# Sequences of resource monotones from modular Hamiltonian polynomials

Arias, R.; de Boer, J.; Di Giulio, G.; Keski-Vakkuri, E.; Tonni, E.

**DOI**
10.1103/PHYSREVRESEARCH.5.043082
**Publication date**
2023
**Document Version**
Final published version
**Published in**
Physical Review Research
**License**
CC BY

# Sequences of resource monotones from modular Hamiltonian polynomials

Raúl Arias [1,2] Jan de Boer [3] Giuseppe Di Giulio [4] Esko Keski-Vakkuri [5,6,7] and Erik Tonni [8]

[1] *Instituto de Física de La Plata, CONICET, Diagonal 113 e/63 y 64, CC67, 1900 La Plata, Argentina*
[2] *Departamento de Física, Universidad Nacional de La Plata, Calle 49 y 115 s/n, CC67, 1900 La Plata, Argentina*
[3] *Institute for Theoretical Physics and Delta Institute for Theoretical Physics, University of Amsterdam, PO Box 94485, 1090 GL Amsterdam, Netherlands*
[4] *Institute for Theoretical Physics and Astrophysics and Würzburg-Dresden Cluster of Excellence ct.qmat, Julius-Maximilians-Universität Würzburg, Am Hubland, 97074 Würzburg, Germany*
[5] *Department of Physics, University of Helsinki, PO Box 64, FIN-00014 Helsinki, Finland*
[6] *Helsinki Institute of Physics, PO Box 64, FIN-00014 Helsinki, Finland*
[7] *InstituteQ - the Finnish Quantum Institute, Helsinki, Finland*
[8] *SISSA and INFN Sezione di Trieste, via Bonomea 265, 34136 Trieste, Italy*

(Received 24 March 2023; accepted 28 September 2023; published 24 October 2023)

We introduce two infinite sequences of entanglement monotones, which are constructed from expectation values of polynomials in the modular Hamiltonian. These monotones yield infinite sequences of inequalities that must be satisfied in majorizing state transitions. We demonstrate this for information erasure, deriving an infinite sequence of "Landauer inequalities" for the work cost, bounded by linear combinations of expectation values of powers of the modular Hamiltonian. These inequalities give improved lower bounds for the work cost in finite-dimensional systems, and depend on more details of the erased state than just on its entropy and variance of modular Hamiltonian. Similarly one can derive lower bounds for marginal entropy production for a system coupled to an environment. These infinite sequences of entanglement monotones also give rise to relative quantifiers that are monotonic in more general processes, namely those involving so-called $\sigma$ majorization with respect to a fixed point full rank state $\sigma$; such quantifiers are called resource monotones. As an application to thermodynamics, one can use them to derive finite-dimension corrections to the Clausius inequality. Finally, in order to gain some intuition for what (if anything) plays the role of majorization in field theory, we compare pairs of states in discretized theories at criticality and study how majorization depends on the size of the bipartition with respect to the size of the entire chain.

DOI: 10.1103/PhysRevResearch.5.043082

## I. INTRODUCTION

Quantum resource theories (see e.g., [1] for a review) have been developed as a general framework to sharpen the distinction between the achievable and the unachievable in various classes of quantum processes. One makes the distinction between "free states", which are generated by the class of allowed quantum operations ("free operations"), and "resource states", which cannot be generated by free operations and must therefore be prepared by an external agent. As an example, entanglement cannot be created by LOCC operations (the free operations in the resource theory of entanglement), and it thus acts as a resource. There are two other important ingredients in resource theories: monotones and partial orders. In the context of entanglement, entanglement entropy is a well-known monotone. The inability to create entanglement in a bipartite system by LOCC operations is quantified by non-increasing (decreasing) entanglement entropy in the reduced state of a subsystem, when LOCC operations are applied to the global state. Similarly, partial orders among quantum states, in particular majorization $\rho \succ \sigma$ defined for a pair of states restrict the ability to convert a state to another state by a single-shot operation. For example, by Nielsen's theorem [2], conversion by a LOCC operation is only possible if the output state majorizes the input state. Conversely, the majorization relation leads to inequalities satisfied by the values of monotones computed for the pair of states.

Generalizing the concept of an entanglement monotone, one can introduce quantifiers to track the loss of a resource under free operations, resource monotones $R$ that have the property [1]

$$R(\rho) \geqslant R(\Phi(\rho)), \tag{1.1}$$

under any free operation $\Phi$ of the resource theory. Likewise, generalizing the concept of majorization, one can define relative notions of majorization involving pairs of states $(\rho_1, \rho_2) \succ (\sigma_1, \sigma_2)$. Such relative majorization arises naturally as a constraint when $\rho_2 = \sigma_2 = \sigma_*$ is a fixed state $\Phi(\sigma_*) = \sigma_*$ of free operations of a resource theory. For example, in the resource theory of quantum thermodynamics, where the fixed state is the thermal equilibrium state, one obtains a thermomajorization relation constraining thermal operations involving nonequilibrium states.

In this paper, we introduce infinite sequences of entanglement and resource monotones. Our constructions are motivated from three different directions. First, Rényi entropies (and their relative generalizations) are a useful tool in many areas of quantum science, and many techniques have been developed to compute them, in particular recently in many-body physics and quantum field theory. However, Rényi entropies in general are not monotones. Here we show how one can use them as generating functions to compute sequences of monotones.

Second, von Neumann entropy and relative entropy are monotones with a central role. Recently [3], inequalities associated with (relative) entropy production in finite-dimensional systems were sharpened with lower bounds involving the variance of surprisal $C(\rho) = \mathrm{Tr}[\rho(-\ln\rho)^2] - [-\mathrm{Tr}(\rho\ln\rho)]^2$. These inequalities were shown to result from quantum operations that imply a majorization relation between the input and output states. We show that our monotones extend these inequalities into infinite sequences involving quantities that generalize $C(\rho)$. These quantities come from our third motivation: the spectrum of the surprisal $K \equiv -\ln\rho$, or as we will be calling it in this paper following [4], the modular Hamiltonian [5].

In the context of entanglement, for the reduced density matrix $\rho_A$ the operator $K_A = -\ln\rho_A$ is known as the *entanglement Hamiltonian*, and its spectrum of eigenvalues $\{\varepsilon_i\}$, the entanglement spectrum, is a useful tool to study many-body quantum systems, e.g., to detect quantum phase transitions [6,7]. It is interesting to study how quantum operations involving majorization relation alter the (entanglement) spectrum; in this paper we characterize the spectrum by its cumulants. Our monotones can be expanded as combinations of arbitrarily higher-order cumulants, and the resulting sequences of inequalities characterize and constrain the spectra by statistically natural quantities generalizing the entropy. After presenting the sequences of monotones, we construct their relative counterparts, which generalize the relative entropy since they involve two states rather than one. Defining these quantities, referred to in the paper as *relative quantifiers*, we obtain sequences of resource monotones.

Most of the results of this paper apply to majorization for finite-dimensional systems. In quantum field theory, less is known about majorization. There have been studies investigating ground state entanglement and the behavior of majorization in the reduced density matrix in a subsystem under renormalization group flow and scaling transformations [6,8–14]. In this paper, our interest is in the possibility of majorization between a pair of states in a quantum field theory. To move the first step towards this goal, we compute some entanglement monotones for states in discretized conformal field theories (CFTs) and study whether the majorization between states is ruled out and how this depends on the size of the bipartition with respect to the one of the entire chain.

Before moving to summarize our main results in Sec. I B, we present some technical background needed to discuss the results.

### A. Some technical background

It is widely known that the Rényi entropies

$$S^{(n)} = \tfrac{1}{1-n}\ln\mathrm{Tr}\rho^n, \tag{1.2}$$

in the limit $n \to 1$ reduce to the Von Neumann entropy

$$S \equiv \lim_{n\to 1} S^{(n)} = -\,\partial_n(\ln\mathrm{Tr}\rho^n)|_{n=1} = -\,\mathrm{Tr}(\rho\ln\rho). \tag{1.3}$$

The definition of $S(\rho)$ allows to interpret it as the expectation value of the modular Hamiltonian, Hermitian and well-defined since $\rho$ is Hermitian and positive definite. The standard normalization is [15] $\mathrm{Tr}(\rho) = \mathrm{Tr}(e^{-K}) = 1$. Given a system bipartite into $A$ and its complement, the entanglement entropy $S_A$ is the average (or the first moment) of the entanglement Hamiltonian, namely $S_A = \langle K_A\rangle$, where the mean value is evaluated through the reduced density matrix $\rho_A$.

In addition to the first moment of modular Hamiltonian, it is natural to explore its higher moments or cumulants as well. The second cumulant, the variance of modular Hamiltonian, is a much less-known quantity, and therefore even its name varies in the literature. It is also known as entropy variance, varentropy, and in the context of entanglement in many-body physics and quantum field theory, as *capacity of entanglement* $C_A(\rho_A)$. In the latter context, it was introduced in [16] and [17], first with a definition modeled after that of a heat capacity, and proposed to detect different phases in topological matter, where entanglement is known to play an important role [7,18,19]. Since heat capacity is related to the variance of thermodynamical entropy, it was realized that capacity of entanglement is equal to the variance of the modular Hamiltonian, and can also be derived from the Rényi entropies [20–22]

$$\begin{aligned} C_A &= \partial_n^2\big(\ln\mathrm{Tr}\rho_A^n\big)\big|_{n=1} = \partial_n^2\big(\mathrm{Tr}\rho_A^n\big)\big|_{n=1} - \big[\partial_n\big(\mathrm{Tr}\rho_A^n\big)\big]^2\big|_{n=1} \\ &= \big\langle K_A^2\big\rangle - \langle K_A\rangle^2. \end{aligned} \tag{1.4}$$

One of the reasons the variance of the modular Hamiltonian or capacity of entanglement is less known is that it is not known to satisfy many interesting properties, unlike the von Neumann entropy does. For instance, the entanglement entropy is known to be an entanglement monotone [23], while the capacity of entanglement is not. Some of its uses in quantum information theory, which we are aware of, are in a finite-dimension correction to the Landauer inequality or more generally in bounding the increase in entropy in state transitions between majorizing states [3,24], in the analysis of catalytic state transformations [3], and in-state interconvertibility in finite systems [25].

This paper is primarily motivated by [3], which considered the capacity of entanglement (there called variance of surprisal) and more generally the relative variance

$$C(\rho||\sigma) = \mathrm{Tr}[\rho(\ln\rho - \ln\sigma)^2] - S(\rho||\sigma)^2, \tag{1.5}$$

where $S(\rho||\sigma) \equiv \mathrm{Tr}[\rho(\ln\rho - \ln\sigma)]$ is the relative entropy. When we consider $\sigma$ in (1.5) to be the maximally mixed state $C(\rho||\sigma)$ reduces to $C(\rho)$. While both the relative variance and the relative entropy have applications in the independent and identically distributed setting involving many copies of a system or operations, the authors of [3] explored their role in a single-shot setting where an operation or protocol is executed only once in one system, proving many new results for the variance and relative variance. Many properties were based on a new quantifier [26]

$$M(\rho) = C(\rho) + [S(\rho) + 1]^2, \tag{1.6}$$

which was shown to be Schur concave, and its relative version, which was shown to be a resource monotone. The quantities $S, C, M$ were shown to be connected by an inequality [27]

$$S(\rho) - S(\sigma) \geqslant \frac{C(\sigma) - C(\rho)}{2\sqrt{M(\sigma)}}, \tag{1.7}$$

when the majorization order $\sigma \succ \rho$ holds for two states in a system with a finite-dimensional Hilbert space. As an application, [3] considered, e.g., information erasure, deriving a new lower bound for the associated work cost that involves both entropy and variance. Related results were proven for the relative quantifiers. In the end they posed a question of whether it is possible to extend this construction of $M$ to a sequence of Schur concave quantifiers that would have similar properties and involve higher cumulants than $C$, perhaps also likewise for the relative quantifiers.

### B. Summary of the main results

In state conversions involving a majorization order between the initial state $\rho$ and the final state $\sigma$, the whole spectrum of eigenvalues is affected, and the majorization order itself can be defined by a sequence of inequalities. Properties of the spectra can be characterized by various quantities, such as Rényi entropies, or moments and cumulants of $-\ln\rho$ and $-\ln\sigma$. Hence it is natural to expect that majorization order may imply a sequence of inequalities involving changes in cumulants beyond the first two. To derive such inequalities, we construct two sequences of entanglement monotones, that can be expanded as combinations of cumulants. We first generalize the construction (1.6) and define the *moments of the shifted modular Hamiltonian*

$$M^{(n)}(\rho; b_n) = \mathrm{Tr}[\rho(-\ln\rho + b_n)^n] - b_n^n, \tag{1.8}$$

for $n \geqslant 1$ [with $M^{(2)}(\rho; b_n = 1) = M(\rho) - 1$]. Explicit formulas for their expansions by higher cumulants are given in Sec. III. For the parameter range $b_n \geqslant n-1$ all of them are concave (see Sec. II for the relevant definitions); hence, from the Vidal's theorem [23], they are pure state entanglement monotones (where $\rho$ and $\sigma$ are the reduced density matrices of global pure states $|\psi\rangle$ and $|\phi\rangle$ respectively), thus yielding inequalities

$$M^{(n)}(\rho; b_n) \geqslant M^{(n)}(\sigma; b_n), \tag{1.9}$$

in local operations assisted with classical communication (LOCC) and other majorizing state transformations [28] with $\sigma \succ \rho$. For example, at second-order $n = 2$ with $b_n = 1$ we obtain the inequality

$$S(\rho) - S(\sigma) \geqslant \frac{C(\sigma) - C(\rho)}{S(\rho) + S(\sigma) + 2}, \tag{1.10}$$

which is slightly sharper than the inequality (1.7). The moments $M^{(n)}$ can be computed from Rényi entropies $S^{(\alpha)}(\rho)$, by using the latter as a generating function

$$M^{(n)}(\rho; b_n) = \left[ e^b (-1)^n \frac{d^n}{d\alpha^n} \exp[-\alpha b + (1-\alpha) S^{(\alpha)}(\rho)] \right]\Bigg|_{\alpha=1, b=b_n} - b_n^n. \tag{1.11}$$

The Rényi entropies $S^{(\alpha)}$ are not concave (for index value $\alpha > 1$); hence, our observation gives a way to repackage their information to an infinite sequence of concave quantifiers, that define entanglement monotones.

Next we identify a basis for monotones, which are polynomial in moments of $-\log\rho$, which allows to construct another infinite sequence that we call *extremal polynomial monotones* $P_E^{(n)}$ (see Sec. III B). All $-P_E^{(n)}$ are also concave, hence define monotones, and any concave polynomial can be written as a linear combination of extremal polynomial monotones with non-negative coefficients. We therefore believe that they provide the tightest inequalities of this type in majorizing state transformations. For example, given two majorizing states $\rho \succ \sigma$, the third monotone $P_E^{(3)}$ yields the inequality

$$\Delta M_3 \geqslant 3\Delta M_2 + \frac{3}{4}\frac{(\Delta M_2)^2}{\Delta M_1}, \tag{1.12}$$

and the fourth-order $P_E^{(4)}$ the inequality

$$\Delta M_4 \geqslant 8\Delta M_3 - 6\Delta M_2 + \frac{8}{9}\frac{(\Delta M_3 - 3\Delta M_2)^2}{\Delta M_2}, \tag{1.13}$$

where $\Delta M_n \equiv M^{(n)}(\sigma; n-1) - M^{(n)}(\rho; n-1)$. Notice that (1.12) and (1.13) are stronger than the inequalities $\Delta M_3 \geqslant 0$ and $\Delta M_4 \geqslant 0$ obtained from (1.9).

As an application of these inequalities we first consider information erasure: We obtain infinite sequences of "Landauer inequalities" for the work cost, bounded by arbitrarily high cumulants of the modular Hamiltonian of the initial state $-\ln\rho$ to be erased, extending the previous result of [3], that involves only the variance. We also derive a slightly sharper inequality for marginal entropy production by applying a unital quantum channel to a system and environment, and outline steps for deriving an infinite sequence of inequalities.

For relative quantifiers, we first generalize a theorem proven in [3], to show how one can construct an infinite class of resource monotones, relative quantifiers based on a concave quantifier $E(\rho) = \mathrm{Tr}[\rho F(\ln\rho)]$. We then apply this construction to the monotones $M^{(n)}$ and $-P_E^{(n)}$, obtaining infinite sequences of resource monotones that involve cumulants of $\ln\rho - \ln\sigma$. An important restriction is that the results only apply in the case where $\rho$ and $\sigma$ commute. The sequences imply inequalities for relative entropy production bounded by changes in the relative cumulants; we consider two examples more explicitly. In particular, as an application to quantum thermodynamics, we derive a finite-dimension correction to the Clausius inequality,

$$S(\gamma_\beta) - S(\rho) \geqslant \frac{1}{k_B T}\left[\langle H\rangle_{\gamma_\beta} - \langle H\rangle_\rho\right] + \frac{C(\rho||\gamma_\beta)}{2 + 2\beta(E_{\max} - F(\beta))}, \tag{1.14}$$

where $\rho$ is a nonequilibrium state commuting with the equilibrium thermal state $\gamma_\beta = \exp(-\beta H)/Z(\beta)$, $E_{\max}$ is the maximum energy eigenvalue of the system, and $F(\beta) = -\beta^{-1}\ln Z(\beta)$ is the Helmholtz free energy. Here we refer to finite-size corrections as corrections that vanish when the dimension of the Hilbert space of the system is infinite. Indeed, $E_{\max}$ is finite only if the dimension of the Hilbert space is finite and, when $E_{\max} \to \infty$, (1.14) reduces to the usual Clausius inequality. Instances where our findings are applicable include systems with a finite number of qubits or spin chains with a finite number of sites.

To gain some insight into the majorization in a field theoretical setting, we consider pairs of states in $(1+1)$-dimensional CFTs, in particular free theories such as a compact boson and a Dirac fermion. As a pair of states we take the ground state and an excited state, then discretize the theory and map it to a fermionic chain, where we find the corresponding pair of states (yielding back to the CFT states in a continuum limit). Furthermore, we take the theory to live on a circle with periodic boundary conditions. The discrete fermionic chain then has a finite-dimensional Hilbert space, so we can use the standard definition of majorization. We bipartite the theory into a line segment and its complement and ask if a majorization order exists between the pair of pure states. The majorization condition involves reduced states, which depends on the bipartition, and thus on the relative size of the subsystem. While it is laborious to directly verify the majorization conditions, it is simpler to show them to be violated by comparing entanglement monotones or Schur concave quantifiers for the pair of states and test if the majorization-implied inequality is falsified for any monotone. In the final part of this paper we perform such comparisons, between the ground state and an excited state in a CFT and the corresponding pair in the periodic fermionic chain. We consider the entanglement entropy, the Rényi entropies $S^{(2)}$, $S^{(3)}$ and the monotone $M^{(2)}$, and compare which quantity gives the most stringent bound ruling out majorization in the largest range of bipartition.

This paper is organized as follows. In Sec. II, we first review some of the relevant basic concepts of quantum information theory. In Sec. III, we introduce two sequences of entanglement monotones: the moments of shifted modular Hamiltonian and the extremal polynomial monotones. We then generalize a theorem in [3], allowing us to construct infinite sequences of relative quantifiers that are also resource monotones. As an application, we derive a finite-size correction to the Clausius inequality. In Sec. IV as an application of the entanglement monotones, we consider Landauer erasure and derive infinite sequences of inequalities for the work cost of the erasure process, involving arbitrarily high cumulants of the state to be erased. We also discuss bounds on marginal entropy production in a system coupled to an environment. We then initiate an exploration of state majorization in CFTs in Sec. V. We compute the quantities $S, C, M$ in some simple free CFTs for the ground state and excited states and in the corresponding discretized fermionic chains, and we examine the monotonicity of some Schur concave quantifiers as proxies for state majorization. We conclude with a discussion and a description of various open problems in Sec. VI and a summary of the results of this paper in Sec. VII.

## II. SOME CONCEPTS OF QUANTUM INFORMATION THEORY

For the benefit of readers who are less familiar with some of the relevant concepts of quantum information theory, we briefly review some relevant background material.

In this paper our focus is on bipartite systems $A \cup B$, where the Hilbert space is decomposed as $\mathcal{H}_{A\cup B} = \mathcal{H}_A \otimes \mathcal{H}_B$. For a pair of quantum states described by the density matrices $\rho$ and $\sigma$, we first review the important concept of majorization (partial) order. Consider a pair of vectors $\boldsymbol{\lambda}, \boldsymbol{\kappa} \in \mathbb{R}^d$ and assume them to be ordered so that the components satisfy $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_d$ and likewise for $\boldsymbol{\kappa}$. We say that $\boldsymbol{\lambda}$ majorizes $\boldsymbol{\kappa}$, and denote $\boldsymbol{\lambda} \succ \boldsymbol{\kappa}$, when [29]

$$\sum_{k=1}^{m} \lambda_k \geqslant \sum_{k=1}^{m} \kappa_k, \quad \forall\, m = 1, 2, \dots, d. \tag{2.1}$$

Majorization defines a partial order in $\mathbb{R}^d$, and the definition easily extends to the case $d \to \infty$, where we have a countably infinite number of inequalities to satisfy. We then define majorization between two density matrices $\rho_1$ and $\rho_2$: $\rho_1 \succ \rho_2$ when $\boldsymbol{\lambda}_1 \succ \boldsymbol{\lambda}_2$, where $\boldsymbol{\lambda}_j$ is the ordered vector of eigenvalues of $\rho_j$. Note that the inequalities (2.1) become trivial for any pair of pure states. However, in a bipartite system, and when this partition is kept fixed, one can define a nontrivial majorization partial order for pure states. Consider a pair of pure states $|\psi\rangle, |\phi\rangle \in \mathcal{H}_{AB}$, and define majorization following that of the reduced density matrices,

$$|\psi\rangle \succ |\phi\rangle \quad \Leftrightarrow \quad \mathrm{Tr}_B(|\psi\rangle\langle\psi|) \succ \mathrm{Tr}_B(|\phi\rangle\langle\phi|). \tag{2.2}$$

Note that it does not matter whether the partial trace is taken over $B$ or $A$ because the resulting reduced density matrices in the two cases have the same eigenvalues, from the Schmidt decomposition. However, we emphasize that the definition depends on the choice of the bipartition $A \cup B$, and it would be more accurate to denote it by $|\psi\rangle \succ_{AB} |\phi\rangle$: An alternative bipartition $A' \cup B'$ in general leads to a partial order $\succ_{A'B'}$ among bipartite pure states, which is not equivalent with $\succ_{AB}$.

We will be interested in quantities that are monotonic under majorization. First, a function $g$ mapping a density matrix to a real number is said to be *Schur concave*, when, for any pair of density matrices, we have

$$\rho \succ \sigma \;\Rightarrow\; g(\rho) \leqslant g(\sigma). \tag{2.3}$$

Conversely, $g$ is Schur convex if $-g$ is Schur concave [29]. A stronger property is concavity, which is a crucial property we employ to construct entanglement monotones. We say that $g$ is *concave*, when, for any $0 \leqslant p \leqslant 1$ and for any pair of operators $\rho$ and $\sigma$, we have

$$g(p\rho + (1-p)\sigma) \geqslant pg(\rho) + (1-p)g(\sigma). \tag{2.4}$$

Conversely, a function $g$ is called convex if $-g$ is concave. A way to construct concave quantities is to begin with a function $f : [0, 1] \to \mathbb{R}$. By applying the function to an operator $\rho$, we obtain another operator $f(\rho)$. Assume that all the involved operators are diagonalised by unitaries $\rho = U\mathrm{diag}(\lambda_i)U^\dagger$, we have that $f(\rho) = U\mathrm{diag}(f(\lambda_i))U^\dagger$, where we assume that $f$ is well defined for all $\lambda_i$. We can now define a function $g$ through

$f$ as

$$g(\rho) = \mathrm{Tr}[f(\rho)]. \tag{2.5}$$

It has been proven that if $f$ is concave (convex) as a single real variable function, then $g$ defined in (2.5) is concave (convex) according to the definition (2.4) [29]. For example, when $f(x) = -x \ln x$ and $\rho$ is a density matrix, the function defined in (2.5) is the von Neumann entropy. Since $f(x) = -x \ln x$ is concave in [0,1], the von Neumann entropy is concave.

As we said, concavity is a stronger property than Schur concavity. Any function that is symmetric in its arguments [such as $g$ defined in (2.5) by the trace, it is symmetric, i.e., invariant with respect to permutations of $\lambda_i$] and concave is also Schur concave, while the opposite does not hold [29]. For example the Rényi entropies $S^{(\alpha)} = \frac{1}{1-\alpha} \ln \mathrm{Tr}(\rho^\alpha)$ are Schur concave when $\alpha > 0$, but concave only when $0 < \alpha \leqslant 1$. Also the so-called min and max entropies are Schur concave but not concave. On the other hand, the von Neumann and Tsallis [30] entropies are both concave and Schur concave (see [31] for a demonstration in the context of unified entropies).

A general way to define quantum operations $\mathcal{E}$ mapping an input state $\rho$ to an output state $\mathcal{E}(\rho)$ is by the operator-sum representation [32]

$$\mathcal{E}(\rho) = \sum_i K_i \rho K_i^\dagger, \tag{2.6}$$

with a collection of Kraus operators $\{K_i\}$ that satisfy the condition

$$\sum_i K_i^\dagger K_i \leqslant \mathbb{1}, \tag{2.7}$$

where, for a Hermitian $A$, $A \leqslant \mathbb{1}$ means that $\mathbb{1} - A$ has only non-negative eigenvalues. If the stronger condition $\sum_i K_i^\dagger K_i = \mathbb{1}$ applies, then $\mathcal{E}$ is trace preserving and it is called a quantum channel. Every quantum channel has a fixed point $\sigma_* = \mathcal{E}(\sigma_*)$ [33]. If the fixed point is the unit matrix, $\mathcal{E}(\mathbb{1}) = \mathbb{1}$, the operation $\mathcal{E}$ is called unital channel. In this case the Kraus operators satisfy the additional condition

$$\sum_i K_i K_i^\dagger = \mathbb{1}. \tag{2.8}$$

For us, an important feature of unital channels is that by Uhlmann's theorem [34,35] they imply majorization between the input and output states,

$$\rho \succ \mathcal{E}(\rho). \tag{2.9}$$

A simple proof (in English) of Uhlmann's theorem, based on the Hardy-Littlewood-Pólya theorem of majorization (which establishes that $\boldsymbol{\lambda} \succ \boldsymbol{\mu}$ iff there exists a bistochastic matrix $T$ such that $\boldsymbol{\mu} = T\boldsymbol{\lambda}$), can be found in [36] (in Appendix B therein). The converse is also true in the sense that if $\rho \succ \sigma$ there exists a unital channel with $\sigma = \mathcal{E}(\rho)$.

Another well-known class of operations are the LOCC. We can think of a LOCC as a process where quantum operations are performed by the two parties $A$ and $B$ separately, while classical communication allows the two parties to correlate their action. We emphasize again that for this process one must first decide on a bipartition, and then keep it fixed. Mathematically, LOCC operations can be represented as separable operations [37]

$$\rho \mapsto \Lambda(\rho) = \sum_i p_i A_i \otimes B_i \, \rho \, A_i^\dagger \otimes B_i^\dagger, \tag{2.10}$$

where $A_i$ and $B_i$ are operators acting on the local subsystems $A$ and $B$ respectively. Note, however, that it is notoriously difficult to characterize the set of operations, which can be a achieved through LOCC and that the class of separable operations (2.10) is strictly larger than LOCC. The LOCC operations can be used to define entanglement; indeed entanglement cannot be created but only decreased by these operations. Moreover, separable states, which are states $\rho$ of the form

$$\rho = \sum_i p_i \rho_i \otimes \sigma_i, \tag{2.11}$$

where $\rho_i, \sigma_i$ are states in the subsystems $A, B$ respectively and $\sum_i p_i = 1$ with $p_i \geqslant 0$, can be prepared from nonentangled pure states $|\psi\rangle_A \otimes |\phi\rangle_B$ by separable operations (2.10), which can easily be seen using an alternative representation of (2.11) as an ensemble of factorized pure states [38]

$$\rho = \sum_j \tilde{p}_j \, |\chi_j\rangle_A A\langle\chi_j| \otimes |\eta_j\rangle_B B\langle\eta_j|, \tag{2.12}$$

which follows from ensemble decompositions of $\rho_i, \sigma_j$ and index relabeling. This leads to the alternative definition of $\rho$ being entangled if and only if it is not separable. Entangled states then act as a resource for LOCC processes. For the simplest case, attempting to convert a pure state $|\Psi\rangle \in \mathcal{H}_{AB}$ to another state $|\Phi\rangle \in \mathcal{H}_{AB}$, Nielsen's theorem [2] provides a necessary and sufficient criterion for the possibility of state transition. It states that it is possible to convert $|\Psi\rangle$ to $|\Phi\rangle$ by LOCC with reference to a bipartition $AB$, namely $|\Psi\rangle \xrightarrow[\text{LOCC}]{} |\Phi\rangle$, if and only if the majorization condition $|\Phi\rangle \succ |\Psi\rangle$ is fulfilled, or equivalently for the corresponding reduced states, $\rho_A \succ \sigma_A$. Unfortunately, the majorization condition (2.1) is somewhat inconvenient to verify. The task of first finding all the eigenvalues and then comparing all the partial sums is in general rather laborious (if not intractable). On the other hand, it is easier to rule out the possibility of the LOCC transition. One can consider a Schur concave function $g$, which is nonincreasing under the transition. Thus, if we find that $\Delta g \equiv g(\rho_A) - g(\sigma_A) > 0$, the transition is ruled out.

More general LOCC processes, where a pure or mixed state is converted to a mixed state $\rho \xrightarrow[\text{LOCC}]{} \sigma = \Lambda(\rho)$, where $\rho$ and $\sigma$ are general (pure or mixed) states in the composite system $A \cup B$, have no simple characterization by majorization. Moreover, Schur concave functions are in general not monotonic under such processes. This leads one to consider entanglement monotones. The key requirement for monotonicity is $E(\rho) \geqslant E(\Lambda(\rho))$. In more detail, an entanglement monotone $E(\rho)$ is defined as a map $\rho \mapsto E(\rho) \in \mathbb{R}$, which satisfies [37]

(1) $E(\rho) \geqslant 0$

(2) $E(\rho) = 0$ if $\rho$ is separable

(3) $E(\rho)$ does not increase on average [39] under LOCC, which means

$$E(\rho) \geqslant \sum_i p_i E\left(\frac{K_i \rho K_i^\dagger}{\mathrm{Tr}(K_i \rho K_i^\dagger)}\right), \tag{2.13}$$

where $K_i = A_i \otimes B_i$ are Kraus operators of a LOCC process as in (2.10), and $p_i = \mathrm{Tr}(K_i \rho K_i^\dagger)$.

Central to this work is Vidal's theorem [23], which provides a way to construct monotones from concave quantities of the type (2.5). Consider a pure state $|\Psi\rangle$ and $\rho_A = \mathrm{Tr}_B|\Psi\rangle\langle\Psi|$, and any function $g(\rho_A)$ such that

(1) $g$ is concave

(2) $g$ is invariant under unitary transformations $U_A$, namely $g(U_A \rho_A U_A^\dagger) = g(\rho_A)$

then Vidal's theorem establishes first that $E_{\mathrm{pure}}(|\Psi\rangle\langle\Psi|) \equiv g(\rho_A)$ is an entanglement monotone for pure states (a pure state entanglement monotone). Moreover, one can extend $E_{\mathrm{pure}}$ to a monotone $E$ for mixed states by using the convex-roof extension, which is defined as follows. Given the density matrix $\sigma$, one considers the minimum over all of its ensemble decompositions $\{p_j, |\Psi_j\rangle\}$ realizing $\sigma = \sum_j p_j |\Psi_j\rangle\langle\Psi_j|$ and defines

$$\begin{aligned} E(\sigma) &\equiv \min_{\{p_j,|\Psi_j\rangle\}} \sum_j p_j\, E_{\mathrm{pure}}(|\Psi_j\rangle\langle\Psi_j|) \\ &= \min_{\{p_j,|\Psi_j\rangle\}} \sum_j p_j\, g(\mathrm{Tr}_B|\Psi_j\rangle\langle\Psi_j|). \end{aligned} \tag{2.14}$$

One can then show that $E$ is an entanglement monotone [40] [23,37,41]. A subtle feature of this construction is that while $g$ is concave with respect to states for the $A$ system, $E$ is convex with respect to states on the $AB$ system (as the name "convex roof extension" implies). While our construction of entanglement monotones follows the above steps, in this paper we do not need to explicitly use the convex roof extension, since we consider only LOCC processes between pure states or processes with unital channels, both implying majorization, where Schur concavity is sufficient to give monotonicity. Since as we mentioned earlier in this section, concavity of $g$ also implies that it is Schur concave, it can therefore directly be used to find necessary criteria for the existence of either type of process.

We now move to construct an infinite sequence of entanglement monotones generalizing $M$ in (1.6) and inequalities generalizing (1.7) for pairs of majorizing states.

## III. RESOURCE MONOTONES AND MAJORIZING STATE TRANSITIONS

In this section we will first construct sequences of entanglement monotones, and then show how they can be applied to define more general resource monotones. As an application, we will briefly consider (the resource theory of) quantum thermodynamics.

### A. An infinite sequence of entanglement monotones

It was established in [3] that $M$ defined in (1.6) is a pure-state entanglement monotone because it is Schur concave, and thus is monotonic for reduced density matrices under majorization. This was proven as a corollary of a more general theorem involving relative quantities, as we will discuss in Sec. III C. Here we present an alternative simple proof, showing that $M(\rho)$ is concave. It is straightforward to rewrite (1.6) as

$$M(\rho) = \mathrm{Tr}[f(\rho)], \tag{3.1}$$

where

$$f(x) = x[-\ln x + 1]^2, \tag{3.2}$$

and it is simple to see that $f(x)$ is a concave function in the unit interval, namely for $x \in [0, 1]$. This implies that $M(\rho)$ is concave, i.e., it satisfies

$$M(p\rho_1 + (1-p)\rho_2) \geqslant pM(\rho_1) + (1-p)M(\rho_2), \tag{3.3}$$

for any pair of density matrices $\rho_1, \rho_2$, and for all $p \in [0, 1]$. Concavity in turn implies Schur concavity, the property of $M$ proven in [3]. Furthermore, by Vidal's theorem [23] concavity implies that $M(\rho) - 1$ can be extended by the convex-roof extension to a proper entanglement monotone for all states. Since $K = -\ln\rho$ is called modular Hamiltonian [4] and we shift it by a constant 1, we call $M$ as the *second moment of shifted modular Hamiltonian*.

It is straightforward to find concave generalizations of $M$ involving higher cumulants. Up to addition of an overall constant term, which has no effect to concavity, we define

$$f_n(x) = x\,(-\ln x + b_n)^n = (-1)^n\, x\,(\ln x - b_n)^n. \tag{3.4}$$

Since (for $n \geqslant 2$) $f''(x) = x^{-1}[n(b_n - \ln x)^{n-2}(\ln x + n - 1 - b_n)] \leqslant 0$ for

$$b_n \geqslant n - 1\,, \tag{3.5}$$

this parameter range ensures that $f_n(x)$ is concave over the unit interval for $n \in \mathbb{N}_+$. We then define a concave quantity (and by Vidal's theorem, a pure state entanglement monotone)

$$M^{(n)}(\rho; b_n) = \mathrm{Tr}[f_n(\rho)] - b_n^n. \tag{3.6}$$

We call $M^{(n)}$ as the $n$th *moment of shifted modular Hamiltonian*. The subtraction of the constant $b_n^n$ in (3.6) ensures that $M^{(n)}(\rho_{\,\mathrm{pure}}; b_n) = 0$, when $\rho_{\,\mathrm{pure}}$ describes a pure state. When $n = 2$ and $b_n = 1$, the expression (3.6) reduces up to an additive constant to (3.1). Moreover, since $M^{(n)}$ has the form (2.5) with $f_n$ concave in [0,1], we can conclude that it is also Schur concave, from the discussion in Sec. II.

To rewrite (3.6) as combination of cumulants, we expand it first as a linear combination of moments $\mu_k(\rho)$ of modular Hamiltonian

$$\begin{aligned} M^{(n)}(\rho; b_n) &= \sum_{k=1}^{n} \binom{n}{k} b_n^{n-k} \mu_k(\rho), \\ \mu_k &= \mathrm{Tr}[\rho\,(-\ln\rho)^k] = \mathrm{Tr}(\rho\, K^k). \end{aligned} \tag{3.7}$$

Then, in turn we can use the relation between moments and cumulants (see e.g., [42]) to write

$$\mu_k(\rho) = \sum_{\{p_j\}}{}' \, k! \prod_j \frac{1}{p_j!(j!)^{p_j}}\, C_j(\rho)^{p_j}, \tag{3.8}$$

where the restricted sum $\sum'$ is over all partitions $\{p_j\}$ of $k = \sum_j j\, p_j$, and $C_j(\rho)$ is the $j$th cumulant of modular Hamiltonian. A more streamlined way is to expand the moments $M^{(n)}(\rho; b_n)$ as cumulants $\tilde{C}_j(\rho)$ of $-\ln\rho + b_n$ as follows:

$$M^{(n)}(\rho; b_n) = \sum_{\{p_j\}}' n! \prod_j \frac{1}{p_j!(n!)^{p_j}} \tilde{C}_j(\rho)^{p_j} - b_n^n, \tag{3.9}$$

where $n = \sum_j j\, p_j$. In the sum over partitions, terms involving cumulants $\tilde{C}_j$ of $-\ln\rho + b_n$ of order $j \geqslant 2$ reduce to $C_j(\rho)$ due to translation invariance. Terms involving $j = 1$ give $\tilde{C}_1^{p_1}(\rho) = [S(\rho) + b_n]^{p_1}$. In this way, setting for simplicity $b_n$ to be the smallest possible value $b_n = n - 1$ for concavity, we obtain the sequence

$$\begin{aligned}
M^{(1)}(\rho; 0) &= S(\rho),\\
M^{(2)}(\rho; 1) &= [S(\rho) + 1]^2 + C(\rho) - 1,\\
M^{(3)}(\rho; 2) &= [S(\rho) + 2]^3 + 3C(\rho)[S(\rho) + 2] + C_3(\rho) - 2^3,\\
M^{(4)}(\rho; 3) &= [S(\rho) + 3]^4 + 6C(\rho)[S(\rho) + 3]^2 + 3C(\rho)^2\\
&\quad + 4C_3(\rho)[S(\rho) + 3] + C_4(\rho) - 3^4, \; \cdots .
\end{aligned} \tag{3.10}$$

The cumulants of modular Hamiltonian can be derived from a generating function as [3,21]

$$C_n(\rho) = (-1)^n \frac{d^n}{d\alpha^n}[(1-\alpha)\, S^{(\alpha)}(\rho)]|_{\alpha=1}, \tag{3.11}$$

where $S^{(\alpha)}(\rho)$ are the Rényi entropies (1.2). The Rényi entropies in turn are determined by the entanglement spectrum. Note that the Rényi entropies themselves are not concave when the Rényi index $\alpha > 1$. On the other hand, the full information about the bipartite entanglement is encapsulated by the entanglement spectrum, which provides the Rényi entropies. They in turn can be converted to the cumulants, which can be converted to the entanglement monotones $M^{(n)}$. To summarize, the above sequence provides a way to convert the full entanglement spectrum into an infinite sequence of entanglement monotones. It is helpful to note that the cumulants of modular Hamiltonian are additive, namely

$$C_n(\rho_1 \otimes \rho_2) = C_n(\rho_1) + C_n(\rho_2). \tag{3.12}$$

This can be verified in different ways, most simply it follows from the additivity of the Rényi entropies and the generating function formula (3.11) for cumulants. Also, for the maximally mixed state $\rho = \mathbb{1}/d$ of a system with a $d$-dimensional Hilbert space, where $\mathbb{1}$ is the $d \times d$ identity matrix, the same formula gives a simple proof of

$$C_n(\mathbb{1}/d) = \begin{cases} \ln d & \text{for } n = 1,\\ 0 & \text{for } n \geqslant 2. \end{cases} \tag{3.13}$$

Notice that $C_n = \mu_n = 0$ for pure states. Also, for maximally mixed state, we have

$$M^{(n)}(\mathbb{1}/d; b_n) = (\ln d + b_n)^n - b_n^n, \tag{3.14}$$

so that one would need to rescale by an overall normalization constant if one wishes to follow the convention [37] that an entanglement monotone is normalized to $\ln d$ for the maximally mixed state (i.e., $M^{(n)}(\mathbb{1}/d; b_n) = \ln d$).

In fact, the sequence $M^{(n)}$ can be derived more straightforwardly from the Rényi entropies, converting the latter to a generating function as follows. Since

$$\mathrm{Tr}(\rho^\alpha) = \exp[(1-\alpha)S^{(\alpha)}(\rho)], \tag{3.15}$$

in terms of the Rényi entropies $S^{(\alpha)}(\rho)$, by defining

$$\begin{aligned}
k_\alpha(\rho; b) &\equiv e^{-\alpha b}\,\mathrm{Tr}(\rho^\alpha) = \mathrm{Tr}\big[e^{\alpha(\ln\rho - b)}\big]\\
&= \exp\big[-\alpha b + (1-\alpha)S^{(\alpha)}(\rho)\big],
\end{aligned} \tag{3.16}$$

we have that

$$M^{(n)}(\rho; b_n) = \left[e^b(-1)^n \frac{d^n}{d\alpha^n} k_\alpha(\rho; b)\right]\bigg|_{\alpha=1, b=b_n} - b_n^n, \tag{3.17}$$

which provides a prescription for converting the Rényi entropies to the infinite sequence of entanglement monotones.

To summarize: The complete information of the bipartite entanglement is encapsulated by the entanglement spectrum and this information can be repackaged first to the set of Rényi entropies, which in turn can be converted to an infinite tower of entanglement monotones $M^{(n)}(\rho; b_n)$.

Note that concavity only gives a lower bound $b_n \geqslant n - 1$. Other conditions may lead to a particular choice for the value of the constant $b_n$.

In Appendix A 1 we report additional comments on how to construct entanglement monotones exploiting the cumulants of the modular Hamiltonian.

### B. Extremal polynomial monotones and inequalities for state transitions

In this section we perform a more general analysis of infinite sequence of monotones constructed from convex polynomials of the moments of $\ln\rho$. With some abuse of terminology, we will refer to them as "polynomial entanglement monotones" for brevity. We find that such monotones form an infinite-dimensional cone, which is determined by extremal rays, defining what we will call for brevity *extremal polynomial monotones*. The extremal polynomial monotones give rise to an infinite sequence of inequalities that must be satisfied in majorizing state transformations.

Our starting point is the general functional

$$P(\rho) = \mathrm{Tr}[\rho\, F(\ln\rho)]. \tag{3.18}$$

The corresponding scalar function $f(x) = x\, F(\ln x)$ is convex if $f'' \geqslant 0$ for $x \in [0, 1]$. For convenience, we focus on convex measures, which can be converted to concave measures by a minus sign. This translates into the condition

$$F'(y) + F''(y) \geqslant 0, \qquad y \leqslant 0. \tag{3.19}$$

For example, $f(y) = y$ clearly meets this criterion, with $P$ in (3.18) being minus the von Neumann entropy. Convex functions yield convex measures, which are monotonic under majorization (Schur convex)

$$\rho \succ \sigma \implies P(\rho) \geqslant P(\sigma). \tag{3.20}$$

In the previous subsections, we have seen that for suitable functions $F$ we get monotones, which we use to test for majorization. We now want to be more systematic and classify all $F$ with the property that $F' + F'' \geqslant 0$ for $y \leqslant 0$.

We restrict our survey here by focusing on measures where $F$ is polynomial in $y = \ln\rho$. Let us introduce

$$G(y) \equiv F'(y) + F''(y). \quad (3.21)$$

From (3.19), consider all polynomials $G(y)$ such that $G(y) \geqslant 0$ for $y \leqslant 0$. For each such polynomial $G$, there is a unique polynomial $F$ such that $F'' + F' = G$ (up to vanishing constant terms, which can be added at will, e.g., changing the value of $F$ and $P$ for pure states).

The space of positive semidefinite polynomials $G(y)$ on negative real axis is a convex cone in the sense that, given a set of functions $G_i$ with this property, a linear combination $\sum_i \alpha_i G_i$ with non-negative coefficients will also have this property [43]. Cones are completely determined by specifying all "extremal" rays, which are functions $G$, which do not admit a nontrivial decomposition of the type $G = \sum_i \alpha_i G_i$. The most general $G$ will then be a linear combination of extremal functions $G$ with non-negative coefficients. In general, there can be finitely or infinitely many extremal $G$.

From the perspective of monotones, the extremal $G_j$ will provide a complete list of nontrivial "extremal polynomial monotones", with all other polynomial monotones being linear combinations of extremal polynomial monotones with non-negative coefficients. It is therefore interesting to classify all such extremal monotones. For this we need to classify all extremal $G_j$. We can use known results from the theory of positive semidefinite polynomials, which can be summarized as the following theorem.

*Theorem 1.* All positive semidefinite polynomials $G(y)$ on the negative half-line $y \in (-\infty, 0]$ have the following form. For polynomials $G(y)$ of degree $2d$ (with $d \geqslant 1$), they are linear combinations with non-negative coefficients of polynomials of the form $G_{\vec{a}}(y) = \prod_{i=1}^{d}(y + a_i)^2$ with all $a_i \geqslant 0$. For polynomials of degree $2d + 1$ they are linear combinations with non-negative coefficients of polynomials of the form $G_{\vec{a}}(y) = -y\prod_{i=1}^{d}(y + a_i)^2$ with again all $a_i \geqslant 0$.

We defer the detailed proof to Appendix A 2. The result is essentially known in mathematics (see [44] for a review of non-negative polynomials).

We emphasize that the higher moments of modular Hamiltonian $M^{(n)}$, introduced in Sec. III A, are in general not extremal monotones. Consider $F(x) = (x - b_k)^k$ so that $G(x) = F' + F'' = k(x - b_k)^{k-2}(x - b_k + k - 1)$. This has an isolated zero at $x = b_k - k + 1$. Since $b_k \geqslant k - 1$, this isolated zero cannot be on the negative real axis. In the limiting case $b_k = k - 1$ one obtains $G = kx(x - k + 1)^{k-2}$. This has many zeros on the positive real line rather than even degeneracy zeros on the negative real line, hence as such it is not extremal. Thus the moments $M^{(n)}(\rho; b_n)$ in general are linear combinations of the extremal monotones $P_E^{(k)}(\rho)$.

Let us now study the lowest degree examples in detail. For $F$ of degree 1, $G$ has degree zero and must be a non-negative constant, which we can take to be 1. Then $F(y) = y$ and the resulting extremal monotone $P_E^{(1)}$ is minus the entropy, namely $P_E^{(1)} = -S(\rho) = -M^{(1)}$.

For $F$ of degree 2, $G$ is of degree one. According to the Theorem 1, there is a unique extremal $G$, which is $-y$. Solving $F'' + F' = -y$, we have $F = y - y^2/2$. We might as well take twice this as extremal functions are defined up to overall normalization only. In that case $F = 2y - y^2$.

Thus, we have proven that

$$\begin{aligned} P_E^{(2)}(\rho) &= \mathrm{Tr}[\rho(2\ln\rho - \ln^2\rho)] \\ &= -C(\rho) - S(\rho)^2 - 2S(\rho) \\ &= -M^{(2)}(\rho, b_2 = 1) \end{aligned} \quad (3.22)$$

is an extremal monotone. Note that, up to second order, the two classes of monotones are related by $M^{(n)}(\rho; n-1) = -P_E^{(n)}(\rho)$, while this is no longer true for $n \geqslant 3$.

Instead of the entropy production inequality with the finite correction (1.7), the extremal monotone appears to give a slightly sharper inequality. The statement

$$\begin{aligned} \rho \succ \sigma \implies P_E^{(2)}(\rho) &= -C(\rho) - S(\rho)^2 - 2S(\rho) \\ &\geqslant -C(\sigma) - S(\sigma)^2 - 2S(\sigma) = P_E^{(2)}(\sigma) \end{aligned} \quad (3.23)$$

can be rewritten as the inequality

$$S(\sigma) - S(\rho) \geqslant \frac{C(\rho) - C(\sigma)}{S(\rho) + S(\sigma) + 2}, \quad (3.24)$$

which appears to be slightly sharper than (1.7) involving $M$, $S$, and $C$.

Until now, the inequalities have been the same as the ones coming from

$$M^{(n)}(\sigma; n-1) \geqslant M^{(n)}(\rho; n-1), \quad (3.25)$$

when $\rho \succ \sigma$. In terms of the cumulants, using (3.9), this sequence has the explicit form

$$\begin{aligned} &S(\sigma) \geqslant S(\rho), \\ &[S(\sigma)+1]^2 + C(\sigma) \geqslant [S(\rho)+1]^2 + C(\rho), \\ &[S(\sigma)+2]^3 + 3C(\sigma)[S(\sigma)+2] + C_3(\sigma) \\ &\quad \geqslant [S(\rho)+2]^3 + 3C(\rho)[S(\rho)+2] + C_3(\rho), \ \cdots . \end{aligned} \quad (3.26)$$

Notice that the "second law" of entropy (claiming that the entropy is nondecreasing in transitions $\rho \mapsto \sigma$ with $\rho \succ \sigma$) becomes refined into an infinite sequence of inequalities that must likewise be satisfied. However, at orders $n \geqslant 3$ the extremal polynomial monotones may give tighter inequalities.

For $F$ of degree 3, $G$ is of degree two and must be of the form $(y + a)^2$ with $a > 0$. So in this case we get a one-parameter family of extremal monotones. Solving $F' + F'' = G$ yields

$$F = \tfrac{1}{3}y^3 + (a-1)y^2 + (a^2 - 2a + 2)y, \quad (3.27)$$

and this gives rise to a one-parameter family of extremal monotones for $a \geqslant 0$,

$$\begin{aligned} P_E^{(3)}(\rho) = \mathrm{Tr}\bigg[\rho\bigg(&\frac{1}{3}\ln^3\rho + (a-1)\ln^2\rho \\ &+ (a^2 - 2a + 2)\ln\rho\bigg)\bigg], \end{aligned} \quad (3.28)$$

and correspondingly to an infinite number of inequivalent inequalities. It is useful to express the coefficients of $a^k$ in terms of the monotones $M_n \equiv M^{(n)}(\rho; n-1)$. Let us first denote $r^{(k)} \equiv \mathrm{Tr}[\rho(-\ln\rho)]$ and $r \equiv r^{(1)}$. We have (including fourth order for future reference)

$$M_1 = \mathrm{Tr}[\rho(-\ln\rho)] = r, \tag{3.29}$$

$$M_2 = \mathrm{Tr}[\rho(-\ln\rho+1)^2] - 1 = r^{(2)} + 2\,r, \tag{3.30}$$

$$M_3 = \mathrm{Tr}[\rho(-\ln\rho+2)^3] - 2^3 = r^{(3)} + 6\,r^{(2)} + 12\,r, \tag{3.31}$$

$$M_4 = \mathrm{Tr}[\rho(-\ln\rho+3)^4] - 3^4 = r^{(4)} + 12\,r^{(3)} + 54\,r^{(2)} + 108\,r. \tag{3.32}$$

We can then express

$$\begin{aligned} P_E^{(3)}(\rho) &= -a^2 r + a(r^{(2)} + 2r) - \tfrac{1}{3}r^{(3)} - r^{(2)} - 2r \\ &= -a^2 M_1 + aM_2 - \tfrac{1}{3}M_3 + M_2. \end{aligned} \tag{3.33}$$

We could thus calculate $P_E^{(3)}$ from the Rényi entropies by using first the generating function formula (3.17) for $M^{(n)}$.

We explore the cubic case a bit more, to find a tight inequality. Assuming $\rho \succ \sigma$, we obtain

$$P_E^{(3)}(\rho) \geqslant P^{(3)}(\sigma), \tag{3.34}$$

which at this stage is an infinite family of inequalities due to the free parameter $a$. The inequality (3.34) can be written more explicitly as

$$\begin{aligned} w_2 a^2 + w_1 a + w_0 \equiv a^2\Delta M_1 + -a\Delta M_2 + \tfrac{1}{3}\Delta M_3 \\ - \Delta M_2 \geqslant 0, \end{aligned} \tag{3.35}$$

where

$$\Delta M_n = M^{(n)}(\sigma; n-1) - M^{(n)}(\rho; n-1) \geqslant 0. \tag{3.36}$$

Notice that $w_2 \geqslant 0$ while $w_1 \leqslant 0$. The quadratic function of $a$ will therefore have a minimum at

$$a_0 = -\frac{w_1}{2w_2} \geqslant 0. \tag{3.37}$$

In order to check whether the inequalities (3.35) are satisfied for all $a \geqslant 0$, we only need to verify it for $a_0$ for which the quadratic polynomial in (3.35) takes its minimum value. This leads to

$$w_0 - \frac{w_1^2}{4w_2} \geqslant 0\,. \tag{3.38}$$

Finally, substituting the $w_n$ from (3.35), the inequality (3.38) takes the compact form

$$\Delta M_3 \geqslant 3\Delta M_2 + \frac{3}{4}\frac{(\Delta M_2)^2}{\Delta M_1}. \tag{3.39}$$

Now we can explicitly see the advantage of (3.39) over the simple inequality $\Delta M_3 \geqslant 0$ from (3.25). We could then substitute the explicit forms of $\Delta M_n$ from (3.36) to explore how entropy production and changes in other cumulants up to third order are bounded by each other. Alternatively, we can study differences in moments $r^{(n)}$ and $s^{(n)} \equiv \mathrm{Tr}[\sigma(-\ln\sigma)^n]$. For example, the difference $r^{(3)} - s^{(3)}$ has a lower bound in terms of $r$, $r^{(2)}$, $s$ and $s^{(2)}$,

$$r^{(3)} - s^{(3)} \leqslant -3(r-s) + \frac{3}{4}\frac{\left(r^{(2)} - s^{(2)}\right)^2}{r-s}. \tag{3.40}$$

Note that (3.49) has an interesting hierarchy where the third-order $\Delta M_3$ is bounded by are bounded by combinations $\Delta M_k$ with $k = 1, 2$, or the next-order inequality is bounded by the previous-order inequalities. This suggests that higher-order inequalities could have some interesting recursive structure. We explore this some more by working out the fourth-order monotone $P_E^{(4)}$ and the resulting inequalities.

To find $P^{(4)}(\rho)$, we first need to solve $F''(y) + F'(y) = G(y)$ with

$$G(y) = -y(y+a)^2, \tag{3.41}$$

where $y = \ln\rho$ and $a > 0$. The polynomial solution is

$$\begin{aligned} F(y) = a^2\left(y - \frac{y^2}{2}\right) + a\left(-4y + 2y^2 - \frac{2y^3}{3}\right) \\ + 6y - 3y^2 + y^3 - \frac{y^4}{4}, \end{aligned} \tag{3.42}$$

collecting coefficients of $a^k$. Then, $P_E^{(4)}(\rho) = \mathrm{Tr}[\rho F(\rho)]$, whose explicit form reads

$$\begin{aligned} P_E^{(4)} = a^2\left(-r - \frac{r^{(2)}}{2}\right) + a\left(4r + 2r^{(2)} + \frac{2r^{(3)}}{3}\right) \\ -6r - 3r^{(2)} - r^{(3)} - \frac{r^{(4)}}{4}. \end{aligned} \tag{3.43}$$

The coefficients of $a^k$ can be expressed in terms of the monotones $M_n \equiv M^{(n)}(\rho; n-1)$ and using (3.32). With some calculation, we find

$$-r - \frac{r^2}{2} = -\frac{1}{2}M_2, \tag{3.44}$$

$$4r + 2r^{(2)} + \frac{2r^{(3)}}{3} = \frac{2}{3}(M_3 - 3M_2), \tag{3.45}$$

$$-6r - 3r^{(2)} - r^{(3)} - \frac{r^{(4)}}{4} = -\frac{1}{4}(M_4 - 8M_3 + 6M_2), \tag{3.46}$$

so that

$$\begin{aligned} P_E^{(4)}(\rho) = \left(-\frac{1}{2}M_2\right)a^2 + \left[\frac{2}{3}(M_3 - 3M_2)\right]a \\ - \frac{1}{4}(M_4 - 8M_3 + 6M_2). \end{aligned} \tag{3.47}$$

Now consider a pair of majorizing states. As before, we have

$$\rho \succ \sigma \;\Rightarrow\; P^{(4)}(\rho) \geqslant P^{(4)}(\sigma). \tag{3.48}$$

Let us denote $\widetilde{M}_n \equiv M^{(n)}(\sigma; n-1)$ and $\Delta M_n = M^{(n)}(\sigma; n-1) - M^{(n)}(\rho; n-1) \geqslant 0$, as in (3.36). As before, we rewrite the inequality (3.48) in the form

$$w_2 a^2 + w_1 a + w_0 \geqslant 0, \tag{3.49}$$

where now

$$w_2 \equiv \tfrac{1}{2}(\widetilde{M}_2 - M_2) \equiv \tfrac{1}{2}\Delta M_2 \geqslant 0, \tag{3.50}$$

$$w_1 \equiv -\tfrac{2}{3}(\Delta M_3 - 3\Delta M_2) \leqslant 0, \tag{3.51}$$

$$w_0 \equiv \tfrac{1}{4}(\Delta M_4 - 8\Delta M_3 + 6\Delta M_2). \tag{3.52}$$

Note that in the above $w_1 \leqslant 0$ since $\Delta M_3 \geqslant 3\Delta M_2$ by the inequality (3.39). We then find the minimum of (3.49) at $a_0$ as in (3.37) and obtain the same $w$ inequality (3.38) as before.

Substituting the $w_k$ from (3.50) by $\Delta M_n$ gives the final form of the $P_E^{(4)}$ inequality,

$$\Delta M_4 \geqslant 8\Delta M_3 - 6\Delta M_2 + \frac{8}{9}\frac{(\Delta M_3 - 3\Delta M_2)^2}{\Delta M_2}. \quad (3.53)$$

The right-hand side is positive since

$$8\Delta M_3 - 6\Delta M_2 = 6\Delta M_3 + 2(\Delta M_3 - 3\Delta M_2) \geqslant 0, \quad (3.54)$$

from (3.39). We could use (3.39) twice in the right-hand side of (3.53) to relax the lower bound a bit to the inequality

$$\Delta M_4 \geqslant 6\Delta M_3 + \frac{3}{2}\frac{(\Delta M_2)^2}{\Delta M_1} + \frac{1}{2}\frac{(\Delta M_2)^3}{\Delta M_1}. \quad (3.55)$$

The upshot is that the fourth-order inequality is manifestly tighter than just $\Delta M_4 \geqslant 0$. There is still some interesting recursive structure, albeit the terms appearing in the right-hand side of (3.53) do not appear in the same combinations as in the previous-order inequalities: both $\Delta M_3$ and $\Delta M_3 - 3\Delta M_2$ appear instead of only $\Delta M_3 - 3\Delta M_2 - \frac{3}{4}\frac{(\Delta M_2)^2}{\Delta M_1}$.

The virtue of $M^{(n)}(\rho; b_n)$ is that they are simple to compute (for example when the Rényi entropies are known) and provide inequalities involving cumulants up to order $n$. The inequalities $\Delta M^{(n)} \geqslant 0$ are expected to be weaker than those derived from the extremal polynomial monotones $P_E^{(n)}(\rho)$. A trade-off is that the latter inequalities are less straightforward to derive, due to the increasingly many free parameters contained in $P_E^{(n)}$, which need to be optimized to make the inequalities as tight as possible.

### C. Other resource monotones from pure state entanglement monotones

In Ref. [3] a more general version of majorization has also been considered, which has applications to other resource theories than that of entanglement, e.g., to quantum thermodynamics. Let $\mathcal{D}$ denote the set of quantum states of a given system. Consider two pairs of states $\rho, \sigma \in \mathcal{D}$ and $\rho', \sigma' \in \mathcal{D}$. If there exists a quantum channel $\mathcal{E}$ in $\mathcal{D}$ such that $\mathcal{E}(\rho) = \rho'$ and $\mathcal{E}(\sigma) = \sigma'$, we denote $(\rho, \sigma) \succ (\rho', \sigma')$, defining a partial order between pairs of states. A special case is the fixed point $\sigma = \sigma' \equiv \sigma_*$ of the free operations, we then write $\rho \succ_{\sigma_*} \rho'$ instead of $(\rho, \sigma_*) \succ (\rho', \sigma_*)$ and say that $\rho$ *$\sigma_*$-majorizes* $\rho'$. An important example is the *thermomajorization*, where the fixed point is the Gaussian thermal state, $\sigma_* = e^{-\beta H}/Z$ (or a generalized Gaussian state). This leads to a partial order in quantum thermodynamics. For the generalized majorization $(\rho, \sigma) \succ (\rho', \sigma')$, the relative entropy $S(\rho||\sigma)$ is monotonic with

$$S(\rho||\sigma) \geqslant S(\rho'||\sigma'), \quad (3.56)$$

and is called a resource monotone in the above context. The monotonicity (3.56) is a special case of the more general contractivity property $S(\rho||\sigma) \geqslant S(\mathcal{N}(\rho)||\mathcal{N}(\sigma))$ for any quantum channel $\mathcal{N}$.

In Ref. [3] a new relative quantifier was introduced, which takes the form

$$M_x(\rho||\sigma) = C(\rho||\sigma) + (1 - \ln(x) - S(\rho||\sigma))^2, \quad (3.57)$$

where $\sigma$ is a full rank state, $C(\rho||\sigma)$ is the variance of the relative modular Hamiltonian [45],

$$C(\rho||\sigma) = \mathrm{Tr}[\rho(\ln\rho - \ln\sigma)^2] - S(\rho||\sigma)^2, \quad (3.58)$$

where $S(\rho||\sigma) \equiv \mathrm{Tr}[\rho(\ln\rho - \ln\sigma)]$ is the relative entropy, which is the expectation value of the relative modular Hamiltonian. They considered pairs of state such that $[\rho, \sigma] = [\rho', \sigma'] = 0$, further assuming that $\sigma, \sigma'$ are both full rank. In this setting, they proved some interesting properties for the quantifier (3.57), and proved a lower bound for the production of relative entropy, involving variations of the relative variance. In this section we generalize this construction of [3] to an infinite class of relative quantifiers, which are also limited to pairs of commuting states. A more general analysis in the noncommuting case remains an important open problem.

Consider first quantifiers of the form (3.18),

$$E(\rho) = \mathrm{Tr}[\rho F(\ln\rho)], \quad (3.59)$$

where we change the notation with respect to (3.18), to emphasize that now we allow $F(x)$ to be any smooth function (not just a polynomial), such that $xF(x)$ is concave in the unit interval $x \in [0, 1]$. By Vidal's theorem, (3.59) defines a pure state entanglement monotone in a bipartite system $A \cup B$, when $\rho$ is the reduced state of a pure state $|\psi\rangle_{A\cup B}$. Then, for a pair of commuting full rank density matrices $\rho, \sigma$, we define a relative quantifier

$$\begin{aligned} E_x(\rho||\sigma) &= E(x\rho\sigma^{-1}) = \mathrm{Tr}[\rho F(\ln(x\rho\sigma^{-1}))] \\ &= \mathrm{Tr}[\rho F(\ln\rho - \ln\sigma + \ln(x))], \end{aligned} \quad (3.60)$$

where $x$ is a real number. We then prove the following theorem, which generalizes Theorem 12 of [3]:

*Theorem 2.* Let $(\rho, \sigma)$ and $(\rho', \sigma')$ be two pairs of commuting states, that is $[\rho, \sigma] = [\rho', \sigma'] = 0$, and $\sigma, \sigma'$ both full rank. If $(\rho, \sigma) \succ (\rho', \sigma')$, namely $\rho \succ \rho'$ and $\sigma \succ \sigma'$, then

$$E_{s_{\min}}(\rho'||\sigma') \geqslant E_{s_{\min}}(\rho||\sigma), \quad (3.61)$$

where $s_{\min}$ denotes the smallest eigenvalue of $\sigma$.

*Proof.* The proof is a simple modification of the proof reported in Appendix G of [3], so we will only present the essential steps here, referring to [3] for details.

First, since $[\rho, \sigma] = 0$ we can diagonalize both states in the same eigenbasis and write $\rho = \sum_i r_i |i\rangle\langle i|$ and $\sigma = \sum_i s_i |i\rangle\langle i|$. Then we write

$$\begin{aligned} E_{s_{\min}}(\rho||\sigma) &= \sum_i r_i F\left(\ln\left(s_{\min}\frac{r_i}{s_i}\right)\right) \\ &= \frac{1}{s_{\min}}\sum_i s_i \left(s_{\min}\frac{r_i}{s_i}\right) F\left(\ln\left(s_{\min}\frac{r_i}{s_i}\right)\right) \\ &\equiv \sum_i s_i g_{s_{\min}}\left(\frac{r_i}{s_i}\right), \end{aligned} \quad (3.62)$$

where

$$g_{s_{\min}}\left(\frac{r_i}{s_i}\right) \equiv \frac{1}{s_{\min}}\left(s_{\min}\frac{r_i}{s_i}\right) F\left(\ln\left(s_{\min}\frac{r_i}{s_i}\right)\right) \quad (3.63)$$

is a concave function in the interval $[\min_i \frac{r_i}{s_i}, \max_i \frac{r_i}{s_i}]$ since $xF(x)$ is concave in the unit interval, $s_{\min} > 0$ since $\sigma$ is full rank and $s_{\min}\frac{r_i}{s_i} \in [0, 1]$. On the other hand, since $(\rho, \sigma) \succ$

$(\rho', \sigma')$, there exists a quantum channel mapping $(\rho, \sigma)$ to $(\rho', \sigma')$, thus by Lemma 20 of [3] there exists a right stochastic matrix $T$ that maps the eigenvalue vectors (denoted with bold symbols) as $\boldsymbol{r}T = \boldsymbol{r}'$, $\boldsymbol{s}T = \boldsymbol{s}'$. Then, by Lemma 34 of [3], it follows that the inequality (3.61) holds. Note also that from the majorization $\sigma \succ \sigma'$, it follows that the respective smallest eigenvalues satisfy $s_{\min} \leqslant s'_{\min}$. ■

For an operation $\Phi$ with a fixed state $\sigma_*$ and $\rho \succ_{\sigma_*} \Phi(\rho)$, the inequality (3.61) implies that

$$R_*(\rho) \equiv -E(\rho||\sigma_*) \geqslant -E(\Phi(\rho)||\sigma_*) \equiv R_*(\Phi(\rho)), \quad (3.64)$$

thus we make contact with the definition of a resource monotone (1.1). As an application of the construction (3.60), we can use our monotones $M^{(n)}$ and $P_E^{(n)}$ to define relative quantifiers as follows:

$$\begin{aligned} M^{(n)}_{b_n,x}(\rho||\sigma) &= (-1)^n \mathrm{Tr}[\rho(\ln(x\rho\sigma^{-1}) - b_n)^n] \\ &= M^{(n)}(x\rho\sigma^{-1}; b_n) + b_n^n \\ &= (-1)^n \mathrm{Tr}[\rho(\ln\rho - \ln\sigma + \ln(x) - b_n)^n]. \end{aligned} \quad (3.65)$$

The first two quantifiers of the sequence reduce to (minus) the relative entropy and $M_x(\rho||\sigma)$ defined in (3.57) by

$$M^{(2)}_{1,x}(\rho||\sigma) = M_x(\rho||\sigma). \quad (3.66)$$

More generally, expanding (3.65) with relative cumulants and setting $a_n = b_n - \ln(x)$, we have

$$M^{(1)}_{b_1,x}(\rho||\sigma) = -S(\rho||\sigma) + a_1, \quad (3.67)$$

$$M^{(2)}_{b_2,x}(\rho||\sigma) = [-S(\rho||\sigma) + a_2]^2 + C(\rho||\sigma), \quad (3.68)$$

$$\begin{aligned} M^{(3)}_{b_3,x}(\rho||\sigma) &= [-S(\rho||\sigma) + a_3]^3 + 3C(\rho||\sigma) \\ &\quad \times [-S(\rho||\sigma) + a_3] - C_3(\rho||\sigma), \end{aligned} \quad (3.69)$$

$$\begin{aligned} M^{(4)}_{b_4,x}(\rho||\sigma) &= [-S(\rho||\sigma) + a_4]^4 + 6C(\rho||\sigma)[-S(\rho||\sigma) + a_4]^2 \\ &\quad + 3C(\rho||\sigma)^2 - 4C_3(\rho||\sigma)[-S(\rho||\sigma) + a_4] \\ &\quad + C_4(\rho||\sigma), \cdots, \end{aligned} \quad (3.70)$$

where $C_n(\rho||\sigma)$ is the $n$th cumulant of $\ln(\rho\sigma^{-1})$. Notice that, up to additive constants, the equations in (3.67) reduce to the ones in (3.10). This follows by choosing $\sigma = \mathbb{1}/d$ giving $S(\rho||\mathbb{1}/d) = -S(\rho) + \ln d$ and more generally $C_j(\rho||\mathbb{1}/d) = (-1)^j C_j(\rho)$, and setting $a_n = n - 1 - \ln s_{\min} = n - 1 + \ln d$.

When $\rho$ and $\sigma$ are commuting density matrices, we can also define a generating function

$$k_{\mathrm{rel}}(\alpha; a) = e^{-\alpha a} \exp[(\alpha - 1) S_\alpha(\rho||\sigma)], \quad (3.71)$$

where

$$S_\alpha(\rho||\sigma) = \frac{1}{\alpha - 1} \ln \mathrm{Tr}[\rho^\alpha \sigma^{1-\alpha}] \quad (3.72)$$

is the Petz-Rényi relative entropy [46]. In this case, we get

$$M^{(n)}_{b_n,x}(\rho||\sigma) = \left[(-1)^n e^a \frac{\partial^n}{\partial\alpha^n} k_{\mathrm{rel}}(\alpha; a)\right]\Bigg|_{\alpha=1, a=b_n-\ln(x)}. \quad (3.73)$$

It would be interesting to generalize the generating function giving (3.73) to the case of noncommuting $\rho$ and $\sigma$. We leave this investigation for a future analysis.

Likewise, from $P_E^{(n)}$ we obtain the relative quantities

$$\begin{aligned} P^{(n)}_{E,x}(\rho||\sigma) &= -\mathrm{Tr}[\rho F_n(\ln\rho - \ln\sigma + \ln(x))] \\ &= -P_E^{(n)}(x\rho\sigma^{-1}), \end{aligned} \quad (3.74)$$

where $F_n$ is an extremal polynomial that is a solution of $F_n''(y) + F_n'(y) = G_n(y)$, where $G_n(y)$ is of the form given in Theorem 1.

In Ref. [3] the monotonicity of $M_{s_{\min}}(\rho||\sigma)$ (see 3.57) has been used to derive a lower bound for relative entropy production. As corollary of the more general theorem (3.61), one can obtain infinitely many inequalities involving the change in relative entropy. If $(\rho, \sigma) \succ (\rho', \sigma')$ with both $\sigma, \sigma'$ full rank, then

$$M^{(n)}_{b_n,s_{\min}}(\rho'||\sigma') \geqslant M^{(n)}_{b_n,s_{\min}}(\rho||\sigma), \quad (3.75)$$

$$P^{(n)}_{E,s_{\min}}(\rho'||\sigma') \geqslant P^{(n)}_{E,s_{\min}}(\rho||\sigma), \quad (3.76)$$

for all $n \geqslant 1$. Substituting the explicit form of $M^{(n)}_{b_n,s_{\min}}$ from (3.65) leads to inequalities resembling the ones in Sec. III B, from which one can solve the relative entropy production with a bound involving relative cumulants to arbitrary order. For example, at order $n = 2$, concavity requires $b_2 \geqslant 1$, leading to [3]

$$\Delta S_{\mathrm{rel}}^2 + 2\chi \Delta S_{\mathrm{rel}} - \Delta C_{\mathrm{rel}} \geqslant 0 \quad (3.77)$$

where $\Delta S_{\mathrm{rel}} \equiv S(\rho||\sigma) - S(\rho'||\sigma')$, $\Delta C_{\mathrm{rel}} \equiv C(\rho||\sigma) - C(\rho'||\sigma')$ and $\chi = a - S(\rho||\sigma)$, with $a \equiv 1 - \ln s_{\min}$. The inequality (3.77) can be rewritten as the following relative entropy production bound:

$$\Delta S_{\mathrm{rel}} \geqslant \frac{\Delta C_{\mathrm{rel}}}{2a - S(\rho||\sigma) - S(\rho'||\sigma')}. \quad (3.78)$$

When $\sigma$ and $\sigma'$ are the maximally mixed state, (3.78) becomes (1.10). The bound in (3.78) can be relaxed obtaining the somewhat less tight inequality (see [3] for the explicit calculation)

$$\Delta S_{\mathrm{rel}} \geqslant \frac{\Delta C_{\mathrm{rel}}}{2\sqrt{M^{(2)}_{1,s_{\min}}(\rho||\sigma)}} \geqslant \frac{\Delta C_{\mathrm{rel}}}{2(1 - \ln s_{\min})}. \quad (3.79)$$

The inequality (3.79) is the same as that reported in [3] with $1 - \ln s_{\min}$ replacing $1/\ln 2 - \log_2(s_{\min})$ due to the use of ln instead of $\log_2$. Notice that the inequality (1.7) can be recovered from (3.79) when $\sigma$ and $\sigma'$ are the maximally mixed state.

Moving to the $P^{(n)}_{E,s_{\min}}$, tighter inequalities can be obtained at orders $n \geqslant 3$. For example, at order $n = 3$, writing

$$\begin{aligned} 0 &\leqslant P^{(3)}_{E,s_{\min}}(\rho'||\sigma') - P^{(3)}_{E,s_{\min}}(\rho||\sigma) \\ &\equiv w_2 a^2 + w_1 a + w_0, \end{aligned} \quad (3.80)$$

where (using $b_n = n - 1$)

$$\begin{aligned} w_2 &= M^{(2)}(s_{\min}\rho'\sigma'^{-1}; 1) - M^{(2)}(s_{\min}\rho\sigma^{-1}; 1) \\ &= M^{(2)}_{1,s_{\min}}(\rho'||\sigma') - M^{(2)}_{1,s_{\min}}(\rho||\sigma), \end{aligned} \quad (3.81)$$

$$w_1 = -M^{(1)}_{0,s_{\min}}(\rho'||\sigma') + M^{(1)}_{0,s_{\min}}(\rho||\sigma), \quad (3.82)$$

$$\begin{aligned} w_0 &= \tfrac{1}{3}\big[M^{(3)}_{2,s_{\min}}(\rho'||\sigma') - M^{(3)}_{2,s_{\min}}(\rho||\sigma)\big] \\ &\quad - \big[M^{(2)}_{1,s_{\min}}(\rho'||\sigma') - M^{(2)}_{1,s_{\min}}(\rho||\sigma)\big], \end{aligned} \quad (3.83)$$

we obtain

$$\Delta M_3^{\rm rel} \geqslant 3\Delta M_2^{\rm rel} + \frac{3}{4}\frac{\left(\Delta M_2^{\rm rel}\right)^2}{\Delta M_1^{\rm rel}}, \tag{3.84}$$

where now

$$\Delta M_n^{\rm rel} \equiv M_{n-1,s_{\min}}^{(n)}(\rho'||\sigma') - M_{n-1,s_{\min}}^{(n)}(\rho||\sigma). \tag{3.85}$$

#### 1. Finite-size correction to Clausius inequality

Rather than moving on to derive more explicit forms of these higher-order inequalities, let us point out an interesting corollary of (3.79). A well-known statement is that the nonnegativity of relative entropy $S(\rho||\sigma) \geqslant 0$ can be rewritten as an inequality

$$\Delta\langle K\rangle \equiv {\rm Tr}(\rho K) - {\rm Tr}(\sigma K) \geqslant S(\rho) - S(\sigma) \equiv \Delta S, \tag{3.86}$$

where [47] $K = -\ln\sigma$. The refined equation (3.79) gives a finite-size correction to the inequality (3.86), assuming $[\rho,\sigma] = 0$. Set $\rho' = \sigma = \sigma'$ in (3.79), which corresponds to consider a mutually commuting pair of states with the $\sigma$-majorization $\rho \succ_\sigma \sigma$. Now (3.79) reduces to

$$S(\rho||\sigma) \geqslant \frac{C(\rho||\sigma)}{2\sqrt{M_{1,s_{\min}}^{(2)}(\rho||\sigma)}} \geqslant \frac{C(\rho||\sigma)}{2(1-\ln s_{\min})}, \tag{3.87}$$

which in turn implies a finite-size correction to (3.86); indeed

$${\rm Tr}(\rho K) - {\rm Tr}(\sigma K) \geqslant S(\rho) - S(\sigma) + \frac{C(\rho||\sigma)}{2\sqrt{M_{1,s_{\min}}^{(2)}(\rho||\sigma)}}. \tag{3.88}$$

In the context of nonequilibrium thermodynamics, we can choose $\sigma = \gamma_\beta \equiv e^{-\beta H}/Z(\beta)$, where $\beta = 1/(k_B T)$ is the inverse temperature, and $\rho$ is a state that thermomajorizes $\gamma_\beta$ and commute with it. From (3.88), we obtain a finite-size correction to a fundamental thermodynamic relation (the Clausius inequality),

$$\begin{aligned} S(\gamma_\beta) - S(\rho) &\geqslant \frac{1}{k_B T}[\langle H\rangle_{\gamma_\beta} - \langle H\rangle_\rho] + \frac{C(\rho||\gamma_\beta)}{2\sqrt{M_{1,s_{\min}}^{(2)}(\rho||\gamma_\beta)}} \\ &\geqslant \frac{1}{k_B T}[\langle H\rangle_{\gamma_\beta} - \langle H\rangle_\rho] \\ &\quad + \frac{C(\rho||\gamma_\beta)}{2+2\beta(E_{\max} - F(\beta))}, \end{aligned} \tag{3.89}$$

where $\langle H\rangle_\rho$ is the nonequilibrium energy of the state $\rho$, $\langle H\rangle_{\gamma_\beta}$ the energy in the thermal state, $S(\rho)$ the nonequilibrium entropy in the state $\rho$, $S(\gamma_\beta)$ the thermal entropy [or, more conventionally, $S(\beta) = k_B S(\gamma_\beta)$], and where $s_{\min}$ is the smallest eigenvalue $\exp(-\beta E_{\max})/Z(\beta)$ of the Gaussian state $\gamma_\beta$ (the finite-dimensional system has a maximum energy eigenvalue $E_{\max}$) and $\beta F(\beta) = -\ln Z(\beta)$ is the Helmholtz free energy. Again, one could attempt to derive a sequence of more refined inequalities from the infinite sequences of resource monotones above constructed. We find it worth stressing that, considering (3.89) when $\gamma_\beta$ is maximally mixed, i.e., when $\beta \to 0$, the inequality becomes

$$\ln d \geqslant S(\rho) + \frac{C(\rho)}{2(1+\ln d)}. \tag{3.90}$$

It is straightforward to check that (1.10), when $\rho$ is maximally mixed, gives an inequality tighter than (3.90).

Let us remark that the finite-size correction obtained above has no effect on the so-called first law of entanglement. Consider a one-parameter family of states $\rho(\lambda)$, with $\rho(0) = \sigma$. Expanding $\rho(\lambda) = \sigma + \lambda\delta\rho + \cdots$, one obtains

$$\begin{aligned} \delta S(\rho||\sigma) \equiv S(\rho(\lambda)||\sigma) - 0 = 0 + \lambda S^{(1)}(\sigma||\sigma) \\ + \lambda^2 S^{(2)}(\sigma||\sigma) + \cdots \geqslant 0. \end{aligned} \tag{3.91}$$

Since $S(\rho||\sigma)$ has a global minimum at $\rho = \sigma$, at first order we have an equality $\delta S(\rho||\sigma) = \lambda S^{(1)}(\sigma||\sigma) = 0$ instead of an inequality, hence (3.86) becomes the equality

$$\delta\langle K\rangle = \delta S, \tag{3.92}$$

for infinitesimal changes, which is known as the first law of entanglement. The reason why it does not receive finite-size corrections from (3.88) is that $\rho = \sigma$ is a local minimum of $C(\rho||\sigma)$, so $\delta C(\rho||\sigma) = 0$ also to first order, so that the first-order equality $\delta S(\rho||\sigma) = 0$ remains intact.

## IV. APPLICATIONS

### A. Information erasure

As an application of the inequalities (3.24) and (3.40), we study a state erasure process, following [3]. The erasure process is described in two steps. Consider first a system in a state $\rho$ to be "erased", i.e., to be converted to some fixed pure state $|\psi\rangle$. This conversion can be accomplished introducing an external system $\mathcal{B}$ involving $n$ qubits, an *information battery*, which is simultaneously converted from a pure state $|0\rangle^{\otimes n}$ to a maximally mixed flat state $(\mathbb{1}/2)^{\otimes n}$. This first step of the process requires the battery to be large enough, meaning that there is a lower bound on the required $n$. Mathematically, this step can be modeled by a unital channel $\mathcal{E}$ mapping the input state $\Omega$ of the composite system to the output state $\Upsilon = \mathcal{E}(\Omega)$. According to the Uhlmann's theorem discussed in Sec. II, the unital channel implies the majorization relation

$$\Omega \equiv \rho \otimes |0\rangle\langle 0|^{\otimes n} \succ |\psi\rangle\langle\psi| \otimes (\mathbb{1}/2)^{\otimes n} \equiv \Upsilon. \tag{4.1}$$

From the Schur concavity of von Neumann entropy it follows that $S(\Omega) \leqslant S(\Upsilon)$, which by additivity of $S$ produces the lower bound

$$n\ln 2 \geqslant S(\rho). \tag{4.2}$$

In the second step, in order to be able to repeat the procedure, one must restore the battery back to the initial state $|0\rangle^{\otimes n}$. This step requires work to be done on the battery [48], and can be performed in different ways (see e.g., [49] for more discussion and references). We consider here the process described in [50], where the battery consists of $n$ identical copies of a tunable two-state system and the energy gap of the two energy eigenstates can be parametrically adjusted to be anything from zero towards infinity. One places the battery in contact with a heat bath at temperature $T$, with the energy gap initially being zero, and then adiabatically increases the gap towards infinity. This requires work $W$ to be performed on the battery. At the end, each two-state system is in the ground state $|0\rangle$ with probability 1, so that the battery is restored to the state $|0\rangle^{\otimes n}$, at the expense of the work cost involved in the adiabatic process, which can be calculated to be

$$W = k_B T n \ln 2\,. \tag{4.3}$$

One then removes the battery from the contact with the heat bath, after which the energy gap can be reduced back to zero while doing no work. The battery is now ready to be used again.

The two equations (4.2) and (4.3) can be combined as the Landauer inequality

$$W \geqslant k_B T S(\rho), \tag{4.4}$$

for the work cost of erasure. Note that erasure deletes all details of the eigenvalue spectrum of $\rho$, so it is natural to ask if additional details of the distribution of the spectrum beyond just the entropy will have an effect on the work cost of erasure. Indeed we find that all cumulants of the spectrum can be used to characterize the cost.

The inequality (4.2) has been derived from $S$, but now we may consider the infinite sequence inequalities following from the monotones $M^{(m)}$. Since $M^{(m)}$ are (Schur) concave, we have that $\Omega \succ \Upsilon \Rightarrow M^{(m)}(\Upsilon, b_m) \geqslant M^{(m)}(\Omega, b_m)$. First we compute

$$M^{(m)}(\Upsilon, b_m) = \sum_{k=1}^{m} \binom{m}{k} b_m^{m-k} \mu_k(\Upsilon),$$
$$\mu_k(\Upsilon) = \sum_{p_j}{}' k! \prod_j \frac{1}{p_j!(j!)^{p_j}} C_j^{p_j}(\Upsilon), \tag{4.5}$$

where $\sum'$ is the restricted sum introduced in (3.8) and, by additivity of cumulants (3.12) and the fact that they are zero for pure states, we have

$$C_j(\Upsilon) = 0 + n C_j(\mathbb{1}/2). \tag{4.6}$$

Since $C_j(\mathbb{1}/2) = 0$ for $j \geqslant 2$, in the above sum only the partition $k = 1 + 1 + \cdots + 1$ remains, hence

$$\begin{aligned} M^{(m)}(\Upsilon, b_m) &= \sum_{k=0}^{m} \binom{m}{k} b_m^{m-k} (n \ln 2)^k \\ &= (n \ln 2 + b_m)^m. \end{aligned} \tag{4.7}$$

Next, we have

$$M^{(m)}(\Omega, b_m) = \sum_{k=0}^{m} \binom{m}{k} b_m^{m-k} \sum_{p_j}{}' k! \prod_j \frac{1}{p_j!(j!)^{p_j}}\, C_j^{p_j}(\Omega), \tag{4.8}$$

where [using again the additivity (3.12)]

$$C_j(\Omega) = C_j(\rho) + 0. \tag{4.9}$$

We thus have confirmed that

$$M^{(m)}(\Omega, b_m) = M^{(m)}(\rho, b_m). \tag{4.10}$$

Now we employ the cumulant expansion (3.9). Setting for simplicity $b_m$ to be the smallest possible value $b_m = m - 1$ allowed by concavity, we obtain the sequence of inequalities for the work cost $W = k_B T n \ln 2$ with [51]

$$n \ln 2 \geqslant S(\rho), \tag{4.11}$$
$$(n \ln 2 + 1)^2 \geqslant [S(\rho) + 1]^2 + C(\rho), \tag{4.12}$$
$$(n \ln 2 + 2)^3 \geqslant [S(\rho) + 2]^3 + 3C(\rho)[S(\rho) + 2] + C_3(\rho), \tag{4.13}$$
$$\begin{aligned}(n \ln 2 + 3)^4 \geqslant{}& [S(\rho) + 3]^4 + 6C(\rho)\,[S(\rho) + 3]^2 + 3C(\rho)^2 \\ &+ 4C_3(\rho)\,[S(\rho) + 3] + C_4(\rho), \cdots . \end{aligned}\tag{4.14}$$

Let us remark on the relative significance of the various terms on the right-hand side of these inequalities. We point out that, while $S_{\max} = \ln d$ by the maximally mixed state $\rho_{\text{flat}} = \frac{1}{d}\mathbb{1}$, we have $C_{\max} \approx \frac{1}{4}\ln^2(d-1)$ [3,21,24] by a different state $\rho_{\max} = \text{diag}(1 - r, \frac{r}{d-1}, \ldots, \frac{r}{d-1})$, where $r$ is the solution of $(1 - 2r)\ln[(1 - r)(d - 1)/r] = 2$. One might anticipate that the higher cumulants have maximum values $C_{n,\max} \propto \ln^n d$ each reached at a different state $\rho_{n,\max}$, but this remains (to our knowledge) an open problem.

Let us now compare these inequalities with those ones obtained from the extremal polynomial monotones. The first two are the same as in the above: $P_E^{(1)} = -S = -M^{(1)}$ gives the Landauer bound (4.11), while the inequality from $P_E^{(2)} = -M^{(2)}$ will be the same as (4.12). Let us verify the latter inequality starting from (3.24), which first gives $n \ln 2 \geqslant S(\rho) + \frac{C(\rho)}{S(\rho) + n\ln 2 + 2}$, or

$$n \ln 2 + 1 \geqslant \sqrt{(S(\rho) + 1)^2 + C(\rho)}, \tag{4.15}$$

which coincides with (4.12). This is already a stronger bound than the one coming from the inequality (1.7) involving $S$, $C$ and $M$, which reads

$$n \ln 2 \geqslant S(\rho) + \frac{C(\rho)}{2\sqrt{M(\rho)}}. \tag{4.16}$$

Consider then the tighter third-order inequality (3.40) applied to the erasure process. We need the moments of modular Hamiltonian $\mu_n$ [see (3.7)] for the states $\Omega$ and $\Upsilon$ in terms of the cumulants. From (3.8), for $\Omega = \rho \otimes |0\rangle\langle 0|^{\otimes n}$ we have

$$\mu_1(\Omega) = S(\rho), \tag{4.17}$$
$$\mu_2(\Omega) = S^2(\rho) + C(\rho), \tag{4.18}$$
$$\begin{aligned}\mu_3(\Omega) &= S^3(\Omega) + 3S(\Omega)C(\Omega) + C_3(\Omega) \\ &= S^3(\rho) + 3S(\rho)C(\rho) + C_3(\rho), \end{aligned}\tag{4.19}$$

and for $\Upsilon = |\psi\rangle\langle\psi| \otimes (\mathbb{1}/2)^{\otimes n}$ we obtain $\mu_k(\Upsilon) = (n \ln 2)^k$. After some calculation, we find

$$\begin{aligned}&(n \ln 2 + 1)^3 - [S(\rho) + 1]^3 + 3[n \ln 2 - S(\rho)] \\ &-[C_3(\rho) + 3C(\rho)(S(\rho) + 1)] \\ &\geqslant \frac{3}{4}\frac{[(n \ln 2 + 1)^2 - (S(\rho) + 1)^2 - C(\rho)]^2}{(n \ln 2 - S(\rho))}. \end{aligned}\tag{4.20}$$

From (3.39), we already know that this is a stronger inequality than (4.13). A weaker form of the inequality (3.39) is

$\Delta M_3 \geqslant 3\Delta M_2$ and it gives

$$(n\ln 2+1)^3+3n\ln 2 \geqslant (S(\rho)+1)^3+3S(\rho) + 3C(\rho)(S(\rho)+1)+C_3(\rho), \tag{4.21}$$

which is still somewhat stronger than (4.13).

Moving to higher-order extremal polynomial monotones, one can derive an infinite sequence of (increasingly more complicated) inequalities, involving higher cumulants of modular Hamiltonian. As already noted, it is possible that the sequence contains some interesting hierarchical structure, e.g., the right-hand side of (4.20) contains a ratio of quantities that appears in the two previous inequalities (4.12) and (4.11).

In the Outlook section of [3], it was commented that it would be interesting to construct a hierarchy of (Schur-) concave functions from cumulants of modular Hamiltonian relevant to single-shot settings. We have proposed two such sequences: The moments of shifted modular Hamiltonian and the extremal polynomial monotones. The latter one could be relevant to provide tight inequalities. It would be interesting to further explore the hierarchical structures contained in the two sequences. Furthermore, besides a "sequence of Landauer inequalities", the majorization relation in the information erasure already contains many inequalities for the partial sums of the two eigenvalue spectra. Since any concave quantifier produces an inequality for the information erasure, and thus some kind of a bound on $n$, there is obviously an uncountable infinity of inequalities that one could consider. The above sequences are special only in the sense that they contain the entropy and higher cumulants.

Majorization gives a partial ordering among density matrices, but one can also define another partial ordering based on the sequences of inequalities from the monotones. It would be interesting to study the relation between these two orderings in more detail and some tentative observations are made in Sec. VI.

Finally, while we discussed an idealized model of state erasure to demonstrate the implications of the monotones, the obvious next step would be to investigate a more realistic scenario. For instance, the initial state $\rho$ of the system and the initial state of the battery could be known only up to some uncertainties $\epsilon_1$ and $\epsilon_2$. The erasure process could be imperfect, modeled by a quantum operation $\mathcal{E}$ involving an additional uncertainty $\delta$. The overall result is an output state approximating $|\psi\rangle\langle\psi|\otimes(\mathbb{1}/2)^{\otimes n}$ only up to a cumulative inaccuracy. In particular, some of the battery qubits could be in an exactly known state, some in a biased state, and some in a completely unknown state, making the work estimation resembling the setting of [52]. One aspect involved is computing cumulants of the probability spectrum from a smoothed distribution. Related questions are whether the monotones $M^{(n)}$ introduced in this paper have interesting physical upper and lower bounds, generalizing the min- and max-entropies, and what happens if a catalytic system is incorporated into the state erasure model. We leave these questions for future study.

### B. Lower bound on marginal entropy production

The inequality (3.24) offers a slight tightening of the lower bound on marginal entropy production derived in [3]. Our discussion follows again [3] but with some small modifications. Consider a quantum channel $\mathcal{E}$ mapping a system S to itself. In order to represent the channel $\mathcal{E}$ acting on a state $\rho_S$ of the system, we introduce an environment E in a state $\rho_E$ and a *unital* channel [53] $\mathcal{U}$ to write

$$\mathcal{E}(\rho_S)=\mathrm{Tr}_E[\mathcal{U}(\rho_S\otimes\rho_E)]. \tag{4.22}$$

We denote by

$$\rho'_{SE}=\mathcal{U}(\rho_S\otimes\rho_E) \tag{4.23}$$

the state of the joint system $S\cup E$ after the application of the unital channel. Since $\mathcal{U}$ is unital, by the Uhlmann's theorem mentioned in Sec. II, we have the majorization $\rho_E\otimes\rho_S\succ\rho'_{SE}$ and therefore, using (3.24), we obtain

$$(S(\rho'_{SE})+1)^2-(S(\rho_S\otimes\rho_E)+1)^2 \geqslant C(\rho_S\otimes\rho_E)-C(\rho'_{SE}), \tag{4.24}$$

which means that the entropy production in the joint system is bounded from below by the decrease of the variance. The right-hand side can be manipulated using a result (Lemma 11) from [3], which takes the form of a correction to the subadditivity of variance. To describe that result we consider a bipartite system $S\cup E$ of dimension $d=d_Sd_E\geqslant 2$, where $\rho_{SE}$ and $\sigma_{SE}$ are two commuting quantum states. If $\sigma_{SE}=\sigma_S\otimes\sigma_E$ and is full rank, then [54]

$$C(\rho_{SE}||\sigma_{SE})\leqslant C(\rho_S||\sigma_S)+C(\rho_E||\sigma_E)+\kappa\, f(I_{SE}/\ln 2), \tag{4.25}$$

where $\rho_S=\mathrm{Tr}_E\,\rho_{SE},\ \rho_E=\mathrm{Tr}_S\,\rho_{SE}$, $\kappa$ is a constant given by

$$\kappa=\sqrt{2\ln 2}\,(12\ln^2 2+\ln^2 s_{\min}+8\ln^2 d), \tag{4.26}$$

where $s_{\min}$ is the smallest eigenvalue of $\sigma$, $I_{SE}$ is the mutual information

$$I_{SE}=S(\rho_{SE}||\rho_S\otimes\rho_E) \tag{4.27}$$

of $\rho_{SE}$ with respect to the bipartition and $f(x)=\max\{x^{1/4},x^{1/2}\}$. As a special case, choosing $\sigma_{SE}=(\mathbb{1}_S/d_S)\otimes(\mathbb{1}_E/d_E)$ with $s_{\min}=\frac{1}{d_Sd_E}=\frac{1}{d}$, we obtain the following correction to the subadditivity of variance:

$$C(\rho_{SE})\leqslant C(\rho_S)+C(\rho_E)+\kappa f(I_{SE}/\ln 2), \tag{4.28}$$

where now

$$\kappa=\sqrt{2\ln 2}\,(12\ln^2 2+9\ln^2 d). \tag{4.29}$$

We can now turn back to (4.24) and bound the right-hand side that equation using the additivity of the variance (3.12), $C(\rho_S\otimes\rho_E)=C(\rho_S)+C(\rho_E)$ and the correction to the subadditivity (4.33) to decompose $C(\rho'_{SE})$, ending up with

$$C(\rho_S\otimes\rho_E)-C(\rho'_{SE})\geqslant-\Delta C_S-\Delta C_E-\kappa\, f(I_{SE}/\ln 2), \tag{4.30}$$

where we have defined $\Delta C_S=C(\rho'_S)-C(\rho_S)$ and $\Delta C_E=C(\rho'_E)-C(\rho_E)$. Note that, for decreasing variance, $-\Delta C\geqslant 0$. In the left-hand side of (4.24) we can use the additivity and subadditivity of entropy to write

$$(S_{S,E}+\Delta S_{S,E}+1)^2-(S_{S,E}+1)^2 \geqslant (S(\rho'_{SE})+1)^2-(S(\rho_S\otimes\rho_E)+1)^2, \tag{4.31}$$

where $S_{\mathrm{S,E}} \equiv S(\rho_{\mathrm{S}}) + S(\rho_{\mathrm{E}})$ and $\Delta S_{\mathrm{S,E}} \equiv \Delta S_{\mathrm{S}} + \Delta S_{\mathrm{E}} \equiv S(\rho'_{\mathrm{S}}) - S(\rho_{\mathrm{S}}) + S(\rho'_{\mathrm{E}}) - S(\rho_{\mathrm{E}})$. Combining everything together, and expressing the resulting inequality for the entropy production, we have

$$\Delta S_{\mathrm{S}} + \Delta S_{\mathrm{E}} \geqslant (S(\rho_{\mathrm{S}}) + S(\rho_{\mathrm{E}}) + 1) \times \left\{\left[1 + \frac{-\Delta C_{\mathrm{S}} - \Delta C_{\mathrm{E}} - \kappa\, f(I_{\mathrm{SE}}/\ln 2)}{(S(\rho_{\mathrm{S}}) + S(\rho_{\mathrm{E}}) + 1)^2}\right]^{1/2} - 1\right\}, \tag{4.32}$$

which is a slightly tighter lower bound for marginal entropy production, compared to Result 4 in [3]. We emphasize that the result is nontrivial only when the numerator in the ratio in the square brackets is positive (the total decrease in variances is sufficiently large). If the ratio in the square bracket in the right-hand side of (4.32) is much smaller than one, we can employ the Taylor expansion, obtaining from the leading approximation

$$\Delta S_{\mathrm{S}} + \Delta S_{\mathrm{E}} \gtrsim \frac{-\Delta C_{\mathrm{S}} - \Delta C_{\mathrm{E}} - \kappa\, f(I_{\mathrm{SE}}/\ln 2)}{2(S(\rho_{\mathrm{S}}) + S(\rho_{\mathrm{E}}) + 1)}, \tag{4.33}$$

which can be easily compared against the inequality in [3] (see also [55]). As an application, [3] considered state transitions with a help of a catalytic system C and derived the lower bound $d_{\mathrm{C}} \geqslant O[\exp(\delta^{-1/8})]$ for the necessary dimension $d_{\mathrm{C}}$ of C for a state transition, where the variation of entropy $\delta$ is small while variance is reduced. Similarly, it would be interesting to consider what are the implications of the other inequalities in the sequences considered in this paper, which involve higher cumulants. For that, one would need to first derive subadditivity properties of higher cumulants $C_n$, generalizing the result (4.33) for the variance.

## V. PAIRS OF BIPARTITE PURE STATES IN 1 + 1 CFT AND PERIODIC CHAINS

Majorization is a central concept in finite-dimensional quantum systems, which provides a classification of bipartite entanglement for pure states. What can be said about pure states in quantum field theories? In this section we develop a strategy to gain insights into this question. To provide explicit examples, we first evaluate $S_A$ and $C_A$ in one-dimensional and translation invariant systems. We review known results for Rényi entropies for a ground state and a class of excited states. From the Rényi entropies we first compute the difference $S_A - C_A$ and study how it changes moving from the ground state to an excited state. We then compute monotones and inequalities, as outlined in Sec. V A, and find the constraints they present for the majorization order between a ground state and an excited state.

### A. Majorization and states of periodic chains

Consider a pair of pure states $|\Phi\rangle$ and $|\Psi\rangle$ in a QFT, is there any meaningful definition for a relation $|\Phi\rangle \succ |\Psi\rangle$? First of all, bipartitioning involves an ultraviolet cutoff, making some entanglement monotones divergent. Also, the definition of majorization via the reduced states and partial sums of ordered eigenvalues becomes cumbersome and very sensitive to the choice of UV regulator. In Sec. V we will explore this question to gain some tentative insight. We consider a pair of states in a CFT, which is a continuum limit of a discrete model at criticality. As a concrete example, we will consider discrete versions of the compact boson at critical radius and free fermion CFT on a circle, which are different critical limits of the periodic anisotropic XY spin chain. The latter can be mapped to free periodic fermionic chains with a finite-dimensional Hilbert space. The original pair of CFT states is then mapped to their counterpart in the fermionic chain. In that case majorization becomes well defined, and we can apply rigorous theorems from quantum information theory and compare bipartite entanglement between the two states. The results will then reflect some properties of the original CFT states, which have to be interpreted with care, as the maps between the spin chains and fermionic chains are typically nonlocal.

Our goal is to take some first steps, leaving more exhaustive studies for future work. We will limit to comparing a ground state $|\Phi\rangle \equiv |0\rangle$ with an excited state $|\Psi\rangle$ and their counterparts in the periodic fermionic chain. We bipartite the periodic chain of finite length $L$ made by $N$ sites (hence the Hilbert space of the full system has dimension $2^N$) into a subregion made by $\ell$ consecutive sites and its complement. Let us denote this subsystem by $A$ and its complement as $B$. We denote the reduced states in the subsystem $A$ by

$$\rho_A \equiv \mathrm{Tr}_B(|\Phi\rangle\langle\Phi|), \qquad \sigma_A \equiv \mathrm{Tr}_B(|\Psi\rangle\langle\Psi|). \tag{5.1}$$

Suppose that $|\Phi\rangle \succ |\Psi\rangle$ or equivalently $\rho_A \succ \sigma_A$. Note that this relation depends on the choice of the bipartition to $A$ and $B$, hence on the relative size of the two lengths $\ell/L$. While it is rather laborious to establish majorization, we ask an easier question of whether it can be ruled out. For this purpose we can use any Schur concave quantity $E_A$ applied to the reduced states: Assuming that $\rho_A \succ \sigma_A$ ($|\Phi\rangle \succ |\Psi\rangle$) then $\Delta E_A \equiv E_A(\sigma_A) - E_A(\rho_A) \geqslant 0$, thus if we instead find $\Delta E_A \leqslant 0$ the assumed majorization order cannot hold. Consequently any process converting one state into the other implying the assumed order is impossible (for example, a LOCC process $|\Psi\rangle \underset{\mathrm{LOCC}}{\longrightarrow} |\Phi\rangle$). The converse analysis naturally can be applied to the opposite assumption $\sigma_A \succ \rho_A$ and opposite processes. In this setting it becomes interesting to compare which quantity $E_A$ gives the most stringent bound.

For the Schur concave quantities $E_A$ to study, we only consider here the entanglement entropy $S_A$, the Rényi entropies $S_A^{(2)}$, $S_A^{(3)}$ and the monotone $M_A^{(2)}(\cdot, 1) = -P_A^{(2)}(\cdot)$. The latter gives the inequality (3.24), which must be satisfied if majorization holds. All in all, in studying the various monotones, we will compare which one gives the tightest bound for the range of $\ell/L$ where the majorization order must be violated. These results are discussed in detail in Sec. V F. Before such comparisons, we will examine the cutoff dependence of $S_A$, $C_A$ and $M_A^{(2)}$. We will also compare the entropy $S_A$ with the capacity $C_A$ for the two states, and examine how they depend on the relative size $\ell/L$.

### B. Some known CFT results

In this section we study $S$, $C$, and $M$ [see (1.4) and (1.6)] in simple bosonic, fermionic conformal field theories and related discrete models (which can be mapped to fermionic chains). Considering a 2D CFT for certain states (including the ground state) and when the subsystem $A$ is a single interval of length $\ell$, it has been found that [56–59]

$$\mathrm{Tr}\rho_A^n = c_n\, e^{-\frac{c}{12}(n-\frac{1}{n})W_A} + \dots, \tag{5.2}$$

where $c$ is the central charge of the CFT and $W_A$ is a function of $\ell$, which diverges as the UV cutoff $\epsilon \to 0$ and depend also on the state and on the geometry of the entire system. The constant $c_n$ is model and boundary condition dependent and $c_1 = 1$ (because of the normalisation of $\rho_A$). For instance, when the system is on the infinite line and in its ground state, when the system is on the circle of length $L$ and in its ground state or when the system is on the infinite line and at finite temperature $1/\beta$, for $W_A$ we have respectively

$$W_A = 2\ln\left(\frac{\ell}{\epsilon}\right), \quad W_A = 2\ln\left(\frac{L}{\pi\epsilon}\sin\frac{\pi\ell}{L}\right),$$
$$W_A = 2\ln\left(\frac{\beta}{\pi\epsilon}\sinh\frac{\pi\ell}{\beta}\right). \tag{5.3}$$

By employing (5.2) into the definitions (1.3) and (1.4), it is straightforward to find that [21]

$$C_A = S_A = \frac{c}{6}W_A + O(1), \tag{5.4}$$

and that $C_A$ and $S_A$ differ at the subleading order $O(1)$ determined by the nonuniversal constant $c_n$ [60].

In the following we report the expression of $M^{(n)}(\rho; b_n)$ defined in (3.6) for a CFT on the line in its ground state and an interval $A$ of length $\ell$. By using (5.2) and (3.17), for the leading term we find

$$M_A^{(n)}(b_n) = \left(\frac{\ln(\ell/\epsilon)}{3}\right)^n + O((\ln(\ell/\epsilon))^{n-1}), \tag{5.5}$$

where the subleading terms in $\ell/\epsilon$ depend both on the nonuniversal constants and on the parameter $b_n$. For instance, in the special case of $n = 2$ we get

$$M_A^{(2)}(b_2) = \left(\frac{\ln(\ell/\epsilon)}{3}\right)^2 + \frac{1}{3}(1 + 2b_2 - 2c_1')\ln(\ell/\epsilon) + O(1), \tag{5.6}$$

where the subleading terms that we have neglected are finite as $\epsilon$ vanishes. In the following, with a slight abuse of notation, we denote by $\ell$ both the number of consecutive sites in a block $A$ and the length of the corresponding interval $A$ in the continuum. This convention is adopted also for the number of sites of a finite chain and for the finite size of the corresponding system in the continuum limit, both denoted by $L$.

### C. Excited states in CFT

Consider a CFT in a circle of length $L$ in the excited state of the form $|\mathrm{ex}\rangle = \mathcal{O}(0,0)|\mathrm{gs}\rangle$, obtained by applying the operator $\mathcal{O}$ on the ground state. The subsystem is an interval $A$ of length $\ell < L$ in a circle of length $L > \ell$.

The Rényi entropies in the low-lying excited states in CFT have been studied in [61,62], finding that the following ratio provides a UV finite scaling function:

$$F_{\mathcal{O}}^{(n)}(\ell/L) \equiv \frac{\mathrm{Tr}(\rho_{\mathcal{O},A}^n)}{\mathrm{Tr}(\rho_{\mathrm{gs},A}^n)} = e^{(1-n)(S_{\mathcal{O},A}^{(n)} - S_{\mathrm{gs},A}^{(n)})}, \tag{5.7}$$

where $S_{\mathcal{O},A}^{(n)}$ and $S_{\mathrm{gs},A}^{(n)}$ denote the Rényi entropies when the system is either in the excited state or in the ground state respectively. The moments $\mathrm{Tr}(\rho_{\mathrm{gs},A}^n)$ in (5.7) are (5.2) with $W_A$ given by the second expression in (5.3), while the Rényi entropies for the excited state read

$$S_{\mathcal{O},A}^{(n)} = \frac{c}{6}\left(1+\frac{1}{n}\right)\ln\left(\frac{L}{\pi\epsilon}\sin\frac{\pi\ell}{L}\right)$$
$$+\frac{1}{1-n}\ln\left[F_{\mathcal{O}}^{(n)}(\ell/L)\right] + \ln c_n. \tag{5.8}$$

In [61,62] it has been found that the ratio (5.7) is obtained from a proper $2n$-point correlator of $\mathcal{O}$. This gives

$$S_{\mathcal{O},A} = S_{\mathrm{gs},A} - \frac{d}{dn}\left(\ln F_{\mathcal{O}}^{(n)}(\ell/L)\right)\Big|_{n=1},$$
$$C_{\mathcal{O},A} = C_{\mathrm{gs},A} + \frac{d^2}{dn^2}\left(\ln F_{\mathcal{O}}^{(n)}(\ell/L)\right)\Big|_{n=1}. \tag{5.9}$$

In the following we explicitly consider only two examples of excited states where [62–64]

$$F_{\mathcal{O}}^{(n)}(\ell/L) = [f_n(\ell/L)]^\gamma, \tag{5.10}$$

where only the exponent $\gamma$ distinguishes the two states and

$$f_n(\ell/L) \equiv \left(\frac{2}{n}\sin(\pi\ell/L)\right)^{2n}$$
$$\times\left(\frac{\Gamma\left(\frac{1}{2}[1+n+n\csc(\pi\ell/L)]\right)}{\Gamma\left(\frac{1}{2}[1-n+n\csc(\pi\ell/L)]\right)}\right)^2. \tag{5.11}$$

From (5.9), (5.10), and (5.4), one obtains the following UV finite combination:

$$\begin{aligned} S_{\mathcal{O},A} - C_{\mathcal{O},A} &= -\gamma\Big(\partial_n[\ln f_n(\ell/L)]|_{n=1} + \partial_n^2[\ln f_n(\ell/L)]|_{n=1}\Big) - c_1' - \left[\partial_n^2(\ln c_n)\right]\Big|_{n=1} \\ &= -2\gamma\left(\ln|2\sin(\pi\ell/L)| + \psi\left(\frac{1}{2\sin(\pi\ell/L)}\right) + \sin(\pi\ell/L)\right) - c_1' \\ &\quad -2\gamma\left(-1 + \frac{1}{\sin(\pi\ell/L)} + \psi'\left(\frac{1}{2\sin(\pi\ell/L)}\right) - [1+\sin(\pi\ell/L)]^2\right) - \left[\partial_n^2(\ln c_n)\right]\Big|_{n=1}, \end{aligned} \tag{5.12}$$

where $\psi(x)$ is the digamma function and $\psi'(x)$ its derivative. When $\ell/L \ll 1$, since $\psi(1/x) \simeq -\ln x$ and $\psi'(1/x) \simeq x$ as $x \to 0$, we have that $\frac{c}{3}\ln(\ell/\epsilon)$ is the leading term of both $S_{\mathcal{O},A}$ and $C_{\mathcal{O},A}$. Furthermore, the combination in (5.12) becomes $-[\partial_n^2(\ln c_n)]|_{n=1} - c_1'$ in this limit. The difference between the excited state and the ground state becomes invisible in the short interval limit. We remark also that the combinations $S_{\mathcal{O},A} - S_{\text{gs},A}$ and $C_{\mathcal{O},A} - C_{\text{gs},A}$ [see (5.9)] are UV finite.

### D. Massless compact boson

Our first example CFT is the massless compactified scalar field, whose central charge is $c = 1$. Its action is

$$I = \frac{g}{4\pi}\int d^2x\, \partial_\mu\phi\, \partial^\mu\phi, \tag{5.13}$$

with a field compactification radius $R$ such that $\phi \sim \phi + 2\pi jR$, $j \in \mathbb{Z}$. Interestingly, for this model it has been found in [62] that, when the excited state is given by a vertex operator $\mathcal{O} = :e^{\mathrm{i}\alpha\phi + \mathrm{i}\bar{\alpha}\bar{\phi}}:$, the scaling function (5.7) is equal to one identically. Therefore this excited state has the curious property that its bipartite entanglement structure is unchanged from the ground state. We will thus move to consider other excited states.

A nontrivial result for (5.7) is found when $\mathcal{O} = \mathrm{i}\partial\phi$ is the current. In this case (5.10) holds with $\gamma = 1$ [62–64]; hence $F^{(n)}_{\mathrm{i}\partial\phi} = f_n$. Although $F^{(n)}_{\mathrm{i}\partial\phi}$ is independent of $R$, for a numerical check we consider a specific value of the compactification radius in order to give an explicit value to the nonuniversal constants in (5.12). At the self-dual point, namely when $gR^2 = 1$, the massless compact boson can be studied as the continuum of a free fermion on the lattice described by the Hamiltonian

$$\widehat{H} = -\sum_{j=-\infty}^{\infty} (\hat{c}^\dagger_{j+1}\hat{c}_j + \hat{c}^\dagger_j\hat{c}_{j+1}), \tag{5.14}$$

where the fermionic operators $\hat{c}_j$ satisfy the anticommutation relations $\{\hat{c}_j, \hat{c}^\dagger_k\} = \delta_{jk}$. Indeed, the continuum limit of this free fermionic chain is the massless Dirac field theory, which in the low-energy regime is formulated, through bosonization techniques, as a massless compact free boson [65,66]. The XX spin chain can be mapped into the free fermionic chain by a Jordan-Wigner transformation. This implies that we can employ the nonuniversal constant term $c_n$ found in [67] through the Fisher-Hartwig theorem, finding that $-c_1' \simeq 0.726$ and $[\partial_n^2(\log c_n)]|_{n=1} \simeq 0.535$, as discussed in the Appendix B.

In Fig. 1 we show $S_{\mathrm{i}\partial\phi,A} - C_{\mathrm{i}\partial\phi,A}$ for a block of $\ell$ consecutive sites in periodic chains of free fermions made by $L$ sites. The top curve is for the XX chain, which corresponds to the compact boson CFT. The numerical data are obtained through the methods described in [61,62]. In the figure, it overlaps with the solid curve, obtained from (5.12) with $\gamma = 1$, where the additive constants are specified above. A very good agreement is observed between the numerical data and the CFT predictions for the compact boson.

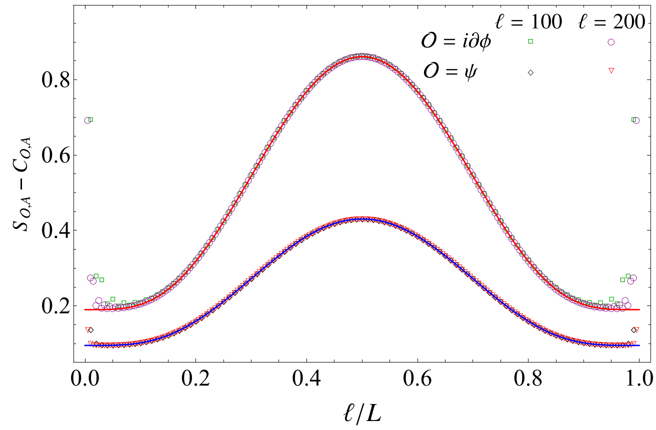


FIG. 1. Difference $S_{\mathcal{O},A} - C_{\mathcal{O},A}$ as function of $\ell/L$. The data in the top curve have been obtained by considering the excited state of the XX chain corresponding to $\mathcal{O} = \mathrm{i}\partial\phi$ in the continuum limit, while the data in the bottom curve have been obtained for the excited state of the Ising chain corresponding to $\mathcal{O} = \psi$ in the continuum limit. The red and blue curves are obtained from (5.12) with $\gamma = 1$ and $\gamma = 1/2$ respectively. The additive constants for the two curves are reported in Secs. V D and V E.

### E. Free fermion

Let us consider the CFT given by free massless fermion (or, equivalently, the Ising CFT) whose central charge is $c = \frac{1}{2}$. In this model, we study the excited states corresponding to the operators $\mathcal{E}$ and $\psi$ [65].

In [62] it has been shown that, for these states, the ratios (5.7) are given by (5.10) with $\gamma = 1$ and $\gamma = 1/2$ respectively, namely

$$F^{(n)}_{\mathcal{E}} = F^{(n)}_{\mathrm{i}\partial\phi}, \quad F^{(n)}_{\psi} = \sqrt{F^{(n)}_{\mathcal{E}}}. \tag{5.15}$$

The discretization of the Ising CFT is provided by the critical Ising spin chain, whose Hamiltonian can be obtained as a particular case of the one of the XY spin chain. It reads

$$H_{\mathrm{XY}} = -\sum_{i=1}^{L}\left(\frac{1+\alpha}{4}\sigma^x_i\sigma^x_{i+1} + \frac{1-\alpha}{4}\sigma^y_i\sigma^y_{i+1} + \frac{\lambda}{2}\sigma^z_i\right), \tag{5.16}$$

in terms of the Pauli matrices $\sigma^{x,y,z}_i$. The Hamiltonian of the critical Ising chain and of the XX chain correspond to (5.17) with $\alpha = \lambda = 1$ and $\alpha = \lambda = 0$ respectively. Performing a Jordan-Wigner transformation, $H_{\mathrm{XY}}$ is mapped into a chain of free fermions; hence the critical Ising chain and the XX chain (and therefore the numerical data of the two panels of Fig. 1) correspond to two different free fermionic models.

In Fig. 1, the bottom curve is $S_{\psi,A} - C_{\psi,A}$ for a block of $\ell$ consecutive sites in a periodic Ising chain made by $L$ spins. The data points are obtained through the procedure detailed in [61,62] and compared with (5.12) evaluated for $\gamma = 1/2$ (blue curve), finding a very good agreement. In this case the additive constants have been fitted, finding $-c_1' \simeq 0.479$ (as already found in [8,68]) and $[\partial_n^2(\log c_n)]|_{n=1} \simeq 0.385$.

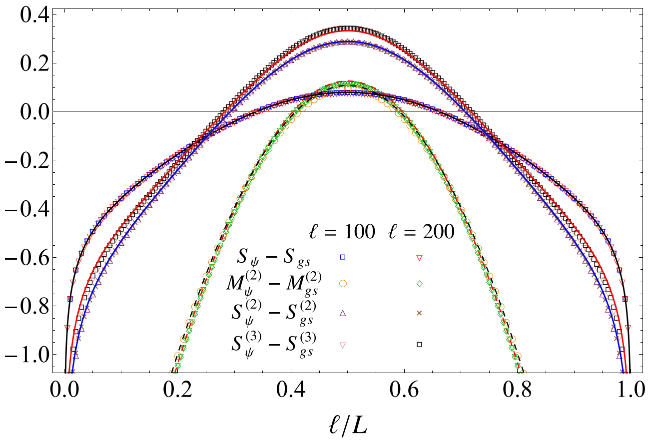


FIG. 2. Changes in the monotones as function of $\ell/L$ for moving from the ground state to the excited state of the Ising chain corresponding to $\mathcal{O} = \psi$ in the continuum limit. The black, blue, and red solid curves are obtained from (5.8), (5.9), (5.10), and (5.11), while the dashed curves are obtained using (5.4) and (5.9), (5.10), and (5.11) into (1.6) with $\ell = 100$ (black dashed curve) and $\ell = 200$ (red dashed curve). The nonuniversal constants have been fixed as specified in Secs. V D and V E.

### F. Constraints for majorization from monotones

Let us consider the results in the light of our exploratory approach to majorization outlined in Sec. V A. As anticipated, we focus on the periodic Ising chain, which maps into a chain of free fermions after a Jordan-Wigner transformation, with a finite-dimensional Hilbert space. We probe possible majorization between the ground state and an excited state. The excited state corresponds to the state created by the operator $\mathcal{O} = \psi$ in the continuum CFT. To see if majorization is ruled out, we consider first the changes in the entanglement monotones $S$, $S^{(2)}$, $S^{(3)}$, and $M^{(2)}(\cdot, 1)$ for the interval $A$ of length $\ell$. The changes in entanglement entropy $\Delta S_A = S_{A,\psi} - S_{A,gs}$, in Rényi entropies $\Delta S_A^{(2)} = S_{A,\psi}^{(2)} - S_{A,gs}^{(2)}$ and $\Delta S_A^{(3)} = S_{A,\psi}^{(3)} - S_{A,gs}^{(3)}$ and in the second moment of modular Hamiltonian $\Delta M_{2,A} = M_{A,\psi}^{(2)}(\cdot, 1) - M_{A,gs}^{(2)}(\cdot, 1)$ increase monotonically from $\ell/L = 0$ to a maximum value when the subsystem is half the size of the chain, $\ell/L = 0.5$, as shown in Fig. 2. Assuming that we want to convert the ground state to an excited state by a process (such as $|gs\rangle \xrightarrow[\text{LOCC}]{} |\psi\rangle$) that involves the order $|\psi\rangle \succ |gs\rangle$, the entanglement monotones cannot increase. Notice that $\Delta S_A$, $\Delta S_A^{(2)}$, $\Delta S_A^{(3)}$, and $\Delta M_{2,A}$ start as negative for small subsystem size, but all of them become positive as the size increases, thus ruling out $|\psi\rangle \succ |gs\rangle$ and the LOCC process. Interestingly, $\Delta M_{2,A}$ becomes positive only at $\ell/L \approx 0.403$, $\Delta S_A$ at $\ell/L \approx 0.337$, $\Delta S_A^{(2)}$ at $\ell/L \approx 0.292$ and $\Delta S_A^{(3)}$ at $\ell/L \approx 0.282$. Thus, in this case the monotone $S_A^{(3)}$ gives the stronger constraint, ruling out $|\psi\rangle \succ |gs\rangle$ in the range $\ell/L \in [0.282, 0.718]$. In the opposite transition from the excited state to the ground state with $|gs\rangle \succ |\psi\rangle$, the signs are reversed with $\Delta M_{2,A} \geqslant 0$ giving a stronger constraint for ruling out the $|gs\rangle \succ |\psi\rangle$ in the regime $\ell/L \in [0, 0.403] \cup [0.597, 1]$.

The general picture is consistent with the naive expectation that it becomes relatively harder to connect two pure states by means of LOCC as the difference between the sizes of the two subsystems $A$ and $B$ becomes smaller. A heuristic argument is that the dimension of the space of allowed local operations decreases as $|d_A - d_B|$ gets smaller. If one, e.g., looks at unitaries, the total dimension of the group of local unitaries $U(d_A) \times U(d_B)$ is $d_A^2 + d_B^2$, which is minimized for $d_A = d_B$. Beyond this, it seems hard to extract general lessons from our preliminary analysis.

## VI. OUTLOOK

*Partial orders among quantum states.* The sequences of inequalities coming from the monotones introduced in Sec. III can be used to define a partial order among quantum states. For example, starting with the sequence $M^{(n)}(\rho; n-1)$, we could define $\rho \succ_n \sigma$ if all monotones up to degree $n$ obey $M^{(k)}(\rho; n-1) \leqslant M^{(k)}(\sigma; n-1)$, and then define $\rho \succ_\infty \sigma$ by $\lim_{k\to\infty} \rho \succ_k \sigma$. With the extremal sequence $P_E^{(n)}(\rho)$ we have even more freedom, due to the infinite range of parameters $\vec{a} = (a_1, \dots, a_{\lceil \frac{n}{2} \rceil - 1})$, where $\lceil x \rceil$ is the smallest integer greater than or equal to $x$. As we did for the order $n = 3, 4$ cases, we could first express $P_E^{(n)}$ in terms of $M^{(k)}$ up to order $n$, and then find the value of parameters $\vec{a}_0$ that produces the tightest inequality. The extremal parameter vector $\vec{a}_0$ can be expressed as a combination of $M^{(k)}$, leading to a nonlinear expression for $P_{E,\vec{a}_0}^{(n)}$ in terms of $M^{(k)}$. We can then define a tighter partial order $\rho \succ_n^E \sigma$ by requiring $P_{E,\vec{a}_0}^{(k)}(\rho) \geqslant P_{E,\vec{a}_0}^{(k)}(\sigma)$ for all $k \leqslant n$, and a limit $\rho \succ_\infty^E \sigma$ as in the above.

It is an interesting question whether any of the above partial orders is equivalent to the majorization partial order. Currently, it is known that the monotonicity of Rényi entropies is not sufficient to imply majorization. All in all, our best hope is perhaps that the order $\succ_\infty^E$ is strong enough to imply majorization. The hope is based on the equivalence (the Hardy-Littlewood-Pólya inequality of majorization [69])

$$\begin{aligned} &\rho \succ \sigma \quad \Leftrightarrow \quad \mathrm{Tr} f(\rho) \geqslant \mathrm{Tr} f(\rho) \\ &\text{for all real valued continuous convex} \\ &\text{functions } f \text{ defined on } [0, 1]. \end{aligned} \tag{6.1}$$

As we will discuss later in this section, one can construct a sufficient condition for majorization, by approximating convex functions with convex Bernstein polynomials. The latter are series expansions in the Bernstein basis constructed from $f$.

One could hope that an analogous series expansion based on the extremal sequence $P_{E,\vec{a}_0}^{(n)}$ would be possible with all coefficients being non-negative, or in other words, that non-negative linear combinations of the extremal polynomials are dense in the space of real values continuous convex functions. Then $\succ_\infty^E$ would imply that the inequality on the right-hand side of (6.1) is satisfied, and be equivalent to majorization. The above mentioned partial order based on monotones and majorization partial order can also be formulated as orderings generated by cones, this concept is discussed, e.g., in [70]. We present this reformulation in Appendix C.

We remark that an efficient criterion, based on a semi-definite program, to determine whether the majorization condition between two states holds has been found in [71]. This algorithm is convenient for confirming the validity of majorization but is somewhat agnostic of the underlying physics, particularly concerning resource theory. Additional insight can be attained by establishing a series of inequalities (derived from monotones) equivalent to the condition of majorization. A set of infinite inequalities with this feature has been found in [71], where the associated monotones are constructed from the conditional min-entropy [72]. A related question is whether the monotones introduced in this paper could also yield a condition equivalent to majorization. Moreover, in our case, the monotones can be expressed in terms of cumulants of the modular Hamiltonian and, therefore, computed from Rényi entropies. Given the advanced technology for computing Rényi entropies in a quantum field theory or a many-body lattice model, our investigation may lead to progress in the study of majorization in the infinite-dimensional setting, which is one of the main motivations of this paper.

Another interesting open question is a version of a moment problem. For a state $\rho$ in a $d$-dimensional system, it is known that $d-1$ first Rényi entropies $S^{(k)}(\rho)$, $k=1,\dots,d-1$, are sufficient to determine the spectrum of $\rho$. Given the spectrum, in small enough dimension an algorithm can then order the eigenvalues and test majorization for a pair of states. The explicit steps and comments on the history of the above observation can be found in [3]. The proof is actually straightforward, as the Rényi entropies yield a basis for the symmetric polynomials in the eigenvalues, which can then directly be used to compute $\det(\lambda-\rho)$ whose roots are the spectrum of $\rho$. In a similar vein, suppose that one knows all $M^{(k)}$ or all $P_E^{(k)}$ up to order $n$, for a state $\rho$ in a $d$-dimensional system. Is it possible to derive the spectrum of $\rho$ for some value of $n$ or in the limit $n\to\infty$? If not, can even a partial spectrum be calculated?

*Extremal convex polynomials.* Functions of the type $\mathrm{Tr}[F(\rho)]$ with complex polynomial $F$ form a cone. The function $F$ is convex when $F''\geqslant 0$, and we can once again find a complete basis of extremal polynomials. In this case, we need to find positive polynomials on the interval [0,1] and these are given by linear combinations (with non-negative coefficients) of polynomials of the form $\prod_i(x-a_i)^2$ or $x(1-x)\prod_i(x-a_i)^2$ with in either case $a_i\in[0,1]$. Notice that linear combinations of such polynomials can yield a polynomial of a lower degree, and one therefore has to be a bit careful to find all polynomials of a particular degree. For example, the most general linear $F''$ is a non-negative linear combination of $x$ and $1-x$, and since $x=x(1-x)+x^2$ and $1-x=x(1-x)+(1-x)^2$ these can indeed both be written as linear combinations of the extremal basis polynomials of higher degree. If $F''=x$ then the monotone is simply $\mathrm{Tr}[\rho^3]$, and for $F''=1-x$ we obtain the monotone $\frac{1}{2}\mathrm{Tr}[\rho^2]-\frac{1}{6}\mathrm{Tr}[\rho^3]$. Going to higher degrees, one could in principle obtain infinite families of monotones. On other hand, as we discuss in Appendix D, there is a simple family of continuous convex functions that provide a sufficient inequality test of majorization, and can be approximated to arbitrary accuracy by convex Bernstein polynomials. In this way one obtains a sequence of polynomial inequalities (D10). Furthermore, convex Bernstein polynomials can be expressed as positive coefficient linear combinations of the extremal convex polynomials. This implies that the inequalities (D10) are equivalent to inequalities satisfied by the extremal convex polynomials. The family of extremal convex polynomial monotones would then be complete; in other words, imposing all of them as inequalities would be equivalent to state majorization. It would be interesting to study this in more detail.

*Inequalities for quantum field theories.* As we discussed, to define majorization in quantum field theory directly requires one to introduce an explicit UV cutoff. It is, however, not obvious that this is a natural construction as the notion of majorization may depend sensitively on the choice of UV cutoff. Since relative entropy, as opposed to entanglement entropy, is well defined for continuum quantum field theories, it is tempting to think that only a relative version of majorization applies in continuum quantum field theories. This leads one to consider the inequality $S(\rho_1||\sigma)\geqslant S(\rho_2||\sigma)$ in quantum field theory. This inequality would follow if there exists a quantum channel $\mathcal{N}$, which maps $\rho_1$ to $\rho_2$ and maps $\sigma$ to itself. For general quantum channels monotonicity of relative entropy is the statement that $S(\rho||\sigma)\geqslant S(\mathcal{N}(\rho)||\mathcal{N}(\sigma))$. Similar monotonicity properties are satisfied by Rényi relative entropies (Rényi divergences). In [73] monotonicity constraints $S_\alpha(\rho(0)||\gamma_\beta)\geqslant S_\alpha(\rho(t)||\gamma_\beta)$ were investigated as additional "second laws" constraining the off-equilibrium dynamical evolution $\rho(t)=\mathcal{N}_t(\rho(0))$ [where the Gaussian state is a fixed point $\gamma_\beta=\mathcal{N}_t(\gamma_\beta)$] in 2d CFTs and their gravity duals.

To speculate, one could try an alternative method to construct inequalities and proceed as follows. In quantum field theory, the definition of relative entropy also requires a choice of algebra, typically associated to a subregion. If we can replace the action of the channel on states by the adjoint action $\mathcal{N}^*$ on the algebra $\mathcal{A}$, defined via $\mathrm{Tr}(\mathcal{N}(\rho)O)=\mathrm{Tr}(\rho\mathcal{N}^*(O))$ for all $\rho$ and $O\in\mathcal{A}$, then we can also write $S_{\mathcal{A}}(\rho||\sigma)\geqslant S_{\mathcal{N}^*\mathcal{A}}(\rho||\sigma)$. This inequality follows from a corresponding operator inequality for the relative modular operators, $\Delta_{\rho|\sigma;\mathcal{A}}\leqslant\Delta_{\rho|\sigma;\mathcal{N}^*\mathcal{A}}$. We could take this inequality to be the fundamental inequality, which defines a quantum field theory counterpart of majorization. It would define a partial ordering for algebras rather than for states. By applying operator monotones [74] we could then derive additional inequalities in the spirit of the paper. We leave a further exploration of these ideas to future work.

*Additional open questions [75].* It is worth asking whether it is possible to generalize from [3] the Result 1 (a sufficient condition for approximate state transition) or the Result 2 (bounds on smoothed min and max entropies) to involve higher cumulants than entropy and variance. As far as we can see, these results rely on the Cantelli-Chebyshev inequality for deviations of a random variable from its mean value, with the bound depending on the variance. One could try to employ a refined inequality involving higher cumulants as well, and then try to construct extensions of the abovementioned results. Finally, it would be interesting to explore if our approach to resource monotones has interesting applications in other quantum resource theories. In particular, it would be interesting to study resource monotones in the context of (un)complexity and its connections to quantum gravity.

## VII. BRIEF SUMMARY

We introduced an infinite sequence of monotones, which are polynomial in the modular Hamiltonian and can be computed from Rényi entropies. Such monotones form a cone, and we identified the sequence of extremal polynomial monotones that span the cone. All these monotones involve the von Neumann entropy in combination with higher cumulants. Our paper generalizes a previous result of [3].

In processes involving majorization between the input and output states, the monotones yield infinite sequences of inequalities that bound the change in entropy. As an example, we considered an idealized model of state erasure, where, in addition to the Landauer inequality, the monotones yield an infinite sequence of more refined inequalities.

Next, we moved to more general processes involving majorization between pairs of states, motivated by resource theories, such as the one of quantum thermodynamics. We showed how to use our monotones to construct infinite sequences of resource monotones, which must be applied to a pair of commuting states. As an application, we derived a finite-dimension correction to the Clausius inequality.

In the end, we discussed whether these sequences of inequalities could be equivalent to majorization. If one instead considers polynomials $\mathrm{Tr}(F(\rho))$ (where $F$ is a polynomial), we sketched a way to perhaps obtain a criterion for majorization based on sequences of convex Bernstein polynomials.

## ACKNOWLEDGMENTS

We thank Henrik Wilming for useful comments and questions. G.D.G., E.K.V., and E.T. acknowledge Galileo Galilei Institute for warm hospitality and financial support (through the program *Reconstructing the Gravitational Hologram with Quantum Information*) during part of this work. J.d.B. is supported by the European Research Council under the European Union's Seventh Framework Programme (Grant No. FP7/2007-2013), ERC Grant Agreement No. ADG 834878. E.K.V.'s research has been conducted within the framework of InstituteQ - the Finnish Quantum Institute, and E.T.'s research within the framework of the Trieste Institute for Theoretical Quantum Technology (TQT).

## APPENDIX A: DETAILS ON THE CONSTRUCTION OF ENTANGLEMENT MONOTONES

In this Appendix, we provide additional discussions on the construction of entanglement monotones for pure states detailed in Sec. II. In particular, in Appendix A 1 we describe a general procedure for obtaining entanglement monotones from the cumulants of the modular Hamiltonian, while in Appendix A 2 we report a detailed proof of the Theorem 1.

### 1. Higher cumulants and pure-state entanglement monotones

In this subsection we point out that there are many ways to construct generalizations of the function $M$ defined in (1.6) to concave quantities (entanglement monotones)

$$M^{(n)}(\boldsymbol{\gamma}) = \mathrm{Tr}[\rho F_n(\rho)], \tag{A1}$$

involving higher cumulants or moments of modular Hamiltonian, where $x\,F_n(x)$ is a concave function of a single real variable $x \in [0, 1]$. With a slight abuse of notation, we denote the quantity in (A1) like the one in (3.6) although the former one is more general because of the occurrence of the parameters $\boldsymbol{\gamma} = (\gamma_0, \dots, \gamma_{n-1})$, as discussed below.

Let us list explicitly the relation (3.8) of moments $\mu_n$ and the cumulants $C_n$ for first moments,

$$\mu_1 = C_1, \tag{A2}$$

$$\mu_2 = C_1^2 + C_2, \tag{A3}$$

$$\mu_3 = C_1^3 + 3\,C_1\,C_2 + C_3, \tag{A4}$$

$$\mu_4 = C_1^4 + 6\,C_1^2\,C_2 + 3\,C_2^2 + 4\,C_1\,C_3 + C_4. \tag{A5}$$

One can invert to obtain the relation of cumulants to moments, for example,

$$C_1 = \mu_1, \tag{A6}$$

$$C_2 = \mu_2 - \mu_1^2, \tag{A7}$$

$$C_3 = \mu_3 - 3\,\mu_2\,\mu_1 + 2\,\mu_1^3, \tag{A8}$$

$$C_4 = \mu_4 - 4\,\mu_3\,\mu_1 + 12\,\mu_2\,\mu_1^2 - 6\,\mu_1^4 - 3\,\mu_2^2. \tag{A9}$$

In our case, given a density matrix $\rho$, we are interested in the moments of modular Hamiltonian $\mu_n = \mathrm{Tr}(\rho(-\ln\rho)^n) = \mathrm{Tr}(\rho K^n)$, where $K = -\ln\rho$ and likewise for the cumulants. The entropy is $C_1 = \mu_1 = S$ and the capacity is $C_2 = C$.

Instead of considering moments of shifted modular Hamiltonian $\mathrm{Tr}[\rho(-\ln\rho + b_n)^n]$ with $b_n > n-1$ as concave generalizations of $M$, there are more general constructions. Here is one way to proceed. For $n = 1$, let us consider

$$M^{(1)} \equiv C_1 + a_1 = \mu_1 + a_1. \tag{A10}$$

Up to an additive constant, $M^{(1)}$ is the entropy. For $n = 2$, let us consider

$$\begin{aligned} M^{(2)} &\equiv C_2 + (\mu_1 + a_1)^2 + a_2 \\ &= \mu_2 + 2a_1\mu_1 + a_1^2 + a_2 = \mu_2 + \gamma_1\mu_1 + \gamma_0, \end{aligned} \tag{A11}$$

which depends on the two parameters $\gamma_0$ and $\gamma_1$. Comparing (A11) with (1.6), we find that, when $\gamma_1 = 2$ and $\gamma_0 = 1$, $M^{(2)}(\gamma_0, \gamma_1)$ reduces to $M$ in (1.6). Notice that the polynomial combination of capacity and entropy in $M^{(2)}$ reduces to a linear combination of moments $\mu_n$ of modular Hamiltonian.

Likewise, for $n = 3$, if we start with a polynomial expression

$$\begin{aligned} M^{(3)} \equiv C_3 + \tfrac{3}{2}(\mu_2 + \mu_1 + b_1)^2 - 2(\mu_1 + c_1)^3 \\ + \alpha_2(\mu_2 + a_2)^2 + \alpha_1(\mu_1 + a_1)^2, \end{aligned} \tag{A12}$$

with the coefficients $3/2$ and $-2$ in (A12) we can cancel the terms $\mu_2\mu_1$ and $\mu_1^3$ respectively [see (A8)]. Imposing the vanishing of the coefficients of $\mu_2^2$ and $\mu_1^2$ in (A12) leads to

$$\alpha_2 = -\tfrac{3}{2}, \quad \alpha_1 = -\tfrac{3}{2} + 6\,c_1. \tag{A13}$$

The point is that we are again lead to a linear combination of moments,

$$M^{(3)}(\gamma_0,\gamma_1,\gamma_2)=\mu_3+\gamma_2\,\mu_2+\gamma_1\,\mu_1+\gamma_0, \tag{A14}$$

with three parameters

$$\begin{aligned}\gamma_2&\equiv 3(b_1-a_2),\\ \gamma_1&\equiv 3\big(b_1-2c_1^2+4c_1a_1-a_1\big),\\ \gamma_0&\equiv -\tfrac{1}{2}\big(4c_1^3-3b_1^2-12c_1a_1^2+3a_2^2+3a_1^2\big).\end{aligned} \tag{A15}$$

Thus, since $\mathrm{Tr}\rho=1$, we have

$$\begin{aligned}M^{(3)}(\gamma_0,\gamma_1,\gamma_2)&=\mathrm{Tr}[\rho\,F_3(\rho)]+\gamma_0,\\ F_3(\rho)&\equiv-(\ln\rho)^3+\gamma_2\,(\ln\rho)^2-\gamma_1\ln\rho.\end{aligned} \tag{A16}$$

It is easy to see that there is a range of parameters $\gamma_1,\gamma_2$ such that $f_3(x)=x\,F_3(x)$ is concave in the unit interval. The difference to the third moment of shifted modular Hamiltonian $\mathrm{Tr}[\rho(-\ln\rho+b_3)^3]$ is that it contains a single parameter $b_3$. It is a special case of (A16) with

$$\gamma_j=(-1)^j\binom{3}{3-j}b_3^{3-j},\quad \text{where } j=0,1,2. \tag{A17}$$

We are thus lead to consider more general linear combinations of moments as an alternative generalization of the measure $M$ [of which the moments of shifted modular Hamiltonian (3.7) are a special case],

$$\begin{aligned}M^{(n)}(\boldsymbol{\gamma})&\equiv\mu_n+\sum_{j=1}^{n-1}\gamma_j\mu_j+\gamma_0\\ &=\mathrm{Tr}\left[\rho\left((-1)^n\ln^n\rho+\sum_{j=1}^{n-1}\gamma_j(-1)^j\ln^j\rho\right)\right]+\gamma_0\\ &=\mathrm{Tr}[\rho F_n(\rho)]+\gamma_0,\end{aligned} \tag{A18}$$

where $\boldsymbol{\gamma}=(\gamma_0,\dots,\gamma_{n-1})$. The range of the parameters $\gamma_{j\neq 0}$ can be chosen so that $xF_n(x)$ is a concave function for $x\in[0,1]$ and $\gamma_0$ so that $M^{(n)}(\boldsymbol{\gamma})\geqslant 0$. It is also clear that these measures can be computed by using the Rényi entropies or $\mathrm{Tr}(\rho^\alpha)$ as a generating function, by applying a combination of derivatives $\sum_k\gamma_k(-1)^k\,\partial_\alpha^k$ and setting $\alpha=1$.

To summarize, there are many ways to construct infinite sequences of entanglement monotones, generalizing $M$, and compute them from Rényi entropies. In the end, the desirable construction depends on the specific physical motivation.

### 2. Proof of Theorem 1

According to Theorem 1, all positive semidefinite polynomials $G(y)$ on the negative half-line $y\in(-\infty,0]$ have the following form. For polynomials $G(y)$ of degree $2d$ (with $d\geqslant 1$) they are linear combinations with positive coefficients of polynomials of the form $G_{\vec a}(y)=\prod_{i=1}^d(y+a_i)^2$, with all $a_i\geqslant 0$. For polynomials of degree $2d+1$ they are linear combinations with positive coefficients of polynomials of the form $G_{\vec a}(y)=-y\prod_{i=1}^d(y+a_i)^2$, with again all $a_i\geqslant 0$.

*Proof.* Consider first a positive polynomial on the entire real line. It can be written as $\prod_i(x-x_i)$ where the roots can be complex. There cannot be an isolated real root, as then the polynomial would be negative somewhere in a small neighborhood of that real root. Similarly, there can not be an odd degeneracy of a real root, because once more the function would be negative in a small neighborhood. Therefore, all real roots need to have even degeneracy. So the polynomial is of the form $q(x)^2r(x)$ where $q(x)$ is real and all other (complex) roots make up $r(x)$. Because the polynomial must be real, the roots must come in complex conjugate pairs. Therefore, $r(x)=|s(x)|^2$ where $s(x)$ contains all the roots in (say) the complex upper half plane. We can write $s=s_0+is_1$ where $s_0$ and $s_1$ are the real and imaginary parts. Then we see that the polynomial is of the form $q(x)^2s_0(x)^2+q(x)^2s_1(x)^2$, which shows that a positive polynomial on the real line must be sum of two squares.

Now consider a polynomial $p(x)$, which is positive on the negative real axis. We can decompose these polynomials again in roots. Negative real roots need to appear with even multiplicity and positive real roots can appear with any multiplicity. Factors of the type $|x-u|^2$ with complex $u$ are positive definite and can appear without restriction. Consider now polynomials of the form $f-xg$ with $f$ and $g$ positive on the entire real axis. These polynomials form a ring (so if you multiply two it will still be of this form). The claim is that $p$ is also of this form (which is manifestly non-negative on the negative real axis). We simply need to check that all factors of $p$ are of this form. A factor with negative real roots with even multiplicity is of the form $f+x\cdot 0$ as it is positive on the entire real axis. A factor with a positive real root can be written as $u-x$, which is also of the required form (with $f=u$ and $g=1$). Finally, factors $|x-u|^2$ are positive definite on the entire real axis and therefore also of the form $f+x\cdot 0$. Using the previous characterization of positive polynomials we conclude that $p(x)$ can be written as

$$p(x)=q(x)^2+r(x)^2-x[s(x)^2+t(x)^2], \tag{A19}$$

for some polynomials $q,r,s,t$. This result is due to Pólya-Szegö [76]. It remains to show that each of these terms can be written as a liner combination of extremal polynomials. Look, e.g., at $q(x)^2$ and expand it in the form

$$q(x)^2=\prod_{i,j}(x-u_i)^2\big[(x-b_j)^2+c_j^2\big]^2. \tag{A20}$$

This is a sum of terms of the form $a_-(x)^2a_+(x)^2$ with positive coefficients, where $a_-(x)$ has zeros on the negative real axis, and $a_+(x)$ has zeros on the positive real axis. We can further expand $a_+(x)^2$ as a power series with alternating coefficients. This shows that $q(x)^2$ is indeed a linear combination of extremal polynomials with non-negative coefficients. The same result applies for the other three terms in $p(x)$. This completes the proof. ∎

## APPENDIX B: CAPACITY OF ENTANGLEMENT IN FERMIONIC CHAINS: CONSTANT TERM

In this Appendix we exploit the method of [67] to determine the nonuniversal constant occurring in the expression of the capacity of entanglement for a block $A$ made by $\ell$ consecutive sites in the infinite free fermionic chain.

The Hamiltonian of the free fermionic chain on the line reads

$$H = -\sum_{n=-\infty}^{+\infty}\left[\hat{c}_n^\dagger\,\hat{c}_{n+1} + \hat{c}_{n+1}^\dagger\,\hat{c}_n - 2h\left(\hat{c}_n^\dagger\,\hat{c}_n - \frac{1}{2}\right)\right], \quad \text{(B1)}$$

where $\{\hat{c}_n^\dagger, \hat{c}_m^\dagger\} = \{\hat{c}_n, \hat{c}_m\} = 0$ and $\{\hat{c}_n, \hat{c}_m^\dagger\} = \delta_{m,n}$ and $h$ is the chemical potential. The ground state of this model is a Fermi sea with a Fermi momentum $k_{\rm F} = \arccos|h|$. A Jordan-Wigner transformation maps the Hamiltonian (B1) into the Hamiltonian of the XX spin chain with magnetic field $h$.

The Toeplitz nature of the correlation matrix restricted to $A$ for this model allows us to write the large $\ell$ expansion of $\ln {\rm Tr}\rho_A^n$ through the Fisher-Hartwig conjecture, which provides the asymptotic behavior of the Fredholm determinant of Toeplitz matrices for large matrix size. The result reads [67]

$$\begin{aligned}\ln {\rm Tr}\rho_A^n &= \frac{1}{6}\left(\frac{1}{n} - n\right)\ln\ell + \ln c_n\\ &= \frac{1}{6}\left(\frac{1}{n} - n\right)[\ln\ell + \ln(2|\sin(k_{\rm F})|)]\\ &\quad + \Upsilon(n) + o(1),\end{aligned} \quad \text{(B2)}$$

where

$$\begin{aligned}\Upsilon(n) = {\rm i}n\int_{-\infty}^{\infty} dw\Bigg\{&[\tanh(\pi w) - \tanh(\pi n w)]\\ &\times \ln\left(\frac{\Gamma\left(\frac{1}{2} + {\rm i}w\right)}{\Gamma\left(\frac{1}{2} - {\rm i}w\right)}\right)\Bigg\}.\end{aligned} \quad \text{(B3)}$$

By introducing $G_n(w) \equiv n[\tanh(\pi w) - \tanh(\pi n w)]$, we need

$$\partial_n G_n(w)\big|_{n=1} = -\frac{\pi w}{\cosh^2(\pi w)}, \quad \text{(B4)}$$

$$\partial_n^2 G_n(w)|_{n=1} = -\frac{2\pi w}{\cosh^2(\pi w)} + \frac{2\pi^2 w^2}{\cosh^2(\pi w)}\tanh(\pi w). \quad \text{(B5)}$$

Plugging (B4) into the derivatives of (B3), one finds the corrections to the entanglement entropy and the capacity of entanglement due to $\Upsilon(n)$ in (B2). This gives

$$S_A = \tfrac{1}{3}\ln\ell + \tfrac{1}{3}\ln(2|\sin(k_{\rm F})|) - \Upsilon'(1) + \dots, \quad \text{(B6)}$$

which has been obtained in [67], and

$$C_A = \tfrac{1}{3}\ln\ell + \tfrac{1}{3}\ln(2|\sin(k_{\rm F})|) + \Upsilon''(1) + \dots, \quad \text{(B7)}$$

where the constant $\Upsilon'(1)$ and $\Upsilon''(1)$ can be evaluated numerically from (B4) in (B3), finding $-\Upsilon'(1) \simeq 0.495018$ and $\Upsilon''(1) \simeq 0.303516$. The subleading terms that we have neglected are vanishing as $\ell \to \infty$ and some of them have been computed in [77] through the generalised Fisher-Hartwig conjecture.

In the main text we have mainly considered (B6) and (B7) in the case of vanishing chemical potential, i.e., for $h = 0$, which means $k_{\rm F} = \frac{\pi}{2}$ (see e.g., all the figures in Sec. V).

## APPENDIX C: SEQUENCES OF MONOTONES AND AN ORDERING GENERATED BY A CONE

In the Sec. VI we discussed partial orders among quantum states based on the sequences of our new monotones. Here we rephrase this question in terms of an ordering generated by a cone [70]. By diagonalizing a density matrix, the space of quantum states in $d+1$ dimensions can be identified with the standard simplex $\Delta^d \subset \mathbb{R}^{d+1}$. Our monotones can be thought as convex functions

$$M : \Delta^d \to \mathbb{R},\ M(\boldsymbol{x}) \equiv \sum_{i=0}^{d} x_i F(\ln x_i), \quad \text{(C1)}$$

such that

$$G(y) \equiv F''(y) + F'(y), \quad \text{(C2)}$$

with $y \equiv \ln x$ is a non-negative polynomial of the order $n-1$ on the negative half-line $(-\infty, 0]$. Thus $F$ is polynomial of degree $n$. The above functions form a convex cone $\mathcal{C}_n$. For our purposes we may identify functions that differ by a constant. Every function $M$ is a linear combination with positive coefficients of the extremal rays of the cone. Let $\Phi_n$ denote the set of extremal rays. We say that the set $\Phi_n$ *generates* the cone $\mathcal{C}_n$. For $n = 1$ there is only one extremal ray with $F(y) = y$, thus $\Phi_1 = \{\sum_i x_i \ln x_i\}$. For $n > 1$ we found in Theorem 1 that the extremal rays correspond to functions $F_{\vec{a}}(y)$ with $G_{\vec{a}}(y) = F_{\vec{a}}''(y) + F_{\vec{a}}(y)$ of the form

$$G_{\vec{a}}(y) \equiv \begin{cases}\prod_{i=1}^{k}(y + a_i)^2, & a_i \geqslant 0\ \forall i, \quad \text{when } n-1 = 2k \geqslant 2,\\ -y\prod_{i=1}^{k}(y + a_i)^2, & a_i \geqslant 0\ \forall i, \quad \text{when } n-1 = 2k+1 \geqslant 3.\end{cases} \quad \text{(C3)}$$

Thus in both cases the generating set $\Phi_n$ is infinite, parameterized by vectors $\vec{a}$ in the hyperorthant of $\mathbb{R}^k$, so the cone $\mathcal{C}_n$ is infinitely generated. Finally let us define the cone

$$\mathcal{C} = \cup_{n=1}^{\infty}\mathcal{C}_n, \quad \text{(C4)}$$

which is infinitely generated by the set

$$\Phi = \cup_{n=1}^{\infty}\Phi_n. \quad \text{(C5)}$$

Majorization $\rho \succ \sigma$ in $d+1$ dimensions can be identified with majorization of vectors $\boldsymbol{\lambda} \succ \boldsymbol{\mu}$ or majorization partial order in the standard simplex $\Delta^d$. Now alternatively [70] we can define an ordering $\succ_{\mathcal{C}}$ based on the function set $\Phi$ that generates the cone $\mathcal{C}$,

$$\boldsymbol{\lambda} \succ_{\mathcal{C}} \boldsymbol{\mu} \Leftrightarrow M(\boldsymbol{\lambda}) \geqslant M(\boldsymbol{\mu})\ \forall M \in \Phi. \quad \text{(C6)}$$

The inequality on the right-hand side is satisfied by every function $M \in \mathcal{C}$, the definition just uses the most economical set of functions generating the cone. The ordering $\succ_{\mathcal{C}}$ is said to be *generated* by the cone $\mathcal{C}$. Such (partial) orderings come with a basic problem. Define the *completion* $\mathcal{C}^*$ of $\mathcal{C}$, the cone of all functions that respect the ordering $\succ_{\mathcal{C}}$,

$$\mathcal{C}^* \equiv \{f : \Delta^d \to R \,||\, \boldsymbol{x} \succ_{\mathcal{C}} \boldsymbol{y} \Rightarrow f(\boldsymbol{x}) \geqslant f(\boldsymbol{y})\}. \tag{C7}$$

A basic problem is to *identify* the completion $\mathcal{C}^*$ of $\mathcal{C}$, which is an important open question for the cone defined above.

We noted the Hardy-Littlewood-Pólya inequality of majorization

$$\rho \succ \sigma \Leftrightarrow \mathrm{Tr} g(\rho) \geqslant \mathrm{Tr} g(\sigma)$$
$$\text{for all continuous convex functions } g : [0,1] \to \mathbb{R}.$$

We could alternatively interpret this as another ordering generated by a cone. Define the convex cone

$$\mathcal{C}_{HLP} = \Bigg\{f : \Delta^d \to R \,||\, f(\boldsymbol{x}) = \sum_{i=1}^{d} g(x_i),\; g : [0,1] \to \mathbb{R} \text{ is continuous, convex}\Bigg\}, \tag{C8}$$

and define the ordering generated by the cone $\mathcal{C}_{HLP}$,

$$\boldsymbol{x} \succ_{\mathcal{C}_{HLP}} \boldsymbol{y} \Leftrightarrow f(\boldsymbol{x}) \geqslant f(\boldsymbol{y}) \,\forall f \in \mathcal{C}_{HLP}. \tag{C9}$$

Then by the HLP inequality we can identify majorization with the cone ordering,

$$\rho \succ \sigma \Leftrightarrow \boldsymbol{\lambda} \succ_{\mathcal{C}_{HLP}} \boldsymbol{\mu}. \tag{C10}$$

Now we can ask if $\mathcal{C}_{HLP}$ could be in the completion $\mathcal{C}^*$ of $\mathcal{C}$ or at least well approximated by $\mathcal{C}^*$. This would mean that $\succ$ and $\succ_{\mathcal{C}}$ are equivalent. We have thus reformulated the question posed in Sec. VI as a problem of comparing orderings generated by cones.

## APPENDIX D: A SUFFICIENT CONDITION FOR MAJORIZATION FROM CONVEX BERNSTEIN POLYNOMIALS

In this paper, we considered convex functions of the type $\mathrm{Tr}[\rho F(\log \rho)]$, which have the feature that expressions of this type include entanglement entropy and moments of shifted modular Hamiltonian. We could also have considered even simpler functions of the type $\mathrm{Tr}[F(\rho)]$ with polynomials $F$, which are essentially linear combinations of exponentials of Rényi entropies with integer powers. In this setting we may construct a sufficient condition for majorization from inequalities of convex polynomials.

From the Hardy-Littlewood-Pólya inequality of majorization (6.1) the first impression may be that the equation

$$\mathrm{Tr} f(\rho) \geqslant \mathrm{Tr} f(\sigma) \tag{D1}$$

must be verified for every continuous convex function $f$ to imply majorization $\rho \succ \sigma$. In fact a sufficient condition for majorization is much simpler. One only needs to consider the continuous convex functions

$$f_a(x) = \begin{cases} 0 & x \leqslant a \\ x - a & x > a \end{cases} \equiv (x-a)\Theta(x-a)\,, \tag{D2}$$

where the parameter $a \in [0, 1]$. Let the pair of ordered eigenvalue vectors of $\rho, \sigma$ of dimension $n$ be $\boldsymbol{\lambda}, \boldsymbol{\mu}$. In [69] it is shown that if

$$\sum_{j=1}^{n} f_a(\lambda_j) \geqslant \sum_{j=1}^{n} f_a(\mu_j) \;\; \forall\, a \in [0,1] \tag{D3}$$

then $\rho \succ \sigma$. Let us now approximate the functions $f_a$ with convex polynomials. Consider the Bernstein basis polynomials

$$b_{k,m}(x) = \binom{m}{k} x^k (1-x)^{m-k}, \tag{D4}$$

with $k = 0, 1, \dots, m$. From any continuous function $g : [0, 1] \to \mathbb{R}$ one can form a general Bernstein polynomial over the basis (D4) by

$$B_m(g)(x) \equiv \sum_{k=0}^{m} g\left(\frac{k}{m}\right) b_{k,m}(x). \tag{D5}$$

It has been shown that

$$\lim_{m\to\infty} B_m(g) = g \tag{D6}$$

uniformly in [0,1]. For the error in the approximation, a simple bound was found by Popoviciu [78] (see also [79]),

$$\max_{x\in[0,1]} |g(x) - B_m(g)(x)| \leqslant \tfrac{5}{4}\omega_g\big(m^{-1/2}\big), \tag{D7}$$

where $\omega_g$ is the uniform modulus of continuity of $g$

$$\omega_g(\delta) = \max\{|f(x) - f(y)| \,|\, x, y \in [0,1], |x-y| \leqslant \delta\}. \tag{D8}$$

Thus we can approximate $g$ to arbitrary precision by a Bernstein polynomial $B_m(g)$. Moreover, it is known that if $g$ is convex, the Bernstein polynomials $B_m(g)$ are also convex. We may thus consider the (convex) Bernstein polynomials derived from $f_a$,

$$\begin{aligned} B_m(f_a)(x) &= \sum_{k=0}^{m} f_a\left(\frac{k}{m}\right) b_{k,m}(x) \\ &= \sum_{k=p}^{m} \left(\frac{k}{m} - a\right)\binom{m}{k} x^k (1-x)^{m-k}, \end{aligned} \tag{D9}$$

where $p$ is the smallest non-negative integer with $\frac{p}{m} > a$. Due to the uniform convergence to $f_a$, instead of (D3) we may reformulate a sufficient condition for majorization, if

$$\sum_{j=1}^{n} B_m(f_a)(\lambda_j) \geqslant \sum_{j=1}^{n} B_m(f_a)(\mu_j) \;\; \forall a \in [0,1]\,, \tag{D10}$$

when $m \to \infty$, then $\rho \succ \sigma$. The virtue of this new criterion is that now we should be able to compute the following quantity without diagonalization:

$$\begin{aligned} \mathrm{Tr}[B_m(f_a)(\rho)] &= \sum_{k=p}^{m} \left(\frac{k}{m} - a\right)\binom{m}{k} \mathrm{Tr}[\rho^k (1-\rho)^{m-k}] \\ &= \sum_{k=0}^{m} \Theta\left(\frac{k}{m} - a\right)\binom{m}{k} \sum_{l=0}^{m-k} (-1)^l \binom{m-k}{l} \\ &\quad \times \mathrm{Tr}[\rho^{k+l}], \end{aligned}$$

and thus check the inequalities (D10) by checking if

$$\Delta^{(m)}_{\rho,\sigma}(a) \equiv \mathrm{Tr}[B_m(f_a)(\rho)] - \mathrm{Tr}[B_m(f_a)(\sigma)]$$
$$= \sum_{k=0}^{m} \Theta\left(\frac{k}{m} - a\right)\binom{m}{k}\sum_{l=0}^{m-k}(-1)^l\binom{m-k}{l}$$
$$\times (\mathrm{Tr}[\rho^{k+l}] - \mathrm{Tr}[\sigma^{k+l}]) \geqslant 0, \qquad \text{(D11)}$$

for all $a \in [0,1]$ for large enough $m$ onwards. This could be done by just plotting $\Delta^{(m)}_{\rho,\sigma}(a)$ in the interval $a \in [0,1]$. According to some preliminary tests, this proposal works well in simple examples. We leave a more extensive investigation for further study. An important question involved is whether there exists a criterion for the order $m$ where the approximation by a Bernstein polynomial becomes good enough to be trusted. The kink in the original function (D2) becomes rounded, and issues may arise when the rounding is too smooth in comparison with the spacing of the spectra of $\rho, \sigma$. Perhaps useful criteria for the approximation can be formulated based on the error estimate (D7) and, e.g., the trace distance of $\rho, \sigma$.


[1] E. Chitambar and G. Gour, Quantum resource theories, Rev. Mod. Phys. **91**, 025001 (2019).

[2] M. A. Nielsen, Conditions for a class of entanglement transformations, Phys. Rev. Lett. **83**, 436 (1999).

[3] P. Boes, N. H. Y. Ng, and H. Wilming, The variance of relative surprisal as single-shot quantifier, PRX Quantum **3**, 010325 (2022).

[4] R. Haag, *Local Quantum Physics* (Springer, New York, 1996).

[5] In quantum information theory $-\ln\rho$ is sometimes called *surprisal*, see e.g., [3]. This name is inherited from its classical equivalent $-\ln p_i$ in the context of a classical discrete probability distribution $p_i$, $\sum_i p_i = 1$: an outcome $i$ with a very low probability is associated with a large surprise—consider, for example, winning the big prize in a lottery. In a quantum version, using a diagonal basis $\rho = \sum_i p_i |i\rangle\langle i|$, if a measurement in the eigenbasis gives as a result an eigenvalue $i$ with very low probability, the surprise is large. Surprisal is also naturally associated with information: Obtaining a rare measurement outcome can be interpreted to reveal a large amount of information about the possible eigenvalues. Indeed another common name for surprisal is *information content.* The von Neumann entropy is thus the expected or average value of surprisal or revealed information, associated with measurement in the eigenbasis.

[6] G. Vidal, J. I. Latorre, E. Rico, and A. Kitaev, Entanglement in quantum critical phenomena, Phys. Rev. Lett. **90**, 227902 (2003).

[7] H. Li and F. D. M. Haldane, Entanglement spectrum as a generalization of entanglement entropy: Identification of topological order in non-Abelian fractional quantum Hall effect states, Phys. Rev. Lett. **101**, 010504 (2008).

[8] J. I. Latorre, E. Rico, and G. Vidal, Ground state entanglement in quantum spin chains, Quantum Inf. Comput. **4**, 48 (2004).

[9] J. I. Latorre, C. A. Lutken, E. Rico, and G. Vidal, Fine grained entanglement loss along renormalization group flows, Phys. Rev. A **71**, 034301 (2005).

[10] R. Orus, Entanglement and majorization in (1+1)-dimensional quantum systems, Phys. Rev. A **73**, 019904(E) (2006).

[11] A. Riera and J. I. Latorre, Area law and vacuum reordering in harmonic networks, Phys. Rev. A **74**, 052326 (2006).

[12] P. Calabrese and A. Lefevre, Entanglement spectrum in one-dimensional systems, Phys. Rev. A **78**, 032329 (2008).

[13] B. Swingle, Entanglement does not generally decrease under renormalization, J. Stat. Mech. (2014) P10041.

[14] M. Mintchev, D. Pontello, A. Sartori, and E. Tonni, Entanglement entropies of an interval in the free Schrödinger field theory at finite density, J. High Energy Phys. 07 (2022) 120.

[15] With this normalization, e.g., for a thermal state $e^{-\beta H}/Z$ we have $K = \beta H + \ln Z$.

[16] H. Yao and X.-L. Qi, Entanglement entropy and entanglement spectrum of the Kitaev model, Phys. Rev. Lett. **105**, 080501 (2010).

[17] J. Schliemann, Entanglement spectrum and entanglement thermodynamics of quantum Hall bilayers at $\nu = 1$, Phys. Rev. B **83**, 115322 (2011).

[18] M. Levin and X.-G. Wen, Detecting topological order in a ground state wave function, Phys. Rev. Lett. **96**, 110405 (2006).

[19] A. Kitaev and J. Preskill, Topological entanglement entropy, Phys. Rev. Lett. **96**, 110404 (2006).

[20] E. Perlmutter, A universal feature of CFT Rényi entropy, J. High Energy Phys. 03 (2014) 117.

[21] J. de Boer, J. Järvelä, and E. Keski-Vakkuri, Aspects of capacity of entanglement, Phys. Rev. D **99**, 066012 (2019).

[22] Y. Nakaguchi and T. Nishioka, A holographic proof of Renyi entropic inequalities, J. High Energy Phys. 12 (2016) 129.

[23] G. Vidal, Entanglement monotones, J. Mod. Opt. **47**, 355 (2000).

[24] D. Reeb and M. M. Wolf, An improved Landauer principle with finite-size corrections, New J. Phys. **16**, 103011 (2014).

[25] C. T. Chubb, M. Tomamichel, and K. Korzekwa, Beyond the thermodynamic limit: Finite-size corrections to state interconversion rates, Quantum **2**, 108 (2018).

[26] Reference [3] uses the convention where definitions involve the binary logarithm $\log_2(x)$ instead of $\ln(x)$. We prefer to follow the physics convention and use the natural logarithm in definitions. Since $\ln(x) = (\ln 2)\log_2(x)$, denoting below quantities defined with $\log_2$ by tildes (e.g., $\tilde{S} = -\mathrm{Tr}[\rho\log_2\rho]$), we have $S = (\ln 2)\tilde{S}$, $C = (\ln 2)^2\tilde{C}$ and $M = (\ln 2)^2\tilde{M}$ with

$$\tilde{M}(\rho) = \tilde{C}(\rho) + \left(\tilde{S}(\rho) + \frac{1}{\ln 2}\right)^2, \qquad (4.12)$$

as given in [3].

[27] This inequality has the same form with the $\log_2$-conventions,

$$\tilde{S}(\rho) - \tilde{S}(\sigma) \geqslant \frac{\tilde{C}(\sigma) - \tilde{C}(\rho)}{2\sqrt{\tilde{M}(\sigma)}}.$$

.

[28] For LOCC transformations of mixed states, we need to apply the convex roof extension to $M^{(n)}$ [see (2.14)].

[29] I. Bengtsson and K. Życzkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement* (Cambridge University Press, Cambridge, 2017).

[30] Given a density matrix $\rho$ and $q > 1$ the Tsallis entropy is $S_q = \frac{1}{1-q}[\mathrm{Tr}\rho^q - 1]$. In the limit $q \to 1$ the Tsallis entropy gives the von Neumann entropy.

[31] X. Hu and Z. Ye, Generalized quantum entropy, J. Math. Phys. **47**, 023502 (2006).
[32] Another equivalent way is by the Stinespring dilation theorem, introducing an environment system $\mathcal{H}_E$ in a reference state $\rho_E$ and then represent $\mathcal{E}(\rho) = \mathrm{Tr}_E[U\rho \otimes \rho_E U^\dagger]$ where $U$ can be chosen to be a unitary operator acting in the composite system $\mathcal{H} \oplus \mathcal{H}_E$ and the partial trace is taken over $\mathcal{H}_E$.
[33] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2010).
[34] A. Uhlmann, Sätze über dichtematrizen, Wiss. Z. Karl-Marx-Univ. Leipzig, Math.-Nat. R. **20**, 633 (1971).
[35] J. Watrous, *The Theory of Quantum Information* (Cambridge University Press, Cambridge, 2018).
[36] D. W. Kribs and R. W. Spekkens, Quantum error-correcting subsystems are unitarily recoverable subsystems, Phys. Rev. A **74**, 042329 (2006).
[37] M. B. Plenio and S. Virmani, An Introduction to entanglement measures, Quantum Inf. Comput. **7**, 1 (2007).
[38] With $|\chi_j\rangle_A = A_j\,|\psi\rangle_A$ , $|\eta_j\rangle_B = B_j\,|\phi\rangle_B$.
[39] This allows for the possibility that some $E(K_i\rho K_i^\dagger)$ in the sum may increase.
[40] A constant term may be added to $g$ to ensure that $E(\rho) = 0$ for separable states.
[41] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Quantum entanglement, Rev. Mod. Phys. **81**, 865 (2009).
[42] M. Kardar, *Statistical Physics of Particles* (Cambridge University Press, Cambridge, 2007).
[43] We remark that there exists a different line of investigation that uses convex cones, to classify and constrain entropy inequalities in holographic gauge-gravity duality, initiated in [80].
[44] V. Powers and B. Reznik, Polynomials that are positive on an interval, Trans. Amer. Math. Soc. **352**, 4677 (2000).
[45] Note that [3] defines $\tilde{M}_x(\rho||\sigma) = \tilde{C}(\rho||\sigma) + (\frac{1}{\ln 2} - \log_2(x) - \tilde{S}(\rho||\sigma))^2$, where we use the tilde notation to emphasize that the quantities are based on binary logarithms, they involve rescalings by $\ln 2$, so that $M_x(\rho||\sigma) = (\ln 2)^2 \tilde{M}_x(\rho||\sigma)$, to compare with (3.57).
[46] M. Ohya and D. Petz, *Quantum Entropy and Its Use* (Springer-Verlag, Berlin, 1993).
[47] The inequality (3.86) is also invariant under a change in the normalization of $\sigma$, obtained through the shift of $K$ by the proper constant.
[48] Note that we assumed an idealized process, following [3], with exact state transitions. In particular, the state of the battery is known exactly before the restoration step. A more realistic and imperfect process would lead to the state of the battery being known only up to some uncertainties. In this case, the work required to restore its state was considered in [52], deriving strategy-dependent bounds on the needed work, involving the smooth min- and max-entropies [81] computed from the state of the battery. We provide additional comments at the end of this section.
[49] P. Faist, Quantum coarse-graining: An information-theoretic approach to thermodynamics, Ph.D. thesis, ETH Zurich, 2016.
[50] O. C. O. Dahlsten, Non-equilibrium statistical mechanics inspired by modern information theory, Entropy **15**, 5346 (2013).
[51] There is again a subtlety with conventions: we get $S((\mathbb{1}/2)^{\otimes n}) = n \ln 2$, whereas (4.11) as quoted in [3] involves $n$, due the convention of using binary logarithms producing $\tilde{S}((\mathbb{1}/2)^{\otimes n}) = n$.
[52] O. C. O. Dahlsten, R. Renner, E. Rieper, and V. Vedral, Inadequacy of von Neumann entropy for characterizing extractable work, New J. Phys. **13**, 053015 (2011).
[53] By the Stinespring dilation one could choose a unitary channel, but we follow [3] and consider this more general possibility.
[54] There are again some rescalings by factors of $\ln 2$, due to our convention of using ln instead of $\log_2$.
[55] Note that we have opposite sign conventions for $\Delta S$, $\Delta C$.
[56] C. G. Callan and F. Wilczek, On geometric entropy, Phys. Lett. B **333**, 55 (1994).
[57] C. Holzhey, F. Larsen, and F. Wilczek, Geometric and renormalized entropy in conformal field theory, Nucl. Phys. B **424**, 443 (1994).
[58] P. Calabrese and J. L. Cardy, Entanglement entropy and quantum field theory, J. Stat. Mech. (2004) P06002.
[59] J. Cardy and E. Tonni, Entanglement Hamiltonians in two-dimensional conformal field theory, J. Stat. Mech. (2016) 123103.
[60] In higher-dimensional CFTs and more general quantum field theories, the relation is more ambiguous; indeed the UV cutoff in the two quantities appears in a power law and the quantities become more dependent on the regularization scheme (see [21] for more discussion).
[61] F. C. Alcaraz, M. I. Berganza, and G. Sierra, Entanglement of low-energy excitations in conformal field theory, Phys. Rev. Lett. **106**, 201601 (2011).
[62] M. I. Berganza, F. C. Alcaraz, and G. Sierra, Entanglement of excited states in critical spin chians, J. Stat. Mech. (2012) P01016.
[63] F. H. L. Essler, A. M. Lauchli, and P. Calabrese, Shell-filling effect in the entanglement entropies of spinful fermions, Phys. Rev. Lett. **110**, 115701 (2013).
[64] P. Calabrese, F. H. L. Essler, and A. M. Lauchli, Entanglement entropies of the quarter filled Hubbard model, J. Stat. Mech. (2014) P09025.
[65] P. Di Francesco, P. Mathieu, and D. Senéchal, *Conformal Field Theory* (Springer, New York, 1997).
[66] A. Gogolin, A. Nersesyan, and A. Tsvelik, *Bosonization and Strongly Correlated Systems* (Cambridge University Press, Cambridge, 2004).
[67] B. Q. Jin and V. E. Korepin, Quantum spin chain, Toeplitz determinants and the Fisher-Hartwig conjecture, J. Stat. Phys. **116**, 79 (2004).
[68] J. L. Cardy, O. A. Castro-Alvaredo, and B. Doyon, Form factors of branch-point twist fields in quantum integrable models and entanglement entropy, J. Stat. Phys. **130**, 129 (2008).
[69] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities* (Cambridge University Press, Cambridge, 1988).
[70] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: Theory of Majorization and Its Applications* (Springer Series in Statistics, New York, 2011).
[71] G. Gour, D. Jennings, F. Buscemi, R. Duan, and I. Marvian, Quantum majorization and a complete set of entropic conditions for quantum thermodynamics, Nat. Commun. **9**, 5352 (2018).
[72] R. Renner, Security of quantum key distribution, Int. J. Quant. Info. **06**, 1 (2008).

[73] A. Bernamonti, F. Galli, R. C. Myers, and J. Oppenheim, Holographic second laws of black hole thermodynamics, J. High Energy Phys. 07 (2018) 111.

[74] A complete classification of operator monotones is known, besides the linear function $f(x) = x$ all other operator monotones are non-negative (possibly infinite) linear combinations of functions of the type $f(x) = x/(x + s)$ with some $s > 0$ [82].

[75] H. Wilming (private communication).

[76] G. Pólya and G. Szegö, *Problems and Theorems in Analysis II* (Springer-Verlag, New York, 1976).

[77] P. Calabrese and F. H. L. Essler, Universal corrections to scaling for block entanglement in spin-1/2 XX chains, J. Stat. Mech. (2010) P08029.

[78] T. Popoviciu, Sur i'approximation des fonctions convexes d'ordre supérieur, Mathematica (Cluj) **10**, 49 (1935).

[79] R. Bojanic and O. Shisha, Degree of $l_1$ approximation to integrable functions by modified Bernstein polynomials, J. Approx. Theory **13**, 66 (1975).

[80] N. Bao, S. Nezami, H. Ooguri, B. Stoica, J. Sully, and M. Walter, The holographic entropy cone, J. High Energy Phys. 09 (2015) 130.

[81] R. Renner and S. Wolf, Smooth Renyi entropy and applications, in *Proceedings of the International Symposium on Information Theory* (IEEE, 2004).

[82] K. Löwner, Über monotone matrixfunktionen, Math. Z. **38**, 177 (1934).