*, which arises naturally in a complexity-constrained variant of information-theoretic decoupling. Assuming that this entropy obeys a conjectured chain rule, we show that the entropy bounds the number of qubits that one can decouple from a reference system, as judged by a computationally bounded referee. Overall, our framework extends the resource-theoretic approach to thermodynamics to integrate a notion of *time*, as quantified by *complexity*.

DOI: [10.1103/PRXQuantum.6.010346](https://doi.org/10.1103/PRXQuantum.6.010346)

## I. INTRODUCTION

Quantum complexity is drawing increasing interest across physics, from many-body physics to quantum gravity [1–4]. For the purposes of our work, the quantum complexity of a unitary operation (respectively, a quantum state) is the minimum number of operations required to

implement the unitary (respectively, to prepare the state). Each operation is chosen from a given set of elementary operations (e.g., a universal set of two-qubit gates). In condensed-matter physics, preparing a topologically ordered state requires a circuit of sufficient complexity to spread information throughout a system [1,2,5–11]. Near-term quantum devices aim to prepare states of sufficient complexities to offer quantum advantages attributable, for example, to the hardness of sampling classically from such states [12–15]. Quantum complexity recently gained significance in the context of the *anti-de Sitter space and conformal field theory* (AdS-CFT) correspondence in high-energy physics: In a prominent conjecture, the complexity of the field-theoretic state dual to a wormhole

\*Contact author: [philippe.faist@fu-berlin.de](mailto:philippe.faist@fu-berlin.de)

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

connecting two black holes is proportional to the worm-hole’s length [3,4,16–30].

The relevance of quantum complexity to AdS-CFT motivates connections to thermodynamics. Brown and Susskind posited that the CFT state’s complexity should tend to increase, formulating a “second law of complexity” [3]. Bai *et al.* extended the second law of complexity by proving fluctuation relations mirroring Jarzynski’s equality in statistical mechanics [31]. A resource theory of *uncomplexity*—a state’s closeness to a simple tensor product—was furthermore established in Ref. [30]. Quantum complexity also appears connected to ergodicity and quantum chaos: complexity is believed to grow linearly for long times under typical quantum chaotic dynamics; complexity would thereby provide a universal measure for how long a chaotic system has evolved [4,17–20,32]. In contrast, standard correlation functions and entanglement entropies typically reach their equilibrium values after short times [4].

Our main goal is to identify which state transformations in quantum thermodynamics can be effected by processes, and probed with measurement effects, utilizing limited computational resources, hence of limited complexity. In particular, we seek to connect quantum complexity and entropy as follows. In conventional thermodynamics, the minimum amount of work needed to transform one equilibrium state into another, via exchanges of heat at a fixed temperature, is determined by the energy and entropy differences between the initial and final states. The complexity of a many-body system’s evolution is upper-bounded by the evolution’s duration. An observer with little computational power typically cannot distinguish a highly complex pure state from a highly entropic mixed state. A phenomenological description of possible state transformations under short-time evolutions, from the viewpoint of such an observer, should not distinguish highly entropic initial states from highly complex pure states. The states become distinguishable if observed over sufficiently long timescales or through sufficiently complex observables. This fact invites us to define thermodynamic potentials for determining which state transformations can be effected under complexity limitations. The roles of these potentials mirror the role of entropy in standard thermodynamics. Using the potentials, we address our general question: *how does one formulate thermodynamics at a given complexity scale—for complexity-constrained processes and observers?*

Our analysis relies on recent information-theoretic frameworks for quantum thermodynamics. The relevance of an observer’s information, or knowledge, in thermodynamics was significantly clarified by the pioneering works of Szilárd [33], Landauer [34], and Bennett [35]. Landauer argued that erasing a bit of information dissipates an amount of heat  $\geq k_B T \log(2)$  [34]. The  $k_B$  denotes Boltzmann’s constant,  $T$  denotes an environment’s

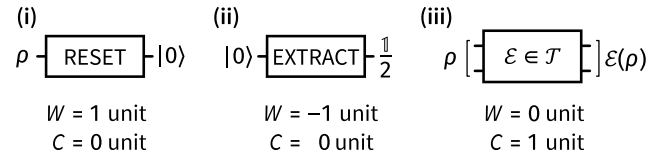


FIG. 1. A simple model combining quantum information thermodynamics and complexity: An  $n$ -qubit memory register is governed by a completely degenerate Hamiltonian. Processes consist of the primitive operations (i)–(iii). Work and complexity costs are defined as the sums of the corresponding costs over the operations. (a) A qubit can be reset from any state  $\rho$  to a standard state  $|0\rangle$ . The resetting costs one unit of work ( $W$ ), by Landauer’s principle [34]. This primitive has no complexity cost ( $C$ ). (b) One unit of work can be extracted from a qubit in the state  $|0\rangle$ . The extraction leaves the qubit maximally mixed [33]. (c) Operations  $\mathcal{E}$  are chosen from a set  $\mathcal{T}$  of elementary computations. Each  $\mathcal{E}$  costs one unit of complexity but no work. A natural choice for  $\mathcal{T}$  is the set of two-qubit unitary operations, potentially subject to connectivity constraints. Choices of  $\mathcal{T}$  natural to thermodynamic applications preserve thermal states.

temperature, and the logarithm is base- $e$ . This observation led to Bennett’s resolution of Maxwell’s demon paradox [35,36] and helped extend quantum thermodynamics to far-from-equilibrium systems [37–40]. A common model for quantum thermodynamics is the resource theory of thermal operations [37,38,41–44]: one assumes that energy-conserving interactions with a fixed-temperature heat bath are the only operations performable without external resources, such as thermodynamic work. Using this model, one can determine whether a state  $\rho$  can transform into a state  $\sigma$ , for large classes of states. The answer can be cast in terms of a family of entropy measures termed the Rényi- $\alpha$  relative entropies [44]. A closely related model captures how many natural, or spontaneous, dynamics have a particular fixed point. Using this model, one can quantify, e.g., the thermodynamic work required to implement a general quantum process [40,45,46]. A key observation is that a process’s work cost is tied fundamentally to its logical irreversibility: information erasure is costly, but energy-conserving unitary operations can, in principle, be implemented at zero work cost [34–36]. We introduce a simple model that captures the main features of the above models and that forms the basis for our analysis. Our model centers on an  $n$ -qubit memory register governed by a completely degenerate Hamiltonian. The following primitive operations are performable (Fig. 1): (i) the reset of one qubit from any state to a standard state  $|0\rangle$ , costing one unit of work; (ii) the preparation of one qubit in the maximally mixed state from  $|0\rangle$ , extracting one unit of work; and (iii) the implementation of one two-qubit unitary gate, costing one unit of complexity. (Our model neglects any complexity cost of single-qubit erasure [47], to separate two types of costs: the thermodynamic

work cost associated with reducing the entropy of a single qubit, and the computational cost of exchanging information between the qubits. The former operations involve thermodynamic work and act on individual qubits. The latter operations involve two-qubit interactions that tend to spread quantum information throughout the system. Our model generalizes to elementary operations that cost both work and complexity.)

We first revisit the standard setting of *information erasure*, or *Landauer erasure*. Erasure has clarified information's role in thermodynamics; we now use erasure to clarify complexity's role in quantum thermodynamics. We consider an  $n$ -qubit system with a product Hamiltonian  $H$  and a zero-energy ground state. We prove that the minimum work required to reset an arbitrary state  $\rho$  to the ground state, using at most a fixed number of thermodynamic processes in our model, is given by the *complexity relative entropy* of  $\rho$  relative to a thermal state [30]. The complexity relative entropy quantifies two states' distinguishability, as judged by an observer with limited computational power. If  $H = 0$ , this quantity reduces to the *complexity entropy* [30] and quantifies how random a state appears to such an observer. A variant of the complexity (relative) entropy appeared in our Ref. [30] to quantify the number of pure qubits extractable from a state, in the resource theory of uncomplexity. The present work further applies the complexity relative entropy to thermodynamics: we extend the pure-qubit extraction protocol of Ref. [30] to general protocols for thermodynamic information erasure under complexity limitations.

Our result quantifies a trade-off between the work and complexity costs of erasing an  $n$ -qubit state  $\rho$  (Fig. 2). Suppose that  $H = 0$ , and consider a highly complex pure state  $|\psi\rangle$ . In standard models for information thermodynamics, one can implement unitary operations on an energy-degenerate memory register at zero work cost, since such operations are logically reversible [43,48,49]. Therefore,  $|\psi\rangle$  can, in principle, be transformed into the all-zero state  $|0^n\rangle$  at no work cost. Yet, this transformation requires many gates, incurring a high *complexity cost*. An alternative procedure would reset each memory qubit with a thermodynamic reset operation. This procedure costs  $n k_B T \log(2)$  units of work but no complexity. The complexity entropy quantifies the trade-off between the work and complexity costs of erasing  $|\psi\rangle$ .

Bennett *et al.* analyzed a trade-off between complexity and work in the context of classical bit erasure [50]. Kolmogorov complexity, rather than quantum complexity, is relevant to their problem. The connection between Kolmogorov complexity and thermodynamics was further cemented in Zurek's work [51] and in Refs. [52–54]. Kolmogorov complexity and quantum complexity quantify the size of a state's "smallest description" in different ways. Kolmogorov complexity quantifies the size of the smallest program that generates the state (regardless of the

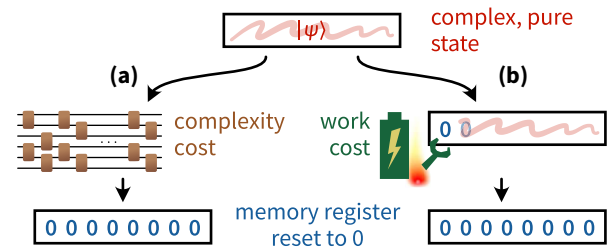


FIG. 2. An example trade-off between the work cost and complexity cost of resetting an  $n$ -qubit memory register initiated in a highly complex pure state  $|\psi\rangle$ . (a) A unitary operation resets the memory to the all-zero state. In principle, the unitary costs no thermodynamic work. Many gates are necessary to implement the unitary, incurring a high complexity cost. (b) To reset the memory without gates, one can erase every qubit [via primitive (i) in Fig. 1]. This operation costs  $n$  units of work total. The *complexity entropy* quantifies the work-complexity trade-off of erasing an arbitrary state  $\rho$ . The complexity entropy was introduced in Ref. [30]; this work elucidates its properties and applications.

program's runtime). In contrast, quantum complexity measures the shortest runtime of a program that generates the state (regardless of whether the program has a compact representation).

The remainder of this work is dedicated to the analysis of more-general thermodynamic processes, including erasure in systems with nontrivial Hamiltonians, and to a deeper study of the complexity entropy's properties, uses, and extensions. We study the work costs of general state transformations in quantum thermodynamics where complexity limitations restrict which processes and measurements one can implement. Furthermore, we evidence the complexity entropy's broad relevance to quantum information theory. First, we present properties of the complexity entropy and bound it using well-known complexity measures and an entanglement measure. Then, we demonstrate the complexity entropy's relevance in random circuits, a simple model of quantum chaotic dynamics. Last, we apply our complexity-entropy measures to information-theoretic tasks with complexity constraints.

What is the work cost of a general state transformation  $\rho \rightarrow \rho'$ , as in the model of Fig. 1? We consider the minimum work cost of any process, of complexity at most  $r \geq 0$ , that maps  $\rho$  to a state indistinguishable from  $\rho'$ . Here, distinguishability is judged by an observer who possesses some bounded computational power  $R \geq 0$ . This work cost extends the complexity relative entropy's use to general state transformations. Quantifying the work cost addresses our main goal of establishing which state transformations are possible in quantum thermodynamics for complexity-constrained agents. Considering regimes in which  $R$  is either extremely high or extremely low, we identify cases in which bounding the work cost is sometimes possible.

Our complexity-entropy measures obey certain properties expected of entropies. For instance, the complexity entropy achieves its maximum value on a maximally mixed state. Also, the complexity relative entropy decreases monotonically under partial traces. Our measures lack other common properties, such as invariance under unitaries and monotonicity under completely positive, trace-preserving maps (a data-processing inequality): by construction, the complexity (relative) entropy is sensitive to a state's complexity, which can change under arbitrary quantum operations. Moreover, the complexity entropy obeys bounds involving known complexity measures: one bound involves the *strong complexity* of Ref. [17]; and the other, the *approximate circuit complexity* (the minimum complexity of any unitary that approximately prepares a target state). For a one-dimensional (1D) chain of qubits, the complexity entropy obeys also a bound involving an entanglement measure defined in terms of the quantum mutual information.

We argue that the complexity entropy can quantify chaotic behavior in a quantum many-body system. Consider a suitably chaotic many-body system initialized in a pure product state. The state's complexity entropy should remain low until the timescale matching an observer's computational power. Beyond this timescale, the complexity entropy should be close to maximal. We employ and extend the results of Refs. [17,55] to prove a corresponding statement about the output of a random quantum circuit applied to  $|0^n\rangle$ , an  $n$ -qubit system's all-zero state. Suppose that an observer can measure only observables of complexities  $\leq r$ . The complexity entropy, we show, transitions from zero to  $n - O(1)$  when the number of gates reaches  $\approx r$  (Fig. 3). For a random circuit of  $\lesssim r$  gates, the observer can apply the inverse circuit to recover the all-zero state and so to ascertain that the random circuit's output has a low entropy. For a random circuit of  $\gtrsim r$  gates, the observer cannot distinguish the circuit's output from a maximally mixed state, using observables of complexities  $\leq r$  [17,32]. Accordingly, the output has a complexity entropy of at least  $n - O(1)$ . Reference [17] quantifies a random circuit's complexity by applying a powerful tool in random-circuit analysis: a *unitary  $k$ -design* [55,57]. The lower bound's linear scaling comes from recently improved bounds for the design order achieved by random circuits [32] over existing bounds [55,56].

We use our complexity-entropy measures to bound the optimal efficiencies of information-theoretic tasks performed under complexity limitations. Thermodynamic erasure is related to data compression, the storage of information in the smallest possible register [58–60]: to erase a memory register, one can first perform data compression to reduce the number of qubits that need resetting. The complexity entropy quantifies the resources required to compress a quantum state into the fewest qubits possible, via any limited-complexity procedure.

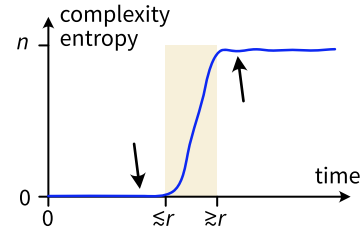


FIG. 3. The complexity entropy of a state outputted by a random circuit, relative to an observer who can implement only measurement effects of complexities  $\leq r$ . We determine the complexity entropy's evolution using techniques from Refs. [17,55,56]. After a time  $t \lesssim r$ , the state has a low enough complexity that the observer can verify the state's purity via measurement. After  $t \gtrsim r$ , the state is complex enough to be nearly indistinguishable from a highly entropic state, according to any measurement whose complexity is  $\leq r$ .

After addressing data compression, we study *decoupling*, or ensuring that an agent's multiqubit system becomes maximally mixed and uncorrelated with a referee's system,  $R$  [61–64]. Assuming a conjecture about a chain rule for the complexity entropy, we lower-bound the number of qubits that an agent must discard to obtain a state indistinguishable, by a complexity-restricted referee, from a maximally mixed state uncorrelated with  $R$ . The lower-bound is given by a conditional variant of the complexity entropy. One can interpret the variant as a *complexity-aware conditional min-entropy*.

Our paper is organized as follows. In Sec. II, we introduce our framework, background information, and the complexity entropy. In Sec. III, we bound the work cost of information erasure subject to complexity restrictions. In Sec. IV, we generalize our analysis to arbitrary state transformations. In Sec. V, we present the complexity entropy's information-theoretic properties and applications. We conclude in Sec. VI.

## II. SETTING

In Sec. II A, we introduce our setup and allowed operations. In Sec. II B, we review *one-shot information theory*, which quantifies the efficiencies of thermodynamic tasks performed on finite-size systems, with arbitrary success probabilities. In Sec. II C, we introduce the complexity entropy and complexity relative entropy.

### A. Thermodynamic framework

We consider a system of  $n$  noninteracting qubits. Qubit  $i$  evolves under a Hamiltonian  $H_i$ , and the entire system under the Hamiltonian  $H = H_1 + H_2 + \dots + H_n$ . For simplicity, we suppose that  $H_i |0\rangle_i = 0$  and  $H_i |1\rangle_i = E_i^{(1)} |1\rangle_i$ , with  $E_i^{(1)} \geq 0$ . The all-zero bit string  $|0^n\rangle$  labels the zero-energy ground state. For a fixed inverse temperature



$\beta = 1/k_B T$ , qubit  $i$  has the thermal state

$$\gamma_i = \Gamma_i / Z_i, \quad (1)$$

wherein  $\Gamma_i = e^{-\beta H_i}$  and  $Z_i = \text{tr}(\Gamma_i) = 1 + e^{-\beta E_i^{(1)}}$ . The thermal qubit's free energy is given by

$$F_i = -k_B T \log(Z_i). \quad (2)$$

The *thermal operations* model possible transformations in quantum thermodynamics, for a system in controlled contact with a heat bath [38,42,43]. Let  $S$  denote a system with a Hamiltonian  $H_S$ . A thermal operation  $\Phi_S$  is defined as any operation of the form

$$\rho \mapsto \Phi_S(\rho) = \text{tr}_B \left( V_{SB} [\rho \otimes \gamma_B] V_{SB}^\dagger \right). \quad (3)$$

$B$  denotes any auxiliary system with a Hamiltonian  $H_B$ .  $V_{SB}$  denotes any unitary that strictly conserves energy:  $[V_{SB}, H_S + H_B] = 0$ . For an arbitrary system  $X$  governed by a Hamiltonian  $H_X$ , the thermal state  $\gamma_X = e^{-\beta H_X} / \text{tr}(e^{-\beta H_X})$  is defined similarly to Eq. (1).

Another general class of operations consists of the *Gibbs-preserving maps* [45]. Let  $S$  denote a system with a Hamiltonian  $H_S$ . A Gibbs-preserving map  $\mathcal{F}$  is any completely positive, trace-preserving map that satisfies  $\mathcal{F}(\gamma_S) = \gamma_S$ . The Gibbs-preserving maps form a larger class than the thermal operations [65]. It remains unclear what resource costs are necessary for implementing a general Gibbs-preserving map, using other standard thermodynamic operations [66]. When  $H_S = 0$ ,  $\gamma_S$  is proportional to the identity operator  $\mathbb{1}_S$ , so every Gibbs-preserving map  $\mathcal{F}$  is a unital map [ $\mathcal{F}(\mathbb{1}_S) = \mathbb{1}_S$ ] and vice versa.

Inspired by the resource theory of thermodynamics, we use a model in which all processes consist of primitive operations. A process is specified as a sequence  $\mathcal{P} := (\mathcal{E}_1, \dots, \mathcal{E}_m)$  of primitive operations. The sequence implements the operation  $\mathcal{E}_{\mathcal{P}} := \mathcal{E}_m \cdots \mathcal{E}_1$ . Each primitive operation  $\mathcal{E}_i$  has a complexity cost  $C(\mathcal{E}_i)$  and a work cost  $W(\mathcal{E}_i)$ . Accordingly, each process  $\mathcal{P}$  has a complexity cost  $C(\mathcal{P})$  and a work cost  $W(\mathcal{P})$ , which are sums of the primitive operations' costs

$$C(\mathcal{P}) := \sum_{i=1}^m C(\mathcal{E}_i), \quad \text{and} \quad W(\mathcal{P}) := \sum_{i=1}^m W(\mathcal{E}_i). \quad (4)$$

One might refer to  $C(\mathcal{P})$  instead as the *circuit size* of  $\mathcal{P}$ . Suppose that  $\mathcal{P}$  implements a unitary operation  $\mathcal{U}$ .  $C(\mathcal{P})$  might differ from the complexity  $C(\mathcal{U})$  of  $\mathcal{U}$ : there might exist a process that also implements  $\mathcal{U}$  but consists of fewer primitive operations.

The primitive operations on an  $n$ -qubit system, given a background inverse temperature  $\beta = 1/k_B T$ , are the following (Fig. 1):

- (i) The RESET operation: Qubit  $i$  can be brought from an arbitrary state to the ground state. This operation has the work cost  $W_{\text{RESET},i} = -F_i = k_B T \log(Z_i)$  and no complexity cost:  $C_{\text{RESET},i} = 0$ .
- (ii) The EXTRACT operation: Work  $|F_i|$  can be extracted as qubit  $i$  is brought from  $|0\rangle_i$  to the thermal state  $\gamma_i$ . This operation has the work cost  $W_{\text{EXTRACT},i} = F_i = -W_{\text{RESET},i}$  and no complexity cost:  $C_{\text{EXTRACT},i} = 0$ .
- (iii) Every operation  $\mathcal{E}$  chosen from a fixed set  $\mathcal{T}$  of elementary computations costs one unit of complexity,  $C(\mathcal{E}) = 1$ , and no work:  $W(\mathcal{E}) = 0$ . If  $H_i = 0$  for all  $i$ , then natural choices of  $\mathcal{T}$  include arbitrary two-qubit unitary gates with arbitrary connectivities. In this case,  $\mathcal{T} \simeq \bigcup_{(i,j)} \text{SU}(4)_{ij}$ , wherein each  $(i,j)$  denotes a pair of qubits. If  $H_i \neq 0$ , natural choices of  $\mathcal{T}$  include the two-qubit thermal operations and the two-qubit Gibbs-preserving maps, with arbitrary connectivities.

The work expended on the RESET, and the work extracted via EXTRACT, naturally generalize Landauer's bound to noninteracting qubits. The RESET and EXTRACT operations' ideal work costs are  $W_{\text{RESET},i} = -W_{\text{EXTRACT},i} = -F_i$  in standard thermodynamic models, including the resource theory of thermal operations [43,44,67–69]. The RESET operation can be applied to any input state, not only a thermal state. Its deterministic work cost,  $-F_i$ , can be viewed as equal to the worst-case work cost of resetting any particular input state. If a qubit has a completely degenerate Hamiltonian,  $H_i = 0$ , these work costs reduce to Landauer's bound:  $-F_i = k_B T \log(2)$ .

The set  $\mathcal{T}$  of elementary computations enables us to define a protocol's complexity cost. To separate elementary thermodynamic operations from elementary computing operations,  $\mathcal{T}$  should contain only operations to which one can reasonably assign no work costs. Crucially, we choose the elementary computations to act nontrivially on just two qubits. This property ensures that several elementary computations are needed to propagate information throughout the system and many more are needed to generate complex quantum states.

For energy-degenerate qubits ( $H_i = 0$  for all  $i$ ), a meaningful choice of  $\mathcal{T}$  is a universal set of two-qubit unitary gates: the primitive operation (iii) then enables quantum computation with unitary circuits. This choice is undesirable if the qubits have nontrivial Hamiltonians, however: some two-qubit unitaries require work and so should be excluded from  $\mathcal{T}$  (for example, a swap of two different-energy levels). One could choose for  $\mathcal{T}$  to consist of two-qubit unitaries that commute with the total Hamiltonian. Unfortunately, this choice typically leads to an unreasonably restricted set. The energy-conserving unitaries on two qubits with distinct Hamiltonians is limited to unitaries that are diagonal with respect to the product

energy eigenbasis. This choice could render impossible some natural operations that might otherwise be implementable with ancillary qubits and nondiagonal two-qubit unitaries. One such operation is a partial thermalization of a qubit:  $\rho \rightarrow \rho/2 + \gamma_i/2$ .

When the qubits have nontrivial Hamiltonians, we include in  $\mathcal{T}$  some nonunitary two-qubit operations to remedy the above problem. We still ensure that the  $\mathcal{T}$  elements cost no work in standard thermodynamic frameworks. Yet  $\mathcal{T}$  can include operations that are not implementable in practice. For instance, they can require unreasonable control. In such a case, the conclusions about the operations' work and complexity costs imply lower bounds applicable to every setting in which the physically implementable elementary computations form a smaller set  $\mathcal{T}_0$ . (These bounds might not be tight.) In fact, choosing a larger  $\mathcal{T}$  ensures that any such lower bounds apply to a broader range of possibilities for  $\mathcal{T}_0$ . In this spirit, we identify two  $\mathcal{T}$ s that are large and likely include all natural possibilities for  $\mathcal{T}_0$ : two-qubit thermal operations and two-qubit Gibbs-preserving maps.

For degenerate Hamiltonians, the choice of  $\mathcal{T}$  as the set of two-qubit unitaries does not suffer the same problems as those that apply to sets of unitary operations for nondegenerate Hamiltonians. Alternative  $\mathcal{T}$  choices include the sets of all the two-qubit noisy operations and of all the two-qubit unital maps. These operations are the two-qubit thermal operations and Gibbs-preserving maps, respectively, if the qubits have degenerate Hamiltonians. However, the two-qubit noisy operations and unital maps likely include operations that demand unreasonable control requirements, making the resource-cost lower bounds loose.

## B. One-shot entropy measures in quantum thermodynamics

A fundamental connection between thermodynamics and statistical mechanics is the identification, under suitable conditions, of Clausius' thermodynamic entropy with the von Neumann entropy  $H(\rho)$  of a quantum state  $\rho$ ,

$$H(\rho) := -\text{tr}(\rho \log \rho). \quad (5)$$

In the information-theoretic approach to thermodynamics, beyond the traditional regime concerning many copies of a system, many thermodynamic tasks have work costs inaccurately represented by the von Neumann entropy and the standard free energy [43,59,60,70–77]. Instead, these work costs are quantified with *one-shot entropy measures*, such as the relative entropies [78–80] and Rényi- $\alpha$  entropies [44], including the min- and max-entropies.

We focus on the *hypothesis-testing relative entropy*, which interpolates between the min- and max-relative entropies [81–87]. The hypothesis-testing relative entropy

is defined for a quantum state  $\rho$ , a positive-semidefinite operator  $\Gamma$ , and an intolerance parameter  $\eta \in (0, 1]$ :

$$D_H^\eta(\rho \parallel \Gamma) := -\log \left( \min_{\substack{0 \leq Q \leq 1 \\ \text{tr}(Q\rho) \geq \eta}} \left\{ \frac{\text{tr}(Q\Gamma)}{\eta} \right\} \right). \quad (6)$$

The *hypothesis-testing entropy* of a quantum state  $\rho$ , for  $\eta \in (0, 1]$ , is defined as

$$H_H^\eta(\rho) := -D_H^\eta(\rho \parallel \mathbb{1}). \quad (7)$$

The entropy measures here have units of nats, rather than bits, because our definitions have base- $e$ , rather than base-2, logarithms. One can convert between bits and nats via (no. of bits) = (no. of nats)/log(2). Our convention introduces factors of log(2) in our results for qubit systems. However, the convention yields familiar forms for thermodynamic relations between entropy and quantities such as the free energy.

The hypothesis-testing entropy quantifies how well one can distinguish between quantum states  $\rho$  and  $\sigma$  via a hypothesis test. Suppose that we receive either  $\rho$  or  $\sigma$ . We must guess which state we obtained, based on the outcome of one measurement. We may choose the measurement, specified by a two-outcome positive-operator-valued measure (POVM)  $\{Q, \mathbb{1} - Q\}$ . The outcome  $Q$  implies that we should guess “ $\rho$ ”; and  $\mathbb{1} - Q$ , that we should guess “ $\sigma$ .” Suppose that, when  $\rho$  is provided, the measurement must identify  $\rho$  correctly with a probability  $\geq \eta$ :  $\text{tr}(Q\rho) \geq \eta$ . When  $\sigma$  is provided, the optimal probability  $\text{tr}(Q\sigma)$  of failing to identify  $\sigma$  is  $\eta e^{-D_H^\eta(\rho \parallel \sigma)}$ . The interpretation of Eq. (6) as a relative-entropy measure arises from certain properties. For example, Eq. (6) is non-negative and obeys a data-processing inequality [82]; Eq. (6) approximates the min- and max-relative entropies when  $\eta \approx 0$  and  $\eta \approx 1$  [82], respectively; and Eq. (6) quantifies the resource costs of communication and thermodynamic tasks [88]. Appendix C contains a more detailed discussion of the hypothesis-testing relative entropy and its properties.

The hypothesis-testing entropy and relative entropy have operational significances in quantum thermodynamics [74,89,90]. These quantities unify quantum thermodynamic results based on the min- and max- (relative) entropies [40,43,46,59,91]: Consider resetting a state  $\rho$  on a memory register  $S$  to a fixed state  $|0\rangle_S$ . Suppose that  $\rho$  evolves under a completely degenerate Hamiltonian  $H_S = 0$  and that the resetting must fail with a probability at most  $\varepsilon \geq 0$ . Absent restrictions on the resetting's complexity, the resetting has a work cost [43,82]

$$W^\varepsilon(\rho_S \rightarrow |0\rangle_S) \approx H_H^{1-\varepsilon}(\rho) \times k_B T. \quad (8)$$

The approximation conceals technical details about how the failure probability is measured. If  $\rho_S = \mathbb{1}_2/2$  is

the single-qubit maximally mixed state, then  $H_{\text{H}}^{1-\varepsilon}(\rho) = \log(2)$ , and we obtain the well-known formula for Landauer erasure [34]

$$W^{\varepsilon}(\mathbb{1}_2/2 \rightarrow |0\rangle_{\text{S}}) \approx k_{\text{B}}T \log(2). \quad (9)$$

### C. Complexity (relative) entropy

The complexity entropy was introduced in the context of the resource theory of uncomplexity [30]. There, a state's complexity entropy quantifies the qubits in the state  $|0\rangle$  extractable via a limited number of unitary gates. In this work, we apply the complexity entropy to quantum information thermodynamics and detail the entropy's properties. We present a version of the complexity entropy that is tailored to  $n$ -qubit quantum circuits wherein each gate is in the set  $\mathcal{T}$  of elementary computations. We introduce the complexity entropy's general form in Appendix D, where we prove general bounds and monotonicity results.

A key motivation for defining the complexity entropy is the following. Consider a hypothesis test between states  $\rho$  and  $\sigma$ . The hypothesis-testing relative entropy quantifies the optimal probability of wrongly rejecting  $\sigma$  with any strategy that correctly accepts  $\rho$  with a probability  $\geq \eta$ . However, an optimal measurement  $\{Q, \mathbb{1} - Q\}$  might be too complex to be executed in a reasonable time. Suppose, for instance, that  $\rho = |\psi\rangle\langle\psi|$  is a highly complex, pure  $n$ -qubit state and that  $\sigma = \mathbb{1}_2^{\otimes n}/2^n$  is maximally mixed. A measurement that distinguishes  $\rho$  from  $\sigma$  may require a complex circuit implementable only in an exponentially long time [92,93]. It is natural to restrict  $Q$  to be implementable in a reasonable time, with a circuit composed of  $\leq r$  gates. The *complexity relative entropy* is defined similarly to the hypothesis-testing relative entropy. The former, however, has a complexity restriction in the optimization over measurement operators.

To define the complexity entropy, we must specify the set of POVM effects that a computationally limited observer can render. Define  $\mathbf{M}^r$  as the set of POVM effects that one can implement by performing  $\leq r$  gates and then applying certain single-qubit projectors. All tensor products of those projectors constitute the set of complexity-zero POVM effects

$$\mathbf{M}^0 := \left\{ \bigotimes_{i=1}^n Q_i : Q_i \in \{|0\rangle\langle 0|, \mathbb{1}_2\} \right\}. \quad (10)$$

For each  $i$ , an effect in  $\mathbf{M}^0$  projects qubit  $i$  onto  $|0\rangle$  or does nothing. Fix any set  $\mathcal{G}$  of elementary quantum operations (completely positive, trace-nonincreasing maps). (We define complexity-restricted POVM elements in terms of  $\mathcal{G}$ , as an information-theoretic quantity independent of the thermodynamic framework in Sec. II A. Hence  $\mathcal{G}$  is independent of the operations introduced there. If applying the POVM elements in that framework, however, one can

choose for  $\mathcal{G}$  to consist of operations introduced in Sec. II A.) For  $r > 0$ , we define

$$\mathbf{M}^r := \left\{ \mathcal{E}_1^\dagger \cdots \mathcal{E}_r^\dagger(P) : P \in \mathbf{M}^0, \mathcal{E}_i \in \mathcal{G} \right\}. \quad (11)$$

One natural choice for  $\mathcal{G}$  might be the set of two-qubit unitary gates. Another choice, in the context of complexity and thermodynamics, is  $\mathcal{G} = \mathcal{T}$ , wherein  $\mathcal{T}$  denotes the set of elementary computations defined in Sec. II A.

Even if a  $Q$  belongs to  $\mathbf{M}^r$ , the complementary effect  $\mathbb{1} - Q$  might not. Indeed, the definition of  $\mathbf{M}^r$  applies in settings where only one POVM effect is relevant in a hypothetical measurement. For instance, suppose we wish to certify that some process outputs a state close to some pure state  $|\psi\rangle$ . One may consider a hypothetical measurement of the output state with a POVM containing the effect  $|\psi\rangle\langle\psi|$ , and ascertain that its outcome probability is close to unity. In such a scenario, the POVM effect need not be implemented in practice, and other effects that complete the POVM may be ignored.

Having introduced complexity-restricted POVM effects, we can define complexity-restricted entropic quantities. Let  $\rho$  denote any quantum state; and  $\Gamma$ , any positive-semidefinite operator. The *complexity relative entropy* of  $\rho$  relative to  $\Gamma$ , at a complexity scale  $r \geq 0$  with respect to  $\mathbf{M}^r$ , and for  $\eta \in (0, 1]$ , is

$$D_{\text{H}}^{r,\eta}(\rho \parallel \Gamma) := -\log \left( \inf_{\substack{Q \in \mathbf{M}^r \\ \text{tr}(Q\rho) \geq \eta}} \left\{ \frac{\text{tr}(Q\Gamma)}{\text{tr}(Q\rho)} \right\} \right). \quad (12)$$

This definition extends the hypothesis-testing relative entropy (6) by restricting the optimization to POVM effects implementable with  $\leq r$  gates. (See Appendix D 1 for a more general definition and for details.) The definition (12) mirrors a construction, based on the trace norm, in Ref. [17].

The complexity relative entropy enjoys an operational interpretation similar to that of the hypothesis-testing relative entropy (6). Consider a hypothesis test between states  $\rho$  and  $\sigma$ . We identify  $\rho$  as the *null hypothesis* and  $\sigma$  as the *alternative hypothesis*. Suppose that one can measure only a POVM  $\{Q, \mathbb{1} - Q\}$  for which  $Q \in \mathbf{M}^r$ . It is useful to allow, apart from a measurement, one toss of a biased classical coin. (We explain why below.) One can freely choose the coin's probability  $q \in (0, 1]$  of landing heads up. Consider guessing " $\rho$ " if the coin shows "heads" and the POVM outcome is  $Q$ , guessing " $\sigma$ " otherwise. This strategy is equivalent to measuring the POVM  $\{qQ, \mathbb{1} - qQ\}$ . A type-I error—incorrectly rejecting the null hypothesis—occurs with a probability  $1 - \text{tr}(qQ\rho)$ . A type-II error—incorrectly rejecting the alternative hypothesis—occurs with a probability  $\text{tr}(qQ\sigma)$ . The following proposition shows how the complexity relative

entropy characterizes hypothesis testing with complexity limitations.

*Proposition 1 (Hypothesis testing with complexity limitations).* Let  $\rho$  and  $\sigma$  denote quantum states. Let  $\eta \in (0, 1]$  and  $\delta \in (0, 1]$ . There exists a  $Q \in \mathcal{M}^r$  and  $q \in (0, 1]$  such that  $\text{tr}[(qQ)\rho] = \eta$  and  $\text{tr}[(qQ)\sigma] \leq \delta$  if and only if

$$D_{\text{H}}^{r,\eta}(\rho \parallel \sigma) \geq -\log(\delta/\eta). \quad (13)$$

The proposition guarantees that, if Eq. (13) holds, then, for any parameters  $\eta, \delta \in (0, 1]$ , some hypothesis test of the sort just described has two properties: a type-I error occurs with a probability  $\leq 1 - \eta$ , and a type-II error occurs with a probability  $\leq \delta$ . One can satisfy Eq. (13) only if  $\delta \leq \eta$ , since the complexity relative entropy is non-negative. The coin toss's inclusion in the hypothesis test enables a precise relationship between the complexity relative entropy and hypothesis testing, for  $\eta$  values far from 1. If  $\eta \approx 1$ , then  $q \approx 1$ , and the agent need not toss the coin. We prove the proposition in Appendix D 3.

Having introduced the complexity relative entropy, we now define the complexity entropy. Let  $\rho$  denote any quantum state. The *complexity entropy* of  $\rho$ , at a complexity scale  $r \geq 0$  with respect to  $\mathcal{M}^r$ , and for  $\eta \in (0, 1]$ , is

$$H_{\text{H}}^{r,\eta}(\rho) := -D_{\text{H}}^{r,\eta}(\rho \parallel \mathbb{1}). \quad (14)$$

This definition mirrors Eq. (7). We follow a standard procedure for defining an entropy using a relative entropy [94]. We show that the complexity entropy is always non-negative [ $H_{\text{H}}^{r,\eta}(\rho) \geq 0$ ]; see Appendix D 2 for a proof.

The normalization factor  $\text{tr}(Q\rho)$  in Eq. (12) guarantees elementary properties of the complexity entropy, such as its having the range  $[0, n \log(2)]$  for all  $\eta$ . In Appendix D, we define an alternative version of the complexity (relative) entropy without the normalization factor. This alternative is technically more convenient, and we bound one version in terms of the other.

### III. THERMODYNAMIC INFORMATION ERASURE WITH COMPLEXITY CONSTRAINTS

We now turn to the erasure of the  $n$ -qubit memory register introduced in Sec. II. What is the optimal work cost of resetting a state  $\rho$  to  $|0^n\rangle$ , using the primitive operations (i)–(iii)?

#### A. Erasure of qubits with a completely degenerate Hamiltonian

We first consider a simpler case: each memory qubit has a degenerate Hamiltonian ( $H_i = 0$  for all  $i$ ). Let the elementary computations  $\mathcal{T}$  be the two-qubit unitary gates with arbitrary connectivities:  $\mathcal{T} \simeq \bigcup_{(i,j)} \text{SU}(4)_{ij}$ .

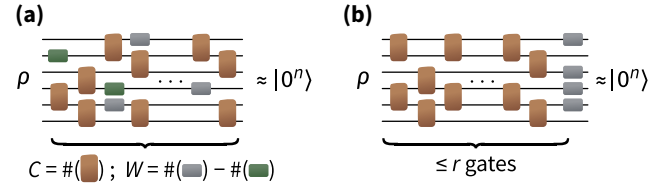


FIG. 4. Protocols for thermodynamic information erasure with complexity constraints. The process consists of RESET gates (gray circuit elements), EXTRACT gates (green), and computational gates (brown) [cf. primitives (i)–(iii) of Fig. 1]. (a) A general protocol interleaves the primitive operations (i)–(iii), preparing a state close to  $|0^n\rangle$ . The protocol's complexity cost  $C$  and work cost  $W$  follow from summing the primitive operations' costs (Fig. 1). (b) In the simplified protocol, one applies  $\leq r$  computational gates [primitive (iii)]. Then, one erases any qubits far from  $|0\rangle$ , using RESET operations [primitive (i)]. The circuit compresses  $\rho$  into fewer qubits, to be erased by RESET operations. Such protocols extract pure  $|0\rangle$  qubits in the resource theory of uncomplexity [30].

The mixed-state fidelity between states  $\sigma$  and  $\tau$  is defined as [95]

$$F(\sigma, \tau) := \text{tr}\left(\sqrt{\sigma^{1/2}\tau\sigma^{1/2}}\right). \quad (15)$$

We define an erasure of  $\rho$  as any composition of the operations (i)–(iii) that transforms  $\rho$  into a state  $\rho'$  satisfying  $F^2(\rho', |0^n\rangle\langle 0^n|) = \langle 0^n | \rho' | 0^n \rangle \geq \eta$ , for some fixed  $\eta \in (0, 1]$  [Fig. 4(a)]. One can interpret  $F^2(\rho', |0^n\rangle\langle 0^n|)$  as the probability of preparing  $|0^n\rangle$ .

We now quantify the trade-off between the work cost and complexity cost of an erasure protocol  $\mathcal{E}$ . Being a process,  $\mathcal{E}$  is a sequence of primitive operations. Yet, we denote by  $\mathcal{E}$  also the operation implemented by the sequence, in a slight abuse of notation. Suppose that  $\mathcal{E}$  has a complexity at most  $r \geq 0$ :  $C(\mathcal{E}) \leq r$ . The optimal work cost for such a protocol is

$$W_r := \min_{\mathcal{E}} \{W(\mathcal{E}) : C(\mathcal{E}) \leq r, F^2(\mathcal{E}(\rho), |0^n\rangle\langle 0^n|) \geq \eta\}. \quad (16)$$

#### 1. Protocols whose RESET operations happen at the end

As an initial step, we consider an erasure protocol divided into two parts [Fig. 4(b)]. First, the state undergoes  $\leq r$  computational gates [primitive (iii)]. Then, RESET operations [primitive (i)] are applied. Such a protocol's optimal work cost is

$$W_r^* = \min_{\mathcal{E}} \left\{ W(\mathcal{E}) : \mathcal{E} = \left( \prod_{i \in \mathcal{W}} \mathcal{E}_{\text{RESET}, i} \right) \mathcal{E}_r \cdots \mathcal{E}_1, \right. \\ \left. \mathcal{W} \subset \{1, 2, \dots, n\}, F^2(\mathcal{E}(\rho), |0^n\rangle\langle 0^n|) \geq \eta \right\}. \quad (17)$$



Here  $\mathcal{W}$  denotes the subset of the qubits that undergo RESET operations.

We now identify how to erase  $\rho$  using the least amount of work. We should compress  $\rho$  into as few qubits as possible, using  $\leq r$  gates, to apply as few RESET operations as possible. In other words, we should perform data compression with  $\leq r$  gates. This information-theoretic interpretation of erasure mirrors two earlier interpretations of erasures: the erasure of  $n$  qubits without complexity restrictions [59] and the erasure of a quantum system using a quantum memory [60].

We analyzed a version of this compression task in the context of extracting pure qubits in the resource theory of uncomplexity [30]. A variant of the complexity entropy, we showed, quantifies the qubits that cannot be reset to  $|0\rangle$  states with  $\leq r$  gates. In this subsection, we adapt that argument to thermodynamic erasure. (We discuss further applications of the complexity entropy to information-theoretic tasks in Sec. V.)

Let  $\mathcal{E}$  denote any protocol that achieves the minimum in Eq. (17):  $W(\mathcal{E}) = W_r^*$ . Assume that the work cost  $W(\mathcal{E})$  corresponds to  $\beta W(\mathcal{E}) / \log(2) = |\mathcal{W}| =: w$  bits. Define the projector  $P := 1_{\mathcal{W}^c} \otimes |0^{n-w}\rangle\langle 0^{n-w}|_{\mathcal{W}^c}$ , wherein  $\mathcal{W}^c$  denotes the complement of  $\mathcal{W}$ . Define the POVM effect

$$Q := \mathcal{E}_1^\dagger \cdots \mathcal{E}_r^\dagger(P), \quad (18)$$

wherein each  $\mathcal{E}_i^\dagger$  denotes the adjoint of the operation  $\mathcal{E}_i$  [defined via  $\text{tr}(\mathcal{E}_i^\dagger(A)B) = \text{tr}(A\mathcal{E}_i(B))$  for all operators  $A$  and  $B$ ].  $P$  projects onto  $|0\rangle$  each qubit not subject to any RESET operation, leaving all other qubits untouched. We choose for the set  $\mathcal{G}$  of elementary operations, used to define  $\mathcal{M}$  in Eq. (11), to be the set  $\mathcal{T}$  of elementary computations, used to define our model in Sec. II A.  $Q$  belongs to  $\mathcal{M}$ . Furthermore,

$$\begin{aligned} \text{tr}(Q\rho) &= \text{tr}(P\mathcal{E}_r \cdots \mathcal{E}_1(\rho)) \\ &= \text{tr}\{|0^{n-w}\rangle\langle 0^{n-w}| \text{tr}_{\mathcal{W}}(\mathcal{E}_r \cdots \mathcal{E}_1(\rho))\} \\ &= \text{tr}(|0^n\rangle\langle 0^n| \mathcal{E}(\rho)) = F^2(\mathcal{E}(\rho), |0^n\rangle\langle 0^n|) \geq \eta, \end{aligned} \quad (19)$$

so  $Q$  is a candidate for the optimization (12) defining the complexity entropy  $H_H^{r,\eta}(\rho) = -D_H^{r,\eta}(\rho \| 1)$ . Moreover, since each  $\mathcal{E}_i$  is unitary,  $\text{tr}(Q) = \text{tr}(P\mathcal{E}_r \cdots \mathcal{E}_1(1)) = \text{tr}(P) = 2^w$ . Hence,

$$\begin{aligned} H_H^{r,\eta}(\rho) - \log(1/\eta) &\leq \log(\text{tr}(Q)/\text{tr}(Q\rho)) - \log(1/\eta) \\ &\leq \log(\text{tr}(Q)) = w \log(2) = \beta W(\mathcal{E}) = \beta W_r^*. \end{aligned} \quad (20)$$

A similar inequality points in the opposite direction, as we show by reversing the steps above. Consider any candidate  $Q \in \mathcal{M}$  for the optimization (12) defining  $H_H^{r,\eta}(\rho) =$

$-D_H^{r,\eta}(\rho \| 1)$ . Let  $Q$  be optimal (or arbitrarily close to optimal). By the definition (11),  $Q = \mathcal{E}_1^\dagger \cdots \mathcal{E}_r^\dagger(P)$ , wherein each  $\mathcal{E}_i \in \mathcal{G}$  and wherein  $P$  projects onto  $|0\rangle$  each qubit in some subset  $\mathcal{W}^c$ , leaving all other qubits untouched. Let  $\mathcal{E} := (\prod_{i \in \mathcal{W}} \mathcal{E}_{\text{RESET},i}) \mathcal{E}_r \cdots \mathcal{E}_1$ .  $\mathcal{E}$  erases  $\rho$ , since

$$F^2(\mathcal{E}(\rho), |0^n\rangle\langle 0^n|) = \text{tr}(|0^n\rangle\langle 0^n| \mathcal{E}(\rho)) = \text{tr}(Q\rho) \geq \eta. \quad (21)$$

Hence,

$$\begin{aligned} \beta W_r^* &\leq \beta W(\mathcal{E}) = |\mathcal{W}| \log(2) \\ &= \log(\text{tr}(Q)) \leq \log(\text{tr}(Q)/\text{tr}(Q\rho)) = H_H^{r,\eta}(\rho). \end{aligned} \quad (22)$$

The last equality holds if  $Q$  is optimal. (If  $Q$  is arbitrarily close to optimal, then the last two quantities are arbitrarily close to each other.) We have therefore proved the following theorem.

*Theorem 1 (Complexity-limited erasure with restricted protocols).* Consider erasing an  $n$ -qubit system governed by a fully degenerate Hamiltonian. Every optimal complexity-limited protocol that uses only primitive (iii) (wherein  $\mathcal{T} = \mathcal{G}$  consists of unitary operations) followed by primitive (i), expends an amount  $W_r^*$  of work that obeys

$$H_H^{r,\eta}(\rho) - \log(1/\eta) \leq \beta W_r^* \leq H_H^{r,\eta}(\rho). \quad (23)$$

Two points merit mentioning. First, one might wonder whether the work cost  $\beta^{-1}[H_H^{r,\eta}(\rho) - \log(1/\eta)]$  is achievable. The answer depends on whether one can implement, with gates in  $\mathcal{G}$ , a POVM effect  $Q \in \mathcal{M}$  that satisfies  $\text{tr}(Q\rho) = \eta$  and is arbitrarily close to optimal for  $H_H^{r,\eta}(\rho)$ . In such a case, the first inequality in Eq. (23) saturates:  $H_H^{r,\eta}(\rho) - \log(1/\eta)$  equals a variant of the complexity entropy, and that variant, we show, equals  $\beta W_r^*$ . (See Appendix D 5 and E for details.) Considering the variant's definition, one can interpret the  $\log(1/\eta)$  in Eq. (23) as a particularity of how we defined the complexity entropy.

Second, the  $\log(1/\eta)$  is proportional to the work won in a successful bet on an event that occurs with a probability  $\eta$  [96]. We can understand this point through a simple example. Consider resetting the single-qubit maximally mixed state  $\rho = 1_2/2$  to  $|0\rangle$ . Suppose that our target success probability  $\eta$  satisfies  $\eta \leq 1/2$ . One successful erasure protocol does nothing: with the probability  $1/2$ ,  $|0\rangle$  is prepared. Yet standard measures of entropy (including the Rényi entropies of one-shot information theory [78]) attribute to  $\rho$  one bit of entropy. The zero work cost of erasing a maximally mixed qubit with  $\eta = 1/2$  can be understood as a sum of (i) one unit of work expended to reduce the entropy of  $\rho$  and (ii) one unit of work extracted by a successful bet with the success probability  $1/2$ .

We now present a few textbook quantum states and the work costs of erasing them, in various regimes of  $r$  and  $\eta$ .

*a. Computational-basis states.* If  $\rho = |0^n\rangle\langle 0^n|$ , then  $H_H^{r,\eta}(\rho) = 0$  for all  $r$  and  $\eta$ . An optimal protocol is to do nothing—apply no RESET operations and no work:  $W_r^* = 0$ .

If  $\rho = |1^n\rangle\langle 1^n|$ , then one resets  $\rho$  to  $|0^n\rangle$  by flipping each qubit to  $|0\rangle$ , incurring a minor complexity cost (minor compared to the exponential complexities expected of most  $n$ -qubit unitaries [4]). Assume that  $n$  is even, for simplicity. One can flip all  $n$  qubits with one layer of  $n/2$  two-qubit gates. Therefore, whenever  $r \geq n/2$ ,  $H_H^{r,\eta}(\rho) = 0$ , and  $W_r^* = 0$ , for all  $\eta$ . If, however,  $r < n/2$ , then one can flip only  $2r < n$  qubits. One must apply RESET operations to the other  $n - 2r$  qubits. In this case,  $H_H^{r,\eta}(\rho) = (n - 2r) \log(2)$ , and  $W_r^* = (n - 2r) k_B T \log(2)$ , for all  $\eta$ .

*b. The maximally mixed state.* If  $\rho$  is maximally mixed, then  $H_H^{r,\eta}(\rho) = n \log(2)$  for all  $r$  and  $\eta$ . Suppose that the error tolerance is insignificant:  $\eta \approx 1$ . An optimal protocol requires  $n$  RESET operations, costing the maximum amount of work:  $W_r^* = n k_B T \log(2)$ .

*c. Greenberger-Horne-Zeilinger (GHZ) state.* Let

$$|\text{GHZ}\rangle := \frac{1}{\sqrt{2}}(|0^n\rangle + |1^n\rangle), \quad (24)$$

and let  $\rho = |\text{GHZ}\rangle\langle \text{GHZ}|$ . One can prepare the GHZ state with a Hadamard gate followed by  $n - 1$  CNOT gates. Therefore,  $H_H^{r,\eta}(\rho) = 0$  if  $r \geq n$ . If  $r < n$ , then one can apply  $r$  CNOT gates to disentangle  $r$  qubits from the other  $n - r$  qubits. The  $n - r$  qubits require RESET operations. Absent any significant error tolerance (if  $\eta \approx 1$ ),  $H_H^{r,\eta}(\rho) = (n - r) \log(2)$ , and  $W_r^* = (n - r) k_B T \log(2)$ .

*d. A Haar-random state.* Let  $|\psi\rangle$  denote a state chosen randomly according to the Haar measure on the pure  $n$ -qubit states. The state  $\rho = |\psi\rangle\langle \psi|$  is indistinguishable from the maximally mixed state, according to  $\text{poly}(n)$ -complex observables [92,93]. Therefore, if  $r \lesssim \text{poly}(n)$ , one cannot distinguish  $\rho$  from the maximally mixed state using  $\leq r$  gates. All  $n$  qubits must undergo RESET operations. Absent any significant error tolerance (if  $\eta \approx 1$ ),  $H_H^{r,\eta}(\rho) = n \log(2)$ , and  $W_r^* = n k_B T \log(2)$ .

This example constitutes a special case of Proposition 3 in Sec. VD. There, we prove a lower bound on the complexity entropy of a state generated by a random circuit. Random circuits effect Haar-random unitaries in the large-circuit-depth limit.

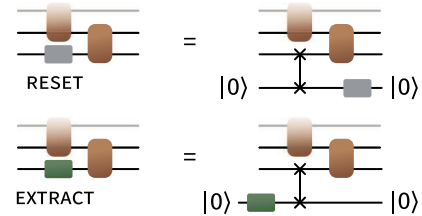
*e. Mixture of different-complexity states.* Let  $\rho$  denote a convex mixture of  $|0^n\rangle\langle 0^n|$  and a high-complexity state  $|\psi\rangle\langle \psi|$ :  $\rho = (1 - \varepsilon)|0^n\rangle\langle 0^n| + \varepsilon|\psi\rangle\langle \psi|$ , wherein  $\varepsilon \in$

$[0, 1 - \varepsilon]$ . The POVM effect  $Q = |0^n\rangle\langle 0^n|$  is a candidate for the optimization (12) defining  $H_H^{r,\eta}(\rho) = -D_H^{r,\eta}(\rho \parallel \mathbb{1})$ . Therefore,  $H_H^{r,\eta}(\rho) \leq -\log(1 - \varepsilon)$ , and  $W_r^* \leq -k_B T \log(1 - \varepsilon) \approx \varepsilon k_B T$  (the approximation holds for small  $\varepsilon$ ).

## 2. General protocols with midcircuit RESET and EXTRACT operations

Suppose that each qubit's Hamiltonian is degenerate. Suppose further that a process can be any sequence of the primitives (i)–(iii) [Fig. 4(a)]. A process may interleave elementary computations with RESET and EXTRACT operations on the same qubit. These midcircuit nonunitary operations can reduce the complexity of information erasure. The significance of midcircuit measurements in *monitored circuits* has only recently been appreciated, both in quantum information and in condensed-matter theory. Experimentalists have recently used midcircuit measurements in quantum error correction and in *monitored circuits* [97–99]. Along similar lines, quantum phases of matter driven by measurements have been explored [100–104]. We seek to bound  $W_r$ , defined in Eq. (16). Our strategy is to map a general erasure protocol to a different protocol involving additional auxiliary systems and whose reset operations all happen at the end. The work cost of the general protocol is then bounded using an argument adapted from the one in Sec. III A 1.

We must transform a protocol  $\mathcal{E}$  into a POVM effect. Our strategy involves ancillary qubits, each initialized to  $|0\rangle$ . Every midcircuit RESET operation on qubit  $i$  is performable with a final RESET operation: we swap qubit  $i$  with an ancilla, then RESET the ancilla. Similarly, every midcircuit EXTRACT operation on qubit  $i$  is performable with an initial EXTRACT operation: we perform an EXTRACT operation on an ancilla, then swap the ancilla with qubit  $i$ . In both processes, every ancilla begins and ends in  $|0\rangle$  (recall that the EXTRACT operation is performable only on a qubit in  $|0\rangle$ ). We assume that the SWAP gate belongs to  $\mathcal{T}$ .



Consider any candidate protocol  $\mathcal{E}$  for the optimization (16). Suppose that  $\mathcal{E}$  consists of  $m_1$  RESET operations,  $m_2$  EXTRACT operations, and  $\leq r$  gates.  $\mathcal{E}$  has the complexity cost  $C(\mathcal{E}) \leq r$  and the work cost  $W(\mathcal{E}) = (m_1 - m_2) k_B T \log(2)$ . Let us transfer all the RESET and EXTRACT operations in  $\mathcal{E}$  to ancillas. We obtain a protocol  $\mathcal{E}'$  of  $\leq r + m_1 + m_2$  gates (each SWAP contributes one gate),

preceded by EXTRACT operations and followed by RESET operations.  $\mathcal{E}'$  acts on  $n + m_1 + m_2$  qubits and has the complexity cost  $C(\mathcal{E}') \leq r + m_1 + m_2$ . Immediately after the initial EXTRACT operations, the input state is

$$\tilde{\rho}_{m_1, m_2} := \rho \otimes |0^{m_1}\rangle\langle 0^{m_1}| \otimes \mathbb{1}_2^{\otimes m_2} / 2^{m_2}. \quad (25)$$

Subsequently,  $\mathcal{E}'$  outputs a state  $\mathcal{E}'(\tilde{\rho}) =: \tilde{\rho}'$  satisfying

$$F^2(\text{tr}_{m_1}(\tilde{\rho}'), |0^{n+m_2}\rangle\langle 0^{n+m_2}|) \geq \eta. \quad (26)$$

We now adapt the argument used for protocols whose RESET operations happen at the end. Suppose the set  $\mathcal{G}$  of elementary operations, used to define  $M^r$  in Eq. (11), equals the set of elementary computations:  $\mathcal{G} = \mathcal{T}$ . Define the POVM effect

$$Q := \mathcal{E}'^\dagger(|0^n\rangle\langle 0^n| \otimes \mathbb{1}_2^{\otimes m_1} \otimes |0^{m_2}\rangle\langle 0^{m_2}|). \quad (27)$$

Then

$$\text{tr}(Q\tilde{\rho}) = \text{tr}(|0^{n+m_2}\rangle\langle 0^{n+m_2}| \text{tr}_{m_1}(\tilde{\rho}')) \geq \eta. \quad (28)$$

Hence,  $Q$  is a candidate for the optimization (12) defining  $H_H^{r+m_1+m_2, \eta}(\tilde{\rho}) = -D_H^{r+m_1+m_2, \eta}(\tilde{\rho} \| \mathbb{1})$ . Therefore,

$$\begin{aligned} H_H^{r+m_1+m_2, \eta}(\tilde{\rho}) &\leq \log(\text{tr}(Q) / \text{tr}(Q\tilde{\rho})) \\ &\leq m_1 \log(2) + \log(1/\eta), \end{aligned} \quad (29)$$

so

$$\beta W(\mathcal{E}) \geq H_H^{r+m_1+m_2, \eta}(\tilde{\rho}) - m_2 \log(2) - \log(1/\eta). \quad (30)$$

We use the shorthand notation

$$g^{r, \eta}(\rho) := \inf_{m_1, m_2 \geq 0} \left\{ H_H^{r+m_1+m_2, \eta}(\tilde{\rho}_{m_1, m_2}) - m_2 \log(2) \right\} \quad (31)$$

and take the infimum of Eq. (30) over all protocols  $\mathcal{E}$ . Combining the previous two equations yields

$$\beta W_r \geq g^{r, \eta}(\rho) - \log(1/\eta). \quad (32)$$

The reverse direction might not hold, generally. For any  $m_1, m_2 \geq 0$ , Eq. (23) guarantees the existence of a protocol  $\mathcal{E}$  that has two properties. First,  $\mathcal{E}$  consists of  $\leq r + m_1 + m_2$  computational gates followed by  $w$  RESET operations. Second,  $\mathcal{E}$  maps  $\tilde{\rho}$  approximately to  $|0^{n+m_1+m_2}\rangle$ , at the work cost  $w \log(2) \leq H_H^{r+m_1+m_2, \eta}(\tilde{\rho})$ . Yet, such a protocol may involve the  $m_1 + m_2$  ancillas in a computation inequivalent to any computation on  $n$  qubits, even if the latter computation includes midcircuit RESET and EXTRACT operations.

Overall, we have bounded the optimal work cost  $W_r$  of erasing  $\rho$  using  $\leq r$  computational gates.

*Theorem 2 (Bound of optimal work cost).*  $W_r$  obeys

$$g^{r, \eta}(\rho) - \log(1/\eta) \leq \beta W_r \leq \beta W_r^* \leq H_H^{r, \eta}(\rho). \quad (33)$$

A large gap may separate  $W_r$  and  $W_r^*$  if, with few pure ancillas, one can uncompute a pure state  $|\chi\rangle$  to  $|0^n\rangle$ , using a circuit much shorter than is possible without ancillas. In other words, an erasure protocol may benefit from early RESET operations: the resulting pure qubits may be used as ancillas for uncomputing the remaining state. For instance, there might exist an  $n$ -qubit state  $|\chi\rangle$ , and an  $\ell \ll n$ , with the following two properties. First,  $|\chi\rangle \otimes |0^\ell\rangle$  is much less complex than  $|\chi\rangle$ :  $C(|\chi\rangle \otimes |0^\ell\rangle) \ll C(|\chi\rangle)$ . Second, in the absence of ancillas,  $r$  gates fail to extract any  $|0\rangle$ 's from  $|\chi\rangle$ , if  $r = C(|\chi\rangle \otimes |0^\ell\rangle)$ . Midcircuit RESET operations would be able to lower the work cost of erasing  $|\chi\rangle \otimes |0^\ell\rangle$  with at most  $r = C(|\chi\rangle \otimes |0^\ell\rangle)$  gates, we now show. The following protocol would employ midcircuit RESET operations and cost work  $k_B T \ell \log(2)$ : RESET the  $\ell$  mixed qubits, paying  $\ell$  units of work. Then, uncompute the remaining state,  $|\chi\rangle \otimes |0^\ell\rangle$ , to  $|0^{n+\ell}\rangle$ , using  $r$  gates. The total work cost would be  $\beta^{-1} \ell \log(2)$ . In contrast, consider a protocol whose RESET operations happen at the end.  $r$  operations would not extract any  $|0\rangle$ 's. Such a protocol would require  $n + \ell$  units of work:  $\beta W_r^* = (n + \ell) \log(2) \gg \beta W_r$ .

## B. Erasure of qubits with a general product Hamiltonian

We now consider a more general setup: each qubit has a not-necessarily-degenerate Hamiltonian  $H_i$ , as per Sec. II. The set  $\mathcal{T}$  of elementary computations can no longer be the set of all two-qubit unitary gates: some gates would require work.

We require that  $\mathcal{T}$  contain only completely positive, trace-preserving maps  $\mathcal{E}$  that satisfy the technical property

$$\mathcal{E}(\Gamma) = \Gamma, \quad \text{wherein } \Gamma = e^{-\beta \sum_i H_i} = \bigotimes_{i=1}^n \Gamma_i. \quad (34)$$

The operators  $\Gamma_i$  are defined in Eq. (1). Examples of such operations include the two-qubit thermal operations and the two-qubit Gibbs-preserving maps (Sec. II). We assume Eq. (34) to ensure a thermodynamically consistent accounting of work costs: elementary computations should cost no work.

For concreteness, we describe another class of operations that satisfy condition (34). These operations can model computations on systems governed by product Hamiltonians, as well as crude control over interactions

with a heat bath. Consider an operation, on qubits  $i$  and  $j$ , of the form

$$\rho \mapsto \mathcal{E}(\rho) = (1 - q)U_{ij}\rho U_{ij}^\dagger + q\gamma_i \otimes \gamma_j. \quad (35)$$

$q$  denotes a probability.  $U_{ij}$  denotes an energy-conserving unitary:  $[U_{ij}, H_i + H_j] = 0$ .  $\gamma_i$  ( $\gamma_j$ ) denotes the thermal state of qubit  $i$  ( $j$ ), defined in Eq. (1). With a probability  $q$ ,  $\mathcal{E}$  replaces its input with a thermal state. Otherwise,  $\mathcal{E}$  effects an energy-conserving unitary.  $\mathcal{E}$  can model an interaction between qubits and a heat bath over short times, during which the bath partially thermalizes the qubits. This interaction may be exploited to implement operations that do not strictly conserve energy, i.e., that do not commute with the total Hamiltonian. Subjecting  $\gamma_i \otimes \gamma_j$  to any operation (35) shows directly that the operation obeys the condition (34).

Consider any protocol  $\mathcal{E}$  in  $\mathcal{T}$  that has two properties: first,  $\mathcal{E}$  consists of  $\leq r$  computational gates followed by RESET operations. Second,  $\mathcal{E}$  satisfies  $F^2(\mathcal{E}(\rho), |0^n\rangle\langle 0^n|) \geq \eta$ . What is the optimal work cost  $W_r^*$  of such a protocol? The following theorem bounds  $W_r^*$  in terms of the complexity relative entropy, generalizing Eq. (23) to product Hamiltonians.

**Theorem 3 (Optimal work cost for product Hamiltonians).** The optimal work cost  $W_r^*$  is quantified by the complexity relative entropy as

$$-D_{H, \mathcal{T}}^{r, \eta}(\rho \parallel \Gamma) - \log(1/\eta) \leq W_r^* \leq -D_{H, \mathcal{T}}^{r, \eta}(\rho \parallel \Gamma). \quad (36)$$

The subscript  $\mathcal{T}$  signifies that the set  $\mathbf{M}'$  in Eq. (11) is defined in terms of the gate set  $\mathcal{G} = \mathcal{T}$ .

The proof, provided in Appendix E, resembles the proof for degenerate Hamiltonians.

#### IV. WORK COSTS OF GENERAL STATE TRANSFORMATIONS

We now turn to our paper's motivating problem. Consider a system that begins in a state  $\rho$  and then undergoes any evolution of complexity at most  $r \geq 0$ .  $\rho$  will transform into a possibly different state. Which transformations appear possible, according to an observer who can distinguish states using only some computational power  $R \geq 0$ ? Alternatively, how much work must one invest to effect an evolution whose complexity is  $\leq r$  and whose output the observer cannot distinguish from a given state  $\rho'$ ?

##### A. Setting

To formalize the questions above, we consider an  $n$ -qubit system initialized in a state  $\rho$ . The state undergoes an evolution  $\mathcal{E}$  composed of primitive operations (Sec. II).

$C(\mathcal{E})$  and  $W(\mathcal{E})$  quantify the complexity cost and work cost of  $\mathcal{E}$ , respectively. To establish whether a transformation  $\rho \rightarrow \rho'$  appears to result from  $\mathcal{E}$ , imagine a referee who must distinguish the output  $\mathcal{E}(\rho)$  from the reference  $\rho'$ . The referee can effect only operations of complexities  $\leq R$ . Using a measurement operator of complexity  $\leq R$ , the referee performs a hypothesis test of the form described in Proposition 1, subject to the type-I and type-II error constraints specified with  $\eta \in (0, 1]$  and  $\delta \in (0, 1]$ , respectively. According to the referee,  $\rho \rightarrow \rho'$  may result from  $\mathcal{E}$  precisely if there exists no hypothesis test by which the referee can correctly accept  $\mathcal{E}(\rho)$  with a probability  $\geq \eta$  and incorrectly accept  $\rho'$  with a probability  $\leq \delta$ . Equivalently, by Proposition 1,  $\rho \rightarrow \rho'$  may result from  $\mathcal{E}$  precisely if

$$D_H^{R, \eta}(\mathcal{E}(\rho) \parallel \rho') \leq -\log(\delta/\eta). \quad (37)$$

Condition (37) can simplify, evoking the previous section's fidelity condition, if the referee lacks computational restrictions. Depending on the POVMs performable, the referee could perform the optimal measurement for distinguishing between  $\rho$  and  $\rho'$ . The measurement's success probability equals the trace distance  $\frac{1}{2}\|\rho - \rho'\|_1$ , which is small when the fidelity is large [95]. Hence condition (37) effectively reduces to a constraint, as in the previous section, on  $F$ .

Let  $\mathcal{E}$  denote any evolution, of a complexity  $\leq r$ , that transforms  $\rho$  into any state  $\mathcal{E}(\rho)$  indistinguishable from  $\rho'$  by the referee. The minimum work cost of any such evolution is

$$\begin{aligned} \mathcal{W}^{r, R, \eta, \delta}[\rho \rightarrow \rho'] \\ := \inf \left\{ W(\mathcal{E}) : C(\mathcal{E}) \leq r, D_H^{R, \eta}(\mathcal{E}(\rho) \parallel \rho') \right. \\ \left. \leq -\log(\delta/\eta) \right\}. \end{aligned} \quad (38)$$

This quantity, in its general form, does not appear amenable to simple analysis. In the following, we specialize to situations where the referee's computational power is extremely high or low. In these situations, we find examples where one of the conditions in Eq. (38) is expected to be tractable, as is needed to upper-bound the work cost.

##### B. Referee with extremely high computational power

First, suppose that  $R$  is extremely large. The first condition in Eq. (38) can be tractable: under certain conditions—see Appendix I— $D_H^{R, \eta}(\mathcal{E}(\rho) \parallel \rho')$  can approximate the hypothesis-testing relative entropy  $D_H^\eta(\mathcal{E}(\rho) \parallel \rho')$  [defined in Eq. (6)]. The latter can be evaluated via a semidefinite program (SDP), in principle [82]. Hence one can, in principle, evaluate each side of condition (37), as is necessary to calculate the work cost, Eq. (38). (As



a caveat, the SDP calculation's time and memory grow exponentially in  $n$ .)

### C. Referee with extremely low computational power

Now, suppose that  $R$  is extremely small. The referee may have great difficulty distinguishing the output  $\mathcal{E}(\rho)$  from the reference  $\rho'$ . Consider the extreme case where  $R = 0$ : the referee may perform only the local projective measurements in Eq. (10). We can apply a general upper bound: denote by  $\sigma$  and  $\tau$  any  $n$ -qubit states; and, by  $\sigma_j$  and  $\tau_j$ , the respective reduced states of qubit  $j$ . The complexity relative entropy obeys

$$D_H^{R=0,\eta}(\sigma \parallel \tau) \lesssim \sum_{j=1}^n D_H^\eta(\sigma_j \parallel \tau_j). \quad (39)$$

The approximation hides error terms that vanish in the limit as  $\eta \rightarrow 1$ . One can apply this inequality to evaluate Eq. (37) and thereby certify that  $\mathcal{E}(\rho)$  is indistinguishable from  $\rho'$  to a referee who lacks computational power and tolerates little error. [Inequality (39) follows from Proposition 18 in Appendix D.]

### D. Simple transformation

Suppose that the transformation  $\mathcal{E}$  involves few primitive operations:  $r < O(n)$ . Also, suppose that the system is 1D.  $\mathcal{E}$  can change the system's entanglement only by an amount proportional to  $r$ . This entanglement is related to the complexity entropy at small  $r$  (see Sec. VC for details). Hence we expect the first condition in Eq. (38) to relax to a constraint on entanglement, which we expect to be more tractable.

## V. INFORMATION-THEORETIC FEATURES OF THE COMPLEXITY ENTROPY

The complexity relative entropy (12) and the complexity entropy (14) are additions to a large cast of information-theoretic entropy measures [94,105]. Intuitively, the complexity entropy quantifies a state's apparent randomness to an agent who can implement only limited-complexity measurement effects. Unlike common entropies, the complexity entropy lacks unitary invariance. Indeed, a computationally limited agent may need more work to erase a state after it has evolved under a complex unitary (cf. examples in Sec. III A).

In this section, we present the complexity entropy's properties, as well as its applications to information-theoretic tasks. Most generally, the complexity (relative) entropy is defined for an abstract family  $\{M^r\}$  of sets. Each  $M^r$  denotes the set of POVM effects of complexities at most  $r \geq 0$ . This formalism generalizes the above,  $n$ -qubit constructions to arbitrary discrete quantum systems. In

Appendix D, we present the general definition of the complexity (relative) entropy and prove the properties stated below. Here, to achieve a simpler and more concrete presentation, we focus on an  $n$ -qubit system and on the family  $\{M^r\}$  defined by Eq. (11).

For information-theoretic applications, it is convenient to measure entropy in units of bits, rather than in nats [1 bit =  $\log(2)$  nats]. In this section, we denote with an overbar entropies expressed in units of bits; these entropies are related to their counterparts by the factor  $\log(2)$ :

$$\begin{aligned} \bar{H}(\rho) &= H(\rho)/\log(2), \\ \bar{D}_H^\eta(\rho \parallel \sigma) &= D_H^\eta(\rho \parallel \sigma)/\log(2), \\ \bar{H}_H^\eta(\rho) &= H_H^\eta(\rho)/\log(2), \\ \bar{D}_H^{r,\eta}(\rho \parallel \sigma) &= D_H^{r,\eta}(\rho \parallel \sigma)/\log(2), \text{ etc.} \end{aligned} \quad (40)$$

The factor  $[\log(2)]^{-1}$  changes natural logarithms to base-2 logarithms in the entropies' definitions. For instance, the von Neumann entropy  $H(\rho) = -\text{tr}(\rho \log \rho)$ , expressed in units of bits, is  $\bar{H}(\rho) = -\text{tr}(\rho \log_2 \rho)$ .

### A. Overview of elementary properties

In this subsection,  $\rho$  denotes any quantum state, and  $\Gamma$  any positive-semidefinite operator, defined on the  $n$ -qubit Hilbert space. Let  $r \geq 0$  and  $\eta \in (0, 1]$ . The complexity relative entropy inherits some properties from the hypothesis-testing relative entropy (6). For example,  $\bar{D}_H^{r,\eta}(\rho \parallel \Gamma)$ , like  $\bar{D}_H^\eta(\rho \parallel \Gamma)$ , monotonically decreases as  $\eta$  increases. The greater an agent's error intolerance (the greater the  $\eta$ ), the more mixed  $\rho$  appears [the greater  $H_H^{r,\eta}(\rho)$  is]. The complexity (relative) entropy also has properties that reflect its sensitivity to state complexity. For example,  $\bar{D}_H^{r,\eta}(\rho \parallel \Gamma)$  monotonically increases as  $r$  increases. The greater an agent's computational power (the greater the  $r$ ), the less mixed  $\rho$  appears [the less  $H_H^{r,\eta}(\rho)$  is]. These monotonicity properties imply that  $\rho$  is more distinguishable from  $\Gamma$  to an observer who has greater computational power and has a higher tolerance for guessing  $\Gamma$  when given  $\rho$ .

Furthermore, the complexity (relative) entropy has the same range as the hypothesis-testing (relative) entropy:

$$\bar{D}_H^{r,\eta}(\rho \parallel \Gamma) \in [-\log_2(\text{tr}(\Gamma)), -\log_2(\|\Gamma\|)], \quad \text{and} \quad (41a)$$

$$\bar{H}_H^{r,\eta}(\rho) \in [0, n]. \quad (41b)$$

The complexity relative entropy, like standard relative entropies, enjoys a scaling property in its second argument: for all  $\lambda > 0$ ,  $\bar{D}_H^{r,\eta}(\rho \parallel \lambda\Gamma) = \bar{D}_H^{r,\eta}(\rho \parallel \Gamma) - \log_2(\lambda)$ .

By construction, the complexity (relative) entropy is not unitarily invariant. On similar grounds, it does not satisfy a data-processing inequality. Indeed, unitary evolution is

a special case of data processing. Furthermore, a unitary can increase or decrease a pure state's complexity. Therefore, a unitary can increase or decrease a state's complexity entropy. Nevertheless, the complexity (relative) entropy monotonically increases (decreases) under partial traces. Tracing out  $k$  qubits from  $\rho$  yields

$$\overline{D}_H^{r,\eta}(\rho \parallel \Gamma) \geq \overline{D}_H^{r,\eta}(\text{tr}_k(\rho) \parallel \text{tr}_k(\Gamma)), \quad (42a)$$

$$\overline{H}_H^{r,\eta}(\rho) \leq \overline{H}_H^{r,\eta}(\text{tr}_k(\rho)) + k. \quad (42b)$$

The complexity (relative) entropy is greater (less) than or equal to the hypothesis-testing (relative) entropy:  $\overline{H}_H^{r,\eta}(\rho) \geq \overline{H}_H^\eta(\rho)$  and  $\overline{D}_H^{r,\eta}(\rho \parallel \Gamma) \leq \overline{D}_H^\eta(\rho \parallel \Gamma)$ . In the large- $r$  limit, one would expect the complexity (relative) entropy to converge to the hypothesis-testing (relative) entropy. Yet, an exact convergence might not occur: some POVM effects might not be well approximated by any effect in  $\mathcal{M}^r$ . This is due to our choice of  $\mathcal{M}^r$  and to the lack of ancillas in our computational model. Nevertheless, we show in Appendix I that, under certain conditions, the complexity (relative) entropy approximates the hypothesis-testing (relative) entropy up to constant error terms. Alternatively, one might hope to ensure convergence to the hypothesis-testing (relative) entropy by allowing the use of one ancillary qubit in implementations of measurement effects, as introduced in Ref. [17].

## B. Bounds from well-known complexity measures

To what extent does the complexity entropy  $\overline{H}_H^{r,\eta}(\psi)$  quantify the complexity of a pure state  $\psi \equiv |\psi\rangle\langle\psi|$ ? We bound the complexity entropy in terms of two complexity measures: the *strong complexity* of Ref. [17] and the *approximate circuit complexity* (Definition 5 in Appendix B). We describe the bounds qualitatively here, deferring rigorous statements, and proofs thereof, to Appendix D4.

The complexity entropy reflects approximate circuit complexity as follows. Assume that the set  $\mathcal{G}$ , used in the definition (11), contains only unitary gates. Consider beginning with some reference state, then applying gates in  $\mathcal{G}$  to prepare any state  $\varepsilon$ -close to  $|\psi\rangle$  in trace distance. Let  $C^\varepsilon(|\psi\rangle)$  denote the least number of gates in any such process.  $C^\varepsilon(|\psi\rangle)$  is the  $\varepsilon$ -approximate circuit complexity of  $|\psi\rangle$ . For all  $r \geq C^\varepsilon(|\psi\rangle)$ ,  $\overline{H}_H^{r,1-\varepsilon^2}(\psi) \approx 0$ . Furthermore,  $\overline{H}_H^{r,1-\varepsilon^2}(\psi) \not\approx 0$  for all  $r < C^{2\varepsilon}(|\psi\rangle)$  (see Proposition 20). The complexity relative entropy also obeys an upper bound involving the *strong complexity* of Ref. [17] (see Proposition 21).

## C. Bounds from entanglement

We lower-bound an  $n$ -qubit state's complexity entropy using a measure of the state's entanglement. Consider a 1D chain  $S$  of  $n \geq 2$  qubits:  $S = S_1 S_2 \dots S_n$ . Let  $\rho$  denote any

state of  $S$ . We define the following entanglement measure:

$$E(\rho) := \frac{1}{n-1} \sum_{j=1}^{n-1} I(S_1 \dots S_j : S_{j+1} \dots S_n)_\rho. \quad (43)$$

$I(A : B)_\rho := H(A)_\rho + H(B)_\rho - H(AB)_\rho$  denotes the quantum mutual information defined in terms of the von Neumann entropy  $H(X)_\rho := -\text{tr}(\rho \log \rho)$ . The mutual information quantifies all correlations—classical and quantum—including correlations due to entanglement.

Assume that the operations  $\mathcal{G}$  in Eq. (11) are unitary gates that can act nontrivially only on two neighboring qubits (the gates are *geometrically local*). We upper-bound the change in  $E(\rho)$  under one gate in  $\mathcal{G}$ .

*Proposition 2 (Change in  $E(\rho)$  under a two-qubit unitary).* Let  $\rho$  denote any quantum state of a 1D chain  $S$  of  $n \geq 2$  qubits. Let  $U$  denote any unitary that can act nontrivially only on two neighboring qubits in  $S$ . Then

$$|E(U\rho U^\dagger) - E(\rho)| \leq \frac{8 \log(2)}{n-1}. \quad (44)$$

To prove Proposition 2, one observes that  $U$  can alter only one mutual information in Eq. (43)—the mutual information associated with (the bipartition that separates) the qubits on which  $U$  can act nontrivially. In Appendix G, we generalize Eq. (44) to account for the *potential entangling power* [27] of  $\mathcal{G}$ . This power quantifies the entanglement generable by one gate in  $\mathcal{G}$ .

By repeatedly applying Eq. (44), we upper-bound the change in  $E(\rho)$  under  $\leq r$  gates in  $\mathcal{G}$ . We then lower-bound the complexity entropy  $H_H^{r,\eta}(\rho)$  in terms of  $E(\rho)$  as

$$H_H^{r,\eta}(\rho) \geq \frac{1}{\eta} \left[ E(\rho) - r \frac{8 \log(2)}{n-1} + H(\rho) + \text{error terms} \right]. \quad (45)$$

The error terms depend on  $\eta$  and  $n$  and vanish in the limit as  $\eta \rightarrow 1$ . See Appendix G for details. In Appendix H, we also investigate similar bounds that arise from natural dynamics under local Hamiltonians. This setting is particularly relevant to locally interacting systems.

## D. Evolution of the complexity entropy under random circuits

Consider an  $n$ -qubit circuit generated with gates sampled Haar randomly from all two-qubit gates. The circuit effects a unitary. As the circuit depth grows, the unitary's complexity grows, until saturating at a value exponential in  $n$ . The complexity growth is at least sublinear [17] and, for some nonrobust complexity measures [18,19,106], is linear. We ask how well the complexity entropy tracks the complexity of a pure state evolving under a random circuit.

We consider *brickwork circuits* [18,56]. A brickwork circuit consists of staggered layers of nearest-neighbor two-qubit gates.

The following proposition reveals the behaviors of the complexity entropy at a fixed complexity scale  $r$ : under short circuits, the complexity entropy vanishes. Around a circuit depth  $t$  commensurate with  $r$ , the complexity entropy transitions to a near-maximal value (Fig. 3). We bound the interval of  $t$  values in which the transition happens to values satisfying  $t \sim r$ . We prove the proposition using a strategy borrowed from Ref. [17], which contains a similar proposition about the strong complexity.

*Proposition 3 (Transition of the complexity entropy under random circuits).* Consider a depth- $t$ , random  $n$ -qubit circuit  $V$  whose gates act in a brickwork layout. Let  $|\psi\rangle := V|0^n\rangle$ ,  $r \geq 0$ , and  $\eta \in (0, 1]$ . If the circuit is short, such that  $t \leq O(r)$ , then

$$\overline{H}_H^{r,\eta}(|\psi\rangle\langle\psi|) = 0. \quad (46)$$

If the circuit is long, such that  $t \geq \Omega(rn^3)$ , then

$$\overline{H}_H^{r,\eta}(|\psi\rangle\langle\psi|) \geq n - \log_2(2/\eta), \quad (47)$$

except with a probability  $e^{-\Omega(n)}$  over the sampling of  $V$ .

The proof of Eq. (46) is immediate: one can use the circuit  $V$  to construct a candidate  $Q = V^\dagger|0^n\rangle\langle 0^n|V \in \mathbf{M}^r$  for the optimization (12) that defines  $\overline{H}_H^{r,\eta}(|\psi\rangle\langle\psi|)$ . The proof of Eq. (47) is presented in Appendix F.

### E. Data compression under complexity limitations

The complexity entropy naturally quantifies the optimal efficiency of data compression under computational limitations. Let  $\rho$  denote any  $n$ -qubit state;  $m$ , a number of qubits; and  $\varepsilon \in [0, 1]$ , an error parameter. To perform *data compression under complexity limitations*, one seeks a unitary  $U$ , composed from  $\leq r$  two-qubit gates, such that some subset  $\mathcal{W} \subset \{1, 2, \dots, n\}$  of  $|\mathcal{W}| = m$  qubits satisfies

$$F^2(\text{tr}_{\mathcal{W}}(U\rho U^\dagger), |0^{n-m}\rangle\langle 0^{n-m}|) \geq 1 - \varepsilon. \quad (48)$$

One *compresses*  $\rho$  onto the qubits of  $\mathcal{W}$ , with an accuracy  $1 - \varepsilon$  and using  $\leq r$  gates. For background, consider an agent who lacks complexity limitations (who operates in the limit as  $r \rightarrow \infty$ ). The hypothesis-testing entropy quantifies the optimal one-shot data-compression size of  $\rho$  [82]

$$m_{\text{optimal}} \approx \overline{H}_H^{1-\varepsilon}(\rho). \quad (49)$$

Every protocol for data compression under complexity limitations can be mapped to a “simple” protocol for erasure. The erasure protocol consists of computational gates

followed by RESET operations [see Fig. 4(b)]. The least number  $m_{\text{optimal},r}$  of qubits onto which  $\rho$  can be compressed, with an accuracy  $1 - \varepsilon$  and using  $\leq r$  gates, satisfies

$$\overline{H}_H^{r,1-\varepsilon}(\rho) - \log_2(1/[1 - \varepsilon]) \leq m_{\text{optimal},r} \leq \overline{H}_H^{r,1-\varepsilon}(\rho). \quad (50)$$

The complexity entropy  $\overline{H}_H^{r,1-\varepsilon}(\rho)$  is defined with respect to the set  $\mathcal{G}$  of all two-qubit gates. In Appendix J, we prove Eq. (50).

### F. Complexity conditional entropy

In information theory, conditional entropy measures quantify the randomness of a system  $A$ , as apparent to an observer given access to a system  $B$  that may share correlations with  $A$ . The conditional entropy appears throughout classical and quantum information theory. Applications include communication [62,63,107], thermodynamic erasure with side information [60], and quantum entropic uncertainty relations [108]. Consider defining a conditional entropy  $\overline{H}_*(A|B)_\rho$  of a state  $\rho_{AB}$ . A standard technique is to identify an appropriate relative entropy  $\overline{D}_*$  and to set  $\overline{H}_*(A|B)_\rho := -\overline{D}_*(\rho_{AB} \| \mathbb{1}_A \otimes \rho_B)$ . Here,  $\rho_B := \text{tr}_A(\rho_{AB})$  is the reduced state of  $\rho$  on  $B$  [94].

$\overline{H}_*(A|B)_\rho$  reflects one’s ability to distinguish  $\rho_{AB}$  from a state that provides the same knowledge about  $B$  ( $B$  is in state  $\rho_B$ ) but minimal knowledge about  $A$  ( $A$  is maximally mixed and uncorrelated with  $B$ ). In this interpretation,  $\overline{D}_*$  quantifies the notion of distinguishability. Conditional entropies defined as above include the conditional min- and max-entropies and the conditional Rényi- $\alpha$  entropies [94].

We employ the complexity relative entropy (12) to define the complexity conditional entropy. Denote any arbitrary bipartite state by  $\rho_{AB}$ . For all  $r \geq 0$  and  $\eta \in (0, 1]$ , the *complexity conditional entropy* of  $\rho_{AB}$  is

$$\overline{H}_H^{r,\eta}(A|B)_\rho := -\overline{D}_H^{r,\eta}(\rho_{AB} \| \mathbb{1}_A \otimes \rho_B). \quad (51)$$

We prove the following properties of the complexity conditional entropy. It, like the complexity entropy, decreases monotonically as  $r$  increases and increases monotonically as  $\eta$  increases. Furthermore, the complexity conditional entropy is bounded in terms of the Hilbert-space dimensionality of  $A$ ,  $d_A$ , as

$$-\log_2(d_A) \leq \overline{H}_H^{r,\eta}(A|B)_\rho \leq \log_2(d_A). \quad (52)$$

The complexity conditional entropy also exhibits strong subadditivity: for every tripartite state  $\rho_{ABC}$ ,

$$\overline{H}_H^{r,\eta}(A|BC)_\rho \leq \overline{H}_H^{r,\eta}(A|B)_\rho. \quad (53)$$

The strong subadditivity follows from the complexity relative entropy’s monotonicity under partial traces. We prove

Eqs. (52) and (53) in Appendix D 6. The complexity conditional entropy has operational significance in a variant of the information-theoretic task of *decoupling*, we show next.

### G. Decoupling from a reference system

First, we define the decoupling of a system  $A$  from a system  $R$ . Define  $d_A$  as the Hilbert-space dimensionality of  $A$  and  $\pi_A := \mathbb{1}_A/d_A$  as a maximally mixed state. We say that  $A$  is maximally mixed and decoupled from  $R$  if the joint state  $\rho_{AR}$  is the product  $\pi_A \otimes \rho_R$  [109]. *Decoupling  $A$  from  $R$*  means transforming a general  $\rho_{AR}$  into a state  $\rho'_{AR}$  close to a state in which  $A$  is maximally mixed and decoupled from  $R$  [62,63]. Decoupling is an information-theoretic primitive applied in randomness extraction [78, 110], quantum communication [62], and thermodynamic erasure with side information [60].

The standard information-theoretic task of decoupling can be defined as follows. Alice possesses an  $n$ -qubit system  $A$  that might be correlated with a reference  $R$ . The joint system begins in a state  $\rho_{AR}$ . First, Alice applies to  $A$  a unitary  $U_0$ , preparing  $(U_0 \otimes \mathbb{1}_R)\rho_{AR}(U_0 \otimes \mathbb{1}_R)^\dagger =: \rho'_{AR}$ . Then, Alice discards a subsystem  $A_1$  formed from  $k \geq 0$  qubits of  $A$ . A subsystem  $A_2$ , formed from  $n - k$  qubits, remains. Alice sends  $A_2$  to a referee, who possesses  $R$ . The referee attempts to distinguish the reduced state  $\rho'_{A_2R}$  from  $\pi_{A_2} \otimes \rho_R$ . Alice succeeds if the referee cannot distinguish the states to within some error tolerance. Alice can ensure that  $\rho'_{A_2R}$  approximates  $\pi_{A_2} \otimes \rho_R$  if [63]

$$k \gtrsim \frac{1}{2} \left[ n - \overline{H}_{\min}^\varepsilon(A | R)_\rho \right]. \quad (54)$$

$\varepsilon > 0$  denotes a tolerance parameter.  $\overline{H}_{\min}^\varepsilon(A | R)_\rho$  denotes the *smooth conditional min-entropy* of  $\rho$  [63,94].

Here, we introduce a complexity-constrained variant of decoupling. Assume that  $R$  consists of qubits. Assume further that both Alice and the referee can implement two-qubit gates that form a set universal for quantum computation. We define our decoupling task as the task above, except we impose two complexity constraints. First, we constrain Alice's computational power, requiring that  $U_0$  be implementable with at most  $r_0 \geq 0$  gates. Second, we constrain the referee's computational power. Using at most  $r_1 \geq 0$  gates, the referee performs a hypothesis test of the form described in Proposition 1, subject to type-I and type-II error constraints specified with  $\eta \in (0, 1]$  and  $\delta \in (0, 1]$ , respectively. Alice succeeds in our decoupling task if there exists no hypothesis test by which the referee can correctly accept  $\rho'_{A_2R}$  with a probability  $\geq \eta$  and incorrectly accept  $\pi_{A_2} \otimes \rho_R$  with a probability  $\leq \delta$ . Equivalently, by Proposition 1, Alice is successful if

$$\overline{D}_{\text{H}}^{r_1, \eta}(\rho'_{A_2R} \parallel \pi_{A_2} \otimes \rho_R) \leq -\log_2(\delta/\eta). \quad (55)$$

In terms of the complexity conditional entropy  $\overline{H}_{\text{H}}^{r_1, \eta}(A_2 | R)_{\rho'}$ , the condition (55) is

$$n - k - \overline{H}_{\text{H}}^{r_1, \eta}(A_2 | R)_{\rho'} \leq -\log_2(\delta/\eta). \quad (56)$$

Thus, Alice succeeds if  $\overline{H}_{\text{H}}^{r_1, \eta}(A_2 | R)_{\rho'}$  is sufficiently close to its maximum value,  $n - k$ .

Before quantifying Alice's decoupling capabilities, we posit two expectations. First, we do not expect Alice's complexity constraint to meaningfully change the number of qubits that she can decouple from  $R$  (unless  $r_0$  is very small). Indeed, one may achieve near-optimal decoupling by choosing for  $U_0$  to be a random circuit of only  $O(n \log^2(n))$  gates, in an all-to-all-coupling model [111].

In contrast, we expect the referee's complexity constraint to substantially increase the number of qubits that Alice can decouple from  $R$ , by enabling Alice to succeed with strategies that would fail at standard decoupling. For instance, consider a highly complex, maximally entangled state  $\rho_{AR}$ . In the standard scenario, Alice cannot decouple any qubits from  $R$ . [We can check this claim by substituting  $\overline{H}_{\min}^\varepsilon(A | R)_\rho = -n$  in Eq. (54).] Now, suppose that the referee is computationally limited. Alice's system is apparently decoupled from  $R$  already: a complex state can be indistinguishable from a maximally mixed state, so the referee cannot distinguish  $\rho_{AB}$  from a decoupled state. More precisely, suppose that  $\rho_{AR}$  is maximally entangled, that  $\eta = 1$ , and that the referee has unlimited computational power. Alice cannot decouple  $n - k$  qubits from  $R$ , if  $\delta \leq 2^{-(n-k)}$ . [If Alice discards  $k$  qubits, then  $\rho'_{A_2R}$  is a uniform mixture of  $2^k$  pure states. In an optimal strategy, the referee transforms  $\rho'_{A_2R}$  into  $\rho''_{A_2R} := |0^{2(n-k)}\rangle\langle 0^{2(n-k)}| \otimes (\mathbb{1}_2^{\otimes k}/2^k)$ , before implementing the measurement effect

$$Q := |0^{2(n-k)}\rangle\langle 0^{2(n-k)}| \otimes \mathbb{1}_2^{\otimes k}. \quad (57)$$

Consequently, the referee can distinguish  $\rho'_{A_2R}$  from  $\pi_{A_2} \otimes \rho_R = \pi_{A_2R}$ :  $\overline{D}_{\text{H}}^{r_1, \eta}(\rho'_{A_2R} \parallel \pi_{A_2} \otimes \rho_R) = -\log_2(\text{tr}(Q\pi_{A_2R})/\text{tr}(Q\rho'_{A_2R})) = 2(n - k) \leq -\log_2(\delta/\eta)$ .] Now, however, suppose that the referee's computational power is substantially limited and that  $\rho_{AR}$  is highly complex. Alice may be able to decouple  $n - k$  (or more) qubits even if  $\delta \leq 2^{-(n-k)}$ : the referee may require more computational power to perform a unitary necessary for distinguishing  $\rho'_{A_2R}$  from  $\pi_{A_2} \otimes \rho_R$ .

How many qubits  $k$  must Alice discard to convince the referee that a maximally mixed, decoupled state was prepared? We present a bound that generalizes Eq. (54) to our decoupling task. The bound holds if one assumes that the complexity conditional entropy obeys an inequality reminiscent of the Rényi-entropy chain rule [112,113].

*Conjecture 1 (Chain rule for the complexity conditional entropy).* Let  $\rho_{ABR}$  denote a quantum state of systems  $A$ ,



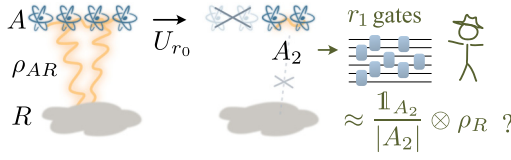


FIG. 5. Decoupling with respect to a computationally bounded referee. Alice's system  $A$  might be correlated with a reference system  $R$ . Alice applies to  $A$  a unitary  $U_0$ , then discards  $k$  qubits of  $A$ . Alice retains an  $(n - k)$ -qubit system  $A_2$ . Alice succeeds in our decoupling task if a referee cannot, using a circuit of  $\leq r_1$  gates, distinguish the final state of  $A_2R$  from  $\pi_{A_2} \otimes \rho_R$ . The success condition depends on the conditional complexity entropy  $\overline{H}_H^{r_1, \eta}(A_2 | R)_{\rho'}$ . Conjecturing a property of that entropy, we bound, in terms of  $\overline{H}_H^{r, \eta}(A | R)_{\rho}$ , the number of qubits that Alice can decouple from  $R$ . Our bound mirrors the bound (54), which holds in the absence of complexity constraints.

$B$ , and  $R$  of  $n_A$ ,  $n_B$ , and  $n_R$  qubits. Let  $r \geq 0$  and  $\eta \in (0, 1]$ . We conjecture that

$$\overline{H}_H^{r, \eta}(B | R)_{\rho} \leq \overline{H}_H^{r, \eta}(AB | R)_{\rho} + n_A. \quad (58)$$

In words, introducing a system  $A$  can decrease the complexity conditional entropy by, at most, the size  $n_A$  of  $A$ . The von Neumann conditional entropy  $\overline{H}(B | R)$  obeys an analogous bound, due to a chain rule and the lower bound  $\overline{H}(A | BR) \geq -n_A$ :  $\overline{H}(B | R) = \overline{H}(AB | R) - \overline{H}(A | BR) \leq \overline{H}(AB | R) + n_A$ . This  $\overline{H}(B | R)$  bound reflects how introducing  $A$  can resolve only as much randomness as  $A$  can store. Indeed, the bound saturates whenever  $A$  is maximally entangled with  $B$  ( $R$ ). In such a case,  $B$  and  $R$  are uncorrelated, and  $A$  purifies a maximally mixed state of  $B$  ( $R$ ). Hence  $A$  resolves  $n_A$  maximally mixed qubits' randomness. Similar bounds follow from chain rules for Rényi entropies [112, 113].

Using Conjecture 1, we prove a bound on the number of qubits that Alice can decouple from  $R$ .

**Theorem 4 (Upper bound on the number of qubits that Alice can decouple).** Consider the complexity-restricted decoupling described above and depicted in Fig. 5. Suppose  $r_1 \geq r_0$ . Assume Conjecture 1. Under the condition (55) (if Alice is successful), then

$$k \geq \frac{1}{2} \left[ n - \overline{H}_H^{r_1 - r_0, \eta}(A | R)_{\rho} + \log_2(\delta/\eta) \right]. \quad (59)$$

We prove Theorem 4 in Appendix K. The bound (59) mirrors the bound (54), which governs decoupling without complexity restrictions. Indeed, if  $\eta \approx 0$ , the conditional hypothesis-testing entropy approximates the smooth conditional min-entropy [82]. As such, one can interpret the conditional complexity entropy, when  $\eta \approx 0$ , as a *complexity-aware version of the conditional min-entropy*.

The general achievability of the bound (59) is unknown. A standard technique for proving standard decoupling's

achievability involves bounding the trace distance between a protocol's output and a decoupled target, using the Hilbert-Schmidt norm [62, 63, 114]. One cannot readily extend this technique to Theorem 4: while the trace distance's operational definition naturally extends to complexity-restricted POVM effects [17], such an extension is unknown for the Hilbert-Schmidt norm.

Conjecture 1 appears to preclude a kind of pseudorandom quantum state that we call a *pseudomixed state*. Let  $A$  and  $B$  denote systems of  $n_A$  and  $n_B$  qubits, respectively. Let  $|\chi\rangle_{AB}$  denote a pure state of  $AB$ . Let  $r \geq 0$  and  $\eta \in (0, 1]$ .  $|\chi\rangle_{AB}$  is an *efficiently preparable pseudomixed state* on  $B$  if the complexity entropy of  $|\chi\rangle$  is low,  $\overline{H}_H^{r, \eta}(|\chi\rangle\langle\chi|) \approx 0$ , while the complexity entropy of the reduced state  $\chi_B := \text{tr}_A(|\chi\rangle\langle\chi|)$  is high:

$$\overline{H}_H^{r, \eta}(\chi_B) \geq \Omega(\exp(n_A)). \quad (60)$$

In other words,  $B$  appears exponentially more mixed without the purifying  $A$ . The condition (60) directly contradicts the conjecture (58). Therefore, any efficiently preparable pseudomixed state would disprove Conjecture 1. We can therefore interpret Conjecture 1 as a *no-efficient-pseudomixedness conjecture*.

A naïve approach to efficiently constructing a pseudomixed state from pseudorandom states [115, 116] may fail. Standard constructions of pseudorandom states take as inputs short, random bit strings  $\kappa$ . The constructions yield efficient circuits for states  $|\phi_{\kappa}\rangle$  that have the following property: no efficient quantum algorithm can distinguish  $\text{poly}(n)$  copies of  $|\phi_{\kappa}\rangle$  from the same number of copies of a Haar-random state. One could naïvely expect to construct a pseudomixed state efficiently in the following way. Consider an  $n$ -qubit system formed from two subsystems: an  $O(\log(n))$ -qubit subsystem  $A$  and an  $O(n)$ -qubit subsystem  $B$ . One can efficiently prepare the state

$$|\psi\rangle_{AB} := \sum_{\kappa} |\kappa\rangle_A \otimes |\phi_{\kappa}\rangle_B \quad (61)$$

from  $|0^n\rangle$ , because pseudorandom states are efficiently preparable. Let  $r_{\text{gen}} \leq \text{poly}(n)$  denote the number of gates required to prepare  $|\psi\rangle_{AB}$ . (One can generate  $|\psi\rangle_{AB}$  by first preparing  $A$  in  $\sum_{\kappa} |\kappa\rangle_A$ , by applying Hadamard gates. Then, one implements a controlled unitary  $\sum_{\kappa} |\kappa\rangle\langle\kappa|_A \otimes U_{\kappa}$ .  $U_{\kappa}$  denotes an efficient unitary that prepares  $|\phi_{\kappa}\rangle$ .) For all  $\eta$ ,  $\overline{H}_H^{r_{\text{gen}}, \eta}(|\psi\rangle\langle\psi|) = 0$ . Denote by  $\psi_B := \text{tr}_A(|\psi\rangle\langle\psi|)$  the reduced state of  $B$ . By discarding  $A$ , one forgets the seed  $\kappa$  used to construct the pseudorandom state  $|\phi_{\kappa}\rangle_B$ . Therefore, no efficient algorithm can distinguish  $\psi_B$  from a Haar-random state.  $|\psi\rangle_{AB}$  may appear to violate Conjecture 1, looking highly mixed to a computationally limited observer. Nevertheless,  $\psi_B$  may have a low complexity entropy, since the pseudorandom states have low complexities. ( $\psi_B$  appears mixed to a computationally limited

observer because the state’s preparation circuit, represented by the seed, is unknown. Evaluating the complexity entropy, we imagine an observer who is given black-box access to a circuit that prepares the state. Under such circumstances, pseudorandom states do not look random. To appear highly mixed with respect to the complexity entropy, a pure state must have high complexity.) Thus,  $|\psi\rangle_{AB}$  does not necessarily constitute a counterexample to Conjecture 1.

## VI. DISCUSSION

Our framework highlights quantum complexity in thermodynamics. Thermodynamics is operational: it concerns how efficiently an agent, given certain resources, can accomplish certain tasks. We incorporate into a version of the theory complexity restrictions on (i) the agent’s operations and (ii) the evaluation of the agent’s output. This “agent” could be human, could consist of a system’s natural dynamics, etc.

Our framework introduces a new role for time in thermodynamics: a complex process requires a long time. Conventional thermodynamics distinguishes only what is possible and impossible. Augmented with complexity, our model of thermodynamics dictates what is practical and impractical. A spin glass illustrates the need to incorporate complexity restrictions into thermodynamics: according to the conventional thermodynamic laws, a spin glass can cool to low-temperature states. However, this cooling process would require a very long time, or a highly complex process.

We illustrated the interplay between complexity and conventional thermodynamics through Landauer erasure. Consider aiming to erase a highly complex pure state. One could “uncompute” the state, paying in complexity. Alternatively, one could erase every qubit, paying work. More generally, consider resetting an arbitrary state  $\rho$  with a success probability  $\eta$ . The complexity entropy  $H_H^{r,\eta}(\rho)$  quantifies the trade-off between the complexity and work costs.

Landauer erasure is one of many thermodynamic and information-processing tasks. We analyzed data compression and decoupling, as well. Many tasks merit analysis in the presence of complexity constraints. We expect complexity-restricted entropies to quantify these tasks’ optimal efficiencies.

Our entropies quantify how mixed a state appears to a computationally limited agent. For instance, even a pure state can have a high complexity entropy, if uncomputable by an agent. Hence the complexity entropy quantifies a variation on pseudorandomness—more specifically, pseudomixedness: randomness apparent in deterministic phenomena, due to the observer’s computational limitations.

Unlike the standard hypothesis-testing entropy, the complexity entropy cannot be approximated via convex

optimization. We expect that computing the complexity entropy is typically hard, given strong evidence that computing the state complexity is hard [117,118]. Yet it might be possible to derive more-tractable bounds for the complexity entropy in specific settings. Examples include low-complexity regimes as well as settings featuring random dynamics. We derived such a bound for pure states evolving under random circuits (Sec. VD).

Other measures have been defined to quantify complexity or apparent randomness: the computational min-entropy [119], the coarse-grained entropy [120], the observational entropy [121], the logical depth [122], and the quantum Kolmogorov complexity [53,123]. These measures reflect different approaches to quantifying complexity, as illustrated in the introduction. Rigorous comparisons between these quantities and the complexity entropies merit future work.

The main text illustrates the complexity entropies’ properties and usefulness; but generalizations are possible, and many are provided in Appendix D. For example, systems can be generalized beyond qubits to qudits. We expect extensions to continuous-variable systems to be achievable. Also, the complexity measure used can be replaced, as with Nielsen’s complexity [124–127]. One could even replace  $r$  with a matrix-product-state bond dimension—anything that constrains the POVM effects implementable. One need only specify those POVM effects to construct general complexity entropies.

Another opportunity is to literally analyze—break apart—our RESET and EXTRACT operations. We have presented these operations as a computation’s basic units. Yet each operation may consist of multiple steps, entailing extra costs. To incorporate these costs into our proofs, one could apply results from Ref. [47]. Furthermore, we attributed to the RESET the ideal work cost of  $k_B T \ln(2)$ . Any realistic RESET will cost more work. Such extra costs can be absorbed straightforwardly into our assumptions. Alternatively, we could consider a different set of primitives  $\mathcal{E}$  that have fixed work costs  $W(\mathcal{E})$  and complexity costs  $C(\mathcal{E})$ . (Any primitive operation could cost a nonzero amount of work and a nonzero amount of complexity.) One possibility is to follow the approach in Ref. [40]. That is, allow as a primitive operation every completely positive, trace-preserving two-qubit map  $\mathcal{E}$ . Then assign to  $\mathcal{E}$  a complexity cost of 1 and a work cost quantified by  $\log \|e^{\beta H/2} \mathcal{E}(e^{-\beta H}) e^{\beta H/2}\|$ , wherein  $H$  denotes the operated-on system’s Hamiltonian. Using this approach, one could lower-bound a process’s thermodynamic costs [40].

Complexity was recently incorporated into thermodynamics alternatively, in a resource theory [30]. In the resource theory of uncomplexity, an agent can perform arbitrarily many free operations. Each operation is slightly noisy, and the agent will have a natural noise tolerance. Hence the agent will naturally limit the number of

operations they perform. In the present paper, the agent's complexity restriction is a hard, external constraint. Just as complexity has recently emerged in resource theories, pseudorandomness has emerged in studies of quantum gravity [24,128,129]. We therefore anticipate the complexity entropy's utility in understanding black holes in the context of the AdS-CFT correspondence, uniting tools of quantum information theory, high-energy physics, and statistical physics.

### ACKNOWLEDGMENTS

The authors thank Matthew Coudron, Richard Kueng, Lorenzo Leone, Yi-Kai Liu, Carl Miller, Ralph Silva, and Twesh Upadhyaya for valuable discussions. This research has been supported by the National Institute of Standards and Technology and the Joint Center for Quantum Information and Computer Science under NIST Grant No. 70NANB21H055\_0, as well as by the Ford Foundation and by the National Science Foundation (QLCI Grant No. OMA-2120757). A.M. thanks Freie Universität Berlin for its hospitality and acknowledges The Ford Foundation for its support. The Berlin team has been funded by the DFG (FOR 2724 and CRC 183), the FQXi, and the ERC (DebuQC). N.B.T.K. acknowledges support by the European Space Agency (EISI Project No. 2021-01250-ESA), and by the Spanish MICIN (Project No. PID2022-141283NB-I00) with the support of FEDER funds. J.H. was funded by the Harvard Quantum Initiative.

### APPENDIX A: PRELIMINARIES, NOTATION, AND USEFUL LEMMAS

Throughout these Appendices, we use the following definitions and conventions. We denote the set of non-negative reals by  $\mathbb{R}_+ := \{x \in \mathbb{R} : x \geq 0\}$ . We consider a system  $S$  equipped with a Hilbert space  $\mathcal{H}_S$  of a finite dimensionality  $d_S$ . For any system  $R$  distinct from  $S$ , we denote the composition of  $S$  and  $R$  by  $SR$ .  $\mathcal{H}_{SR}$ , as well as spaces of operators and of superoperators on  $\mathcal{H}_S$ , are equipped with the standard topology of  $\mathbb{C}^{d_S}$ . We denote the identity operator on  $\mathcal{H}_S$  by  $\mathbb{1}_S$ . We denote the identity superoperator on  $\mathcal{H}_S$  by  $\text{id}_S$ . We denote the single-qubit identity operator by  $\mathbb{1}_2$ . For any Hermitian operators  $A$  and  $B$ , we write  $A \leq B$  if  $B - A$  is positive semidefinite. A *subnormalized quantum state*  $\rho_S$  is a positive-semidefinite operator that satisfies  $\text{tr}(\rho_S) \leq 1$ . If  $\text{tr}(\rho_S) = 1$ , then  $\rho_S$  is normalized and represents a physical quantum state. In the absence of ambiguity, we omit system subscripts from operators. We denote the projector onto a pure state  $|\psi\rangle$  by  $\psi := |\psi\rangle\langle\psi|$ . We denote by  $\pi_S := \mathbb{1}_S/d_S$  the maximally mixed state of  $S$ . For any sets  $X$  and  $Y$  of operators, we define  $X \otimes Y := \{A \otimes B : A \in X, B \in Y\}$ . An operator  $A$  has the operator norm  $\|A\|$ , defined as the greatest singular value of  $A$ . The trace norm  $\|A\|_1$  is the sum of the singular values of  $A$ . We denote by  $|0\rangle$  the eigenvalue-1

eigenvector of the Pauli- $z$  operator. Tensor products are notated as  $|0^k\rangle := |0\rangle^{\otimes k}$ . A *POVM effect*  $Q$  on a system  $S$  is a (Hermitian) positive-semidefinite operator satisfying  $0 \leq Q \leq \mathbb{1}_S$ . Finally, all logarithms are base- $e$ , unless otherwise indicated.

The *diamond norm* of a superoperator  $\mathcal{X}$  on  $\mathcal{H}_S$  is defined as

$$\|\mathcal{X}\|_\diamond := \max_{\rho_{SR}} \{ \|(\mathcal{X} \otimes \text{id}_R)(\rho_{SR})\|_1 \}. \quad (\text{A1})$$

$R$  denotes any system isomorphic to  $S$ , and the maximization is defined with respect to all subnormalized states  $\rho_{SR}$ . For any subnormalized state  $\rho$  on  $S$ , the diamond norm  $\|\mathcal{X}\|_\diamond$  upper-bounds the trace norm  $\|\mathcal{X}(\rho)\|_1$  as

$$\|\mathcal{X}(\rho)\|_1 = \|(\mathcal{X} \otimes \text{id}_R)(\rho_S \otimes \rho_R)\|_1 \leq \|\mathcal{X}\|_\diamond. \quad (\text{A2})$$

For any completely positive, trace-preserving maps  $\mathcal{E}$  and  $\mathcal{F}$ , the diamond distance  $\frac{1}{2}\|\mathcal{E} - \mathcal{F}\|_\diamond$  quantifies the one-shot distinguishability of  $\mathcal{E}$  and  $\mathcal{F}$  [109].

For any state  $\rho$  of  $S$ , the *von Neumann entropy* of  $\rho$  is  $H(\rho) \equiv H(S)_\rho := -\text{tr}(\rho \log(\rho))$ . For any positive-semidefinite operator  $\Gamma$  on  $\mathcal{H}_S$ , the (Umegaki) *quantum relative entropy* of  $\rho$  with respect to  $\Gamma$  is  $D(\rho \parallel \Gamma) := \text{tr}(\rho[\log(\rho) - \log(\Gamma)])$ . If  $S$  is a joint system,  $S = XY$ , the *conditional quantum entropy*  $H(X \mid Y)_\rho$  is  $H(X \mid Y)_\rho := H(XY)_\rho - H(Y)_\rho$ .

The following inequality is used often in quantum information theory. The inequality is equivalent to a general lower bound on a version of the conditional min-entropy [78,94].

*Lemma 1 (Partial order for joint states).* Let  $X$  and  $Y$  denote distinct quantum systems. If  $\rho_{XY}$  denotes any subnormalized state on  $XY$ , then

$$\frac{1}{d_X} \rho_{XY} \leq \mathbb{1}_X \otimes \rho_Y. \quad (\text{A3})$$

*Proof.* Let  $C$  denote a system isomorphic to  $X$ . Let  $\{|j\rangle_X\}$  and  $\{|j\rangle_C\}$  denote bases for the eigenspaces of  $X$  and  $C$ . If  $|\chi\rangle_C := \sum_j |j\rangle_C / \sqrt{d_X}$ , then  $\rho_{XY} \otimes \chi_C \leq \rho_{XY} \otimes \mathbb{1}_C$ , since  $\chi_C \leq \mathbb{1}_C$ . Every positive map preserves the partial order on operators. [If  $\mathcal{N}$  is a positive (linear) map acting on operators  $A$  and  $B \geq A$ , then  $\mathcal{N}(B) \geq \mathcal{N}(A)$ :  $\mathcal{N}(B - A) = \mathcal{N}(B) - \mathcal{N}(A)$  is positive-semidefinite, since  $B - A$  is positive-semidefinite.] Hence, conjugating both sides of  $\rho_{XY} \otimes \chi_C \leq \rho_{XY} \otimes \mathbb{1}_C$  by  $\sum_j |j\rangle_X \langle j| \otimes |j\rangle_C$  yields the

pinching inequality

$$\begin{aligned} \rho_{XY} &= d_X \left( \sum_j |j\rangle\langle j|_X \otimes |j\rangle\langle j|_C \right) (\rho_{XY} \otimes \chi_C) \\ &\quad \times \left( \sum_k |k\rangle\langle k|_X \otimes |k\rangle\langle k|_C \right) \\ &\leq d_X \sum_j |j\rangle\langle j|_X \langle j|_X \rho_{XY} |j\rangle_X. \end{aligned} \quad (\text{A4})$$

In terms of the Fourier-transformed states  $|\tilde{l}\rangle_X := \sum_j e^{i2\pi lj/d_X} |j\rangle_X / \sqrt{d_X}$ , the pinching inequality is

$$\rho_{XY} \leq d_X \sum_l |\tilde{l}\rangle\langle\tilde{l}|_X \langle\tilde{l}|_X \rho_{XY} |\tilde{l}\rangle_X. \quad (\text{A5})$$

Applying the positive map  $d_X \sum_j |j\rangle\langle j|_X \langle j|_X (\cdot) |j\rangle_X$  to both sides of Eq. (A5) yields

$$\begin{aligned} d_X \sum_j |j\rangle\langle j|_X \langle j|_X \rho_{XY} |j\rangle_X \\ \leq d_X^2 \sum_{jl} |j\rangle\langle j|_X \left| \langle j| \tilde{l}_X \right|^2 \langle\tilde{l}|_X \rho_{XY} |\tilde{l}\rangle_X \\ = d_X \mathbb{1}_X \otimes \text{tr}_X(\rho_{XY}) = d_X \mathbb{1}_X \otimes \rho_Y. \end{aligned} \quad (\text{A6})$$

We have used  $|\langle j| \tilde{l}_X \rangle|^2 = 1/d_X$ . Chaining together Eqs. (A4) and (A6) proves the lemma. ■

The following lemma provides a useful expression for, and bound on, the diamond distance between a unitary operation and the identity process.

*Lemma 2 (Diamond distance between a unitary operation and the identity process).* Let  $U$  denote any unitary operator, and let  $\mathcal{U}(\cdot) := U(\cdot)U^\dagger$ . It holds that

$$\frac{1}{2} \|\mathcal{U} - \text{id}\|_\diamond = \sin\left(\min\left\{e(U), \frac{\pi}{2}\right\}\right), \quad (\text{A7})$$

where

$$\begin{aligned} e(U) &:= \min \left\{ \|H\| : H = H^\dagger, \exists \chi \in \mathbb{R} \right. \\ &\quad \left. \text{such that } e^{-iH} = e^{-i\chi} U \right\}. \end{aligned} \quad (\text{A8})$$

Furthermore,

$$\frac{1}{2} \|\mathcal{U} - \text{id}\|_\diamond \leq \|U - \mathbb{1}\|. \quad (\text{A9})$$

The quantity  $e(U)$  is called the *potential entangling power* of  $U$ ; a similar quantity is defined in Ref. [27].  $e(U)$

is insensitive to global phases: for all  $\chi \in \mathbb{R}$ ,  $e(e^{i\chi}U) = e(U)$ .

Inequality Eq. (A9) automatically yields an upper bound on the diamond distance between arbitrary unitary operations. For all unitaries  $U$  and  $V$  acting on the same Hilbert space, the operations  $\mathcal{U}(\cdot) := U(\cdot)U^\dagger$  and  $\mathcal{V}(\cdot) := V(\cdot)V^\dagger$  obey

$$\begin{aligned} \frac{1}{2} \|\mathcal{U} - \mathcal{V}\|_\diamond \\ = \frac{1}{2} \|\mathcal{V}^\dagger \mathcal{U} - \text{id}\|_\diamond \leq \|V^\dagger U - \mathbb{1}\| = \|U - V\|. \end{aligned} \quad (\text{A10})$$

The first (second) equality follows from the diamond (operator) norm's invariance under unitary operations (operators).

*Proof of Lemma 2.* By Proposition 18 of Ref. [130],  $\frac{1}{2} \|\mathcal{U} - \text{id}\|_\diamond = \sin(\min\{\alpha, \pi/2\})$ , wherein  $\alpha$  denotes half the angle of the shortest arc  $A$ , on the unit circle, that contains every eigenvalue (eigenphase) of  $U$ . Let  $\chi_* \in \mathbb{R}$  such that  $e^{i\chi_*}$  is the midpoint of  $A$ . (Since  $U$  has only finitely many eigenphases,  $A$  does not encompass the unit circle, so  $A$  has endpoints and a midpoint.) By definition,  $A = \{e^{i\chi_*} e^{i\varphi} : \varphi \in [-\alpha, \alpha]\}$ .  $A$  has the endpoints  $\theta_\pm := e^{i\chi_*} e^{\pm i\alpha}$ .  $\theta_\pm$  are eigenphases of  $U$ . Otherwise, one could obtain an arc shorter than  $A$  that contains every eigenphase of  $U$ .

To prove Eq. (A7), we show that  $e(U) = \alpha$ . Let  $-iH_*$  denote the unique logarithm of  $e^{-i\chi_*}U$  such that every eigenvalue of  $H_*$  lies in  $(-\pi, \pi]$ . Since  $H_*$  is Hermitian and  $e^{-iH_*} = e^{-i\chi_*}U$ ,  $e(U) \leq \|H_*\|$ . Because every eigenvalue of  $U$  lies in  $A$ , every eigenvalue of  $e^{-iH_*}$  lies in the arc  $A_* := \{e^{i\varphi} : \varphi \in [-\alpha, \alpha]\}$ , which one obtains by multiplying every point in  $A$  by  $e^{-i\chi_*}$ . Consequently, every eigenvalue of  $H_*$  lies in  $[-\alpha, \alpha]$ , so  $\|H_*\| \leq \alpha$ . Hence,  $e(U) \leq \alpha$ .

For the sake of contradiction, suppose  $e(U) < \alpha$ . Let  $H$  denote any Hermitian operator that achieves the minimum in Eq. (A8).  $e^{-iH} = e^{-i\chi}U$  for some  $\chi \in \mathbb{R}$ , and  $\|H\| = e(U) < \alpha$ . Since every eigenvalue of  $H$  lies in  $[-\|H\|, \|H\|]$ , the arc  $A_H := \{e^{i\chi_*} e^{i\varphi} : \varphi \in [-\|H\|, \|H\|]\}$  contains every eigenvalue of  $e^{i\chi_*} e^{-iH} = U$ . Having the half-angle  $\|H\| < \alpha$ ,  $A_H$  is shorter than  $A$ , a contradiction. Therefore,  $e(U) = \alpha$ .

We now prove Eq. (A9). Suppose  $\alpha \geq \pi/2$ . Then  $A$  contains a half-circle and thus has an endpoint  $\theta \in \{\theta_\pm\}$  whose real part is nonpositive:  $\text{Re}(\theta) \leq 0$ .  $\theta$  is an eigenphase of  $U$ , so  $\theta - 1$  is an eigenvalue of  $U - \mathbb{1}$ . Hence,

$$\begin{aligned} \frac{1}{2} \|\mathcal{U} - \text{id}\|_\diamond &= \sin\left(\min\left\{\alpha, \frac{\pi}{2}\right\}\right) = 1 \leq |\text{Re}(\theta) - 1| \\ &= |\text{Re}(\theta - 1)| \leq |\theta - 1| \leq \|U - \mathbb{1}\|. \end{aligned}$$



Now, suppose  $\alpha < \pi/2$ . Then

$$\begin{aligned}
\frac{1}{2} \|\mathcal{U} - \text{id}\|_\diamond &= \sin\left(\min\left\{\alpha, \frac{\pi}{2}\right\}\right) = \sin(\alpha) = |e^{i\chi_*} \sin(\alpha)| \\
&= \frac{1}{2} |e^{i\chi_*} (e^{i\alpha} - e^{-i\alpha})| = \frac{1}{2} |\theta_+ - \theta_-| \\
&= \frac{1}{2} |(\theta_+ - 1) - (\theta_- - 1)| \\
&\leq \frac{1}{2} |\theta_+ - 1| + \frac{1}{2} |\theta_- - 1| \\
&\leq \|U - \mathbb{1}\|.
\end{aligned} \tag{A11}$$

The second inequality follows because  $U - \mathbb{1}$  has the eigenvalues  $\theta_\pm - 1$ . ■

## APPENDIX B: MEASURES OF COMPLEXITY AND POVM-EFFECT COMPLEXITY

We define the complexity (relative) entropy in terms of POVM-effect complexity. This section's main purpose is to introduce general assumptions that define POVM-effect complexity. Our setting's generality enables one to define the complexity (relative) entropy for a range of physical systems, including qubits, qudits, and continuous-variable systems.

### 1. Complexity of superoperators and unitaries

Using conservative assumptions, we define complexity measures for superoperators and unitaries. We later use these measures to construct measures of POVM-effect complexity.

*Definition 1 (Superoperator-complexity measure).* Let  $S$  denote a composition of  $N$  quantum subsystems:  $S = S_1 S_2 \dots S_N$ . For every subset  $I \subset \{1, 2, \dots, N\}$ , let  $S_I$  denote the composition of the subsystems in  $\{S_i\}_{i \in I}$ . A superoperator-complexity measure on  $S$  is a family of functions  $C_{S_I}$  that map completely positive, trace-nonincreasing maps to  $\mathbb{R}_+ \cup \{\infty\}$ , for all  $I \subset \{1, 2, \dots, N\}$ . The functions must have the following properties, for every  $I \subset \{1, 2, \dots, N\}$ :

- (i) The identity operation has zero complexity:  $C_{S_I}(\text{id}_{S_I}) = 0$ .
- (ii) No sequential composition of operations has a complexity that exceeds the sum of the individual operations' complexities: for any operations  $\mathcal{E}_1$  and  $\mathcal{E}_2$  on  $S_I$ ,  $C_{S_I}(\mathcal{E}_2 \mathcal{E}_1) \leq C_{S_I}(\mathcal{E}_1) + C_{S_I}(\mathcal{E}_2)$ .
- (iii) No parallel composition of operations has a complexity that exceeds the sum of the individual operations' complexities: for any  $J$  disjoint from  $I$ , for any operation  $\mathcal{E}_1$  on  $S_I$ , and for any operation  $\mathcal{E}_2$  on  $S_J$ ,  $C_{S_I S_J}(\mathcal{E}_1 \otimes \mathcal{E}_2) \leq C_{S_I}(\mathcal{E}_1) + C_{S_J}(\mathcal{E}_2)$ .

For every operation  $\mathcal{E}$  on any  $S_I$ , we employ the shorthand notation  $C_S(\mathcal{E}) := C_{S_I}(\mathcal{E})$ . Accordingly, we denote a superoperator-complexity measure by its function  $C_S$ , for convenience.

The following is a natural way to construct a superoperator-complexity measure: Choose a set  $\mathcal{G}$  of elementary operations. Define the complexity of an arbitrary operation  $\mathcal{E}$  as the minimum number of operations in  $\mathcal{G}$  that, under composition, effect  $\mathcal{E}$ .

*Definition 2 (Circuit-superoperator-complexity measure).* Let  $S = S_1 S_2 \dots S_N$  denote a quantum system. We employ the notation of Definition 1. For every subset  $I \subset \{1, 2, \dots, N\}$ , let  $\mathcal{G}_{S_I}$  denote an (arbitrary) set of completely positive, trace-nonincreasing maps such that  $\text{id}_{S_I} \in \mathcal{G}_{S_I}$  and  $\mathcal{G}_{S_I} \subset \mathcal{G}_{S_J}$  for all  $J \supset I$ . Denote the set  $\mathcal{G}_S$  by  $\mathcal{G}$ . A circuit-superoperator-complexity measure on  $S$  associated with  $\mathcal{G}$  is a family of functions  $C_{\mathcal{G}, S_I}$ , for any  $I \subset \{1, 2, \dots, N\}$ , defined as

$$\begin{aligned}
C_{\mathcal{G}, S_I}(\mathcal{E}) &:= \min\{r \in \mathbb{N} : \mathcal{E}_r \dots \mathcal{E}_2 \mathcal{E}_1 = \mathcal{E}, \\
&\quad \mathcal{E}_i \in \mathcal{G}_{S_I} \forall i \in \{1, 2, \dots, r\}\}.
\end{aligned} \tag{B1}$$

We introduce three conventions. First, we denote a circuit-superoperator-complexity measure by its function  $C_{\mathcal{G}, S}$ . Second, we designate the composition of a zero number of operations on  $S_I$  as the identity operation  $\text{id}_{S_I}$ . Third, we set  $C_{S_I}(\mathcal{E}) = \infty$  whenever no finite sequence of operations from  $\mathcal{G}_{S_I}$  implements an operation  $\mathcal{E}$  exactly.

The following proposition confirms that every circuit-complexity measure (Definition 2) is a superoperator-complexity measure (Definition 1).

*Proposition 4 (Compatibility of circuit-superoperator complexity and general-superoperator complexity).* Define  $S$  and the gate set  $\mathcal{G}$  as in Definition 2. Every circuit-superoperator-complexity measure on  $S$  associated with  $\mathcal{G}$  is a superoperator-complexity measure.

*Proof.* We verify that every circuit-superoperator-complexity measure  $C_{\mathcal{G}, S}$  has the three properties in Definition 1. Property (i) holds because, by definition,  $\text{id}_{S_I} \in \mathcal{G}_{S_I}$ ; and, by convention,  $\text{id}_{S_I}$  is the sequential composition of zero operations. Property (ii) automatically characterizes every sequential composition that involves an infinite-complexity operation. Now, we prove that property (ii) characterizes finite-complexity operations  $\mathcal{E}$  and  $\mathcal{E}'$  on  $S_I$ . If the operations' complexities are  $C_{S_I}(\mathcal{E}) = r$  and  $C_{S_I}(\mathcal{E}') = r'$ , we can decompose  $\mathcal{E}$  and  $\mathcal{E}'$  as sequences of  $r$  and  $r'$  operations in  $\mathcal{G}_{S_I}$ , respectively. Therefore, we can decompose  $\mathcal{E}' \mathcal{E}$  as a sequence of  $r + r'$  operations in  $\mathcal{G}_{S_I}$ . Hence  $C_{S_I}(\mathcal{E}' \mathcal{E}) \leq r + r' = C_{S_I}(\mathcal{E}) + C_{S_I}(\mathcal{E}')$ . Property (iii) follows as a special case of Property (ii): one can express every operation  $\mathcal{E}$  on  $S_I$  as the operation  $\mathcal{E} \otimes \text{id}_{S_J}$  on  $S_I S_J$  and vice versa. ■

Many complexity measures in the literature are defined only for unitary operations. The following definition specializes Definition 1 to such situations.

**Definition 3 (Unitary-complexity measure).** A unitary-complexity measure is a superoperator-complexity measure that assigns finite values only to unitary operations. A unitary-complexity measure  $C_S$  is called *adjoint-invariant* if  $C_S(\mathcal{U}) = C_S(\mathcal{U}^\dagger)$  for every unitary operation  $\mathcal{U}$  on  $S$ .

Let  $U$  denote any unitary operator on  $\mathcal{H}_S$ , and let  $\mathcal{U}(\cdot) := U(\cdot)U^\dagger$ . We write  $C_S(U)$  as a shorthand for  $C_S(\mathcal{U})$ , for convenience. Correspondingly, a *unitary-circuit-complexity measure* is a unitary-complexity measure that is also a circuit-complexity measure. Equivalently, a unitary-circuit-complexity measure is a circuit-complexity measure associated with a set of unitary gates.

## 2. Quantum state complexity

Here, we introduce state-complexity measures in general, for completeness and for use in later appendices.

**Definition 4 (Exact-state-complexity measure).** Let  $S$  denote a quantum system. Let  $C_S$  denote any superoperator-complexity measure; and  $\rho_0$ , any state of  $S$ . The *exact-state-complexity measure* with respect to  $C_S$  and  $\rho_0$  is

$$C_S^{\rho_0}(\rho) := \inf\{C_S(\mathcal{E}) : \mathcal{E}(\rho_0) = \rho\}. \quad (\text{B2})$$

An exact-state-complexity measure  $C_S^{\rho_0}$  is called an *exact-pure-state-complexity measure* if  $\rho_0$  is pure and if  $C_S^{\rho_0}$  assigns finite values only to pure states.  $C_S^{\rho_0}$  is such a measure if, for instance,  $\rho_0$  is pure and  $C_S$  is a unitary-complexity measure.

**Definition 5 (Approximate-state-complexity measure).** Let  $S$  denote a quantum system. Let  $C_S$  denote any superoperator-complexity measure; and  $\rho_0$ , any state of  $S$ . Let  $\delta \geq 0$ . The  $\delta$ -*approximate-state-complexity measure* with respect to  $C_S$  and  $\rho_0$  is

$$C_S^{\rho_0, \delta}(\rho) := \lim_{\zeta \rightarrow 0^+} \inf\left\{C_S(\mathcal{E}) : \frac{1}{2}\|\mathcal{E}(\rho_0) - \rho\|_1 \leq \delta + \zeta\right\}. \quad (\text{B3})$$

The stabilization  $\lim_{\zeta \rightarrow 0^+}$  in Eq. (B3) helps smooth  $C_S^{\rho_0, \delta}$  at unusual discontinuities in superoperator-complexity measures. For instance, consider any set  $\mathcal{G}_0$  of two-qubit unitary gates that is universal for quantum computation. Let  $\mathcal{G}$  denote the set of  $n$ -qubit unitaries realizable as finite sequences of gates in  $\mathcal{G}_0$ .  $\mathcal{G}$  is closed under composition. Let  $C_{\mathcal{G}}$  denote the unitary-circuit-complexity measure associated with  $\mathcal{G}$  (Definition 2). By construction,

the identity  $1_2^{\otimes n}$  has zero complexity, and every unitary  $U \in \mathcal{G}$  has unit complexity:  $C_{\mathcal{G}}(U) = 1$ . Any other unitary has infinite complexity, since it can only be approximated by a sequence of gates in  $\mathcal{G}_0$ —and thus by a unitary in  $\mathcal{G}$  [95]. Let  $|\psi_0\rangle$  denote any pure  $n$ -qubit state. The stabilization ensures that  $C_S^{\psi_0, 0}(\psi) = 1$  for all pure states  $|\psi\rangle \neq |\psi_0\rangle$ . Without the stabilization, every state  $|\psi\rangle$  inaccessible from  $|\psi_0\rangle$  by any unitary in  $\mathcal{G}$  would have an infinite complexity:  $C_S^{\psi_0, 0}(\psi) = \infty$ .

As an immediate consequence of Eq. (B3),  $C_S^{\rho_0, \delta}$  is right-continuous in  $\delta$ :

$$\lim_{\zeta \rightarrow 0^+} C_S^{\rho_0, \delta + \zeta}(\rho) = C_S^{\rho_0, \delta}(\rho). \quad (\text{B4})$$

An approximate-state-complexity measure  $C_S^{\rho_0, \delta}$  is called an *approximate-pure-state-complexity measure* if  $\rho_0$  is pure and if  $C_S^{\rho_0, 0}$  assigns finite values only to pure states.  $C_S^{\rho_0, \delta}$  is such a measure if, for instance,  $\rho_0$  is pure and  $C_S$  is a unitary-complexity measure.

The following proposition provides an alternative expression for Eq. (B3). From this expression, we derive a condition under which the stabilization in Eq. (B3) becomes inconsequential.

**Proposition 5 (Alternative expression for approximate-state-complexity measure).** Let  $S$  denote a quantum system. Let  $C_S$  denote any superoperator-complexity measure; and  $\rho_0$ , any state of  $S$ . Let  $\delta \geq 0$ . It holds that

$$C_S^{\rho_0, \delta}(\rho) = \inf\{r : \text{dist}_1(\rho, \mathfrak{G}_r[\rho_0]) \leq \delta\}, \quad (\text{B5})$$

where

$$\mathfrak{G}_r[\rho_0] := \{\mathcal{E}(\rho_0) : C_S(\mathcal{E}) \leq r\} \quad \text{and} \quad \text{dist}_1(\rho, T) := \inf_{\rho' \in T} \left\{ \frac{1}{2} \|\rho - \rho'\|_1 \right\}, \quad (\text{B6})$$

and wherein  $T$  denotes any set of states.

We now show that, under simple conditions, we can forgo the stabilization in the definition of the approximate-state-complexity measure.

**Corollary 1 (No stabilization of the approximate-state-complexity measure required for compact gate sets).** Consider the setting in Proposition 5. Suppose that, for all  $r \geq 0$ , the sets  $\mathfrak{G}_r[\rho_0]$  are compact. Then

$$C_S^{\rho_0, \delta}(\rho) = \inf\left\{C_S(\mathcal{E}) : \frac{1}{2}\|\mathcal{E}(\rho_0) - \rho\|_1 \leq \delta\right\}. \quad (\text{B7})$$

Furthermore, the sets  $\mathfrak{G}_r[\rho_0]$  are compact if  $C_S$  is a circuit-complexity measure (Definition 2) associated with a compact gate set  $\mathcal{G}_S$ .

*Proof of Proposition 5.* Let  $\rho$  denote any state of  $S$ . We show that the expressions in Eq. (B3) and (B5) equal

$$r^* := \inf \left\{ r : \forall \zeta > 0, \exists \mathcal{E} \text{ such that } C_S(\mathcal{E}) \leq r \text{ and } \frac{1}{2} \|\mathcal{E}(\rho_0) - \rho\| \leq \delta + \zeta \right\}. \quad (\text{B8})$$

The infimum in Eq. (B5) equals  $r^*$ , since

$$\begin{aligned} & \inf \{ r : \text{dist}_1(\rho, \mathfrak{G}_r[\rho_0]) \leq \delta \} \\ &= \inf \{ r : \forall \zeta > 0, \exists \rho' \in \mathfrak{G}_r[\rho_0] \\ & \text{such that } \frac{1}{2} \|\rho - \rho'\|_1 \leq \delta + \zeta \} = r^*. \end{aligned} \quad (\text{B9})$$

Now, consider  $C_S^{\rho_0, \delta}(\rho)$ , as defined in Eq. (B3). We show that  $C_S^{\rho_0, \delta}(\rho) = r^*$  by showing first that  $r^* \leq C_S^{\rho_0, \delta}(\rho)$  and then that  $C_S^{\rho_0, \delta}(\rho) \leq r^*$ . Consider any  $\zeta, \zeta' > 0$ . By (B3), there exists an  $\mathcal{E}$  such that  $C_S(\mathcal{E}) \leq C_S^{\rho_0, \delta}(\rho) + \zeta'$  and  $\frac{1}{2} \|\mathcal{E}(\rho_0) - \rho\|_1 \leq \delta + \zeta$ . Hence,  $r = C_S^{\rho_0, \delta}(\rho) + \zeta'$  is a candidate for the optimization in Eq. (B8), so

$$r^* \leq C_S^{\rho_0, \delta}(\rho) + \zeta'. \quad (\text{B10})$$

Thus,  $r^* \leq C_S^{\rho_0, \delta}(\rho)$ , since  $\zeta' > 0$  is arbitrary. By the infimum in Eq. (B8), there exists an  $\mathcal{E}$  and an  $r \leq r^* + \zeta'$  such that  $C_S(\mathcal{E}) \leq r$  and  $\frac{1}{2} \|\mathcal{E}(\rho_0) - \rho\|_1 \leq \delta + \zeta$ . Hence,  $r$  is lower-bounded as

$$\inf \left\{ C_S(\mathcal{E}) : \frac{1}{2} \|\mathcal{E}(\rho_0) - \rho\|_1 \leq \delta + \zeta \right\} \leq r \leq r^* + \zeta'. \quad (\text{B11})$$

Thus, the infimum in Eq. (B11) lower-bounds  $r^*$ , since  $\zeta' > 0$  is arbitrary. Consequently,

$$\begin{aligned} & C_S^{\rho_0, \delta}(\rho) \\ &= \lim_{\zeta \rightarrow 0^+} \inf \left\{ C_S(\mathcal{E}) : \frac{1}{2} \|\mathcal{E}(\rho_0) - \rho\|_1 \leq \delta + \zeta \right\} \leq r^*. \end{aligned} \quad (\text{B12})$$

■

*Proof of Corollary 1.* Suppose that  $\mathfrak{G}_r[\rho_0]$  is compact for all  $r \geq 0$ . The infimum defining  $\text{dist}_1(\rho, \mathfrak{G}_r[\rho_0])$ , in Eq. (B6), is achieved by some state in  $\mathfrak{G}_r[\rho_0]$ . Thus,

$$\text{dist}_1(\rho, \mathfrak{G}_r[\rho_0]) = \min \left\{ \frac{1}{2} \|\mathcal{E}(\rho_0) - \rho\|_1 : C_S(\mathcal{E}) \leq r \right\}. \quad (\text{B13})$$

Substituting Eq. (B13) into Eq. (B5) yields Eq. (B7).

Now, suppose that  $C_S$  is a circuit-complexity measure (Definition 2) associated with a gate set  $\mathcal{G}_S$ . Assume that  $\mathcal{G}_{S_I}$  is compact for all  $I \subset \{1, 2, \dots, N\}$ . Consider the function  $F$  that maps every sequence  $(\mathcal{E}_1, \dots, \mathcal{E}_r)$  of gates in  $\mathcal{G}_S$  to an action on  $\rho_0$ , as follows:

$$F(\mathcal{E}_1, \mathcal{E}_2, \dots, \mathcal{E}_r) := (\mathcal{E}_r \cdots \mathcal{E}_2 \mathcal{E}_1)(\rho_0). \quad (\text{B14})$$

By definition,  $F$  is a map from  $(\mathcal{G}_S)^{\times r}$  to  $\mathfrak{G}_r[\rho_0]$ .  $F$  is continuous, since it is a polynomial map.  $(\mathcal{G}_S)^{\times r}$  is compact, since the finite product of compact sets is compact. Therefore,  $\mathfrak{G}_r[\rho_0]$  is compact, since continuous functions preserve compactness. ■

### 3. POVM-effect complexity

A central ingredient in the complexity relative entropy's definition is a family  $\{\mathbf{M}^r\}$  of sets of POVM effects.  $\{\mathbf{M}^r\}$  is a formal parameter that can accommodate any notion of POVM-effect complexity. Each set  $\mathbf{M}^r$  contains precisely the POVM effects of complexities  $\leq r$ . An example of such a family of sets is defined via Eq. (11).

Here, we present conservative assumptions about every family  $\{\mathbf{M}^r\}$  to ensure that it has properties necessary for our construction of the complexity (relative) entropy.

*Definition 6 (POVM-effect-complexity sets).* Let  $S$  denote a composition of  $N$  quantum subsystems:  $S = S_1 S_2 \dots S_N$ . For every subset  $I \subset \{1, 2, \dots, N\}$ , let  $S_I$  denote the composition of the subsystems in  $\{S_i\}_{i \in I}$ . A family of POVM-effect-complexity sets on  $\mathcal{H}_S$  is a family  $\{\mathbf{M}_{S_I}^r\}_{r \in \mathbb{R}_+}$  of sets of POVM effects, for each  $S_I$ . Each  $\mathbf{M}_{S_I}^r$  is a POVM-effect-complexity set. The family must satisfy the following axioms, for every  $I \subset \{1, 2, \dots, N\}$ :

- (i) The identity operator has zero complexity:  $\mathbb{1}_{S_I} \in \mathbf{M}_{S_I}^0$ .
- (ii) For all  $r' \geq r \geq 0$ ,  $\mathbf{M}_{S_I}^r \subset \mathbf{M}_{S_I}^{r'}$ .
- (iii) For all  $r, r' \geq 0$  and for all  $J$  disjoint from  $I$ ,  $\mathbf{M}_{S_I}^r \otimes \mathbf{M}_{S_J}^{r'} \subset \mathbf{M}_{S_{IJ}}^{r+r'}$ .

For convenience, we write  $\{\mathbf{M}_S^r\}$  as shorthand for a family  $\{\mathbf{M}_{S_I}^r\}_{r, I}$ .

### 4. POVM-effect-complexity sets from simple POVM effects and a superoperator-complexity measure

A natural way to construct POVM-effect-complexity sets is as follows: Designate a set of “simple,” zero-complexity POVM effects. Define the complexity of an effect  $Q$  as the minimum complexity of any superoperator that maps a simple effect to  $Q$ . In the main text, we use this procedure to construct the sets  $\mathbf{M}^r$  in Eq. (11) from single-qubit projectors and a circuit-complexity measure.

The conditions in Definition 6 imply that zero-complexity effects must satisfy two conditions: (i) the

identity operator is a zero-complexity effect. (ii) The parallel composition of two zero-complexity effects is a zero-complexity effect. Any set of “simple” POVM effects must satisfy these criteria, which we formalize in the following definition.

**Definition 7 (Simple POVM effects).** Let  $S$  denote a composition of  $N$  quantum subsystems:  $S = S_1 S_2 \dots S_N$ . For every subset  $I \subset \{1, 2, \dots, N\}$ , let  $S_I$  denote the composition of the subsystems in  $\{S_i\}_{i \in I}$ . Let  $\{\mathbf{P}_{S_I}\}$  denote a family of sets of POVM effects. The  $\mathbf{P}_{S_I}$  are sets of *simple POVM effects* if, for all  $I \subset \{1, 2, \dots, N\}$ , the following conditions hold:

- (i)  $\mathbf{P}_{S_I}$  contains the identity operator:  $\mathbb{1}_{S_I} \in \mathbf{P}_{S_I}$ .
- (ii)  $\mathbf{P}_{S_I} \otimes \mathbf{P}_{S_J} \subset \mathbf{P}_{S_I S_J}$  for all  $J$  disjoint from  $I$ .

For convenience, we denote a family  $\{\mathbf{P}_{S_I}\}$  by its set  $\mathbf{P}_S$ .

We now construct POVM-effect-complexity sets from simple POVM effects and a superoperator-complexity measure.

**Definition 8 (POVM-effect-complexity sets from simple effects and a superoperator-complexity measure).** Let  $S = S_1 S_2 \dots S_N$  denote a quantum system as in Definition 7. Let  $\mathbf{P}_S$  denote any set of simple POVM effects (Definition 7); and  $C_S$ , any superoperator-complexity measure (Definition 1). For all  $r \geq 0$  and  $I \subset \{1, 2, \dots, N\}$ , we define

$$\mathbf{M}_{S_I}^r(\mathbf{P}_{S_I}, C_{S_I}) := \{\mathcal{E}^\dagger(P) : P \in \mathbf{P}_{S_I}, C_{S_I}(\mathcal{E}) \leq r\}. \quad (\text{B15})$$

**Proposition 6.** The sets  $\mathbf{M}_{S_I}^r(\mathbf{P}_{S_I}, C_{S_I})$  constructed in Eq. (B15) define a family of POVM-effect-complexity sets (Definition 6).

*Proof.* We write  $\mathbf{M}_{S_I}^r = \mathbf{M}_{S_I}^r(\mathbf{P}_{S_I}, C_{S_I})$  for short. We verify that the sets  $\mathbf{M}_{S_I}^r$  have the three properties in Definition 6. Property (i) holds because  $\mathbb{1}_{S_I} = \text{id}_{S_I}^\dagger(\mathbb{1}_{S_I})$ ,  $\mathbb{1}_{S_I} \in \mathbf{P}_{S_I}$  (Definition 7), and  $C_{S_I}(\text{id}_{S_I}) = 0$  (Definition 1). Property (ii) holds because, for all  $r' \geq r \geq 0$  and for all operations  $\mathcal{E}$  on  $S_I$ ,  $C_{S_I}(\mathcal{E}) \leq r$  implies  $C_{S_I}(\mathcal{E}) \leq r'$ . To show that Property (iii) holds, we consider any POVM effects  $Q_1 \in \mathbf{M}_{S_I}^r$  and  $Q_2 \in \mathbf{M}_{S_J}^{r'}$ , with  $I$  and  $J$  disjoint.  $Q_1 = \mathcal{E}_1^\dagger(P_1)$  for some effect  $P_1 \in \mathbf{P}_{S_I}$  and for some operation  $\mathcal{E}_1$  satisfying  $C_{S_I}(\mathcal{E}_1) \leq r$ . Similarly,  $Q_2 = \mathcal{E}_2^\dagger(P_2)$  for some effect  $P_2 \in \mathbf{P}_{S_J}$  and for some operation  $\mathcal{E}_2$  satisfying  $C_{S_J}(\mathcal{E}_2) \leq r'$ . By Definition 7,  $P_1 \otimes P_2 \in \mathbf{P}_{S_I S_J}$ . By Definition 1,  $C_{S_I S_J}(\mathcal{E}_1 \otimes \mathcal{E}_2) \leq C_{S_I}(\mathcal{E}_1) + C_{S_J}(\mathcal{E}_2) \leq r + r'$ . Thus,  $Q_1 \otimes Q_2 = (\mathcal{E}_1 \otimes \mathcal{E}_2)^\dagger(P_1 \otimes P_2) \in \mathbf{M}_{S_I S_J}^{r+r'}$ . ■

By Eq. (B15), every simple POVM effect is a zero-complexity effect, since the identity process has zero complexity. In general, the converse is false: not every

zero-complexity effect is a simple POVM effect. Indeed, a nontrivial zero-complexity operation applied to a simple effect may yield a zero-complexity effect that is not a simple effect. Let  $\mathbf{P}$  denote a set of simple POVM effects (Definition 7); and  $C$ , a superoperator-complexity measure (Definition 1). Suppose that there exists a zero-complexity operation  $\mathcal{E} \neq \text{id}$ :  $C(\mathcal{E}) = 0$ . Suppose, further, that  $\mathcal{E}^\dagger(P) \notin \mathbf{P}$  for some effect  $P \in \mathbf{P}$ . Then  $\mathbf{M}^{r=0}(\mathbf{P}, C) \supsetneq \mathbf{P}$ , since  $\mathcal{E}^\dagger(P) \in \mathbf{M}^{r=0}(\mathbf{P}, C)$ , by Eq. (B15). If  $\text{id}$  is the unique zero-complexity operation—or, more generally, if  $\mathbf{P}$  is closed under (the adjoint of) every zero-complexity operation—then  $\mathbf{M}^{r=0}(\mathbf{P}, C) = \mathbf{P}$ . In particular,  $\text{id}$  is the unique zero-complexity operation whenever  $C$  is a circuit-complexity measure (Definition 2).

## 5. POVM-effect-complexity sets from a POVM-effect-complexity measure

An alternative, equivalent formulation of a POVM-effect-complexity set (Definition 6) arises from a complexity measure on individual POVM effects. This formulation has the advantage of using a primitive formally similar to a superoperator-complexity measure (Definition 1). Consequently, some readers may find POVM-effect-complexity sets more intuitive in this formulation.

**Definition 9 (POVM-effect-complexity measure).** Let  $S$  denote a composition of  $N$  quantum subsystems:  $S = S_1 S_2 \dots S_N$ . For every subset  $I \subset \{1, 2, \dots, N\}$ , let  $S_I$  denote the composition of the subsystems in  $\{S_i\}_{i \in I}$ . A *POVM-effect-complexity measure* on  $S$  is a family of functions  $F_{S_I}$  that map POVM effects on  $S_I$  to  $\mathbb{R}_+ \cup \{\infty\}$ , for all  $I \subset \{1, 2, \dots, N\}$ . The functions have the following properties, for every  $I \subset \{1, 2, \dots, N\}$ :

- (i) The identity effect has zero complexity:  $F_{S_I}(\mathbb{1}_{S_I}) = 0$ .
- (ii) No parallel composition of POVM effects has a complexity greater than the sum of the individual effects' complexities:  $F_{S_I S_J}(Q_1 \otimes Q_2) \leq F_{S_I}(Q_1) + F_{S_J}(Q_2)$  for all  $J$  disjoint from  $I$ .

For convenience, we denote a POVM-effect-complexity measure by its function  $F_S$ . Every POVM-effect-complexity measure induces a family of POVM-effect-complexity sets (Definition 6), as per the following definition.

**Definition 10 (POVM-effect-complexity sets from a POVM-effect-complexity measure).** Let  $S = S_1 S_2 \dots S_N$  denote a quantum system; and  $F_S$ , a POVM-effect-complexity measure, as in Definition 9. A *family of POVM-effect-complexity sets* on  $\mathcal{H}_S$ , associated with  $F_S$ , is a family  $\{\mathbf{M}_{S_I}^r(F_{S_I})\}_{r \in \mathbb{R}_+}$  of sets of POVM effects. Each  $\mathbf{M}_{S_I}^r(F_{S_I})$



denotes the set of POVM effects on  $S_I$  of complexities  $\leq r$ :

$$M_{S_I}^r(F_{S_I}) := \{Q : F_{S_I}(Q) \leq r\}. \quad (\text{B16})$$

One can straightforwardly check that the sets  $M_{S_I}^r(F_{S_I})$  in Eq. (B16) satisfy the conditions of Definition 6.

We now construct a POVM-effect-complexity measure from simple POVM effects and a superoperator-complexity measure, providing an alternative way to define the sets  $M_{S_I}^r(P_{S_I}, C_{S_I})$  in Eq. (B15). Let  $S$  denote a quantum system; and  $F_S$ , a POVM-effect-complexity measure, as in Definition 9. Let  $P_S$  denote a set of simple POVM effects (Definition 7). The following prescription, for all subsets  $I \subset \{1, 2, \dots, N\}$ , defines a POVM-effect-complexity measure:

$$F_{S_I}(Q) := \inf\{C_{S_I}(\mathcal{E}) : P \in P_{S_I}, \mathcal{E}^\dagger(P) = Q\}. \quad (\text{B17})$$

**Proposition 7.** The sets  $F_{S_I}$  constructed in Eq. (B17) define a POVM-effect-complexity measure.

*Proof.* We verify that the sets  $F_{S_I}$  have both the properties in Definition 9. Property (i) holds because  $C_{S_I}(\text{id}_{S_I}) = 0$  (Definition 1),  $\mathbb{1}_{S_I} \in P_{S_I}$  (Definition 7), and  $\text{id}_{S_I}^\dagger(\mathbb{1}_{S_I}) = \mathbb{1}_{S_I}$ . To show that Property (ii) holds, we consider any POVM effects  $Q_1$  on  $S_I$  and  $Q_2$  on  $S_J$ , with  $I$  and  $J$  disjoint. Let  $P_1 \in P_{S_I}$ , and let  $\mathcal{E}_1$  denote an operation on  $S_I$  such that  $\mathcal{E}_1^\dagger(P_1) = Q_1$ . Similarly, let  $P_2 \in P_{S_J}$ , and let  $\mathcal{E}_2$  denote an operation on  $S_J$  such that  $\mathcal{E}_2^\dagger(P_2) = Q_2$ .  $\mathcal{E}_1 \otimes \mathcal{E}_2$  is a candidate for the  $F_{S_I S_J}(Q_1 \otimes Q_2)$  optimization in Eq. (B17), since  $P_1 \otimes P_2 \in P_{S_I S_J}$  (Definition 7) and  $(\mathcal{E}_1 \otimes \mathcal{E}_2)^\dagger(P_1 \otimes P_2) = Q_1 \otimes Q_2$ . Hence,  $F_{S_I S_J}(Q_1 \otimes Q_2) \leq C_{S_I S_J}(\mathcal{E}_1 \otimes \mathcal{E}_2)$ . Furthermore,  $C_{S_I S_J}(\mathcal{E}_1 \otimes \mathcal{E}_2) \leq C_{S_I}(\mathcal{E}_1) + C_{S_J}(\mathcal{E}_2)$  (Definition 1), so  $F_{S_I S_J}(Q_1 \otimes Q_2) \leq C_{S_I}(\mathcal{E}_1) + C_{S_J}(\mathcal{E}_2)$ . We obtain  $F_{S_I S_J}(Q_1 \otimes Q_2) \leq F_{S_I}(Q_1) + F_{S_J}(Q_2)$  by taking the infimum of the last inequality with respect to all candidates  $\mathcal{E}_1$  and  $\mathcal{E}_2$  for the  $F_{S_I}(Q_1)$  and  $F_{S_J}(Q_2)$  optimizations in Eq. (B17), respectively. ■

A POVM-effect-complexity measure constructed as in Eq. (B17), with a circuit-superoperator-complexity measure (Definition 2), is called a *circuit-POVM-effect-complexity measure*. The POVM-effect-complexity measure in Eq. (B17) induces, via Eq. (B16), the sets  $M_{S_I}^r(P_{S_I}, C_{S_I})$  in Eq. (B15).

## 6. Examples of POVM-effect-complexity sets

Any set of simple POVM effects (Definition 7) and any unitary-complexity measure (Definition 3) determine a family of POVM-effect-complexity sets, as per Eq. (B15). In this section, we provide natural examples of simple POVM effects and unitary-complexity measures on  $n$  qubits.

### a. Examples of simple POVM effects

Among the POVM effects that one can directly implement in a laboratory, local effects are often the most feasible. The following sets of simple POVM effects contain only tensor products.

- (i) The set of tensor products of the single-qubit projectors  $|0\rangle\langle 0|$  and  $\mathbb{1}_2$ . We use this choice throughout the main text [see Eq. (10)], as is natural in the context of erasure.
- (ii) The set of tensor products of single-qubit projectors. This choice is natural if one can easily perform any single-qubit gate before implementing a projector onto a single-qubit state, e.g.,  $|0\rangle$ .
- (iii) The set of tensor-product POVM effects. If performing nonlocal operations is hard, local operations are a natural choice for simple effects. Tensor-product POVM effects generalize tensor-product projectors.

Importantly, there exist POVM effects that one cannot implement by first applying a unitary and then implementing a tensor-product POVM effect. For instance, the two-qubit effect  $\frac{1}{2}(|00\rangle\langle 00| + |01\rangle\langle 01| + |10\rangle\langle 10|)$  is a rank-3 operator, unlike any tensor-product two-qubit effect: operator rank is invariant under unitary transformation, so the given effect is not unitarily equivalent to any tensor-product effect.

### b. Examples of unitary-complexity measures

A natural unitary-complexity measure on  $n$  qubits is the circuit complexity associated with a two-qubit gate set  $\mathcal{G}$  (Definition 2). Depending on one's experimental capabilities, one can consider circuits with all-to-all connectivity (any gate in  $\mathcal{G}$  can act on any two qubits); geometric locality (any gate in  $\mathcal{G}$  can act on only two neighboring qubits); or intermediate connectivity (some gates in  $\mathcal{G}$  can act on any two qubits, and other gates can act on only two neighboring qubits). Convenient choices for  $\mathcal{G}$  include the following:

- (i) The set of two-qubit gates:  $\mathcal{G} = \text{SU}(4)$ . This choice is widespread in studies of random circuits [55, 131]. For instance, this choice has elucidated the linear growth, with time, of exact circuit complexity under random circuits [18, 19].
- (ii) The following Clifford gates: Hadamard,  $S$ , and CNOT gates. This set generates quantum circuits that can be efficiently simulated classically [132]. The set is not universal, and almost all unitaries have infinite complexities with respect to the set.
- (iii) Any finite, universal set of two-qubit gates, such as the Clifford +  $T$  gate set. This set can approximate every unitary arbitrarily well with a finite-complexity unitary [95]. In this setting, the circuit

complexity grows algebraically under random quantum circuits with gates drawn from a finite gate set [17,55,56].

Our definition of unitary complexity accommodates other notions of complexity, such as the Nielsen complexity [124–127]. Another example is the minimum  $T$ -count of any Clifford +  $T$  circuit that approximates a given unitary. (In the latter example, different gates contribute different amounts of complexity. Each  $T$  gate contributes one unit of complexity; and each Clifford gate, zero units.

### APPENDIX C: HYPOTHESIS-TESTING (RELATIVE) ENTROPY

In a hypothesis test, one receives either  $\rho$  or  $\sigma$ , and guesses which state one received. In the most general strategy to distinguish  $\rho$  and  $\sigma$ , one performs a two-outcome measurement  $\{Q, \mathbb{1} - Q\}$ . One guesses  $\rho$  if  $Q$  obtains and guesses  $\sigma$  otherwise. Suppose that one must, if the state is  $\rho$ , guess  $\rho$  with a probability  $\geq \eta \in (0, 1]$ . The minimum probability of wrongly guessing  $\rho$  defines the *hypothesis-testing relative entropy*.

*Definition 11 (Hypothesis-testing relative entropy).* Let  $\rho$  denote any subnormalized state, and  $\Gamma$  any positive-semidefinite operator, that act on the same Hilbert space. Let  $\eta \in (0, \text{tr}(\rho)]$ . The *hypothesis-testing relative entropy* is defined as [82,83,86]

$$D_H^\eta(\rho \parallel \Gamma) := -\log \left( \min_{\substack{0 \leq Q \leq \mathbb{1} \\ \text{tr}(Q\rho) \geq \eta}} \left\{ \frac{\text{tr}(Q\Gamma)}{\eta} \right\} \right). \quad (\text{C1})$$

In a hypothesis test, the least probability of erroneously rejecting  $\sigma$ , using a strategy that successfully accepts  $\rho$  with a probability  $\geq \eta$ , is  $\eta \exp(-D_H^\eta(\rho \parallel \sigma))$ . Every optimal POVM effect  $Q$  satisfies  $\text{tr}(Q\rho) = \eta$ . If any such  $Q$  violated this equation, one could achieve a better objective value with  $Q' := \eta Q / \text{tr}(Q\rho) \leq Q$ . Consequently, if  $\Gamma = \rho$ , then  $D_H^\eta(\rho \parallel \Gamma) = 0$ . In this case, every  $Q$  satisfying  $\text{tr}(Q\rho) = \eta$  is optimal.

There always exists an optimal effect for the  $D_H^\eta$  optimization, since the domain of optimization is compact. This fact justifies the use of a minimum, instead of an infimum, in Eq. (C1).

*Definition 12 (Hypothesis-testing entropy).* Let  $\rho$  denote any subnormalized state. Let  $\eta \in (0, \text{tr}(\rho)]$ . The *hypothesis-testing entropy* is defined as [82]

$$H_H^\eta(\rho) := -D_H^\eta(\rho \parallel \mathbb{1}) = \log \left( \min_{\substack{0 \leq Q \leq \mathbb{1} \\ \text{tr}(Q\rho) \geq \eta}} \left\{ \frac{\text{tr}(Q)}{\eta} \right\} \right). \quad (\text{C2})$$

The hypothesis-testing entropy and relative entropy are related to smooth entropies, including the min- and max-relative entropies [79,82]. The smooth entropies quantify the optimal efficiencies of operational tasks performed in the absence of complexity restrictions. Such tasks feature finitely many copies of a quantum state or random variable, as well as finite failure probabilities.

One can express the hypothesis-testing relative entropy (C1) as a semidefinite program. The dual problem takes the form [82]

$$D_H^\eta(\rho \parallel \Gamma) = -\log \left( \max_{\substack{X \geq 0, \mu \geq 0 \\ \mu \rho \leq \Gamma + X}} \left\{ \mu - \frac{\text{tr}(X)}{\eta} \right\} \right). \quad (\text{C3})$$

We now reformulate the hypothesis-testing relative entropy. The rewriting reveals the entropy's formal similarity to the complexity relative entropy (12).

*Proposition 8 (Alternative expression for hypothesis-testing relative entropy).* Let  $\rho$  denote any subnormalized state; and  $\Gamma$ , any positive-semidefinite operator. Let  $\eta \in (0, \text{tr}(\rho)]$ . The hypothesis-testing relative entropy obeys

$$D_H^\eta(\rho \parallel \Gamma) = -\log \left( \inf_{\substack{0 \leq Q \leq \mathbb{1} \\ \text{tr}(Q\rho) \geq \eta}} \left\{ \frac{\text{tr}(Q\Gamma)}{\text{tr}(Q\rho)} \right\} \right). \quad (\text{C4})$$

Without loss of generality, we can assume that every candidate effect  $Q$  for the optimization satisfies  $\|Q\| = 1$ . If a candidate  $Q$  did not, we could replace  $Q$  with a candidate  $Q' := Q/\|Q\| \geq Q$  that achieves the same objective value as  $Q$ .

*Proof.* Let  $Q$  denote any POVM effect such that  $\text{tr}(Q\rho) \geq \eta$ . Let  $Q' := \eta Q / \text{tr}(Q\rho)$ . By definition,  $Q'$  satisfies  $\text{tr}(Q'\rho) = \eta$ .  $Q$  and  $Q'$  are candidates for the  $D_H^\eta$  optimization in Eq. (C1) and achieve the same objective value:

$$\frac{\text{tr}(Q\Gamma)}{\text{tr}(Q\rho)} = \frac{\text{tr}(Q'\Gamma)}{\text{tr}(Q'\rho)}. \quad (\text{C5})$$

Consequently, the set of candidate objective values in Eq. (C4) is unaltered if we restrict to POVM effects  $Q'$  satisfying  $\text{tr}(Q'\rho) = \eta$ :

$$\begin{aligned} & \left\{ \frac{\text{tr}(Q\Gamma)}{\text{tr}(Q\rho)} : 0 \leq Q \leq \mathbb{1}, \text{tr}(Q\rho) \geq \eta \right\} \\ &= \left\{ \frac{\text{tr}(Q'\Gamma)}{\text{tr}(Q'\rho)} : 0 \leq Q' \leq \mathbb{1}, \text{tr}(Q'\rho) = \eta \right\}. \end{aligned} \quad (\text{C6})$$

Therefore,

$$\begin{aligned} \inf_{\substack{0 \leq Q \leq \mathbb{1} \\ \text{tr}(Q\rho) \geq \eta}} \left\{ \frac{\text{tr}(Q\Gamma)}{\text{tr}(Q\rho)} \right\} &= \inf_{\substack{0 \leq Q' \leq \mathbb{1} \\ \text{tr}(Q'\rho) = \eta}} \left\{ \frac{\text{tr}(Q'\Gamma)}{\text{tr}(Q'\rho)} \right\} \\ &= \inf_{\substack{0 \leq Q' \leq \mathbb{1} \\ \text{tr}(Q'\rho) = \eta}} \left\{ \frac{\text{tr}(Q'\Gamma)}{\eta} \right\}. \end{aligned} \quad (\text{C7})$$

The second equality follows because  $\text{tr}(Q'\rho) = \eta$  for every  $Q'$  in the second infimum's domain.

Consider, again, any effect  $Q$  satisfying  $\text{tr}(Q\rho) \geq \eta$  and the effect  $Q' := \eta Q / \text{tr}(Q\rho)$ . Clearly  $Q' \leq Q$ , so  $\text{tr}(Q'\Gamma) / \eta \leq \text{tr}(Q\Gamma) / \eta$ . Hence, the final infimum in Eq. (C7) does not decrease if we extend its domain to all  $Q$  satisfying  $\text{tr}(Q\rho) \geq \eta$ :

$$\inf_{\substack{0 \leq Q' \leq \mathbb{1} \\ \text{tr}(Q'\rho) = \eta}} \left\{ \frac{\text{tr}(Q'\Gamma)}{\eta} \right\} = \inf_{\substack{0 \leq Q \leq \mathbb{1} \\ \text{tr}(Q\rho) \geq \eta}} \left\{ \frac{\text{tr}(Q\Gamma)}{\eta} \right\}. \quad (\text{C8})$$

Finally, the second infimum in Eq. (C8) is a minimum, since the infimum's domain is compact. Hence,

$$\inf_{\substack{0 \leq Q \leq \mathbb{1} \\ \text{tr}(Q\rho) \geq \eta}} \left\{ \frac{\text{tr}(Q\Gamma)}{\eta} \right\} = \min_{\substack{0 \leq Q \leq \mathbb{1} \\ \text{tr}(Q\rho) \geq \eta}} \left\{ \frac{\text{tr}(Q\Gamma)}{\eta} \right\} = e^{-D_{\text{H}}^{\eta}(\rho \parallel \Gamma)}. \quad (\text{C9})$$

The second equality follows from Definition 11. Chaining together Eqs. (C7), (C8), and (C9) yields an equality equivalent to Eq. (C4). ■

## APPENDIX D: COMPLEXITY (RELATIVE) ENTROPY AND ITS VARIANTS

Imagine a hypothesis test performed by an observer able to render only limited-complexity measurement effects. The hypothesis-testing relative entropy is ill-suited for such a test: the optimization in Eq. (C1) is appropriate only for an observer who can render all POVM effects. In Appendix D 1, we introduce the complexity (relative) entropy as a hypothesis-testing (relative) entropy tailored for the complexity-limited observer. In the optimization defining the complexity (relative) entropy, candidate POVM effects cannot exceed a given complexity  $r \geq 0$ . Appendix D 2 details the complexity (relative) entropy's elementary properties. In Appendix D 3, we apply the complexity relative entropy to hypothesis testing with complexity limitations. In Appendix D 4, we bound the complexity (relative) entropy using state-complexity measures. Appendix D 5 mainly concerns a variant of the complexity (relative) entropy, the reduced complexity (relative) entropy. The reduced complexity entropy features in our data-compression results and appears in Ref. [30] as

“the complexity entropy.” Last, in Appendix D 6, we define the complexity conditional entropy and study some of its elementary properties. In the following,  $\{M^r\}$  denotes any family of POVM-effect-complexity sets (Definition 6). The complexity (relative) entropy and its variants are defined with respect to this set, unless further specified.

### 1. Definition of the complexity (relative) entropy

*Definition 13 (Complexity relative entropy).* Let  $\rho$  denote any subnormalized state; and  $\Gamma$ , any positive-semidefinite operator. Let  $r \geq 0$  and  $\eta \in (0, \text{tr}(\rho)]$ . Here,  $\rho$ ,  $\Gamma$ , and every  $Q \in M^r$  act on the same Hilbert space. The *complexity-restricted hypothesis-testing relative entropy*, or simply the *complexity relative entropy*, is

$$D_{\text{H}}^{r,\eta}(\rho \parallel \Gamma) := -\log \left( \inf_{\substack{Q \in M^r \\ \text{tr}(Q\rho) \geq \eta}} \left\{ \frac{\text{tr}(Q\Gamma)}{\text{tr}(Q\rho)} \right\} \right). \quad (\text{D1})$$

The denominator  $\text{tr}(Q\rho)$  is a normalization factor ensuring that the complexity (relative) entropy has some desired properties. First, the complexity relative entropy satisfies  $D_{\text{H}}^{r,\eta}(\rho \parallel \rho) = 0$  (Proposition 9). Second, if  $\rho$  acts on a Hilbert space of dimensionality  $d$ , the complexity entropy assumes values in the range  $[0, \log(d)]$  (Proposition 10). The hypothesis-testing relative entropy is the same with a denominator  $\text{tr}(Q\rho)$  [as in Eq. (C4)] or  $\eta$  [as in Eq. (C1)]. In contrast, the complexity relative entropy might increase if we replace  $\text{tr}(Q\rho)$  with  $\eta$  in Eq. (D1): it may be impossible to find a  $Q \in M^r$  such that  $\text{tr}(Q\rho) = \eta$  [or such that  $\text{tr}(Q\rho)$  approximates  $\eta$  arbitrarily well]. This is typically the case, for instance, for a family of sets that contains only projectors; such a family is defined by Eq. (11).

*Definition 14 (Complexity entropy).* Let  $\rho$  denote any subnormalized state. Let  $r \geq 0$  and  $\eta \in (0, \text{tr}(\rho)]$ . The *complexity-restricted hypothesis-testing entropy*, or simply the *complexity entropy*, is

$$H_{\text{H}}^{r,\eta}(\rho) := -D_{\text{H}}^{r,\eta}(\rho \parallel \mathbb{1}) = \log \left( \inf_{\substack{Q \in M^r \\ \text{tr}(Q\rho) \geq \eta}} \left\{ \frac{\text{tr}(Q)}{\text{tr}(Q\rho)} \right\} \right). \quad (\text{D2})$$

The complexity entropy measures how well one can distinguish  $\rho$  from the maximally mixed state, using a limited-complexity measurement effect. Indeed, if  $\rho$  acts on a Hilbert space of dimensionality  $d$ , then  $H_{\text{H}}^{r,\eta}(\rho) = \log(d) - D_{\text{H}}^{r,\eta}(\rho \parallel \pi)$ . The  $\pi := \mathbb{1}/d$  denotes the maximally mixed state. Two special cases offer insight. First,  $H_{\text{H}}^{r,\eta}(\rho) = 0$  if  $\rho$  is pure and  $\rho \in M^r$ . Second,  $H_{\text{H}}^{r,\eta}(\rho) = \log(d)$  if  $\rho$  is normalized and  $\mathbb{1} \in M^r$  is the only candidate for the  $H_{\text{H}}^{r,\eta}(\rho)$  optimization.

The complexity entropy is analogous to the strong complexity of Ref. [17]. Compared to the strong complexity, the complexity entropy quantifies distinguishability in terms of a hypothesis test, rather than a modified trace distance. Also, the complexity entropy bears a meaningful interpretation for mixed states.

## 2. Elementary properties of the complexity (relative) entropy

The complexity (relative) entropy possesses several basic properties.

*Proposition 9 (Vanishing complexity relative entropy).* Let  $\rho$  denote any subnormalized state. Let  $r \geq 0$  and  $\eta \in (0, \text{tr}(\rho)]$ . It holds that

$$D_{\text{H}}^{r,\eta}(\rho \parallel \rho) = 0. \quad (\text{D3})$$

*Proof.* Every candidate effect  $Q \in \mathcal{M}^r$  for the optimization in Eq. (D1), including  $\mathbb{1} \in \mathcal{M}^{r=0} \subset \mathcal{M}^r$ , achieves the objective value  $\text{tr}(Q\rho)/\text{tr}(Q\rho) = 1$ . Hence,  $D_{\text{H}}^{r,\eta}(\rho \parallel \rho) = \log(1) = 0$ . ■

In general, the converse is false: there may exist a state  $\sigma \neq \rho$  such that  $D_{\text{H}}^{r,\eta}(\rho \parallel \sigma) = 0$ . This situation arises if the measurement effects in  $\mathcal{M}^r$  are too crude to distinguish  $\rho$  from  $\sigma$ . For instance, consider the sets  $\mathcal{M}^r$  defined in Eq. (11) for  $n$  qubits, with respect to the set of two-qubit gates. Let  $\rho = |1^n\rangle\langle 1^n|$ . For any  $\eta \in (0, 1]$ , the only effect  $Q \in \mathcal{M}^{r=0}$  satisfying  $\text{tr}(Q\rho) \geq \eta$  is  $Q = \mathbb{1}$ . Therefore, trivially,  $D_{\text{H}}^{r=0,\eta}(\rho \parallel \sigma) = 0$  for all  $\sigma$ .

The complexity relative entropy  $D_{\text{H}}^{r,\eta}(\rho \parallel \Gamma)$  lies within a fixed range dependent on only  $\text{tr}(\rho)$  and the eigenvalues of  $\Gamma$ . The range takes a simple form if  $\rho$  is normalized and  $\Gamma$  is positive definite.

*Proposition 10 (General bounds).* Let  $\rho$  denote any subnormalized state, and  $\Gamma$  any positive-semidefinite operator, that act on a Hilbert space of dimensionality  $d$ . Let  $r \geq 0$  and  $\eta \in (0, \text{tr}(\rho)]$ . It holds that

$$D_{\text{H}}^{r,\eta}(\rho \parallel \Gamma) \geq -\log(\text{tr}(\Gamma)) + \log(\text{tr}(\rho)). \quad (\text{D4a})$$

Furthermore, if  $\Gamma$  is positive definite (has full rank),

$$D_{\text{H}}^{r,\eta}(\rho \parallel \Gamma) \leq \log(\|\Gamma^{-1}\|) + \log(\text{tr}(\rho)). \quad (\text{D4b})$$

In particular,

$$-\log(\text{tr}(\rho)) \leq H_{\text{H}}^{r,\eta}(\rho) \leq \log(d) - \log(\text{tr}(\rho)). \quad (\text{D5})$$

Consequently, if  $\rho$  is normalized and  $\Gamma$  is positive definite,

$$-\log(\text{tr}(\Gamma)) \leq D_{\text{H}}^{r,\eta}(\rho \parallel \Gamma) \leq \log(\|\Gamma^{-1}\|), \quad (\text{D6a})$$

$$\text{and } 0 \leq H_{\text{H}}^{r,\eta}(\rho) \leq \log(d). \quad (\text{D6b})$$

Equation (D4a) implies that  $D_{\text{H}}^{r,\eta}(\rho \parallel \sigma) \geq 0$  for normalized states  $\rho$  and  $\sigma$ .

*Proof.* Equation (D4a) follows because  $\mathbb{1} \in \mathcal{M}^r$  is a candidate effect for the  $D_{\text{H}}^{r,\eta}(\rho \parallel \Gamma)$  optimization. To prove Eq. (D4b), let  $Q \in \mathcal{M}^r$  denote any candidate for the  $D_{\text{H}}^{r,\eta}(\rho \parallel \Gamma)$  optimization. If  $\Gamma$  is positive definite,  $\Gamma \geq \|\Gamma^{-1}\|^{-1}\mathbb{1}$ , wherein  $\|\Gamma^{-1}\|^{-1}$  is the smallest eigenvalue of  $\Gamma$ . Furthermore,  $\rho \leq \text{tr}(\rho)\mathbb{1}$ , since  $\text{tr}(\rho) \leq 1$ . Thus,  $\text{tr}(Q\Gamma) \geq \|\Gamma^{-1}\|^{-1}\text{tr}(Q)$ , and  $\text{tr}(Q\rho) \leq \text{tr}(\rho)\text{tr}(Q)$ . These two inequalities, together, imply that

$$\frac{\text{tr}(Q\Gamma)}{\text{tr}(Q\rho)} \geq \frac{\|\Gamma^{-1}\|^{-1}\text{tr}(Q)}{\text{tr}(\rho)\text{tr}(Q)} = (\|\Gamma^{-1}\|\text{tr}(\rho))^{-1}. \quad (\text{D7})$$

Therefore,  $(\|\Gamma^{-1}\|\text{tr}(\rho))^{-1}$  lower-bounds the objective value  $\text{tr}(Q\Gamma)/\text{tr}(Q\rho)$  of every candidate effect  $Q$ . By the definition of an infimum,

$$(\|\Gamma^{-1}\|\text{tr}(\rho))^{-1} \leq \inf_{\substack{Q \in \mathcal{M}^r \\ \text{tr}(Q\rho) \geq \eta}} \left\{ \frac{\text{tr}(Q\Gamma)}{\text{tr}(Q\rho)} \right\} = e^{-D_{\text{H}}^{r,\eta}(\rho \parallel \Gamma)}, \quad (\text{D8})$$

which is equivalent to Eq. (D4b). One obtains Eq. (D5) by setting  $\Gamma = \mathbb{1}$  in Eqs. (D4a) and (D4b). ■

The complexity relative entropy never exceeds the hypothesis-testing relative entropy.

*Proposition 11 (Upper bound by hypothesis-testing relative entropy).* Let  $\rho$  denote any subnormalized state; and  $\Gamma$ , any positive-semidefinite operator. Let  $r \geq 0$  and  $\eta \in (0, \text{tr}(\rho)]$ . It holds that

$$D_{\text{H}}^{r,\eta}(\rho \parallel \Gamma) \leq D_{\text{H}}^{\eta}(\rho \parallel \Gamma) \quad \text{and} \quad H_{\text{H}}^{r,\eta}(\rho) \geq H_{\text{H}}^{\eta}(\rho). \quad (\text{D9})$$

*Proof.* The  $D_{\text{H}}^{\eta}(\rho \parallel \Gamma)$  optimization ranges over all POVM effects, while the  $D_{\text{H}}^{r,\eta}(\rho \parallel \Gamma)$  optimization ranges over only effects in  $\mathcal{M}^r$ . ■

The greater an agent's computational power (the greater the  $r$ ), the less mixed  $\rho$  appears [the less  $H_{\text{H}}^{r,\eta}(\rho)$  is]. Likewise, the greater an agent's error intolerance (the greater the  $\eta$ ), the more mixed  $\rho$  appears [the greater  $H_{\text{H}}^{r,\eta}(\rho)$  is]. Consequently, the complexity entropy monotonically decreases as  $r$  increases and monotonically increases as  $\eta$  increases. The complexity entropy inherits its monotonicity in  $\eta$  from the hypothesis-testing entropy.

*Proposition 12 (Monotonicity in  $r$  and  $\eta$ ).* Let  $\rho$  denote any subnormalized state; and  $\Gamma$ , any positive-semidefinite operator. Let  $r \geq 0$  and  $\eta \in (0, \text{tr}(\rho)]$ .



- (i) For all  $r' \geq r$ ,  $D_{\text{H}}^{r',\eta}(\rho \parallel \Gamma) \geq D_{\text{H}}^{r,\eta}(\rho \parallel \Gamma)$ , and  $H_{\text{H}}^{r',\eta}(\rho) \leq H_{\text{H}}^{r,\eta}(\rho)$ .
- (ii) For all  $\eta' \in (0, \eta]$ ,  $D_{\text{H}}^{r,\eta'}(\rho \parallel \Gamma) \geq D_{\text{H}}^{r,\eta}(\rho \parallel \Gamma)$ , and  $H_{\text{H}}^{r,\eta'}(\rho) \leq H_{\text{H}}^{r,\eta}(\rho)$ .

*Proof.* The monotonicity in  $r$  follows because  $\mathbf{M}^r \subset \mathbf{M}^{r'}$ : every candidate  $Q \in \mathbf{M}^r$  for the  $D_{\text{H}}^{r,\eta}(\rho \parallel \Gamma)$  optimization belongs to  $\mathbf{M}^{r'}$  and is therefore a candidate for the  $D_{\text{H}}^{r',\eta}(\rho \parallel \Gamma)$  optimization. The monotonicity in  $\eta$  follows because every candidate  $Q$  for the  $D_{\text{H}}^{r,\eta}(\rho \parallel \Gamma)$  optimization satisfies  $\text{tr}(Q\rho) \geq \eta \geq \eta'$  and is therefore a candidate for the  $D_{\text{H}}^{r,\eta'}(\rho \parallel \Gamma)$  optimization. ■

Like standard relative entropies, the complexity relative entropy enjoys a scaling property in its second argument.

*Proposition 13 (Scaling property in the second argument).* Let  $\rho$  denote any subnormalized state; and  $\Gamma$ , any positive-semidefinite operator. Let  $r \geq 0$ ,  $\eta \in (0, \text{tr}(\rho)]$ , and  $a > 0$ . It holds that

$$D_{\text{H}}^{r,\eta}(\rho \parallel a\Gamma) = D_{\text{H}}^{r,\eta}(\rho \parallel \Gamma) - \log(a). \quad (\text{D10})$$

*Proof.* The equality holds because  $\log(a)$  factorizes out of the optimizer in Eq. (D1). ■

The complexity (relative) entropy inverts (preserves) the partial order of positive-semidefinite operators.

*Proposition 14 (Monotonicity under operator ordering).* Let  $\rho$  and  $\rho' \geq \rho$  denote any subnormalized states. Let  $\Gamma$  and  $\Gamma' \leq \Gamma$  denote any positive-semidefinite operators. Let  $r \geq 0$  and  $\eta \in (0, \text{tr}(\rho)]$ . It holds that

$$D_{\text{H}}^{r,\eta}(\rho \parallel \Gamma) \leq D_{\text{H}}^{r,\eta}(\rho' \parallel \Gamma') \quad (\text{D11a})$$

$$\text{and } H_{\text{H}}^{r,\eta}(\rho) \geq H_{\text{H}}^{r,\eta}(\rho'). \quad (\text{D11b})$$

*Proof.* For every POVM effect  $Q$ ,  $\text{tr}(Q\rho) \leq \text{tr}(Q\rho')$ , and  $\text{tr}(Q\Gamma) \geq \text{tr}(Q\Gamma')$ . Therefore,

$$\begin{aligned} e^{-D_{\text{H}}^{r,\eta}(\rho \parallel \Gamma)} &= \inf_{\substack{Q \in \mathbf{M}^r \\ \text{tr}(Q\rho) \geq \eta}} \left\{ \frac{\text{tr}(Q\Gamma)}{\text{tr}(Q\rho)} \right\} \geq \inf_{\substack{Q \in \mathbf{M}^r \\ \text{tr}(Q\rho') \geq \eta}} \left\{ \frac{\text{tr}(Q\Gamma')}{\text{tr}(Q\rho')} \right\} \\ &= e^{-D_{\text{H}}^{r,\eta}(\rho' \parallel \Gamma')}. \end{aligned} \quad (\text{D12})$$

One obtains Eq. (D11b) by setting  $\Gamma = \Gamma' = \mathbb{1}$  in Eq. (D11a). ■

For tensor-product states, the complexity entropy has a property similar to the von Neumann entropy's subadditivity.

*Proposition 15 (Subadditivity for tensor-product states).* Let  $S$  and  $S'$  denote distinct quantum systems. Let  $\rho_S$  and  $\rho_{S'}$  denote any subnormalized states, and let  $\Gamma_S$  and  $\Gamma_{S'}$  denote any positive-semidefinite operators. Let  $r, r' \geq 0$ ,  $\eta \in (0, \text{tr}(\rho)]$ , and  $\eta' \in (0, \text{tr}(\rho'))$ . It holds that

$$D_{\text{H}}^{r+r',\eta+\eta'}(\rho \otimes \rho' \parallel \Gamma \otimes \Gamma') \geq D_{\text{H}}^{r,\eta}(\rho \parallel \Gamma) + D_{\text{H}}^{r',\eta'}(\rho' \parallel \Gamma') \quad (\text{D13a})$$

$$\text{and } H_{\text{H}}^{r+r',\eta+\eta'}(\rho \otimes \rho') \leq H_{\text{H}}^{r,\eta}(\rho) + H_{\text{H}}^{r',\eta'}(\rho'). \quad (\text{D13b})$$

A consequence of Proposition 15 is, ancillas cannot decrease the complexity relative entropy. For all subnormalized states  $\rho$  and  $\sigma$  defined on the same Hilbert space, and for all subnormalized states  $\tau$ ,

$$D_{\text{H}}^{r,\eta}(\rho \otimes \tau \parallel \sigma \otimes \tau) \geq D_{\text{H}}^{r,\eta}(\rho \parallel \sigma). \quad (\text{D14})$$

Inequality (D14) follows from setting  $r' = 0$  and  $\eta' = 1$  in Eq. (D13a) and applying Proposition 9 [ $D_{\text{H}}^{0,1}(\tau \parallel \tau) = 0$ ]. Inequality (D14) is compatible with observations about the power of the one-clean-qubit computational model (DQC1) [133] and about how tossing an extra pure qubit into a black hole is expected to decrease the black hole's complexity [3,4].

*Proof of Proposition 15.* Consider any  $\zeta, \zeta' > 0$ . By the definitions of  $D_{\text{H}}^{r,\eta}(\rho \parallel \Gamma)$  and  $D_{\text{H}}^{r',\eta'}(\rho' \parallel \Gamma')$ , there exist  $Q \in \mathbf{M}_S^r$ , with  $\text{tr}(Q\rho) \geq \eta$ , and  $Q' \in \mathbf{M}_{S'}^{r'}$ , with  $\text{tr}(Q'\rho') \geq \eta'$ , such that

$$\begin{aligned} \frac{\text{tr}(Q\Gamma)}{\text{tr}(Q\rho)} &\leq e^{-D_{\text{H}}^{r,\eta}(\rho \parallel \Gamma)} + \zeta \quad \text{and} \quad \frac{\text{tr}(Q'\Gamma')}{\text{tr}(Q'\rho')} \\ &\leq e^{-D_{\text{H}}^{r',\eta'}(\rho' \parallel \Gamma')} + \zeta'. \end{aligned} \quad (\text{D15})$$

By Definition 6,  $\mathbf{M}_S^r \otimes \mathbf{M}_{S'}^{r'} \subseteq \mathbf{M}_{SS'}^{r+r'}$ , so  $Q \otimes Q' \in \mathbf{M}_{SS'}^{r+r'}$ . Moreover,

$$\text{tr}([Q \otimes Q'][\rho \otimes \rho']) = \text{tr}(Q\rho) \text{tr}(Q'\rho') \geq \eta\eta'. \quad (\text{D16})$$

Therefore,  $Q \otimes Q'$  is a candidate for the  $D_{\text{H}}^{r+r',\eta+\eta'}(\rho \otimes \rho' \parallel \Gamma \otimes \Gamma')$  optimization. Consequently,

$$\begin{aligned} e^{-D_{\text{H}}^{r+r',\eta+\eta'}(\rho \otimes \rho' \parallel \Gamma \otimes \Gamma')} &\leq \frac{\text{tr}(Q\Gamma) \text{tr}(Q'\Gamma')}{\text{tr}(Q\rho) \text{tr}(Q'\rho')} \\ &\leq \left( e^{-D_{\text{H}}^{r,\eta}(\rho \parallel \Gamma)} + \zeta \right) \left( e^{-D_{\text{H}}^{r',\eta'}(\rho' \parallel \Gamma')} + \zeta' \right), \end{aligned} \quad (\text{D17})$$

which implies Eq. (D13a), since  $\zeta$  and  $\zeta'$  are arbitrary. One obtains Eq. (D13b) by setting  $\Gamma = \mathbb{1}_S$  and  $\Gamma' = \mathbb{1}_{S'}$  in Eq. (D13a). ■

The complexity relative entropy never increases under any partial trace.

*Proposition 16 (Monotonicity under partial traces).* Let  $A$  and  $B$  denote distinct quantum systems. Let  $\rho_{AB}$  denote any subnormalized quantum state of  $AB$ . Let  $\Gamma_{AB}$  denote any positive-semidefinite operator. Let  $r \geq 0$  and  $\eta \in (0, \text{tr}(\rho)]$ . It holds that

$$D_H^{r,\eta}(\rho_{AB} \parallel \Gamma_{AB}) \geq D_H^{r,\eta}(\rho_A \parallel \Gamma_A), \quad (\text{D18})$$

wherein  $\rho_A := \text{tr}_B(\rho_{AB})$  and  $\Gamma_A := \text{tr}_B(\Gamma_{AB})$ . Furthermore,

$$H_H^{r,\eta}(\rho_{AB}) \leq H_H^{r,\eta}(\rho_A) + \log(d_B). \quad (\text{D19})$$

*Proof.* Consider any  $\zeta > 0$ . There exists a  $Q_A \in \mathcal{M}_A^r$  such that  $\text{tr}(Q_A \rho_A) \geq \eta$  and  $\text{tr}(Q_A \Gamma_A) / \text{tr}(Q_A \rho_A) \leq \exp(-D_H^{r,\eta}(\rho_A \parallel \Gamma_A)) + \zeta$ . By Properties (i) and (iii) of Definition 6,  $\tilde{Q}_{AB} := Q_A \otimes \mathbb{1}_B \in \mathcal{M}_{AB}^r$ . Moreover,  $\text{tr}(\tilde{Q}_{AB} \rho_{AB}) = \text{tr}(Q_A \rho_A) \geq \eta$ . Therefore,  $\tilde{Q}_{AB}$  is a candidate for the  $D_H^{r,\eta}(\rho_{AB} \parallel \Gamma_{AB})$  optimization. Consequently,

$$\begin{aligned} e^{-D_H^{r,\eta}(\rho_{AB} \parallel \Gamma_{AB})} &\leq \frac{\text{tr}(\tilde{Q}_{AB} \Gamma_{AB})}{\text{tr}(\tilde{Q}_{AB} \rho_{AB})} = \frac{\text{tr}(Q_A \Gamma_A)}{\text{tr}(Q_A \rho_A)} \\ &\leq e^{-D_H^{r,\eta}(\rho_A \parallel \Gamma_A)} + \zeta, \end{aligned} \quad (\text{D20})$$

which implies (D18), since  $\zeta$  is arbitrary. Equation (D19) follows from applying first Eq. (D18) and then Proposition 13, to get

$$\begin{aligned} H_H^{r,\eta}(\rho_{AB}) &= -D_H^{r,\eta}(\rho_{AB} \parallel \mathbb{1}_{AB}) \leq -D_H^{r,\eta}(\rho_A \parallel d_B \mathbb{1}_A) \\ &= H_H^{r,\eta}(\rho_A) + \log(d_B). \end{aligned} \quad (\text{D21})$$

■

We can bound the complexity (relative) entropy in some cases where its argument undergoes a limited-complexity unitary.

*Proposition 17 (Unitary operations on arguments).* Let  $\rho$  denote any subnormalized state; and  $\Gamma$ , any positive-semidefinite operator. Let  $\mathcal{P}$  denote any set of simple POVM effects (Definition 7) and  $C$  any adjoint-invariant unitary-complexity measure (Definition 3). Let  $\{\mathcal{M}^r = \mathcal{M}^r(\mathcal{P}, C)\}$  denote the family of POVM-effect-complexity sets defined by Eq. (B15). Let  $r, r' \geq 0$  and  $\eta \in (0, \text{tr}(\rho)]$ . Let  $U$  denote any unitary satisfying  $C(U) \leq r'$ . It holds that

$$D_H^{r+r',\eta}(U\rho U^\dagger \parallel U\Gamma U^\dagger) \geq D_H^{r,\eta}(\rho \parallel \Gamma) \quad (\text{D22a})$$

$$\text{and } H_H^{r+r',\eta}(U\rho U^\dagger) \leq H_H^{r,\eta}(\rho). \quad (\text{D22b})$$

*Proof.* Consider any  $\zeta > 0$ . There exists a  $Q \in \mathcal{M}^r$  such that  $\text{tr}(Q\rho) \geq \eta$  and  $\text{tr}(Q\Gamma) / \text{tr}(Q\rho) \leq \exp(-D_H^{r,\eta}(\rho \parallel \Gamma)) + \zeta$ .  $Q = U_0^\dagger P U_0$  for some effect  $P \in \mathcal{P}$  and for some unitary  $U_0$  satisfying  $C(U_0) \leq r$ . Let  $Q' := UQ U^\dagger = (U_0 U^\dagger)^\dagger P (U_0 U^\dagger)$ . By Definition 1 and by the adjoint invariance of  $C$ ,  $C(U_0 U^\dagger) \leq C(U_0) + C(U^\dagger) = C(U_0) + C(U) \leq r + r'$ , so  $Q' \in \mathcal{M}^{r+r'}$ . Moreover,  $\text{tr}(Q'[U\rho U^\dagger]) = \text{tr}(Q\rho) \geq \eta$ . Therefore,  $Q'$  is a candidate for the  $D_H^{r+r',\eta}(U\rho U^\dagger \parallel U\Gamma U^\dagger)$  optimization, so

$$\begin{aligned} e^{-D_H^{r+r',\eta}(U\rho U^\dagger \parallel U\Gamma U^\dagger)} &\leq \frac{\text{tr}(Q'[U\Gamma U^\dagger])}{\text{tr}(Q'[U\rho U^\dagger])} = \frac{\text{tr}(Q\Gamma)}{\text{tr}(Q\rho)} \\ &\leq e^{-D_H^{r,\eta}(\rho \parallel \Gamma)} + \zeta, \end{aligned} \quad (\text{D23})$$

which implies Eq. (D22a), since  $\zeta$  is arbitrary. ■

For composite systems, one can bound the complexity (relative) entropy in terms of hypothesis-testing (relative) entropies on each subsystem, if  $\mathcal{M}^r$  contains only tensor-product POVM effects. In most cases,  $\mathcal{M}^r$  contains nonlocal effects for all  $r > 0$ ; in such cases, only  $\mathcal{M}^{r=0}$  may consist solely of tensor-product effects. Importantly, the set  $\mathcal{M}^{r=0}$  defined in Eq. (10), for  $n$  qubits, contains only tensor-product effects.

*Proposition 18 (Upper bound by hypothesis-testing relative entropies of subsystems).* Let  $S$  denote a composition of  $N$  quantum subsystems:  $S = S_1 S_2 \dots S_N$ . Let  $\rho_S$  denote any subnormalized state. For all  $i = 1, 2, \dots, N$ , let  $\Gamma_{S_i}$  denote any positive-semidefinite operator. Let  $\Gamma_S := \bigotimes_{i=1}^N \Gamma_{S_i}$ . Let  $r \geq 0$  and  $\eta \in (0, \text{tr}(\rho_S)]$ . Suppose that  $\mathcal{M}_S^r$  contains only tensor-product POVM effects: every  $Q_S \in \mathcal{M}_S^r$  is of the form  $\bigotimes_{i=1}^N Q_{S_i}$ . Each  $Q_{S_i}$  is an effect on  $S_i$ . It holds that

$$D_H^{r,\eta}(\rho_S \parallel \Gamma_S) \leq \sum_{i=1}^N D_H^\eta(\rho_{S_i} \parallel \Gamma_{S_i}) \quad (\text{D24a})$$

$$\text{and } H_H^{r,\eta}(\rho_S) \geq \sum_{i=1}^N H_H^\eta(\rho_{S_i}). \quad (\text{D24b})$$

*Proof.* Consider any  $\zeta > 0$ . There exists a  $Q_S = \bigotimes_{i=1}^N Q_{S_i} \in \mathcal{M}_S^r$ , with each  $Q_{S_i}$  an effect on  $S_i$ , such that  $\text{tr}(Q_S \rho_S) \geq \eta$  and  $\text{tr}(Q_S \Gamma_S) / \text{tr}(Q_S \rho_S) \leq \exp(-D_H^{r,\eta}(\rho_S \parallel \Gamma_S)) + \zeta$ . For all  $i = 1, 2, \dots, N$ , let  $Q_S^{(i)} := (\mathbb{1}_{S_1} \otimes \dots \otimes \mathbb{1}_{S_{i-1}}) \otimes Q_{S_i} \otimes (\mathbb{1}_{S_{i+1}} \otimes \dots \otimes \mathbb{1}_{S_N})$ . For all  $i$ ,  $Q_S^{(i)} \geq Q_S$ , since  $\mathbb{1}_{S_j} \geq Q_{S_j}$  for all  $j \neq i$ ; consequently,  $\text{tr}(Q_S^{(i)} \rho_S) = \text{tr}(Q_S^{(i)} \rho_S) \geq \text{tr}(Q_S \rho_S) \geq \eta$ . Thus, for all  $i$ ,  $Q_S^{(i)}$  is a candidate for the  $D_H^\eta(\rho_{S_i} \parallel \Gamma_{S_i})$  optimization, so

$\exp(-D_H^\eta(\rho_{S_i} \parallel \Gamma_{S_i})) \leq \text{tr}(Q_{S_i} \Gamma_{S_i})$ . Hence,

$$\begin{aligned} e^{-\sum_{i=1}^N D_H^\eta(\rho_{S_i} \parallel \Gamma_{S_i})} &\leq \prod_{i=1}^N \frac{\text{tr}(Q_{S_i} \Gamma_{S_i})}{\text{tr}(Q_{S_i} \rho_{S_i})} \leq \prod_{i=1}^N \frac{\text{tr}(Q_{S_i} \Gamma_{S_i})}{\text{tr}(Q_S \rho_S)} \\ &= \frac{\text{tr}(Q_S \Gamma_S)}{\text{tr}(Q_S \rho_S)} \leq e^{-D_H^{r,\eta}(\rho_S \parallel \Gamma_S)} + \zeta, \end{aligned} \quad (\text{D25})$$

which implies Eq. (D24a), since  $\zeta$  is arbitrary. ■

### 3. Complexity relative entropy and hypothesis testing

We prove Proposition 1 of the main text, quantifying the type-I and type-II errors of a hypothesis test that involves limited-complexity measurement effects.

*Proposition 19 (Hypothesis testing with complexity limitations).* Let  $\rho$  and  $\sigma$  denote any quantum states. Let  $r \geq 0$ ,  $\eta \in (0, 1]$ , and  $\delta \in (0, 1]$ . The following statements are equivalent:

- (i) For all  $\zeta > 0$ , there exist a  $Q \in \mathcal{M}^r$  and a  $q \in [\eta, 1]$  such that  $\text{tr}([qQ]\rho) = \eta$  and  $\text{tr}([qQ]\sigma) \leq \delta + \zeta$ .
- (ii) It holds that

$$D_H^{r,\eta}(\rho \parallel \sigma) \geq -\log\left(\frac{\delta}{\eta}\right). \quad (\text{D26})$$

Consequently, there exist a  $Q \in \mathcal{M}^r$  and a  $q \in (0, 1]$  such that  $\text{tr}([qQ]\rho) = \eta$  and  $\text{tr}([qQ]\sigma) < \delta$  if and only if

$$D_H^{r,\eta}(\rho \parallel \sigma) > -\log\left(\frac{\delta}{\eta}\right). \quad (\text{D27})$$

*Corollary 2.* Let  $r \geq 0$ . If  $\mathcal{M}^r$  is compact, then statements (i) and (ii) in Proposition 19 are equivalent to the following:

- (i) There exist a  $Q \in \mathcal{M}^r$  and a  $q \in [\eta, 1]$  such that  $\text{tr}([qQ]\rho) = \eta$  and  $\text{tr}([qQ]\sigma) \leq \delta$ .

In particular, suppose that  $\mathcal{P}$  is a set of simple POVM effects (Definition 7) that is compact, and suppose that  $C_{\mathcal{G}}$  is a circuit-superoperator-complexity measure (Definition 2) associated with a compact gate set  $\mathcal{G}$ . Then  $\mathcal{M}^r(\mathcal{P}, C_{\mathcal{G}})$  (Definition 8) is compact.

In particular, the sets  $\mathcal{M}^r$  defined in Eq. (11) are compact.

*Proof.* To prove the first part of the proposition, consider any  $\zeta > 0$ . Suppose there exist a  $Q \in \mathcal{M}^r$  and a  $q \in [\eta, 1]$  such that  $\text{tr}([qQ]\rho) = \eta$  and  $\text{tr}([qQ]\sigma) \leq \delta + \zeta$ .

Then  $\text{tr}(Q\rho) \geq \text{tr}([qQ]\rho) = \eta$ , and

$$e^{-D_H^{r,\eta}(\rho \parallel \sigma)} \leq \frac{\text{tr}(Q\sigma)}{\text{tr}(Q\rho)} = \frac{\text{tr}([qQ]\sigma)}{\text{tr}([qQ]\rho)} \leq \frac{\delta + \zeta}{\eta}. \quad (\text{D28})$$

Therefore,  $e^{-D_H^{r,\eta}(\rho \parallel \sigma)} \leq \delta/\eta$ , since  $\zeta$  is arbitrary. Conversely, suppose that  $\exp(-D_H^{r,\eta}(\rho \parallel \sigma)) \leq \delta/\eta$ . Consider again any  $\zeta > 0$ . There exists a  $Q \in \mathcal{M}^r$  such that  $\text{tr}(Q\rho) \geq \eta$  and  $\text{tr}(Q\sigma)/\text{tr}(Q\rho) \leq (\delta + \zeta)/\eta$ . Let  $q := \eta/\text{tr}(Q\rho)$ .  $q \in [\eta, 1]$ , since  $\eta \leq \text{tr}(Q\rho) \leq 1$ . Then  $\text{tr}([qQ]\rho) = \eta$ , and

$$\begin{aligned} \text{tr}([qQ]\sigma) &= \text{tr}([qQ]\rho) \cdot \frac{\text{tr}(Q\sigma)}{\text{tr}(Q\rho)} \leq \text{tr}([qQ]\rho) \cdot \frac{\delta + \zeta}{\eta} \\ &= \delta + \zeta. \end{aligned} \quad (\text{D29})$$

We now prove the second part of the proposition. Suppose there exist a  $Q \in \mathcal{M}^r$  and a  $q \in (0, 1]$  such that  $\text{tr}([qQ]\rho) = \eta$  and  $\text{tr}([qQ]\sigma) < \delta$ . Let  $\delta' := \text{tr}([qQ]\sigma) < \delta$ .  $Q$ ,  $q$ , and  $\delta'$  satisfy the conditions equivalent to Eq. (D26), so  $\exp(-D_H^{r,\eta}(\rho \parallel \sigma)) \leq \delta'/\eta < \delta/\eta$ . Conversely, suppose that  $\exp(-D_H^{r,\eta}(\rho \parallel \sigma)) < \delta/\eta$ . Let  $\delta'' := \eta \exp(-D_H^{r,\eta}(\rho \parallel \sigma)) < \delta$ . Then  $D_H^{r,\eta}(\rho \parallel \sigma) \leq -\log(\delta''/\eta)$ , as in Eq. (D26), so there exist  $Q \in \mathcal{M}^r$  and  $q \in (0, 1]$  such that  $\text{tr}([qQ]\rho) = \eta$  and  $\text{tr}([qQ]\sigma) \leq \delta'' + \zeta$ . By considering  $\zeta$  sufficiently small such that  $\delta'' + \zeta < \delta$ , one obtains  $\text{tr}([qQ]\sigma) < \delta$ . ■

*Proof of Corollary 2.* The implication (iii)  $\Rightarrow$  (i) is straightforward; we need to show that (i)  $\Rightarrow$  (iii). Assume (i). Consider the continuous functions  $f, g : [\eta, 1] \times \mathcal{M}^r \rightarrow [0, 1]$  defined as  $f(q, Q) := \text{tr}([qQ]\rho)$  and  $g(q, Q) := \text{tr}([qQ]\sigma)$ . Let  $A$  denote the preimage of  $\eta$  under  $f$ :  $A = f^{-1}(\{\eta\})$ .  $A$  is nonempty, since, by assumption, (i) holds.  $A$  is closed, since  $f$  is continuous and  $\{\eta\}$  is closed. Therefore,  $A$  is compact, being a closed subspace of the compact set  $[\eta, 1] \times \mathcal{M}^r$ . Let

$$v := \inf\{g(A)\} = \inf\{\text{tr}([qQ]\sigma) : (q, Q) \in A\}. \quad (\text{D30})$$

$g(A)$  is compact, since it is the image of a compact set under a continuous function. Therefore, the infimum is attained, and there exists a  $(q, Q) \in A$  such that  $g(q, Q) = v$ . Since (i) holds,  $v \leq \delta + \zeta$  for all  $\zeta > 0$ ; hence,  $v \leq \delta$ . Thus, there exists  $(q, Q) \in A$  such that  $g(q, Q) \leq \delta$ , so (iii) holds.

Finally, suppose that  $\mathcal{P}$  and  $\mathcal{G}$  are compact. Consider the continuous function  $h(P, \mathcal{E}_1, \dots, \mathcal{E}_r) := (\mathcal{E}_1^\dagger \dots \mathcal{E}_r^\dagger)(P)$ .  $\mathcal{M}^r(\mathcal{P}, C_{\mathcal{G}})$  is compact, since it is the image of a compact set under a continuous function. ■

### 4. Complexity (relative) entropy and state-complexity measures

The complexity entropy relates to pure-state complexity as follows.

*Proposition 20 (Relation between complexity entropy and pure-state complexity).* Let  $\mathbf{P}$  denote any set of simple POVM effects (Definition 7); and  $C$ , any adjoint-invariant unitary-complexity measure (Definition 3). Let  $r \geq 0$ , and let  $\mathbf{M}^r = \mathbf{M}^r(\mathbf{P}, C)$  denote the POVM-effect-complexity set defined in Eq. (B15). Let  $\delta \geq 0$ , and let  $|\psi_0\rangle$  denote any pure reference state such that  $\psi_0 \in \mathbf{P}$ . Let  $C^{\psi_0, \delta}$  denote the  $\delta$ -approximate-pure-state-complexity measure with respect to  $C$  and  $\psi_0$  (Definition 5). Let  $|\psi\rangle$  denote any pure state. The following statements hold:

(i) Suppose that  $\delta < 1$  and  $r > C^{\psi_0, \delta}(\psi)$ . Then

$$H_{\mathbf{H}}^{r, 1-\delta^2}(\psi) \leq \log\left(\frac{1}{1-\delta^2}\right). \quad (\text{D31})$$

(ii) Let  $r_0 \geq 0$  and  $\varepsilon \in [0, 1)$ . Suppose that  $H_{\mathbf{H}}^{r, 1-\varepsilon^2}(\rho) \leq \log(1/[1-\varepsilon^2])$ . Then

$$C^{\psi_0, \delta}(\psi) \leq r + r_0, \quad (\text{D32})$$

where

$$\begin{aligned} \delta &= \varepsilon + \sqrt{\alpha \varepsilon^2 + (1-\alpha)}; \\ \alpha &:= \lim_{\zeta \rightarrow 0^+} \inf_{P \in \mathbf{P}} \frac{\text{tr}(P^2)}{[\text{tr}(P)]^2} \geq 1-8\varepsilon^2-\zeta \\ &\times \left\{ \sup_{\substack{U_0: \\ C(U_0) \leq r_0}} \left\{ \langle \psi_0 | U_0 P U_0^\dagger | \psi_0 \rangle \right\} \right\}. \end{aligned} \quad (\text{D33})$$

The error parameter  $\alpha$  is necessary for the second statement of the proposition.  $\alpha$  assumes simple values for suitable choices of  $\mathbf{P}$ ,  $C$ ,  $|\psi_0\rangle$ , and  $r_0$ . For instance, consider an  $n$ -qubit system, and let  $|\psi_0\rangle = |0^n\rangle$ . Suppose that  $\mathbf{P}$  equals the set of single-qubit projectors in Eq. (10). For every  $P \in \mathbf{P}$ , the supremum in Eq. (D33) achieves its maximum value with  $U_0 = \mathbb{1}$ :  $\langle 0^n | U_0 P U_0^\dagger | 0^n \rangle = \langle 0^n | 0^n \rangle = 1$ . [For any  $C$  and any  $r_0$ ,  $C(\mathbb{1}) = 0 \leq r_0$ .] Having unit purity,  $P = |0^n\rangle\langle 0^n|$  always satisfies the constraint on the infimum in Eq. (D33), so  $\alpha = 1$ , and  $\delta = 2\varepsilon$ . Now, suppose that  $\mathbf{P}$  equals the set of tensor products of all single-qubit projectors. Suppose further that  $C$  assigns the value 1 (0) to every single-qubit unitary and that  $r_0 = n$  ( $r_0 = 0$ ). Then, for all  $P \in \mathbf{P}$ , there exists a tensor product  $U_0$  of single-qubit unitaries such that  $\langle \psi_0 | U_0 P U_0^\dagger | \psi_0 \rangle = 1$ .  $U_0$  satisfies  $C(U_0) \leq r_0$ . Having unit purity, every rank-1 projector in  $\mathbf{P}$  always satisfies the constraint on the infimum in Eq. (D33); so, again,  $\alpha = 1$ , and  $\delta = 2\varepsilon$ .

We now relate the complexity relative entropy to the strong complexity of Ref. [17]. Let

$$\beta^r(\rho, \sigma) := \max_{M \in \mathbf{M}^r} \{|\text{tr}(M[\rho - \sigma])|\}. \quad (\text{D34})$$

Let  $\delta \geq 0$ , and let  $|\psi\rangle$  denote a pure state in a Hilbert space of dimensionality  $d$ . The *strong complexity* of  $|\psi\rangle$  is defined as

$$C_{\text{strong}}^\delta(|\psi\rangle) := \inf \left\{ r \geq 0 : \beta^r(\psi, \pi) \geq 1 - \frac{1}{d} - \delta \right\}. \quad (\text{D35})$$

The  $\pi := \mathbb{1}/d$  denotes the maximally mixed state. Our definition of  $\beta^r(\psi, \pi)$  differs slightly from that in Ref. [17], wherein  $\mathbf{M}^r$  can be defined using an auxiliary system.

The following proposition presents a bound relating the complexity entropy and the strong complexity.

*Proposition 21 (Complexity entropy and strong complexity).* Let  $\rho$  and  $\sigma$  denote any quantum states. Let  $r \geq 0$  and  $\eta \in (0, 1]$ . It holds that

$$\begin{aligned} D_{\mathbf{H}}^{r, \eta}(\rho \parallel \sigma) &\leq -\log\left(1 - \frac{\beta^r(\rho, \sigma)}{\eta}\right) \\ &= \frac{\beta^r(\rho, \sigma)}{\eta} + O\left[\left(\frac{\beta^r(\rho, \sigma)}{\eta}\right)^2\right]. \end{aligned} \quad (\text{D36})$$

Furthermore, let  $\delta \geq 0$ , and let  $|\psi\rangle$  denote any pure state in a Hilbert space of dimensionality  $d$ . Suppose that  $r < C_{\text{strong}}^\delta(|\psi\rangle)$  and  $\eta > 1 - d^{-1} - \delta$ . Then

$$\begin{aligned} H_{\mathbf{H}}^{r, \eta}(\psi) &> \log(d) - \log\left(\frac{1}{1-c}\right); \\ c &:= \frac{1 - d^{-1} - \delta}{\eta} < 1. \end{aligned} \quad (\text{D37})$$

We now prove Propositions 20 and 21.

*Proof of Proposition 20.* We prove the first statement. Let  $\zeta := r - C^{\psi_0, \delta}(\psi) > 0$ . There exists a unitary  $U$  such that  $\frac{1}{2}\|U\psi_0 U^\dagger - \psi\|_1 \leq \delta$  and  $C(U) \leq r = C^{\psi_0, \delta}(\psi) + \zeta$ . Let  $Q := U\psi_0 U^\dagger$ .  $Q \in \mathbf{M}^r$ , since  $\psi_0 \in \mathbf{P}$  and  $C(U^\dagger) = C(U) \leq r$ , by the adjoint-invariance of  $C$ . Moreover,

$$\begin{aligned} \text{tr}(Q\psi) &= \text{tr}(U\psi_0 U^\dagger \psi) = |\langle \psi | U | \psi_0 \rangle|^2 \\ &= 1 - \left(\frac{1}{2}\|U\psi_0 U^\dagger - \psi\|_1\right)^2 \geq 1 - \delta^2. \end{aligned} \quad (\text{D38})$$

The last equality expresses a general relation for the trace norm [95]: for any two pure states  $|\phi\rangle$  and  $|\varphi\rangle$ ,



$\frac{1}{2}\|\phi - \varphi\|_1 = \sqrt{1 - |\langle \phi | \varphi \rangle|^2}$ . Thus,  $Q$  is a candidate for the  $H_H^{r, 1-\delta^2}(\psi)$  optimization, so

$$H_H^{r, 1-\delta^2}(\psi) \leq \log\left(\frac{\text{tr}(Q)}{\text{tr}(Q\psi)}\right) \leq \log\left(\frac{1}{1-\delta^2}\right). \quad (\text{D39})$$

We now prove the second statement. Suppose that  $H_H^{r, 1-\varepsilon^2}(\psi) \leq \log(1/[1-\varepsilon^2])$ . Let  $\zeta > 0$ . There exists a  $Q \in \mathbf{M}^r$  such that  $\text{tr}(Q\psi) \geq 1 - \varepsilon^2$  and

$$\frac{\text{tr}(Q)}{\text{tr}(Q\psi)} \leq e^{H_H^{r, 1-\varepsilon^2}(\psi)} + \zeta \leq \frac{1}{1-\varepsilon^2} + \zeta. \quad (\text{D40})$$

$Q = U^\dagger P U$  for some effect  $P \in \mathbf{P}$  and for some unitary  $U$  satisfying  $C(U) \leq r$ . Let  $|\psi'\rangle := U|\psi\rangle$ . It holds that

$$\begin{aligned} F^2\left(\psi', \frac{P}{\text{tr}(P)}\right) &= \langle \psi' | \frac{P}{\text{tr}(P)} | \psi' \rangle = \frac{\text{tr}(P\psi')}{\text{tr}(P)} \\ &= \left[ \frac{\text{tr}(Q)}{\text{tr}(Q\psi)} \right]^{-1} \geq 1 - \varepsilon^2 - \zeta_1. \end{aligned} \quad (\text{D41})$$

Here and in the following, each  $\zeta_i \equiv \zeta_i(\zeta) > 0$  is an error parameter such that  $\lim_{\zeta \rightarrow 0^+} \zeta_i = 0$ . By a Fuchs-van de Graaf inequality [134],

$$\frac{1}{2} \left\| \psi' - \frac{P}{\text{tr}(P)} \right\|_1 \leq \sqrt{1 - F^2\left(\psi', \frac{P}{\text{tr}(P)}\right)} \leq \sqrt{\varepsilon^2 + \zeta_1}. \quad (\text{D42})$$

Let

$$\alpha_P := \sup_{\substack{U_0: \\ C(U_0) \leq r_0}} \left\{ \langle \psi_0 | U_0 P U_0^\dagger | \psi_0 \rangle \right\}. \quad (\text{D43})$$

There exists a unitary  $U_0$  such that  $C(U_0) \leq r_0$  and

$$\langle \psi_0 | U_0 P U_0^\dagger | \psi_0 \rangle \geq \alpha_P - \zeta. \quad (\text{D44})$$

Consequently,

$$\begin{aligned} F^2\left(U_0^\dagger \psi_0 U_0, \frac{P}{\text{tr}(P)}\right) &= \langle \psi_0 | U_0 \frac{P}{\text{tr}(P)} U_0^\dagger | \psi_0 \rangle \geq \frac{\alpha_P - \zeta}{\text{tr}(P)} \\ &\geq (\alpha_P - \zeta)(1 - \varepsilon^2 - \zeta_1) = \alpha_P(1 - \varepsilon^2) - \zeta_2. \end{aligned} \quad (\text{D45})$$

The second inequality follows from Eq. (D41)

$$\frac{1}{\text{tr}(P)} \geq \frac{\text{tr}(P\psi')}{\text{tr}(P)} \geq 1 - \varepsilon^2 - \zeta_1. \quad (\text{D46})$$

The largest eigenvalue of  $P$ ,  $\|P\|$ , satisfies  $\|P\| \geq \langle \psi' | P | \psi' \rangle = \langle \psi | Q | \psi \rangle \geq 1 - \varepsilon^2$ , so  $\text{tr}(P^2) \geq \|P\|^2 =$

$\|P\|^2 \geq (1 - \varepsilon^2)^2$ , because  $P$  is positive semidefinite. Hence,

$$\begin{aligned} \frac{\text{tr}(P^2)}{[\text{tr}(P)]^2} &\geq (1 - \varepsilon^2)^2 (1 - \varepsilon^2 - \zeta_1)^2 = (1 - \varepsilon^2)^4 - \zeta_3 \\ &\geq 1 - 4\varepsilon^2 - 4(\varepsilon^2)^3 - \zeta_3 \geq 1 - 8\varepsilon^2 - \zeta_3, \end{aligned} \quad (\text{D47})$$

so

$$\alpha_P \geq \alpha_{\zeta_3} := \inf_{\substack{P \in \mathbf{P} \\ \frac{\text{tr}(P^2)}{[\text{tr}(P)]^2} \geq 1 - 8\varepsilon^2 - \zeta_3}} \{\alpha_P\}. \quad (\text{D48})$$

Thus,

$$F^2\left(U_0^\dagger \psi_0 U_0, \frac{P}{\text{tr}(P)}\right) \geq \alpha_{\zeta_3}(1 - \varepsilon^2) - \zeta_2. \quad (\text{D49})$$

By a Fuchs-van de Graaf inequality,

$$\begin{aligned} \frac{1}{2} \left\| U_0^\dagger \psi_0 U_0 - \frac{P}{\text{tr}(P)} \right\|_1 &\leq \sqrt{1 - F^2\left(U_0^\dagger \psi_0 U_0, \frac{P}{\text{tr}(P)}\right)} \\ &\leq \sqrt{1 - \alpha_{\zeta_3}(1 - \varepsilon^2) - \zeta_2}. \end{aligned} \quad (\text{D50})$$

Since the trace norm is unitarily invariant and obeys the triangle inequality, Eqs. (D42) and (D50) imply that

$$\begin{aligned} \frac{1}{2} \left\| \psi - U^\dagger U_0^\dagger \psi_0 U_0 U^\dagger \right\|_1 &= \frac{1}{2} \left\| U \psi U^\dagger - U_0^\dagger \psi_0 U_0 \right\|_1 \\ &\leq \sqrt{\varepsilon^2 + \zeta_1} + \sqrt{1 - \alpha_{\zeta_3}(1 - \varepsilon^2) - \zeta_2} =: \delta_\zeta. \end{aligned} \quad (\text{D51})$$

By the adjoint invariance of  $C$ ,  $C(U^\dagger U_0^\dagger) \leq C(U^\dagger) + C(U_0^\dagger) = C(U) + C(U_0) \leq r + r_0$ , so

$$C^{\psi_0, \delta_\zeta}(\psi) \leq r + r_0. \quad (\text{D52})$$

Since  $\lim_{\zeta \rightarrow 0^+} \delta_\zeta = \delta$  and since  $C$  is right-continuous [as per Eq. (B4)],

$$C^{\psi_0, \delta}(\psi) = \lim_{\zeta \rightarrow 0^+} C^{\psi_0, \delta_\zeta}(\psi) \leq r + r_0. \quad (\text{D53})$$

■

*Proof of Proposition 21.* Consider any  $\zeta > 0$ . There exists a  $Q \in \mathbf{M}^r$  such that  $\text{tr}(Q\rho) \geq \eta$  and  $\text{tr}(Q\sigma)/\text{tr}(Q\rho) \leq$

$\exp(-D_{\text{H}}^{r,\eta}(\rho \parallel \sigma)) + \zeta$ . It holds that

$$\begin{aligned} \frac{\text{tr}(Q\sigma)}{\text{tr}(Q\rho)} &= 1 - \frac{\text{tr}(Q[\rho - \sigma])}{\text{tr}(Q\rho)} \geq 1 - \frac{\beta^r(\rho, \sigma)}{\text{tr}(Q\rho)} \\ &\geq 1 - \frac{\beta^r(\rho, \sigma)}{\eta}, \end{aligned} \quad (\text{D54})$$

so

$$e^{-D_{\text{H}}^{r,\eta}(\rho \parallel \sigma)} \geq 1 - \frac{\beta^r(\rho, \sigma)}{\eta} + \zeta, \quad (\text{D55})$$

which implies Eq. (D36), since  $\zeta$  is arbitrary.

We now prove Eq. (D37). Suppose that  $r < C_{\text{strong}}^{\delta}(|\psi\rangle)$  and  $\eta > 1 - d^{-1} - \delta$ .  $r < C_{\text{strong}}^{\delta}(|\psi\rangle)$  implies that  $\beta^r(\psi, \pi) < 1 - d^{-1} - \delta < \eta$ , so

$$\begin{aligned} H_{\text{H}}^{r,\eta}(\psi) &= \log(d) - D_{\text{H}}^{r,\eta}(\psi \parallel \pi) \\ &\geq \log(d) + \log\left(1 - \frac{\beta^r(\psi, \pi)}{\eta}\right) \\ &> \log(d) + \log\left(1 - \frac{1 - d^{-1} - \delta}{\eta}\right) \\ &= \log(d) - \log\left(\frac{1}{1 - c}\right). \end{aligned} \quad (\text{D56})$$

■

## 5. Reduced complexity (relative) entropy and related quantities

### a. Definition of the reduced complexity (relative) entropy

Some situations call for a variant of the complexity (relative) entropy—a variant that lacks the denominator in Eq. (D2) [in Eq. (D1)]. We call these variants the *reduced complexity relative entropy* and the *reduced complexity entropy*. The variants lack a desirable property possessed by the original versions: the reduced complexity (relative) entropy may diverge as  $\eta \rightarrow 0$ , whereas the complexity (relative) entropy will not whenever  $\Gamma$  is positive definite (see Proposition 10). Nevertheless, the variants bear simpler definitions than the originals and more readily facilitate certain technical proofs. In any case, the reduced complexity (relative) entropy and the complexity (relative) entropy differ by at most  $\log(1/\eta)$  (Proposition 22). Therefore, the quantities are interchangeable if the error tolerance is insignificant:  $\eta \approx 1$ .

**Definition 15 (Reduced complexity relative entropy).** Let  $\rho$  denote any subnormalized state; and  $\Gamma$ , any positive-semidefinite operator. Let  $r \geq 0$  and  $\eta \in (0, \text{tr}(\rho)]$ . Here,  $\rho$ ,  $\Gamma$ , and every  $Q \in \mathbb{M}^r$  act on the same Hilbert space.

The *reduced complexity-restricted hypothesis-testing relative entropy*, or simply the *reduced complexity relative entropy*, is

$$D_{\text{h}}^{r,\eta}(\rho \parallel \Gamma) := -\log\left(\inf_{\substack{Q \in \mathbb{M}^r \\ \text{tr}(Q\rho) \geq \eta}} \{\text{tr}(Q\Gamma)\}\right). \quad (\text{D57})$$

**Definition 16 (Reduced complexity entropy).** Let  $\rho$  denote any subnormalized state. Let  $r \geq 0$  and  $\eta \in (0, \text{tr}(\rho)]$ . The *reduced complexity-restricted hypothesis-testing entropy*, or simply the *reduced complexity entropy*, is

$$H_{\text{h}}^{r,\eta}(\rho) := -D_{\text{h}}^{r,\eta}(\rho \parallel \mathbb{1}) = \log\left(\inf_{\substack{Q \in \mathbb{M}^r \\ \text{tr}(Q\rho) \geq \eta}} \{\text{tr}(Q)\}\right). \quad (\text{D58})$$

By the constraint  $\eta \leq \text{tr}(Q\rho) \leq 1$ , the reduced complexity (relative) entropy differs from the complexity (relative) entropy by at most  $\log(1/\eta)$ :

**Proposition 22 (Difference between  $D_{\text{h}}^{r,\eta}$  and  $D_{\text{H}}^{r,\eta}$ ).** Let  $\rho$  denote any subnormalized state; and  $\Gamma$ , any positive-semidefinite operator. Let  $r \geq 0$  and  $\eta \in (0, \text{tr}(\rho)]$ . It holds that

$$0 \leq D_{\text{h}}^{r,\eta}(\rho \parallel \Gamma) - D_{\text{H}}^{r,\eta}(\rho \parallel \Gamma) \leq \log\left(\frac{1}{\eta}\right) \quad (\text{D59a})$$

$$\text{and } 0 \leq H_{\text{H}}^{r,\eta}(\rho) - H_{\text{h}}^{r,\eta}(\rho) \leq \log\left(\frac{1}{\eta}\right). \quad (\text{D59b})$$

If  $\rho$  is normalized,  $D_{\text{h}}^{r,\eta=1}(\rho \parallel \Gamma) = D_{\text{H}}^{r,\eta=1}(\rho \parallel \Gamma)$ , and  $H_{\text{h}}^{r,\eta=1}(\rho) = H_{\text{H}}^{r,\eta=1}(\rho)$ .

**Proof.** The denominator  $\text{tr}(Q\rho)$  in Definition 13 is constrained to  $\eta \leq \text{tr}(Q\rho) \leq 1$ , leading to the claimed bounds. The second part of the proposition follows by setting  $\eta = 1$  in the bounds. ■

The reduced complexity (relative) entropy also obeys bounds similar to those in Proposition 10.

**Proposition 23 (General bounds for  $D_{\text{h}}^{r,\eta}$ ).** Let  $\rho$  denote any subnormalized state, and  $\Gamma$  any positive-semidefinite operator, that act on a Hilbert space of dimensionality  $d$ .

Let  $r \geq 0$  and  $\eta \in (0, \text{tr}(\rho)]$ . It holds that

$$D_h^{r,\eta}(\rho \parallel \Gamma) \geq -\log(\text{tr}(\Gamma)). \quad (\text{D60a})$$

Furthermore, if  $\Gamma$  is positive definite (has full rank),

$$D_h^{r,\eta}(\rho \parallel \Gamma) \leq \log(\|\Gamma^{-1}\|) + \log(\text{tr}(\rho)) + \log\left(\frac{1}{\eta}\right). \quad (\text{D60b})$$

In particular,

$$-\log(\text{tr}(\rho)) - \log\left(\frac{1}{\eta}\right) \leq H_h^{r,\eta}(\rho) \leq \log(d). \quad (\text{D61})$$

Consequently, if  $\rho$  is normalized and  $\Gamma$  is positive definite,

$$-\log(\text{tr}(\Gamma)) \leq D_h^{r,\eta}(\rho \parallel \Gamma) \leq \log(\|\Gamma^{-1}\|) + \log\left(\frac{1}{\eta}\right), \quad (\text{D62a})$$

$$\text{and } -\log\left(\frac{1}{\eta}\right) \leq H_h^{r,\eta}(\rho) \leq \log(d). \quad (\text{D62b})$$

*Proof.* Equation (D60a) follows because  $\mathbb{1} \in \mathbf{M}^r$  is a candidate effect for the  $D_h^{r,\eta}(\rho \parallel \Gamma)$  optimization. Equation (D60b) follows from combining Eq. (D4b) of Proposition 10 with Eq. (D59a) of Proposition 22. ■

Equation (D60a) implies that  $D_h^{r,\eta}(\rho \parallel \sigma) \geq 0$  for normalized states  $\rho$  and  $\sigma$ .

The reduced complexity (relative) entropy possesses several elementary properties of the complexity (relative) entropy.

**Proposition 24 (Properties of  $D_h^{r,\eta}$ ).** The following propositions hold if, everywhere therein, one replaces the complexity (relative) entropy with the reduced complexity (relative) entropy: Proposition 12, Proposition 13, Proposition 14, Proposition 15, Proposition 16, and Proposition 17.

*Proof.* Adapt the propositions' proofs as appropriate. ■

In general, the reduced complexity (relative) entropy lacks the properties stated in Proposition 9 and in Proposition 11. Proposition 9 does not hold because  $D_h^{r,\eta}(\rho \parallel \rho) > 0$  if  $\mathbf{M}^r = \{\mathbb{1}\}$  and  $\text{tr}(\mathbb{1}) > 1$ . Proposition 11 does not hold because  $H_h^{r,\eta}(\rho)$ , but not  $H_h^\eta(\rho)$ , can be negative.

### b. Complexity success probability

It is convenient to introduce a quantity closely related to the reduced complexity entropy. We call this quantity the *complexity success probability*. Given any subnormalized

state  $\rho$  and any  $r \geq 0$ , we have defined the reduced complexity entropy as the infimum of  $\log(\text{tr}(Q))$  over all  $Q \in \mathbf{M}^r$ , subject to a constraint (a lower bound) on  $\text{tr}(Q\rho)$ , the probability of successfully identifying  $\rho$ . Reciprocally, we define the complexity success probability as the supremum of  $\text{tr}(Q\rho)$  over all  $Q \in \mathbf{M}^r$ , subject to a constraint (an upper bound) on  $\log(\text{tr}(Q))$ .

**Definition 17 (Complexity success probability).** Let  $\rho$  denote any subnormalized state. Let  $r \geq 0$ . Let  $m_- := \inf_\eta \{H_h^{r,\eta}(\rho)\}$  and  $m_+ := \sup_\eta \{H_h^{r,\eta}(\rho)\} = H_h^{r,\eta=\text{tr}(\rho)}(\rho)$ , with  $\eta$  ranging over  $(0, \text{tr}(\rho)]$ . Let  $m \in [m_-, m_+]$  if  $m_-$  is the minimum of  $\{H_h^{r,\eta}(\rho)\}_\eta$ ; otherwise, let  $m \in (m_-, m_+]$ . The *reduced-complexity-entropy success probability*, or simply the *complexity success probability*, is

$$\eta_h^{r,m}(\rho) := \sup\{\eta : H_h^{r,\eta}(\rho) \leq m\} = \sup_{\substack{Q \in \mathbf{M}^r \\ \log(\text{tr}(Q)) \leq m}} \{\text{tr}(Q\rho)\}. \quad (\text{D63})$$

$\eta_h^{r,m}(\rho)$  appears in the proof of Proposition 30 (Appendix F).

$\eta_h^{r,m}(\rho)$  takes a simple form in the special case where  $\mathbf{M}^r$  is defined for  $n$  qubits, as in Eq. (11). Consider any candidate  $Q \in \mathbf{M}^r$ .  $Q = U^\dagger P U$  for some unitary  $U$  and for some effect  $P$  that, up to a permutation of qubits, equals  $\mathbb{1}_2^{\otimes k_Q} \otimes |0^{n-k_Q}\rangle\langle 0^{n-k_Q}|$ , wherein  $k_Q := \log_2(\text{tr}(Q))$ .  $\log(\text{tr}(Q)) \leq m$  precisely if  $k_Q \leq \lfloor m/\log(2) \rfloor$ , which holds precisely if  $k_Q \leq \lfloor m/\log(2) \rfloor$ . Moreover, if  $k_Q < \lfloor m/\log(2) \rfloor$ , then there exists a  $\tilde{Q} \in \mathbf{M}^r$  such that  $k_{\tilde{Q}} = \lfloor m/\log(2) \rfloor$  and  $\text{tr}(\tilde{Q}\rho) \geq \text{tr}(Q\rho)$ . (Indeed, let  $\tilde{Q} := U^\dagger \tilde{P} U$ , wherein  $\tilde{P}$ , up to the aforementioned permutation, equals  $\mathbb{1}_2^{\otimes k_{\tilde{Q}}} \otimes |0^{n-k_{\tilde{Q}}}\rangle\langle 0^{n-k_{\tilde{Q}}}|$ .) Thus, we can replace the constraint  $k_Q \leq \lfloor m/\log(2) \rfloor$  with  $k_Q = \lfloor m/\log(2) \rfloor$ :

$$\eta_h^{r,m}(\rho) = \sup_{\substack{Q \in \mathbf{M}^r \\ \log_2(\text{tr}(Q)) = \lfloor m/\log(2) \rfloor}} \{\text{tr}(Q\rho)\}. \quad (\text{D64})$$

The complexity success probability monotonically increases as  $r$  increases and monotonically increases as  $m$  increases.

**Proposition 25 (Monotonicity of  $\eta_h^{r,m}$  in  $r$  and  $m$ ).** Let  $\rho$  denote any subnormalized state. Let  $r \geq 0$ . Let  $m$  and  $m_+$  be as in Definition 17.

- (i) For all  $r' \geq r$ ,  $\eta_h^{r',m}(\rho) \geq \eta_h^{r,m}(\rho)$ .
- (ii) For all  $m' \in [m, m_+]$ ,  $\eta_h^{r,m'}(\rho) \geq \eta_h^{r,m}(\rho)$ .

*Proof.* Consider the definition (D63) of  $\eta_h^{r,m}(\rho)$ . Consider any  $\eta \in (0, \text{tr}(\rho)]$ . The monotonicity in  $r$  follows because  $H_h^{r,\eta}(\rho) \leq m$  implies that  $H_h^{r',\eta}(\rho) \leq m$ , by Proposition 12 (via Proposition 24). The monotonicity in

$m$  follows because  $H_h^{r,\eta}(\rho) \leq m$  implies that  $H_h^{r,\eta}(\rho) \leq m'$ . ■

Moreover, the complexity success probability is convex.

**Proposition 26 (Convexity of  $\eta_h^{r,m}$ ).** Let  $\{p_k\}$  denote any probability distribution; and  $\{\rho_k\}$ , any associated collection of subnormalized states. Let  $r \geq 0$ . Let  $m$  be as in Definition 17. It holds that

$$\eta_h^{r,m} \left( \sum_k p_k \rho_k \right) \leq \sum_k p_k \eta_h^{r,m}(\rho_k). \quad (\text{D65})$$

*Proof.* Consider any  $\zeta > 0$ . There exists a  $Q \in \mathbf{M}^r$  such that  $\log(\text{tr}(Q)) \leq m$  and

$$\begin{aligned} \eta_h^{r,m} \left( \sum_k p_k \rho_k \right) - \zeta &\leq \text{tr} \left( Q \left[ \sum_k p_k \rho_k \right] \right) = \sum_k p_k \text{tr}(Q \rho_k) \\ &\leq \sum_k p_k \eta_h^{r,m}(\rho_k). \end{aligned} \quad (\text{D66})$$

The second inequality holds because, for each  $k$ ,  $Q$  is a candidate for the  $\eta_h^{r,m}(\rho_k)$  optimization and thus satisfies  $\text{tr}(Q \rho_k) \leq \eta_h^{r,m}(\rho_k)$ . Equation (D65) now follows because  $\zeta$  is arbitrary. ■

### c. Interrelating reduced complexity (relative) entropies under two computational models

Under certain conditions, a simple bound relates reduced complexity (relative) entropies defined with respect to different computational models.

**Proposition 27 (Reduced complexity relative entropies under two computational models).** Let  $\rho$  denote any subnormalized state; and  $\Gamma$ , any positive-semidefinite operator. Let  $\mathbf{P}$  denote any set of simple POVM effects (Definition 7), and let  $C^{(1)}$  and  $C^{(2)}$  denote superoperator-complexity measures (Definition 1). For each  $i = 1, 2$ , let  $\{\mathbf{M}^{r,(i)} := \mathbf{M}^r(\mathbf{P}, C^{(i)})\}$  denote the family of POVM-effect-complexity sets defined by Eq. (B15). Let  $r \geq 0$ ,  $\alpha > 0$ ,  $\varepsilon \in [0, 1]$ , and  $\eta \in (\varepsilon, 1]$ . Assume there exists a monotonically increasing function  $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  with the following property: for every operation  $\mathcal{E}$  satisfying  $C^{(1)}(\mathcal{E}) < \infty$ , there exists an operation  $\mathcal{F}$  such that

$$\begin{aligned} \frac{1}{2} \|\mathcal{E} - \mathcal{F}\|_\diamond &\leq \varepsilon, \quad \mathcal{F}(\Gamma) \leq \alpha \mathcal{E}(\Gamma), \\ \text{and } C^{(2)}(\mathcal{F}) &\leq f(C^{(1)}(\mathcal{E})). \end{aligned} \quad (\text{D67})$$

It holds that

$$D_{h,(2)}^{f(r),\eta-\varepsilon}(\rho \parallel \Gamma) \geq D_{h,(1)}^{r,\eta}(\rho \parallel \Gamma) - \log(\alpha). \quad (\text{D68})$$

The reduced complexity relative entropies  $D_{h,(1)}^{r,\eta}$  and  $D_{h,(2)}^{f(r),\eta-\varepsilon}$  are defined with respect to  $\mathbf{M}^{r,(1)}$  and  $\mathbf{M}^{f(r),(2)}$ , respectively.

**Corollary 3 (Reduced complexity entropies under two unitary computational models).** Let  $\rho$  denote any subnormalized state. Let  $\mathbf{P}$  denote any set of simple POVM effects (Definition 7), and let  $C^{(1)}$  and  $C^{(2)}$  denote unitary-complexity measures (Definition 3). For each  $i = 1, 2$ , let  $\{\mathbf{M}^{r,(i)} := \mathbf{M}^r(\mathbf{P}, C^{(i)})\}$  denote the family of POVM-effect-complexity sets defined by Eq. (B15). Let  $r \geq 0$ ,  $\varepsilon \in [0, 1]$ , and  $\eta \in (\varepsilon, 1]$ . Assume there exists a monotonically increasing function  $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  with the following property: for every unitary  $U$  satisfying  $C^{(1)}(U) < \infty$ , there exists a unitary  $V$  such that

$$\|U - V\| \leq \varepsilon \quad \text{and} \quad C^{(2)}(V) \leq f(C^{(1)}(U)). \quad (\text{D69})$$

It holds that

$$H_{h,(2)}^{f(r),\eta-\varepsilon}(\rho) \leq H_{h,(1)}^{r,\eta}(\rho). \quad (\text{D70})$$

The reduced complexity entropies  $H_{h,(1)}^{r,\eta}$  and  $H_{h,(2)}^{f(r),\eta-\varepsilon}$  are defined with respect to  $\mathbf{M}^{r,(1)}$  and  $\mathbf{M}^{f(r),(2)}$ , respectively.

*Proof of Proposition 27.* Consider any  $\zeta > 0$ . There exists a  $Q \in \mathbf{M}^{r,(1)}$  such that  $D_{h,(1)}^{r,\eta}(\rho \parallel \Gamma) \leq -\log(\text{tr}(Q\Gamma)) + \zeta$ .  $Q = \mathcal{E}^\dagger(P)$  for some effect  $P \in \mathbf{P}$  and for some operation  $\mathcal{E}$  satisfying  $C^{(1)}(\mathcal{E}) \leq r$ . By assumption, there exists an operation  $\mathcal{F}$  satisfying the conditions (D67). Hence,  $C^{(2)}(\mathcal{F}) \leq f(C^{(1)}(\mathcal{E})) \leq f(r)$ . The second inequality follows because  $f$  is monotonically increasing. Let  $Q' := \mathcal{F}^\dagger(P) \in \mathbf{M}^{f(r),(2)}$ . It holds that

$$\begin{aligned} \text{tr}(Q'\rho) &= \text{tr}(P\mathcal{F}(\rho)) \geq \text{tr}(P\mathcal{E}(\rho)) - \varepsilon \\ &= \text{tr}(Q\rho) - \varepsilon \geq \eta - \varepsilon. \end{aligned} \quad (\text{D71})$$

The first inequality follows because  $\text{tr}(P\mathcal{F}(\rho)) \geq \text{tr}(P\mathcal{E}(\rho)) - \frac{1}{2}\|\mathcal{E}(\rho) - \mathcal{F}(\rho)\|_1$ , by Corollary 9.1.1 of Ref. [109], and because  $\frac{1}{2}\|\mathcal{E} - \mathcal{F}\|_\diamond \leq \varepsilon$  implies that  $\frac{1}{2}\|\mathcal{E}(\rho) - \mathcal{F}(\rho)\|_1 \leq \varepsilon$ , by Eq. (A2). Therefore,  $Q'$  is a candidate for the  $D_{h,(2)}^{f(r),\eta-\varepsilon}(\rho \parallel \Gamma)$  optimization. Moreover,  $\text{tr}(Q'\Gamma) = \text{tr}(P\mathcal{F}(\Gamma)) \leq \text{tr}(P[\alpha\mathcal{E}(\Gamma)]) = \alpha \text{tr}(Q\Gamma)$ , so

$$\begin{aligned} D_{h,(2)}^{f(r),\eta-\varepsilon}(\rho \parallel \Gamma) &\geq -\log(\text{tr}(Q'\Gamma)) \geq -\log(\text{tr}(Q\Gamma)) \\ &\quad - \log(\alpha) \geq D_{h,(1)}^{r,\eta}(\rho \parallel \Gamma) - \log(\alpha) - \zeta, \end{aligned} \quad (\text{D72})$$

which implies Eq. (D68), since  $\zeta$  is arbitrary. ■

*Proof of Corollary 3.* Equation (D70) follows from applying Proposition 27 when  $\Gamma = \mathbb{1}$  and  $\alpha = 1$ . We need to show only that the conditions (D67) hold in



this case. Let  $\Gamma = \mathbb{1}$  and  $\alpha = 1$ . Consider any operation  $\mathcal{E}$  satisfying  $C^{(1)}(\mathcal{E}) < \infty$ .  $\mathcal{E}$  is a unitary operation, since  $C^{(1)}$  is a unitary-complexity measure. Hence,  $\mathcal{E}(\cdot) = U(\cdot)U^\dagger$  for some unitary operator  $U$  satisfying  $C^{(1)}(U) < \infty$ . By assumption, there exists a unitary operator  $V$  such that  $\|U - V\| \leq \varepsilon$  and  $C^{(2)}(V) \leq f(C^{(1)}(U))$ . Let  $\mathcal{F}(\cdot) := V(\cdot)V^\dagger$ . By Lemma 2,  $\frac{1}{2}\|\mathcal{E} - \mathcal{F}\|_\diamond \leq \|U - V\| \leq \varepsilon$ . Also,  $\mathcal{F}(\Gamma) \leq \alpha \mathcal{E}(\Gamma)$ , since  $\mathcal{F}(\mathbb{1}) = \mathbb{1} = \mathcal{E}(\mathbb{1})$ . Last,  $C^{(2)}(\mathcal{F}) \leq f(C^{(1)}(\mathcal{E}))$  is a restatement of  $C^{(2)}(V) \leq f(C^{(1)}(U))$ . ■

## 6. Complexity conditional entropy

We here define the *complexity conditional entropy* in terms of the complexity relative entropy, just as one defines one-shot conditional entropies in terms of one-shot relative entropies [94]. In Sec. V G, the complexity conditional entropy quantifies the maximally mixed qubits one can decouple from a reference system.

**Definition 18 (Complexity conditional entropy).** Let  $A$  and  $B$  denote distinct quantum systems. Let  $\rho_{AB}$  denote any subnormalized state of  $AB$ . Let  $r \geq 0$  and  $\eta \in (0, \text{tr}(\rho))$ . The *complexity conditional entropy* of  $A$  conditioned on  $B$  is

$$H_H^{r,\eta}(A|B)_\rho := -D_H^{r,\eta}(\rho_{AB} \parallel \mathbb{1}_A \otimes \rho_B). \quad (\text{D73})$$

**Proposition 28 (General bounds for the complexity conditional entropy).** Let  $A$  and  $B$  denote distinct quantum systems. Let  $\rho_{AB}$  denote any subnormalized state of  $AB$ . Let  $r \geq 0$  and  $\eta \in (0, \text{tr}(\rho))$ . It holds that

$$-\log(d_A) \leq H_H^{r,\eta}(A|B)_\rho \leq \log(d_A). \quad (\text{D74})$$

Equivalently,

$$0 \leq D_H^{r,\eta}(\rho_{AB} \parallel \pi_A \otimes \rho_B) \leq 2 \log(d_A). \quad (\text{D75})$$

*Proof.*  $H_H^{r,\eta}(A|B)_\rho \leq \log(d_A)$  follows from applying Eq. (D4a) in Proposition 10 with  $\Gamma = \mathbb{1}_A \otimes \rho_B$ :

$$\begin{aligned} -D_H^{r,\eta}(\rho_{AB} \parallel \mathbb{1}_A \otimes \rho_B) &\leq \log(\text{tr}(\mathbb{1}_A \otimes \rho_B)) - \log(\text{tr}(\rho_{AB})) \\ &= \log(d_A \text{tr}(\rho_B)) - \log(\text{tr}(\rho_B)) \\ &= \log(d_A). \end{aligned} \quad (\text{D76})$$

The first equality follows because  $\text{tr}(\rho_{AB}) = \text{tr}_B(\text{tr}_A(\rho_{AB})) = \text{tr}(\rho_B)$ .  $H_H^{r,\eta}(A|B)_\rho \geq -\log(d_A)$  follows from applying Proposition 14 to the inequality  $d_A^{-1} \rho_{AB} \leq \mathbb{1}_A \otimes \rho_B$  (Lemma 1):

$$\begin{aligned} -D_H^{r,\eta}(\rho_{AB} \parallel \mathbb{1}_A \otimes \rho_B) &\geq -D_H^{r,\eta}(\rho_{AB} \parallel d_A^{-1} \rho_{AB}) \\ &= -D_H^{r,\eta}(\rho_{AB} \parallel \rho_{AB}) - \log(d_A) \\ &= -\log(d_A). \end{aligned} \quad (\text{D77})$$

The first equality follows from Proposition 13 and the second equality from Proposition 9.

Last, Eqs. (D74) and (D75) are equivalent because Proposition 13 implies that  $H_H^{r,\eta}(A|B)_\rho = \log(d_A) - D_H^{r,\eta}(\rho_{AB} \parallel \pi_A \otimes \rho_B)$ :

$$\begin{aligned} H_H^{r,\eta}(A|B)_\rho &= -D_H^{r,\eta}(\rho_{AB} \parallel \mathbb{1}_A \otimes \rho_B) \\ &= -D_H^{r,\eta}(\rho_{AB} \parallel d_A \pi_A \otimes \rho_B) \\ &= \log(d_A) - D_H^{r,\eta}(\rho_{AB} \parallel \pi_A \otimes \rho_B). \end{aligned} \quad (\text{D78})$$

**Proposition 29 (Strong subadditivity of the complexity conditional entropy).** Let  $A$ ,  $B$ , and  $C$  denote distinct quantum systems. Let  $\rho_{ABC}$  denote any subnormalized state of  $ABC$ . Let  $r \geq 0$  and  $\eta \in (0, \text{tr}(\rho))$ . It holds that

$$H_H^{r,\eta}(A|BC)_\rho \leq H_H^{r,\eta}(A|B)_\rho. \quad (\text{D79})$$

*Proof.* By Proposition 16, the complexity relative entropy never increases under a partial trace, so

$$\begin{aligned} H_H^{r,\eta}(A|BC)_\rho &= -D_H^{r,\eta}(\rho_{ABC} \parallel \mathbb{1}_A \otimes \rho_{BC}) \\ &\leq -D_H^{r,\eta}(\rho_{AB} \parallel \mathbb{1}_A \otimes \rho_B) \\ &= H_H^{r,\eta}(A|B)_\rho. \end{aligned} \quad (\text{D80})$$

## APPENDIX E: THERMODYNAMIC ERASURE OF QUBITS GOVERNED BY A PRODUCT HAMILTONIAN

Here, we generalize the arguments in Sec. III A to prove Theorem 3 (Sec. III B). Recall the theorem's setting. We consider a system of  $n$  noninteracting qubits. Qubit  $i$  evolves under a Hamiltonian  $H_i$  that has zero ground-state energy,  $H_i|0\rangle_i = 0$ . We fix an inverse temperature  $\beta > 0$ . Qubit  $i$  has the Gibbs-weight operator  $\Gamma_i = e^{-\beta H_i}$ ; and the  $n$ -qubit system, the Gibbs-weight operator  $\Gamma = \bigotimes_{i=1}^n \Gamma_i = e^{-\beta \sum_i H_i}$ .

Let  $\mathcal{T}$  denote a set of elementary computations on the  $n$ -qubit system. Each computation in  $\mathcal{T}$  is a completely positive, trace-preserving map that sends  $\Gamma$  to itself and costs no work to implement. After implementing elementary computations, one can apply RESET operations to a selection  $\mathcal{W} \subset \{1, 2, \dots, n\}$  of qubits. Each RESET operation  $\mathcal{E}_{\text{RESET},i}$  initializes qubit  $i$  in the state  $|0\rangle_i$  and costs an amount of work  $W_{\text{RESET},i} = \beta^{-1} \log(\text{tr}(\Gamma_i)) \geq 0$ . Consider a general protocol  $\mathcal{E}$  consisting of  $r \geq 0$  elementary computations followed by RESET operations on the qubits of  $\mathcal{W}$ .  $\mathcal{E}$  has the form  $\mathcal{E} = (\prod_{i \in \mathcal{W}} \mathcal{E}_{\text{RESET},i}) \mathcal{E}_r \cdots \mathcal{E}_2 \mathcal{E}_1$ . Each  $\mathcal{E}_i$  is in  $\mathcal{T}$ , and the work cost  $W(\mathcal{E}) = \sum_{i \in \mathcal{W}} W_{\text{RESET},i} = \beta^{-1} \log(\prod_{i \in \mathcal{W}} \text{tr}(\Gamma_i))$ .

**Theorem 5.** Let  $\rho$  denote any quantum state of  $n$  qubits. Let  $\mathbf{P}$  denote the set of simple POVM effects (Definition 7)

defined in Eq. (10). Let  $C_{\mathcal{T}}$  denote the circuit-complexity measure associated with  $\mathcal{T}$  (Definition 2). Let  $r \geq 0$  denote an integer, and let  $\eta \in (0, 1]$ . Let  $M^r = M^r(\mathbf{P}, C_{\mathcal{T}})$

denote the POVM-effect-complexity set defined in Eq. (B15). Let

$$W_r^* := \min \left\{ W(\mathcal{E}) : \mathcal{E} = \left( \prod_{i \in \mathcal{W}} \mathcal{E}_{\text{RESET}, i} \right) \mathcal{E}_r \cdots \mathcal{E}_2 \mathcal{E}_1, \mathcal{W} \subset \{1, 2, \dots, n\}, F^2(\mathcal{E}(\rho), |0^n\rangle\langle 0^n|) \geq \eta \right\}. \quad (\text{E1})$$

Each  $\mathcal{E}_i$  is in  $\mathcal{T}$ , and  $F$  denotes the fidelity between quantum states. It holds that

$$\beta W_r^* = -D_{h, \mathcal{T}}^{r, \eta}(\rho \parallel \Gamma), \quad (\text{E2})$$

wherein the reduced complexity relative entropy  $D_{h, \mathcal{T}}^{r, \eta}$  (Definition 15) is defined with respect to  $M^r$ . Consequently,

$$-D_{h, \mathcal{T}}^{r, \eta}(\rho \parallel \Gamma) - \log\left(\frac{1}{\eta}\right) \leq \beta W_r^* \leq -D_{h, \mathcal{T}}^{r, \eta}(\rho \parallel \Gamma). \quad (\text{E3})$$

Consider the special case where each Hamiltonian  $H_i$  is degenerate:  $H_i = 0$ . In this case,  $\Gamma = \mathbb{1}$ . By Eq. (E2),  $\beta W_r^*$  equals the reduced complexity entropy (Definition 16):  $\beta W_r^* = -D_{h, \mathcal{T}}^{r, \eta}(\rho \parallel \mathbb{1}) = H_{h, \mathcal{T}}^{r, \eta}(\rho)$ . If, furthermore,  $\mathcal{T}$  contains only unitary operations, then Eq. (E3) yields the bounds in Theorem 1 (Sec. III A).

*Proof of Theorem 5.* We first prove that  $\beta W_r^* \geq -D_{h, \mathcal{T}}^{r, \eta}(\rho \parallel \Gamma)$ . Consider any operation  $\mathcal{E}$  that achieves the minimum in Eq. (E1) and thus satisfies  $W(\mathcal{E}) = W_r^*$ .  $\mathcal{E} = (\prod_{i \in \mathcal{W}} \mathcal{E}_{\text{RESET}, i}) \mathcal{E}_r \cdots \mathcal{E}_2 \mathcal{E}_1$  for some subset  $\mathcal{W} \subset \{1, 2, \dots, n\}$  and some operations  $\mathcal{E}_i \in \mathcal{T}$ . Moreover,  $F^2(\mathcal{E}(\rho), |0^n\rangle\langle 0^n|) \geq \eta$ . Let  $P := \mathbb{1}_{\mathcal{W}} \otimes |0^{n-|\mathcal{W}|}\rangle\langle 0^{n-|\mathcal{W}|}|_{\mathcal{W}^c} \in \mathbf{P}$ , wherein  $\mathcal{W}^c$  denotes the set complement of  $\mathcal{W}$ . Let  $Q := \mathcal{E}_1^\dagger \mathcal{E}_2^\dagger \cdots \mathcal{E}_r^\dagger(P) \in M^r$ . It holds that

$$\begin{aligned} \text{tr}(Q\rho) &= \text{tr}(P \mathcal{E}_r \cdots \mathcal{E}_2 \mathcal{E}_1(\rho)) \\ &= \text{tr}\{|0^{n-|\mathcal{W}|}\rangle\langle 0^{n-|\mathcal{W}|}|_{\mathcal{W}^c} \text{tr}_{\mathcal{W}}(\mathcal{E}_r \cdots \mathcal{E}_2 \mathcal{E}_1(\rho))\} \\ &= \text{tr}(|0^n\rangle\langle 0^n| \mathcal{E}(\rho)) = F^2(\mathcal{E}(\rho), |0^n\rangle\langle 0^n|). \end{aligned} \quad (\text{E4})$$

The third equality follows because  $\mathcal{E}(\rho) = |0^{|\mathcal{W}|}\rangle\langle 0^{|\mathcal{W}|}|_{\mathcal{W}} \otimes \text{tr}_{\mathcal{W}}(\mathcal{E}_r \cdots \mathcal{E}_2 \mathcal{E}_1(\rho))$ . Hence,  $\text{tr}(Q\rho) \geq \eta$ , so  $Q$  is a candidate for the  $D_{h, \mathcal{T}}^{r, \eta}(\rho \parallel \Gamma)$  optimization. Furthermore,

$$\begin{aligned} \text{tr}(Q\Gamma) &= \text{tr}(P \mathcal{E}_r \cdots \mathcal{E}_1(\Gamma)) = \text{tr}(P\Gamma) \\ &= \text{tr}\left(\left[\prod_{i \in \mathcal{W}} \Gamma_i\right] \otimes |0^{n-|\mathcal{W}|}\rangle\langle 0^{n-|\mathcal{W}|}|_{\mathcal{W}^c}\right) \\ &= \prod_{i \in \mathcal{W}} \text{tr}(\Gamma_i) = e^{\beta W(\mathcal{E})}. \end{aligned} \quad (\text{E5})$$

The second equality follows because  $\mathcal{E}_r \cdots \mathcal{E}_2 \mathcal{E}_1(\Gamma) = \Gamma$ , since each  $\mathcal{E}_i$  satisfies  $\mathcal{E}_i(\Gamma) = \Gamma$ . Thus,  $\beta W_r^* = \beta W(\mathcal{E}) = \log(\text{tr}(Q\Gamma)) \geq -D_{h, \mathcal{T}}^{r, \eta}(\rho \parallel \Gamma)$ .

We now prove that  $\beta W_r^* \leq -D_{h, \mathcal{T}}^{r, \eta}(\rho \parallel \Gamma)$ . Consider any  $\zeta > 0$ . There exists a  $Q \in M^r$  such that  $\text{tr}(Q\rho) \geq \eta$  and  $\log(\text{tr}(Q\Gamma)) \leq -D_{h, \mathcal{T}}^{r, \eta}(\rho \parallel \Gamma) + \zeta$ . Since  $M^r = M^r(\mathbf{P}, C_{\mathcal{T}})$ ,  $Q = \mathcal{E}_1^\dagger \mathcal{E}_2^\dagger \cdots \mathcal{E}_r^\dagger(P)$  for some operations  $\mathcal{E}_i \in \mathcal{T}$  and for some effect  $P \in \mathbf{P}$ , with  $P = \mathbb{1}_{\mathcal{W}} \otimes |0^{n-|\mathcal{W}|}\rangle\langle 0^{n-|\mathcal{W}|}|_{\mathcal{W}^c}$  for some subset  $\mathcal{W} \subset \{1, 2, \dots, n\}$ . Let  $\mathcal{E} := (\prod_{i \in \mathcal{W}} \mathcal{E}_{\text{RESET}, i}) \mathcal{E}_r \cdots \mathcal{E}_2 \mathcal{E}_1$ . As per Eq. (E4),  $F^2(\mathcal{E}(\rho), |0^n\rangle\langle 0^n|) = \text{tr}(Q\rho) \geq \eta$ , so  $\mathcal{E}$  is a candidate for the  $W_r^*$  optimization in Eq. (E1). As per Eq. (E5),  $\text{tr}(Q\Gamma) = e^{\beta W(\mathcal{E})}$ , so  $\beta W_r^* \leq \beta W(\mathcal{E}) = \log(\text{tr}(Q\Gamma)) \leq -D_{h, \mathcal{T}}^{r, \eta}(\rho \parallel \Gamma) + \zeta$ . Thus,  $\beta W_r^* \leq -D_{h, \mathcal{T}}^{r, \eta}(\rho \parallel \Gamma)$ , since  $\zeta$  is arbitrary. Finally, Eq. (E3) follows from substituting Eq. (E2) into Eq. (D59a). ■

## APPENDIX F: EVOLUTION OF THE COMPLEXITY ENTROPY UNDER RANDOM CIRCUITS

Here, we prove the bound Eq. (47) to complete the proof of Proposition 3 (Sec. V D), which describes the complexity entropy's evolution under random circuits. Random circuits are often used as proxies for Hamiltonian quantum chaotic dynamics. We adapt the proof of Theorem 8 in Ref. [17]; the theorem describes the strong complexity's growth under random circuits.

We consider a system of  $n \geq 2$  qubits. Let  $G$  denote any set of two-qubit unitary gates:  $G \subset \text{SU}(4)$ . Let  $\{M_G^r\}$  denote the family of POVM-effect-complexity sets (Definition 6) defined by Eq. (11), wherein the computational gate set  $\mathcal{G}$  equals the set of gates on  $n$  qubits, in  $G$ , with arbitrary (fixed) connectivity.

Let  $\mathbb{E} = \{p_j, U_j\}$  denote a random ensemble of unitary operators. Each unitary  $U_j \in \text{U}(2^n)$  is chosen with the probability  $p_j$ . For each integer  $k > 0$ , we define the  $k$ -twirling superoperator associated with  $\mathbb{E}$  as

$$\mathcal{M}_{\mathbb{E}}^{(k)}(\cdot) := \sum_j p_j U_j^{\otimes k}(\cdot) U_j^{\dagger \otimes k}. \quad (\text{F1})$$

The  $k$ -twirling superoperator associated with the Haar measure is

$$\mathcal{M}_{\text{Haar}}^{(k)}(\cdot) := \int dU U^{\otimes k}(\cdot) U^{\dagger \otimes k}. \quad (\text{F2})$$

$dU$  denotes the Haar measure on the unitary group  $U(2^n)$ . If  $\mathcal{M}_{\text{E}}^{(k)}$  and  $\mathcal{M}_{\text{Haar}}^{(k)}$  are sufficiently close in diamond distance,  $\text{E}$  is an approximate  $k$  design.

*Definition 19 (Approximate unitary  $k$  design).* Let  $\varepsilon \geq 0$ , and let  $k > 0$  denote an integer. An ensemble  $\text{E} = \{p_j, U_j\}$  is an  $\varepsilon$ -approximate unitary  $k$  design if

$$\left\| \mathcal{M}_{\text{E}}^{(k)} - \mathcal{M}_{\text{Haar}}^{(k)} \right\|_{\diamond} \leq \frac{\varepsilon}{2^{nk}}. \quad (\text{F3})$$

We employ the definition of an approximate  $k$  design used in Ref. [56]; this definition differs from that used in Ref. [17]. In Ref. [17],  $(k!/2^{2nk})\varepsilon$  replaces the upper bound in Eq. (F3). Therefore, an  $\varepsilon$ -approximate  $k$  design, according to Definition 19, is a  $(2^{nk}\varepsilon/k!)$ -approximate  $k$ -design, according to Ref. [17].

Following Ref. [17], we rely on the fact that random circuits generate approximate  $k$ -designs [32,55,56,135]. In particular, we utilize the following result, which shows that certain random circuits of a depth  $t > 0$  are approximate  $k$  designs, with  $k \sim t$ .

*Theorem 6 ([32]).* Consider an  $n$ -qubit system. Let  $\tilde{\varepsilon} > 0$  and  $c \geq 0$  be independent of  $n$ . Let  $V$  denote a circuit of a depth  $t > 0$ , with staggered layers of nearest-neighbor two-qubit gates (the “brickwork” layout). Suppose each of the

circuit’s gates is chosen at random from the Haar measure on  $\text{SU}(4)$ . Then  $V$  is a  $(2^{-cnk}\tilde{\varepsilon})$ -approximate  $k$  design, with

$$k = \min \left\{ 2^{n/2 - O(\sqrt{n})}, \frac{t}{\text{poly}(n)} \right\}. \quad (\text{F4})$$

The dependence of  $k$  on  $\tilde{\varepsilon}$  is hidden in the coefficient implied by the  $\text{poly}(n)$  notation.

We briefly show how condition (F4) arises from the results in Ref. [32]. Theorem 1.5 in Ref. [56] implies the following statement. Let  $0 < k \leq 2^{n/2 - \sqrt{n} - 2}$  and let  $\varepsilon > 0$ . Then there exists  $C > 0$  such that depth- $t$  brickwork random circuits form an  $\varepsilon$ -approximate  $k$  design whenever

$$t \geq Cn^3[2nk + \log_2(1/\varepsilon)] =: t_0(k). \quad (\text{F5})$$

Setting  $\varepsilon = 2^{-cnk}\tilde{\varepsilon}$ , with  $\tilde{\varepsilon} \in (0, 1)$  a constant of  $n$ , and using  $k \leq 2^{n/2}$ , we find that

$$t_0(k) \leq \text{poly}(n)k. \quad (\text{F6})$$

Consider now the setting of Theorem 6. The  $k$  specified in Eq. (F4) implies  $t \geq t_0(k)$ , ensuring that Corollary 1 in Ref. [56] applies.

The bound (47)—and hence Proposition 3—follows from Theorem 6 and the following proposition.

*Proposition 30 (Complexity-entropy lower bound with an approximate  $k$  design (simple)).* Let  $|\psi_0\rangle$  denote any pure  $n$ -qubit state. Let  $\tilde{\varepsilon} \in (0, 1)$ . Assume that  $k$  is even. Let  $V$  denote a circuit sampled from a  $(2^{-nk}\tilde{\varepsilon})$ -approximate  $k$  design. Let  $r \geq 0$  and  $\eta \in (0, 1]$ . Suppose that

$$r \leq \frac{k \left[ \frac{n}{2} - \log_2(k) - \log_2(4/\eta) \right] - \{3n + \log_2(n)[1 + \log_2(4/\eta)]\}}{c_1 [2n + \log_2(2/\eta)]^4 + \log_2(n+1) + 1}, \quad (\text{F7})$$

wherein  $c_1 > 0$  denotes a constant independent of  $n$ . Then

$$H_{\text{H}}^{r,\eta}(V\psi_0 V^\dagger) \geq H_{\text{h}}^{r,\eta}(V\psi_0 V^\dagger) \geq n \log(2) - \log\left(\frac{4}{\eta}\right), \quad (\text{F8})$$

except with a probability  $\leq e^{-\Omega(n)}$  over the sampling of  $V$ . Here,  $H_{\text{H}}^{r,\eta}$  is defined with respect to  $\mathbf{M}_{\text{SU}(4)}^r$ .

We specify the constant  $c_1$  in the proof of Proposition 30. To prove the proposition, we first use the

Solovay-Kitaev theorem [95,136–138] to approximate  $\text{SU}(4)$  with a finite, universal gate set. We then apply the following lemma, whose proof closely follows that of Theorem 8 of Ref. [17].

*Lemma 3 (Gate-set complexity-entropy lower bound for approximate  $k$ -designs).* Let  $|\psi_0\rangle$  denote any pure  $n$ -qubit state. Let  $\tilde{\varepsilon} \in (0, 1)$ . Assume that  $k$  is even. Let  $V$  denote a circuit sampled from a  $(2^{-nk}\tilde{\varepsilon})$ -approximate  $k$  design. Let  $r \geq 0$ ,  $\eta \in (0, 1]$ ,  $c \in (0, 1)$ , and  $g > 0$ . Suppose that

$$r \leq \frac{1}{\log_2[(n+1)G]} \left( k \left[ \frac{n}{2} - \log_2(k) - \log_2\left(\frac{1}{\eta(1-c)}\right) \right] - \left\{ (g+1)n + 2 + \log_2(n) \left[ 1 + \log_2\left(\frac{1}{c\eta}\right) \right] \right\} \right). \quad (\text{F9})$$

Then

$$H_{h,G}^{r,\bar{\eta}}(V\psi_0 V^\dagger) \geq n \log(2) - \log\left(\frac{1}{c\bar{\eta}}\right), \quad (\text{F10})$$

except with a probability  $\leq 2^{-gn}$  over the sampling of  $V$ . Here, the reduced complexity entropy  $H_{h,G}^{r,\bar{\eta}}$  (Definition 16) is defined with respect to  $\mathbf{M}_G^r$ , wherein  $G$  is finite.

*Proof.* Let

$$m := \left\lfloor n - \log_2\left(\frac{1}{c\bar{\eta}}\right) \right\rfloor = n - \left\lceil \log_2\left(\frac{1}{c\bar{\eta}}\right) \right\rceil. \quad (\text{F11})$$

Equation (F10) holds, except with the probability

$$\begin{aligned} p_{<} &:= \Pr_V \left[ H_{h,G}^{r,\bar{\eta}}(V\psi_0 V^\dagger) < n \log(2) - \log\left(\frac{1}{c\bar{\eta}}\right) \right] \\ &\leq \Pr_V \left[ H_{h,G}^{r,\bar{\eta}}(V\psi_0 V^\dagger) \leq n \log(2) - \log\left(\frac{1}{c\bar{\eta}}\right) \right] \\ &= \Pr_V \left[ H_{h,G}^{r,\bar{\eta}}(V\psi_0 V^\dagger) \leq m \log(2) \right] \\ &\leq \Pr_V \left[ \eta_{h,G}^{r,m \log(2)}(V\psi_0 V^\dagger) \geq \bar{\eta} \right] \\ &= \Pr_V \left[ \max_{\substack{Q \in \mathbf{M}^r \\ \log_2(\text{tr}(Q))=m}} \{ \text{tr}(Q V\psi_0 V^\dagger) \} \geq \bar{\eta} \right] \\ &\leq \sum_{\substack{Q \in \mathbf{M}^r \\ \log_2(\text{tr}(Q))=m}} \Pr_V [\text{tr}(Q V\psi_0 V^\dagger) \geq \bar{\eta}] \\ &= \sum_{\substack{Q \in \mathbf{M}^r \\ \log_2(\text{tr}(Q))=m}} \Pr_V [\text{tr}([Q - 2^{m-n} \mathbb{1}] V\psi_0 V^\dagger) \geq \bar{\eta} - 2^{m-n}] \\ &\leq \sum_{\substack{Q \in \mathbf{M}^r \\ \log_2(\text{tr}(Q))=m}} \Pr_V [|\text{tr}([Q - 2^{m-n} \mathbb{1}] V\psi_0 V^\dagger)| \\ &\geq |(1-c)\bar{\eta}|] \\ &= \sum_{\substack{Q \in \mathbf{M}^r \\ \log_2(\text{tr}(Q))=m}} \Pr_V [\{ \text{tr}([Q - 2^{m-n} \mathbb{1}] V\psi_0 V^\dagger) \}^k \\ &\geq \{(1-c)\bar{\eta}\}^k] \\ &\leq \sum_{\substack{Q \in \mathbf{M}^r \\ \log_2(\text{tr}(Q))=m}} \frac{\mathbb{E}_V [\{ \text{tr}([Q - 2^{m-n} \mathbb{1}] V\psi_0 V^\dagger) \}^k]}{[(1-c)\bar{\eta}]^k} \end{aligned}$$

$$\begin{aligned} &\leq \sum_{\substack{Q \in \mathbf{M}^r \\ \log_2(\text{tr}(Q))=m}} 2 \left( \frac{k}{2^{n/2}(1-c)\bar{\eta}} \right)^k \\ &= 2 \left( \frac{k}{2^{n/2}(1-c)\bar{\eta}} \right)^k |\{Q \in \mathbf{M}^r : \log_2(\text{tr}(Q))=m\}|. \end{aligned} \quad (\text{F12})$$

The second equality follows because, by the form of  $\mathbf{M}^r$ ,  $H_{h,G}^{r,\bar{\eta}}(V\psi_0 V^\dagger)$  is an integer multiple of  $\log(2)$ . The second inequality follows from Definition 17:  $H_{h,G}^{r,\bar{\eta}}(V\psi_0 V^\dagger) \leq m \log(2)$  implies that  $\eta_{h,G}^{r,m \log(2)}(V\psi_0 V^\dagger) \geq \bar{\eta}$ . The third equality follows from Eq. (D64); here, the supremum is a maximum because  $G$ —and hence  $\mathbf{M}_G^r$ —is finite. The third inequality is a crude union bound. The fourth equality introduces the traceless part of each  $Q$ :  $Q - 2^{m-n} \mathbb{1} = Q - \text{tr}(Q) \cdot 2^{-n} \mathbb{1}$ . The fourth inequality follows because

$$|\text{tr}([Q - 2^{m-n} \mathbb{1}] V\psi_0 V^\dagger)| \geq \text{tr}([Q - 2^{m-n} \mathbb{1}] V\psi_0 V^\dagger) \quad (\text{F13})$$

and because  $\bar{\eta} - 2^{m-n} \geq (1-c)\bar{\eta} > 0$ , since  $2^{m-n} = 2^{-\lceil \log_2(1/c\bar{\eta}) \rceil} \leq 2^{-\log_2(1/c\bar{\eta})} = c\bar{\eta}$ . The fifth equality follows because, by assumption,  $k$  is even. The fifth inequality follows from Markov's inequality: for any non-negative random variable  $X$  and any  $\tau > 0$ ,  $\Pr[X \geq \tau] = \mathbb{E}[X]/\tau$ . The last inequality follows from Corollary 24 of Ref. [17]:  $\mathbb{E}_V [\{ \text{tr}([Q - 2^{m-n} \mathbb{1}] V\psi_0 V^\dagger) \}^k] \leq (1+\varepsilon) (k/2^{n/2})^k \leq 2(k/2^{n/2})^k$ , wherein  $\varepsilon = (2^{nk}/k!)2^{-nk}\tilde{\varepsilon} \leq 1$  (recall the alternative convention used for  $\varepsilon$ -approximate designs in Ref. [17]).

By the form of  $\mathbf{M}^r$ ,  $|\{Q \in \mathbf{M}^r : \log_2(\text{tr}(Q))=m\}| \leq N_r \binom{n}{m}$ . The  $N_r$  denotes the number of circuits that one can compose from  $\leq r$  gates in  $\mathcal{G}$ .  $\binom{n}{m}$  is the number of projectors that, up to a permutation of qubits, equal  $\mathbb{1}^{\otimes m} \otimes |0^{n-m}\rangle\langle 0^{n-m}|$ . By Eq. (B21) of Ref. [17],  $N_r \leq 2^{n+1} (n+1)^r |G|^r$ . Moreover,  $\binom{n}{m} = \binom{n}{n-m} \leq n^{n-m} = n^{\lceil \log_2(1/c\bar{\eta}) \rceil} \leq n^{1+\log_2(1/c\bar{\eta})}$ . Hence,

$$|\{Q \in \mathbf{M}^r : \log_2(\text{tr}(Q))=m\}| \leq 2^{n+1} (n+1)^r |G|^r \cdot n^{1+\log_2(1/c\bar{\eta})}, \quad (\text{F14})$$

so  $p_{<} \leq 2^{a+rb-ky}$ , where

$$\begin{aligned} a &:= n + 2 + \log_2(n) \left[ 1 + \log_2\left(\frac{1}{c\bar{\eta}}\right) \right], \\ b &:= \log_2([n+1]|G|), \quad \text{and} \\ y &:= \frac{n}{2} - \log_2(k) - \log_2\left(\frac{1}{(1-c)\bar{\eta}}\right). \end{aligned} \quad (\text{F15})$$

By Eq. (F9),  $r \leq \frac{1}{b}[ky - (gn+a)]$ , so  $a+rb-ky \leq -gn$ . Thus,  $p_{<} \leq 2^{-gn}$ . ■



*Proof of Proposition 30.* Let  $\varepsilon' := \eta/2$ . Let  $G \subset \text{SU}(4)$  denote any finite gate set that forms an  $(\varepsilon'/4^n)$ -net on  $\text{SU}(4)$ . One can build such a gate set using the Solovay-Kitaev theorem [95,136–138] as follows. Consider any finite, universal gate set  $G_0 \subset \text{SU}(4)$ , such as the Clifford +  $T$  gate set (see Appendix B 6b). By the Solovay-Kitaev theorem, one can approximate any unitary  $U \in \text{SU}(4)$  to the error  $\varepsilon'$  with at most  $N := c_{G_0}[\log_2(4^n/\varepsilon')]^4 = c_{G_0}[2n + \log_2(2/\eta)]^4$  gates in  $G_0$ . The  $c_{G_0} > 0$  is a universal constant depending only on  $G_0$ . Now, let  $G$  consist of all the circuits that one can compose from  $\leq N$  gates in  $G_0$ . One can approximate any unitary  $U \in \text{SU}(4)$  to the error  $\varepsilon'$  by some circuit in  $G$ . The number of circuits that one can compose from exactly  $t > 0$  gates in  $G_0$  is  $|G_0|^t$ , so

$$|G| \leq \sum_{t=1}^N |G_0|^t = \frac{|G_0|}{|G_0| - 1} (|G_0|^N - 1) \leq 2|G_0|^N. \quad (\text{F16})$$

Hence,

$$\log_2(|G|) \leq N \log_2(|G_0|) + 1 = c_1 \left[ 2n + \log_2\left(\frac{2}{\eta}\right) \right]^4 + 1, \quad (\text{F17})$$

wherein  $c_1 := c_{G_0} \log_2(|G_0|)$ .

We now apply Lemma 3. Set

$$g = 1, \quad c = \frac{1}{2}, \quad \text{and} \quad \bar{\eta} = \eta - \varepsilon' = \frac{\eta}{2}. \quad (\text{F18})$$

$r$  satisfies condition (F9) in the lemma, since the right-hand side of Eq. (F7) lower-bounds the right-hand side of Eq. (F9), as one can verify using the following inequalities:

$$\begin{aligned} \log_2([n+1]|G|) &\leq c_1 \left[ 2n + \log_2\left(\frac{2}{\eta}\right) \right]^4 \\ &\quad + \log_2(n+1) + 1, \end{aligned} \quad (\text{F19})$$

$$\log_2\left(\frac{1}{(1-c)\bar{\eta}}\right) = \log_2\left(\frac{4}{\eta}\right), \quad (\text{F20})$$

$$\begin{aligned} \text{and } 2 \leq n \quad \Rightarrow \quad (g+1)n + 2 + \log_2\left(\frac{1}{c\bar{\eta}}\right) \\ \leq 3n + \log_2\left(\frac{4}{\eta}\right). \end{aligned} \quad (\text{F21})$$

Thus, by Lemma 3,

$$\begin{aligned} H_{h,G}^{r,\eta-\varepsilon'}(V\psi_0 V^\dagger) &= H_{h,G}^{r,\bar{\eta}}(V\psi_0 V^\dagger) \geq n \log(2) - \log\left(\frac{1}{c\bar{\eta}}\right) \\ &= n \log(2) - \log\left(\frac{4}{\eta}\right), \end{aligned} \quad (\text{F22})$$

except with a probability  $\leq 2^{-n}$ . Furthermore,

$$\begin{aligned} H_H^{r,\eta}(V\psi_0 V^\dagger) &\geq H_h^{r,\eta}(V\psi_0 V^\dagger) \geq H_{h,G}^{r,\eta}(V\psi_0 V^\dagger) \\ &\geq H_{h,G}^{r,\eta-\varepsilon'}(V\psi_0 V^\dagger), \end{aligned} \quad (\text{F23})$$

wherein  $H_h^{r,\eta}$  is defined with respect to  $\mathbf{M}_{\text{SU}(4)}^r$ . The first inequality follows from Proposition 22. The second inequality follows because  $\mathbf{M}_G^r \subset \mathbf{M}_{\text{SU}(4)}^r$ : every candidate  $Q \in \mathbf{M}_G^r$  for the  $H_{h,G}^{r,\eta}(V\psi_0 V^\dagger)$  optimization belongs to  $\mathbf{M}_{\text{SU}(4)}^r$  and is therefore a candidate for the  $H_h^{r,\eta}(V\psi_0 V^\dagger)$  optimization. The third inequality follows from Proposition 12 (via Proposition 24). Therefore,

$$H_H^{r,\eta}(V\psi_0 V^\dagger) \geq H_h^{r,\eta}(V\psi_0 V^\dagger) \geq n \log(2) - \log\left(\frac{4}{\eta}\right), \quad (\text{F24})$$

except with a probability at most exponentially small in  $n$ . ■

## APPENDIX G: ENTANGLEMENT BOUNDS ON THE COMPLEXITY ENTROPY

Here, we use entanglement measures to bound the complexity entropy, proving the results in Sec. VC. We restrict our attention to a spatially one-dimensional chain  $S$  of  $n \geq 2$  qubits:  $S = S_1 S_2 \dots S_n$ . Let  $\{\mathbf{M}^r\}$  denote the family of POVM-effect-complexity sets (Definition 6) defined by Eq. (11), wherein the computational gate set  $\mathcal{G}$  contains only unitary operations that can act nontrivially only on two neighboring qubits in  $S$ . Furthermore, we define the *potential entangling power* of  $\mathcal{G}$  (cf. Ref. [27]) as

$$\begin{aligned} e(\mathcal{G}) &:= \sup_{U \in \mathcal{G}} \{e(U)\}, \quad \text{wherein } e(U) := \min\{\|H\| : \\ &\quad H = H^\dagger, \exists \chi \in \mathbb{R} \text{ such that } e^{-iH} = e^{-i\chi} U\}. \end{aligned} \quad (\text{G1})$$

In general,  $e(\mathcal{G})$  can be as large as  $\pi$ . However, our entanglement bounds are most interesting when  $\mathcal{G}$  contains only operations that act weakly and remain near the identity operator (up to an overall phase). In this case,  $e(\mathcal{G}) \approx e(\mathbb{1}) = 0$ .

In the following, we use the *quantum mutual information*. Let  $A$  and  $B$  denote distinct quantum systems, and let  $\sigma_{AB}$  denote any state of  $AB$ . The quantum mutual information of  $\sigma_{AB}$  is defined in terms of the von Neumann entropy as

$$\begin{aligned} I(A : B)_\sigma &:= H(A)_\sigma + H(B)_\sigma - H(AB)_\sigma \\ &= H(A)_\sigma - H(A|B)_\sigma. \end{aligned} \quad (\text{G2})$$

For any state  $\rho$  of the qubit chain  $S$ , we define the entanglement measure

$$E(\rho) := \frac{1}{n-1} \sum_{j=1}^{n-1} I(S_1 \dots S_j : S_{j+1} \dots S_n)_\rho. \quad (\text{G3})$$

$E(\rho)$  quantifies the average amount of correlation, as quantified by the mutual information, between any bipartition of  $S$  into subsystems of neighboring qubits. We now prove the bounds in Proposition 2 and in Eq. (45).

*Theorem 7 (Continuity bound on  $E(\rho)$  under one operation).* Let  $\rho$  denote any quantum state of  $S$ . Let  $U$  denote any unitary that can act nontrivially only on two neighboring qubits in  $S$ . It holds that

$$\begin{aligned} & |E(U\rho U^\dagger) - E(\rho)| \\ & \leq \frac{1}{n-1} \min \left\{ 8 \log(2), 8\nu \log(2) \right. \\ & \quad \left. + 3(1+\nu) h\left(\frac{\nu}{1+\nu}\right) \right\}. \end{aligned} \quad (\text{G4})$$

The  $h(x) := -x \log(x) - (1-x) \log(1-x)$  denotes the binary entropy function, and

$$\nu := \sin\left(\min\left\{e(U), \frac{\pi}{2}\right\}\right). \quad (\text{G5})$$

*Proof.* Let  $\rho'$  denote the state that results from evolving  $\rho$  under  $U$ :  $\rho' := U\rho U^\dagger$ . For all  $j \in \{1, 2, \dots, n-$

$1\}$ , we write, for short,  $A_j := S_1 S_2 \dots S_j$  and  $B_{j+1} := S_{j+1} S_{j+2} \dots S_n$ . By Eq. (G3),

$$\begin{aligned} & E(\rho') - E(\rho) \\ & = \frac{1}{n-1} \sum_{j=1}^{n-1} [I(A_j : B_{j+1})_{\rho'} - I(A_j : B_{j+1})_\rho]. \end{aligned} \quad (\text{G6})$$

Let  $k, k+1 \in \{1, 2, \dots, n\}$  denote two qubits on which  $U$  can act nontrivially. For all  $j \neq k$ ,  $I(A_j : B_{j+1})_{\rho'} = I(A_j : B_{j+1})_\rho$ ;  $U$  can act nontrivially on qubits only in  $A_j$  or only in  $B_{j+1}$ , so

$$\begin{aligned} I(A_j : B_{j+1})_\rho &= H(A_j)_\rho + H(B_{j+1})_\rho - H(A_j B_{j+1})_\rho \\ &= H(A_j)_{\rho'} + H(B_{j+1})_{\rho'} - H(A_j B_{j+1})_{\rho'} \\ &= I(A_j : B_{j+1})_{\rho'}. \end{aligned} \quad (\text{G7})$$

The first and last equalities follow from Eq. (G2). The middle equality follows because the von Neumann entropy is unitarily invariant. Thus,

$$E(\rho') - E(\rho) = \frac{1}{n-1} [I(A_k : B_{k+1})_{\rho'} - I(A_k : B_{k+1})_\rho]. \quad (\text{G8})$$

Using the chain rule for the von Neumann entropy, we simultaneously re-express  $I(A_k : B_{k+1})_\rho$  and  $I(A_k : B_{k+1})_{\rho'}$ :

$$\begin{aligned} & I(A_k : B_{k+1}) \\ &= H(A_k) + H(B_{k+1}) - H(A_k B_{k+1}) \\ &= H(A_{k-1}) + H(S_k | A_{k-1}) + H(B_{k+2}) + H(S_{k+1} | B_{k+2}) - H(A_{k-1} B_{k+2}) - H(S_k S_{k+1} | A_{k-1} B_{k+2}) \\ &= \underbrace{H(S_k | A_{k-1}) + H(S_{k+1} | B_{k+2}) - H(S_k S_{k+1} | A_{k-1} B_{k+2})}_{(\text{I})} + \underbrace{H(A_{k-1}) + H(B_{k+2}) - H(A_{k-1} B_{k+2})}_{(\text{II})}. \end{aligned} \quad (\text{G9})$$

$A_{k-1}(B_{k+2})$  is trivial if  $k=1$  ( $k=n-1$ ). Since  $U$  can act nontrivially only on  $S_k S_{k+1}$ , the reduced states of  $\rho$  and  $\rho'$  on  $A_{k-1} B_{k+2}$  coincide. Hence, the terms in (II) have the same value whether evaluated on  $\rho$  or on  $\rho'$ , so (G8) equals the difference between the terms in (I) evaluated on  $\rho$  and on  $\rho'$ . We now upper-bound the absolute value of this difference. Let  $\mathcal{U}(\cdot) := U(\cdot)U^\dagger$ . By the diamond norm's definition (A1) and by Lemma 2,

$$\begin{aligned} & \frac{1}{2} \|\rho' - \rho\|_1 = \frac{1}{2} \|(\mathcal{U} - \text{id})\rho\|_1 \leq \frac{1}{2} \|\mathcal{U} - \text{id}\|_\diamond \\ & = \sin\left(\min\left\{e(U), \frac{\pi}{2}\right\}\right) = \nu. \end{aligned} \quad (\text{G10})$$

By the Alicki-Fannes-Winter continuity bound on the conditional entropy [139,140] (see Lemma 2 of Ref. [140]),

Eq. (G10) implies that the terms in (I) obey the following bounds:

$$\begin{aligned} & |H(S_k | A_{k-1})_{\rho'} - H(S_k | A_{k-1})_\rho| \\ & \leq 2\nu \log(d_{S_k}) + (1+\nu) h\left(\frac{\nu}{1+\nu}\right), \end{aligned} \quad (\text{G11a})$$

$$\begin{aligned} & |H(S_{k+1} | B_{k+2})_{\rho'} - H(S_{k+1} | B_{k+2})_\rho| \\ & \leq 2\nu \log(d_{S_{k+1}}) + (1+\nu) h\left(\frac{\nu}{1+\nu}\right), \end{aligned} \quad (\text{G11b})$$

$$\begin{aligned} & \text{and } |H(S_k S_{k+1} | A_{k-1} B_{k+2})_{\rho'} - H(S_k S_{k+1} | A_{k-1} B_{k+2})_\rho| \\ & \leq 2\nu \log(d_{S_k} d_{S_{k+1}}) + (1+\nu) h\left(\frac{\nu}{1+\nu}\right). \end{aligned} \quad (\text{G11c})$$

Therefore,

$$\begin{aligned} & |I(A_k : B_{k+1})_{\rho'} - I(A_k : B_{k+1})_{\rho}| \\ & \leq 2\nu \log(d_{S_k}^2 d_{S_{k+1}}^2) + 3(1 + \nu) h\left(\frac{\nu}{1 + \nu}\right) \\ & = 8\nu \log(2) + 3(1 + \nu) h\left(\frac{\nu}{1 + \nu}\right). \end{aligned} \quad (\text{G12})$$

The equality follows because  $S_k$  and  $S_{k+1}$  are qubits:  $d_{S_k} = d_{S_{k+1}} = 2$ . Moreover, each term in (I) can vary by at most the range of the term's corresponding conditional entropy, so

$$\begin{aligned} & |I(A_k : B_{k+1})_{\rho'} - I(A_k : B_{k+1})_{\rho}| \\ & \leq 2 \log(d_{S_k}) + 2 \log(d_{S_{k+1}}) + 2 \log(d_{S_k} d_{S_{k+1}}) \\ & = 2 \log(d_{S_k}^2 d_{S_{k+1}}^2) = 8 \log(2). \end{aligned} \quad (\text{G13})$$

Combining Eqs. (G8), (G12), and (G13) yields

$$\begin{aligned} & |E(\rho') - E(\rho)| \leq \frac{1}{n-1} \\ & \min \left\{ 8 \log(2), 8\nu \log(2) + 3(1 + \nu) h\left(\frac{\nu}{1 + \nu}\right) \right\}. \end{aligned} \quad (\text{G14})$$

■

It is convenient to define the worst-case upper bound in Eq. (G4) among all the gates in  $\mathcal{G}$ . Namely, let

$$\begin{aligned} \mu(\mathcal{G}) &:= \frac{1}{n-1} \\ & \min \left\{ 8 \log(2), 8\nu_{\mathcal{G}} \log(2) + 3(1 + \nu_{\mathcal{G}}) h\left(\frac{\nu_{\mathcal{G}}}{1 + \nu_{\mathcal{G}}}\right) \right\}, \\ & \text{where } \nu_{\mathcal{G}} := \sin\left(\min\left\{e(\mathcal{G}), \frac{\pi}{2}\right\}\right). \end{aligned} \quad (\text{G15})$$

By definition,  $\mu(\mathcal{G})$  always satisfies  $\mu(\mathcal{G}) \leq 8 \log(2)/(n-1)$ .

**Theorem 8 (Entanglement bound on the complexity entropy).** Let  $\rho$  denote any quantum state of  $S$ . Let  $r \geq 0$  and  $\eta \in (0, 1]$ . It holds that

$$\begin{aligned} H_{\text{H}}^{r, \eta}(\rho) & \geq \frac{1}{\eta} [E(\rho) - r \mu(\mathcal{G}) + H(\rho) \\ & - 2h(\eta) - (1 - \eta) n \log(2)] - 2 \log(\eta). \end{aligned} \quad (\text{G16})$$

The  $h(x) := -x \log(x) - (1-x) \log(1-x)$  denotes the binary entropy function, and  $H_{\text{H}}^{r, \eta}$  is defined with respect to  $\mathcal{M}^r$ .

*Proof.* Consider any  $\zeta > 0$ . There exists a  $Q \in \mathcal{M}^r$  such that  $\text{tr}(Q\rho) \geq \eta$  and  $\log(\text{tr}(Q)/\text{tr}(Q\rho)) \leq H_{\text{H}}^{r, \eta}(\rho) + \zeta$ .  $Q = U^\dagger P U$  for some effect  $P \in \mathcal{M}^{r=0}$  and for some unitary  $U$  satisfying  $C_{\mathcal{G}}(U_0) \leq r$ . Let  $\rho' := U\rho U^\dagger$ .  $\text{tr}(P\rho') = \text{tr}(Q\rho) \geq \eta$ , so  $P$  is a candidate for the  $H_{\text{H}}^{r=0, \eta}(\rho')$  optimization:

$$\begin{aligned} H_{\text{H}}^{r=0, \eta}(\rho') & \leq \log\left(\frac{\text{tr}(P)}{\text{tr}(P\rho')}\right) = \log\left(\frac{\text{tr}(Q)}{\text{tr}(Q\rho)}\right) \\ & \leq H_{\text{H}}^{r, \eta}(\rho) + \zeta. \end{aligned} \quad (\text{G17})$$

For all  $j \in \{1, 2, \dots, n-1\}$ , we write, for short,  $A_j := S_1 S_2 \dots S_j$  and  $B_{j+1} := S_{j+1} S_{j+2} \dots S_n$ . Since every POVM effect in  $\mathcal{M}^{r=0}$  is tensor-product, Proposition 18 implies that, for all  $j$ ,

$$\begin{aligned} H_{\text{H}}^{r=0, \eta}(\rho') & = -D_{\text{H}}^{r=0, \eta}(\rho' \parallel \mathbb{1}_S) \geq -D_{\text{H}}^{\eta}(\rho'_{A_j} \parallel \mathbb{1}_{A_j}) \\ & \quad - D_{\text{H}}^{\eta}(\rho'_{B_{j+1}} \parallel \mathbb{1}_{B_{j+1}}). \end{aligned} \quad (\text{G18})$$

Consequently,

$$\begin{aligned} & H_{\text{H}}^{r=0, \eta}(\rho') \\ & \geq \frac{1}{n-1} \sum_{j=1}^{n-1} \left[ -D_{\text{H}}^{\eta}(\rho'_{A_j} \parallel \mathbb{1}_{A_j}) - D_{\text{H}}^{\eta}(\rho'_{B_{j+1}} \parallel \mathbb{1}_{B_{j+1}}) \right]. \end{aligned} \quad (\text{G19})$$

By Proposition 4.67 of Ref. [88], one can bound each hypothesis-testing relative entropy in Eq. (G19) in terms of a standard (Umegaki) quantum relative entropy: for each  $A_j$ ,

$$\begin{aligned} & -D_{\text{H}}^{\eta}(\rho'_{A_j} \parallel \mathbb{1}_{A_j}) \\ & \geq -\frac{1}{\eta} \left[ D(\rho'_{A_j} \parallel \mathbb{1}_{A_j}) + h(\eta) + (1 - \eta) \log(\text{tr}(\mathbb{1}_{A_j})) \right] \\ & \quad - \log(\eta) \\ & = \frac{1}{\eta} \left[ H(\rho'_{A_j}) - h(\eta) - j(1 - \eta) \log(2) \right] - \log(\eta). \end{aligned} \quad (\text{G20})$$

Similarly, for each  $B_{j+1}$ ,

$$\begin{aligned} & -D_{\text{H}}^{\eta}(\rho'_{B_{j+1}} \parallel \mathbb{1}_{B_{j+1}}) \\ & \geq \frac{1}{\eta} \left[ H(\rho'_{B_{j+1}}) - h(\eta) - (n-j)(1 - \eta) \log(2) \right] \\ & \quad - \log(\eta). \end{aligned} \quad (\text{G21})$$

Thus,

$$\begin{aligned}
H_H^{r,\eta}(\rho) + \zeta &\geq H_H^{r=0,\eta}(\rho') \\
&\geq \frac{1}{n-1} \sum_{j=1}^{n-1} \left\{ \frac{1}{\eta} \left[ H(\rho'_{A_j}) + H(\rho'_{B_{j+1}}) - 2h(\eta) - (1-\eta)n \log(2) \right] - 2 \log(\eta) \right\} \\
&= \frac{1}{\eta} \left\{ \frac{1}{n-1} \sum_{j=1}^{n-1} [I(A_j : B_{j+1})_{\rho'} + H(\rho')] - 2h(\eta) - (1-\eta)n \log(2) \right\} - 2 \log(\eta) \\
&= \frac{1}{\eta} [E(\rho') + H(\rho') - 2h(\eta) - (1-\eta)n \log(2)] - 2 \log(\eta) \\
&\geq \frac{1}{\eta} [E(\rho) - r\mu(\mathcal{G}) + H(\rho) - 2h(\eta) - (1-\eta)n \log(2)] - 2 \log(\eta). \tag{G22}
\end{aligned}$$

The first and second equalities follow from the definitions (G2) and (G3), respectively. The last inequality follows because  $E(\rho') - E(\rho) \geq -r\mu(\mathcal{G})$ , by Theorem 7, and because  $H(\rho') = H(\rho)$ , since the von Neumann entropy is unitarily invariant. Finally, Eq. (G22) implies Eq. (G16), since  $\zeta$  is arbitrary. ■

## APPENDIX H: BOUNDS FROM SHORT-TIME HAMILTONIAN DYNAMICS

Here, we apply the bounds in Appendix G to short-time Hamiltonian dynamics. We consider a nearest-neighbor Hamiltonian  $H$  acting on a 1D chain  $S$  of  $n \geq 2$  qubits. The chain is initially in a state  $\rho$ . We show that the entanglement measure  $E(\rho)$  can grow only slowly in time.

*Corollary 4 (Continuity bound on  $E(\rho)$  under local Hamiltonian evolution).* Let  $\rho$  denote any quantum state of  $S$ . Let  $H$  denote a nearest-neighbor, translationally invariant Hamiltonian governed by local-interaction terms  $h$  that can act on only two qubits at a time. It holds that

$$\frac{d}{dt} |E(e^{-iHt}\rho e^{iHt}) - \rho| \leq C(n-1)\|h\|, \tag{H1}$$

for some universal constant  $C > 0$ .

*Proof.* The Hamiltonian has the form

$$H = \sum_{j=1}^{n-1} \tau_j(h), \tag{H2}$$

wherein  $h$  denotes a fixed interaction term that acts non-trivially only on two neighboring qubits in  $S$ . For all  $j \in \{1, 2, \dots, n-1\}$ ,  $\tau_j$  is the shift operator that places  $\tau_j(h)$  onto qubits  $j$  and  $j+1$ . By the Trotter formula, it is manifest that for suitably small times  $t > 0$ ,  $e^{-iHt}\rho e^{iHt}$

can be arbitrarily well approximated by a brickwork circuit involving nearest-neighbor quantum gates, even for noncommuting  $\tau_j(h)$  for different  $j$ . For each cut specified by  $j$ , the small incremental entangling bound for the von-Neumann entropy [27,141] gives

$$\frac{d}{dt} I(S_1 \dots S_j : S_{j+1} \dots S_n)_\rho \leq C\|h\|. \tag{H3}$$

Summing this inequality over all cuts yields Eq. (H1). The constant  $C$  is proven in Ref. [141] to be  $C = 22 \log(2)$ . The 2 originates from the qubits' local dimension. In numerical studies, it is found to be rather tightly given by  $C = 22 \log(2)$ . ■

The linear growth of  $E(\rho)$  in time is tight, in the sense that there exist local nearest-neighbor Hamiltonians that exhibit such growth for suitably short times. This claim largely follows from the results of Ref. [142], which build on those in Ref. [143]. Intuitively speaking, the linear growth originates from a light-cone-like entanglement dynamics governed by Lieb-Robinson bounds. For free bosons and fermions (called noninteracting bosons and fermions in this context), such bounds are readily computable.

*Corollary 5 (Lower bound on  $E(\rho)$  under local Hamiltonian evolution).* Let  $\rho$  denote any quantum state of  $S$ . There exist local Hamiltonians and product initial conditions  $\rho$ , with  $E(\rho) = 0$ , such that

$$E(e^{-iHt}\rho e^{iHt}) \geq (n-1) \left[ \frac{4}{3\pi} - \frac{1}{2} \log(t) - 1 \right]. \tag{H4}$$

*Proof.* Let  $S$  consist of an odd number  $n \geq 21$  of qubits. Let  $H$  be the Ising Hamiltonian with periodic boundary conditions. Let the initial state be  $\rho =$



$|1, 1, \dots, 1\rangle\langle 1, 1, \dots, 1|$ , so that  $E(\rho) = 0$ . The time-evolved state is  $\rho(t) = e^{-iHt}\rho e^{iHt}$ . For this state, bounds on entanglement entropies for sufficiently large subsystems are known. Since we consider only bipartite cuts of  $S$ , the subsystem of length  $L$  considered in Ref. [142] always satisfies  $L = (n - 1)/2$ . Hence, the constraint  $L \geq 10$  is always satisfied. It follows from Theorem 1 in Ref. [142], applied to the  $n - 1$  cuts of  $S$ , that

$$E(\rho(t)) \geq (n - 1) \left[ \frac{4}{3\pi} - \frac{1}{2} \log(t) - 1 \right]. \quad (\text{H5})$$

■

## APPENDIX I: RECOVERING THE HYPOTHESIS-TESTING (RELATIVE) ENTROPY AT HIGH COMPLEXITIES

Consider an “unrestricted” agent who can render POVM effects of arbitrarily high complexities. In this Appendix, we formulate conditions under which the complexity entropy approximates the hypothesis-testing entropy. We first identify the set of measurements that an unrestricted agent can render with arbitrary precision.

*Definition 20 (POVM effects accessible with unbounded complexity).* Let  $\{\mathbf{M}^r\}$  denote any family of POVM-effect-complexity sets (Definition 6). The set of *POVM effects accessible with unbounded complexity* is

$$\mathbf{M}^\infty := \overline{\bigcup_{r \geq 0} \mathbf{M}^r}. \quad (\text{I1})$$

The  $\bar{A}$  denotes the topological closure of a set  $A$ .

*Definition 21 (“Complexity-unrestricted” entropy and relative entropy).* Let  $\mathbf{M}^\infty$  denote any set of POVM effects defined by Eq. (I1). Let  $\rho$  denote any subnormalized state; and  $\Gamma$ , any positive-semidefinite operator. Let  $\eta \in (0, \text{tr}(\rho)]$ . We define the “complexity-unrestricted” relative entropy  $D_{\text{H}}^{\infty, \eta}(\rho \parallel \Gamma)$  by replacing  $\mathbf{M}^r$  with  $\mathbf{M}^\infty$  in the definition of  $D_{\text{H}}^{r, \eta}(\rho \parallel \Gamma)$  (Definition 13). The “complexity-unrestricted” entropy is  $H_{\text{H}}^{\infty, \eta}(\rho) := -D_{\text{H}}^{\infty, \eta}(\rho \parallel \mathbb{1})$ .

If  $\mathbf{M}^\infty$  contains all POVM effects, we recover the hypothesis-testing (relative) entropy exactly:  $H_{\text{H}}^{\infty, \eta}(\rho) = H_{\text{H}}^{\eta}(\rho)$  [ $D_{\text{H}}^{\infty, \eta}(\rho \parallel \Gamma) = D_{\text{H}}^{\eta}(\rho \parallel \Gamma)$ ]. Yet, for many natural families  $\{\mathbf{M}^r\}$  of POVM-effect-complexity sets,  $\mathbf{M}^\infty$  contains only a limited portion of all POVM effects. For instance, let  $C$  denote a unitary-complexity measure (Definition 3), and let  $\mathbf{P}$  denote a set of simple POVM effects (Definition 7) consisting only of projectors. Suppose that  $\{\mathbf{M}^r\} = \{\mathbf{M}^r(\mathbf{P}, C)\}$ , as per Eq. (B15). Since every unitary transformation maps projectors to projectors,

$\mathbf{M}^\infty$  is a set of projectors and therefore excludes almost all POVM effects.

However, even when limited in form,  $\mathbf{M}^\infty$  may contain enough effects to ensure that  $H_{\text{H}}^{r, \eta}(\rho)$  converges to a value near  $H_{\text{H}}^{\eta}(\rho)$  as  $r \rightarrow \infty$ . We now formulate conditions under which  $H_{\text{H}}^{\infty, \eta}(\rho)$  [ $D_{\text{H}}^{\infty, \eta}(\rho \parallel \Gamma)$ ] lies close to the hypothesis-testing (relative) entropy,  $H_{\text{H}}^{\eta}(\rho)$  [ $D_{\text{H}}^{\eta}(\rho \parallel \Gamma)$ ].

*Definition 22 ( $q$ -quasiuniversal sets of POVM effects).*

Let  $q \geq 0$ , and let  $\Gamma$  denote any positive-semidefinite operator. Let  $\mathbf{M}^\infty$  denote any set of POVM effects defined by Eq. (I1).  $\mathbf{M}^\infty$  is  *$q$ -quasiuniversal with respect to  $\Gamma$*  if, for all POVM effects  $Q$  with  $\|Q\| = 1$  and for all probability weights  $g \in (0, 1)$ , there exist accessible effects  $\tilde{Q}, \tilde{Q}' \in \mathbf{M}^\infty$  such that

$$g\tilde{Q} \leq Q \leq (1 - g)\tilde{Q}' + g\mathbb{1} \quad (\text{I2a})$$

$$\text{and } \log(\text{tr}(\tilde{Q}'\Gamma)) - \log(\text{tr}(\tilde{Q}\Gamma)) \leq q. \quad (\text{I2b})$$

Consider any  $Q$  with  $\|Q\| = 1$ . Whenever  $\mathbf{M}^\infty$  is quasiuniversal (for any  $q$ ), one can “underestimate”  $Q$  by forming a convex mixture from an accessible effect  $\tilde{Q} \in \mathbf{M}^\infty$  and the trivial effect 0. Similarly, one can “overestimate”  $Q$  by mixing an accessible  $\tilde{Q}'$  and the trivial effect  $\mathbb{1}$ . One can estimate  $Q$  with such mixtures, no matter the mixing probability  $g$ . The variable  $g$  continuously parametrizes inequalities that approach  $0 \leq Q \leq \tilde{Q}'$  as  $g \rightarrow 0$  and  $\tilde{Q} \leq Q \leq \mathbb{1}$  as  $g \rightarrow 1$ . These boundaries provide intuition about the parametrization: as one requirement (that  $\tilde{Q}$  lower-bound  $Q$ ) tightens, the other requirement (that  $\tilde{Q}'$  upper-bound  $Q$ ) relaxes and vice versa. We exclude the probability weight  $g = 0$  ( $g = 1$ ) to avoid trivializing the role of  $\tilde{Q}$  ( $\tilde{Q}'$ ).

The estimations  $\tilde{Q}$  and  $\tilde{Q}'$  lie close to  $Q$  because they lie close to one another, by condition (I2b): their candidate hypothesis-testing-entropy values  $\log(\text{tr}(\tilde{Q}\Gamma)/\eta)$  and  $\log(\text{tr}(\tilde{Q}'\Gamma)/\eta)$  differ by  $\leq q$  for every error intolerance  $\eta \in (0, 1]$ . Without Eq. (I2b), one might trivially satisfy Eq. (I2a) with  $\tilde{Q} = 0$  and  $\tilde{Q}' = \mathbb{1}$ . Finally, as expected,  $\mathbf{M}^\infty$  is  $q$ -quasiuniversal for every  $q$  and with respect to every  $\Gamma$ , whenever the set contains all POVM effects. In this case, one can choose  $\tilde{Q} = \tilde{Q}' = Q \in \mathbf{M}^\infty$ .

The hypothesis-testing (relative) entropy bounds the complexity-unrestricted (relative) entropy from above and below.

*Theorem 9 (Recoverability of the hypothesis-testing entropy).* Let  $\rho$  denote any subnormalized state; and  $\Gamma$ , any positive-semidefinite operator. Let  $\mathbf{M}^\infty$  denote any set of POVM effects defined by (I1). Let  $q \geq 0$ ,  $\eta \in (0, \text{tr}(\rho)]$ , and  $g \in (0, 1)$ . Suppose that  $\mathbf{M}^\infty$  is  $q$ -quasiuniversal with

respect to  $\Gamma$  (Definition 22). Then

$$D_H^\eta(\rho \parallel \Gamma) \geq D_H^{\infty, \eta}(\rho \parallel \Gamma) \geq D_H^{\eta_g}(\rho \parallel \Gamma) - q - \log\left(\frac{\text{tr}(\rho)}{g\eta}\right), \quad (13)$$

wherein  $\eta_g := (1 - g)\eta + g \text{tr}(\rho) \geq \eta$ . If  $\mathbf{M}^\infty$  is  $q$ -quasiuniversal with respect to  $\mathbb{1}$ , then

$$H_H^\eta(\rho) \leq H_H^{\infty, \eta}(\rho) \leq H_H^{\eta_g}(\rho) + q + \log\left(\frac{\text{tr}(\rho)}{g\eta}\right). \quad (14)$$

*Proof.* The  $D_H^\eta(\rho \parallel \Gamma)$  optimization ranges over all POVM effects, while the  $D_H^{\infty, \eta}(\rho \parallel \Gamma)$  optimization ranges over only effects in  $\mathbf{M}^\infty$ . Therefore,  $D_H^\eta(\rho \parallel \Gamma) \geq D_H^{\infty, \eta}(\rho \parallel \Gamma)$ . Now, we prove the second inequality in Eq. (13). Let  $Q$  denote an optimal POVM effect for  $D_H^{\eta_g}(\rho \parallel \Gamma)$  in Eq. (C4). Without loss of generality,  $\|Q\| = 1$ , since we could otherwise replace  $Q$  by another candidate  $\tilde{Q} := Q/\|Q\| \geq Q$  that achieves the same objective value as  $Q$ . Moreover,  $\text{tr}(Q\rho) \geq \eta_g$  and  $\text{tr}(Q\Gamma)/\text{tr}(Q\rho) = \exp(-D_H^{\eta_g}(\rho \parallel \Gamma))$ . Since  $\mathbf{M}^\infty$  is  $q$ -quasiuniversal, some  $\tilde{Q}, \tilde{Q}' \in \mathbf{M}^\infty$  satisfy the conditions (I2). Condition (I2a) implies that

$$(1 - g) \text{tr}(\tilde{Q}'\rho) \geq \text{tr}(Q\rho) - g \text{tr}(\rho) \geq \eta_g - g \text{tr}(\rho) = (1 - g)\eta. \quad (15)$$

Equivalently,  $\text{tr}(\tilde{Q}'\rho) \geq \eta$ . Therefore,  $\tilde{Q}'$  is a candidate for the  $D_H^{\infty, \eta}(\rho \parallel \Gamma)$  optimization in Eq. (D1), so

$$\begin{aligned} D_H^{\infty, \eta}(\rho \parallel \Gamma) &\geq -\log\left(\frac{\text{tr}(\tilde{Q}'\Gamma)}{\text{tr}(\tilde{Q}'\rho)}\right) \geq -\log\left(\frac{\text{tr}(\tilde{Q}'\Gamma)}{\eta}\right) \\ &\geq -\log\left(e^q \frac{\text{tr}(\tilde{Q}'\Gamma)}{\eta}\right) \geq -\log\left(\frac{\text{tr}(Q\Gamma)}{g\eta}\right) - q \\ &\geq -\log\left(\frac{\text{tr}(Q\Gamma)}{\text{tr}(Q\rho)} \frac{\text{tr}(\rho)}{g\eta}\right) - q \\ &= D_H^{\eta_g}(\rho \parallel \Gamma) - q - \log\left(\frac{\text{tr}(\rho)}{g\eta}\right). \end{aligned} \quad (16)$$

The third inequality holds because  $\text{tr}(\tilde{Q}'\Gamma) \leq e^q \text{tr}(\tilde{Q}\Gamma)$ , by condition (I2b). The fourth inequality follows from condition (I2a); the fifth from  $\text{tr}(Q\rho) \leq \text{tr}(\rho)$ . ■

We now provide an example of quasiuniversal set that arises naturally in quantum computation.

*Proposition 31 (Example of  $q$ -quasiuniversal set).* Let  $\mathcal{G}$  denote any set of unitary gates that is universal for quantum computation. Let  $C_{\mathcal{G}}$  be the unitary-circuit-complexity

measure associated with  $\mathcal{G}$  (Definition 2). Let  $\mathbf{P}$  denote a set of simple POVM effects (Definition 7) that consists only of projectors and contains at least one rank-1 projector. Let  $\{\mathbf{M}^r = \mathbf{M}^r(\mathbf{P}, C_{\mathcal{G}})\}$  denote the family of POVM-effect-complexity sets defined by Eq. (B15). Let  $\mathbf{M}^\infty$  denote the set of POVM effects defined in Eq. (II), and assume that the effects in  $\mathbf{M}^\infty$  act on a Hilbert space of dimensionality  $d$ .  $\mathbf{M}^\infty$  is  $q$ -quasiuniversal with respect to  $\mathbb{1}$  (Definition 22). The  $q$  denotes the largest gap within the set  $\mathcal{X} := \{\log(\text{tr}(P)) : P \in \mathbf{P}\} \subset [0, \log(d)]$ :

$$q := \sup\{|b - a| : (a, b) \subset [0, \log(d)] \setminus \mathcal{X}\}. \quad (17)$$

$q$  is the greatest length of any interval  $(a, b) \subset [0, \log(d)]$  disjoint from  $\mathcal{X}$ .

*Proof.* Let  $Q$  denote any POVM effect with  $\|Q\| = 1$ . Let us fix a matrix representation of  $Q$  that is diagonal and has decreasing diagonal elements:  $Q = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_d)$ , with  $1 = \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d \geq 0$ . Consider any  $g \in (0, 1)$ , and let  $\ell_g := \max\{i : \lambda_i \geq g\}$  denote the number of  $Q$  eigenvalues not less than  $g$ .  $\ell_g$  satisfies  $1 \leq \ell_g \leq d$ , since  $\lambda_1 = 1 \geq g$ . Let

$$\begin{aligned} \ell_g^- &:= \max\{\text{rank}(P) : P \in \mathbf{P}, \text{rank}(P) \leq \ell_g\} \quad \text{and} \\ \ell_g^+ &:= \min\{\text{rank}(P) : P \in \mathbf{P}, \text{rank}(P) \geq \ell_g\}. \end{aligned} \quad (18)$$

As per the definitions of  $\ell_g^-$  and  $\ell_g^+$ , there exist projectors  $\tilde{P}, \tilde{P}' \in \mathbf{P}$  such that  $\text{rank}(\tilde{P}) = \ell_g^-$  and  $\text{rank}(\tilde{P}') = \ell_g^+$ . Moreover,  $1 \leq \ell_g^- \leq \ell_g^+ \leq d$ , since  $\mathbf{P}$  contains both a rank-1 projector and  $\mathbb{1}$ . Let  $U$  denote a unitary such that  $\tilde{Q} = U^\dagger \tilde{P} U$  is diagonal:  $\tilde{Q} = \text{diag}(1, \dots, 1, 0, \dots, 0)$ , with  $\text{rank}(\tilde{Q}) = \ell_g^-$ . Likewise, let  $U'$  denote a unitary such that  $\tilde{Q}' = U'^\dagger \tilde{P}' U'$  is diagonal, with  $\text{rank}(\tilde{Q}') = \ell_g^+$ . By assumption,  $U$  and  $U'$  are arbitrarily well approximated by unitaries of finite complexity, so  $\tilde{Q}, \tilde{Q}' \in \mathbf{M}^\infty$ . We now show that  $\tilde{Q}$  and  $\tilde{Q}'$  satisfy the conditions (I2) in the definition of a  $q$ -quasiuniversal set.

Consider the operator  $g\tilde{Q}$ . The first  $\ell_g \geq \ell_g^-$  diagonal entries of  $g\tilde{Q}$  do not exceed  $g$  and, hence, do not exceed the first  $\ell_g$  diagonal entries of  $Q$ . Moreover, the last  $\ell_g$  diagonal entries of  $g\tilde{Q}$  equal 0 and, hence, do not exceed the last  $\ell_g$  diagonal entries of  $Q$ . Thus,  $g\tilde{Q} \leq Q$ . Now, consider the operator  $(1 - g)\tilde{Q}' + g\mathbb{1}$ . The first  $\ell_g \leq \ell_g^+$  diagonal entries of  $(1 - g)\tilde{Q}' + g\mathbb{1}$  equal 1 and, hence, are not less than the first  $\ell_g$  diagonal entries of  $Q$ . Moreover, the last  $\ell_g$  diagonal entries of  $g\tilde{Q}$  are not less than  $g$  and, hence, are not less than the last  $\ell_g$  diagonal entries of  $Q$ . Thus,  $Q \leq (1 - g)\tilde{Q}' + g\mathbb{1}$ . As such,  $\tilde{Q}$  and  $\tilde{Q}'$  satisfy condition (I2a).

Compare the definitions of  $\ell_g^-$  and  $\ell_g^+$ . There exists no projector  $P \in \mathbf{P}$  such that  $\ell_g^- < \text{rank}(P) < \ell_g^+$ .

$\ell_g^+$ . Accordingly, the interval  $(\log(\ell_g^-), \log(\ell_g^+))$  is disjoint from  $\mathcal{X}$ . [If  $\ell_g^- = \ell_g^+$ , then  $(\log(\ell_g^-), \log(\ell_g^+)) = \emptyset$  by convention.] Thus,  $\log(\text{tr}(\tilde{Q}')) - \log(\text{tr}(\tilde{Q})) = \log(\ell_g^+) - \log(\ell_g^-) \leq q$ , with  $q$  defined as in Eq. (I7), so  $\tilde{Q}$  and  $\tilde{Q}'$  satisfy condition (I2b). ■

## APPENDIX J: DATA COMPRESSION UNDER COMPLEXITY LIMITATIONS

Here, we establish optimal protocols for data compression under complexity limitations. We thereby prove the bounds (50) in Sec. V E.

Let  $\rho$  denote any state, and  $C$  any unitary-complexity measure, of  $n$  qubits. Let  $r \geq 0$ . Let  $m \leq n$  denote a number of qubits; and  $\varepsilon \in [0, 1)$ , an error parameter. We define an  $(m, r, \varepsilon)$ -data-compression protocol as a unitary  $U$  that satisfies  $C(U) \leq r$  and that can compress  $\rho$  into  $m$  qubits. The compression succeeds if the other  $n - m$  qubits end in a state  $\varepsilon$ -close to  $|0^{n-m}\rangle$  in fidelity: there exists a subset  $\mathcal{W} \subset \{1, 2, \dots, n\}$  of qubits such that  $|\mathcal{W}| = m$  and

$$\text{tr}(\text{tr}_{\mathcal{W}}(U\rho U^\dagger) |0^{n-m}\rangle\langle 0^{n-m}|) \geq 1 - \varepsilon. \quad (\text{J1})$$

We first prove that some protocol compresses  $\rho$  into a number of qubits proportional to the reduced complexity entropy (Definition 16).

*Theorem 10 (Achievability of data compression).* Let  $\rho$  denote any state, and  $C$  any unitary-complexity measure (Definition 3), of  $n$  qubits. Let  $\mathbf{P}$  denote the set of simple POVM effects (Definition 7) defined in Eq. (10). Let  $r \geq 0$  and  $\varepsilon \in [0, 1)$ . Let  $\mathbf{M}^r = \mathbf{M}^r(\mathbf{P}, C)$  denote the POVM-effect-complexity set defined in Eq. (B15). There exists an  $(m, r, \varepsilon)$ -data-compression protocol for  $\rho$  such that

$$m = \frac{1}{\log(2)} H_{\text{h}}^{r, 1-\varepsilon}(\rho). \quad (\text{J2})$$

The reduced complexity entropy  $H_{\text{h}}^{r, 1-\varepsilon}$  is defined with respect to  $\mathbf{M}^r$ .

The above theorem can be reformulated in terms of the complexity entropy  $H_{\text{H}}^{r, 1-\varepsilon}(\rho)$ . Proposition 22 shows that the reduced complexity entropy  $H_{\text{h}}^{r, 1-\varepsilon}(\rho)$  differs from the complexity entropy  $H_{\text{H}}^{r, 1-\varepsilon}(\rho)$  by at most  $\log(1/[1 - \varepsilon])$ . Combining this fact with Theorem 10 yields the bounds (50) in Sec. V E. If  $\varepsilon \approx 0$ , then the optimal number of qubits in Eq. (J2) is approximately proportional to the complexity entropy:  $m \approx H_{\text{H}}^{r, 1-\varepsilon}(\rho) / \log(2)$ .

We further prove that the protocol in Theorem 10 is optimal.

*Theorem 11 (Optimality of data compression).* Let  $\rho$  denote any state of  $n$  qubits. Let  $r \geq 0$  and  $\varepsilon \in [0, 1)$ . Let

$\mathbf{M}^r$  denote a POVM-effect-complexity set as in Theorem 10. Every  $(m, r, \varepsilon)$ -data-compression protocol satisfies

$$m \geq \frac{1}{\log(2)} H_{\text{h}}^{r, 1-\varepsilon}(\rho). \quad (\text{J3})$$

The reduced complexity entropy  $H_{\text{h}}^{r, 1-\varepsilon}$  is defined with respect to  $\mathbf{M}^r$ .

The above data-compression problem maps directly onto the thermodynamic-erasure problem in Sec. III A. The proofs of Theorems 10 and 11 are closely related to the arguments in Sec. III A and to the proof of Theorem 5.

*Proof of Theorem 10.* Consider any  $\zeta > 0$ . There exists a  $Q \in \mathbf{M}^r$  such that  $\text{tr}(Q\rho) \geq 1 - \varepsilon$  and  $\log(\text{tr}(Q)) \leq H_{\text{h}}^{r, 1-\varepsilon}(\rho) + \zeta$ .  $Q = U^\dagger P U$  for some unitary  $U$  satisfying  $C(U) \leq r$  and for some simple POVM effect  $P$  acting as  $\mathbb{1}_2$  on every qubit in a subset  $\mathcal{W} \subset \{1, 2, \dots, n\}$  and as  $|0\rangle\langle 0|$  on every other qubit. Without loss of generality, assume that  $\mathcal{W}$  consists of the first  $m := |\mathcal{W}|$  of the  $n$  qubits. Hence  $P = \mathbb{1}_2^{\otimes m} \otimes |0^{n-m}\rangle\langle 0^{n-m}|$ .

$U$  is an  $(m, r, \varepsilon)$ -data-compression protocol for  $\rho$ , since  $C(U) \leq r$  and condition (J1) holds:

$$\begin{aligned} & \text{tr}(\text{tr}_{\mathcal{W}}(U\rho U^\dagger) |0^{n-m}\rangle\langle 0^{n-m}|) \\ &= \text{tr}([U\rho U^\dagger] [\mathbb{1}_2^{\otimes m} \otimes |0^{n-m}\rangle\langle 0^{n-m}|]) \\ &= \text{tr}([U\rho U^\dagger] P) = \text{tr}(Q\rho) \\ &\geq 1 - \varepsilon. \end{aligned} \quad (\text{J4})$$

Moreover,

$$\begin{aligned} m &= \log_2(\text{tr}(P)) = \log_2(\text{tr}(Q)) \\ &= \frac{1}{\log(2)} \log(\text{tr}(Q)) \leq \frac{1}{\log(2)} \left( H_{\text{h}}^{r, 1-\varepsilon}(\rho) + \zeta \right), \end{aligned} \quad (\text{J5})$$

which implies Eq. (J2), since  $\zeta$  is arbitrary. ■

*Proof.* Let  $U$  denote any  $(m, r, \varepsilon)$ -data-compression protocol for  $\rho$ . By definition, there exists a subset  $\mathcal{W} \subset \{1, 2, \dots, n\}$  such that  $|\mathcal{W}| = m$  and condition (J1) holds. Without loss of generality, assume that  $\mathcal{W}$  consists of the first  $m$  of  $n$  qubits. Let  $P := \mathbb{1}_2^{\otimes m} \otimes |0^{n-m}\rangle\langle 0^{n-m}|$  and  $Q := U^\dagger P U$ . By construction,  $Q$  is in  $\mathbf{M}^r$  and satisfies

$$\begin{aligned} \text{tr}(Q\rho) &= \text{tr}(P U\rho U^\dagger) \\ &= \text{tr}(\text{tr}_{\mathcal{W}}(U\rho U^\dagger) |0^{n-m}\rangle\langle 0^{n-m}|) \geq 1 - \varepsilon. \end{aligned} \quad (\text{J6})$$

$Q$  is therefore a candidate for the  $H_h^{r,1-\varepsilon}(\rho)$  optimization, so

$$\begin{aligned} m &= |\mathcal{W}| = \log_2(\text{tr}(P)) = \log_2(\text{tr}(Q)) \\ &\geq \frac{1}{\log(2)} H_h^{r,1-\varepsilon}(\rho). \end{aligned} \quad (\text{J7})$$

■

## APPENDIX K: DECOUPLING FROM A REFERENCE SYSTEM

Here, we prove the decoupling bound of Theorem 4. Our proof relies on the following conjectured property of the complexity conditional entropy.

*Conjecture 2.* Let  $A$ ,  $B$ , and  $R$  denote distinct quantum systems. Let  $\rho_{ABR}$  denote any state of  $ABR$ . Let  $r \geq 0$  and  $\eta \in (0, 1]$ . It holds that

$$H_H^{r,\eta}(B|R)_\rho \leq H_H^{r,\eta}(AB|R)_\rho + \log(d_A). \quad (\text{K1})$$

Equivalently,

$$\begin{aligned} D_H^{r,\eta}(\rho_{BR} \parallel \pi_B \otimes \rho_R) &\geq D_H^{r,\eta}(\rho_{ABR} \parallel \pi_{AB} \otimes \rho_R) \\ &\quad - 2 \log(d_A). \end{aligned} \quad (\text{K2})$$

The equivalence between Eqs. (K1) and (K2) holds because  $H_H^{r,\eta}(B|R)_\rho = \log(d_B) - D_H^{r,\eta}(\rho_{BR} \parallel \pi_B \otimes \rho_R)$  and because  $H_H^{r,\eta}(AB|R)_\rho = \log(d_A d_B) - D_H^{r,\eta}(\rho_{ABR} \parallel \pi_{AB} \otimes \rho_R)$ , by the argument in Eq. (D78). Equation (K1) mirrors a known inequality for the von Neumann conditional entropy  $H(AB|R)_\rho$ . The latter inequality follows from a chain rule and the lower bound  $H(A|BR)_\rho \geq -\log(d_A)$ :

$$\begin{aligned} H(B|R)_\rho &= H(AB|R)_\rho - H(A|BR)_\rho \leq H(AB|R)_\rho \\ &\quad + \log(d_A). \end{aligned} \quad (\text{K3})$$

*Theorem 12 (Upper bound on the number of qubits Alice can decouple under complexity limitations).* Let  $A$  denote an  $n$ -qubit system, and let  $R$  denote a quantum system distinct from  $A$ . Let  $\rho_{AR}$  denote any state of  $AR$ . Let  $\mathbf{P} = \mathbf{P}_{AR}$  denote any set of simple POVM effects (Definition 7); and  $C = C_{AR}$ , any adjoint-invariant unitary-complexity measure (Definition 3). Let  $\{M^r = M^r(\mathbf{P}, C)\}$  denote the family of POVM-effect-complexity sets defined by Eq. (B15). Let  $r_0 \geq 0$  and  $r_1 \geq r_0$ . Let  $\eta \in (0, 1]$  and  $\delta \in (0, 1]$ . Let  $U_0$  denote any unitary on  $A$  satisfying  $C(U_0) \leq r_0$ . Let  $A_1$  denote a subsystem of  $k \geq 0$  qubits in  $A$ ; and  $A_2$ , the

subsystem of the other  $n - k$  qubits. Let

$$\rho'_{AR} := (U_0 \otimes \mathbb{1}_R) \rho_{AR} (U_0 \otimes \mathbb{1}_R)^\dagger. \quad (\text{K4})$$

Suppose that

$$D_H^{r_1,\eta}(\rho'_{A_2R} \parallel \pi_{A_2} \otimes \rho_R) \leq -\log\left(\frac{\delta}{\eta}\right), \quad (\text{K5})$$

wherein  $D_H^{r_1,\eta}$  is defined with respect to  $M^{r_1}$ . Assume Conjecture 2 holds. It holds that

$$k \geq \frac{1}{2} \left[ n - \frac{1}{\log(2)} H_H^{r_1-r_0,\eta}(A|R)_\rho + \log_2\left(\frac{\delta}{\eta}\right) \right], \quad (\text{K6})$$

wherein  $H_H^{r_1-r_0,\eta}(A|R)$  is defined with respect to  $M^{r_1-r_0}$ .

*Proof.* The following chain of inequalities implies (K6):

$$\begin{aligned} -\log\left(\frac{\delta}{\eta}\right) &\geq D_H^{r_1,\eta}(\rho'_{A_2R} \parallel \pi_{A_2} \otimes \rho_R) \\ &= \log(d_{A_2}) - H_H^{r_1,\eta}(A_2|R)_{\rho'} \\ &\geq \log(d_{A_2}) - H_H^{r_1,\eta}(A|R)_{\rho'} - \log(d_{A_1}) \\ &= (n - 2k) \log(2) - H_H^{r_1,\eta}(A|R)_{\rho'} \\ &\geq (n - 2k) \log(2) - H_H^{r_1-r_0,\eta}(A|R)_\rho. \end{aligned} \quad (\text{K7})$$

The first equality follows by the argument in Eq. (D78); the second inequality, by the assumption that Conjecture 2 holds; and the last inequality, by Proposition 17. ■

- 
- [1] D. Aharonov and L. Eldar, On the complexity of commuting local hamiltonians, and tight conditions for topological order in such systems, [arXiv:1102.0770](#).
  - [2] P. Caputa and S. Liu, Quantum complexity and topological phases of matter, *Phys. Rev. B* **106**, 195125 (2022).
  - [3] A. R. Brown and L. Susskind, Second law of quantum complexity, *Phys. Rev. D* **97**, 086015 (2018).
  - [4] L. Susskind, Three Lectures on Complexity and Black Holes, [arXiv:1810.11563](#).
  - [5] X. Chen, Z.-C. Gu, and X.-G. Wen, Classification of gapped symmetric phases in one-dimensional spin systems, *Phys. Rev. B* **83**, 035107 (2011).
  - [6] M. B. Hastings, Topological order at nonzero temperature, *Phys. Rev. Lett.* **107**, 210501 (2011).
  - [7] M. Schwarz, K. Temme, F. Verstraete, D. Perez-Garcia, and T. S. Cubitt, Preparing topological projected entangled pair states on a quantum computer, *Phys. Rev. A* **88**, 032321 (2013).
  - [8] J. Miller and A. Miyake, Latent computational complexity of symmetry-protected topological order with fractional symmetry, *Phys. Rev. Lett.* **120**, 170503 (2018).



- [9] T. Ali, A. Bhattacharyya, S. Shajidul Haque, E. H. Kim, and N. Moynihan, Post-quench evolution of complexity and entanglement in a topological system, *Phys. Lett. B* **811**, 135919 (2020).
- [10] F. Liu, S. Whitsitt, J. B. Curtis, R. Lundgren, P. Titum, Z.-C. Yang, J. R. Garrison, and A. V. Gorshkov, Circuit complexity across a topological phase transition, *Phys. Rev. Res.* **2**, 013323 (2020).
- [11] Z. Xiong, D.-X. Yao, and Z. Yan, Nonanalyticity of circuit complexity across topological phase transitions, *Phys. Rev. B* **101**, 174305 (2020).
- [12] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, M. J. Bremner, J. M. Martinis, and H. Neven, Characterizing quantum supremacy in near-term devices, *Nat. Phys.* **14**, 595 (2018).
- [13] F. Arute *et al.* Quantum supremacy using a programmable superconducting processor, *Nature* **574**, 505 (2019).
- [14] D. Hangleiter and J. Eisert, Computational advantage of quantum random sampling, *Rev. Mod. Phys.* **95**, 035001 (2023).
- [15] A. Deshpande, B. Fefferman, M. C. Tran, M. Foss-Feig, and A. V. Gorshkov, Dynamical phase transitions in sampling complexity, *Phys. Rev. Lett.* **121**, 030501 (2018).
- [16] L. Susskind, Entanglement is not enough, [arXiv:1411.0690](https://arxiv.org/abs/1411.0690).
- [17] F. G. Brandão, W. Chemsy, N. Hunter-Jones, R. Kueng, and J. Preskill, Models of quantum complexity growth, *PRX Quantum* **2**, 030316 (2021).
- [18] J. Haferkamp, P. Faist, N. B. T. Kothakonda, J. Eisert, and N. Yunger Halpern, Linear growth of quantum circuit complexity, *Nat. Phys.* **18**, 528 (2022).
- [19] Z. Li, Short proofs of linear growth of quantum circuit complexity, [arXiv:2205.05668](https://arxiv.org/abs/2205.05668).
- [20] A. R. Brown, A quantum complexity lower bound from differential geometry, *Nat. Phys.* **19**, 401 (2023).
- [21] L. Susskind, Computational complexity and black hole horizons, *Fortschritte der Physik* **64**, 24 (2016).
- [22] D. Stanford and L. Susskind, Complexity and shock wave geometries, *Phys. Rev. D* **90**, 126007 (2014).
- [23] A. R. Brown, D. A. Roberts, L. Susskind, B. Swingle, and Y. Zhao, Complexity, action, and black holes, *Phys. Rev. D* **93**, 086006 (2016).
- [24] A. Bouland, B. Fefferman, and U. Vazirani, Computational pseudorandomness, the wormhole growth paradox, and constraints on the AdS/CFT duality, [arXiv:1910.14646](https://arxiv.org/abs/1910.14646).
- [25] A. R. Brown, D. A. Roberts, L. Susskind, B. Swingle, and Y. Zhao, Holographic complexity equals bulk action? *Phys. Rev. Lett.* **116**, 191301 (2016).
- [26] S. Chapman, J. Eisert, L. Hackl, M. P. Heller, R. Jefferson, H. Marrochio, and R. C. Myers, Complexity and entanglement for thermofield double states, *SciPost Phys.* **6**, 034 (2019).
- [27] J. Eisert, Entangling power and quantum circuit complexity, *Phys. Rev. Lett.* **127**, 020501 (2021).
- [28] A. Belin, R. C. Myers, S.-M. Ruan, G. Sárosi, and A. J. Speranza, Does complexity equal anything? *Phys. Rev. Lett.* **128**, 081602 (2022).
- [29] V. Balasubramanian, M. DeCross, A. Kar, Y. C. Li, and O. Parrikar, Complexity growth in integrable and chaotic models, *J. High Energy Phys.* **2021**, 11 (2021).
- [30] N. Yunger Halpern, N. B. T. Kothakonda, J. Haferkamp, A. Munson, J. Eisert, and P. Faist, Resource theory of quantum uncomplexity, *Phys. Rev. A* **106**, 062417 (2022).
- [31] C. Bai, W.-H. Li, and X.-H. Ge, Toward the nonequilibrium thermodynamic analog of complexity and the Jarzynski identity, *J. High Energy Phys.* **2022**, 131 (2022).
- [32] C.-F. Chen, J. Haah, J. Haferkamp, Y. Liu, T. Metger, and X. Tan, Incompressibility and spectral gaps of random circuits, [arXiv:2406.07478](https://arxiv.org/abs/2406.07478).
- [33] L. Szilard, Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen, *Z. Phys.* **53**, 840 (1929).
- [34] R. Landauer, Irreversibility and heat generation in the computing process, *IBM J. Res. Dev.* **5**, 183 (1961).
- [35] C. H. Bennett, The thermodynamics of computation—a review, *Int. J. Theor. Phys.* **21**, 905 (1982).
- [36] C. H. Bennett, Notes on Landauer’s principle, reversible computation and Maxwell’s demon, *Stud. Hist. Philos. Mod. Phys.* **34**, 501 (2003).
- [37] J. Goold, M. Huber, A. Riera, L. d. Rio, and P. Skrzypczyk, The role of quantum information in thermodynamics—a topical review, *J. Phys. A* **49**, 143001 (2016).
- [38] F. Binder, L. A. Correa, C. Gogolin, J. Anders, and G. Adesso, Eds., *Thermodynamics in the Quantum Regime: Fundamental Aspects and New Directions*, Fundamental Theories of Physics (Springer International Publishing, Cham, 2018), Vol. 195.
- [39] T. Sagawa, *Entropy, Divergence, and Majorization in Classical and Quantum Thermodynamics*, Springer Briefs in Mathematical Physics (Springer, Singapore, 2022), Vol. 16.
- [40] P. Faist and R. Renner, Fundamental work cost of quantum processes, *Phys. Rev. X* **8**, 021011 (2018).
- [41] M. Horodecki, P. Horodecki, and J. Oppenheim, Reversible transformations from pure to mixed states and the unique measure of information, *Phys. Rev. A* **67**, 062104 (2003).
- [42] F. G. S. L. Brandão, M. Horodecki, J. Oppenheim, J. M. Renes, and R. W. Spekkens, Resource theory of quantum states out of thermal equilibrium, *Phys. Rev. Lett.* **111**, 250404 (2013).
- [43] M. Horodecki and J. Oppenheim, Fundamental limitations for quantum and nanoscale thermodynamics, *Nat. Commun.* **4**, 2059 (2013).
- [44] F. Brandão, M. Horodecki, N. Ng, J. Oppenheim, and S. Wehner, The second laws of quantum thermodynamics, *Proc. Natl. Acad. Sci. USA* **112**, 3275 (2015).
- [45] D. Janzing, P. Wocjan, R. Zeier, R. Geiss, and T. Beth, Thermodynamic cost of reliability and low temperatures: Tightening Landauer’s principle and the second law, *Int. J. Theor. Phys.* **39**, 2717 (2000).
- [46] P. Faist, F. Dupuis, J. Oppenheim, and R. Renner, The minimal work cost of information processing, *Nat. Commun.* **6**, 7669 (2015).
- [47] P. Taranto, F. Bakhshinezhad, A. Bluhm, R. Silva, N. Friis, M. P. Lock, G. Vitagliano, F. C. Binder, T. Debarba, E. Schwarzhans, F. Clivaz, and M. Huber, Landauer vs. Nernst: What is the true cost of cooling a quantum system? *PRX Quantum* **4**, 010332 (2023).

- [48] C. H. Bennett, Logical reversibility of computation, *IBM J. Res. Dev.* **17**, 525 (1973).
- [49] D. Egloff, O. C. O. Dahlsten, R. Renner, and V. Vedral, A measure of majorization emerging from single-shot statistical mechanics, *New J. Phys.* **17**, 073001 (2015).
- [50] C. H. Bennett, P. Gács, M. Li, P. M. B. Vitányi, and W. H. Zurek, in *Proceedings of the Twenty-Fifth Annual ACM Symposium on Theory of Computing*, STOC '93 (Association for Computing Machinery, New York, NY, USA, 1993), pp. 21–30.
- [51] W. H. Zurek, Thermodynamic cost of computation, algorithmic complexity and the information metric, *Nature* **341**, 119 (1989).
- [52] C. M. Caves, Information and entropy, *Phys. Rev. E* **47**, 4010 (1993).
- [53] C. Mora, H. Briegel, and B. Kraus, Quantum Kolmogorov complexity and its applications, [arXiv:quant-ph/0610109](https://arxiv.org/abs/quant-ph/0610109).
- [54] J. Baez and M. Stay, Algorithmic thermodynamics, *Math. Struct. Comput. Sci.* **22**, 771 (2012).
- [55] F. G. S. L. Brandão, A. W. Harrow, and M. Horodecki, Local random quantum circuits are approximate polynomial-designs, *Commun. Math. Phys.* **346**, 397 (2016).
- [56] J. Haferkamp, Random quantum circuits are approximate unitary  $t$ -designs in depth  $o(nt^{5+o(1)})$ , *Quantum* **6**, 795 (2022).
- [57] D. A. Roberts and B. Yoshida, Chaos and complexity by design, *J. High Energy Phys.* **2017**, 121 (2017).
- [58] R. P. Feynman, in *Feynman Lectures on Computation*, edited by A. J. G. Hey and R. W. Allen (Addison-Wesley, Reading, Massachusetts, 1996).
- [59] O. C. O. Dahlsten, R. Renner, E. Rieper, and V. Vedral, Inadequacy of von Neumann entropy for characterizing extractable work, *New J. Phys.* **13**, 53015 (2011).
- [60] L. del Rio, J. Åberg, R. Renner, O. Dahlsten, and V. Vedral, The thermodynamic meaning of negative entropy, *Nature* **474**, 61 (2011).
- [61] P. Hayden, M. Horodecki, A. Winter, and J. Yard, A decoupling approach to the quantum capacity, *Open Syst. Inf. Dyn.* **15**, 7 (2008).
- [62] A. Abeyesinghe, I. Devetak, P. Hayden, and A. Winter, The mother of all protocols: restructuring quantum information's family tree, *Proc. R. Soc. Lond. A* **465**, 2537 (2009).
- [63] F. Dupuis, M. Berta, J. Wullschleger, and R. Renner, One-shot decoupling, *Commun. Math. Phys.* **328**, 251 (2014).
- [64] O. Szehr, F. Dupuis, M. Tomamichel, and R. Renner, Decoupling with unitary approximate two-designs, *New J. Phys.* **15**, 053022 (2013).
- [65] P. Faist, J. Oppenheim, and R. Renner, Gibbs-preserving maps outperform thermal operations in the quantum regime, *New J. Phys.* **17**, 043003 (2015).
- [66] I. Vienna, Problem 46 (2023), <https://oqp.iqoqi.oeaw.ac.at/thermodynamic-implementation-of-gibbs-preserving-maps>.
- [67] R. Alicki, M. Horodecki, P. Horodecki, and R. Horodecki, Thermodynamics of quantum information systems — Hamiltonian description, *Open Syst. Inf. Dyn.* **11**, 205 (2004).
- [68] M. F. Frenzel, D. Jennings, and T. Rudolph, Reexamination of pure qubit work extraction, *Phys. Rev. E* **90**, 052136 (2014).
- [69] P. Skrzypczyk, A. J. Short, and S. Popescu, Work extraction and thermodynamics for individual quantum systems, *Nat. Commun.* **5**, 4185 (2014).
- [70] O. Dahlsten, Non-equilibrium statistical mechanics inspired by modern information theory, *Entropy* **15**, 5346 (2013).
- [71] M. Lostaglio, D. Jennings, and T. Rudolph, Description of quantum coherence in thermodynamic processes requires constraints beyond free energy, *Nat. Commun.* **6**, 6383 (2015).
- [72] N. Yunger Halpern, A. J. P. Garner, O. C. O. Dahlsten, and V. Vedral, Maximum one-shot dissipated work from Rényi divergences, *Phys. Rev. E* **97**, 052135 (2018).
- [73] B.-B. Wei and M. B. Plenio, Relations between dissipated work in non-equilibrium process and the family of Rényi divergences, *New J. Phys.* **19**, 023002 (2017).
- [74] N. Yunger Halpern and J. M. Renes, Beyond heat baths: Generalized resource theories for small-scale thermodynamics, *Phys. Rev. E* **93**, 022126 (2016).
- [75] R. van der Meer, N. H. Y. Ng, and S. Wehner, Smoothed generalized free energies for thermodynamics, *Phys. Rev. A* **96**, 062135 (2017).
- [76] N. H. Y. Ng, M. P. Woods, and S. Wehner, Surpassing the Carnot efficiency by extracting imperfect work, *New J. Phys.* **19**, 113005 (2017).
- [77] G. Guarnieri, N. H. Y. Ng, K. Modi, J. Eisert, M. Paternostro, and J. Goold, Quantum work statistics and resource theories: Bridging the gap through Rényi divergences, *Phys. Rev. E* **99**, 050101 (2019).
- [78] R. Renner, Security of Quantum Key Distribution, Ph.D. thesis, ETH Zürich (2005).
- [79] N. Datta, Min- and max-relative entropies and a new entanglement monotone, *IEEE Trans. Inf. Theory* **55**, 2816 (2009).
- [80] N. Yunger Halpern, A. J. P. Garner, O. C. O. Dahlsten, and V. Vedral, Introducing one-shot work into fluctuation relations, *New J. Phys.* **17**, 095003 (2015).
- [81] F. Hiai and D. Petz, The proper formula for relative entropy and its asymptotics in quantum probability, *Commun. Math. Phys.* **143**, 99 (1991).
- [82] F. Dupuis, L. Kraemer, P. Faist, J. M. Renes, and R. Renner, in *XVIIth International Congress on Mathematical Physics* (World Scientific, Singapore, 2013), pp. 134–153.
- [83] N. Datta, M. Mosonyi, M.-H. Hsieh, and F. G. S. L. Brandão, A smooth entropy approach to quantum hypothesis testing and the classical capacity of quantum channels, *IEEE Trans. Inf. Theory* **59**, 8014 (2013).
- [84] M. Hayashi, Optimal sequence of quantum measurements in the sense of Stein's Lemma in quantum hypothesis testing, *J. Phys. A* **35**, 10759 (2002).
- [85] J. Watrous, Lecture notes on Advanced Topics in Quantum Information Theory, University of Waterloo (2020), <https://cs.uwaterloo.ca/~watrous/QIT-notes/>.
- [86] L. Wang and R. Renner, One-shot classical-quantum capacity and hypothesis testing, *Phys. Rev. Lett.* **108**, 200501 (2012).

- [87] M. Tomamichel and M. Hayashi, A hierarchy of information quantities for finite block length analysis of quantum tasks, *IEEE Trans. Inf. Theory* **59**, 7693 (2013).
- [88] S. Khatri and M. M. Wilde, Principles of quantum communication theory: A modern approach, [arXiv:2011.04672](https://arxiv.org/abs/2011.04672).
- [89] F. Buscemi and G. Gour, Quantum relative Lorenz curves, *Phys. Rev. A* **95**, 012110 (2017).
- [90] N. Yunger Halpern, Beyond heat baths II: Framework for generalized thermodynamic resource theories, *J. Phys. A* **51**, 094001 (2018).
- [91] J. Åberg, Truly work-like work extraction via a single-shot analysis, *Nat. Commun.* **4**, 1925 (2013).
- [92] D. Gross, S. T. Flammia, and J. Eisert, Most quantum states are too entangled to be useful as computational resources, *Phys. Rev. Lett.* **102**, 190501 (2009).
- [93] M. J. Bremner, C. Mora, and A. Winter, Are random pure states useful for quantum computation? *Phys. Rev. Lett.* **102**, 190502 (2009).
- [94] M. Tomamichel, *Quantum Information Processing with Finite Resources* (Springer International Publishing, Cham, 2016), Vol. 5.
- [95] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, UK, 2010).
- [96] Á. M. Alhambra, J. Oppenheim, and C. Perry, Fluctuating states: What is the probability of a thermodynamical transition? *Phys. Rev. X* **6**, 041016 (2016).
- [97] T. M. Graham, L. Phuttitarn, R. Chinnarasu, Y. Song, C. Poole, K. Jooya, J. Scott, A. Scott, P. Eichler, and M. Saffman, Mid-circuit measurements on a single species neutral alkali atom quantum processor, *Phys. Rev. X* **13**, 041051 (2023).
- [98] J. M. Koh, S.-N. Sun, M. Motta, and A. J. Minnich, Measurement-induced entanglement phase transition on a superconducting quantum processor with mid-circuit readout, *Nat. Phys.* **19**, 1314 (2023).
- [99] D. Zhu, G. D. Kahanamoku-Meyer, L. Lewis, C. Noel, O. Katz, B. Harraz, Q. Wang, A. Risinger, L. Feng, D. Biswas, L. Egan, A. Gheorghiu, Y. Nam, T. Vidick, U. Vazirani, N. Y. Yao, M. Cetina, and C. Monroe, Interactive cryptographic proofs of quantumness using mid-circuit measurements, *Nat. Phys.* **19**, 1725 (2023).
- [100] B. Skinner, J. Ruhman, and A. Nahum, Measurement-induced phase transitions in the dynamics of entanglement, *Phys. Rev. X* **9**, 031009 (2019).
- [101] A. Chan, R. M. Nandkishore, M. Pretko, and G. Smith, Unitary-projective entanglement dynamics, *Phys. Rev. B* **99**, 224307 (2019).
- [102] G.-Y. Zhu, N. Tantivasadakarn, A. Vishwanath, S. Trebst, and R. Verresen, Nishimori's cat: Stable long-range entanglement from finite-depth unitaries and weak measurements, *Phys. Rev. Lett.* **131**, 200201 (2023).
- [103] B. Yoshida, Decoding the entanglement structure of monitored quantum circuits, [arXiv:2109.08691](https://arxiv.org/abs/2109.08691).
- [104] R. Suzuki, J. Haferkamp, J. Eisert, and P. Faist, Quantum complexity phase transitions in monitored random circuits, *Quantum* **9**, 1627 (2025).
- [105] P. Faist, The Entropy Zoo (2015), <https://phfaist.com/entropyzoo>.
- [106] J. Haferkamp, On the moments of random quantum circuits and robust quantum complexity, [arXiv:2303.16944](https://arxiv.org/abs/2303.16944).
- [107] D. Slepian and J. Wolf, Noiseless coding of correlated information sources, *IEEE Trans. Inf. Theory* **19**, 471 (1973).
- [108] P. J. Coles, M. Berta, M. Tomamichel, and S. Wehner, Entropic uncertainty relations and their applications, *Rev. Mod. Phys.* **89**, 015002 (2017).
- [109] M. M. Wilde, *Quantum Information Theory* (Cambridge University Press, Cambridge, UK, 2013).
- [110] M. Tomamichel, C. Schaffner, A. Smith, and R. Renner, Leftover hashing against quantum side information, *IEEE Trans. Inf. Theory* **57**, 5524 (2011).
- [111] W. Brown and O. Fawzi, Decoupling with random quantum circuits, *Commun. Math. Phys.* **340**, 867 (2015).
- [112] A. Vitanov, F. Dupuis, M. Tomamichel, and R. Renner, Chain rules for smooth min- and max-entropies, *IEEE Trans. Inf. Theory* **59**, 2603 (2013).
- [113] F. Dupuis, Chain rules for quantum Rényi entropies, *J. Math. Phys.* **56**, 022203 (2015).
- [114] J. T. Yard and I. Devetak, Optimal quantum source coding with quantum side information at the encoder and decoder, *IEEE Trans. Inf. Theory* **55**, 5339 (2009).
- [115] Z. Ji, Y.-K. Liu, and F. Song, in *Advances in Cryptology – CRYPTO 2018*, Lecture Notes in Computer Science, edited by H. Shacham and A. Boldyreva (Springer International Publishing, Cham, 2018), pp. 126–152.
- [116] Z. Brakerski and O. Shmueli, (pseudo) random quantum states with binary phase, [arXiv:1906.10611](https://arxiv.org/abs/1906.10611).
- [117] D. Gosset, V. Kliuchnikov, M. Mosca, and V. Russo, An algorithm for the  $T$ -count, *Quantum Inf. Comput.* **14**, 1261 (2014).
- [118] J. van de Wetering and M. Amy, Optimising  $T$ -count is NP-hard, [arXiv:2310.05958](https://arxiv.org/abs/2310.05958).
- [119] Y.-H. Chen, K.-M. Chung, C.-Y. Lai, S. P. Vadhan, and X. Wu, Computational notions of quantum min-entropy, [arXiv:1704.07309](https://arxiv.org/abs/1704.07309).
- [120] M. Gell-Mann and J. B. Hartle, Quasiclassical coarse graining and thermodynamic entropy, *Phys. Rev. A* **76**, 022104 (2007).
- [121] D. Šafránek, J. M. Deutsch, and A. Aguirre, Quantum coarse-grained entropy and thermodynamics, *Phys. Rev. A* **99**, 010101 (2019).
- [122] C. H. Bennett, in *The Universal Turing Machine: A Half-Century Survey* (Oxford University Press, New York, New York, 1988), pp. 227–257.
- [123] A. Berthiaume, W. van Dam, and S. Laplante, Quantum Kolmogorov complexity, *J. Comput. Syst. Sci.* **63**, 201 (2001).
- [124] M. A. Nielsen, A geometric approach to quantum circuit lower bounds, [arXiv:quant-ph/0502070](https://arxiv.org/abs/quant-ph/0502070).
- [125] M. A. Nielsen, M. R. Dowling, M. Gu, and A. C. Doherty, Quantum computation as geometry, *Science* **311**, 1133 (2006).
- [126] M. A. Nielsen, M. R. Dowling, M. Gu, and A. C. Doherty, Optimal control, geometry, and quantum computing, *Phys. Rev. A* **73**, 062323 (2006).

- [127] M. R. Dowling and M. A. Nielsen, The geometry of quantum computation, [Quantum Inf. Comput.](#) **8**, 861 (2008).
- [128] I. Kim, E. Tang, and J. Preskill, The ghost in the radiation: Robust encodings of the black hole interior, [J. High Energy Phys.](#) **2020**, 31 (2020).
- [129] L. Susskind, Black holes at exp-time, [arXiv:2006.01280](#).
- [130] I. Nechita, Z. Puchała, Ł. Paweła, and K. Życzkowski, Almost all quantum channels are equidistant, [J. Math. Phys.](#) **59**, 052201 (2018).
- [131] A. M. Dalzell, N. Hunter-Jones, and F. G. S. L. Brandão, Random quantum circuits anticoncentrate in log depth, [PRX Quantum](#) **3**, 010333 (2022).
- [132] D. Gottesman, in *Group22: Proceedings of the XXII International Colloquium on Group Theoretical Methods in Physics* (International Press of Boston, Inc., Somerville, Massachusetts, 1998), Vol. 32.
- [133] E. Knill and R. Laflamme, Power of one bit of quantum information, [Phys. Rev. Lett.](#) **81**, 5672 (1998).
- [134] C. A. Fuchs and J. van de Graaf, Cryptographic distinguishability measures for quantum-mechanical states, [IEEE Trans. Inf. Theory](#) **45**, 1216 (1999).
- [135] J. Haferkamp and N. Hunter-Jones, Improved spectral gaps for random quantum circuits: Large local dimensions and all-to-all interactions, [Phys. Rev. A](#) **104**, 022417 (2021).
- [136] C. M. Dawson and M. A. Nielsen, The Solovay-Kitaev algorithm, [arXiv:quant-ph/0505030](#).
- [137] A. Y. Kitaev, A. Shen, and M. N. Vyalyi, *Classical and Quantum Computation* (American Mathematical Society, Providence, Rhode Island, 2002).
- [138] M. Oszmaniec, A. Sawicki, and M. Horodecki, Epsilon-nets, unitary designs and random quantum circuits, [arXiv:2007.10885](#).
- [139] R. Alicki and M. Fannes, Continuity of quantum conditional information, [J. Phys. A](#) **37**, L55 (2004).
- [140] A. Winter, Tight uniform continuity bounds for quantum entropies: Conditional entropy, relative entropy distance and energy constraints, [Commun. Math. Phys.](#) **347**, 291 (2016).
- [141] M. Mariën, K. M. R. Audenaert, K. Van Acoleyen, and F. Verstraete, Entanglement rates and the stability of the area law for the entanglement entropy, [Commun. Math. Phys.](#) **346**, 35 (2016).
- [142] N. Schuch, M. M. Wolf, K. G. H. Vollbrecht, and J. I. Cirac, On entropy growth and the hardness of simulating time evolution, [New J. Phys.](#) **10**, 033032 (2008).
- [143] M. Cramer, C. M. Dawson, J. Eisert, and T. J. Osborne, Exact relaxation in a class of non-equilibrium quantum lattice systems, [Phys. Rev. Lett.](#) **100**, 030602 (2008).