Fusion rules from entanglement

Bowen Shi

Department of Physics, The Ohio State University, Columbus, OH 43210, USA

Kohtaro Kato

Institute for Quantum Information and Matter, California Institute of Technology, Pasadena, CA, USA

Isaac H. Kim

Perimeter Institute for Theoretical Physics, Waterloo ON N2L 2Y5, Canada

Institute for Quantum Computing, University of Waterloo, Waterloo ON N2L 3G1, Canada

IBM T. J. Watson Research Center, Yorktown Heights NY, USA

Stanford University, Stanford CA 94305, USA

Abstract

We derive some of the axioms of the algebraic theory of anyon [A. Kitaev, Ann. Phys., 321, 2 (2006)] from a conjectured form of entanglement area law for two-dimensional gapped systems. We derive the fusion rules of topological charges and show that the multiplicities of the fusion rules satisfy these axioms. Moreover, even though we make no assumption about the exact value of the constant sub-leading term of the entanglement entropy of a disk-like region, this term is shown to be equal to $\ln \mathcal{D}$, where $\mathcal{D}$ is the total quantum dimension of the underlying anyon theory. These derivations are rigorous and follow from the entanglement area law alone. More precisely, our framework starts from two local entropic constraints which are implied by the area law. From these constraints, we prove what we refer to as the "isomorphism theorem." The existence of superselection sectors and fusion multiplicities follow from this theorem, even without assuming anything about the parent Hamiltonian. These objects and the axioms of the anyon theory are shown to emerge from the structure and the internal self-consistency relations of the information convex sets.

1. Introduction

One of the outstanding questions in modern physics concerns the classification of quantum phases. Many attempts have been already made to classify quantum phases over the past decade. For instance, gapped free-electron systems are completely classified [1, 2]. For more general short-range entangled states, an approach based on cobordism was proposed [3]. One-dimensional (1D) gapped systems are completely classified at this point [4, 5, 6, 7]. A general gapped two-dimensional (2D) systems are expected to be described within the framework topological quantum field theory; see [8], for example.

This whole slew of different approaches raises a natural question. Why are there so many different approaches, and how can we ever be sure that the classification is complete? The main difficulty lies in identifying the correct framework. In the presence of interaction, one often needs to make a nontrivial assumption. The only exception so far is the one-dimensional (1D)

gapped system. Hastings' theorem [9] implies that any gapped 1D system obeys an area law. This subsequently implies that a matrix product state can approximate the ground state with a moderate bond dimension. It is this result from which a classification of quantum phases of 1D gapped system [4, 5, 6] follows.

However, in higher dimensions, an analog of Hastings' theorem is unknown. This is mainly because proving area law in 2D gapped systems remains challenging. Furthermore, even if area law turns out to be correct, the states that satisfy area law may not be well-approximated by an efficient tensor network [10]. These facts suggest that a classification program in 2D cannot merely mimic the classification program for 1D gapped systems. In fact, in any classification proposal based on tensor networks, there will always be a lingering question on whether we are not missing any unknown phases.

While it is widely believed at this point that topological quantum field theory (TQFT) describes all possible gapped phases in 2D, there is currently no rigorous argument that supports this belief. The existence of a three-dimensional (3D) gapped phase outside of the TQFT framework [11] shows that there may be gapped phases of matter that lie outside of the TQFT framework. Even if TQFT turns out to be the correct framework in 2D, understanding of where this framework comes from remains as an important fundamental problem.

Motivated by this state of affairs, we initiate a program in which a familiar set of axioms of TQFT can be derived from a seemingly innocuous assumption about entanglement. We show that some of the basic concepts of the algebraic theory of anyon [12], i.e., superselection sectors and fusion multiplicities, emerge from a familiar form of entanglement area law [13, 14]:

$$S(A) = \alpha\ell - \gamma, \tag{1}$$

where $S(A)$ is the von Neumann entropy of a simply connected region A, ℓ is the perimeter of A, and γ is a constant correction term[1] that only depends on the topology of A. The sub-leading correction, which vanishes in the $\ell \to \infty$ limit, is suppressed here.

We then show that our definition of the fusion multiplicities satisfies all the properties one would have expected from the algebraic theory of anyon. Again, these properties are derived from Eq. (1). Moreover, we further derive the following well-known formula:

$$\gamma = \ln \mathcal{D},$$

where $\mathcal{D}$ is the total quantum dimension of the anyon theory we defined. Our derivation is rigorous under the assumption (Eq. (1)) and is completely independent from the previous approaches, i.e., an approach based on an effective field theory description [13] and explicit calculations in exactly solvable models [14].

While our assumption is not as rigorous as Hastings' proof of the 1D area law, it is something that is widely accepted at this point. Therefore, we believe this would be a reasonable starting point to obtain a general understanding of gapped phases. A similar, but a markedly different starting point of our work would be the two axioms we have identified. These two axioms are entropic conditions on bounded-radius disks (Axiom **A0** and **A1** in Sec. 2.1). We can show that these two axioms follow from Eq. (1), but after that, we never use Eq. (1) explicitly. All of our results follow directly from the axioms.

---

[1]Nontrivial sectors, e.g., non-Abelian anyons and topological defects, can modify the constant term. It is a widely-adopted notation that γ is the constant term for a disk in the absence of nontrivial sectors.

In other words, the axioms of the anyon fusion theory can be derived from Eq. (1). The same conclusion follows from our axioms as well; see Axiom **A0** and **A1**. While the conclusion would be the same either way, we would like to advocate for the use of the axioms over Eq. (1) for the following reasons. The first reason is that Axiom **A0** and **A1** are assumed to hold on patches whose size is independent of the system size. Therefore, in principle, one can verify these axioms in time that scales linearly with the system size. Under a promise that the state is translation-invariant, the time can be reduced to a constant. In contrast, Eq. (1) is defined over length scales that are comparable to the system size. Verifying this assumption will incur an exponential computational cost. Secondly, in the continuum limit, the leading term of Eq. (1) depends on the ultraviolet-cutoff. On the other hand, the axioms manifestly cancel out this divergent piece.

Our framework is completely Hamiltonian-independent, in the sense that we only require the existence of a global state on the system satisfying the two local entropic constraints. This work is motivated from a number of recent observations: that local reduced density matrices of topological quantum phases often have a quantum Markov chain structure [15, 16, 17, 18, 19, 20, 21, 22, 23]. The key overarching concept is a convex set of density matrices introduced in [18], which is later rediscovered and studied under the name "information convex (set)" [22, 23]. Roughly speaking, this is a set of density matrices which are locally indistinguishable from some reference state. In our context, this reference state would be the ground state of some local Hamiltonian. However, we do not use the fact that the state is a ground state.

Our framework opens up a concrete route to classify gapped quantum phases without resorting to ad-hoc assumptions. In addition, we believe our framework is capable of answering a long-standing open question about topological phases. The question is if a single ground state contains all the data necessary to define a topological phase. Given that we can define a notion of topological charges and fusion multiplicities from a single ground state, progress may be made by using our framework. Our approach can be generalized to a broader context, e.g., to higher dimensions and to setups in which a topological defect [24] or a boundary is present [25]. We will discuss these applications in our upcoming work.

The rest of this paper is organized as follows. In Sec. 2, we specify our formal setup and summarize our main results. In Sec. 3, we prove fundamental properties of the information convex sets, which are the key to obtaining some of the axioms of the algebraic theory of anyon. We shall refer to this part of the full algebraic theory of anyon as the *anyon fusion theory* from now on. In Sec. 4, we define the notion of superselection sectors and fusion multiplicities in our framework and prove that the definition satisfies all the axioms of the anyon fusion theory. In Sec. 5, we show that the constant term γ in the area law equals the logarithm of the total quantum dimension. In Sec. 6, we conclude with a discussion.

2. Setup and Summary

Let us begin with a general setup and state our physical assumptions. Before we delve into the details, it will be instructive to discuss the physical system we have in mind. We are envisioning a gapped system in 2D, which consists of microscopic degrees of freedom, e.g., spins. We would like to coarse-grain these microscopic degrees of freedom so that we can view non-overlapping blocks of spins as gigantic "supersites," see Fig. 1. We can consider the limit in which the length scale of each block is large compared to the correlation length. We would like to define a sensible notion of distance between the subsystems as well as their topologies.
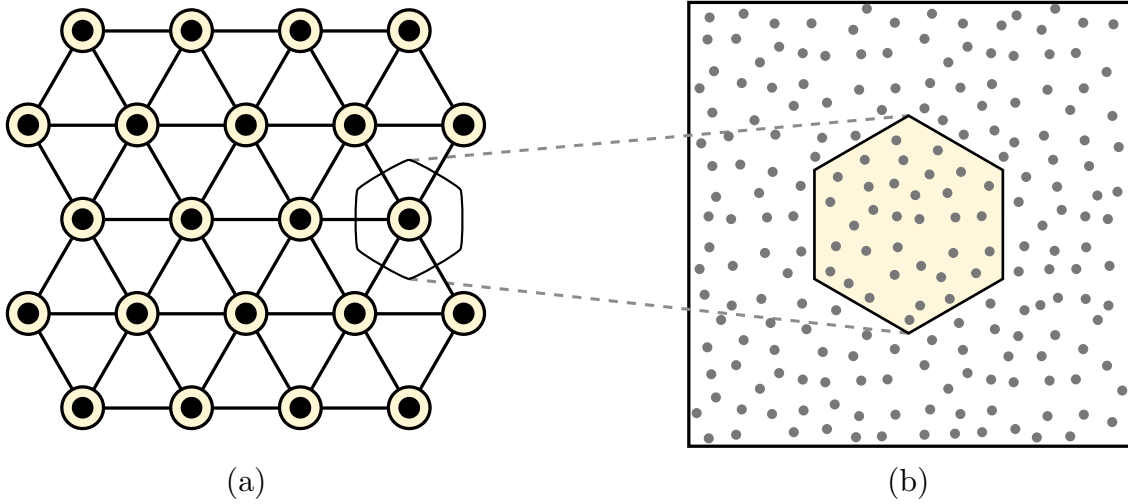
Figure 1: An illustration of the setup. (a) A graph $G = (V, E)$ with each vertex represents a "supersite" which contains a cluster of microscopic degrees of freedom, e.g. spins in real space. The edges encode the locality of the underlying physical system and it allows us to define a notion of topology for a set of vertices. (b) A zoomed-in depiction of a supersite. It contains a block of physical spins. The length scale of each block is a constant that is large compared to the correlation length.

More concretely, we can consider a quantum many-body spin system with a tensor product structure $\mathcal{H} = \otimes_{v \in V} \mathcal{H}_v$, where $\mathcal{H}_v$ is a finite-dimensional Hilbert space and V is a set of vertices of a finite graph $G = (V, E)$ defined on a 2D closed manifold[2]. By specifying the set of edges E, we can define a natural notion of distance on this graph (the graph distance). We denote the state space of $\mathcal{H}$ by $\mathcal{S}(\mathcal{H})$, which is the set of all density matrices on $\mathcal{H}$. We say operator O has support on $X \subseteq V$ if $O = O_X \otimes I_{V \setminus X}$ where $I_{V \setminus X}$ is the identity operator.

We assume that there is a partition of the manifold into simply connected subsystems so that each $v \in V$ is associated with one of these subsystems. Furthermore, there is an edge between $v_1, v_2 \in V$ if and only if the subsystems associated with the vertices are adjacent to each other. This assumption lets us define a notion of topology for a set of vertices. Without loss of generality, let U be a set of vertices and $\mathcal{U}$ be a union of the subsystems associated with the vertices in U. The topology of U is defined as the topology of $\mathcal{U}$. By construction, a single vertex is topologically a disk. However, a more general topology, e.g., an annulus, can be built out of a union of the vertices.

Throughout this paper, we consider a state $\sigma \in \mathcal{S}(\mathcal{H})$ satisfying two axioms shown below. We will call this state the *(global) reference state*. We always use σ to refer to the same reference state unless specified otherwise. We use σ_A to represent the reduced density matrix of σ on a region A. Physically interesting examples of the reference state are the ground states of gapped local Hamiltonians. However, our derivations are only based on the properties of the quantum state. Technically, we are allowed to assume the global state to be pure $\sigma = |\psi\rangle\langle\psi|$ without loss of generality. This is because one can always show the existence of a pure state which has the same local reduced density matrices as the given reference state (see Theorem B.1).

For the readers' convenience, we have summarized the key concepts in Table 1.[3]

---

[2]Our results work for both orientable and non-orientable manifolds provided that a reference state σ exists.
[3]We thank the referee for this suggestion.

Notation	Brief description	Reference
σ	The global reference state that satisfies Axiom **A0** and **A1**	N/A
σ_b	Reduced density matrix of σ over region b	
$\mu(r)$	A set of σ_b over b whose radius is smaller than r	Eq. (2)
$\Sigma(\Omega)$	The information convex set of a region Ω	Definition 3.1

Table 1: A list of notations and their descriptions.

2.1. Axioms

We start by defining a set of density matrices

$$\mu(r) = \{\sigma_b | b \in \mathcal{B}(r)\}, \tag{2}$$

where $\mathcal{B}(r)$ is a set of balls of radius less or equal to r and σ_b is the reduced density matrix of the reference state (σ) on b. Because r will be chosen to be a constant independent of the system size, we will simply denote $\mu(r)$ by μ. We will refer to the set of $b \in \mathcal{B}(r)$ as the set of μ-disks.

The axioms of our framework concern two entropic constraints on the set of μ-disks.[4] Let $S(\rho) = -\text{Tr}(\rho \ln \rho)$ be the von Neumann entropy of a state ρ. We assume that Axiom **A0** and **A1** hold for all μ-disks.

Axiom A0. For any $\sigma_b \in \mu$, for any configuration of subsystems $BC \subseteq b$ topologically equivalent to the one described in Fig. 2,

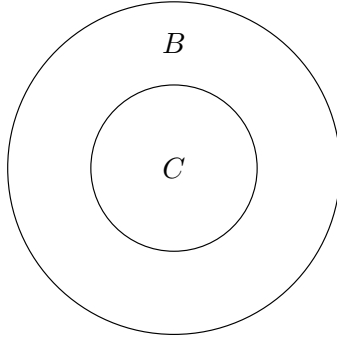$$S(\sigma_{BC}) + S(\sigma_C) - S(\sigma_B) = 0. \tag{3}$$

Figure 2: A disk is divided into its "core" (C) and outer boundary (B). The boundary (B) is chosen to be thick enough so that correlation between C and the complement of BC is negligible.

Axiom A1. For any $\sigma_b \in \mu$, for any configuration of subsystems $BCD \subseteq b$ topologically equivalent to the one described in Fig. 3,

$$S(\sigma_{BC}) + S(\sigma_{CD}) - S(\sigma_B) - S(\sigma_D) = 0. \tag{4}$$

---

[4]The two local entropic constraints **A0** and **A1** are originally proposed by one of us in Ref. [17] The first attempt at deriving the axioms of anyon theory from these conditions was presented in Ref. [18].
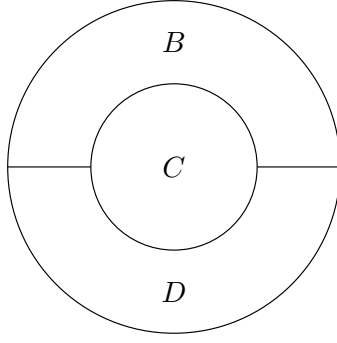
Figure 3: A disk is divided into its "core" (C) and its outer boundary, which is further divided into two pieces (B and D).

To see the physical meaning of these axioms, we observe that Eq. (3) implies

$$I(A : C)_\sigma \equiv (S_A + S_C - S_{AC})_\sigma = 0, \tag{5}$$

where A is contained in the complement of BC and $I(A : C)$ is the mutual information. This result follows from the strong subadditivity (SSA) of von Neumann entropy [26]. The mutual information is a measure of the total amount of bipartite correlation. Therefore, Axiom **A0** can be viewed as a formalization of the intuition that long-range two-point correlation vanishes in gapped ground states. However, note that the assumption is strictly stronger than the vanishing of mutual information itself. For instance, an infinite temperature Gibbs state satisfies Eq. (5) but does not satisfy **A0**. Eq. (4) implies that the *quantum conditional mutual information* vanishes:

$$I(A : C|B)_\sigma \equiv (S_{AB} + S_{BC} - S_B - S_{ABC})_\sigma = 0, \tag{6}$$

where A is contained in the complement of BCD. Again, **A1** is a stronger condition than Eq. (6).

While these axioms can be derived from Eq. (1), they are weaker assumptions and may hold in more general settings. Moreover, we expect these axioms to be a more well-defined way of formulating our assumptions compared to Eq. (1) because they manifestly get rid of the leading divergent term in the area law. In reality, we expect our assumptions to be satisfied only approximately, up to an error that decays exponentially with $\frac{r}{\xi}$, where r is the length scale of the subsystems and ξ is the correlation length. We believe our framework has a natural extension to these cases since every theoretical tool we use can be generalized to such situations (see also Discussion).

How we use these axioms will be explained later in this paper, starting from Section 3. However, for the hasty readers, we can briefly summarize our intuition as follows. The intuition behind **A0** is that the correlation between two sufficiently separated subsystems is negligible. This fact allows us to decouple two such subsystems without affecting either of them too much. The intuition behind **A1** is more profound and subtle. Let us consider the following example. Suppose we have a density matrix over AB and another density matrix over BC. Our goal is to "merge" these two density matrices; we want to construct a density matrix over ABC whose marginal density matrices on AB and BC are consistent with the given data. It is well-known that one cannot generally do this in quantum systems. For example, if AB is a maximally entangled state and BC is also (the same) maximally entangled state, there cannot be an extension of these density matrices into a single state acting on ABC. By assuming Axiom

A1, we can bypass this difficulty and merge the density matrices that are given to us. The majority of our key results follow from this merging process. In particular, we shall extensively use nontrivial identities involving the entropies of the merged subsystems.

2.2. Main results

Our framework employs the notion of *information convex set* [18, 22], the structure of which has been recently conjectured [23]. Loosely speaking, this is a set of reduced density matrices on a given subsystem; these density matrices are locally indistinguishable from the global reference state, but with an extra structure on this set in order to facilitate our analysis; see Sec. 3.2 for the details. Crucially, information convex sets are defined from a single state. As such, we do not need to invoke any assumption about the parent Hamiltonian.

The following results follow from Axiom **A0** and **A1**. Again, we do not assume anything about the parent Hamiltonian.

1. **Isomorphism theorem**
   Let A and B be two subsystems which can be smoothly deformed from one to another. We show that the information convex sets associated with A and B are *isomorphic* (Theorem 3.10). These sets can be mapped onto each other by a linear bijective map. Moreover, these maps preserve the distance and the entropy difference between the elements of the information convex set. Concretely, let Φ be one such map. Then $D(\rho, \rho') = D(\Phi(\rho), \Phi(\rho'))$ for any distance measure $D(\cdot, \cdot)$. Moreover, $S(\rho) - S(\rho') = S(\Phi(\rho)) - S(\Phi(\rho'))$. Therefore, the structure of the information convex set only depends on the topology of the region associated with it.

2. **A well-defined notion of topological charges**
   We show that the information convex set of an annulus is a simplex whose extreme points are orthogonal to each other. That is, any state ρ in the information convex set of an annulus must have the following form (Theorem 4.1):

$$\rho = \bigoplus_a p_a \, \sigma^a \,,$$

   where $\{p_a\}$ is a probability distribution over a finite set, and σ^a is a state only depending on the choice of the region. We define the label a as a topological charge/superselection sector of the system. In exactly solvable models (e.g., toric code), σ^a corresponds to a reduced density matrix of an annulus that surrounds a topological charge a. Different superselection sectors are perfectly distinguishable from each other. Moreover, the isomorphism theorem (see the first main result) implies that the charge is globally well-defined. We furthermore prove that for each sector, there exists a unique anti-sector (Sec. 4.3).

3. **Extracting fusion multiplicities**
   We completely characterize the information convex set of a 2-hole disk. We show that any element in this set can be expressed as

$$\rho = \bigoplus_{a,b,c \in \mathcal{C}} p_{ab}^c \, \rho^{abc},$$

   where $\mathcal{C}$ is the set of topological charges, $\{p_{ab}^c\}$ is a probability distribution, and $\{\rho^{abc}\}$ are mutually orthogonal quantum states labeled by a, b, and c. We show that, for each

choice of a, b and c, the set $\{\rho^{abc}\}$ forms a state space of a finite-dimensional Hilbert space; see Theorem 4.5. The dimensions of these Hilbert spaces are identified as the fusion multiplicities of the underlying anyon theory. In the context of the anyon theory, a and b label the two charges associated with the holes and c represents the total charge of the disk.

4. **Axioms of the anyon fusion theory**
   Our definition of the fusion multiplicities (see the third main result) satisfies all the axioms of the anyon fusion theory. The proof is based on the merging technique [21]. We derived several consistency relations by "merging" two density matrices and comparing the entropy of the reduced density matrices before and after the merge. An equation that relates these entropies leads to the axioms of the anyon fusion theory.

5. **Topological entanglement entropy**
   From our definition of the fusion multiplicities, we further show that the sub-leading constant term γ in Eq. (1) is the logarithm of the total quantum dimension. Unlike in [13], our derivation makes no assumption about the underlying effective field theory.

As a corollary, many of the anyon data can, in principle, be extracted from the local information of a single ground state. Importantly, we do not need to assume that the low-energy excitations can be described by a unitary modular tensor category. (Instead, we need **A0** and **A1**.) In contrast, an oft-used numerical method [27] requires global information of multiple ground states and the assumption that the system is described by a modular tensor category. Also, Haah [28] has attempted to extract the topological S-matrix from a single ground state. Haah's final argument that his invariant matches the S-matrix, however, still relies on the modular tensor category assumptions.

Finally, we can also define a unitary string operator that creates an anyon pair. The support of the string can deform freely as long as the endpoints are fixed (Appendix H).

3. Axiom extension, Information convex set and Isomorphism Theorem

We have proclaimed in Section 2.2 that we can, among many things, establish a *globally well-defined* notion of topological charge. Because our axioms (Axiom **A0** and **A1**) are assumed to hold only on bounded-radius disks, the fact that such a notion can even exist in the first place is not obvious at all.

In order to explain how this works, we choose the main theme of this section to be "local to global." Starting from our local axioms (Axiom **A0** and **A1**), we will see how we can infer some of the global statements we made in Section 2.2. The first step in this direction lies in extending our local axioms to larger regions. Note that our axioms are assumed to hold only on (bounded-sized) μ-disks. In Section 3.1, we will show that the axioms in fact hold on arbitrarily large disks.

Then we move onto a concept that plays the central role in this paper: information convex set. In order to define an information convex set, we need to fix a subsystem and consider a set of states on this subsystem which are locally indistinguishable from the given reference state. Let us refer to this subsystem as Ω' and the set of states as $\tilde{\Sigma}(\Omega')$.

The set $\tilde{\Sigma}(\Omega')$ is *not* the information convex set, but we are getting close. For every state in $\tilde{\Sigma}(\Omega')$, trace out part of the subsystem that lies at the "outer edge" of Ω'; see Fig. 6. This way,

we get a set of reduced density matrix over a subset of Ω'. Let this subset be Ω. What we have obtained is the information convex set of Ω, which we refer to as $\Sigma(\Omega)$. As one can see, this is quite an involved definition! However, in our opinion, we have a sensible justification. Why we choose $\Sigma(\Omega)$ over $\tilde{\Sigma}(\Omega')$ as a main actor of our story will be the main topic of discussion in Section 3.2.

In Section 3.3, we derive the key technical result of this section: the *isomorphism theorem* (Theorem 3.10). This theorem establishes an equivalence between information convex sets for topologically equivalent subsystems connected by a path.

This section may seem a bit abstract at first, so it will be useful to have a concrete physical picture in mind. Consider a topologically ordered medium [29] that can host anyons [30, 31]. It is well-known that, within such a medium, there is a *globally* well-defined notion of superselection sectors and fusion space. This is because one can adiabatically transport anyons from different regions and compare them. For example, suppose we have two anyons that are well-separated from each other. How would we able to decide if they are the same anyon type or not? One can adiabatically bring either of the anyons to some fixed location and perform an Aharonov-Bohm type interference experiment. If the underlying anyon theory is unitary, there must be some experiment that can distinguish two different types of anyons.

In the above illustrative example, we observed that there is a physical process by which we can compare anyons or even a collection of anyons with each other. This comparison is possible if one can transport these objects from one place to another. The isomorphism theorem formalizes this transportation process in a Hamiltonian-independent manner.

3.1. Extension of the axioms

In this Section, we show that our axioms (Axiom **A0** and **A1**) hold on arbitrarily large disks. Results in this section have been discussed in Ref. [17]. In order to understand this section, one should become familiar with the notion of *quantum Markov state* [32, 15]. A quantum Markov state is a tripartite state, say over subsystems A, B, and C, such that

$$I(A:C|B) = S(\rho_{AB}) + S(\rho_{BC}) - S(\rho_B) - S(\rho_{ABC})$$

is equal to 0. These states enjoy several nontrivial properties, which we make extensive use of.

The following two lemmas will be used frequently.

Lemma 3.1. *Let ρ_{ABC} and σ_{ABC} be density matrices such that (1) $\rho_{AB} = \sigma_{AB}$ and $\rho_{BC} = \sigma_{BC}$; (2) $I(A:C|B)_\rho = I(A:C|B)_\sigma = 0$. Then $\rho_{ABC} = \sigma_{ABC}$.*

The proof follows from Ref. [33]. Suppose we know σ_{AB} and σ_{BC} for the partition in Fig. 4. Then **A1** implies $I(A:C|B)_\sigma \leq (S_{BC} + S_{CD} - S_B - S_D)_\sigma = 0$. This lemma implies that the state σ_{ABC} is uniquely determined by its reduced density matrices. Moreover, there is a CPTP map which can recover σ_{ABC} from its reduced density matrices. This map is known as the Petz recovery map.

Lemma 3.2. *(Petz recovery map [33]) For any tripartite state ρ_{ABC}, $I(A:C|B)_\rho = 0$ if and only if*

$$\rho_{ABC} = \mathcal{E}^\rho_{B \to BC}(\rho_{AB})\,, \tag{7}$$

where $\mathcal{E}^\rho_{B \to BC}$ (the Petz recovery map) is defined as

$$\mathcal{E}^\rho_{B \to BC}(X_B) = \rho_{BC}^{\frac{1}{2}} \rho_B^{-\frac{1}{2}} X_B \rho_B^{-\frac{1}{2}} \rho_{BC}^{\frac{1}{2}}\,.$$
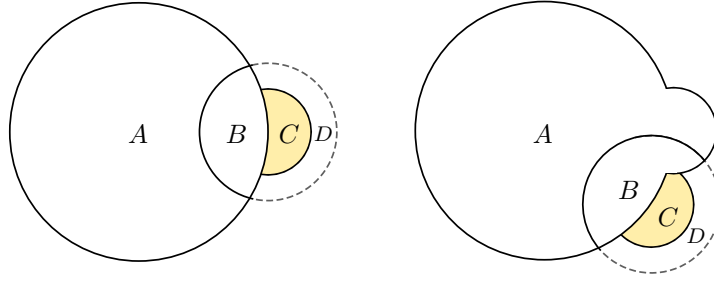
9

Figure 4: An illustration of the growth procedure of a disk from AB to ABC. Here A can be large and BCD is contained in a μ-disk in a manner similar to Fig. 3.
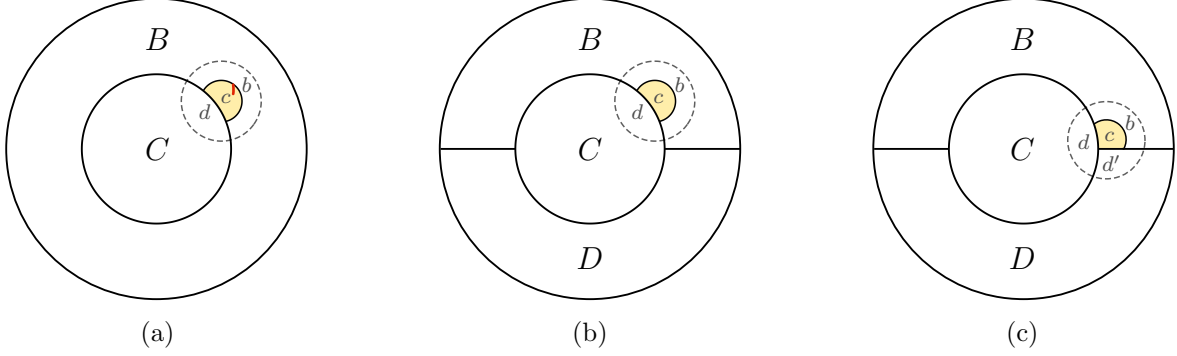

(a)   (b)   (c)

Figure 5: The extension of the axioms. A disk is divided into either BC or BCD. A μ-disk is on a smaller length scale, i.e., the small dashed circle surrounding the colored region. These figures represent three ways of enlarging C by a small step. (a) $bc \subseteq B$ and $d \subseteq C$; (b) $bc \subseteq B$ and $d \subseteq C$; (c) $bc \subseteq B$, $d \subseteq C$ and $d' \subseteq D$.

Now we are in a position to state the main result of this section. In Proposition 3.3, we show that Axioms **A0** and **A1** hold for all disks. By assumption, these axioms hold for μ-disks. The nontrivial part of the statement is that the axioms hold at a larger length scale.

Proposition 3.3. *For a reference state satisfying axioms **A0** and **A1**, the entropic conditions Eq. (3) and Eq. (4) are satisfied on all larger disk-like subsystems.*

Proof. We shall extensively use the following two inequalities:

$$S_{BC} + S_C - S_B \;\geq\; S_{BB'C} + S_C - S_{BB'}, \tag{8}$$
$$S_{BC} + S_{CD} - S_B - S_D \;\geq\; S_{BB'C} + S_{CDD'} - S_{BB'} - S_{DD'}. \tag{9}$$

Both of them follow from SSA.

Let us first extend Axiom **A0** to a larger scale. Without loss of generality, consider a disk and its subsystem B and C, shown in Fig. 5(a). Assume that BC is not contained in any μ-disk. One can consider a sequence of (enlarged) subsystems to obtain this disk (BC) from a μ-disk.

For this purpose, it suffices to consider the following two moves. The first move is to enlarge B while keeping C fixed. The second move is to enlarge C while keeping BC fixed. Our goal is to show that, for both of these steps, the linear combination of entropy $S_{BC} + S_C - S_B$ is non-increasing. The first move preserves Eq. (3) because of Eq. (8) where we set $B \sqcup B'$ to be the enlarged B. To understand why the second move preserves Eq. (3), consider a deformation depicted in Fig. 5(a). We need to show $S_{BC} + S_C - S_B$ is non-increasing when we deform C to include the colored region of Fig. 5(a). A variation of Eq. (9), which involves the μ-disk bcd, is

$$S_B + S_{Cc} - S_{B\backslash c} - S_C \leq S_{bc} + S_{cd} - S_b - S_d = 0. \qquad \textcolor{red}{(9)} \tag{10}$$

10

Therefore, both moves preserve Eq. (3). Because any disk can be enlarged from a μ-disk by applying a sequence of these moves, Axiom **A0** holds for any disk.

Now, let us move on to Axiom **A1**. Similarly, we can obtain a larger disk, say BCD, from a μ-disk by making use of the following moves. As usual, the choice of the subsystems are similar to the one used in Axiom **A1**. The first move is to enlarge B and D while keeping C fixed. The second move is to enlarge C while keeping BCD fixed.

Let us show that the entropic condition (4) holds at every step. The first move preserves Eq. (4). This is because inequality (9) can be applied to enlarged region BB' and DD'. Notably, B' and D' can be chosen to touch each other. This allows us to show Eq. (4) for the arbitrary BB' and DD' satisfying the topology condition, by suitably choosing B and D in the initial μ-disk. For the second move, we can consider the following sequence of small steps shown in Fig. 5(b)(c) as well as the same steps in which the choice of B and D are switched. Here is the justification of each small step. It is sufficient to justify the steps shown in Fig. 5(b) and (c). Regarding the enlargement process described in Fig. 5(b), we have

$$S_B + S_{CDc} - S_{B\backslash c} - S_{CD} \le S_{bc} + S_{cd} - S_b - S_d = 0. \tag{11}$$

For the enlargement process in Fig. 5(c), we have

$$S_B + S_{CDc} - S_{B\backslash c} - S_{CD} \le S_{bc} + S_{cdd'} - S_b - S_{dd'} = 0. \tag{12}$$

Therefore, both Eq. (3) and Eq. (4) hold at larger length scales. This completes the proof. $\qquad\square$

3.2. Information convex set

Care must be taken in reading this section, for we are about to explain the most important concept in this paper: information convex set. Let us begin with some definitions. We say that two density matrices ρ and ρ' are *consistent* with each other if they have identical density matrices on the overlapping support, i.e., $\rho_A = \rho'_A$ where A is the intersection of the support of ρ and that of ρ', and denote it by

$$\rho \overset{c}{=} \rho'.$$

For each subsystem $\Omega \subseteq V$, we can define its information convex set by first considering the "thickening" of Ω. Let $\partial\Omega \subseteq V$ be a set of vertices that are (graph) distance 1 away from Ω. If the subsystem $\Omega \sqcup \partial\Omega$ is topologically equivalent to Ω, we refer to that subsystem as a thickening of Ω. Let there exists a thickening of Ω; we denote the thickening as Ω'. Intuitively, Ω' is a subsystem that can be smoothly deformed into Ω such that $\Omega'\backslash\Omega$ is a boundary of Ω with a non-vanishing thickness. This thickness must be chosen to be sufficiently large compared to the correlation length of the underlying state.

Now, we are in a position to define the information convex set.

Definition 3.1. Let $\Omega \subseteq V$ be a subsystem and let Ω' be the thickening of Ω. The information convex set of Ω is defined as

$$\Sigma(\Omega) \equiv \left\{\rho_\Omega | \rho_\Omega = \text{Tr}_{\Omega'\backslash\Omega} \; \rho_{\Omega'}, \rho_{\Omega'} \in \tilde{\Sigma}(\Omega')\right\}, \tag{13}$$

where $\tilde{\Sigma}(\Omega')$ is defined as

$$\tilde{\Sigma}(\Omega') = \{\rho_{\Omega'} | \rho_{\Omega'} \overset{c}{=} \sigma_b, \quad \forall \sigma_b \in \mu\}.$$

This definition was first advocated in [18]. A related Hamiltonian-based definition was introduced in Ref. [22].

One may have thought that $\tilde{\Sigma}(\Omega)$ would be a more natural definition of information convex set. However, it turns out that the local indistinguishability constraints are insufficient to fix some of the degrees of freedom at the boundaries of Ω. For this reason, it is possible for some of the elements in $\tilde{\Sigma}(\Omega)$ to possess entanglement between two distant μ-disks in the vicinity of the boundary of Ω. Moreover, even a multipartite correlation may arise along the boundary. These exceptional elements cannot be regarded as the reduced state of a state on a larger support that is also locally indistinguishable from the reference state. In particular, they contain unnecessary extra information that has nothing to do with the anyon theory. The additional partial trace operation in the definition of $\Sigma(\Omega)$ removes such irrelevant correlations around the boundaries.

As it stands, the definition of information convex set is independent of our axioms (Axiom **A0** and **A1**). However, once we impose these axioms, one can come up with a more restrictive definition of information convex set that does not involve Ω'. This definition is formulated in Definition C.1. Proposition C.4 establishes the equivalence of the two.
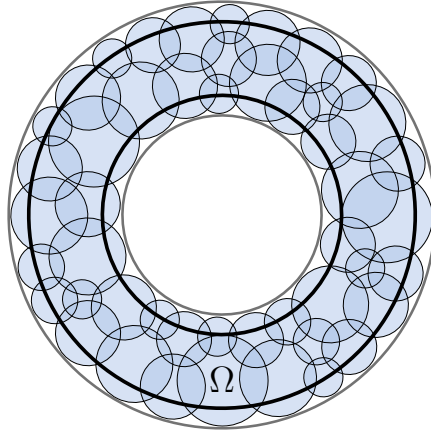

Figure 6: This figure is a schematic depiction of regions involved in the definition of information convex set $\Sigma(\Omega)$. Here Ω is the annulus between the black circles, annulus $\Omega' \supseteq \Omega$ is the thickening of Ω, i.e., the region between gray circles. Any element in $\tilde{\Sigma}(\Omega')$ is consistent with the reference state σ on every μ-disk contained in Ω' (the blue disks). We chose Ω to be an annulus for illustration purposes. Other topologies are allowed as well.

Below, we derive properties of the information convex set. First of all, information convex set is a convex subset of the state space. This follows straightforwardly from the definition.

Proposition 3.4. *For any nonempty $\Omega \subseteq V$, $\Sigma(\Omega)$ is a nonempty finite-dimensional compact convex set. Furthermore, if $\Omega \subseteq \Omega'$ and $\rho_{\Omega'} \in \Sigma(\Omega')$, then $\mathrm{Tr}_{\Omega' \setminus \Omega} \rho_{\Omega'} \in \Sigma(\Omega)$.*

Proof. The state space of a finite-dimensional Hilbert space is a finite-dimensional compact convex set. The partial trace operation $\mathrm{Tr}_{\Omega' \setminus \Omega}$ is linear and bounded. Therefore, the image of the partial trace is compact and convex. The last statement is a direct consequence of Definition 3.1. $\qquad\square$

Secondly, we show that the information convex set of a disk contains a single element. We use this result throughout this paper, primarily for identifying the uniqueness of the global state

on a sphere and for defining the vacuum sector.[5]

Proposition 3.5. *For any disk-like subsystem ω, we have*

$$\Sigma(\omega) = \{\sigma_\omega\}, \tag{14}$$

where $\sigma_\omega \equiv \mathrm{Tr}_{\bar{\omega}} |\psi\rangle\langle\psi|$ and $\bar{\omega}$ is the complement of ω.

Proof. The idea is illustrated in Fig. 7, which involves a repetition of the "growth process" described in Fig. 4. Starting from a μ-disk, one can grow the disk until the enlarged disk covers ω.

Let us get into the details. First, recall that $\Sigma(\omega)$ is nonempty since it contains σ_ω. Let us pick another element $\sigma'_\omega \in \Sigma(\omega)$. According to Definition 3.1, σ_ω and σ'_ω must be identical on a μ-disk. Therefore, we can repeatedly use the conditional independence condition, i.e., $I(A : C|B)_\sigma = I(A : C|B)_{\sigma'} = 0$, to show that the reduced density matrices of the two states are identical on increasingly larger disks; see Lemma 3.1. The disk can grow to the point it covers ω. Thus, we can conclude that $\sigma'_\omega = \sigma_\omega$. $\square$
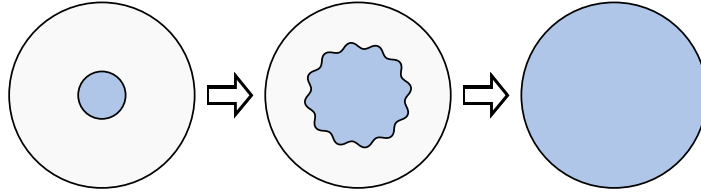

Figure 7: A μ-disk can grow until it covers a larger disk ω.

More generally, we can consider the information convex set of a subsystem which is *not* topologically a disk. We shall discuss the structure of such set in Section 4. The following fact will be important for that discussion. Let Ω be an arbitrary subsystem. We show that, for any element in the information convex set of Ω, its reduced density matrix on a disk-like region is equal to the reduced density matrix of the reference state on the same region.

Proposition 3.6. *Any state $\rho_\Omega \in \Sigma(\Omega)$ satisfies*

$$\mathrm{Tr}_{\Omega\backslash\omega}\, \rho_\Omega = \sigma_\omega \tag{15}$$

on any disk-like subsystem $\omega \subseteq \Omega$.

Proof. By Proposition 3.4, for any state $\rho_\Omega \in \Sigma(\Omega)$, $\mathrm{Tr}_{\Omega\backslash\omega}\rho_\Omega$ is an element of $\Sigma(\omega)$. By Proposition 3.5, the set $\Sigma(\omega)$ contains only one element. $\square$

As a side note, we can always choose the reference state to be pure even if the initially given reference state is not. This is because the information convex set of the entire system is the state space of a finite-dimensional Hilbert space (see Theorem B.1); one can simply choose one of the pure states in this space to be the reference state. This pure reference state, restricted to the μ-disks, would be consistent with the initially given reference state.

An interesting special case is when the underlying manifold is a sphere (S^2). In this case, the global state is unique and thus pure.

---

[5] As a side note, let us point out that this result is based only on **A1**. Therefore, the uniqueness of the element for the information convex set of a disk can hold more generally, even when **A0** breaks down.

Proposition 3.7. $\Sigma(S^2) = \{|\psi\rangle\langle\psi|\}$.

Proof. Let us begin by setting up an appropriate set of subsystems. Partition the sphere into three subsystems, A, B, and C. We choose C to be a disk and B to be an annulus that surrounds C; see Fig. 2. The complement of BC, i.e., A, is a disk.

Without loss of generality, suppose we have two states $\rho_{ABC}, \sigma_{ABC} \in \Sigma(S^2)$. We show that they must be equal. By Proposition 3.5, their reduced density matrices are identical on AB and BC. SSA implies that for both states,

$$I(A:C|B) \leq S(BC) + S(C) - S(B),$$

where we suppressed the dependence on ρ and σ. Either way, the right hand side is 0 because our axioms hold at any scale; see Proposition 3.3. Because ρ_{ABC} and σ_{ABC} both have vanishing conditional mutual information (conditioned on B) and have identical density matrices (on AB and BC), we can use Lemma 3.1. Therefore, $\rho_{ABC} = \sigma_{ABC}$.

By Theorem B.1, $\Sigma(S^2)$ is isomorphic to a state space of a finite-dimensional Hilbert space. Because this set has a unique element, the global reference state must be pure. $\square$

3.3. Elementary steps and isomorphism theorem

In this section, we establish an isomorphism between two information convex sets. This isomorphism exists if the subsystems associated with the two sets are topologically equivalent and can be smoothly deformed into each other. More precisely, in order to establish an isomorphism, we require that the two subsystems to be connected by a *path* (Defintion 3.2).

Crucial to this analysis is the concept of state merging. Suppose we have two quantum states ρ and σ which share an overlapping support and consistent. The question is whether one can consistently "sew" them together. Namely, can we find a state which is consistent with both ρ and σ? This is known as the quantum marginal problem. In general, even deciding whether there is such a state or not is known to be extremely difficult [34]. There are several nontrivial necessary conditions [16, 17], but sufficient conditions are rare.

However, one of us has found a nontrivial sufficient condition [21]. We restate the result for the reader's convenience.

Lemma 3.8. *(Merging Lemma [21]) Given a set of density matrices $\mathcal{S} \equiv \{\rho_{ABC}\}$ and a density matrix σ_{BCD} such that $\rho_{BC} = \sigma_{BC}$ and*

$$I(A:C|B)_\rho = I(B:D|C)_\sigma = 0, \quad \forall \rho \in \mathcal{S}, \tag{16}$$

there exists a unique set of "merged" states $\{\tau^\rho_{ABCD} = \mathcal{E}^\sigma_{C \to CD}(\rho_{ABC})\}$ which satisfy the following properties.

(1) τ^ρ is consistent with ρ and σ, i.e.

$$\tau^\rho_{ABC} = \rho_{ABC} \quad and \quad \tau^\rho_{BCD} = \sigma_{BCD}. \tag{17}$$

(2) Vanishing conditional mutual information,

$$I(A:CD|B)_{\tau^\rho} = I(AB:D|C)_{\tau^\rho} = 0, \quad \forall \rho. \tag{18}$$

14

(3) The conservation of von Neumann entropy difference, for arbitrary $\rho, \rho' \in \mathcal{S}$,

$$S(\tau_{ABCD}^{\rho}) - S(\tau_{ABCD}^{\rho'}) = S(\rho_{ABC}) - S(\rho'_{ABC}). \qquad (19)$$

The significance of this lemma lies on the fact that one can guarantee the existence of a global state from a (relatively) local information. What is given to us are the density matrices over ABC and BCD, together with the conditions that can be verified on ABC and BCD. In particular, these conditions can be directly verified from the given states. Once the conditions are verified, one can guarantee the existence of a state over $ABCD$ that is consistent with the given density matrices.

The merging lemma (Lemma 3.8), together with our axioms (Axiom **A0** and **A1**) underpin the majority of our technical work. The interplay between the two is what allows us to start from strictly local information and conclude something nontrivial at a larger scale. Roughly speaking, such analysis is carried out as follows. Our axioms allow us to upper bound certain conditional mutual information by 0. We can then apply this fact to Lemma 3.8 repeatedly to merge (many) density matrices. In particular, we can merge elements of multiple information convex sets into an element of yet another information convex set. This not only allows us to smoothly deform the boundary of a subsystem (Fig. 8) but also allows us to consider merging processes with nontrivial topology changes; see Sec. 4.3.

Now, we are in a position to prove the *isomorphism theorem*. This theorem establishes an equivalence between two information convex sets whose underlying subsystems can be smoothly deformed into each other. We first explain a method to establish the equivalence when one subsystem is merely an "infinitesimal" deformation of the other. Of course, the word infinitesimal should not be taken literally, because we are considering a quantum many-body system on a lattice. What we mean is that one subsystem can be obtained from the other by either attaching or removing a region whose size is comparable to that of the μ-disks.

Imagine zooming into the region in which this deformation occurs. Without loss of generality, we can consider two subsystems $\Omega = ABC$ and $\Omega' = ABCD$ depicted in Fig. 8, where CD is contained in a μ-disk. We can show that there exists a bijection between $\Sigma(\Omega)$ and $\Sigma(\Omega')$. This is the content of Proposition 3.9. The proof is in Appendix C.
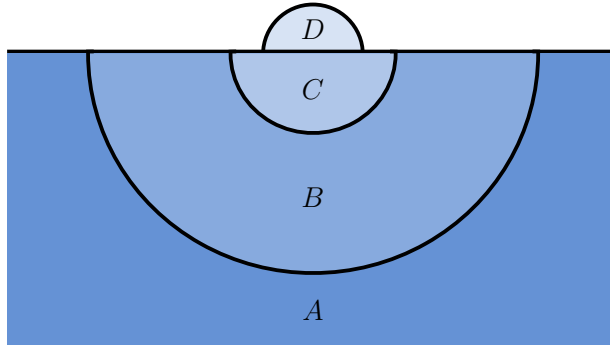

Figure 8: Two slightly different subsystems $\Omega = ABC$ and $\Omega' = ABCD$ (only a part of A is shown). BCD is a disk and CD is contained in a μ-disk. The topology of A can be arbitrary. We require B to be thick enough, so that A and D are separated by at least $2r + 1$. (Recall that r is the scale used in defining the μ-disks; see Eq. (2).)

Proposition 3.9. *Consider the partition in Fig. 8. (Note: A and D are assumed to be separated by at least $2r+1$ in Fig. 8.) Let the domain of Tr_D and $\mathcal{E}^{\sigma}_{C \to CD}$ to be $\Sigma(\Omega')$ and $\Sigma(\Omega)$ respectively.*

$$\begin{aligned}
\mathrm{Im}\,\mathrm{Tr}_D &\subseteq \Sigma(\Omega) \\
\mathrm{Im}\,\mathcal{E}^{\sigma}_{C \to CD} &\subseteq \Sigma(\Omega').
\end{aligned} \tag{20}$$

Moreover, for all $\rho_\Omega \in \Sigma(\Omega)$ and $\rho_{\Omega'} \in \Sigma(\Omega')$,

$$\mathrm{Tr}_D \circ \mathcal{E}^{\sigma}_{C \to CD}(\rho_\Omega) = \rho_\Omega, \tag{21}$$

$$\mathcal{E}^{\sigma}_{C \to CD} \circ \mathrm{Tr}_D(\rho_{\Omega'}) = \rho_{\Omega'}. \tag{22}$$

In particular, Tr_D and $\mathcal{E}^{\sigma}_{C \to CD}$ are bijections.

As a side note, Proposition 3.9 further implies that the isomorphism $\mathcal{E}^{\sigma}_{C \to CD} : \Sigma(\Omega) \to \Sigma(\Omega')$ is independent of the choice of B and C. To see why, consider two choices, say $BC \subseteq \Omega$ and $B'C' \subseteq \Omega$. Consider two maps $\mathcal{E}^{\sigma}_{C \to CD}$ and $\mathcal{E}^{\sigma}_{C' \to C'D}$. Suppose there exists an element of $\Sigma(\Omega)$ which, under these two maps, is mapped into two different elements of $\Sigma(\Omega')$. Upon applying Tr_D to these two states, by Proposition 3.9, both states must be mapped back to the same state. This implies that Tr_D is not injective, which contradicts Proposition 3.9.

There are two simple corollaries of Proposition 3.9, which will become handy in the rest of the paper. First, the bijective map Tr_D and $\mathcal{E}^{\sigma}_{C \to CD}$ preserves the distance between two states.

Corollary 3.9.1. *(Distance preservation) Let $\rho_\Omega, \rho'_\Omega \in \Sigma(\Omega)$ and $\rho_{\Omega'}, \rho'_{\Omega'} \in \Sigma(\Omega')$. For any distance measure $D(\cdot, \cdot)$ between quantum states,*

$$D(\rho_\Omega, \rho'_\Omega) = D(\mathcal{E}^{\sigma}_{C \to CD}(\rho_\Omega), \mathcal{E}^{\sigma}_{C \to CD}(\rho'_\Omega)) \tag{23}$$

and

$$D(\rho_{\Omega'}, \rho'_{\Omega'}) = D(\mathrm{Tr}_D(\rho_{\Omega'}), \mathrm{Tr}_D(\rho'_{\Omega'})). \tag{24}$$

Proof. For both identities, the proof is practically identical. So we only discuss the proof of the first identity. Because both $\mathcal{E}^{\sigma}_{C \to CD}$ and Tr_D are CPTP maps, distance is nonincreasing under these maps. Therefore,

$$\begin{aligned}
D(\rho_\Omega, \rho'_\Omega) &\geq D(\mathcal{E}^{\sigma}_{C \to CD}(\rho_\Omega), \mathcal{E}^{\sigma}_{C \to CD}(\rho'_\Omega)) \\
&\geq D(\mathrm{Tr}_D \circ \mathcal{E}^{\sigma}_{C \to CD}(\rho_\Omega), \mathrm{Tr}_D \circ \mathcal{E}^{\sigma}_{C \to CD}(\rho'_\Omega)) \\
&= D(\rho_\Omega, \rho'_\Omega),
\end{aligned} \tag{25}$$

where in the last line we used Proposition 3.9. Therefore, $D(\rho_\Omega, \rho'_\Omega) = D(\mathcal{E}^{\sigma}_{C \to CD}(\rho_\Omega), \mathcal{E}^{\sigma}_{C \to CD}(\rho'_\Omega))$. $\square$

We note that, while we only considered distances between two quantum states, the same proof applies to the preservation of the fidelity $F(\rho, \tau) = (\mathrm{Tr}\sqrt{\sqrt{\rho}\,\tau\,\sqrt{\rho}})^2$. While fidelity is not a distance measure, its behavior is monotonic under application of CPTP maps. Therefore, the proof of Corollary 3.9.1 still applies.

In fact, we can show more. Even the entropy difference is preserved under Tr_D and $\mathcal{E}^{\sigma}_{C \to CD}$. The proof of this statement is a simple byproduct of the proof of Proposition 3.9. The proof follows immediately from property (3) of Lemma 3.8.

Corollary 3.9.2. *(Entropy difference preservation) Let $\rho_\Omega, \rho'_\Omega \in \Sigma(\Omega)$ and $\rho_{\Omega'}, \rho'_{\Omega'} \in \Sigma(\Omega')$. The von Neumann entropies satisfy*

$$S(\rho_\Omega) - S(\rho'_\Omega) = S(\mathcal{E}^\sigma_{C \to CD}(\rho_\Omega)) - S(\mathcal{E}^\sigma_{C \to CD}(\rho'_\Omega)) \tag{26}$$

and

$$S(\rho_{\Omega'}) - S(\rho'_{\Omega'}) = S(\text{Tr}_D(\rho_{\Omega'})) - S(\text{Tr}_D(\rho'_{\Omega'})). \tag{27}$$

Therefore, given a subsystem Ω and its information convex set $\Sigma(\Omega)$ we can establish a bijection between $\Sigma(\Omega)$ and $\Sigma(\Omega')$ where Ω' is a slight deformation of Ω. In order to apply Proposition 3.9, Ω' must be topologically equivalent to Ω. We refer to the process of subtracting/adding a disk-like region to a given subsystem as the *elementary step* of the deformation.

The isomorphism between two information convex sets can be established by repeating these elementary steps. However, we have to be careful on two points. First, for two given topologically equivalent subsystems, there can be more than one way to deform one to the other. Second, even if the underlying subsystems are topologically equivalent, there may not be a smooth deformation between the two. As a trivial example, suppose we have two spheres. We can place two subsystems on each of these spheres. Even if these subsystems are topologically equivalent to each other, there is no sequence of subsystems that smoothly deforms one to the other. Even on a connected space, one cannot make such a statement; see Fig. 10.

Therefore, these (potentially different) isomorphisms must be labeled by their paths. Let us formalize this notion below.

Definition 3.2. (Path) A finite sequence of subsystems $\{\Omega^t\}$ with $t = i/N$ and $i = 0, 1, 2, \cdots, N$, (N is a positive integer), is a *path* connecting Ω^0 and Ω^1 if each pair of nearby subsystems in the sequence are related by an elementary step of deformation, illustrated in Fig. 8.

Because a path is built up from elementary steps, we obtain the following theorem.

Theorem 3.10. *(Isomorphism Theorem) If Ω^0 and Ω^1 are connected by a path $\{\Omega^t\}$, then there is an isomorphism*

$$\Phi_{\{\Omega^t\}} : \quad \Sigma(\Omega^0) \to \Sigma(\Omega^1) \tag{28}$$

uniquely determined by the path $\{\Omega^t\}$. Moreover, it preserves the distance and the entropy difference between elements

$$D(\rho, \sigma) = D\left(\Phi_{\{\Omega^t\}}(\rho), \Phi_{\{\Omega^t\}}(\sigma)\right) \tag{29}$$

$$S(\rho) - S(\sigma) = S\left(\Phi_{\{\Omega^t\}}(\rho)\right) - S\left(\Phi_{\{\Omega^t\}}(\sigma)\right), \tag{30}$$

where $D(\cdot, \cdot)$ is any distance measure which is non-increasing under CPTP-maps.

We omit the proof since it straightforwardly follows by applying Proposition 3.9 repeatedly. For any path $\{\Omega^t\}$, we can define an inverted path $\{\Omega^{1-t}\}$ which reverses the sequence of subsystems. This leads to the inverse isomorphism $\Phi_{\{\Omega^{1-t}\}} : \Sigma(\Omega^1) \to \Sigma(\Omega^0)$.

Generally speaking, different paths may give rise to different isomorphisms. That is, under two different isomorphisms, an element of the information convex set may be mapped to two distinct elements. However, sometimes, we merely need the existence of an isomorphism. In that case, we will use a notation

$$\Sigma(\Omega^0) \cong \Sigma(\Omega^1)$$

to indicate the existence of such an isomorphism. Under this condition, any distance measure and entropy difference is preserved; see Eqs. (29) and (30).

4. Fusion data from information convex sets

The isomorphism theorem (Theorem 3.10) guarantees that the structure of the information convex set only depends on the topology, as long as the underlying subsystems can be smoothly deformed from one to another along some path.

We now focus on how to extract the information of the topological charges and the corresponding fusion rules from the information convex set. We do this by studying how the geometry of the information convex set depends on the topology of the underlying subsystem. We then use the merging technique (Lemma 3.8) to relate subsystems with different topologies and obtain several consistency equations. We then define the fusion rules and show that they satisfy all the constraints expected from the known algebraic theory of anyon [12]. The result of this study is summarized in Table 2.

Physical data	Number of holes
Superselection sectors	1
Fusion multiplicities	2
Axioms of the fusion theory	$1, 2, 3, 4$ (merging)

Table 2: Physical data that can be extracted from disks with different number of holes.

4.1. Superselection sectors/topological charges

Let us define a notion of superselection sectors, which is one of the key ingredients of the algebraic theory of anyon [12]. Historically, the notion of superselection sectors was introduced in the context of local field theory; see [35, 36]. In the context of topologically ordered systems which is most relevant to our discussion, a nontrivial superselection sector corresponds to an anyon type that cannot be created by any local operator.

There are several recent attempts to rigorously formulate the superselection sectors based on operator algebra assumptions. One approach is based on the operator algebra on an annulus [28, 37], and another approach is based on the operator algebra on a cone-like subsystem of an infinite lattice [38, 39]. Our approach to characterize the superselection sectors is similar to the one based on the operator algebra on annuli. However, these two approaches differ in their assumptions and their range of validity.

We will identify a well-defined information-theoretic object and find that this object coincides with the conventional notion of superselection sectors in anyon theory. Importantly, we find that the information convex set of an annulus forms a simplex (Theorem 4.1). The simplex has a finite number of extreme points. Moreover, these extreme points are orthogonal to each other. See Fig. 9(c) for an illustration. (See Appendix A.1 for the general definition of extreme points.) We will define these extreme points as the superselection sectors.

Theorem 4.1. *(Simplex Theorem) For an annulus X, the information convex set is the convex hull of a finite number of orthogonal extreme points, $\{\sigma_X^a\}$, i.e.*

$$\Sigma(X) = \left\{ \rho_X \, \middle| \, \rho_X = \bigoplus_a p_a \sigma_X^a \right\}, \tag{31}$$

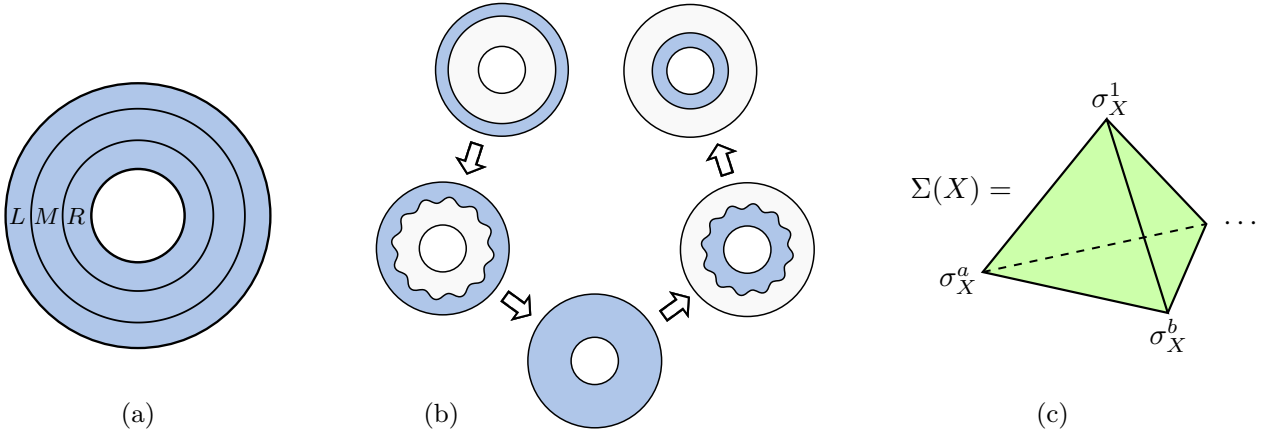*where $\{a\}$ is a finite set of labels and $\{p_a\}$ is a probability distribution.*

Figure 9: (a) Division of an annulus X into three thinner annuli L, M, R. (b) A path (extensions-extensions-restrictions-restrictions) which generates an isomorphism $\Sigma(L) \cong \Sigma(X) \cong \Sigma(R)$. (c) A schematic depiction of the simplex structure of $\Sigma(X)$. The extreme points are the "corners" of the simplex. If the annulus X is contained in a disk, then one of the extreme points has the vacuum label "1".

Here we show a sketch of the proof of Theorem 4.1. (See Appendix D.1 for the full proof). The orthogonality follows from the factorization property of the fidelity $F(\rho, \tau) = (\mathrm{Tr}\sqrt{\sqrt{\rho}\,\tau\sqrt{\rho}})^2$. Let F_X be the fidelity of two extreme points in the information convex set of $X = LMR$ in Fig. 9(a). (We use the same convention for subsystems.) By using the fact that any extreme point has $I(L:R) = 0$ (Corollary D.5.1), we find

$$F_{LR} = F_L F_R. \tag{32}$$

Because the fidelity is non-decreasing under a partial trace, we have $F_{LMR} \leq F_{LR}$. Since X and L, R are annuli connected by paths, see Fig. 9(b), the isomorphism theorem implies $F = F_L = F_R = F_{LMR}$ and thus

$$F \leq F^2. \tag{33}$$

$F \in [0, 1]$. Therefore, the two extreme points are either the same ($F = 1$) or orthogonal ($F = 0$). This derivation also shows that we can copy the information of the extreme point to L and R simultaneously. The finiteness of the label set follows from the orthogonality and the setup that the Hilbert space is finite-dimensional.

Theorem 4.1 implies that $\Sigma(X)$ forms a simplex in the state space. It has a finite number of extreme points $\{\sigma_X^a\}$, which can be perfectly distinguished from each other by a projective measurement supported on the annulus. The simplex structure also implies that its elements can only store classical information in the probability distribution $\{p_a\}$. The isomorphism theorem 3.10 guarantees the universality of the label set, i.e., the fact that the same set of labels applies to all annuli, which could be connected to each other by a path. Note that there could be annuli not connected by any path, e.g., the X^0 and X^1 in Fig. 10. Theorem 4.1 is still applicable for both annuli, but the label sets for them can be different. This is related to the existence of topological defects [24].

One of the extreme points is special. Consider a contractible annulus (see X^0 in Fig. 10 for example). The information convex set of such annulus has a special extreme point which we label as "1"; physically, this corresponds to the vacuum sector.
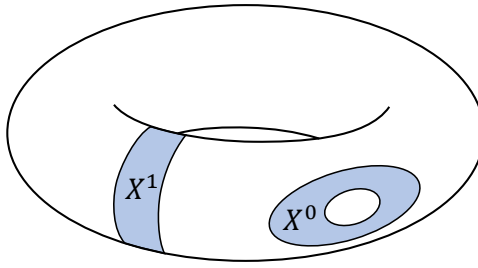
Figure 10: A pair of annuli X^0 and X^1 on a torus. They cannot be connected by any smooth path because X^0 is contractible and X^1 is non-contractible.

Proposition 4.2. *Let ω be a disk. For any annulus $X \subseteq \omega$*

$$\sigma_X^1 \equiv \mathrm{Tr}_{\omega \setminus X}\, \sigma_\omega, \tag{34}$$

is an extreme point of $\Sigma(X)$.

See Appendix D.2 for the proof. Importantly, the notion of vacuum sector is unambiguous because $\Sigma(\omega)$ has a unique element if ω is a disk; see Proposition 3.5.

Now we are ready to define the superselection sectors in our framework. When there is a pair of anyons, the topological charge can be measured by an Aharonov-Bohm type interferometry measurement by braiding the anyons [40]. Indeed, the projective measurement used for distinguishing different σ_X^a corresponds to this interferometry measurement for several exactly solvable models. Based on this observation, we identify each label of the extreme points as a superselection sector of the system.

Definition 4.1. Let X be a contractible annulus. The set of superselection sectors is a set of extreme points in $\Sigma(X)$.

Except for the vacuum sector, we label each extreme points with the lower-case Roman letters:

$$\mathcal{C} = \{1, a, b, c, \cdots\}. \tag{35}$$

Several authors have already made attempts to define superselection sectors in 2D gapped phases. A statement analogous to the simplex theorem was obtained recently in [37] for models with a local commuting parent Hamiltonian. This proof is based on the operator-algebraic framework of Haah [28].

We expect our derivation to hold more generally, because we make no assumption about the parent Hamiltonian. If the area law (Eq.(1)) holds, our results follow. In particular, if we can prove approximate versions of our statements, we may be able to rigorously define a notion of superselection sectors for models with non-zero Hall conductance or non-zero chiral central charge. These models cannot have a commuting projector parent Hamiltonian [41, 42].

The isomorphism theorem guarantees the label set of the superselection sector to be independent of the details of the annulus. However, this theorem in itself does not imply there is a well-defined way to compare the topological charges for two annuli. Indeed, in the presence of a topological defect, transporting the same superselection in two different ways may result in different sector [24].

In order to show that the isomorphisms associated with different paths are identical, one necessarily has to invoke an extra condition on the paths. Lemma 4.3 establishes one such

20

condition. Roughly speaking, it is possible to compare two annuli unambiguously independent of the path, as long as both paths lie in a single disk.

Lemma 4.3. *Let X^0 and X^1 be two annuli contained in a disk C; see Fig. 11 for example. Let $\{X_{(1)}^t\}$ and $\{X_{(2)}^t\}$ be two paths connecting X^0 and X^1 such that $X_{(i)}^0 = X^0$, $X_{(i)}^1 = X^1$ for $i = 1, 2$. Moreover, assume that $\cup_t X_{(1)}^t \subseteq C$, $\cup_t X_{(2)}^t \subseteq C$. Then, the isomorphisms*

$$\Phi_{\{X_{(1)}^t\}} : \Sigma(X^0) \to \Sigma(X^1)$$

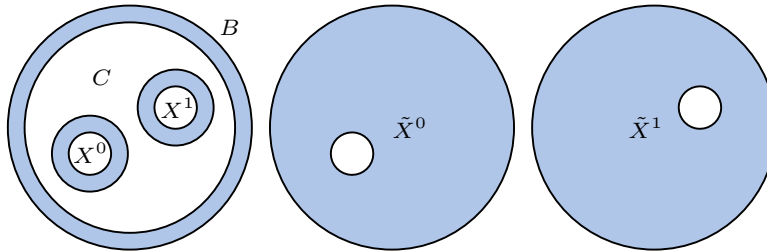$$and \quad \Phi_{\{X_{(2)}^t\}} : \Sigma(X^0) \to \Sigma(X^1)$$

are identical.

Figure 11: Both C and BC are disks. B is an annulus. X^0, X^1, $\tilde{X}^0$ and $\tilde{X}^1$ are annuli. Note that X^0 and X^1 are subsets of C. In the proof of Lemma 4.3, we construct an extension $X^t \to \tilde{X}^t$ with $B \subseteq \tilde{X}^t$.

Lemma 4.3 implies that we can always treat $\Phi_{\{X_{(1)}^t\}}(\sigma_{X^0}^a)$ as the label a for X^1. The key idea behind the proof is that one can copy the information about which superselection sector lies inside an annulus to a common annulus; see Fig. 11 for illustration. The proof is left in Appendix D.2.

Now that we have defined a notion of superselection sectors, we can define their quantum dimensions. We will do so by investigating a contribution to the entanglement entropy that depends on the choice of this sector. We will use the following definition. Later, we will be able to determine their value.

Definition 4.2 (Entropy contribution of superselection sector)**.** For a contractible annulus X, we define the universal contribution to von Neumann entropy from superselection sector a as

$$f(a) \equiv \frac{S(\sigma_X^a) - S(\sigma_X^1)}{2}. \tag{36}$$

The denominator 2 is introduced to take into account that X has two boundaries. For a connected 2D manifold, $f(a)$ is a real number that does not depend on the choice of the contractible annulus. This is because the entropy difference is preserved by an isomorphism. Furthermore, $f(1) = 0$ by definition.

Later, we shall study similar contributions for a n-hole disk with $n \geq 2$. We will find that $f(a)$ appears generically, even for these more generic subsystems. The repeated appearance of these objects hint at a possibility that there may be nontrivial relations concerning $f(a)$. Indeed, we will later see that $f(a) = \ln d_a$, where d_a is the quantum dimension of an anyon/superselection sector a. We will also be able to derive the fusion axioms and an expression for topological entanglement entropy.

4.2. Fusion rules and fusion spaces

The fusion rules determine the possible choice of the total composite topological charge of two given topological charges. In the algebraic theory of anyons, the fusion rule for charges a and b is formally written as

$$a \times b = \sum_c N_{ab}^c \, c \,,$$

where $N_{ab}^c \in \mathbb{Z}_{\geq 0}$ is the fusion multiplicity. This is analogous to the fact that two spin-$\frac{1}{2}$ particles can fuse into spin-0 or spin-1 particle. Nevertheless, there is a fundamental difference between the fusion of spins and that of anyons. For the definition of particle spins and their fusion rules, rotational symmetry is often needed either in the Hamiltonian or the Lagrangian. In contrast, the notion of topological charges and their fusions are expected to emerge from the collective properties of a many-body quantum system [43, 44, 12]. Indeed, we emphasize that our axioms (Axiom **A0** and **A1**) are unrelated to any symmetry.

In our framework, the superselection sectors were identified from annuli. Thus, one may expect the fusion rules to be extracted from 2-hole disks. In this section, we show that this is indeed the case. Let us consider a 2-hole disk Y, which we depict in Fig. 12. Let B_1, B_2 and B_3 be the three annuli around the boundaries of Y. The information convex set of each annulus has the same simplex structure, and we can label the extreme points by the same label set. Let $\Sigma_{ab}^c(Y)$ be a convex subset of $\Sigma(Y)$, defined as

$$\Sigma_{ab}^c(Y) \equiv \left\{ \rho_Y \in \Sigma(Y) \, \middle| \, \begin{array}{l} \mathrm{Tr}_{Y \backslash B_1} \, \rho_Y = \sigma_{B_1}^a \\ \mathrm{Tr}_{Y \backslash B_2} \, \rho_Y = \sigma_{B_2}^b \\ \mathrm{Tr}_{Y \backslash B_3} \, \rho_Y = \sigma_{B_3}^c \end{array} \right\} \,, \tag{37}$$

where $\sigma_{B_i}^a$ is an extreme point of $\Sigma(B_i)$, $i = 1, 2, 3$. The convention of charge labeling among the different annuli is fixed by Lemma 4.3. $\Sigma_{ab}^c(Y)$ may be empty if there is no state satisfying all the conditions. We call such a combination of (a, b, c) *forbidden*.
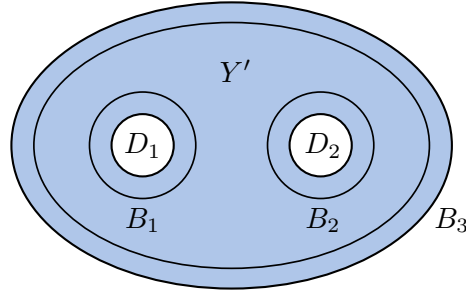

Figure 12: A 2-hole disk $Y = BY'$, with $B = B_1 B_2 B_3$. B_1, B_2 and B_3 are three annuli surrounding the three boundaries of Y. YD is a disk, where $D = D_1 D_2$. D_1 and D_2 are the two disks surrounded by annuli B_1 and B_2.

We show that every extreme point of $\Sigma(Y)$ is contained in some $\Sigma_{ab}^c(Y)$. Because $\Sigma(Y)$ is a convex set, the entire set can be characterized by $\{\Sigma_{ab}^c(Y)\}$. Each $\Sigma_{ab}^c(Y)$ is isomorphic to the state space of a finite-dimensional Hilbert space $\mathbb{V}_{ab}^c$. These two results are summarized below in Theorems 4.4 and 4.5.

Theorem 4.4. *For a 2-hole disk Y, the information convex set $\Sigma(Y)$ is the following convex combination*

$$\Sigma(Y) = \left\{ \rho_Y = \bigoplus_{a,b,c \in \mathcal{C}} p_{ab}^c \, \rho_Y^{abc} \,\middle|\, \rho_Y^{abc} \in \Sigma_{ab}^c(Y) \right\}, \tag{38}$$

where $\{p_{ab}^c\}$ is a probability distribution.

Proof. After taking a partial trace, the reduced density matrix of an extreme point of $\Sigma(Y)$ reduces to an extreme point of $\Sigma(B_1)$, $\Sigma(B_2)$ and $\Sigma(B_3)$. This fact follows from Lemma D.7 in the appendix. Therefore, every extreme point of $\Sigma(Y)$ is in $\Sigma_{ab}^c(Y)$ for some a, b, and c. This implies Eq. (38). $\qquad\square$

This theorem implies that one can classify the extreme points by a triple of labels (a, b, c). Furthermore, the convex combination in Eq. (38) is orthogonal, since one can perfectly distinguish these labels by projective measurements on the three distinct annuli.

Now we study the geometric structure of each $\Sigma_{ab}^c(Y)$. We should emphasize an important difference between $\Sigma_{ab}^c(Y)$ and the information convex set of an annulus. On an annulus, the information convex set has a classical structure specified by a probability distribution $\{p_a\}_{a \in \mathcal{C}}$. In contrast, $\Sigma_{ab}^c(Y)$ is coherent in the sense that it is isomorphic to the state space $\mathcal{S}(\mathbb{V}_{ab}^c)$ of a certain finite dimensional Hilbert space $\mathbb{V}_{ab}^c$. If the dimension of $\mathbb{V}_{ab}^c$ is greater or equal to 2, the structure of $\Sigma_{ab}^c(Y)$ allows the storage of quantum information. This structure is established by the following theorem.

Theorem 4.5. *Consider a 2-hole disk Y. $\forall a, b, c \in \mathcal{C}$,*

$$\Sigma_{ab}^c(Y) \cong \mathcal{S}(\mathbb{V}_{ab}^c), \tag{39}$$

where $\mathbb{V}_{ab}^c$ is a finite-dimensional Hilbert space.

A particular choice of (a, b, c) is forbidden, when $\dim \mathbb{V}_{ab}^c = 0$. See Appendix E for the proof of Theorem 4.5. The key idea is to show that there is a quantum channel which simultaneously purifies every extreme point of $\Sigma_{ab}^c(Y)$ into a state in Hilbert space $\mathcal{H}_{EY}$ (E is an auxiliary system). We then show that any superposition of the purified states reduces to an extreme point of $\Sigma_{ab}^c(Y)$ on Y. It follows that the quantum channel which achieves the purification provides an isomorphism between $\Sigma_{ab}^c(Y)$ and $\mathcal{S}(\mathbb{V}_{ab}^c)$.

We call the Hilbert space $\mathbb{V}_{ab}^c$ defined in Theorem 4.5 as the *fusion space*. Physically, this Hilbert space is nonempty if the superselection sectors a, b has a total charge of c. We thus define the fusion rule using the dimension of the corresponding fusion space.

Definition 4.3. We define the *fusion rule* of labels a, b in $\mathcal{C}$ by the formal product

$$a \times b = \sum_{c \in \mathcal{C}} N_{ab}^c \, c \,, \tag{40}$$

where $N_{ab}^c \equiv \dim \mathbb{V}_{ab}^c$.

The results in Theorem 4.4 and Theorem 4.5 generalize to n-hole disks with $n \geq 3$. The same applies to the concepts of fusion spaces and fusion rules.

4.3. Derivation of the axioms of the fusion rules

In this section, we show how the axioms of the anyon fusion theory emerge from our axioms. This derivation includes the existence of antiparticles and a set of rules that $\{N_{ab}^c\}$ has to satisfy.

We have defined the set of superselection sectors

$$\mathcal{C} = \{1, a, b, c, \cdots\}$$

in terms of the extreme points of $\Sigma(X)$, where X is a contractible annulus. $\mathcal{C}$ is always a finite set and there is a unique sector $1 \in \mathcal{C}$ which we refer to as the *vacuum*. We have also identified a set of non-negative integers $\{N_{ab}^c\}$ encoded in the structure of $\Sigma(Y)$ with a 2-hole disk Y.

The following is a list of the results we are going to prove under our definitions.

1. $N_{ab}^c = N_{ba}^c$. (Proposition 4.6).

2. $N_{a1}^c = N_{1a}^c = \delta_{a,c}$. (Proposition 4.7.)

3. The existence of an anti-sector $\bar{a} \in \mathcal{C}$ for $\forall a \in \mathcal{C}$ such that $N_{ab}^1 = \delta_{b,\bar{a}}$. (Proposition 4.9 and Definition 4.4.)

4. $N_{ab}^c = N_{\bar{b}\bar{a}}^{\bar{c}}$. (Proposition 4.10.)

5. $\sum_i N_{ab}^i N_{ic}^d = \sum_j N_{aj}^d N_{bc}^j$. (Proposition 4.11.)

Together, these properties form a subset of the axioms of the algebraic theory of anyon outlined in Appendix E of [12], also known as the unitary modular tensor category (UMTC). Concretely, what we derive in this section is the set of axioms of *fusion rule algebra* [45] which is also known under the name *commutative fusion ring* [46]. It contains slightly less axioms than a fusion category because we have not defined the F-symbols.[6]

In the derivation of the axioms of the fusion rules, we will extensively use the merging technique to relate subsystems of different topologies. From our axioms, we can infer that elements in the information convex sets are quantum Markov states with respect to many relevant partitions of certain subsystems. Because quantum Markov states can be merged together (Lemma 3.8), we can merge many of the elements of the information convex sets together. Moreover, this merging process can be repeated many times. With this process, we can generate elements of an information convex set over some region from information convex sets of its subregions. We refer the readers to Proposition C.5 in the appendix for the technical details.

Proposition 4.6.
$$N_{ab}^c = N_{ba}^c. \tag{41}$$

Proof. Let us consider a path which maps a 2-hole disk Y back to itself by exchanging the two internal holes. Associated with this path, there is an automorphism $\Sigma(Y) \cong \Sigma(Y)$. The automorphism permutes the labeling and induces an isomorphism $\Sigma_{ab}^c(Y) \cong \Sigma_{ba}^c(Y)$ for each a, b, c. Thus, $N_{ab}^c = N_{ba}^c$. $\qquad\square$

Proposition 4.7.
$$N_{1a}^c = N_{a1}^c = \delta_{a,c}. \tag{42}$$

---

[6]F-symbols will give further constraints to the fusion multiplicities, the simplest examples are studied in [47].
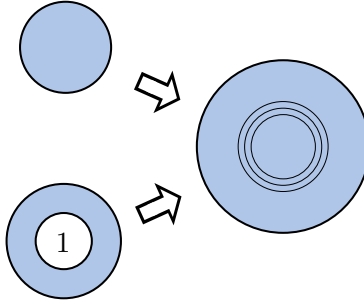
Figure 13: A hole with the vacuum charge can be merged with a disk. The case shown in this diagram involves an annulus and a disk, but the idea works for any n-hole disk with $n \geq 1$. The left side shows the topology of the subsystems before they are merged. Also, the number "1" is the vacuum sector. The merged subsystem is shown on the right. The three concentric lines partition the disk into the four subsystems used in the merging process.

Proof. Suppose $\rho_Y \in \Sigma^c_{1a}(Y)$ for 2-hole disk Y. Then the hole with the vacuum charge can be merged with a disk, see Fig. 13. After the merging process, we obtain an annulus X. The density matrix obtained from the merging process belongs to $\Sigma(X)$. The isomorphism theorem implies that the two boundaries of X detect the same topological charge. Therefore, $N^c_{1a} = \delta_{a,c}$. Then, $N^c_{a1} = \delta_{a,c}$ follows from Proposition 4.6. □

One implication of this result is that $\Sigma^1_{11}(Y)$ contains a unique element, which we call σ^1_Y. This statement generalizes to n-hole disks with $n \geq 3$. The following lemma, which is about the universal contribution to the von Neumann entropy, will be useful for the rest of the proofs. Moreover, this lemma will be one of the key results that establish a connection between this contribution and the quantum dimension.

Lemma 4.8. *Let ρ_Y be an extreme point of $\Sigma^c_{ab}(Y)$ and σ^1_Y be the unique element of $\Sigma^1_{11}(Y)$, then*

$$S(\rho_Y) - S(\sigma^1_Y) = f(a) + f(b) + f(c), \tag{43}$$

where $f(\cdot)$ is the function defined in Definition 4.2.

The proof is in Appendix E. The key idea is that for an extreme point of $\Sigma^c_{ab}(Y)$, we can prove a condition similar to that in **A0**, which converts the entropy of a pair of 2-hole disks into that of the three annuli around the three disjoint boundaries of Y. The result generalizes easily to n-hole disks for any $n \geq 3$.

Compared to the previous proofs, the proofs of the rest of the properties requires a new technique. The key idea lies in deriving consistency equations of the entropy difference, obtained by the following four steps:

(i) Obtain an element of an information convex set by merging two (or three) extreme points of the information convex sets associated with the subsystems.

(ii) Compute the entropy of the merged element from the entropy formulas with respect to the pre-merged regions.

(iii) Compute the entropy of the merged element from the entropy formula with respect to the post-merged regions.

(iv) The entropy obtained from these two perspectives must yield the same result. This leads to a set of consistency equations, which leads to a set of nontrivial relations.
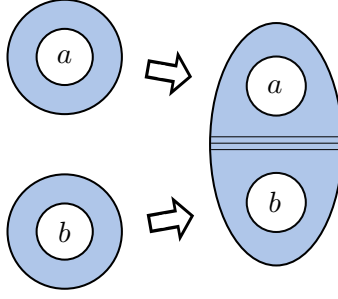
Figure 14: Merging two annuli and obtain a 2-hole disk. On the right side, there are two thin disk-like regions in the middle, which are chosen to be the B and C subsystem in the merging lemma (Lemma 3.8).

For a concrete example of the method, let us study the case shown in Fig. 14. The cases in Fig. 15, 16, 18, 19, 21 employ a similar logic. Let us explain the idea, which is broken down into four steps.

(i) We can merge the pair of annuli for any chosen charge pair $a, b \in \mathcal{C}$. This is possible because the conditions required for merging are satisfied. Let us call the merged state as $\sigma_Y^{a \times b}$. It follows that $\sigma_Y^{a \times b} \in \mathrm{conv}\left(\bigcup_c \Sigma_{ab}^c(Y)\right)$. Since the merged state exists, the set $\mathrm{conv}\left(\bigcup_c \Sigma_{ab}^c(Y)\right)$ is nonempty. This implies that $\sum_c N_{ab}^c \geq 1$, $\forall a, b \in \mathcal{C}$. Moreover, $\sigma_Y^{1 \times 1}$ is equal to σ_Y^1, which is the unique element of $\Sigma_{11}^1(Y)$.

(ii) From the perspective of the two annuli, the von Neumann entropy difference can be expressed as:

$$S(\sigma_Y^{a \times b}) - S(\sigma_Y^1) = 2f(a) + 2f(b). \tag{44}$$

This result follows from the fact that merging preserves the entropy difference; see property (3) of Lemma 3.8. More explicitly, this result follows from conditional independence condition $(I(A : D|BC) = 0)$ of the merged state; see Fig. 14. Here A (D) is the upper (lower) annuli, with charge a (b); the two annuli are separated by disk-like region BC in the middle.

(iii) From the perspective of the 2-hole disk Y, the von Neumann entropy difference is

$$\begin{aligned} & S(\sigma_Y^{a \times b}) - S(\sigma_Y^1) \\ &= f(a) + f(b) + \ln\left(\sum_c N_{ab}^c e^{f(c)}\right). \end{aligned} \tag{45}$$

To derive this result, note that the merged state is the maximal entropy element in $\mathrm{conv}\left(\bigcup_c \Sigma_{ab}^c(Y)\right)$. This is because the entropy of any state in $\mathrm{conv}\left(\bigcup_c \Sigma_{ab}^c(Y)\right)$ can be upper bounded by its marginals by the SSA and the merged state saturates this bound. Given the structure of $\Sigma(Y)$, it is easy to find the maximal entropy in terms of $\{N_{ab}^c\}$ and $f(\cdot)$. We calculated the maximal entropy and obtained Eq. (45) [23].

(iv) By comparing the two perspectives in Eq. (44) and Eq. (45), we find

$$e^{f(a)} e^{f(b)} = \sum_c N_{ab}^c e^{f(c)}. \tag{46}$$

Readers well-versed in the fusion theory of anyon may have noticed the similarity between $e^{f(a)}$ and the quantum dimension [12]. This is not a coincidence. In Sec. 5, we shall see that they are, in fact, the same thing. To establish their equivalence, we will derive a few more identities involving $\{N_{ab}^c\}$.

Moreover, we can calculate the probability of having charge c on the third boundary

$$P_{(a \times b \to c)} = \frac{N_{ab}^c e^{f(c)}}{e^{f(a)} e^{f(b)}}. \tag{47}$$

Its physical meaning is the probability to have an outcome c from the fusion of two independently created charges a and b. In terms of the density matrices, $P_{(a \times b \to c)}$ is the coefficient of the element in the center of $\Sigma_{ab}^c(Y)$ when writing $\sigma_Y^{a \times b}$ in terms of a convex combination.

It is worth noting that the same function $f(\cdot)$ appears in the entropy of the annulus and the 2-hole disk. This is crucial for the comparing the two perspectives (Eq. (44) and (45)). With this equivalence, we are in a position to derive more properties of $\{N_{ab}^c\}$. In deriving these properties, we will curtail our explanation a bit, because the argument is essentially the same.
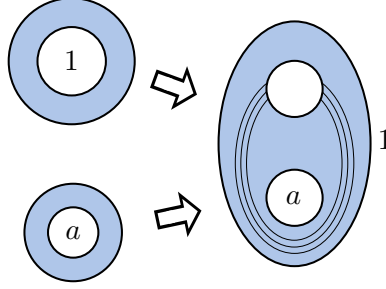

Figure 15: Merging a pair of annuli to obtain a 2-hole disk. We first deform the annulus associated to 1 so that it becomes "longer" vertically. Then, the annulus associated to a is merged into the interior of this deformed annulus. The two thin U-shaped disk-like regions are chosen to be the subsystem B and C in the merging lemma (Lemma 3.8).

Proposition 4.9. *For each charge sector $a \in \mathcal{C}$, there is a unique sector $\bar{a} \in \mathcal{C}$ such that*

$$N_{ab}^1 = \delta_{b,\bar{a}}. \tag{48}$$

It further satisfies the following properties.

$$\bar{1} = 1, \quad \bar{\bar{a}} = a, \quad f(\bar{a}) = f(a). \tag{49}$$

Proof. From the merging of two annuli with charges 1 and a shown in Fig. 15, we can derive that $\sum_b N_{ab}^1 \geq 1, \forall a$ because the merged state always exists. Furthermore, $e^{f(a)} = \sum_b N_{ab}^1 e^{f(b)}$.

Let us pick a sector b such that $N_{ab}^1 \geq 1$. We can see that $e^{f(a)} \geq e^{f(b)}$. However, since $N_{ab}^1 = N_{ba}^1$, by repeating the same logic we obtain $e^{f(b)} \geq e^{f(a)}$. For both of them to be true, we must have a unique sector $\bar{a}$ such that $N_{ab}^1 = \delta_{b,\bar{a}}$ and $f(a) = f(\bar{a})$. Then it follows from $N_{11}^1 = 1$ that $\bar{1} = 1$. Since $N_{\bar{a}a}^1 = N_{a\bar{a}}^1 = 1$, we have $\bar{\bar{a}} = a$. $\quad\square$

Definition 4.4 (Antiparticle). We define the antiparticle of $a \in \mathcal{C}$ as the unique sector $\bar{a} \in \mathcal{C}$ established in Proposition 4.9.

The definition of $\bar{a}$ is universal and insensitive to the choice of the subsystem. Furthermore, on a sphere, one could alternatively define $\bar{a}$ according to a nontrivial automorphism of $\Sigma(X)$, see Appendix G.

Proposition 4.10.

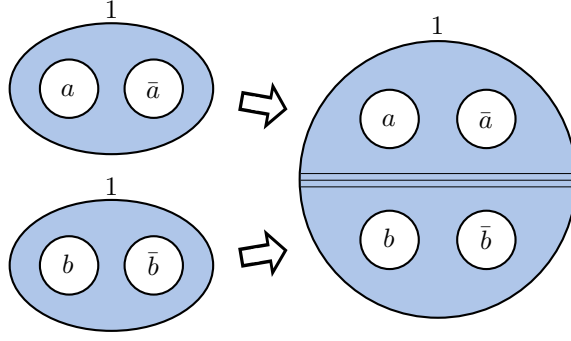$$N_{ab}^c = N_{\bar{b}\bar{a}}^{\bar{c}}. \tag{50}$$

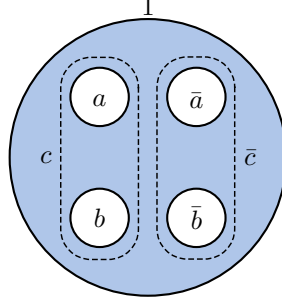Figure 16: Merging a pair of 2-hole disks to obtain a 4-hole disk.

Figure 17: The fusion of (a, b) and $(\bar{b}, \bar{a})$, and matching the fusion probabilities $P_{(a \times b \to c)}$ and $P_{(\bar{b} \times \bar{a} \to \bar{c})}$.

Proof. We consider the merging process in Fig. 16. Before merging, the density matrices are two extreme points from $\Sigma^1_{a\bar{a}}(Y_u)$ and $\Sigma^1_{b\bar{b}}(Y_d)$, if we call the pair of 2-hole disks as Y_u and Y_d. Since $N^a_{11} = \delta_{1,a}$, the outermost boundary of the merged subsystem must have charge 1.

Now let us view the merged state in a different way, as depicted in Fig. 17. We have derived that $N^1_{cd} = \delta_{d,\bar{c}}$, which implies that in the merged state, the fusion outcome of $a \times b$ and that of $\bar{b} \times \bar{a}$ are perfectly correlated. Whenever we get the outcome c from the fusion of a and b, we must get $\bar{c}$ from the fusion of $\bar{b}$ and $\bar{a}$.

Furthermore, a and b in this state are "independently created" in the sense that fusion probability $P_{(a \times b \to c)}$ obeys Eq. (47). To see why, consider partial trace operations over (i) a region which connects the hole with charge $\bar{a}$ to the outer boundary and (ii) a region which connects the hole with charge $\bar{b}$ to the outer boundary. These regions are chosen so that the remaining subsystems are topologically equivalent to the ones appearing on the right side of Fig. 14. Recalling the general inequality $I(AA' : CC'|B) \geq I(A : C|B)$, we observe that the annulus associated with a and the annulus associated with b are independent conditioned on a disk-like region in between them that separates the two annuli. As we have already discussed above, this conditional independence condition implies that $P_{(a \times b \to c)}$ obeys Eq. (47). Of course, an analogous argument can be applied to $P_{(\bar{b} \times \bar{a} \to \bar{c})}$.

Because c and $\bar{c}$ are completely correlated,

$$P_{(a \times b \to c)} = P_{(\bar{b} \times \bar{a} \to \bar{c})}. \tag{51}$$

Then, noticing $f(a) = f(\bar{a})$ from Eq. (49), we can derive $N^c_{ab} = N^{\bar{c}}_{\bar{b}\bar{a}}$. $\qquad \square$

As mentioned earlier, the results in Theorem 4.4 and Theorem 4.5 generalize to n-hole disks with $n \geq 3$, and the same applies to the concepts of fusion space and fusion rules. Let us

introduce a few notations for $n = 3$ which are useful for the next proof. For a 3-hole disk Z, we use $\Sigma(Z)$ to denote its information convex set, $\Sigma_{abc}^d(Z)$ to denote the convex subset of $\Sigma(Z)$ with fixed charges a, b, c, d on the boundaries (Fig. 18). The corresponding fusion space $\mathbb{V}_{abc}^d$ has a finite dimension $N_{abc}^d \in \mathbb{Z}_{\geq 0}$.

Proposition 4.11. *The fusion rules are associative, i.e.,*

$$N_{abc}^d = \sum_i N_{ab}^i N_{ic}^d = \sum_j N_{aj}^d N_{bc}^j. \tag{52}$$

Proof. The key idea is to obtain a 3-hole disk Z in two different ways; see Fig. 18 and 19. The first method gives us a lower bound of N_{abc}^d in terms of N_{ab}^c, and the second method shows the bound saturates.
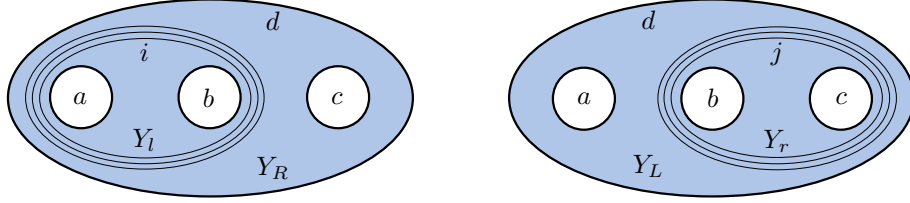

Figure 18: Merging a pair of 2-hole disks to obtain a 3-hole disk. Here $Z = Y_l \cup Y_R = Y_L \cup Y_r$. Here a, b, c, d, i, j are labels of the topological charges.

Let us consider the merging of a pair of 2-hole disks to obtain a 3-hole disk shown in Fig. 18. We summarize the logic in a streamlined fashion in (i), (ii), (iii) below.

(i) Let us consider the left side of Fig. 18, which describes the merging of Y_l and Y_R. We pick an orthonormal basis of $\mathbb{V}_{ab}^i$, which can be chosen to be the extreme points of $\Sigma_{ab}^i(Y_l)$. The number of such extreme points is equal to N_{ab}^i, which is the dimension of the Hilbert space $\mathbb{V}_{ab}^i$. Applying the same logic to Y_R, we see that the number of these extreme points is N_{ic}^d.

(ii) Let us pick two arbitrary extreme points from the sets discussed above (one from $\Sigma_{ab}^i(Y_l)$ and another from $\Sigma_{ic}^d(Y_R)$) and merge them. We get an element in $\Sigma_{abc}^d(Z)$. It is an extreme point. This fact is verified by calculating the von Neumann entropy and making use of the 3-hole version of Lemma 4.8. This way, we get $N_{ab}^i N_{ic}^d$ number of extreme points, and any two of them are orthogonal. This follows from the fact that fidelity is nondecreasing under a CPTP map.

(iii) By applying the merging process for all i, we find $\sum_i N_{ab}^i N_{ic}^d$ mutually orthogonal extreme points of $\Sigma_{abc}^d(Z)$. Therefore, we must have

$$N_{abc}^d \geq \sum_i N_{ab}^i N_{ic}^d. \tag{53}$$

The reason is N_{abc}^d is the maximal number of mutually orthogonal extreme points in $\Sigma_{abc}^d(Z)$. Similarly, from the right side of Fig. 18, we have

$$N_{abc}^d \geq \sum_j N_{aj}^d N_{bc}^j. \tag{54}$$

We did not find a way to turn "$\geq$" into "$=$" from Fig. 18 alone. However, we can show "$=$" by considering a different way of merging subsystems; see Fig. 19. The merged element, which we call $\sigma_Z^{a \times b \times c}$, is the maximal entropy element of $\text{conv}\left(\bigcup_d \Sigma_{abc}^d(Z)\right)$. Comparing with
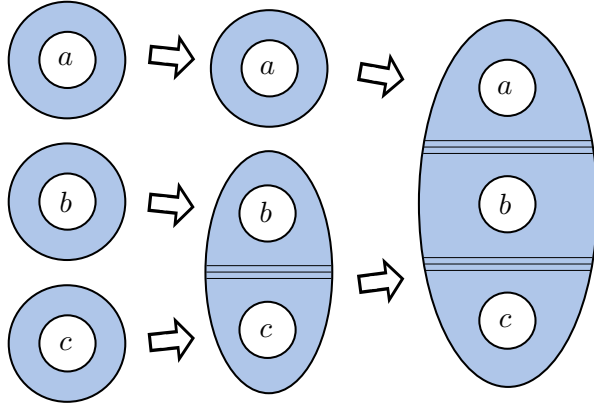
Figure 19: Merging three annuli to obtain a 3-hole disk.

σ_Z^1, the unique element of $\Sigma_{111}^1(Z)$, it has an extra $2(f(a) + f(b) + f(c))$ contribution to the von Neumann entropy. On the other hand, we use the structure of $\Sigma(Z)$ to calculate the maximal entropy in the sector with charge a, b, c in terms of $\{N_{abc}^d\}$, we find a contribution equals to $f(a) + f(b) + f(c) + \ln(\sum_d N_{abc}^d e^{f(d)})$. These two perspectives must provide a consistent answer. Thus,

$$e^{f(a)} e^{f(b)} e^{f(c)} = \sum_d N_{abc}^d e^{f(d)}. \tag{55}$$

However, from Eq. (46) we know that

$$e^{f(a)} e^{f(b)} e^{f(c)} = \sum_d (\sum_i N_{ab}^i N_{ic}^d) e^{f(d)} \tag{56}$$

and $e^{f(\cdot)}$ is positive since $f(\cdot)$ is real. So the "$\geq$" in Eq. (53) must be replaced by "$=$" and the same replacement works for Eq. (54). Thus, we conclude that Eq. (52) holds. $\qquad\square$

The result and proof of proposition 4.11 generalizes to n-hole disks with $n > 3$.

5. Topological entanglement entropy

In this section, we show that the sub-leading term γ of the area law (1) for a disk is given by the well-known formula

$$\gamma = \ln \mathcal{D}, \tag{57}$$

where $\mathcal{D}$ is the total quantum dimension defined from our definition of the fusion multiplicities $\{N_{ab}^c\}$. We show this result by calculating two different linear combinations of subsystem entropies[7] respectively proposed by Kitaev-Preskill [13] and Levin-Wen [14], see Fig. 20.

The sub-leading term γ is called the topological entanglement entropy (TEE) [13]. There are two known methods for deriving TEE: assuming an underlying field theory description or explicitly calculating entropy in an exactly solvable model. Our method, on the other hand,

---

[7]More precisely speaking, we show that a certain linear combination of entanglement entropy must be $\ln \mathcal{D}$. This result implies that $\gamma = \ln \mathcal{D}$ is the only consistent value of the sub-leading term in the area law formula Eq. (1).
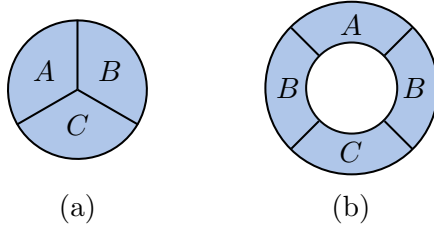
Figure 20: (a) The Kitaev-Preskill partition; (b) the Levin-Wen partition.

shows that the area law formula itself implies the equivalence of TEE and $\ln \mathcal{D}$, which may be applicable to a larger class of systems. The ingredients behind this proof are scattered in literature [16, 19, 21, 22]. Recently, one of us showed that the quantum dimension must show up in the von Neumann entropy if the fusion space is coherently encoded in the 2-hole disk [23]. In this work, we further reduce the assumption to our Axiom **A0** and **A1**. The end result is the same.

We begin by defining the quantum dimensions in our framework.

Definition 5.1. We define the set of quantum dimensions $\{d_a\}$ as the unique positive solution of the equation set

$$d_a d_b = \sum_c N_{ab}^c \, d_c \,, \tag{58}$$

where N_{ab}^c is defined in Definition 4.3. We also define the total quantum dimension $\mathcal{D}$ by $\mathcal{D} = \sqrt{\sum_{a \in \mathcal{C}} d_a^2}$.

Note that given the results in Sec. 4.3, the uniqueness of Eq. (58) is guaranteed by the Perron-Frobenius theorem, see e.g. appendix of [23] for a self-contained derivation. Furthermore,

$$d_{\bar{a}} = d_a, \quad d_a \geq 1, \quad d_1 = 1. \tag{59}$$

Recall that, from the merging in Fig. 14, we have obtained Eq. (46), and since $e^{f(a)}$ is positive, we must have

$$f(a) = \ln d_a. \tag{60}$$

In Ref. [14], it is proposed that the conditional mutual information $I(A : C|B)$ for the partition in Fig. 20(b) matches to $2 \ln \mathcal{D}$. In the paper, it is proven for a class of exactly solvable model called the Levin-Wen model (also known as the string-net model) [48]. Here we show that the same formula also holds in our framework.

Proposition 5.1. *For the Levin-Wen partition (Fig. 20(b)), it holds that*

$$I(A : C|B)_{\sigma^1} = 2 \ln \mathcal{D}. \tag{61}$$

Proof. Let us consider the merging process in Fig. 21, which obtains an annulus X from a pair of disks. Let $\tilde{\sigma}_X \in \Sigma(X)$ be the element obtained from merging. It is in the center of $\Sigma(X)$, i.e. the maximal entropy element. Dividing X according to the Levin-Wen partition in Fig. 20(b), gives $I(A : C|B)_{\tilde{\sigma}} = 0$ because of the property of the merged state; see Lemma 3.8.
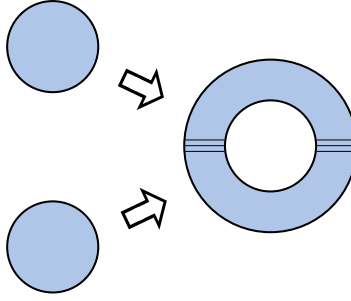
Figure 21: Merging a pair of disks to obtain an annulus. Two disks are deformed so that, once merged together, they form an annulus.

Because of the simplex structure of $\Sigma(X)$ (see theorem 4.1) and the fact that $f(a)$ is equal to $\ln d_a$ (see Eq. (60)), we can express $\tilde{\sigma}_X$ as a convex combination of extreme points

$$\tilde{\sigma}_X = \sum_a \frac{d_a^2}{\mathcal{D}^2} \sigma_X^a.$$

This formula is obtained by maximalizing the von Neumann entropy. From it, one derives $S(\tilde{\sigma}_X) - S(\sigma_X^1) = 2\ln \mathcal{D}$. It follows that Eq. (61) is true. $\qquad\square$

Proposition 5.2. *For the Kitaev-Preskill partition,*

$$\gamma \equiv (S_{AB} + S_{BC} + S_{CA} - S_A - S_B - S_C - S_{ABC})_{\sigma_\omega} \tag{62}$$

where $\omega = ABC$, see Fig. 20(a), then $\gamma = \ln \mathcal{D}$.

The idea of the proof is to relate the Levin-Wen combination with two copies of Kitaev-Preskill combinations. See Appendix F for the proof.

Because there are gapped systems in which these axioms are violated by *spurious* contributions to the area law [49, 50], one should not expect our result is applicable to every gapped system. These violations may be pathological unless certain symmetries are imposed. However, they can persist in certain subsystem symmetry-protected phases [51]. Reconciling our framework with these systems remains as an outstanding open problem. Nonetheless, our result does shed some light on a related issue: if we check the quantum state on a finite length scale and verify **A0** and **A1**, then it is guaranteed that TEE will not suffer from any spurious contribution on all larger length scales.

6. Summary and discussions

In this paper, we have initiated a derivation of the axioms of the algebraic theory of anyon from a conjectured form of entanglement area law for the ground states of 2D gapped phases. Our framework is based on two entropic constraints (axiom **A0** and **A1**), which are implied by the area law formula. We have defined the superselection sectors and the fusion spaces through the geometry of the information convex sets. The axioms of the anyon fusion theory are derived from the internal self-consistency relation of the information convex sets. Moreover, we have provided a rigorous derivation of the well-known formula for TEE, $\gamma = \ln \mathcal{D}$. While our main physical motivation was to consider ground states of 2D gapped phases, we only required a single

quantum state satisfying our axioms as the input. Our result implies that many of the anyon data can be extracted from local information of a single ground state alone.

Some of the readers may contest that our exact area law assumption is unrealistic. It would be desirable to relax this assumption to something that is less restrictive. We expect our framework to have a natural extension to the case in which Axiom **A0** and **A1** holds approximately. This is because every technical tool we have used in this paper has an analog for such situations. For instance, the merging lemma can be generalized by using the approximate recovery map [52].

It should be noted that there are gapped systems in which Axiom **A1** is violated. Such corrections are known as the *spurious* contributions to the area law [49, 50]. The existence of the spurious contribution implies that one should not expect our result to hold in every gapped system. While we do not have a solution to this problem, one may hope to take one of the following approaches. First, one may show that the notion of superselection sectors and the fusion rules are stable under a finite-depth quantum circuit when starting from a state that satisfies our axioms. Alternatively, one may attempt to show that there is always a finite-depth quantum circuit that can remove the spurious contribution.

While we have proved a set of axioms pertinent to the anyon fusion theory, further work is necessary to reproduce the anyon theory in its known form. It would be interesting to investigate whether our axioms give rise to a well-defined notion of R and F-symbols. Also, could the S and T-matrices be extracted from a single ground state? Can we prove that every anyon theory consistent with our axioms are modular? To tackle these questions, one may need to recover a certain $U(1)$ phase that is missing in the density matrix formulation, perhaps with the help of the string operator shown to exist in Appendix H.

A more ambitious question is whether we can arrive at a complete classification of two-dimensional gapped phases from our axioms. The current conjecture [12] is that two systems are in the same phase if and only if their underlying anyon theory and the chiral central charges are identical. In Ref. [12], Kitaev speculated: "To prove or disprove this statement, a mathematical notion of equivalence between topological phases is necessary. It may be based on local (or quasilocal) isomorphisms between operator algebras." Our framework seems to be the step in the right direction, given that we have a sensible definition of isomorphism between different subsystems and that we could derive axioms of the anyon fusion theory from a reasonable physical assumption. Such a feat will be a complete and rigorous justification of the point of view that 2D gapped quantum phases can be described by the anyon theory [12, 44].

Compared to the existing approaches that compute entanglement measures on ground states, information convex set leads to a more incisive understanding of the underlying topological phase. Upon calculating entanglement entropy from a given ground state, one often obtains an order parameter that reveals partial information about the underlying quantum phase. An oft-cited example is the total quantum dimension [13, 14]. Remarkably, we found that a much more refined set of data can be extracted by studying the information convex set of the ground state. This includes quantum dimensions and fusion multiplicities. Because two different anyon theories can give rise to the same total quantum dimension but different (individual) quantum dimensions and fusion multiplicities, information convex set clearly contains richer data than entanglement entropy.

Moreover, starting from our entanglement-based assumptions (Axiom **A0** and **A1**), we were able to derive the basic concepts of anyon theory from the internal consistency relations between the elements of the information convex set. What is surprising is that the emergent physical laws that appear in these topologically ordered systems were logical consequences of these axioms;

we did not need to make any further assumptions. It would be interesting to understand how widely this approach can be applied in other circumstances.

Obvious areas to explore further would be higher dimensions and setups in which a topological defect [24] or a boundary is present [25]. Such studies may be an ideal framework to classify topological phases in 3D, which remains as an outstanding open problem. We will discuss these applications in our upcoming work.

Acknowledgments

A. Notations and useful facts

In this appendix, we summarize basic notations of convex analysis and quantum information theory. Well-known facts in quantum information theory will be summarized in a self-contained manner. This section can be skipped for readers who are familiar with convex analysis and strong subadditivity of entropy.

A.1. Convex sets

Here we review facts about convex sets. We consider a subset of a finite-dimensional real space $\mathbb{R}^N$ closed under convex combinations, where $N \in \mathbb{Z}_{\geq 0}$. The convex set is compact if it is a compact subset of $\mathbb{R}^N$. For our purpose, for an N dimensional Hilbert space, the real space $\mathbb{R}^{2N^2}$ could be identified as the $2N^2$ real components of an operator acting on the Hilbert space.

We use $\text{conv}(\mathcal{X})$ to denote the convex hull of a set $\mathcal{X} \subseteq \mathbb{R}^N$, which is the smallest convex set that contains set $\mathcal{X}$. In other words, it is the set of all convex combinations of elements in $\mathcal{X}$.

An *extreme point* of a convex set $\mathcal{S}$ is a point in $\mathcal{S}$, which does not lie in any open line segment joining two points of $\mathcal{S}$. We use $\text{ext}(\mathcal{S})$ to denote the set of extreme point of a convex set $\mathcal{S}$.

Finally, we notice the *Minkowski-Caratheodory theorem*, which states that: *Let $\mathcal{S}$ be a compact convex subset of $\mathbb{R}^N$ of dimension n. Then any point in $\mathcal{S}$ is a convex combination of at most $n+1$ extreme points.* This is the reason we often talk about extreme points. Note that, without compactness, an element of a convex set sometimes cannot be written as a convex combination of extreme points.

A.2. Quantum information facts

We shall use Greek letters, e.g., ρ, σ for density matrices. Subsystems are specified in the subscript, e.g. ρ_A, σ_B. We frequently use AB as a shorthand notation of the disjoint union of A and B (i.e., $A \sqcup B$) when $A \cap B = \emptyset$. We will sometimes call a (reduced) density matrix as a state for short. Also, we will sometimes refer to the reduced density matrix of a state as its marginal. We use $\mathcal{S}(\mathcal{H})$ to denote the *state space* of a Hilbert space $\mathcal{H}$. It is the set of all density

matrices on $\mathcal{H}$. A direct sum is denoted as $\bigoplus$. It is a sum with objects living on orthogonal supports.

In order to quantify the distance between two quantum states, we use the trace distance, defined as

$$\|\rho - \sigma\|_1 \equiv \text{Tr}\sqrt{(\rho - \sigma)^2}$$

for any pair of density matrices ρ and σ. This is a reasonable notion of distance because two states close in trace distance cannot be distinguished well by any measurement.

We shall use a variety of quantum mechanical entropies in our discussion. These are all defined in terms of the von Neumann entropy of a state, which is defined as

$$S(\rho) \equiv -\text{Tr}(\rho \ln \rho).$$

Depending on the context, we shall use the following shorthand notations to denote the von Neumann entropy of the reduced density matrix over some subsystem: S_A, $(S_A + S_B)_\rho$. In the first case, the global state should be obvious from the context. In the second case, the global state is ρ.

There are two information-theoretic quantities that will play an important role in this paper:

$$I(A:B) \equiv S_A + S_B - S_{AB},$$
$$I(A:C|B) \equiv S_{AB} + S_{BC} - S_B - S_{ABC}.$$

The first quantity, known as the mutual information between A and B, quantifies a correlation between A and B. The second quantity, known as the conditional mutual information between A and C conditioned on B, quantifies the correlation between A and C given a knowledge on B. By the strong subadditivity of entropy [26], $I(A:C|B) \geq 0$ for any quantum state.

Below, we summarize the basic facts about quantum states and entropies. Most of these statements can be found in [53].

A.2.1. Fidelity

Let us begin with the fidelity between two quantum states, which is defined as

$$F(\rho, \sigma) = \left(\text{Tr}\sqrt{\rho^{\frac{1}{2}}\sigma\rho^{\frac{1}{2}}}\right)^2. \tag{A.1}$$

This is a natural generalization of the absolute value of inner product into mixed state (and from the whole system to subsystems). Indeed, one can easily verify $F(\rho, \sigma) = |\langle\psi|\varphi\rangle|^2$ when $\rho = |\psi\rangle\langle\psi|$ and $\sigma = |\varphi\rangle\langle\varphi|$. Furthermore, $F(\rho, \sigma) \in [0, 1]$ and $F(\rho, \sigma) = F(\sigma, \rho)$.

Why do we care about fidelity? It is because fidelity enjoys several useful properties. As a starter, $F(\rho, \sigma)$ can give a reasonably tight upper and lower bound on $\|\rho - \sigma\|_1$. In particular, $F(\rho, \sigma) = 1$ if and only if $\|\rho - \sigma\|_1 = 0$ and $F(\rho, \sigma) = 0$ if and only if $\|\rho - \sigma\|_1 = 2$. Moreover, fidelity has a rather special property:

$$F(\rho_A \otimes \rho_B, \sigma_A \otimes \sigma_B) = F(\rho_A, \sigma_A) \cdot F(\rho_B, \sigma_B), \tag{A.2}$$

which will play an important role in the proof of Theorem 4.1. There is an intuitive explanation for both of these facts. The first fact says that two density matrices have zero (unit) overlap if and only if the two states are orthogonal (identical).

Lastly, if two quantum states have unit fidelity, their purifications are identical up to a unitary operator acting on the purifying space [54]. In other words, two states $|\psi_{AB}\rangle$ and $|\varphi_{AB}\rangle$ have the same reduced density matrix on subsystem B if only if there is a unitary operator U_A such that

$$|\varphi_{AB}\rangle = U_A \otimes I_B |\psi_{AB}\rangle. \tag{A.3}$$

A.2.2. Quantum channels

Quantum channel, also known as complete-positive trace-preserving (CPTP) map, is the most general form of physical operation that can be applied to a quantum state. It is a linear map from bounded operators on $\mathcal{H}_A$ to bounded operators on $\mathcal{H}_{A'}$. It preserves positivity, even in the presence of any ancillary system, and also preserves trace and hermiticity.[8] In particular, it maps density matrices to density matrices. It can be written in an explicit form using a set of Kraus operators $\{M_a\}$:

$$\mathcal{E}_{A\to A'}(X_A) = \sum_a M_a X_A M_a^\dagger, \tag{A.4}$$

where $\sum_a M_a^\dagger M_a = I_A$ and I_A is the identity operator on $\mathcal{H}_A$. With the definition of CPTP map, we could discuss some additional properties.

Quantum channels do not make two quantum states more distinguishable than they already are. This means that the trace distance is nonincreasing under quantum channels. More relevant to us is the fact that the fidelity is nondecreasing under quantum channels.

$$F(\mathcal{E}(\rho), \mathcal{E}(\sigma)) \geq F(\rho, \sigma), \tag{A.5}$$

for any quantum channel $\mathcal{E}$. Since partial trace is also a quantum channel, we have

$$F(\rho_A, \sigma_A) \geq F(\rho_{AB}, \sigma_{AB}). \tag{A.6}$$

A.2.3. Properties of entropies

Let us begin with a few elementary facts about entropy. First, $S_A = S_B$ for an arbitrary pure state $|\varphi_{AB}\rangle$. Secondly, suppose a set of density matrices $\{\rho^i\}$ has mutually orthogonal supports, i.e. $\rho^i \perp \rho^j, \forall i \neq j$, then

$$S(\sum_i p_i \rho^i) = \sum_i p_i (S(\rho^i) - \ln p_i), \tag{A.7}$$

where $\{p_i\}$ is a probability distribution.

For a bipartite quantum state, we have the following set of well-known inequalities:

$$\begin{aligned} I(A:B)_\rho &\geq 0 \\ S_{BC} + S_C - S_B &\geq 0. \end{aligned} \tag{A.8}$$

The first inequality is known as the subadditivity of entropy, and the second inequality is known as the Araki-Lieb inequality [55]. It is interesting to study the conditions under which these inequalities are satisfied with equality. The mutual information is 0 if and only if the underlying state is a product state over A and B. The condition for saturating the Araki-Lieb inequality is

---

[8]That we require the map to preserve positivity in the presence of any ancillary system is important. Otherwise, there are operations, e.g., transpose of a matrix, that preserves the positivity in the absence of an ancillary system but may not if the ancillary system is entangled with the system of interest.

more subtle and interesting. We will revisit this condition later after we discuss inequalities for tripartite quantum states.

The most important inequality involving a tripartite quantum state is the strong subadditivity (SSA) of entropy [26]:

$$I(A : C|B) \geq 0. \tag{A.9}$$

This inequality is surprisingly powerful in that inequalities that may look "stronger" than this inequality are in fact implied by SSA. Here is a list of such inequalities:

$$
\begin{aligned}
&I(AA' : BB') \geq I(A : B) \\
&I(AA' : CC'|B) \geq I(A : C|B) \\
&I(AA' : CC'|B) \geq I(A : C|A'BC') \\
&S_{BC} + S_C - S_B \geq I(A : C) \\
&S_{BC} + S_C - S_B \geq I(A : C|B) \\
&S_{BC} + S_C - S_B \geq S_{BB'C} + S_C - S_{BB'} \\
&S_{BC} + S_{CD} - S_B - S_D \geq I(A : C|B) \\
&S_{BC} + S_{CD} - S_B - S_D \geq S_{BB'C} + S_{CDD'} - S_{BB'} - S_{DD'}.
\end{aligned}
$$

Also, let $\{\rho_{AB}^i\}$ be a set of density matrices and $\{p^i\}$ is a probability distribution, then

$$\sum_i p_i (S_{AB} - S_B)_{\rho^i} \leq (S_{AB} - S_B)_{\sum_i p_i \rho^i}. \tag{A.10}$$

To see why, let us introduce an auxiliary system C with an orthonormal basis $\{|i_C\rangle\}$. Let $\rho_{ABC} \equiv \sum_i p_i \, \rho_{AB}^i \otimes |i_C\rangle\langle i_C|$ and notice $I(A : C|B)_\rho \geq 0$.

A.2.4. The structure of quantum Markov states

If a tripartite quantum state satisfies SSA with equality, then such a state has a rather special property. Such a state is referred to as a quantum Markov state [32, 15]. Let ρ be a quantum Markov state such that it has a vanishing conditional mutual information $I(A : C|B)_\rho = 0$. Then the following facts hold.

1. ρ_{ABC} is uniquely determined by its marginals ρ_{AB} and ρ_{BC}. The recovery can be done with a quantum channel, see Lemma 3.1 and 3.2 of the main text.

2. There is a decomposition of the Hilbert space $\mathcal{H}_B$ into a direct sum of tensor products $\mathcal{H}_B = \bigoplus_j \mathcal{H}_{b_j^L} \otimes \mathcal{H}_{b_j^R}$ such that

$$\rho_{ABC} = \bigoplus_j p_j \, \rho_{Ab_j^L} \otimes \rho_{b_j^R C}, \tag{A.11}$$

   where $\{p_j\}$ is a probability distribution, $\rho_{Ab_j^L}$ is a density matrix on $\mathcal{H}_A \otimes \mathcal{H}_{b_j^L}$ and $\rho_{Ab_j^R}$ is a density matrix on $\mathcal{H}_{b_i^R} \otimes \mathcal{H}_C$.

3. Eq. (A.11) implies that

$$\mathrm{Tr}_B \, \rho_{ABC} = \sum_j p_j \rho_A^j \otimes \rho_C^j. \tag{A.12}$$

   Note that it is separable and therefore subsystem A and C have only classical correlations (no quantum correlation).

Now, we can make an important connection between the states saturating the Araki-Lieb inequality and quantum Markov states. The density matrices which saturate Araki-Lieb inequality have the properties summarized in Lemma A.1.

Lemma A.1. *The following conditions about density matrix ρ_{BC} are equivalent.*

(1) $(S_{BC} + S_C - S_B)_\rho = 0$, *(saturated Araki-Lieb).*

(2) *Any state ρ_{ABC} which reduces to ρ_{BC} on BC has $I(A:C)_\rho = 0$ and $I(A:C|B)_\rho = 0$.*

(3) *For any expression of the form $\rho_{BC} = \sum_i q_i \rho^i_{BC}$, where $\{q_i\}$ is a probability distribution with $q_i > 0$, $\forall i$ and $\{\rho^i_{BC}\}$ is a set of density matrices, we have*

$$\rho_C = \mathrm{Tr}_B \, \rho^i_{BC}, \quad \forall i. \tag{A.13}$$

(4) *Let $\rho_{BC} = \sum_i q_i |i_{BC}\rangle\langle i_{BC}|$, with $q_i > 0$, $\forall i$ and $\langle i_{BC}|j_{BC}\rangle = \delta_{i,j}$, $\forall i, j$, we have*

$$\mathrm{Tr}_B \, |i_{BC}\rangle\langle j_{BC}| = \delta_{i,j} \, \rho_C, \quad \forall i, j. \tag{A.14}$$

Proof. (1) $\Rightarrow$ (2). Let us purify ρ_{ABC} and obtain $|\Psi_{A'ABC}\rangle$. Condition (1) implies $I(A'A : C)_{|\Psi\rangle\langle\Psi|} = 0$ and $I(A'A : C|B)_{|\Psi\rangle\langle\Psi|} = 0$. Then, we use the facts $I(A'A : C) \geq I(A : C)$ and $I(A'A : C|B) \geq I(A : C|B)$.

(2) $\Rightarrow$ (1). Simply consider a pure ρ_{ABC}.

(1), (2) $\Rightarrow$ (3). Let $\rho_{ABC} = \sum_i q_i |i_A\rangle\langle i_A| \otimes \rho^i_{BC}$ with an orthonormal set of vectors $\{|i_A\rangle\}$. Since $I(A : C)_\rho = 0$, a projective measurement in $\{|i_A\rangle\}$ will not change the density matrix on C. It follows that $\rho_C = \mathrm{Tr}_B \, \rho^i_{BC}, \forall i$.

(3) $\Rightarrow$ (4). Let us introduce an auxiliary system A to purify $\rho_{BC} = \sum_i q_i |i_{BC}\rangle\langle i_{BC}|$ into $|\Psi_{ABC}\rangle = \sum_i \sqrt{q_i} |i_A\rangle \otimes |i_{BC}\rangle$. Here $\{|i_A\rangle\}$ is an orthonormal basis of the Hilbert space $\mathcal{H}_A$. Therefore, one may obtain ρ_{BC} from $|\Psi_{ABC}\rangle\langle\Psi_{ABC}|$ by taking a partial trace Tr_A. For an arbitrary orthonormal basis $\{|\phi^i_A\rangle\}$ we could define $p_i \rho^i_{BC} \equiv \langle\phi^i_A|\Psi_{ABC}\rangle\langle\Psi_{ABC}|\phi^i_A\rangle$. Recall that the condition Eq. (A.13) applies to any basis. For the "diagonal" basis $\{|i_A\rangle\}$, one derives $\mathrm{Tr}_B \, |i_{BC}\rangle\langle i_{BC}| = \rho_C, \forall i$. Then, consider an "off-diagonal" basis which contains a basis vector $|\varphi^\theta_{ij}\rangle = \frac{1}{\sqrt{2}}(|i_A\rangle + e^{i\theta}|j_A\rangle)$ with $i \neq j$ and $\theta \in [0, 2\pi]$. One derives that, for $\forall i \neq j$ and $\forall \theta \in [0, 2\pi]$,

$$e^{i\theta}\mathrm{Tr}_B \, |i_{BC}\rangle\langle j_{BC}| + e^{-i\theta}\mathrm{Tr}_B \, |j_{BC}\rangle\langle i_{BC}| = 0.$$

Therefore, $\mathrm{Tr}_B \, |i_{BC}\rangle\langle j_{BC}| = 0$ for $i \neq j$. Thus, Eq. (A.14) holds.

(4) $\Rightarrow$ (1). Let us introduce an auxiliary system A to purify ρ_{BC} into $|\Psi_{ABC}\rangle = \sum_i \sqrt{q_i}|i_A\rangle \otimes |i_{BC}\rangle$. Here $\{|i_A\rangle\}$ is an orthonormal basis of the Hilbert space $\mathcal{H}_A$. Then, it follows from Eq. (A.14) that $I(A : C)_{|\Psi\rangle\langle\Psi|} = 0$ which implies condition (1). $\qquad\square$

B. Information convex set on a closed manifold

In this section, we show that the information convex set of a closed manifold is isomorphic to a (state space of a) finite-dimensional Hilbert space; see Theorem B.1.

Theorem B.1. *Let M be a closed 2D manifold. Let σ be a reference state on this manifold satisfying axiom **A0** and **A1**. With respect to this reference state,*

$$\Sigma(M) = \mathcal{S}(\mathbb{V}) \tag{B.1}$$

for some finite dimensional Hilbert space $\mathbb{V} \subseteq \mathcal{H}$. Moreover, $\mathbb{V}$ is nonempty.

Proof. Because $\Sigma(M)$ is a subset of the state space of a finite-dimensional Hilbert space, one can represent an element $\rho \in \Sigma(M)$ as

$$\rho = \sum_i p_i |i\rangle\langle i|$$

for some $p_i > 0$ and an orthonormal set of vectors $\{|i\rangle\}$. The state ρ is consistent with σ on any disk. This fact follows from Proposition 3.6.

We show below that the reduced density matrices of $|i\rangle\langle i|$ on the μ-disks are equal to those obtained from the reference state σ. Without loss of generality, consider a μ-disk, say c. Consider an annulus that surrounds c, which we denote as b. This annulus is chosen so that bc is again a disk. By Proposition 3.3, $(S_{bc} + S_c - S_b)_\rho = 0$. By Lemma A.1 (1) $\Rightarrow$ (3), for $\rho_{bc} = \sigma_{bc} = \sum_i p_i \text{Tr}_{M\backslash bc}(|i\rangle\langle i|)$, we find that

$$\sigma_c = \text{Tr}_b(\text{Tr}_{M\backslash bc}(|i\rangle\langle i|)). \tag{B.2}$$

Therefore, we conclude

$$\text{Tr}_{M\backslash c}(|i\rangle\langle i|) = \sigma_c \tag{B.3}$$

for all $|i\rangle$.

Now we show that any state in the span of $\{|i\rangle\}$ lies in $\Sigma(M)$. Let $|\{|i\rangle\}| = N$. Note that $\frac{1}{N}\sum_i |i\rangle\langle i| \in \Sigma(M)$. Moreover, this state is equal to a uniform mixture over $\{U|i\rangle\}$ for any unitary U acting on the span of $\{|i\rangle\}$. Therefore, for any such U, we find

$$\text{Tr}_{M\backslash c}(U|i\rangle\langle i|U^\dagger) = \sigma_c \tag{B.4}$$

for all $|i\rangle$. This follows from the same logic that leads to Eq. (B.3). The fact that $\Sigma(M)$ is nonempty follows from the existence of $\sigma \in \Sigma(M)$. This completes the proof. $\qquad\square$

C. Elementary step of the isomorphism theorem

Our goal here is to prove the elementary step of the isomorphism theorem; see Proposition 3.9. The proof involves an alternative formulation of the information convex set, which we denote as $\hat{\Sigma}(\Omega)$. Under Axiom **A0** and **A1**, $\Sigma(\Omega)$ becomes equivalent to $\hat{\Sigma}(\Omega)$. However, generally, we do not expect these two sets to be equivalent.

C.1. An alternative formulation of the information convex set

The set $\hat{\Sigma}(\Omega)$, which we define in Definition C.1, enjoys a number of properties which are not evident from the definition of $\Sigma(\Omega)$. We use these facts to prove Proposition 3.9. The entire proof is admittedly lengthy and circuitous. A more succinct proof is left for the readers to work on.

The key difference between the definition of $\hat{\Sigma}(\Omega)$ and $\Sigma(\Omega)$ is that the latter involves an extended subsystem[9] whereas the former does not.

Definition C.1. Let $\hat{\Sigma}(\Omega)$ be a set such that $\forall \rho \in \hat{\Sigma}(\Omega)$

1. $\rho \overset{c}{=} \sigma_b \ \forall \sigma_b \in \mu$.
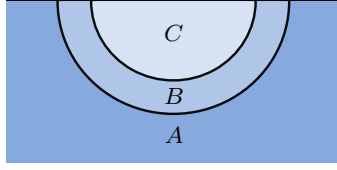
2. $I(A:C)_\rho = 0$ for the partition in Fig. 22.

Figure 22: A partition of the subsystem Ω for defining $\hat{\Sigma}(\Omega)$; see the second condition of Definition C.1. Let $\Omega = ABC$ where BC is a subsystem contained in a μ-disk. The horizontal line is the boundary of Ω. Only part of A is shown for illustration purposes.
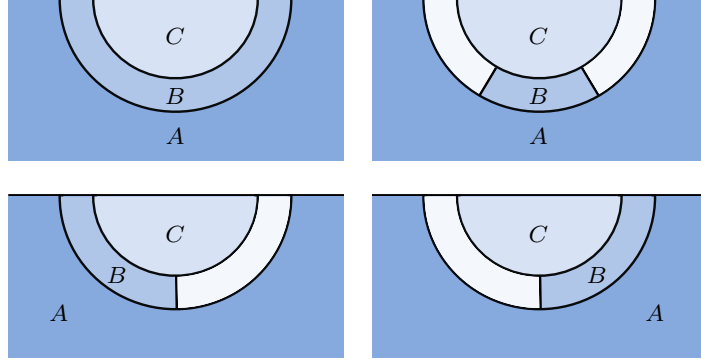

Figure 23: A partition of the subsystem Ω for defining $\hat{\Sigma}(\Omega)$; see the third condition of Definition C.1. Let $\Omega \supseteq ABC$ where BC is a subsystem contained in a μ-disk. The horizontal line is the boundary of Ω. Only part of A is shown for illustration purposes.

3. $I(A:C|B)_\rho = 0$ for the partitions in Fig. 23.

This definition does not make it clear why $\hat{\Sigma}(\Omega)$ is convex. The following proposition establishes this fact.

Proposition C.1. $\hat{\Sigma}(\Omega)$ *is convex.*

Proof. Let $\rho_\Omega, \lambda_\Omega \in \hat{\Sigma}(\Omega)$. We wish to show that their convex combination also satisfies the three conditions in Definition C.1. Without loss of generality, consider the convex combination $p\rho_\Omega + (1-p)\lambda_\Omega$ where $p \in [0,1]$.

The first condition in Definition C.1 is trivially true. For the second condition, note that $\rho_{AC} = \rho_A \otimes \rho_C$ and $\lambda_{AC} = \lambda_A \otimes \lambda_C$. Moreover, because C is in a μ-disk, $\rho_C = \lambda_C = \sigma_C$. Therefore, any convex combination of ρ and λ is also factorized over AC. Lastly, for the third condition, let us consider the conditional mutual information of the convex combination.

$$
\begin{aligned}
I(A:C|B)_{p\rho+(1-p)\lambda} &= (S_{BC} - S_B)_{p\rho+(1-p)\lambda} + (S_{AB} - S_{ABC})_{p\rho+(1-p)\lambda} \\
&\leq p(S_{BC} - S_B)_\rho + (1-p)(S_{BC} - S_B)_\lambda \\
&\quad + p(S_{AB} - S_{ABC})_\rho + (1-p)(S_{AB} - S_{ABC})_\lambda \\
&= p\,I(A:C|B)_\rho + (1-p)I(A:C|B)_\lambda \\
&= 0.
\end{aligned}
\tag{C.1}
$$

To derive this bound, we used the following two facts. First, BC is in a μ-disk. Therefore, $S_{BC} - S_B$ is the same for ρ, λ, as well as their convex combinations. Second, conditional entropy is concave; see Eq.(A.10).

---

[9]In the main text, we referred to the extended subsystems as Ω'.

Therefore, $I(A : C|B) = 0$ for the state $p\rho + (1 - p)\lambda$, for any $p \in [0, 1]$. This completes the proof. $\qquad \square$
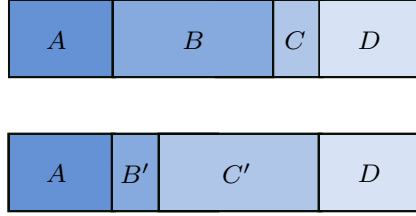

Figure 24: A schematic depiction of subsystem $ABCD$. The partition $B'C' = BC$ is chosen such that no μ-disk overlaps with both AB' and CD. Note that, the subsystems A, B, C, D are allowed to take a variety of topologies.

In a variety of circumstances, elements in $\hat{\Sigma}(ABC)$ and $\hat{\Sigma}(BCD)$ can be merged, and the merging result is an element of $\hat{\Sigma}(ABCD)$. This is the content of the following proposition.

Proposition C.2. *Consider two density matrices $\rho_{ABC} \in \hat{\Sigma}(ABC)$ and $\lambda_{BCD} \in \hat{\Sigma}(BCD)$. If the following conditions hold, ρ_{ABC} and λ_{BCD} can be merged. Moreover, the resulting density matrix is an element of $\hat{\Sigma}(ABCD)$.*

1. *There exists a partition $B'C' = BC$, such that no μ-disk overlaps with both AB' and CD; see Fig. 24.*

2. $\rho \overset{c}{=} \lambda$.

3. $I(A : C|B)_\rho = I(B : D|C)_\lambda = 0$.

4. $I(A : C'|B')_\rho = I(B' : D|C')_\lambda = 0$.

Proof. It follows from the conditions of the proposition that ρ_{ABC} and λ_{BCD} can be merged. In fact, there are two different ways to merge these density matrices. Using the third condition, we get $\tau_{ABCD} \equiv \mathcal{E}^\lambda_{C \to CD}(\rho_{ABC})$ and using the fourth condition, we get $\tau'_{ABCD} \equiv \mathcal{E}^\rho_{B' \to AB'}(\lambda_{BCD})$.

Note that $\tau_{ABCD} = \tau'_{ABCD}$. This is because both of them satisfy $\boxed{I(A : D|BC)}_{=0}$ and they have the same marginal on ABC and BCD. Therefore, Lemma 3.1 implies that $\tau_{ABCD} = \tau'_{ABCD}$.

Below, we show that τ_{ABCD} is an element of $\hat{\Sigma}(ABCD)$. In order to prove this claim, we need to show that τ_{ABCD} satisfies the condition 1, 2, and 3 in Definition C.1. Condition 1 is easy to check. Because no μ-disk overlaps with both AB' and CD, the overlap between a μ-disk and $ABCD$ is either contained in ABC or BCD.

For the second and the third condition, the key observation is that we have the freedom to choose the quantum channel:

$$\tau_{ABCD} \equiv \mathcal{E}^\lambda_{C \to CD}(\rho_{ABC}) = \mathcal{E}^\rho_{B' \to AB'}(\lambda_{BCD}). \tag{C.2}$$

For every μ-disk near the boundary of subsystem $ABCD$, we could pick a suitable quantum channel, (either $\mathcal{E}^\lambda_{C \to CD}$ or $\mathcal{E}^\rho_{B' \to AB'}$), which has no overlap with the μ-disk.

For the second condition, we use the following fact. Suppose, without loss of generality, we are given a mutual information between two subsystems say X and Y. The mutual information does not increase under a quantum channel acting only on either X or Y. Indeed, the channels

we described above are instances of such quantum channels. Therefore, the mutual information in the second condition is upper bounded by 0. This subsequently implies that the second condition holds.

For the third condition, we use a similar fact. Now, suppose we are given a conditional mutual information $I(X : Y|Z)$. This also does not increase under a quantum channel acting only on either X or Y. Therefore, with the same logic, the third condition holds as well. $\qquad\square$

Proposition C.3. *Consider $\Omega = ABC$ and $\Omega' = ABCD$ whose partitions are depicted in Fig. 8, (note that A and D are assumed to be separated by at least $2r + 1$). Let $\mathcal{E}^\sigma_{C\to CD}$ be the Petz map constructed from the reference state density matrix σ_{BCD}. Then,*

$$\mathrm{Tr}_D \circ \mathcal{E}^\sigma_{C\to CD}(\rho_\Omega) \;=\; \rho_\Omega, \quad \forall\, \rho_\Omega \in \hat{\Sigma}(\Omega), \tag{C.3}$$

$$\mathcal{E}^\sigma_{C\to CD}(\rho_\Omega) \;\in\; \hat{\Sigma}(\Omega'), \quad \forall\, \rho_\Omega \in \hat{\Sigma}(\Omega). \tag{C.4}$$

Proof. Eq. (C.3) follows directly from Lemma 3.8. To see why, let us observe that ρ_Ω and σ_{BCD} can be merged using Lemma 3.8. With condition 1, 2, 3 in Definition C.1, one could verify the two conditions required in Lemma 3.8. First, ρ_Ω is consistent with the global state on any disk $\omega \subseteq \Omega$. In particular, $\rho_\Omega \overset{c}{=} \sigma_{BCD}$. Second, $I(A : C|B)_\rho = I(B : D|C)_\sigma = 0$.

Eq. (C.4) is essentially a corollary of Proposition C.2. It is straightforward to construct the $B'C'$ required in Proposition C.2 and check all the conditions. The $2r + 1$ separation between A and D is large enough for the construction. This completes the proof. $\qquad\square$

C.2. Equivalence of the definitions

Now we can show that $\hat{\Sigma}(\Omega)$ is equivalent to $\Sigma(\Omega)$. This justifies our choice of calling $\hat{\Sigma}(\Omega)$ as the information convex set.

Proposition C.4. $\Sigma(\Omega) = \hat{\Sigma}(\Omega), \forall\, \Omega$.

Proof. If Ω is a closed manifold, then it is obvious that $\Sigma(\Omega) = \hat{\Sigma}(\Omega)$. If Ω has boundaries, it is easy to show that $\Sigma(\Omega) \subseteq \hat{\Sigma}(\Omega)$ from the assumptions. On the other hand, Proposition C.3 implies that any $\rho_\Omega \in \hat{\Sigma}(\Omega)$ can be written as $\rho_\Omega = \mathrm{Tr}_{\Omega_\epsilon \setminus \Omega}\, \rho_{\Omega_\epsilon}$ for some element $\rho_{\Omega_\epsilon} \in \hat{\Sigma}(\Omega_\epsilon)$. This is because of Eq. (C.4) and that Ω and Ω_ϵ are connected by a path which consists of a sequence of elementary extensions. It follows that $\rho_\Omega \in \Sigma(\Omega)$ and therefore $\Sigma(\Omega) \supseteq \hat{\Sigma}(\Omega)$. Thus, $\Sigma(\Omega) = \hat{\Sigma}(\Omega), \forall\, \Omega$. This completes the proof. $\qquad\square$

C.3. Merging with $\Sigma(\Omega)$

In a variety of circumstances, we can merge elements in multiple information convex sets into an element of yet another information convex set. This follows from Proposition C.2 and the established equivalence between $\Sigma(\Omega)$ and $\hat{\Sigma}(\Omega)$ (Proposition C.4).

Proposition C.5. *Consider two density matrices $\rho_{ABC} \in \Sigma(ABC)$ and $\lambda_{BCD} \in \Sigma(BCD)$. If the following conditions hold, ρ_{ABC} and λ_{BCD} can be merged. Moreover, the resulting density matrix belongs to $\Sigma(ABCD)$.*

1. *There exists a partition $B'C' = BC$, such that no μ-disk overlaps with both AB' and CD; see Fig. 24.*

2. *$\rho \overset{c}{=} \lambda$.*

3. $I(A:C|B)_\rho = I(B:D|C)_\lambda = 0$.

4. $I(A:C'|B')_\rho = I(B':D|C')_\lambda = 0$.

Proof. The proof directly follows from Proposition C.2 and C.4. $\qquad\square$

Remark. Importantly, Proposition C.5 implies that the merged results in Fig. 13, 14, 15, 16, 18, 19, 21 are elements of some information convex sets.

A special case relevant to the proof of isomorphism theorem is the following corollary.

Corollary C.5.1. *The merging process in Fig. 8 generates a map from $\Sigma(\Omega)$ to $\Sigma(\Omega')$, i.e.*

$$\mathcal{E}^\sigma_{C\to CD}(\rho_\Omega) \in \Sigma(\Omega'), \quad \forall \rho_\Omega \in \Sigma(\Omega), \tag{C.5}$$

where $\mathcal{E}^\sigma_{C\to CD}$ is the Petz map constructed from the reference state density matrix σ_{BCD}.

Proof. The proof follows from Proposition C.3 and C.4. $\qquad\square$

C.4. Proof of Proposition 3.9

Now, we are in a position to prove Proposition 3.9.

Proposition 3.9. *Consider the partition in Fig. 8. (Note: A and D are assumed to be separated by at least $2r+1$ in Fig. 8.) Let the domain of Tr_D and $\mathcal{E}^\sigma_{C\to CD}$ to be $\Sigma(\Omega')$ and $\Sigma(\Omega)$ respectively.*

$$\begin{aligned} \mathrm{Im}\,\mathrm{Tr}_D &\subseteq \Sigma(\Omega) \\ \mathrm{Im}\,\mathcal{E}^\sigma_{C\to CD} &\subseteq \Sigma(\Omega'). \end{aligned} \tag{20}$$

Moreover, for all $\rho_\Omega \in \Sigma(\Omega)$ and $\rho_{\Omega'} \in \Sigma(\Omega')$,

$$\begin{aligned} \mathrm{Tr}_D \circ \mathcal{E}^\sigma_{C\to CD}(\rho_\Omega) &= \rho_\Omega, \tag{21} \\ \mathcal{E}^\sigma_{C\to CD} \circ \mathrm{Tr}_D(\rho_{\Omega'}) &= \rho_{\Omega'}. \tag{22} \end{aligned}$$

In particular, Tr_D and $\mathcal{E}^\sigma_{C\to CD}$ are bijections.

Proof. We have shown that Tr_D is a linear map which maps elements of $\Sigma(\Omega')$ into elements of $\Sigma(\Omega)$ (Proposition 3.4). Conversely, upon applying $\mathcal{E}^\sigma_{C\to CD}$ to ρ_Ω we obtain a merged state τ^ρ_{ABCD} of ρ_Ω and σ_{BCD}. (Here σ_{BCD} is the reduced density matrix of the reference state.) We can merge them using the merging lemma (Lemma 3.8) because (1) $\rho_{BC} = \sigma_{BC}$ follows from Proposition 3.6 and (2) the requisite conditional independence conditions follow from **A1**. While the merging lemma guarantees the existence of $\tau^\rho_{\Omega'}$, it remains to show that $\tau^\rho_{\Omega'}$ is an element of $\Sigma(\Omega')$. This fact follows from Corollary C.5.1. This step requires that BC is large enough, or equivalently, A and D are separated by enough distance.

Now, it remains to prove Eqs. (21) and (22). Eq. (21) holds because the merged state is consistent with the given marginals; see Eq. (17). Eq. (22) follows from the fact that the state on both sides of the equation obey $I(AB:D|C) = 0$ and that they have the same reduced density matrices over ABC and CD; they are equal to ρ_{ABC} and σ_{CD}. According to Lemma 3.1, the two global states must be identical. $\qquad\square$

D. Extreme points

In this section, we prove various properties of the extreme points of the information convex set. Throughout this section, we shall often consider a slight "thickening" of a subsystem. Like the convention we used in the main text, a thickening of a subsystem Ω is an enlarged subsystem Ω' which is obtained by expanding the boundaries of Ω.

If the boundary of Ω is expanded by a thickness of δ, we shall refer to that subsystem as Ω_δ. For the ensuing analysis, it will be convenient to consider a length scale ϵ, which is comparable to a single lattice spacing for the convention used in the main text.

Let us begin with the following lemma.

Lemma D.1. *Suppose $\rho_{\Omega_{2\epsilon}} \in \Sigma(\Omega_{2\epsilon})$ can be written as*

$$\rho_{\Omega_{2\epsilon}} = \sum_i q_i \, \rho^i_{\Omega_{2\epsilon}}, \tag{D.1}$$

where $\{q_i\}$ is a probability distribution with $q_i > 0$, $\forall i$ and $\{\rho^i_{\Omega_{2\epsilon}}\}$ is a set of density matrices. Then,

$$\mathrm{Tr}_{\Omega_{2\epsilon}\backslash\Omega} \, \rho^i_{\Omega_{2\epsilon}} \in \Sigma(\Omega). \tag{D.2}$$

Proof. It suffices to show that every $\rho^i_{\Omega_{2\epsilon}}$ reduces to $\sigma_b \in \mu$ on any μ-disk $b \subseteq \Omega_\epsilon$. In order to show this fact, consider $b_\epsilon \subseteq \Omega_{2\epsilon}$. Let $C = b$ and choose B such that $b_\epsilon = BC$. Then, the topology of BC is identical to the one shown in Fig. 2.

Because our axioms hold at a scale larger than the μ-disks(Proposition 3.3), and that $\rho_{BC} = \sigma_{BC}$(Proposition 3.6), we conclude that $(S_{BC}+S_C-S_B)_\rho = 0$. Now, apply Lemma A.1 (1) $\Rightarrow$ (3). We conclude

$$\rho^i_b = \rho_b = \sigma_b$$

for any i and any μ-disk $b \in \Omega_\epsilon$. This completes the proof. $\qquad\square$

The following lemma shows that any element in the information convex set has zero conditional mutual information for onion-like partitions; see Fig. 25.

Lemma D.2. *Let $\Omega = ABC$. Suppose B and C are concentric annuli described in Fig. 25. Then,*

$$I(A:C|B)_\rho = 0, \quad \forall \rho_{ABC} \in \Sigma(ABC). \tag{D.3}$$

Proof. Let us consider a sequence of regions C_m with $m = 0, 1, 2, 3, \cdots M$ with $C_0 = \emptyset$, $C_M = C$, and $C_i \subseteq C_{i+1}$. Here BC_{i+1} is an "infinitesimal deformation" of BC_i as in Fig. 4.

We want to show that

$$(S_{ABC_i} - S_{BC_i})_\rho = (S_{ABC_{i+1}} - S_{BC_{i+1}})_\rho, \tag{D.4}$$

for $i = 0, \cdots, M-1$. Equivalently, we can show that

$$I(A:\delta C_{i+1}|BC_i) = 0, \tag{D.5}$$

where $\delta C_{i+1} = C_{i+1}\backslash C_i$. One can upper bound the left hand side of Eq. (D.5) by 0. To see why, first use $I(AA':C|B) \geq I(A:C|A'B)$ so that

$$I(A:\delta C_{i+1}|BC_i) \leq I(ABC_i\backslash D_i:\delta C_{i+1}|D_i), \tag{D.6}$$

where $D_i \subseteq BC_i$ is a disk-like subsystem that separates δC_{i+1} from $ABC_i \backslash D_i$. In particular, we choose D_i such that $D_i \delta C_{i+1}$ is contained in a μ-disk. Then, one can upper bound the right-hand side of Eq. (D.6) by 0, by using Axiom **A1**. Therefore,

$$(S_{ABC_0} - S_{BC_0})_\rho = (S_{ABC_M} - S_{BC_M})_\rho,$$
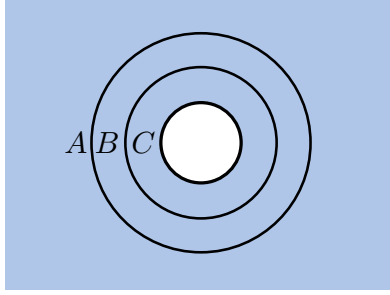
which justifies Eq. (D.3). $\qquad\square$

Figure 25: Here $\Omega = ABC$ is an arbitrary subsystem with a boundary. B and C are concentric annuli covering the boundary.

Lemma D.3. *Consider an extreme point $\sigma_{\Omega_{2\epsilon}}^{\langle e \rangle} \in \Sigma(\Omega_{2\epsilon})$ written as*

$$\sigma_{\Omega_{2\epsilon}}^{\langle e \rangle} = \sum_i q_i\, \rho_{\Omega_{2\epsilon}}^i \tag{D.7}$$

where $\{q_i\}$ is a probability distribution with $q_i > 0, \forall i$ and $\{\rho_{\Omega_{2\epsilon}}^i\}$ is a set of density matrices. Then,

$$\mathrm{Tr}_{\Omega_{2\epsilon}\backslash\Omega}\, \rho_{\Omega_{2\epsilon}}^i = \mathrm{Tr}_{\Omega_{2\epsilon}\backslash\Omega}\, \sigma_{\Omega_{2\epsilon}}^{\langle e \rangle}, \quad \forall i$$

is the same extreme point of $\Sigma(\Omega)$.

Proof. It follows from Lemma D.1 that $\mathrm{Tr}_{\Omega_{2\epsilon}\backslash\Omega}\, \rho_{\Omega_{2\epsilon}}^i$ is an element of $\Sigma(\Omega)$ for all i. The nontrivial statement is that the reduced state on Ω is an extreme point and that the reduced state does not depend on i.

Suppose there is a dependences on i. Then $\mathrm{Tr}_{\Omega_{2\epsilon}\backslash\Omega}\, \sigma_{\Omega_{2\epsilon}}^{\langle e \rangle}$ must be a convex combination of these distinct elements, so this density matrix cannot be an extreme point. This contradicts with the isomorphism theorem (Theorem 3.10), which implies that $\mathrm{Tr}_{\Omega_{2\epsilon}\backslash\Omega}\, \sigma_{\Omega_{2\epsilon}}^{\langle e \rangle}$ is an extreme point of $\Sigma(\Omega)$. (Recall that any linear bijective map between two convex sets must map extreme points to extreme points.) Therefore, the density matrix $\mathrm{Tr}_{\Omega_{2\epsilon}\backslash\Omega}\, \rho_{\Omega_{2\epsilon}}^i$ is independent of i and it follows that it is an extreme point of $\Sigma(\Omega)$. $\qquad\square$

Lemma D.4. *Consider an extreme point $\sigma_{\Omega_{2\epsilon}}^{\langle e \rangle} \in \Sigma(\Omega_{2\epsilon})$ and let $B = \Omega_{2\epsilon}\backslash\Omega$, then*

$$(S_{B\Omega} + S_\Omega - S_B)_{\sigma^{\langle e \rangle}} = 0. \tag{D.8}$$

Proof. This result follows from Lemma D.3 and $(3) \Rightarrow (1)$ of Lemma A.1. $\qquad\square$

The significance of Lemma D.4 is that it applies to any subsystem. In particular, for 2-hole disks, we obtain the following corollary.

Corollary D.4.1. *Let Y be a 2-hole disk divided according to Fig. 12, i.e. $Y = BY'$. Let $\sigma_Y^{\langle e \rangle}$ be an extreme point of $\Sigma(Y)$, then*

$$(S_{BY'} + S_{Y'} - S_B)_{\sigma^{\langle e \rangle}} = 0. \tag{D.9}$$

Lemma D.5. *Let $\Omega = ABC$ with a choice of subsystems described in Fig. 25. If $\sigma_\Omega^{\langle e \rangle}$ is an extreme point of $\Sigma(\Omega)$:*

$$I(A : C)_{\sigma^{\langle e \rangle}} = 0. \tag{D.10}$$

Proof. Since $\sigma_\Omega^{\langle e \rangle} \in \Sigma(\Omega)$, from Lemma D.2,

$$I(A : C | B)_{\sigma^{\langle e \rangle}} = 0. \tag{D.11}$$

Then, it follows from the explicit structure of quantum Markov state Eq. (A.11) and (A.12) that

$$\mathrm{Tr}_B \, \sigma_\Omega^{\langle e \rangle} = \sum_i p_i \, \rho_A^i \otimes \rho_C^i, \tag{D.12}$$

where ρ_A^i and ρ_C^i are density matrices (which may or may not belong to $\Sigma(A)$ and $\Sigma(C)$). $\{p_i\}$ is a probability distribution. We know from the isomorphism theorem 3.10 that $\sigma_A^{\langle e \rangle} \equiv \sum_i p_i \rho_A^i$ is an extreme point of $\Sigma(A)$.

Since A is thick enough, let $A = A'_{2\epsilon}$. Here A' has the same topology as A but it is thinner. From Lemma D.3 we know that $\mathrm{Tr}_{A \backslash A'} \rho_A^i = \sigma_{A'}^{\langle e \rangle}$, $\forall i$. Therefore, $I(A' : C)_{\sigma^{\langle e \rangle}} = 0$. Since we could enlarge A' (as that in Fig. 4) until it recovers A without changing the mutual information, we conclude that $I(A : C)_{\sigma^{\langle e \rangle}} = 0$. $\square$

As an immediate application of Lemma D.5, we can prove the following factorization property between subsystem L and R described in Fig. 9. This plays an important role in the proof of the orthogonality of the extreme points.

Corollary D.5.1. *For the annulus $X = LMR$ in Fig. 9(a), for any extreme point $\sigma_X^a \in \Sigma(X)$,*

$$\mathrm{Tr}_M \, \sigma_X^a = \sigma_L^a \otimes \sigma_R^a, \tag{D.13}$$

where σ_L^a and σ_R^a are the reduced density matrices of σ_X^a on L and R respectively.

Corollary D.5.2. *Consider the partition of a 2-hole disk Y in Fig. 12, i.e. $Y = Y'B$ and $B = B_1 B_2 B_3$. Let $\sigma_Y^{\langle e \rangle}$ be an extreme point of $\Sigma(Y)$, then*

$$(S_{B_1} + S_{B_2} + S_{B_3} - S_B)_{\sigma^{\langle e \rangle}} = 0. \tag{D.14}$$

Note that Eq. (D.14) is equivalent to saying that $\sigma_{B_1 B_2 B_3}^{\langle e \rangle}$ is a tripartite product state.

D.1. Orthogonality of extreme points

Below, we present the proof of Theorem 4.1, which establishes the orthogonality of extreme points.

Theorem 4.1. *(Simplex Theorem) For an annulus X, the information convex set is the convex hull of a finite number of orthogonal extreme points, $\{\sigma_X^a\}$, i.e.*

$$\Sigma(X) = \left\{ \rho_X \, \middle| \, \rho_X = \bigoplus_a p_a \sigma_X^a \right\}, \tag{31}$$

where $\{a\}$ is a finite set of labels and $\{p_a\}$ is a probability distribution.

Proof. Let us divide an annulus X according to Fig. 9(a) and consider the path in Fig. 9(b) which defines an isomorphism $\Sigma(L) \cong \Sigma(X) \cong \Sigma(R)$. Let us consider an extreme point σ_X^a and let its image in $\Sigma(L)$ and $\Sigma(R)$ be σ_L^a and σ_R^a. It follows from the isomorphism theorem (Theorem 3.10) that (1) they are extreme points; (2) they are the reduced elements of σ_X^a. For a pair of extreme points σ_X^a and σ_X^b,

$$F(\sigma_X^a, \sigma_X^b) = F(\sigma_L^a, \sigma_L^b) = F(\sigma_R^a, \sigma_R^b), \tag{D.15}$$

since the isomorphism preserves fidelity. According to Corollary D.5.1, reducing each extreme point (σ_X^a) to LR gives a tensor product structure $(\sigma_L^a \otimes \sigma_R^a)$. Thus,

$$\begin{aligned} F(\sigma_X^a, \sigma_X^b) &\leq F(\sigma_L^a \otimes \sigma_R^a, \sigma_L^b \otimes \sigma_R^b) \\ &= F(\sigma_L^a, \sigma_L^b) \cdot F(\sigma_R^a, \sigma_R^b). \end{aligned} \tag{D.16}$$

The first line follows from the monotonicity of fidelity, namely, the fact that fidelity is nondecreasing when restricted to a smaller region ($LR \subseteq X$). Eqs. (D.15) and (D.16) imply that $F(\sigma_X^a, \sigma_X^b)$ is either 0 or 1. If $F(\sigma_X^a, \sigma_X^b) = 1$ then $\sigma_X^a = \sigma_X^b$, so $a = b$. If $F(\sigma_X^a, \sigma_X^b) = 0$ then $\sigma_X^a \cdot \sigma_X^b = 0$, i.e. σ_X^a and σ_X^b live on orthogonal subspaces. This justifies the direct sum structure. Since a finite dimensional Hilbert space could only accommodate a finite number of orthogonal subspaces, the extreme points form a finite set. $\qquad \square$

D.2. Implication of the orthogonality

Based on the orthogonality of the extreme points, we can prove several new facts about the elements of the information convex set. In the remainder of this section, we use both the isomorphism theorem (Theorem 3.10) and the simplex theorem (Theorem 4.1). Let us begin with a succinct formula for the mutual information.

Proposition D.6. *Let $\rho_X = \sum_a p_a \sigma_X^a$ be an element of $\Sigma(X)$, written in terms of the orthogonal extreme points. Let $X = LMR$ be a subsystem described in Fig. 9(a). Then,*

$$I(L:R)_\rho = -\sum_a p_a \ln p_a. \tag{D.17}$$

A similar result has been obtained in [20] using Chern-Simons theory. We obtained the same result as a consequence of **A0** and **A1**.

Proof. From Theorem 3.10 we know that the reduced elements of σ_X^a on L and R, which we call as σ_L^a and σ_R^a, are extreme points of $\Sigma(L)$ and $\Sigma(R)$ respectively. Moreover, it follows from Corollary D.5.1 that $\text{Tr}_M \sigma_X^a = \sigma_L^a \otimes \sigma_R^a$. From the orthogonality relation (Eq. (31)) of Theorem 4.1. We obtain

$$(S_L)_\rho = H(p) + \sum_a p_a S(\sigma_L^a),$$

$$(S_R)_\rho = H(p) + \sum_a p_a S(\sigma_R^a),$$

$$(S_{LR})_\rho = H(p) + \sum_a p_a (S(\sigma_L^a) + S(\sigma_R^a)),$$

where $H(p) \equiv -\sum_a p_a \ln p_a$ is the Shannon entropy of the probability distribution $\{p_a\}$. The claim follows straightforwardly from these equations. $\qquad \square$

As a consequence, we can identify the extreme points of $\Sigma(X)$ as the elements with zero correlation between L and R.

Corollary D.6.1. *An element $\rho_X \in \Sigma(X)$ is an extreme point if only if $I(L : R)_\rho = 0$ for the partition in Fig. 9(a).*

Lemma D.7. *Let $\Omega = ABC$ with a choice of subsystems described in Fig. 25. If $\sigma_\Omega^{\langle e \rangle}$ is an extreme point of $\Sigma(\Omega)$, $\mathrm{Tr}_{\Omega \backslash C}\, \sigma_\Omega^{\langle e \rangle}$ is an extreme point of $\Sigma(C)$.*

Proof. Let us consider an annulus subset $A' \subseteq A$ next to BC. $I(A' : C)_{\sigma^{\langle e \rangle}} = 0$ follows from Lemma D.5. Then, because $A'BC$ is a partition of annulus similar to LMR in Fig. 9(a), $I(A' : C) = 0$ guarantees that $\mathrm{Tr}_{\Omega \backslash C}\, \sigma_\Omega^{\langle e \rangle}$ is an extreme point of $\Sigma(C)$. This step follows from Corollary D.6.1. □

The proof of Proposition 4.2 and Lemma 4.3 are also discussed below.

Proposition 4.2. *Let ω be a disk. For any annulus $X \subseteq \omega$*

$$\sigma_X^1 \equiv \mathrm{Tr}_{\omega \backslash X}\, \sigma_\omega, \tag{34}$$

is an extreme point of $\Sigma(X)$.

Proof. Without loss of generality, assume there is a disk D surrounded by R (in Fig. 9(a)). Let us consider the disk $\omega = LMRD$. There is a constraint $S_{MRD} + S_{RD} - S_M = 0$ for the unique element of $\Sigma(\omega)$, because of the enlarged version of **A0** (Proposition 3.3). This implies that $I(L : R)_{\sigma^1} = 0$. Then it follows from Corollary D.6.1 that σ_X^1 is an extreme point of $\Sigma(X)$. □

Lemma 4.3. *Let X^0 and X^1 be two annuli contained in a disk C; see Fig. 11 for example. Let $\{X_{(1)}^t\}$ and $\{X_{(2)}^t\}$ be two paths connecting X^0 and X^1 such that $X_{(i)}^0 = X^0$, $X_{(i)}^1 = X^1$ for $i = 1, 2$. Moreover, assume that $\cup_t X_{(1)}^t \subseteq C$, $\cup_t X_{(2)}^t \subseteq C$. Then, the isomorphisms*

$$\Phi_{\{X_{(1)}^t\}} : \Sigma(X^0) \to \Sigma(X^1)$$

$$and \quad \Phi_{\{X_{(2)}^t\}} : \Sigma(X^0) \to \Sigma(X^1)$$

are identical.

Proof. The key of the proof is the existence of an extension $X_{(i)}^t \to \tilde{X}_{(i)}^t$, where $B \subseteq \tilde{X}_{(i)}^t$ for each path label $i = 1, 2$ and for any time step t; see Fig. 11. For both $i = 1, 2$, the sequence of configurations $\{\tilde{X}_{(i)}^t\}$ forms a path.

Let us show that two isomorphisms on the "extended paths", $\Phi_{\{\tilde{X}_{(1)}^t\}}$ and $\Phi_{\{\tilde{X}_{(2)}^t\}}$, are identical. (Each of them maps $\Sigma(\tilde{X}^0)$ to $\Sigma(\tilde{X}^1)$.) Because every subsystem in the paths contains annulus C, the reduced density matrix on C is unchanged during the process, (for any t). Since an element in $\Sigma(\tilde{X}^0)$ (or $\Sigma(\tilde{X}^1)$) is uniquely determined from its reduced density matrix on C independent of the chosen path, the two isomorphisms $\Phi_{\{\tilde{X}_{(1)}^t\}}$ and $\Phi_{\{\tilde{X}_{(2)}^t\}}$ must be identical.

To complete the proof, note that $\Phi_{\{X_{(1)}^t\}}$ and $\Phi_{\{X_{(2)}^t\}}$ are determined from $\Phi_{\{\tilde{X}_{(1)}^t\}}$ and $\Phi_{\{\tilde{X}_{(2)}^t\}}$ by taking a partial trace. This completes the proof. □

E. Fusion space

Here, we prove that the convex set $\Sigma_{ab}^c(Y)$ is isomorphic to the state space of a finite-dimensional fusion space $\mathbb{V}_{ab}^c$. Recall that a, b, and c are superselection sectors and Y is a 2-hole disk. We begin with the following lemma, which characterizes the extreme points of $\Sigma_{ab}^c(Y)$.

Lemma E.1. *Let Y be a 2-hole disk. Then, every extreme point of $\Sigma_{ab}^c(Y)$ has the same von Neumann entropy.*

Proof. Consider a partition of Y into $Y = BY'$ as described in Fig. 12. Let λ_Y and ρ_Y be two extreme points of $\Sigma_{ab}^c(Y)$ and $\delta \equiv S(\lambda_Y) - S(\rho_Y)$ be the entropy difference between the two states. Eq. (D.9) and (D.14) imply that $(S_Y + S_{Y'})_\lambda$ and $(S_Y + S_{Y'})_\rho$ are identical. Also, by the isomorphism theorem $(S_Y + S_{Y'})_\lambda - (S_Y + S_{Y'})_\rho = 2\delta$. Thus, $\delta = 0$. $\qquad\square$

We have seen that all the extreme points of $\Sigma_{ab}^c(Y)$ have the same entropy. It follows that a non-extreme point, which is a convex combination of multiple extreme points, must have higher entropy. This fact follows from the general property of von Neumann entropy, $S(\sum_i p_i \rho^i) \geq \sum_i p_i S(\rho^i)$, where $\{p_i\}$ is a probability distribution with $p_i > 0$. The equality is achieved if and only if all ρ^i are identical. Thus, we have the following corollary.

Corollary E.1.1. *If the entropy of an element $\rho_Y \in \Sigma_{ab}^c(Y)$ is identical to that of an extreme point of $\Sigma_{ab}^c(Y)$, then ρ_Y itself is an extreme point of $\Sigma_{ab}^c(Y)$.*

Theorem 4.5. *Consider a 2-hole disk Y. $\forall a, b, c \in \mathcal{C}$,*

$$\Sigma_{ab}^c(Y) \cong \mathcal{S}(\mathbb{V}_{ab}^c), \tag{39}$$

where $\mathbb{V}_{ab}^c$ is a finite-dimensional Hilbert space.

Proof. Recall that we have already partitioned Y into BY'; see Fig. 12. We shall consider two different partitions of B. In Fig. 12, we have already considered a partition of B into $B = B_1 B_2 B_3$, which is a (disjoint) union of three annuli (B_1, B_2, and B_3). We shall also consider a different partition of $B = B_L B_M B_R$. Here B_L is a (disjoint) union of three "outermost" annuli and B_R is a (disjoint) union of three "innermost" annuli; see Fig. 26.

In total, we are considering 9 disjoint subsets of B, Y', and E; see Table E.3 for a detailed discussion on the partition of B. Here E is an auxiliary Hilbert space used to purify a density matrix supported on $Y = BY'$.

Partitions	B_1	B_2	B_3
B_R (Inner)	B_{1R}	B_{2R}	B_{3R}
B_M (Middle)	B_{1M}	B_{2M}	B_{3M}
B_L (Outer)	B_{1L}	B_{2L}	B_{3L}

Table E.3: A partition of B used in the proof of Theorem 4.5.

The statement is trivially true if $\Sigma_{ab}^c(Y)$ is empty. In this case $\dim \mathbb{V}_{ab}^c = 0$. For a nonempty $\Sigma_{ab}^c(Y)$, we use $\{\sigma_Y^{\langle e \rangle_x}\}$ to denote the set of extreme points of $\Sigma_{ab}^c(Y)$ and use λ_Y for a generic element of $\Sigma_{ab}^c(Y)$. For the extreme points, the alphabet e signifies that the density matrix is an extreme point. They are labeled by x, y, and z in this proof.
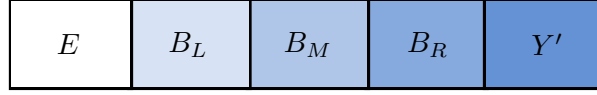
Figure 26: A partition of B used in the proof of Theorem 4.5. This figure does not represent the actual underlying geometry. Rather, it represents the relative distance between the "inner" part of Y(i.e., Y') and the annuli surrounding Y'(i.e., B_R, B_M, and B_L). Auxiliary system E is introduced to purify the extreme points. Here B_R is the "innermost" part of B that is directly in contact with Y', B_M is a disjoint union of annuli surrounding B_R, and B_L surrounds B_M.

Pick an extreme point, say $\sigma_Y^{\langle e \rangle x}$. Let us purify $\sigma_Y^{\langle e \rangle x}$ into $|\varphi_{EY}^x\rangle$ and let ρ_{EB} be its reduced density matrix on EB. From Corollary D.4.1, one can verify $I(E : B_M B_R | B_L)_\rho = 0$. Moreover, for $\forall \lambda_Y \in \Sigma_{ab}^c(Y)$ we have $I(B_L : Y'|B_M B_R)_\lambda = 0$ and $\rho_{EB} \stackrel{c}{=} \lambda_Y$. From the merging lemma (Lemma 3.8), there is a quantum channel $\mathcal{E}_{B_L \to EB_L}^\rho$ which defines a set of states

$$\mathcal{S}_{EY} \equiv \{\mathcal{E}_{B_L \to EB_L}^\rho(\lambda_Y) \,|\, \lambda_Y \in \Sigma_{ab}^c(Y)\} \tag{E.1}$$

obtained from merging ρ_{EB} with λ_Y. It follows that $\Sigma_{ab}^c(Y) \cong \mathcal{S}_{EY}$. Below, we will determine the structure of $\mathcal{S}_{EY}$.

Let us first show that the extreme points of $\mathcal{S}_{EY}$ are pure states. For the particular extreme point we have already considered, i.e., $\sigma_Y^{\langle e \rangle x}$, the merged state is obviously the pure state $|\varphi_{EY}^x\rangle$. Because all the extreme points in $\Sigma_{ab}^c(Y)$ have the same entropy (Lemma E.1) and because the entropy difference is preserved under the map $\mathcal{E}_{B_L \to EB_L}^\rho$ (property (3) of Lemma 3.8), all the other extreme points are also mapped to pure states. This means the quantum channel $\mathcal{E}_{B_L \to EB_L}^\rho$ purifies all the extreme points of $\Sigma_{ab}^c(Y)$ simultaneously.

Now, we show $\mathcal{S}_{EY}$ is the state space of a finite dimensional Hilbert space. The nontrivial statement is that any superposition of pure states in $\mathcal{S}_{EY}$ is again in $\mathcal{S}_{EY}$. Once this statement is verified, the finiteness of dimension follows straightforwardly from the fact that $\Sigma_{ab}^c(Y)$ is finite dimensional. To prove this claim, we consider a normalized state

$$|\varphi_{EY}^z\rangle \equiv \sum_i z_i |\varphi_{EY}^{y_i}\rangle, \tag{E.2}$$

where $\{|\varphi_{EY}^{y_i}\rangle\}$ is the set of purifications (in $\mathcal{S}_{EY}$) of a finite (sub)set of extreme points $\{\sigma_Y^{\langle e \rangle y_i} \in \Sigma_{ab}^c(Y)\}$. The complex numbers z_i can be arbitrary as long as $|\varphi_{EY}^z\rangle$ is normalized. It is sufficient to show $|\varphi_{EY}^z\rangle\langle\varphi_{EY}^z| \in \mathcal{S}_{EY}$. A proof is done by the following steps.

1. The reduced density matrix of $|\varphi_{EY}^z\rangle$ on $EB_L B_M$ is $\rho_{EB_L B_M}$. This is because (1) $I(EB_L B_M : Y') = 0$ on $|\varphi_{EY}^z\rangle$ and (2) $|\varphi_{EY}^z\rangle = O_{Y'}|\varphi_{EY}^x\rangle$ for some operator $O_{Y'}$ supported on Y'. The first equation follows from Corollary D.4.1. The second equation follows from the fact that $|\varphi_{EY}^{y_i}\rangle$ and $|\varphi_{EY}^x\rangle$ have the same reduced density matrix ρ_{EB} $\forall y_i$. In fact, Eq. (A.3) implies an explicit choice $O_{Y'} = \sum_i z_i U_{Y'}^i$, where $U_{Y'}^i$ are unitary operators.

2. The reduced density matrix of $|\varphi_{EY}^z\rangle$ on $B_M B_R Y'$, which we denote as $\sigma_{B_M B_R Y'}^{\langle e \rangle z}$, is an extreme point of $\Sigma_{ab}^c(B_M B_R Y')$. To see this, we first observe that $|\varphi_{EY}^z\rangle\langle\varphi_{EY}^z| \stackrel{c}{=} \sigma_b$ for any μ-disk $b \subseteq (B_M B_R Y')_\epsilon$. The logic to establish this fact is similar to that leads to the point made above: (1) $|\varphi_{EY}^x\rangle$ has vanishing correlation between b and $EY \backslash b_\epsilon$ and (2) $|\varphi_{EY}^z\rangle = O_{EY \backslash b_\epsilon}|\varphi_{EY}^x\rangle$ for some $O_{EY \backslash b_\epsilon}$. Thus, $\sigma_{B_M B_R Y'}^{\langle e \rangle z} \in \Sigma(B_M B_R Y')$. Its reduced density

matrix ρ_{B_M} determines the charge sectors a, b, c. The entropy of $\sigma^{\langle e \rangle_z}_{B_M B_R Y'}$ is identical to that of known extreme points of $\Sigma^c_{ab}(B_M B_R Y')$, e.g., $\sigma^{\langle e \rangle_x}_{B_M B_R Y'}$. Therefore, according to Corollary E.1.1, $\sigma^{\langle e \rangle_z}_{B_M B_R Y'}$ is an extreme point of $\Sigma^c_{ab}(B_M B_R Y')$.

3. The state $|\varphi^z_{EY}\rangle$ has vanishing conditional mutual information $I(B_L : B_R Y' | B_M) = 0$. Therefore, its reduced density matrix on Y is uniquely determined from its reduced density matrices $\rho_{B_L B_M}$ and $\sigma^{\langle e \rangle_z}_{B_M B_R Y'}$ (by Lemma 3.1). Therefore, $\mathrm{Tr}_E |\varphi^z_{EY}\rangle\langle\varphi^z_{EY}|$ is the extreme point of $\Sigma^c_{ab}(Y)$ obtained from an extension of $\sigma^{\langle e \rangle_z}_{B_M B_R Y'}$. We denote this extreme point as $\sigma^{\langle e \rangle_z}_Y$.

4. From the discussion above one can see, for any $|\varphi^z_{EY}\rangle$ of the form (E.2), there exists an extreme point $\sigma^{\langle e \rangle_z}_Y$ of $\Sigma^c_{ab}(Y)$ such that

$$|\varphi^z_{EY}\rangle\langle\varphi^z_{EY}| = \mathcal{E}^\rho_{B_L \to E B_L}(\sigma^{\langle e \rangle_z}_Y). \tag{E.3}$$

Thus, $|\varphi^z_{EY}\rangle\langle\varphi^z_{EY}| \in \mathcal{S}_{EY}$.

We have proved that the set $\mathcal{S}_{EY}$ in Eq. (E.1) is the state space of some finite dimensional Hilbert space. The Hilbert space depends on the purification, but its dimension cannot depend on this detail. The reason is that the state spaces of two finite-dimensional Hilbert spaces are isomorphic if and only if the dimension of the Hilbert spaces are the same and that $\mathcal{S}_{EY} \cong \Sigma^c_{ab}(Y)$. Therefore, we can assign an abstract finite dimensional Hilbert space $\mathbb{V}^c_{ab}$ with $\dim \mathbb{V}^c_{ab} = N^c_{ab} \in \mathbb{Z}_{\geq 0}$, such that

$$\Sigma^c_{ab}(Y) \cong \mathcal{S}_{EY} \cong \mathcal{S}(\mathbb{V}^c_{ab}). \tag{E.4}$$

Here $\mathcal{S}(\mathbb{V}^c_{ab})$ is the state space of $\mathbb{V}^c_{ab}$. This completes the proof. $\qquad\square$

Lemma 4.8. *Let ρ_Y be an extreme point of $\Sigma^c_{ab}(Y)$ and σ^1_Y be the unique element of $\Sigma^1_{11}(Y)$, then*

$$S(\rho_Y) - S(\sigma^1_Y) = f(a) + f(b) + f(c), \tag{43}$$

where $f(\cdot)$ is the function defined in Definition 4.2.

Proof. From Lemma E.1, we know that all the extreme points of $\Sigma^c_{ab}(Y)$ have the same entropy. It follows that $f(a, b, c) \equiv S(\rho_Y) - S(\sigma^1_Y)$ depends only on the sector (a, b, c). Let us determine $f(a, b, c)$ in terms of the universal contributions $f(\cdot)$ in Definition 4.2. Let $Y = BY'$ according to Fig. 12.

$$\begin{aligned} 2f(a, b, c) &= (S_Y + S_{Y'})_\rho - (S_Y + S_{Y'})_{\sigma^1} \\ &= (S_B)_\rho - (S_B)_{\sigma^1} \\ &= (S(\sigma^a_{B_1}) + S(\sigma^b_{B_2}) + S(\sigma^c_{B_3})) \\ &\quad - (S(\sigma^1_{B_1}) + S(\sigma^1_{B_2}) + S(\sigma^1_{B_3})) \\ &= 2(f(a) + f(b) + f(c)). \end{aligned} \tag{E.5}$$

In the first line, we used the isomorphism theorem. Recall that the isomorphism preserves the entropy difference. In the second line, we applied Eq. (D.9). In the third line, we applied Eq. (D.14). The fifth line follows from Definition 4.2. This completes the proof. $\qquad\square$

F. Topological entanglement entropy for the Kitaev-Preskill partition

In this section, we provide a proof of Proposition 5.2.

Proposition 5.2. *For the Kitaev-Preskill partition,*

$$\gamma \equiv (S_{AB} + S_{BC} + S_{CA} - S_A - S_B - S_C - S_{ABC})_{\sigma_\omega} \tag{62}$$

where $\omega = ABC$, see Fig. 20(a), then $\gamma = \ln \mathcal{D}$.

Proof. In the following, all the von Neumann entropies are calculated for the reference state σ. By deforming the subsystems using the idea in Fig. 4, one shows that the value γ in Eq. (62) is invariant under small deformations of subsystem A, B, C. Since large deformations can be built up from small ones, γ is a topological invariant. In the following, we calculate its value.
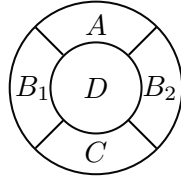

Figure 27: The partition of a disk in the proof, where $B = B_1 B_2$ and that ABC is the Levin-Wen partition.

First, let us consider subsystem choice A, B_1, D in Fig. 27, and find

$$\gamma = S_{AD} + S_{DB_1} + S_{AB_1} - S_A - S_D - S_{B_1} - S_{ADB_1}. \tag{F.1}$$

Second, we consider subsystem choice AD, B_1, B_2C in Fig. 27, and find

$$\begin{aligned}
\gamma &= S_{BC} + S_{ADB_1} + S_{ADB_2C} \\
&\quad - S_{B_2C} - S_{B_1} - S_{AD} - S_{ABCD}.
\end{aligned} \tag{F.2}$$

Both partitions are equivalent to the Kitaev-Preskill partition Fig. 20(a). Adding up Eq. (F.1) and Eq. (F.2), and using

$$\begin{aligned}
S_{ABCD} + S_D &= S_{ABC}, \\
S_{DB_1} + S_{ADB_2C} &= S_{B_1} + S_{AB_2C}, \\
S_{AB_2C} &= S_{AB_2} + S_{B_2C} - S_{B_2}, \\
S_{B_1} + S_{B_2} &= S_B, \\
S_{AB_1} + S_{AB_2} - S_A &= S_{AB},
\end{aligned}$$

(these results follow from the quantum Markov chain structure of the global reference state, i.e. the fact that Eqs. (3) and (4) are satisfied on all length scales larger than a constant value), we find

$$\begin{aligned}
2\gamma &= S_{AB} + S_{BC} - S_B - S_{ABC} \\
&= 2\ln \mathcal{D}.
\end{aligned}$$

Therefore, $\gamma = \ln \mathcal{D}$. $\qquad\square$

G. Antiparticles and automorphism of annulus on a sphere

In this appendix, we discuss a connection of antiparticle with the automorphisms of the information convex set of an annulus on a sphere. Intuitively, the connection comes from two facts. First, the automorphism only depends on the topological class of the path that maps the annulus back to itself. This fact is established by Lemma 4.3. Second, on a sphere, the topological class of the paths is described by the braid group on a sphere [10]. In general, we use $\mathcal{B}_n(V)$ to denote the n-string braid group of manifold V. Physically, this is related to the spacetime diagram of n particles braiding on a manifold V. In our framework, it is related to the deformation of a subsystem (V with n holes) by a path and then goes back to itself. For our purpose, an annulus is a 2-hole sphere, and the relevant result is the 2-string braid group on a sphere:

$$\mathcal{B}_2(S^2) = \mathbb{Z}_2. \tag{G.1}$$

For an automorphism of $\Sigma(X)$ generated by a path $\{X^t\}$ with $X^0 = X^1 = X$, where X is an annulus on a sphere, we could draw a spacetime diagram corresponding to the path. The spacetime diagram shows the braiding of two holes. The braiding belongs to one of the two classes in $\mathcal{B}_2(S^2) = \mathbb{Z}_2$, so does the path.

The path in the trivial class could be smoothly deformed into the path $X^t = X$, $\forall t$. The corresponding automorphism of $\Sigma(X)$ preserves the superselection sectors, i.e., it maps each extreme point back to itself.

On the other hand, a path in the nontrivial class generates an automorphism of $\Sigma(X)$ which permutes the extreme points according to

$$\Phi(\sigma_X^a) = \sigma_X^{\bar{a}}, \quad \forall\, a \in \mathcal{C}.$$

Intuitively, a nontrivial path switches the pair of holes. Furthermore, if one introduces an oriented loop to the annulus X, which deforms smoothly with X, then the loop will end up in the opposite orientation after X is mapped back to itself according to the nontrivial path.

H. String operators

Proposition H.1. *Given a pure reference state $|\psi\rangle$, two holes within a disk and $\forall\, a \in \mathcal{C}$, there exists a deformable unitary string operator $U^{(a,\bar{a})}$ supported within the disk, such that the state*

$$|\varphi^{(a,\bar{a})}\rangle \equiv U^{(a,\bar{a})}|\psi\rangle, \tag{H.1}$$

has topological charges a and $\bar{a}$ in the two holes.

Here, deformable means the support of $U^{(a,\bar{a})}$ can be deformed smoothly while keeping its endpoints fixed. It is easy to see, for $a = 1$, the string can be chosen to be the identity operator while for $a \neq 1$, the string cannot break apart. While this result is not directly useful in the current work, we expect it to play a role in deriving more advanced fusion and braiding properties, e.g., the S-matrix.

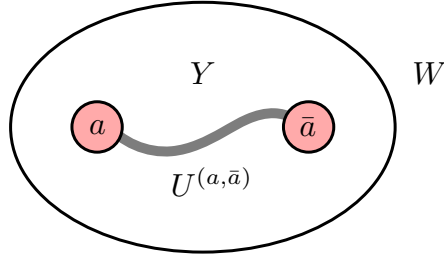---

[10]https://homepages.warwick.ac.uk/ maseay/doc/braids.pdf

Figure 28: Disk ω is the union of 2-hole disk Y and its two holes. W is the complement of ω. The topological charges a and $\bar{a}$ within the two holes are created by unitary string operator $U^{(a,\bar{a})}$. The support of $U^{(a,\bar{a})}$ is the union of the deformable gray area and the two holes shown in red.

Proof. Let ω be the disk required in the proposition and W is its complement, see Fig. 28. One can verify that $\sigma_W \equiv \mathrm{Tr}_\omega |\psi\rangle\langle\psi| \in \Sigma(W)$ and that σ_W is an extreme point. The 2-hole disk $Y \subseteq \omega$ is obtained by erasing the two holes from disk ω. From Proposition 4.9, $\Sigma^1_{a\bar{a}}(Y_{2\epsilon})$ contains a unique element which we denote as $\sigma^{a\bar{a}1}_{Y_{2\epsilon}}$, where $Y_{2\epsilon}$ is a thickening of Y. Here the subscript 2ϵ means that Y is expanded along its boundary by two lattice spacings; see the beginning of Appendix D for a related discussion.

The elements σ_W and $\sigma^{a\bar{a}1}_{Y_{2\epsilon}}$, whose supports are overlapping around the boundary of ω_ϵ, can be merged and the resulting state is an extreme point of $\Sigma(WY_{2\epsilon})$, where $WY_{2\epsilon}$ is again the thickening of WY by two unit lattice spacing. Let $|\varphi^{(a,\bar{a})}\rangle = |\varphi^{(a,\bar{a})}_{WY_{2\epsilon}}\rangle \otimes |\varphi_{V\backslash WY_{2\epsilon}}\rangle$, where $|\varphi^{(a,\bar{a})}_{WY_{2\epsilon}}\rangle$ is an eigenvector (with nonzero eigenvalue) of the merged state and $|\varphi_{V\backslash WY_{2\epsilon}}\rangle$ is an arbitrary pure state. According to Lemma D.3, the reduced density matrix of $|\varphi^{(a,\bar{a})}\rangle$ on WY is identical to that of the merged state. Therefore, $|\varphi^{(a,\bar{a})}\rangle$ has topological charges a and $\bar{a}$ within the two holes, and $|\varphi^{(a,\bar{a})}\rangle$ is identical to the reference state $|\psi\rangle$ on any subsystem $W' \subseteq WY$ which is connected to W by a path, where the path is within WY. In particular, choose W' to be the complement of disk ω', where ω' is the union of the gray string and the two holes in Fig. 28. Since $|\psi\rangle$ and $|\varphi^{(a,\bar{a})}\rangle$ are identical on W', by applying Eq. (A.3), we have $|\varphi^{(a,\bar{a})}\rangle = U^{(a,\bar{a})}|\psi\rangle$ for a unitary operator $U^{(a,\bar{a})}$ supported on ω'. Because we may deform W' and ω', the support of $U^{(a,\bar{a})}$ can be deformed smoothly. This completes the proof. $\qquad\square$

References

[1] A. Kitaev, AIP Conference Proceedings **1134**, 22 (2009), `https://aip.scitation.org/doi/pdf/10.1063/1.3149495`, URL `https://aip.scitation.org/doi/abs/10.1063/1.3149495`.

[2] A. P. Schnyder, S. Ryu, A. Furusaki, and A. W. W. Ludwig, Phys. Rev. B **78**, 195125 (2008), URL `https://link.aps.org/doi/10.1103/PhysRevB.78.195125`.

[3] A. Kapustin, ArXiv e-prints (2015), `1403.1467`.

[4] X. Chen, Z.-C. Gu, and X.-G. Wen, Phys. Rev. B **83**, 035107 (2011), URL `https://link.aps.org/doi/10.1103/PhysRevB.83.035107`.

[5] N. Schuch, D. Pérez-García, and I. Cirac, Phys. Rev. B **84**, 165139 (2011), URL `https://link.aps.org/doi/10.1103/PhysRevB.84.165139`.

[6] O. Szehr and M. M. Wolf, J. Math. Phys. **57**, 081901 (2016), `https://doi.org/10.1063/1.4960557`, URL `https://doi.org/10.1063/1.4960557`.

[7] L. Fidkowski and A. Kitaev, Phys. Rev. B **83**, 075103 (2011), URL `https://link.aps.org/doi/10.1103/PhysRevB.83.075103`.

[8] M. Barkeshli, P. Bonderson, M. Cheng, and Z. Wang, ArXiv e-prints (2015), `1410.4540`.

[9] M. B. Hastings, JSTAT **2007**, P08024 (2007), URL `https://doi.org/10.1088%2F1742-5468%2F2007%2F08%2Fp08024`.

[10] Y. Ge and J. Eisert, New J. Phys. **18**, 083026 (2016), URL `https://doi.org/10.1088%2F1367-2630%2F18%2F8%2F083026`.

[11] J. Haah, Phys. Rev. A **83**, 042330 (2011), URL `https://link.aps.org/doi/10.1103/PhysRevA.83.042330`.

[12] A. Kitaev, Ann. Phys. **321**, 2 (2006), `cond-mat/0506438`.

[13] A. Kitaev and J. Preskill, Phys. Rev. Lett. **96**, 110404 (2006), URL `https://link.aps.org/doi/10.1103/PhysRevLett.96.110404`.

[14] M. Levin and X.-G. Wen, Phys. Rev. Lett. **96**, 110405 (2006), URL `https://link.aps.org/doi/10.1103/PhysRevLett.96.110405`.

[15] P. Hayden, R. Jozsa, D. Petz, and A. Winter, Comm. Math. Phys. **246**, 359 (2004), `quant-ph/0304007`.

[16] I. H. Kim, Phys. Rev. Lett. **111**, 080503 (2013), URL `https://link.aps.org/doi/10.1103/PhysRevLett.111.080503`.

[17] I. H. Kim, ArXiv e-prints arXiv:1405.0137 (2014), `1405.0137`.

[18] I. H. Kim, Sydney Quantum Information Theory Workshop, (Jan. 22, 2015), URL `www.physics.usyd.edu.au/quantum/Coogee2015/Presentations/Kim.pdf`.

[19] I. H. Kim and B. J. Brown, Phys. Rev. B **92**, 115139 (2015), URL `https://link.aps.org/doi/10.1103/PhysRevB.92.115139`.

[20] C.-M. Jian, I. H. Kim, and X.-L. Qi, ArXiv e-prints (2015), `1508.07006`.

[21] K. Kato, F. Furrer, and M. Murao, Phys. Rev. A **93**, 022317 (2016), `1505.01917`.

[22] B. Shi and Y.-M. Lu, Phys. Rev. B **99**, 035112 (2019), `1801.01519`.

[23] B. Shi, Phys. Rev. Res. **1**, 033048 (2019), `1810.01986`.

[24] H. Bombin, Phys. Rev. Lett. **105**, 030403 (2010), URL `https://link.aps.org/doi/10.1103/PhysRevLett.105.030403`.

[25] S. Bravyi and A. Kitaev, ArXiv e-prints (1998), `quant-ph/9811052`.

[26] E. H. Lieb and M. B. Ruskai, J. Math. Phys. **14**, 1938 (1973).

[27] Y. Zhang, T. Grover, A. Turner, M. Oshikawa, and A. Vishwanath, Phys. Rev. B **85**, 235151 (2012), URL `https://link.aps.org/doi/10.1103/PhysRevB.85.235151`.

[28] J. Haah, Comm. Math. Phys. **342**, 771 (2016), ISSN 1432-0916, URL `https://doi.org/10.1007/s00220-016-2594-y`.

[29] X.-G. Wen, Int. J. Mod. Phys. B **04**, 239 (1990), `https://doi.org/10.1142/S0217979290000139`, URL `https://doi.org/10.1142/S0217979290000139`.

[30] J. M. Leinaas and J. Myrheim, Il Nuovo Cimento B (1971-1996) **37**, 1 (1977), ISSN 1826-9877, URL `https://doi.org/10.1007/BF02727953`.

[31] F. Wilczek, Phys. Rev. Lett. **49**, 957 (1982), URL `https://link.aps.org/doi/10.1103/PhysRevLett.49.957`.

[32] D. Petz, Q J MATH **39**, 97 (1988), ISSN 0033-5606, `http://oup.prod.sis.lan/qjmath/article-pdf/39/1/97/4559225/39-1-97.pdf`, URL `https://doi.org/10.1093/qmath/39.1.97`.

[33] D. Petz, Rev. Math. Phys. **15**, 79 (2003), `quant-ph/0209053`.

[34] Y.-K. Liu, in *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, edited by J. Díaz, K. Jansen, J. D. P. Rolim, and U. Zwick (Springer Berlin Heidelberg, Berlin, Heidelberg, 2006), pp. 438–449, ISBN 978-3-540-38045-0.

[35] K. Fredenhagen, K. H. Rehren, and B. Schroer, Comm. Math. Phys. **125**, 201 (1989).

[36] R. Haag, *Local Quantum Physics* (Springer Berlin Heidelberg, 1996).

[37] K. Kato and P. Naaijkens, ArXiv e-prints (2018), `1810.02376`.

[38] P. Naaijkens, *Quantum Spin Systems on Infinite Lattices: A Concise Introduction*, Lecture Notes in Physics (Springer International Publishing, 2017), ISBN 9783319514581, URL `https://books.google.co.jp/books?id=8YVtDgAAQBAJ`.

[39] M. Cha, P. Naaijkens, and B. Nachtergaele, Comm. Math. Phys. **373**, 219 (2019), `1804.03203`.

[40] P. Bonderson, K. Shtengel, and J. Slingerland, Annals of Physics **323**, 2709 (2008), ISSN 0003-4916, URL `http://www.sciencedirect.com/science/article/pii/S0003491608000171`.

[41] A. Kapustin and L. Fidkowski, Comm. Math. Phys. p. 257 (2019), `1810.07756`.

[42] A. Kapustin and L. Spodyneiko, arXiv e-prints arXiv:1905.06488 (2019), `1905.06488`.

[43] P. W. Anderson, Science **177**, 393 (1972).

[44] X.-G. Wen, *Quantum field theory of many-body systems: from the origin of sound to an origin of light and electrons* (Oxford University Press on Demand, 2004).

[45] J. Fröhlich and T. Kerler, *Quantum groups, quantum categories and quantum field theory* (Springer, 2006).

[46] P. Etingof, S. Gelaki, D. Nikshych, and V. Ostrik, *Tensor categories*, vol. 205 (American Mathematical Soc., 2016).

[47] V. Ostrik, arXiv preprint math/0203255 (2002).

[48] M. A. Levin and X.-G. Wen, Phys. Rev. B **71**, 045110 (2005), `cond-mat/0404617`.

[49] S. Bravyi, unpublished (2008).

[50] L. Zou and J. Haah, Phys. Rev. B **94**, 075151 (2016), `1604.06101`.

[51] D. J. Williamson, A. Dua, and M. Cheng, Phys. Rev. Lett. **122**, 140506 (2019), `1808.05221`.

[52] O. Fawzi and R. Renner, Comm. Math. Phys. **340**, 575 (2015).

[53] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, New York, NY, USA, 2011), 10th ed., ISBN 1107002176, 9781107002173.

[54] A. Uhlmann, Rep. Math. Phys. **9**, 273 (1976), ISSN 0034-4877, URL `http://www.sciencedirect.com/science/article/pii/0034487776900604`.

[55] H. Araki and E. H. Lieb, Comm. Math. Phys. **18**, 160 (1970).