学位論文

Correlations in quantum many-body systems and quantum information theoretic measures

(量子多体系における相関と量子情報的測度)

平成 28 年 12 月博士（理学）申請

東京大学大学院理学系研究科
物理学専攻

加藤　晃太郎

Acknowledgment

First and foremost, I would like to express my great appreciations to my supervisor, Prof. Mio Murao, for her guidance and kind supports throughout the past five years. She gave me the freedom to choose my research topic, and supported to pursue it in my own way. My special thanks go to Dr. Fabian Furrer, who has warmly helped me a lot from when I had started to do scientific research. Their insightful advices had significant impact on both my research and career. Especially I thank them for improving my presentation and writing skills patiently. I should thank to Dr. Akihito Soeda for supporting to improve my applications, papers, presentations and this thesis.

I also owe many thanks to my collaborator, Prof. Fernando Brandao, for being my mentor during the twelve-week internship at Quantum Architectures and Computation Group (QuArC) at Microsoft Research. I was always impressed with his interesting ideas, deep insight and creativity. It was very fortunate to work with him. I sincerely thank the researchers at QuArC, including Dr. Krysta Svore, Dr. Nathan Wiebe and Dr. Matthew Hastings for their hospitality. I also thank other interns at Microsoft Research, especially Aleksander Kubica, Michael Beverland, Serwan Asaad, Shawn Cui and Bettina Heim for being good friends. It was a great pleasure for me to have an opportunity to work with these people in such a wonderful environment.

I am grateful to all past/current members of the research group of Prof. Mio Murao, including Dr. Yoshifumi Nakata, Dr. Takanori Sugiyama, Dr. Shojun Nakayama, Dr. Eyuri Wakakuwa, Dr. Seiseki Akibue, Jisho Miyazaki, Kosuke Nakago, Atsushi Shimbo, Yuki Mori, Hayata Yamasaki, Ryosuke Sakai, Matthias Strodtkotter and Qingxiuxiong Dong. All conversations, travels and parties together with my colleagues made the past five years really fruitful and enjoyable.

I am indebted to Dr. Isaac Kim for accepting me to visit to the Perimeter Institute and all the discussions during that period. I also wish to thank Dr. Beni Yoshida for having valuable discussions during the visit, and also in the other conferences.

This work was supported by JSPS Research Fellowships and Advanced Leading Graduate Course for Photon Science. I am grateful to these programs for making completing my Ph.D course possible with financial supports.

Finally, I would like to express my gratitude to my family and friends for being my best emotional support.

Abstract

Correlations between constituents are sources of various phenomena in matters. Generic properties of multipartite correlations in quantum many-body systems are far from fully understood in contrast to bipartite correlations. In this thesis, we study multipartite correlations in quantum many-body systems by using theoretical tools developed in quantum information theory. We especially focus on (i) ground states of gapped short-range Hamiltonians of two-dimensional (2D) spin systems and (ii) Gibbs states of short-range Hamiltonians of one-dimensional (1D) spin systems.

2D gapped systems can exhibit exotic quantum phases known as topologically ordered phases. One characteristic feature of topologically ordered phases is the existence of large-scale multipartite correlations which appear on topologically non-trivial regions. The first aim of this thesis is establishing information-theoretic characterization of such correlations. A quantity called the topological entanglement entropy has been considered to characterize the large-scale correlations, but it is unclear whether the quantity has information-theoretical meanings. To address this issue, we consider a classification of multipartite correlations based on a geometric measure called the irreducible correlation. We show the equivalence between the topological entanglement entropy and the irreducible correlation in states with zero correlation length, such as ground states in exactly solvable models. This suggests that the characteristic large-scale correlation can be well-classified by using the irreducible correlation. We then apply this equivalence to study of the entanglement spectrum on a half cylinder, and show that the value of the topological entanglement entropy provides a restriction on the entanglement spectrum. We further show that the value of the topological entanglement entropy can be interpreted in terms of the optimal rate of a quantum information protocol referred as secret sharing. This result establishes an operational characterization of the characteristic correlations in topologically ordered phases. In these investigations, a special property of multipartite correlations in many-body states called the Markov property has been extensively analyzed. States satisfying the Markov property are known as Markov chains or Markov networks, and their properties have been substantially studied. Theoretical techniques to treat states approximately satisfying the Markov property has been recently developed. Motivated by these recent developments, we extend some of our results for zero-correlation length to the case of finite-correlation length.

The second aim of this thesis is characterizing states which approximately satisfy the Markov property. In classical information theory, distributions satisfying the Markov property are characterized by the Gibbs distributions of short-range Hamiltonians and vice versa. This relation has been extended to quantum systems as well, but only pertains to exact Markov property and commuting Hamiltonians. In this thesis, we extend the relation to 1D states approximately satisfying the Markov property to and 1D Gibbs states of general short-range Hamiltonians. Our results establish an alternative characterization of states approximately satisfying the Markov property, and also show a generic property of multipartite correlations for 1D Gibbs states.

Contents

Chapter 1

Introduction

1.1 Overview

Quantum information theory is an interdisciplinary field bridging quantum physics and information theory. The central idea of quantum information theory is utilizing well-controlled quantum systems for performing information-processing such as computation and cryptography. One can naturally expect that quantum information processing outperforms classical counterparts by using characteristic properties of quantum systems not existing in classical ones, e.g., existence of superposition and quantum correlations. In 1994, Shor discovered the celebrated factoring algorithm [1] for quantum computers that can solve the factoring problem in polynomial time. This quantum algorithm exceeds, with sub-exponential speed up, the performance of the current best known classical algorithm. For cryptography, quantum key distribution [2, 3, 4] enables to share secret random bits for communication by utilizing superpositions of quantum states. In contrast to classical protocols relying on the limitation of computational power of eavesdroppers such as the RSA public-key cryptosystem [5], the security of quantum key distribution is guaranteed by the law of quantum mechanics. Another emerging application of quantum information theory is quantum metrology [6, 7], which exploits quantum correlations to perform high-precision measurements. Motivated by these potential advantages of using quantum systems for information processing, many efforts have been made to analyze and implement quantum information processing.

Information theory provides a theoretical framework and tools to analyze information, quantitatively and independently of actual implementations of information processing. These information-theoretical methods have been utilized in quantum information theory to clarify the differences between quantum and classical systems and understand the origin of quantum advantages in information processing.

In quantum physics, correlations are usually quantified by the correlation functions defined in terms of expectation values of observables. In contrast, correlations in quantum information theory are mainly referred to properties of quantum states that determine statistical correlations between measurement outcomes. Especially a class of quantum correlations between distinct subsystems called *entanglement* has been intensively investigated. Such correlations in states are analyzed as the necessary *resource* for performing various quantum information tasks, for example, quantum teleportation [8]. For this reason, these correlations in states are quantified by *operationally defined* functions via some specific information-processing tasks. Aside from their practical usefulness, operational characterizations of correlations have been also studied to understand fundamental aspects of quantum mechanics [9, 10].

Simultaneously, theoretical tools developed in quantum information theory have been applied to other fields of physics, e.g., gravity theory [11, 12, 13], chemical physics [14, 15], statistical mechanics [16, 17, 18] and condensed matter physics [19, 20]. A notable example is the applications of entanglement theory for analyzing ground states of a gapped system (gapped ground states, for short) in quantum many-body physics. A commonly used measure of entanglement in these analysis is the entanglement entropy for bipartite pure states. The entanglement entropy is used to quantify the amount of entanglement between a subsystem of a many-body system and its complement. It has been observed that for gapped ground states, the entanglement entropy typically scales proportionally to the size of the perimeter of the subsystem, not the volume of the subsystem (such a scaling is first analyzed in the study of entropy of black holes [21]). This behavior of entanglement is called the area law. Intuitively, the area law asserts that entanglement contained in the state is only short-ranged and thus correlations across the boundary are dominated by the contribution from the degrees of freedom near the boundary. The area law of gapped ground states is in contrast to the entanglement property of almost all pure states in the many-body Hilbert space, which behave according to the volume law [22].

An area law is rigorously proven for the ground states of one-dimensional (1D) gapped systems [23, 24]. Besides of the fundamental interest, the 1D area law is also related to numerical simulatability of many-body states. It is proven that the area law of a 1D ground state implies that the state can be well-approximated by a matrix product state (MPS) [25]. MPS is a state where the coefficients of the state in a product basis are written as a product of small matrices. An example (and the first appearance) of MPS is the Affleck-Lieb-Kennedy-Tasaki (AKLT) state [26]. In practice, MPS is used as an ansatz state in the density matrix renormalization-group (DMRG) [27, 28], which has been successful at simulating 1D ground states of gapped systems. This observation is a theoretical consequence of the area law

and its relation to the MPS approximation.

For gapless (i.e., critical) 1D systems, generic properties of entanglement are less known. Several numerical and analytical calculations show that the area law is violated up to a logarithm factor [29, 30, 31]. A similar logarithmic violation of the area law is observed in conformal field theory (CFT) [32, 33], and the prefactor is related to the central charge which characterizes the universality class of the theory [29]. The logarithmic violation of the area law also holds for 1D states described by the multi-scale entanglement renormalization ansatz (MERA) [34]. The agreement of entanglement scaling suggests that 1D critical systems are effectively well-described by CFT and MERA.

In higher-dimensional systems, the area law for general gapped ground states has not been rigorously established. However, it is proven for several specific models or settings [35, 36, 37, 38]. In one example, projected entangled pair states (PEPS) [39, 40] which are extensions of MPS to higher-dimensional systems satisfy the area law by construction. The logarithmic violation of the area law is observed for specific models of critical systems as in 1D critical systems [41, 42, 43]. There also exists a critical 2D model exhibiting the area law [40].

When a many-body state is in thermal equilibrium, namely, when the state is given by a Gibbs state, the entanglement entropy is no longer a valid measure of entanglement or correlations. Instead, the mutual information is used to quantify the total amount of bipartite correlations which includes both quantum and classical correlations. For a system modeled by a short-range Hamiltonian, it is proven for any dimension that the scaling of the mutual information obeys an area law [44]. This situation differs from the area law of entanglement for any dimension being still a conjecture. The area law of the mutual information strengthens the intuition that correlations in Gibbs states of short-range Hamiltonians are dominated by "short-range" correlations.

Most analysis of correlations in many-body systems to date have been focused on two-point or bipartite correlations, although more general correlation properties are necessary to fully characterize many-body states. Generic properties of multipartite correlations are far from being well-understood, but investigating multipartite correlations may open a new avenue of characterizing quantum many-body systems. Analyzing or computing a measure of multipartite correlations is in general a demanding task especially when the number of subsystems is large. However, as represented by the area laws for gapped ground states and Gibbs states, one might expect that the structure of most of physically realizable states is strongly restricted. Indeed, the area laws of the entanglement entropy and the mutual information guarantee that a particular tripartite correlation measure called conditional mutual information is small, or sometimes exactly zero for a certain settings in these states. When a state has zero conditional mutual in-

formation, it is called a (spatial) *Markov chain* which has been widely utilized in statistics, information theory, physics and beyond. The structural characterization of Markov chains in classical systems has been well studied and established. The first characterization of the quantum version of Markov chains is given in Ref. [45], where all quantum Markov chains are characterized by a particular decomposition of the state and a property called local recoverability, similarly to classical Markov chains. By using the restriction on the conditional mutual information, several generic properties of gapped ground states have been revealed [46, 47, 48, 49].

A particularly interesting class of states containing non-trivial multipartite correlations in gapped systems is ground states in *topologically ordered phases.* Any ground state in a topologically ordered phase should be robust against any local perturbations. This makes such a ground state a promising candidate for storing and processing quantum information coherently [50, 51]. Ground states in topologically ordered phases have only short-range two-point correlations and exhibit the area law. However, observations from known models suggest that they contain large-scale multipartite correlations in loop-like regions, not exhibited by gapped ground states in topologically-trivial phases. In addition, these characteristic multipartite correlations are considered to be a necessary resource for topological protection of quantum information [48].

Interestingly, the existence of the characteristic multipartite correlations affects the value of the entanglement entropy. It has been observed [53] that the entanglement entropy of states in topologically ordered phases contains a constant term that is believed to characterize the phase. The constant is called the topological entanglement entropy [55] (TEE) and has been used as an indicator of topologically ordered phases in a number of numerical researches [56, 57, 58, 59, 60]. The TEE is extracted by taking a linear combination of the entanglement entropy associated to subsystems by considering a particular choice of the subsystems. The linear combination of the entanglement entropy is independently proposed by Levin and Wen [54], and called the topological entropy. Therefore, the TEE is equivalent to the topological entropy up to irrelevant technical conditions. The topological entropy can also be considered as a quantum analog of a function called multivariate mutual information or interaction information in classical information theory [61], which is used to analyze multipartite correlations.

It may be that the TEE provides a rigorous definition of the characteristic multipartite correlations in topologically ordered phases via the equivalence to the topological entropy. However, information-theoretic meaning of the multivariate mutual information has not been established even in classical cases [62, 63]. For this reason, the present definition poses a challenge when attempting to further reveal other consequences of the existence of the characteristic multipartite correlations. As example, an extension of the TEE for mixed states remains unsettled,

which however would be necessary when analyzing quantum information stored at finite temperature. The TEE is originally defined via the entanglement entropy, but it loses its meaning as a measure of entanglement for mixed states. A possible extension of the TEE based on the entanglement entropy is proposed by using the mutual information [64]. The TEE can also be interpreted as the multivariate mutual information, which is applicable for mixed states but with no known information-theoretic meaning. In general, these two extensions of the TEE give different values, however both quantities still lack an acceptable justification much less any reason to regard either of these as more appropriate. These disadvantages of the TEE make it difficult to accept it as a proper classification/quantification of the characteristic multipartite correlations.

To fully understand the role of the characteristic multipartite correlations in topologically ordered phases, we propose to characterize these correlations by quantities which are already known to have various applications in information theory. Since information-theoretic quantities are often defined for general quantum states including mixed states, they are naturally applicable for thermal states. On the other hand, the TEE has been successfully employed in practice, suggesting that an information-theoretically meaningful interpretation of the TEE is possible under specific settings. If we can obtain an information-theoretical meaning of the TEE, we can specify what class of multipartite correlations is represented by the value of the TEE. In this thesis, we investigate the information-theoretical characterization of the TEE motivated by these considerations.

To analyze complex multipartite correlations, one strategy is to decompose into different classes of multipartite correlations. In classical information geometry, Amari treats a hierarchical decomposition of parametrized Gibbs distributions in Ref. [66]. The hierarchical decomposition is given by a classification of patterns of interactions in the corresponding "Hamiltonians" of the Gibbs distributions. For instance, if a Gibbs distribution corresponds to a Hamiltonian only containing bipartite interactions, then we can say that the multipartite correlation in the distribution has a bipartite origin. Amari has shown that these hierarchical structures can be used to decompose the total correlations in a given multivariate distribution into a sum of correlations originating from pairwise, triplewise, and further higher-order interactions. Each function quantifying the kth-order contribution is called the kth-order effect or connected correlation of order k [67]. The quantum analog of Amari's framework has been discussed by several authors [69, 70]. The analog of k-th order effect is sometimes called the *irreducible correlation* of order k. Their mathematical foundation is discussed in Refs. [71, 72, 73].

The irreducible correlation and similar functions have been employed to quantify the characteristic multipartite correlations under a conjecture of the equivalence to the TEE for gapped 2D ground states [74, 75]. If this conjecture is true,

the TEE has an information-theoretical meaning in terms of interaction patterns of Hamiltonians. To investigate the conjecture, we first focus on many-body states with an exactly vanishing correlation length as an ideal case. This condition is satisfied when the system is described by a gapped exactly solvable models such as the quantum double models [50] or the Levin-Wen models [76], for example. We show that the TEE and the irreducible correlation are equivalent for these cases. Via the connection between the irreducible correlation and the interaction patterns of Hamiltonians, this result implies that the entanglement Hamiltonian, i.e., the logarithm of a reduced state of the ground state, contains a global many-body interaction in topologically ordered phases. We further provide another information-theoretic characterization of the TEE by connecting the TEE to a quantum information processing task called secret sharing. Certain reduced states of gapped ground state with a zero correlation length form Markov chains. We employ the structure of Markov chains to obtain these results.

In more realistic cases, the conditional mutual information of states is generally small but not exactly zero. In the case of classical systems, it is easy to show that states with small conditional mutual information is close to a Markov chain. Therefore, all properties of Markov chains approximately hold for such states. However, in quantum systems this is not the case [77] and a characterization of these states had been a long standing problem in quantum information theory. Recently, Fawzi and Renner have shown [78] that the local recoverability approximately holds even in quantum case. We extend some of our results in the zero correlation length setting to allow sufficiently small error due to non-zero correlation length by applying the Fawzi-Renner result.

In classical information theory, another characterization of Markov chains is given by the Hammersley-Clifford theorem [79] that states all Markov chains (or more generally, Markov networks) are equivalent to Gibbs states of short-range Hamiltonians. This theorem is generalized to quantum Markov chains as well [80, 81], which gives equivalence between quantum Markov chains and quantum Gibbs states of short-range commuting Hamiltonians. Therefore, these theorems also indicate that any Gibbs states with short-range commuting Hamiltonians have always vanishing conditional mutual information under certain choice of regions. This property of Gibbs states is especially important to numerically simulate Gibbs states, since it implies that Gibbs states can be simulated by locally "patching" small regions [82]. A remaining question is whether we can still characterize the states with small conditional mutual information in terms of Gibbs states of short-range Hamiltonians.

We provide an affirmative answer to this question by generalizing the quantum version of the Hammersley-Clifford theorem. As already mentioned, the area law of the mutual information indicates one direction of the Hammersley-Clifford

theorem, i.e., any Gibbs state of a short-range Hamiltonian has small conditional mutual information. However, the area law is too crude in practice, that is, it does not show how small the conditional mutual information is. As a result, we provide the very first explicit upper bound on the conditional mutual information for general 1D Gibbs states of short-range Hamiltonians. We further show that any state with small conditional mutual information for certain settings can be well-approximated by a 1D Gibbs state of a short-range Hamiltonian. This gives a generalization of the Hammersley-Clifford theorem to general 1D quantum systems.

1.2 Organization of This Thesis

The structure of this thesis is as follows. We introduce basic notions in quantum information theory in Chap. 2. Especially theoretical tools in Sec. 2.2.3 and in Sec. 2.3 are frequently used in Chap. 4 and Chap. 5. We review basic preliminaries for topologically ordered phases in Chap. 3. In Chapter 4, we study the irreducible correlation and the optimal rate of a secret sharing protocol in topologically ordered phases. We investigate the conditional mutual information of 1D Gibbs state and its relation to an approximate version of Markov chains in Chap. 5. We summarize our results in Chap. 6.

1.3 Notations

We will use the following notation throughout this thesis.

$[n]$	the set $\{1, 2, ..., n\}$
X^c	the complement of a set X
$A^\dagger$	the conjugate transpose of A.
$\lambda(A)$	the set of the spectrum of A.
$\mathcal{H}_A$	a finite dimensional Hilbert space corresponding to a system A.
$\mathcal{S}(\mathcal{H})$	the set of all states on $\mathcal{H}$.
$\mathbb{1}_A$	the identity operator on $\mathcal{H}_A$.
Tr_A	the partial trace over $\mathcal{H}_A$.
$\mathrm{id_A}$	the identity map acting on linear operators on $\mathcal{H}_A$.
$\|A\|_1$	the trace norm of operator A.
$\|A\|$	the operator norm of operator A.
$G = (V, E)$	a graph with the set of vertices V and the set of edges E.
$d(A, B)$	the number of edges in a shortest path connecting two sets A, B in G.
ρ_A	the reduced state on A.
ρ^H	the Gibbs state $\rho^H := \frac{e^{-\beta H}}{\mathrm{Tr}e^{-\beta H}}$ of a Hamiltonian H at temperature β.

Chapter 2

Preliminaries of Quantum Information

In this chapter, we introduce the terminology and notations widely used in quantum information. In Sec. 2.1, we briefly summarize the basic descriptions of quantum states and physical operations on state spaces. In Sec. 2.2, we introduce entropic measures of correlations commonly used in quantum information. We explain the definition and equivalent conditions of quantum Markov chains in Sec. 2.2.3 We review useful information-geometrical properties of quantum Gibbs states in Sec. 2.3, which are intensively used in this thesis.

2.1 States and Operations

In this section, we summarize an axiomatic formalism of quantum mechanics for finite-dimensional systems. Note that another formulation is required to describe infinite-dimensional systems (see, e.g., [83]).

2.1.1 Quantum Systems and Quantum States

The first postulate of quantum mechanics is the following:

Postulate 1. *For a given system, there exists a corresponding Hilbert space* $\mathcal{H} = \mathbb{C}^d$ *with dimension* $d < \infty$*. The state of the system is represented by a positive semidefinite Hermitian operator* $\rho \geq 0$ *with unit trace* $\mathrm{Tr}\rho = 1$ *acting on* $\mathcal{H}$*.*

The operator ρ representing a state is also called a density matrix. In the following, by *system* we do not only refer to the physical system but also to the corresponding Hilbert space for simplicity.

When a state ρ is a rank-1 operator, there exists a complex unit vector $|\psi\rangle \in \mathcal{H}$ such that $\rho = |\psi\rangle\langle\psi|$. We refer such a state to as a *pure state* and interchangeably represent it by vector $|\psi\rangle$. When the rank of ρ is strictly larger than 1, we call it *mixed state*. In particular, we refer a mixed state $\frac{1}{d}\mathbb{1}$ on $\mathbb{C}^d$ as a *completely mixed state*. We denote the set of all states on $\mathcal{H}$ by $\mathcal{S}(\mathcal{H})$.

Any state can be regarded as a probabilistic mixture of different pure states. Therefore, for any density matrix $\rho \in \mathcal{S}(\mathcal{H})$, there exist a probability distribution $\{p_i\}_i$ and a set of pure states $\{|\psi_i\rangle\}_i$ such that

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \,. \tag{2.1}$$

Note that the decomposition of (2.1) is only unique for pure states. For general mixed states, there are infinite numbers of ensembles which describe the same state.

The second postulate states about composite physical systems.

Postulate 2. *The Hilbert space of a composite physical system consisting of subsystems $\mathcal{H}_A$ and $\mathcal{H}_B$ is given by $\mathcal{H}_A \otimes \mathcal{H}_B$.*

If a state of one of the systems is $\rho_1 \in \mathcal{S}(\mathcal{H}_1)$ and another system is $\rho_2 \in \mathcal{S}(H_2)$, and they are independent, the total state of the combined system is represented as

$$\rho_1 \otimes \rho_2 \,. \tag{2.2}$$

We call such a state a *product state*.

A probabilistic mixture of product states represented by

$$\sum_i p_i \rho_1^i \otimes \rho_2^i \tag{2.3}$$

is called *separable* states. A state which is not a separable state is called an *entangled* state. A generalization to N-composite systems is straightforward. For example, a separable state is a state represented by

$$\sum_i p_i \rho_1^i \otimes \cdots \otimes \rho_N^i \,. \tag{2.4}$$

When $\rho = |\psi\rangle\langle\psi| \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is a pure state, it is separable if and only if

$$|\psi\rangle = |\psi\rangle_A \otimes |\phi\rangle_B \,, \tag{2.5}$$

where $|\psi\rangle_A$ ($|\psi\rangle_B$) is a pure state on $\mathcal{H}_A$ ($\mathcal{H}_B$). This implies that all separable pure states are product states.

Entanglement is regarded as an important "resource" in quantum information, which is consumed to perform certain quantum information processing, such as quantum teleportation [8]. According to this aspect of entanglement, quantifying entanglement of quantum states is necessary. We will discuss about an approach to this problem using entropic functions in Sec. 2.2.2.

2.1.2 Operations

Transformation of quantum states are caused by time evolutions or measurements. We refer a map describing state transformations to as *quantum operation*. In the following we introduce mathematical representations for quantum operations.

CPTP-maps

In this thesis, time evolutions denote deterministic quantum operations such as unitary evolutions of closed systems and also non-unitary evolutions of open systems. Mathematically, such a deterministic quantum operation is represented by a linear map from a state space to state space satisfying completely-positive (CP) and trace-preserving (TP) conditions. For this reason, we call a map representing a quantum operation a CPTP-map. Consider a linear map $\Lambda : \mathcal{S}(\mathcal{H}_A) \to \mathcal{S}(\mathcal{H}_B)$ and some Hilbert space $\mathcal{H}$. CP means that for any n, $\Lambda \otimes \mathrm{id}_{\mathcal{H}}^{\otimes \mathrm{n}} : \mathcal{S}(\mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}^{\otimes \mathrm{n}}) \to \mathcal{S}(\mathcal{H}_{\mathrm{B}} \otimes \mathcal{H}^{\otimes \mathrm{n}})$ is positive, i.e., it maps a positive semidefinite operator to a positive semidefinite operator. TP means that the trace is conserved which ensures that the normalization condition at the output holds.

The most fundamental class of CPTP-maps is unitary evolutions in a closed system.

Postulate 3. *Any time evolution of a state in a closed system is described by a unitary operator acting on the corresponding Hilbert space.*

General CPTP-maps have several representations. For example, all quantum operations can be composed from three elementary operations $(O1) - (O3)$.

$(O1)$ Adding an uncorrelated ancilla[1]:
$\Lambda_\sigma : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H} \otimes \mathcal{H}_a)$, $\Lambda_\sigma(\rho) = \rho \otimes \sigma$, where $\mathcal{H}_a$ is the ancilla system and σ is the state of the ancilla.

$(O2)$ Tracing out a part of the system:
$\Lambda_B : \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B) \to \mathcal{S}(\mathcal{H}_A)$, $\Lambda_B(\rho_{AB}) = \mathrm{Tr}_B \rho_{AB}$, where $\rho_{AB} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$

[1] An ancilla is a fixed state of an auxiliary (ancilla) system.

and Tr_B denotes the partial trace over the system H_B which is defined as

$$\mathrm{Tr}_B \rho_{AB} = \sum_i \left(I \otimes \langle i|\right) \rho_{AB} \left(I \otimes |i\rangle\right) , \tag{2.6}$$

where $\{|i\rangle\}$ is an arbitrary orthonormal basis of $\mathcal{H}_B$. We will denote *reduced state* $\mathrm{Tr}_B \rho_{AB}$ by ρ_A.

($O3$) Unitary transformations:
$\Lambda_U : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H})$, $\Lambda_U(\rho) = U\rho U^\dagger$, where U is a unitary operation on $\mathcal{H}$.

For any CPTP-map $\Lambda : \mathcal{S}(\mathcal{H}_A) \to \mathcal{S}(\mathcal{H}_B)$, there exist a Hilbert space $\mathcal{H}_C$, a pure state $\rho_0 \in \mathcal{S}(\mathcal{H}_B \otimes \mathcal{H}_C)$ and a unitary U_Λ on $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$ such that [84]

$$\Lambda(\rho) = \mathrm{Tr}_{AC} U_\Lambda \left(\rho \otimes \rho_0\right) U_\Lambda \, . \tag{2.7}$$

Furthermore, Λ has another representation called the Kraus representation:

$$\Lambda(\rho) = \sum_{i=1}^{d} K_i \rho K_i^\dagger \, , \tag{2.8}$$

where $d \leq \dim \mathcal{H}_A \dim \mathcal{H}_B$ and each K_i is a $\dim \mathcal{H}_A \times \dim \mathcal{H}_B$ matrix such that $\sum_i K_i^\dagger K_i = \mathbb{1}_A$.

Quantum Measurement

To extract information about states, we need to perform measurements. If one is not interested in the post-measurement state, the description of a quantum measurement is given in terms of a set of positive operators representing probabilistic maps.

Postulate 4. *A quantum measurement is described by a set of positive operators* $\{M_m\}$ *such that*

$$\sum_m M_m = \mathbb{1} \, . \tag{2.9}$$

Each measurement outcome m *is obtained with probability given by*

$$p(m) = \mathrm{Tr}[M_m \rho] \, . \tag{2.10}$$

A set of operators $\{M_m\}$ corresponding to a measurement is called a *positive operator-valued measure (POVM).*

The general description of a quantum measurement describing probabilistic state transformations is given by *instruments.* An instrument $\{\kappa_m\}_m$ is a set

of CP-maps κ_m such that $\sum_m \kappa_m$ is TP. Consider we perform a measurement corresponding to an instrument $\{\kappa_m\}$ on a state ρ. A post-measurement state

$$\rho_m := \frac{\kappa_m(\rho)}{\mathrm{Tr}[\kappa_m(\rho)]}\,, \tag{2.11}$$

is obtained with a probability $\mathrm{Tr}[\kappa_m(\rho)]$.

Distance measures

To quantify "closeness" of two quantum states ρ and σ, the trace distance:

$$D(\rho,\sigma) := \frac{1}{2}\|\rho-\sigma\|_1 \tag{2.12}$$

$$= \frac{1}{2}\sqrt{(\rho-\sigma)^\dagger(\rho-\sigma)} \tag{2.13}$$

is often used in quantum information. For simplicity, we shall consider the difference in terms of trace norm $\|\rho-\sigma\|_1$ rather than the half of it. The trace distance (and therefore $\|\rho-\sigma\|_1$) satisfies the conditions for a distance measure: the non-negativity, symmetry under the exchange of the entries and the triangle inequality. Another important property of the trace distance is the monotonicity under CPTP-maps. Performing the same CPTP-map on two states makes the states more indistinguishable, since it basically adds "noise" to them. This is mathematically represented as

$$D(\rho,\sigma) \geq D(\Lambda(\rho),\Lambda(\sigma))\,, \tag{2.14}$$

where $\rho,\sigma \in \mathcal{S}(\mathcal{H})$ and $\Lambda : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H}')$ is a CPTP-map. The equality holds if and only if Λ is a unitary transformation, or an isometry in general.

Another measure of the distance between two quantum states often used in quantum information is the fidelity:

$$F(\rho,\sigma) := \|\sqrt{\rho}\sqrt{\sigma}\|_1\,. \tag{2.15}$$

The fidelity is essentially equivalent to the trace distance in the sense that

$$1-F(\rho,\sigma) \leq D(\rho,\sigma) \leq \sqrt{1-F(\rho,\sigma)^2} \tag{2.16}$$

holds.

2.2 Entropy and Correlation Measures

2.2.1 Entropy

Entropy is a key concept in information theory. Operationally, it represents the asymptotic optimal rate of data compression [85]. It is also used to define various measures of correlations between random variables/quantum systems. In this section, we introduce definitions and properties of entropy and related functions.

The Shannon entropy of a discrete probability distribution $p = \{p_i\}$ is defined as

$$H(p) := -\sum_i p_i \log_2 p_i \,. \tag{2.17}$$

On the other hand, the von Neumann entropy for a quantum state ρ is similarly defined as

$$S(\rho) := -\mathrm{Tr}\rho \log_2 \rho \,. \tag{2.18}$$

The von Neumann entropy is a non-negative and concave function as well as the Shannon entropy and coincides to the Shannon entropy of the eigenvalues of the state $\{\lambda_i\}_{\lambda_i \in \lambda(\rho)}$. This function is not only mathematically useful but also operationally meaningful. It was shown that if N copies of a state ρ are given, they can be compressed into about $NS(\rho)$ qubits [2] with an error vanishing in the limit of $N \to \infty$ [85].

For a conditional probability distribution $p_X(x|y)$ of two random variables X and Y, one can consider the average of the Shannon entropy over $p_Y(y)$ as the conditional entropy:

$$H(X|Y)_p := \sum_y p_Y(y) \left(-\sum_x p_X(x|y) \log_2 p_X(x|y) \right) \tag{2.19}$$

$$= H(XY) - H(Y) \,, \tag{2.20}$$

where $H(XY)$ is the Shannon entropy of the joint distribution $p(x, y)$. Unfortunately, the concept of the conditional distribution may not be straightforwardly extended to quantum states [86]. We define a quantum version of the conditional entropy for a quantum state $\rho_{AB} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ by

$$S(A|B)_\rho := S(AB)_\rho - S(B)_\rho \,, \tag{2.21}$$

where $S(X)_\rho$ is the von Neumann entropy of the (reduced) state on system X. This function can be negative while the classical analog $H(X|Y)_p$ is always positive.

[2] A qubit is a two-dimensional quantum system.

This is due to the presence of entanglement [87]. For two quantum states ρ_{AB} and σ_{AB} satisfying $\|\rho_{AB} - \sigma_{AB}\|_1 \leq \epsilon \leq 1$, the (Alicki-)Fannes inequality:

$$|S(A|B)_\rho - S(A|B)_\sigma| \leq 4\epsilon \log_2 d_A + 2h_2(\epsilon) \tag{2.22}$$

$$\leq 6\sqrt{\epsilon} \log_2 d_A \tag{2.23}$$

holds, where $d_A = \dim \mathcal{H}_A$ and $h_2(\epsilon)$ is the Shannon entropy of the probability distribution $\{\epsilon, 1-\epsilon\}$ (the binary entropy function). This inequality guarantees that if two states are close, their conditional entropy are also close. Note that if ρ_{AB} is a product state, the conditional entropy reduces to the von Neumann entropy on system A.

The (quantum) relative entropy is an entropic function describing "closeness" of two quantum states. For states ρ and σ, it is defined as

$$S(\rho\|\sigma) := \begin{cases} \mathrm{Tr}[\rho(\log_2 \rho - \log_2 \sigma)] & \text{if } \mathrm{supp}(\rho) \subset \mathrm{supp}(\sigma)\,, \\ \infty & otherwise \end{cases} \tag{2.24}$$

Here, $\mathrm{supp}(\rho)$ is the subspace spanned by the eigenvectors of ρ with non-zero eigenvalues. We will use "supp" in the different definition in Chap. 5. The relative entropy is a non-negative function and is zero if and only if two states are the same states. For this reason, the relative entropy is often treated as a (psudo) distance between two states although it is not symmetric and does not satisfy the triangle inequality. An operational meaning of the relative entropy is provided via hypothesis testing [88]. In that context, the relative entropy represents how two quantum states are (asymptotically) distinguishable by hypothesis testing. As well as the trace distance, monotonicity of the relative entropy holds, namely, for arbitrary states $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ and any CPTP-map $\Lambda : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H}')$, it holds that

$$S(\rho\|\sigma) \geq S\left(\Lambda(\rho)\|\,\Lambda(\sigma)\right)\,. \tag{2.25}$$

Importantly, the relative entropy is related to the trace norm through the Pinsker inequality:

$$S(\rho\|\sigma) \geq \frac{1}{2}\|\rho - \sigma\|_1^2\,. \tag{2.26}$$

Therefore, if the relative entropy of two states is small, the two states are close in the sense of the trace norm too. Note that in general one cannot upper bound the relative entropy in terms of the trace norm.

2.2.2 Correlation Measures

In standard quantum mechanics, the (connected) correlation function for two observables X and Y is defined as

$$Cor(X,Y)_\rho := \langle XY \rangle - \langle X \rangle \langle Y \rangle\,, \tag{2.27}$$

where $\langle X \rangle := tr[\rho X]$. The correlation function is easy to calculate and characterizes different phases of matter.

If a multipartite state is not a product state, there exists a set of observables of which corresponding measurement outcomes are statistically correlated. In this sense, we say a state is correlated or correlations are contained in a state if the state is not a product state. This is in contrast to correlations in terms of the correlation function, which depends on observables. To quantify correlations in states, entropic functions are commonly used in information theory.

A common measure of correlations contained in a bipartite state ρ_{AB} is the mutual information:

$$I(A:B)_\rho := S(A)_\rho + S(B)_\rho - S(AB)_\rho \,. \tag{2.28}$$

In terms of the relative entropy, this function represents the distance from the set of all product states, i.e., it holds that

$$\min_{\sigma_A,\sigma_B} S(\rho_{AB}\|\sigma_A \otimes \sigma_B) = S(\rho_{AB}\|\rho_A \otimes \rho_B) \tag{2.29}$$

$$= I(A:B)_\rho \,. \tag{2.30}$$

The first line follows from an inequality $\mathrm{Tr}(\rho \log \sigma) \leq \mathrm{Tr}(\rho \log \rho)$ and simple calculations. From this representation, it is clear that the mutual information of a state ρ_{AB} is a symmetric and non-negative function which vanishes if and only if the state is the product state $\rho_{AB} = \rho_A \otimes \rho_B$. By monotonicity of the relative entropy (2.25), the mutual information is non-increasing under local CPTP-maps acting on A or B. The mutual information represents how much randomness do we need to remove the correlation. See Ref. [89] for details. For arbitrary observables M_A and M_B on system A and B individually, the mutual information of a state ρ_{AB} satisfies [44]

$$I(A:B)_\rho \geq \frac{Cor(M_A, M_B)^2}{\|M_A\|^2 \|M_B\|^2} \,. \tag{2.31}$$

In general, there are states with an arbitrary small values of the right hand side but highly correlated in terms of the mutual information. Such states are called data hiding states [90].

The conditional mutual information quantifies correlations between two subsystems in the presence of the third system. For a probability distribution $p_{XYZ}(x, y, z)$, it can be defined as the average of the mutual information of a conditional distribution:

$$H(X:Z|Y)_p := \sum_y p_Y(y) H(X:Z)_{p_{XZ}(x,z|Y=y)} \tag{2.32}$$

$$= H(XY)_p + H(YZ)_p - H(Y)_p - H(XYZ)_p \,. \tag{2.33}$$

By definition, it is clear that $H(X:Z|Y)_p$ is a non-negative function. The quantum version of the conditional mutual information is then defined as

$$I(A:C|B)_\rho := S(AB)_\rho + S(BC)_\rho - S(B)_\rho - S(ABC)_\rho \,. \tag{2.34}$$

In contrast to the classical case, showing non-negativity of this function is a non-trivial problem. The non-negativity of the quantum conditional mutual information is guaranteed by a non-trivial inequality called the strong subadditivity:

$$S(AB)_\rho + S(BC)_\rho \geq S(B)_\rho + S(ABC)_\rho \,. \tag{2.35}$$

A useful property of the conditional mutual information is the chain rule. For the conditional mutual information $I(A_1A_2...A_n:C|B)_\rho$ of $\rho_{A_1A_2...A_nBC}$, it holds that

$$I(A_1A_2...A_n:C|B)_\rho = I(A_1:C|B)_\rho + I(A_2:C|BA_1)_\rho + \cdots + I(A_n:C|BA_1...A_{n-1})_\rho \,. \tag{2.36}$$

We will investigate properties of the states with vanishing conditional mutual information in Sec. 2.2.3.

There are several ways to generalize the mutual information for states on more than two subsystems. One generalization is given by the total correlation [91]. For a state $\rho_{A_1...A_n}$, it is defined as

$$T(A_1:A_2:...:A_n)_\rho := \sum_i S(A_i)_\rho - S(A_1A_2...A_n)_\rho \,. \tag{2.37}$$

In terms of the relative entropy, the total correlation can be written as

$$T(A_1:A_2:...:A_n)_\rho = \min_{\sigma_{A_1},\cdots,\sigma_{A_n}} S(\rho_{A_1A_2...A_n} \| \sigma_{A_1} \otimes \cdots \otimes \sigma_{A_n}) \,. \tag{2.38}$$

The total correlation is the distance of the state from the set of all product states. The total correlation vanishes if and only if the state is a product state.

Entropic measures of entanglement

Entropy is also useful for quantifying entanglement. For a pure bipartite state $\rho_{AB} = |\psi_{AB}\rangle\langle\psi_{AB}|$, the entanglement entropy:

$$S(A)_\rho = -\mathrm{Tr}\rho_A \log_2 \rho_A \tag{2.39}$$

is the unique measure satisfying axioms of entanglement measures introduced by Ref. [92].

For mixed bipartite states, there are various inequivalent measures of entanglement such as the distillable entanglement [93], the entanglement cost [93] and

the entanglement of formation [94]. There are several requirements for "good" entanglement measures [95, 92], however checking these requirements is difficult in general. One entanglement measure which satisfies these requirements is the squashed entanglement [96]:

$$E_{sq}(\rho_{AB}) := \frac{1}{2} \inf_{\tau_{ABE}:\tau_{AB}=\rho_{AB}} I(A:B|E)_\tau \,, \tag{2.40}$$

where the infinitum is taken over all E and τ_{ABE} satisfying $\mathrm{Tr}_E \tau_{ABE} = \rho_{AB}$. We will analyze the squashed entanglement of Gibbs states in Chapter 5.

When the number of subsystems increase, classifying entanglement is getting much harder. One possible entropic measure of multipartite entanglement for a n-partite state $\rho_{1\ldots n}$ is the relative entropy of entanglement defined as

$$E_R(\rho_{1\ldots n}) := \min_{\sigma_{1\ldots n}:SEP} S(\rho_{1\ldots n} \| \sigma_{1\ldots n}) \,, \tag{2.41}$$

where the minimum is taken over all separable states $\sigma_{1\ldots n}$ in the state space.

Calculating entanglement measures is a computationally hard problem, since it often involves to solve optimization problems. For an example, evaluating the relative entropy of entanglement is known to be in a class of computational problems called NP-complete and evaluating the squashed entanglement is in NP-hard [97], which are suspected that even quantum computers cannot solve efficiently.

2.2.3 Quantum Markov Chains

A sequence of random variables X, Y, Z is called a (short) *Markov chain* if the corresponding probability distribution $p_{XYZ}(x, y, z)$ satisfies

$$p_X(x|y, z) = p_X(x|y) \,. \tag{2.42}$$

We call such a distribution a Markov distribution conditioned on Y. In terms of the conditional mutual information, this is equivalent to

$$I(X:Z|Y)_p = 0 \,. \tag{2.43}$$

More generally, the conditional mutual information for any distribution $p_{XYZ}(x, y, z)$ can be written as

$$I(X:Z|Y)_p = \min_{q:\mathrm{Markov}} S(p_{XYZ} \| q_{XYZ}) \,, \tag{2.44}$$

where the minimum is over all Markov distributions conditioned on Y. Therefore, the conditional mutual information quantifies how far is the distribution from Markov distributions.

In a similar mannar, we define a quantum Markov chain as a tripartite state ρ_{ABC} satisfying

$$I(A:C|B)_\rho = 0\,. \tag{2.45}$$

By definition, quantum Markov chains saturate the strong subadditivity (2.35). Saturating the strong subadditivity $\rho_{ABC} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$ is equivalent to the following conditions [45]:

- There exists a decomposition $\mathcal{H}_B = \bigoplus_i \mathcal{H}_{B_i^L} \otimes \mathcal{H}_{B_i^R}$ such that

$$\rho_{ABC} = \bigoplus_i p_i \rho_{AB_i^L} \otimes \rho_{B_i^R C}\,, \tag{2.46}$$

  where p_i is a probability distribution.

- There exists a CPTP-map $\Lambda_{B\to BC} : \mathcal{S}(\mathcal{H}_B) \to \mathcal{S}(\mathcal{H}_B \otimes \mathcal{H}_C)$ such that

$$\rho_{ABC} = (\mathrm{id_A} \otimes \Lambda_{\mathrm{B}\to\mathrm{BC}})(\rho_{\mathrm{AB}})\,. \tag{2.47}$$

  The map $\Lambda_{B\to BC}$ is called a recovery map.

One would expect that the quantum conditional mutual information also quantifies how far is the state from quantum Markov chains. However, in contrast to the classical case, the quantum conditional mutual information does not satisfy [77]

$$I(A:C|B)_\rho = \min_{\sigma:\mathrm{Markov}} S(\rho_{ABC}\|\sigma_{ABC})\,, \tag{2.48}$$

which is an analog of Eq. (2.44). A structural characterization of quantum states with small (quantum) conditional mutual information is a long-standing problem in quantum information (cf. Ref. [78] and references therein). An important consequence of a small quantum conditional mutual information was recently discovered in Ref. [78]. It is shown that the quantum conditional mutual information satisfies an inequality

$$I(A:C|B)_\rho \geq \min_{\Lambda_{B\to BC}} \frac{1}{4\ln(2)} \|\rho_{ABC} - (\mathbb{1}_A \otimes \Lambda_{B\to BC})(\rho_{AB})\|_1^2\,, \tag{2.49}$$

where the minimum is taken over all CPTP-maps from $\mathcal{S}(\mathcal{H}_B)$ to $\mathcal{S}(\mathcal{H}_B \otimes \mathcal{H}_C)$. This means that a state with small conditional mutual information is locally recoverable in the sense that there exists a CPTP-map, acting only on the conditioning system, which approximately recovers the total state from its reduced state.

Markov chains are generalized to more longer chains. Consider a state $\rho_{12\ldots n}$ on a system $\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$ defined on a one-dimensional chain labeled by $\{1, ..., n\}$. Then, we say $\rho_{12\ldots n}$ is a Markov chain if

$$I(A:C|B)_\rho = 0 \tag{2.50}$$

holds for any tripartition ABC of the system such that $d(A,C) \geq 1$. Note that this definition of the Markov chain includes the case of closed chains as well. In a similar manner, we say a state $\rho_{12\ldots n}$ is an ε-approximate Markov chain if

$$I(A:C|B)_\rho \leq \varepsilon \tag{2.51}$$

holds for any tripartition ABC of the system such that $d(A,C) \geq 1$.

2.3 Gibbs States and Maximum Entropy Principle

In this section, we review information-geometric properties of Gibbs states. Gibbs states are closely connected to the maximum entropy principle introduced by Jaynes [68], a principle to choose an inference from given partial knowledge of a state. Let us consider a set of observables $\mathcal{A} = \{A_i\}_{i=1}^m$ on $\mathcal{H}$. Suppose that we have partial information of a state given by a set of linear constraints:

$$\mathrm{Tr}(A_i\rho) = m_i\,, \quad i = 1, ..., m\,. \tag{2.52}$$

The possible candidates of the state is a set of states

$$\mathcal{C}_{\mathbf{m}}(\mathcal{A}) := \{\sigma \in \mathcal{S}(\mathcal{H}) \,|\, \mathrm{Tr}(A_i\sigma) = m_i,\ i = 1, ..., m\}\,. \tag{2.53}$$

We only deal in the case of the nonempty set. According to the maximum entropy principle, the most "unbiased" inference is given by the maximum entropy state

$$\tilde{\rho}_{\mathbf{m}} = \arg\max_{\sigma \in \mathcal{C}_{\mathbf{m}}(\mathcal{A})} S(\sigma)\,, \tag{2.54}$$

since $\tilde{\rho}_{\mathbf{m}}$ has the maximum uncertainty (entropy) under the constraints (2.52).

This type of problem is well-known in the literature of statistical mechanics. When the constraint is given by

$$\mathrm{Tr}(H\rho) = E \tag{2.55}$$

for a "Hamiltonian" H, the corresponding maximum entropy state is given by the *Gibbs state*

$$\tilde{\rho}_E = \frac{e^{-\beta H}}{Z}\,, \tag{2.56}$$

where $Z = \mathrm{Tr}(e^{-\beta H})$ and the Lagrange multiplier $\beta > 0$ is chosen so that $\mathrm{Tr}(H\tilde{\rho}_E) = E$ [3]. Therefore, the maximum entropy state is a Gibbs state.

[3]The parameter β can be ∞ if E is the lowest eigenvalue of H.

The equivalence of the maximum entropy states and Gibbs states can be extended to more general situations. Associated to $\mathcal{A}$, there exists a set of Gibbs states

$$\mathcal{E}(\mathcal{A}) := \left\{ \sigma \in \mathcal{S}(\mathcal{H}) \,\middle|\, \sigma = \frac{1}{Z} e^{-\sum_{i=1}^{m} a_i A_i} \right\}, \tag{2.57}$$

where Z is the normalizer and each $a_i \in \mathbb{R}$. This set only includes full-rank states. To treat non-full-rank states, we consider a closure of $\mathcal{E}(\mathcal{A})$. For a technical reason, the suitable closure is given by

$$\overline{\mathcal{E}}^{rI}(\mathcal{A}) := \left\{ \omega \in \mathcal{S}(\mathcal{H}) \,\middle|\, \inf_{\sigma \in \mathcal{E}(\mathcal{A})} S(\omega \| \sigma) = 0 \right\}, \tag{2.58}$$

which is called the reverse-information closure. For states $\rho \in \mathcal{C}_{\mathbf{m}}(\mathcal{A})$ and $\omega \in \overline{\mathcal{E}}^{rI}(\mathcal{A})$, the Pythagorean theorem [72]

$$S(\rho \| \omega) = S(\rho \| \tilde{\rho}_{\mathbf{m}}) + S(\tilde{\rho}_{\mathbf{m}} \| \omega) \tag{2.59}$$

holds (Fig. 2.1). Moreover, it was shown that the maximum entropy state $\tilde{\rho}_{\mathbf{m}}$ is the unique element in $\mathcal{C}_{\mathbf{m}}(\mathcal{A}) \cap \overline{\mathcal{E}}^{rI}(\mathcal{A})$ [72]. Therefore, we have

$$\begin{aligned} \inf_{\omega \in \mathcal{E}(\mathcal{A})} S(\rho \| \omega) &= \min_{\omega \in \overline{\mathcal{E}}^{rI}(\mathcal{A})} S(\rho \| \omega) && (2.60) \\ &= S(\rho \| \tilde{\rho}_{\mathbf{m}}) && (2.61) \\ &= S(\tilde{\rho}_{\mathbf{m}}) - S(\rho) \,. && (2.62) \end{aligned}$$

The last line follows from the fact that the completely mixed state $\tau_{\mathcal{H}}$ is an element of $\mathcal{E}(\mathcal{A})$ and thus Eq. (2.59) implies

$$S(\tau_{\mathcal{H}}) - S(\rho) = S(\rho \| \tilde{\rho}_{\mathbf{m}}) + S(\tau_{\mathcal{H}}) - S(\tilde{\rho}_{\mathbf{m}}) \,. \tag{2.63}$$

2.3.1 Irreducible Correlation

In general, multipartite states are correlated in very complicated ways and classifying multipartite correlations is a hard problem. The amount of correlations contained in a multipartite state is quantified by e.g., the total correlation introduced in Sec. 2.2.2, but it has only one value. In Ref. [66], Amari provided finer analysis of multipartite correlations by employing a hierarchical structure of Gibbs distributions. Recently, this framework has been extended to quantum systems as well [70, 98] which is reviewed in the following. In Chap. 4, we shall use this extended framework to analyze multipartite correlations in many-body quantum systems.

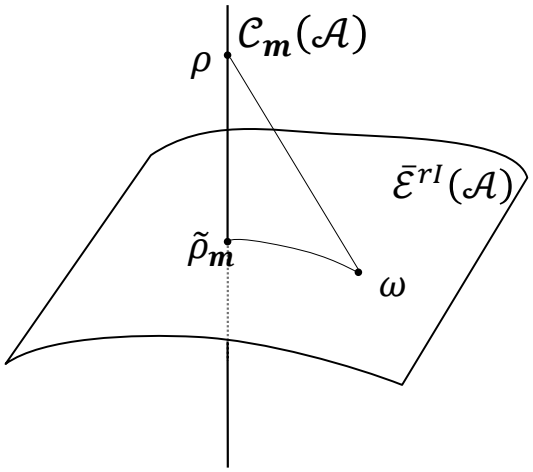

Figure 2.1: A schematic picture of the Pythagorean theorem (2.59). Manifolds $\mathcal{C}_{\mathbf{m}}(\mathcal{A})$ and $\overline{\mathcal{E}}^{rI}(\mathcal{A})$ are "orthogonal" each other in terms of the relative entropy. The projection of a point in $\mathcal{C}_{\mathbf{m}}(\mathcal{A})$ to $\overline{\mathcal{E}}^{rI}(\mathcal{A})$ is given by $\tilde{\rho}_{\mathbf{m}}$.

Since any quantum state ρ is positive semidefinite, we can write the state as

$$\rho = e^{-H_\rho} \tag{2.64}$$

in terms of a (possibly unbounded) Hermitian operator corresponding to ρ denoted by H_ρ. A key idea of Amari's framework is characterizing multipartite correlations in ρ via the structure of H_ρ. For an n-partite system $\mathcal{H} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$, we say a bounded operator H on $\mathcal{H}$ is *k-local* if it is an element of

$$\mathscr{H}_k := \left\{ H = \sum_{S_k \subset [n]} h_{S_k} \otimes \mathbb{1}_{S_k^c} , \right\} \tag{2.65}$$

where the sum is taken over all subsystems S_k such that $|S_k| \leq k$ and $S_k^c = [1,n] \backslash S_k$. Then, we define the set of its corresponding Gibbs state as

$$\mathcal{E}_k := \left\{ \sigma \in \mathcal{S}(\mathcal{H}) \,\middle|\, \sigma = \frac{1}{Z} e^{-H}, H \in \mathscr{H}_k \right\} . \tag{2.66}$$

To include non-full-rank states, we consider the closure $\overline{\mathcal{E}}_k^{rI}$. Then, we obtain a hierarchical decomposition of $\mathcal{S}(\mathcal{H})$ as

$$\overline{\mathcal{E}}_1^{rI} \subset \overline{\mathcal{E}}_2^{rI} \subset \cdots \subset \overline{\mathcal{E}}_n^{rI} = \mathcal{S}(\mathcal{H}) . \tag{2.67}$$

Note that $\overline{\mathcal{E}}_1^{rI}$ is the set of all product states $\rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_n$. When a state $\rho \in \mathcal{S}(\mathcal{H})$ is in $\overline{\mathcal{E}}_k^{rI}$, we can interpret that the state only contains multipartite

correlations "generated" by up to k-local interactions. Therefore, the distance from $\overline{\mathcal{E}}_k^{rI}$ represents how much correlations are contained in the state which cannot be generated by k-local interactions.

We use the relative entropy to quantify these multipartite correlations. Define the distance-like function $D^{(k)}$ as

$$D^{(k)}(\rho) := \inf_{\sigma\in\mathcal{E}_k} S(\rho\|\sigma) = \min_{\sigma\in\overline{\mathcal{E}}_k^{rI}} S(\rho\|\sigma)\,. \tag{2.68}$$

Since $\overline{\mathcal{E}}_1^{rI}$ is the set of all product states, $D^{(1)}(\rho)$ is the total correlation, i.e.,

$$D^{(1)}(\rho) = T(1:2:\ldots:n)_\rho\,. \tag{2.69}$$

$\mathcal{E}_k$ corresponds to $\mathcal{E}(\mathcal{A}^k)$ for $\mathcal{A}^k = \{X^i_{S_k}\otimes\mathbb{1}_{S_k^c}\}_{i,S_k}$, where $\{X^i_{S_k}\}_i$ is the generalized Pauli operators on $\bigotimes_{j\in S_k}\mathcal{H}_j$. Therefore, when we impose $\mathrm{Tr}(\sigma A_i) = \mathrm{Tr}(\rho A_i)$ for all $A_i\in\mathcal{A}^k$, it is equivalent to say $\sigma_{S_k} = \rho_{S_k}$ for all S_k, i.e., σ and ρ has exactly same k-partite reduced states. Then the minimum in Eq. (2.68) is achieved by the maximum entropy state $\tilde{\rho}^{(k)}\in\overline{\mathcal{E}}_k^{rI}$, which is also an element of

$$\mathcal{C}_\rho^{(k)} := \{\sigma\in\mathcal{S}(\mathcal{H})|\sigma_{S_k} = \rho_{S_k}\,,\ \forall S_k\subset[n]\}\ . \tag{2.70}$$

By using the properties of the maximum entropy state, we can obtain another representation of $D^{(k)}(\rho)$, that is,

$$D^{(k)}(\rho) = S(\tilde{\rho}^{(k)}) - S(\rho)\,. \tag{2.71}$$

To extract the kth-order effect, we define the *kth-order irreducible correlation* as

$$\begin{aligned} C^{(k)}(\rho) &:= D^{(k-1)}(\rho) - D^{(k)}(\rho) && (2.72)\\ &= S(\tilde{\rho}^{(k)})\|\tilde{\rho}^{(k-1)}) && (2.73)\\ &= S(\tilde{\rho}^{(k-1)}) - S(\tilde{\rho}^{(k)})\,. && (2.74)\end{aligned}$$

By definition, $C^{(k)}(\rho)$ is always non-negative. For so-called stabilizer states, the irreducible correlation can be calculated via a formula obtained by Ref. [70]. $C^{(k)}(\rho)$ quantifies information which is contained in the k-partite reduced states, but not in the $(k-1)$-partite reduced states. A numerical algorithms to calculate the irreducible correlation is known [99]. Eq. (2.69) implies that the irreducible correlation provides a decomposition of the total correlation:

$$T(1:2:\ldots:n)_\rho = \sum_{k=2}^{n} C^{(k)}(\rho)\,. \tag{2.75}$$

A geometric illustration of $D^{(k)}$ and $C^{(k)}$ is depicted in Fig. 2.2.

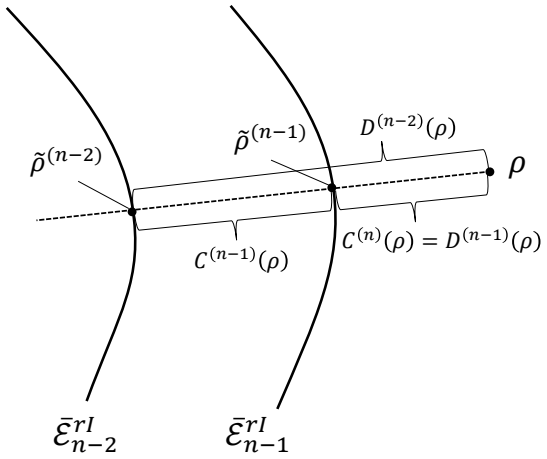

Figure 2.2: A geometrical illustration of the functions $D^(k)$ and $C^{(k)}$. $D^{(k)}$ is the distance from the set of states only containing kth-order correlation $\overline{\mathcal{E}}^{rI}_k$ and $C^{(k)}$ is the difference between $D^{(k)}$ and $D^{(k-1)}$.

Example 5. In the case of three-qubit pure states, calculating the irreducible correlation is an easy task, since only states in the form

$$\alpha|000\rangle + \beta|111\rangle \tag{2.76}$$

have non-zero 3rd-order irreducible correlation [69]. Let us consider the irreducible correlation of the GHZ state $|GHZ_3\rangle$ and the W state $|W_3\rangle$ in a three-qubit system defined by

$$|GHZ_3\rangle := \frac{1}{\sqrt{2}}\left[|000\rangle + |111\rangle\right], \qquad |W_3\rangle = \frac{1}{\sqrt{3}}\left[|001\rangle + |010\rangle + |100\rangle\right]. \tag{2.77}$$

For the GHZ state, the maximum entropy state $\tilde{\rho}^{(2)}_{GHZ}$ is given by

$$\tilde{\rho}^{(2)}_{GHZ} = \frac{1}{2}(|000\rangle\langle 000| + |111\rangle\langle 111|) \tag{2.78}$$

and the W state is the maximum entropy state for the bipartite marginals itself. Then, the irreducible correlations are given by

$$C^{(3)}(GHZ_3) = 1\,,\ C^{(2)}(GHZ_3) = 2\,,\ C^{(3)}(W_3) = 0\,,\ C^{(2)}(W_3) = 3h\left(\frac{2}{3}\right), \tag{2.79}$$

where $h(p) := -p\log_2 p - (1-p)\log_2(1-p)$.

The irreducible correlation is an non-negative, local unitary invariant and additive function $C^{(k)}(\rho\otimes\sigma) = C^{(k)}(\rho) + C^{(k)}(\sigma)$ for any $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ [4]. It is clear

[4]Here, we regard the $2n$-partite state $\rho\otimes\sigma$ as a n-partite state on Hilbert space $\mathcal{H}' = \mathcal{H}'_1 \otimes \cdots \otimes \mathcal{H}'_n$, where $\mathcal{H}'_i \equiv \mathcal{H}_i \otimes \mathcal{H}_i$.

that $\tilde{\rho}^{(k)} \otimes \tilde{\sigma}^{(k)}$ is included in $\mathcal{C}^{(k)}_{\rho\otimes\sigma}$. To avoid confusion, let us denote $\mathcal{H} \otimes \mathcal{H}$ by $\mathcal{H}_1 \otimes \mathcal{H}_2$ with $\rho \in \mathcal{S}(\mathcal{H}_1)$ and $\sigma \in \mathcal{S}(\mathcal{H}_2)$. For any state $\omega_{12} \in \mathcal{C}^{(k)}_{\rho\otimes\sigma}$, it holds that

$$S(\omega_{12}) \leq S(\omega_1) + S(\omega_2) \tag{2.80}$$
$$\leq S(\tilde{\rho}^{(k)}) + S(\tilde{\sigma}^{(k)}) \tag{2.81}$$
$$= S(\tilde{\rho}^{(k)} \otimes \tilde{\sigma}^{(k)}), \tag{2.82}$$

where we used the subadditivity of the von Neumann entropy and the fact that $\omega_{12} \in \mathcal{C}^{(k)}_{\rho\otimes\sigma}$ implies $\omega_1 \in \mathcal{C}^{(k)}_{\rho}$ and $\omega_2 \in \mathcal{C}^{(k)}_{\omega}$. Therefore, $\tilde{\rho}^{(k)} \otimes \tilde{\sigma}^{(k)}$ is the kth-maximum entropy state corresponding to $\rho \otimes \sigma$. By definition, this implies that the irreducible correlation is additive.

However, the irreducible correlation is known to lack properties called continuity and monotonicity under local operations. The irreducible correlation is not continuous, that is, there exists a one-parameter family of states ρ_ϵ such that $\epsilon \to +0$ but not $C^{(3)}(\rho_\epsilon) \to C^{(3)}(\rho_0)$ even if the dimension of the system is finite [71, 72]. Interestingly, this discontinuity only happens for quantum states [71]. The lack of monotonicity under local operations is found in both classical [67] and quantum [98] systems. It means that there exists a CPTP-map $\mathcal{E}$, which acts on only one subsystem, such that

$$C^{(k)}(\rho) < C^{(k)}\left(\mathcal{E}(\rho)\right) \tag{2.83}$$

for some k. It sounds counterintuitive if one can increase correlations by applying local operations. However, since the total correlation, i.e., the sum of the irreducible correlations is monotone under local operations, local operations can increase the kth-order irreducible correlation by *consuming* the irreducible correlation of other orders. To recover the monotonicity, a possible modification is proposed in Ref. [100]. In that paper, the authors use the local operation orbit of $\mathcal{E}_k$ instead of $\mathcal{E}_k$.

2.3.2 The multivariate mutual information

The irreducible correlation is not the unique way to decompose the total correlation into different "levels". For instance, the multivariate mutual information [61] introduced in classical information theory also decompose the total correlation [5]. A quantum analog of the multivariate mutual information $I^{(k)}(\rho)$ for an n-partite state $\rho \in \mathcal{S}(\mathcal{H})$, which is defined as

$$I^{(k)}(\rho) := -\sum_{X\subset[n]} (-1)^{n-|X|} S(X)_\rho \,, \tag{2.84}$$

[5] $I^{(k)}(\rho)$ and its minus are called in several different ways e.g., the interaction information [61] and the co-information [101].

where the sum is taken over all subsets X of $[n]$, satisfies

$$T(1:2:...:n)_\rho = \sum_{k=2}^{n} I^{(k)}(\rho)\,. \tag{2.85}$$

Although $I^{(k)}(\rho)$ is widely used in classical and quantum information theory, it does not have any clear operational or geometrical interpretation. One reason which makes interpretation difficult is that the multivariate mutual information can be negative if the number of subsystems is odd. Therefore, $T(1:2:3) = I^{(2)}(\rho) + I^{(3)}(\rho)$ can be strictly smaller than $I^{(2)}(\rho)$ when $I^{(3)}(\rho) < 0$. Moreover, another problem arises in quantum case, e.g., $I^{(3)}(\rho) = 0$ for all tripartite pure states. See, e.g., Ref. [102] for an approach to understand the multivariate mutual information in terms of "redundancy" and "synergy".

Chapter 3

Topologically Ordered Phases and Entanglement

A central problem in condensed matter physics is classifying various phases of matters. In 1937, Landau developed the theory of phase transitions to describe transitions in ferromagnets [103]. This theory has been generalized and commonly used to understand various phases, such as superconducting phase and superfluids. The key ideas of Landau's theory are *symmetry-breaking* and corresponding *local order parameters*. When a phase-transition occurs, the symmetry of the system changes and the change can be detected by the expectation value of an appropriate operator acting on a local region.

Starting from 1980's, it has been gradually realized that there are phases which cannot be described by Landau's framework. One kind of such phases is *topologically ordered phases* which were observed in for example, the fractional quantum Hall effect (FQHE) [104, 105, 106]. Topologically ordered phases have distinguished properties which do not appear in symmetry-breaking phases described by Landau's theory. Examples of such properties are the ground state degeneracy depending on the topology of the manifold supporting the system [107], anyonic excitations [105] and protected gapless edge modes on open boundaries [108]. Topologically ordered phases are stable against any local perturbation and thus the ground states are promising candidates for the fault-tolerant quantum memory [50]. Moreover, anyonic excitations enable fault-tolerant quantum computation by "braiding" anyons [51]. Due to these reasons, study of topologically ordered phases has been an actively investigated interdisciplinary research field bridging condensed matter and quantum information.

In this chapter, we introduce basic preliminaries for understanding topologically ordered phases. We begin with defining topologically ordered phases as a class of gapped phases in Sec. 3.1 based on the definitions introduced in Refs. [109, 110]. See these references for the more precise definition of topologically ordered phases.

Another way to define that states are in topologically ordered phases is proposed in Ref. [111] by using local indistinguishablity of degenerated ground states. Note that in this thesis we only consider many-body quantum systems defined on 2D spin lattices as topologically ordered systems while topologically ordered phases exist in higher dimensional systems. We then review the toric code model [50] as a simple example of exactly solvable topological models. We introduce the definition of the topological entanglement entropy and the entanglement spectrum in Sec. 3.2.

3.1 Gapped Phases and Topological Order

Consider a set of short-range Hamiltonians H_{N_k} of systems defined on graphs with N_k spins, where $N_k \to \infty$ as $k \to \infty$ [1]. The system is said to be gapped if there exists a constant $\delta > 0$ such that each H_{N_k} does not have eigenvalues smaller than δ above the ground state energy. Here, we allow energy splitting between (candidates of) the ground states which vanishes in the limit of $k \to \infty$. When we say a Hamiltonian is gapped, we assume that the Hamiltonian corresponds to H_{N_k} of a gapped system with certain N_k.

Let us introduce the notion of *gapped quantum phases*. Consider an one-parameter family of gapped short-range Hamiltonians $H(g) = \sum_i H_i(g)$, where each i labels a lattice site and $H_i(g)$ acts only on spins j with $d(i,j) \leq r$ for all $g \in [0,1]$. We assume that the Hamiltonians have finite interaction strength and smoothly depend on g, i.e., for any i and $g \in [0,1]$, it holds that $\|H_i(g)\| \leq J$ and $\|\partial_g H_i(g)\| \leq K$ for some constants J and K. Let $|\psi(0)\rangle$ be the ground state of $H(0)$ and $|\psi(1)\rangle$ be the ground state of $H(1)$. We say two states $|\psi(0)\rangle$ and $|\psi(1)\rangle$ are in the same gapped quantum phase if $H(g)$ has a finite gap between $g \in [0,1]$, more precisely, if $|\psi(0)\rangle$ and $|\psi(1)\rangle$ are connected via an adiabatic change without closing the gap for all system sizes [2].

By using the technique of quasi-adiabatic evolution, one can show that this definition is equivalent to consider a unitary evolution $|\psi(1)\rangle = U|\psi(0)\rangle$ described by

$$U := \mathcal{T} \exp\left[-i \int_0^1 dg \tilde{H}(g)\right], \tag{3.1}$$

where $\tilde{H}(g)$ is a short-range Hamiltonian defined as

$$\tilde{H}(g) := i \int_{-\infty}^{\infty} dt F(t) e^{iH(g)t} \left(\partial_g H(g)\right) e^{-iH(g)t} \tag{3.2}$$

[1] Here, system sizes $\{N_k\}$ do not have to cover all natural numbers.

[2] Furthermore, we assume two additional constraints on gapped Hamiltonians which allow to connect Hamiltonians of different systems smoothly, but we skip to explain details here since they are not important in this thesis. See Ref. [110] for details.

for a function $F(t)$ satisfying several conditions [111]. The unitary operator U is called *local unitary evolution* and can be simulated by a constant-depth local unitary circuit with a constant error [109]. A constant-depth local unitary circuit is defined as a product of unitaries

$$U = U^{(1)}U^{(2)}\cdots U^{(M)}\,, \tag{3.3}$$

where M is a constant independent of the system size and each $U^{(i)} = \bigotimes_l U_l^{(i)}$ is a tensor product of unitaries acting on non-overlapping local regions (Fig. 3.1). Conversely, any constant-depth local unitary circuit can be simulated by such local unitary evolutions [109]. In this sense, we can say that two (sets of) states are in the same phase if one can transform one state to the other by a constant-depth local unitary circuit.

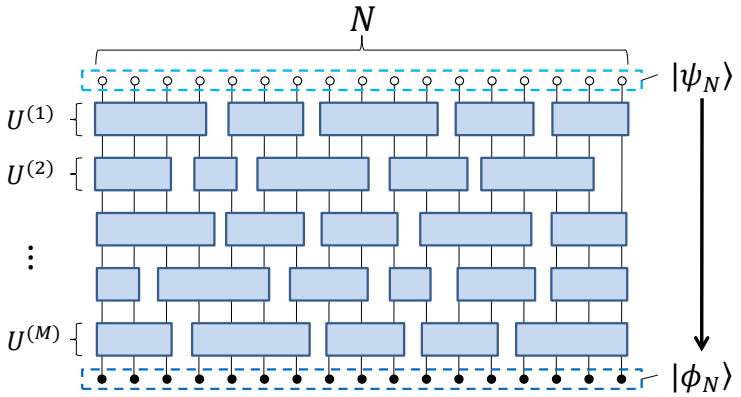

Figure 3.1: A illustration of a transformation by a constant-depth local unitary circuit. A set of states $\{|\psi_N\rangle\}_N$ can be converted to $\{|\phi_N\rangle\}_N$ by a constant-depth local unitary circuit if for each N there exists a unitary U in the form of Eq. (3.3) with $M = \mathcal{O}(1)$ which transforms $|\psi_N\rangle$ to $|\phi_N\rangle$.

According to this classification, ground states of gapped short-range Hamiltonians are classified in different types of phases. We say a system obeys a *topologically-ordered phase* if the ground states are not in the equivalence class of a product state and the ground degeneracy is stable under local perturbations. Otherwise, we say the system is in the (topologically) trivial phase.

When the system obeys a symmetry, we can divide the equivalence classes according to the symmetry. In this case, we introduce a finer equivalence relation by imposing that the generators $\tilde{H}(g)$ of local unitary evolutions are invariant under the actions of the symmetry group. Then, the trivial phase can be classified to the conventional symmetry-breaking phases or the symmetry-protected topologically ordered (SPT) phases. In a SPT phase, ground states do not break any symmetry

of the Hamiltonian, but they cannot be connected to a product state by symmetric local unitary evolutions without closing the gap. The famous examples of SPT phases are the Haldane phase of the AKLT-model [26] and the topological insulators [112]. While we will not investigate SPT phases in this thesis, the ground states in the SPT phases have been shown to be a useful resource to perform universal quantum computation [113, 114].

3.1.1 Example: Toric Code Model

A simple example of topologically ordered phases is given by Kitaev's toric code model [50]. The toric code model is defined on a 2D spin lattice where each edge of the lattice corresponds to a $\frac{1}{2}$-spin system (Fig. 3.2). The Hamiltonian of the model is given by

$$H = -\sum_v A_v - \sum_p B_p \,, \tag{3.4}$$

with A_v and B_p defined as

$$A_v := \bigotimes_i X_{v_i} \,, \quad B_p := \bigotimes_i Z_{p_i} \,, \tag{3.5}$$

where X_{v_i} is the Pauli-X operator acting on edge v_i which includes vertex v and Z_{p_i} is the Pauli-Z operator acting on edge p_i around face p.

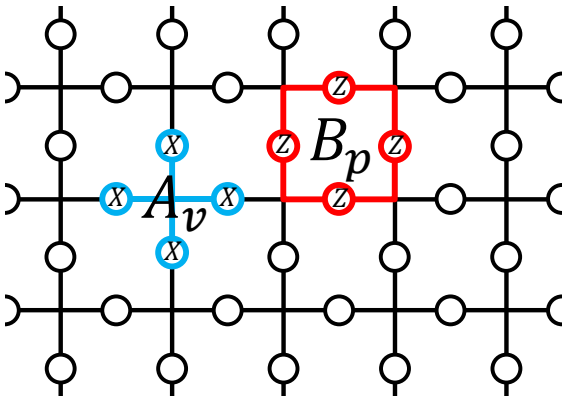

Figure 3.2: The toric code on a square lattice. A_v is the tensor product of Pauli-X operator acting on spins (depicted as circles) around vertex v (the center of the blue cross) and B_p is the tensor product of Pauli-Z operator acting around face p (the red square).

Neighboring a face and a vertex share two edges and $[A_v, B_p] = 0$ holds for any pair of v and p. Therefore, the ground states of the Hamiltonian are eigenstates

of all interaction terms. The ground state degeneracy of the toric code model depends on the topology of the system. If the system is defined on a 2D closed manifold with genus g, the degeneracy is given by 4^g. Since each Pauli operator has ± 1 eigenvalues, one ground state $|\psi\rangle$ is obtained by

$$|\psi\rangle = \prod_v \frac{\mathbb{1}_v + A_v}{\sqrt{2}} |000....\rangle \,, \tag{3.6}$$

where $|0\rangle$ is the eigenstate of Z with eigenvalue $+1$.

Let us consider a string C consisting of connecting edges. Define a Z-string operator $W_Z(C)$ as

$$W_Z(C) := \bigotimes_{e \in C} Z_e \,, \tag{3.7}$$

where e is an edge on the string C (Fig. 3.3). In a similar way, we can define a X-string operator $W_X(\tilde{C})$ as the tensor product of X operators along a string on the dual lattice (dual string). One can deform $W_Z(C)$ ($W_X(\tilde{C})$) by applying A_v (B_p) operator neighboring the string, since a product of same Pauli operators cancels out.

When C is a contractible loop, $W_Z(C)$ can be written as a product of B_p, and therefore $\langle\psi|W_Z(C)|\psi\rangle = 1$ for arbitrary loop C, but$\langle\psi|W_Z(C)|\psi\rangle = 0$ for any open string C (A similar relation holds for a dual loop $\tilde{C}$ and $W_X(\tilde{C})$). In this sense, we say that ground states of the toric code model exhibit multipartite loop-like correlations. By expanding the product in Eq. (3.6), we obtain another representation of the ground state in the form

$$|\psi\rangle = \frac{1}{\sqrt{|\mathcal{E}|}} \sum_{c \in \mathcal{E}} |c\rangle \,, \tag{3.8}$$

where $\mathcal{E}$ is the all possible loop configurations of X-strings, and $|c\rangle$ is a product state where only spins along the loop are $|1\rangle$ and otherwise $|0\rangle$. Thus, we can regard the ground state as a "loop gas" (A similar expression exists for a dual loop $\tilde{C}$ and $W_X(\tilde{C})$).

By applying the $W_Z(C)$ operator on an open string C, one can create a pair of quasiparticle excitations on the two vertices at the endpoints of C (Fig. 3.3). Quasiparticles created by the Z-string operator are called e-anyons. When we apply $W_X(\tilde{C})$, it creates a pair of anyons called m-anyons located at the endpoint of the dual string $\tilde{C}$. A neighboring pair of e and m anyons is also considered as another anyon, called ϵ-anyon. The existence of an e-anyon, which is created by $W_Z(C)$, can be checked by performing a X-string operator on a dual loop $\tilde{C}$ enclosing the anyon, since $W_Z(C)$ and $W_X(\tilde{C})$ anticommutes each other. In the same way, the existence of a m-anyon can be detected by a Z-string operator enclosing the m-anyon. ϵ-anyons can be detected by using both types of operators.

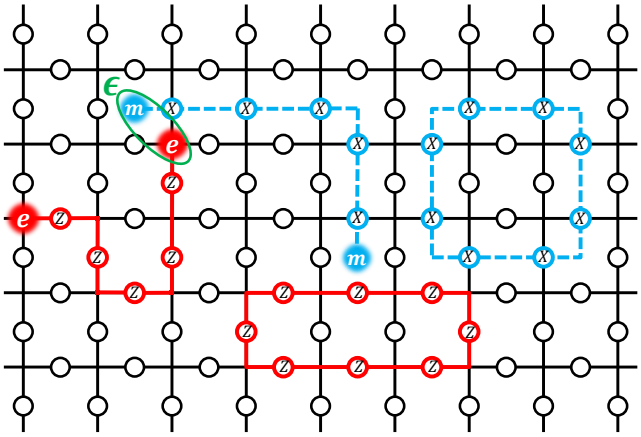

Figure 3.3: String operators in the toric code. The ground states in the toric code exhibits loop-like correlations which are characterized by loop operators (the blue and red closed loops). For a open-string, there are anyonic excitations at the end points associating the type of the string. e-anyons (red circles) corresponding to the Z-string operator appear at vertices, and m-anyons (blue circles) corresponding to the X-string operator appear at faces. ϵ-anyons are a pair of neighboring e and m-anyon (green circle).

When a system is defined on a 2D closed manifold with non-trivial topology, e.g., a torus, C can be a non-contractible loop on the manifold. Then, $W_Z(C)$ cannot be a product of B_p but still commutes with the Hamiltonian. Therefore, the $+1$ eigenstates and -1 eigenstates of $W_Z(C)$ are orthogonal ground states. The same argument holds for $W_X(\tilde{C})$ and we obtain 4 orthogonal ground states (Fig. 3.4). These states are unchanged when we deform C or $\tilde{C}$ in the same homology class, and considering another homology class corresponds to a change of the basis of the ground state subspace.

Toric Code as A Quantum Memory

A large obstacle to implement quantum information processing is decoherence caused by interactions with the environment. When we perform large-scale quantum information-processing, we often need to create a complex entangled state at some point. However, such complicated entanglement is extremely fragile and collapses in very short-time. One idea to overcome this difficulty is encoding quantum information by using quantum error-correcting codes [115]. When an error happens in the code, we can detect the existence of the error by performing a so-called syndrome measurement without breaking the stored information. The error is then corrected by performing appropriate error-corrections.

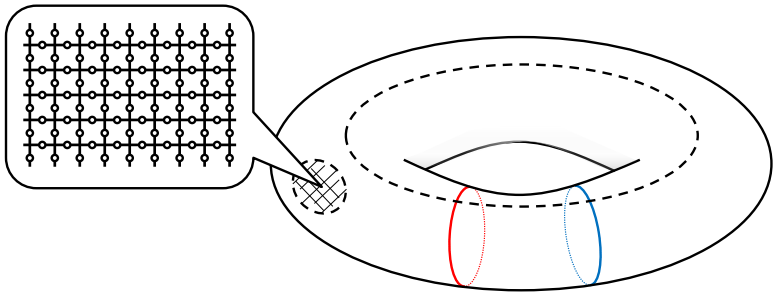

Figure 3.4: The ground states in the toric code defined on a torus. A basis of the ground subspace is given by the eigenstates of the non-contractible Z- and X-loop operators (the red and blue loop). These eigenstates are unchanged when we choose other loops which are in the same homotopy class. Another choice of the homology class (the dashed loop) gives a different choice of the basis.

A notable property of degenerated ground states $\{|\psi_i\rangle\}_i$ of the toric code on a torus is

$$\langle\psi_i|O|\psi_j\rangle = C_O\delta_{ij}\,, \tag{3.9}$$

where O is an operator acting on a local (namely, not wrapping the torus) region and C_O is a constant depending on O. This is the condition of states being a quantum error correcting code which is robust against local noises. In this sense, the toric code is a quantum error-correcting code. We can encode quantum information as a ground state in the ground subspace of the toric code. An local error creates a pair of excitations and the positions of the excitations can be detected by using A_v and B_p. After the error detection, we annihilate the excitations by performing appropriate string operators. The encoded information is changed only if the creating and annihilating process make a non-contractible loop on the manifold. Therefore, by taking the system size to be sufficiently large, the toric code is a good candidate for storing of quantum information.

A system is called quantum self-correcting memory if we can preserve quantum information in the system over a long time without actively performing syndrome measurement and error corrections. This is related to thermal stability of topologically ordered phases, since their ground subspaces are good candidates for quantum memories. Unfortunately, it has shown that a large class of exactly solvable models on 2D lattices are fragile against thermal noises [116, 117], while a variant of the toric code in 4D spin lattice has shown to be a quantum self-correcting memory [118]. Investigating realizable quantum self-correcting memory is currently a central problem in the field of quantum error-correction.

3.2 Topological Entanglement Entropy and Entanglement Spectrum

Since topologically ordered phases do not allow classifications in terms of symmetry and local order parameters, we need a new way to characterize these phases. One can clarify the existence of topologically order by checking their topological degeneracy, anyonic excitations or edge states. Besides these methods, the topological entanglement entropy (TEE) [55, 54] is an indicator of topologically ordered phases calculated by a reduced state of a ground state. The TEE is drawn from the entanglement entropy of the ground state, and therefore information about phases is understood to be contained in the reduced state of the ground state. The logarithm of the reduced state of some region is called the entanglement Hamiltonian and its spectrum, called the entanglement spectrum (ES), has been investigated as well for characterizing quantum phases [119]. In this section, we introduce the TEE and the ES in more detail.

3.2.1 The Area Law of Entanglement and The Topological Entanglement Entropy

As presented in Sec. 2.2.2, the universal measure of entanglement for pure bipartite states is the entanglement entropy. Analyzing the entanglement entropy in many-body state provides a new way to characterize different phases. For a sufficiently smooth region A, a ground state in a 2D gapped system typically obeys the area law

$$S(A)_\rho = \alpha|\partial A| - n_A\gamma + o(1)\,, \tag{3.10}$$

where α, γ are positive constants and n_A is the number of disconnected boundaries of A. $o(1)$ comprises the corrections due to short-range correlations which vanishes at the thermodynamic limit or fixed points of renormalization groups, such as ground states of the toric code model [50] or the Levin-Wen models [76]. The universal constant γ (or -γ) only depends on the phase of the states and called the topological entanglement entropy [55]. In general, distinguishing the constant γ from the value of the entanglement entropy contains ambiguity due to other constant terms. To avoid this problem, one can extract the TEE by taking a suitable linear combination of entropies of (at least) three subsystems. For regions ABC depicted in Fig. 3.5, the topological entropy [54] is defined as

$$S_{topo} := S(AB)_\rho + S(BC)_\rho + S(CA)_\rho - S(A)_\rho - S(B)_\rho - S(C)_\rho - S(ABC)_\rho\,. \tag{3.11}$$

Note that $S_{topo} = I(A:C|B)_\rho$ when $I(A:C)_\rho = 0$. When the state ρ obeys the area law (3.10) and the boundary terms cancels out, the topological entropy is

equivalent to the TEE up to a small correction, i.e.,

$$S_{topo} = n_{ABC}\gamma + o(1)\,. \tag{3.12}$$

In the following we call both quantities the topological entanglement entropy.

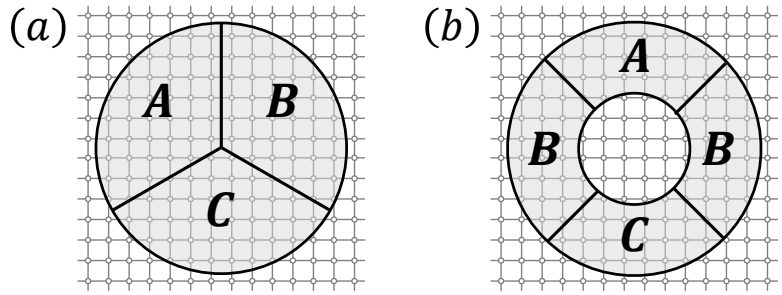

Figure 3.5: Examples of regions used for calculating the topological (entanglement) entropy given by Eq. (3.11). (a): a choice of region proposed in Ref. [55]. (b): a choice of the region proposed in Ref. [54]. In the case of (b), S_{topo} is approximately equivalent to conditional mutual information $I(A:C|B)$ since $I(A:C) \approx 0$ for typical gapped ground states.

The value of the topological entanglement entropy is related to the corresponding anyon model. For several classes of gapped models, such as systems decribed by topological quantum field theory (TQFT), it is proven that the topological entanglement entropy is given by

$$\gamma = \log_2 \sqrt{\sum_i d_i^2}\,, \tag{3.13}$$

where d_i is called the quantum dimension associated to anyon type i and the sum is taken over all anyon types appearing in the model. When there exists a pair of anyonic excitations labeled by a between a region A and its complement, the area law is known to have an extra term $\log_2 d_a$ due to entanglement shared by the anyons [55, 120]. An operational interpretation of this extra term in terms of entanglement distillation and dilution is discussed in Ref. [121].

Topologically ordered phases are often considered stable against weak local perturbations. Although the rigorous definition of topologically ordered phases is still arguable, the stability of the gap under these perturbations is rigorously proven for one of the definitions [111]. The existence of loop operators are also proven to be an universal property in the whole phase, i.e., the algebraic structure of loop operators are preserved under any quasi-adiabatic evolution [122]. The stability of the topological entanglement entropy is discussed in Ref. [47], and

rigorously proven up to the first order of perturbation under reasonable assumptions. However, whether it is stable against all possible quasi-adiabatic evolutions has been an important open problem. Actually, there exists a state in the trivial phase with non-trivial topological entanglement entropy for a particular choice of regions [65, 123].

As we discussed in the previous section, investigating thermal stability of topologically ordered phases is an important problem for not only condensed matter physics but also quantum information science. In that situation, we need to consider mixed states as well. The entanglement entropy is a valid measure of entanglement only for pure states, and therefore we need an alternative quantity instead of the TEE. In Ref. [64], the topological mutual information γ_I is introduced through the area law of the mutual information:

$$I(A : A^c) = \alpha|\partial A| - \gamma_I + o(1)\,. \tag{3.14}$$

γ_I can be extracted by taking an appropriate linear combination of the mutual information of subsystems as in Fig. 3.5, that is given by

$$\gamma_I = \sum_{X=AB,BC,CA} I(X : X^c) - \sum_{X=A,B,C} I(X : X^c) - I(ABC : (ABC)^c)\,. \tag{3.15}$$

At $T = 0$, the mutual information is equivalent to twice of the entanglement entropy and therefore this is a generalization of the TEE for arbitrary temperature. It has been observed for the toric code that the topological mutual information survives in finite temperature with an appropriate limit of the system size and the coupling constant.

While the TEE is originally introduced by considering the area law of the entanglement entropy, it also appears in a measure of multipartite entanglement. In Ref. [52], a function which is equivalent to the relative entropy of entanglement are calculated for several gapped models where $o(1)$ terms of the area law exactly vanish. The authors of Ref. [52] divide the spin lattice supporting the system into several blocks consisting of neighboring spins, and investigate the scaling property. As a result, the authors found that the relative entropy of entanglement obeys the area law with respect to the size of the blocks as

$$E_R = cn_bL - \gamma\,, \tag{3.16}$$

where c is a constant, n_b is the number of blocks and L is the size of each block. Therefore, the TEE also appears as the constant term in the multipartite settings.

3.2.2 Entanglement Hamiltonian and Entanglement Spectrum

The entanglement entropy is a single value function calculated from the spectrum (or equivalently, the Schmidt coefficients) $\{\lambda_A^i\}_i$ of the reduced state ρ_A on a region A. Therefore, in principle the spectrum contains more information about the phase than the entanglement entropy. Based on this idea, Li and Haldane proposed to use *the entanglement spectrum* [119]

$$\{\xi_i := -\ln \lambda_A^i\} \tag{3.17}$$

to characterize the phase [3]. By definition, the entanglement spectrum is interpreted as the spectrum of *the entanglement Hamiltonian* H_A defined via

$$\rho_A = e^{-H_{\rho_A}} . \tag{3.18}$$

Indeed, a remarkable observation has been obtained for FQHE states that the low-energy part of the entanglement spectrum has one-to-one correspondence with the low-lying physical edge-state spectrum described by CFT [119]. A substantial literature has been followed for chiral topologically ordered phases and SPT phases [124, 125, 126, 127]. Note that nonchiral topologically ordered phases can exhibit gapped boundaries [128] and the correspondence between the ES and the edge physics is unclear [129].

[3] In the original paper, the partition A, A^c is introduced in orbital degrees of freedom of electrons.

Chapter 4

Topological Entanglement Entropy and Multipartite Correlations

As represented by the toric code model, ground states in topologically ordered phases contains characteristic multipartite correlations in loop-like regions. The main topic of this chapter is establishing information-theoretic characterization of such multipartite correlations. The topological entanglement entropy (TEE) [55, 54] has been considered to be an indicator of the characteristic correlations in topologically ordered phases [54]. Indeed, the TEE is equivalent to the topological entropy defined in Eq. (3.11) that can be interpreted as a quantum analog of the multivariate mutual information (Eq. (2.84)) of order 3 for a certain tripartite region. The multivariate mutual information provides a decomposition of the total correlation [91] into different levels, and the TEE represents the amount of 3rd-order correlations in this sense. However, several properties of the multivariate mutual information, such as negativity, make difficult to interpret the value of this function as the amount of a particular correlation (see also Sec. 2.3.1). Functions in information theory often have a geometrical or operational meaning, but the multivariate mutual information still lacks both meaning.

To address this issue, we investigate an alternative way to characterize the "characteristic correlations" in topologically ordered phases. One candidate of functions which properly quantify the characteristic correlations is the irreducible correlation which has information-geometrical meaning [69, 70]. Indeed, the irreducible correlation has recently been investigated in topologically ordered systems [74, 75] (see also Ref. [130] for a similar function developed independently). It has observed that the highest-order irreducible correlation takes non-zero value in a topologically ordered phases and moreover the value coincides to the TEE [75]. Although the TEE and the irreducible correlations are conceptually different, one

can expect that they coincide under a particular setting, especially for gapped ground states, as conjectured in Ref. [74, 75]. If this conjecture is true, we can classify the characteristic correlations in topologically ordered phases in terms of information-geometry and also affirm the fact that the TEE has been successfully detect topologically ordered phases in variational models. However, under what conditions the conjecture holds has not been clarified so far.

In this chapter, we show this conjecture holds for 2D many-body states with exactly zero correlation length as ground states of exactly solvable gapped models. These states can be interpreted as fixed-point of (spatial) renormalization flow, and therefore reflect property of phases independent of local properties due to non-zero correlation length. This property is suitable for our analysis, since we are interested in the characteristic correlations which are considered to be a universal property of topologically ordered phases. We will also discuss an extension to states with finite correlation length in Sec. 4.4. Formally, a state with the "zero correlation length" in this thesis refers to a state obeying a specific area law given by

$$S(A)_\rho = \alpha|\partial A| - n_A\gamma\,, \tag{4.1}$$

where α is independent of the choice of the region. This assumption of zero correlation length implies that

- (I) If two regions A and B are separated, then the reduced state is a product state $\rho_{AB} = \rho_A \otimes \rho_B$, i.e., $I(A:B)_\rho = 0$.
- (II) For a simply connected region ABC such that B separates A from C, $I(A:C|B)_\rho = 0$.

In the following, we assume these two conditions rather than the area law of Eq. (4.1). Although our main concern is ground states, our results also hold for mixed states. Note that assumption (II) implies assumption (I) if the whole state is pure. This is because when $(AB)^c$ separates A from B, $I(A:B)_\rho = I(A:B|(AB)^c)_\rho = 0$ for separated regions A and B. Therefore, our results hold even if the global state on the whole lattice is a mixed state, and only exploits the properties of states, not the Hamiltonian.

In general, direct calculation of the irreducible correlation is a computationally hard problem. However, assumption (II) indicates that certain reduced states in a many-body system are quantum Markov chains. We overcome the difficulty and show the conjecture by explicitly constructing the maximum entropy state by using properties of quantum Markov chains. We further consider a finer partition of the regions in which the topological entropy is defined and show that the TEE is equivalent to the highest-order irreducible correlation in this case. As an application of this result, we show that the value of the TEE provides a restriction

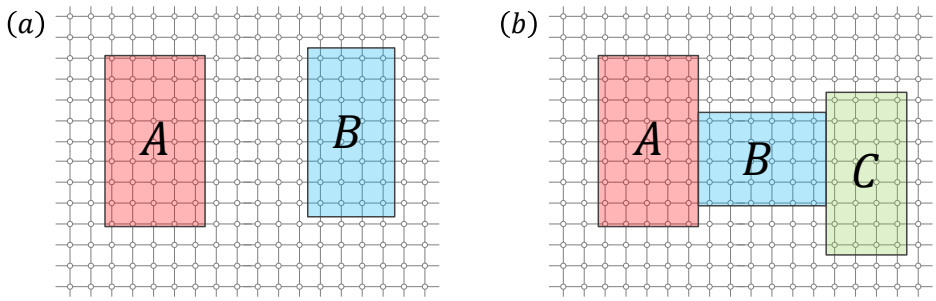

Figure 4.1: Examples of regions for our assumptions. (a): assumption (I) means that the state has no correlation between A and B. (b): assumption (II) means the reduced state on ABC is a Markov chain if B separates A from C.

to the ES of the half cylinder a spin lattice on a 2D cylinder. More precisely, we show that the ES is equivalent to the spectrum of a short-range Hamiltonian when the TEE is zero, while it is equivalent to the spectrum of a non-local Hamiltonian otherwise. The non-locality of the ES has been also observed in by using the PEPS formalism [131, 132], and we prove a correspondence between the value of the TEE and this observation.

We also provide an *operational characterization* of the characteristic correlations in topologically ordered phases. In information theory, an operational meaning of a function is a fundamental concept linking the quantitative characterization of a property and information processing. A function is said to have an operational meaning when we can interpret the value of the function as a rate or efficiency of some information theoretical task. Such an interpretation helps better understanding of the value of the function and provides new theoretical-tools. For a specific class of states called stabilizer states [133], the irreducible correlation is known to be equivalent to the asymptotic optimal rate of an information theoretical protocol called *secret sharing* [134, 70]. Here, we use completely different technique used in Refs. [134, 70] and show that the equivalence also holds for more general states satisfying our assumptions. Combining with the equivalence between the TEE and the irreducible correlation provides an operational meaning to the TEE.

This chapter is organized as follows. In Sec. 4.1, we calculate the irreducible correlation and show that the highest-order irreducible correlation coincides to the TEE. We also show an application of this result to the ES on a half cylinder. We provide a proof of these result in Sec. 4.2. In Sec. 4.3, we show the equivalence between the irreducible correlation and the optimal rate of a secret sharing protocol. The proof is given in Sec. 4.3.3. We extend results in Sec. 4.1 to states with the finite correlation length in Sec. 4.4. Finally, we discuss our results in Sec. 4.5.

4.1 The Irreducible Correlations in 2D Gapped Ground States

Let us consider a many-body system defined on a 2D spin lattice. The TEE is then defined on regions with appropriate tripartitions as depicted in Fig. 3.5. Recall that the TEE can be written as

$$S_{topo} = S(AB)_\rho + S(BC)_\rho + S(CA)_\rho - S(A)_\rho - S(B)_\rho - S(C)_\rho - S(ABC)_\rho \,.$$

The first result in this section is that if the many-body state satisfies assumptions (I) and (II), the TEE is equivalent to the 3rd-order irreducible correlation of the reduced state.

Theorem 6. *If assumptions (I) and (II) are satisfied,*

$$C^{(3)}(\rho_{ABC}) = S_{topo} \tag{4.2}$$

for all choices of regions A, B and C as depicted in Fig. 3.5.

This result shows that at least for exactly solvable models, the TEE has a clear information-geometrical meaning through the irreducible correlation, namely,

$$S_{topo} = \inf_{\sigma \in \mathcal{E}_2} S(\rho \| \sigma) \,. \tag{4.3}$$

The irreducible correlations decompose the total correlation into the class of correlations based on the order. Thus we obtain an explicit formula for the 2nd-order irreducible correlation as a corollary.

Corollary 7. *Under the same setting of Theorem 6,*

$$C^{(2)}(\rho_{ABC}) = I(A:B)_\rho + I(B:C)_\rho + I(C:A)_\rho \,, \tag{4.4}$$

where the right hand side is $I^{(2)}$, the multivariate mutual information of order 2.

We emphasize that in general the 2nd-order irreducible correlation is not the sum of the mutual information of bipartite reduced states and the relation (4.4) is a special property of ground states with zero correlation length.

The TEE can be generalized to more finer partitions. For instance, we can divide an annular region into n subregions as depicted in Fig. 4.2. In this case, the TEE γ is equivalent to S_{topo} which is defined as $I^{(m)}$ in Eq. (2.84). By using the assumptions, it can be reduced to a simpler form of

$$S_{topo} = \sum_{i=1}^{n} (S(X_i X_{i+1})_\rho - S(X_i)_\rho) - S(X_1 ... X_n)_\rho \,, \tag{4.5}$$

where we set $X_{n+1} = X_1$. We can also calculate the irreducible correlation of each order for these cases.

Theorem 8. *If assumptions (I) and (II) are satisfied, we have*

$$C^{(k)}(\rho) = \begin{cases} \sum_{i=1}^{m} I(X_i : X_{i+1})_\rho & k = 2\,, \\ 0 & 2 < k < m\,, \\ S_{topo} & k = m\,. \end{cases} \tag{4.6}$$

for regions depicted in Fig. 4.2.

Therefore, the highest-order irreducible correlation is the TEE. This result is a generalization of the previously obtained results on the irreducible correlation of the toric code model [75]. Theorem 8 shows that in ground states of gapped exactly solvable models, the total correlation contained in the region is divided to the correlations caused by nearest-neighbor (bipartite) interactions and that by genuinely m-body correlations.

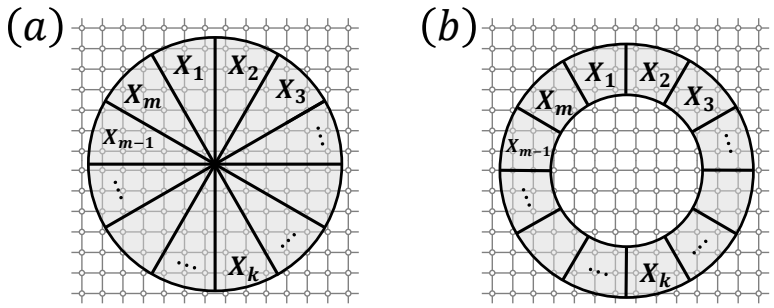

Figure 4.2: Examples of regions with finer partitions. Note that we arrange regions $X_1X_2...X_m$ in a cyclic way.

The non-zero value of $C^{(m)}(\rho)$ implies that there exists a non-local operator whose expectation value does not match to the expectation value of the maximum entropy state constructed from reduced states. Using the information-geometrical definition of the irreducible correlation, Eq. (4.6) implies that the reduced state ρ_X in the region can be written as

$$\rho_X = e^{-H_2 - H_m}\,, \tag{4.7}$$

where H_2 is a nearest-neighbor Hamiltonian and H_m is a m-body operator. If $S_{topo} > 0$, the expectation value $\langle H_m \rangle_\rho$ is different from the expectation values for $\tilde{\rho}_X^{(m-1)}$, and therefore it cannot be estimated by the maximum entropy principle.

The globalness of $\log_2 \rho_{ABC}$ is also discussed in Ref. [135] under the setting of an approximate version of assumption (II). In Ref. [135], an operator

$$H_{A:C|B} := \log \rho_{ABC} - \log \rho_{AB} - \log \rho_{BC} + \log \rho_B \tag{4.8}$$

is found to have a small correlation $\langle H_{A:C|B}, X\rangle - \langle H_{A:C|B}\rangle\langle X\rangle$ for any local operator X supported on one of A, B and C. Our result provides a finer analysis of the structure of $\log \rho_{ABC}$ under more strict conditions.

Since the expectation value of the non-local operator cannot be determined locally, it can be used to encode a "secret" which is hidden from parties having only local information of the state. We will discuss an application of this property to secret-sharing protocols in Sec. 4.3.

4.1.1 Relation to Entanglement Spectrum of A Cylinder

As an application, we connect the results in Sec. 4.1 to the ES of a ground state defined on a cylinder. In Refs. [131, 132], the ES of a region wrapping a cylinder has studied by the PEPS formalism [39]. It has been observed that the ES of a ground state in the topologically trivial is given by the spectrum of a short-range Hamiltonian acting on the virtual boundary. If the ground state is in a topologically ordered phase, the ES is the spectrum of a Hamiltonian containing universal (i.e., independent of local properties of the system) non-local interactions. These results reveal a new structural characterization of the ES apart from the original work of the entanglement spectrum by Li and Haldane [119].

Here, we consider a similar setting outside of the PEPS formalism, and show that a similar property holds under our assumptions (I) and (II). We further show that the value of the TEE of the wrapping region characterizes the non-locality of the ES.

Let us consider a system as depicted in Fig. 4.3. We are interested in the ES of the half cylinder, i.e., the spectrum of the entanglement Hamiltonian

$$H_{\rho_Y} := -\ln \rho_Y \,. \tag{4.9}$$

To measure the difference between entanglement Hamiltonians, we introduce a "weighted norm"

$$\|A\|_\rho := \mathrm{Tr}\left(|\rho^{\frac{1}{2}} A \rho^{\frac{1}{2}}|\right) \tag{4.10}$$

which is a proper norm only for operators on the support of ρ. We denote the set of all bounded nearest-neighbor Hamiltonians by

$$\mathcal{E}_{nn} := \left\{ H = \sum_{i=1}^{n} H_{X_i X_{i+1}} \right\} . \tag{4.11}$$

Note that $\mathcal{E}_{nn}$ is a subset of $\mathcal{E}_2$ defined in Sec. 2.3.1. For a pure bipartite states, the spectrum of the reduced state on a subsystem matches the spectrum on the complement. This property leads to the following theorem.

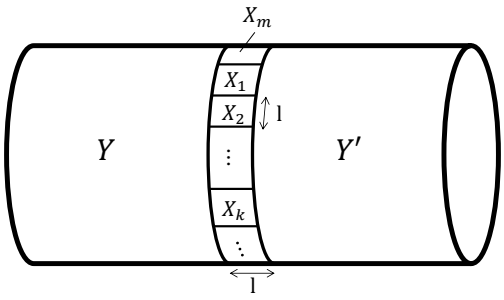

Figure 4.3: A 2D spin lattice on a cylinder. We choose $m = \Theta\left(|X|/(\log_2|X|)^2\right)$. The horizontal length of Y and Y' can be arbitrary large. The region X can be viewed as a 1D "boundary" of Y if we divide the cylinder into two halves.

Theorem 9. *Suppose that a pure state $\rho = |\psi\rangle\langle\psi|_{YXY'}$ defined on the cylinder depicted in Fig. 4.3 satisfies assumption (I) and (II). Moreover, we assume that the system obeys the reflection symmetry such that*

$$\rho_Y = \rho_{Y'}\,. \tag{4.12}$$

Then, there exists a Hamiltonian H_X on $X = X_1...X_m$ such that for any $\Lambda > 0$, it holds that

$$\lambda\left(H^{(2)}_{\rho_Y}\right) = \lambda(H_X)\,, \tag{4.13}$$

where

$$H^{(2)}_{\rho_Y} = H_{\rho_Y}\otimes I + I\otimes H_{\rho_Y} \tag{4.14}$$

acting on $\mathcal{H}_Y^{\otimes 2}$. Moreover, H_X satisfies

$$\inf_{H'\in\mathscr{H}_{n-1}}\|H_X - H'\|_\rho = \inf_{H'\in\mathscr{H}_{nn}}\|H_X - H'\|_\rho = S_{topo}\,. \tag{4.15}$$

Proof. Since $|\psi_{YXY'}\rangle$ is pure, it holds that

$$\lambda(\rho_{YY'}) = \lambda(\rho_X)\,. \tag{4.16}$$

Assumption (II) implies

$$I(Y:Y')_\rho = I(Y:Y'|X)_\psi = 0\,. \tag{4.17}$$

Therefore, $\rho_{YY'} = \rho_Y^{\otimes 2}$ and by taking the logarithm of both sides, we obtain Eq. (4.13). Equation (4.15) follows from Theorem 8. □

Theorem 9 implies that in topologically ordered phases, the (double of) ES on Y is the spectrum of some non-local Hamiltonian defined on the "boundary" X. In

other words, there exists an isometry V converting $H^{(2)}_{\rho_Y}$ defined on a 2D system to a 1D non-local Hamiltonian. The value of γ (and therefore S_{topo}) depends on not only the type of the topological order, but also on the choice of the ground state for non-contractible region, e.g., region X presented in Fig. 4.3 or a similar region on a torus [136]. Moreover, there exists a specific choice of the ground state of the topological model with vanishing TEE on a such region [136]. Our theorem implies that even in a topologically ordered phase, the non-local part of H_X disappears and the ES coincides to the spectrum of a short-range Hamiltonian if we choose this kind of ground states. The disappearance of the non-local part has also been observed in the analysis of the PEPS formalism [132], though completely different method.

4.2 Proof: The Irreducible Correlation in States with Zero Correlation Length

In this section, we provide a proof for the main theorems presented in the previous section. We first show the proof of Theorem 6 for regions with partitions depicted by Fig. 4.4(b), which we call the Levin-Wen type partition. The proof explicitly exploits the structure of quantum Markov chains discussed in Sec. 2.2.3. Then we show the proof for the regions with another type of partitions depicted by Fig. 4.4(a), the Kitaev-Preskill type partition, in a slightly different procedure. The sketch of the proofs is as follows. We first divide each subregion that connects two different subregions into two halves (see Fig. 4.4(a') and (b')). Then, each reduced state on three consecutive regions becomes a quantum Markov chain due to assumption (II). We then construct a global state by "merging" such local Markov states. Since the global state is constructed from local reduced states of the original state, it is a candidate for the maximum entropy state. Indeed, we prove that the constructed state is the maximum entropy state with the same local reduced state as the original. The entropy of the maximum entropy state can be calculated by a property of quantum Markov chains which establishes the theorem.

4.2.1 Levin-Wen Type Partitions

Proof. Let us consider regions as depicted in Fig. 4.4. From assumption (I), we have $I(A : C)_\rho = 0$ and thus $S_{topo} = I(A : C|B)_\rho$. By dividing region B as in Fig. 4.4(b'), we obtain that $I(A : B_2|B_1)_\rho = 0$ and $I(B_1 : C|B_2)_\rho = 0$ by assumption (II). From the property of quantum Markov chains (2.47), there exist

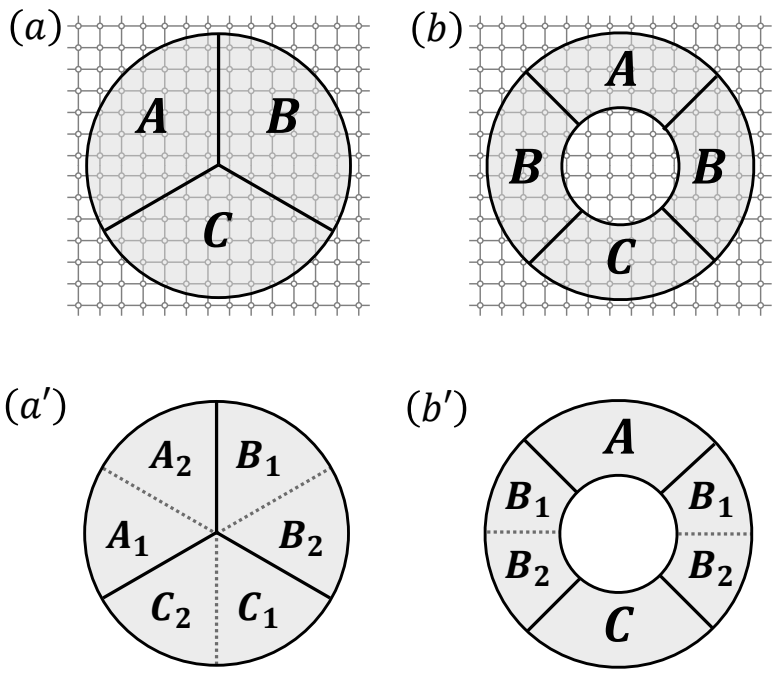

Figure 4.4: Regions with appropriate divisions. To prove the main theorems, we divide regions (a) and (b) into (a') and (b'), respectively.

recovery maps $\Lambda_{B_1\to AB_1}$ and $\Lambda_{B_2\to B_2C}$ such that

$$\rho_{AB_1B_2} = \Lambda_{B_1\to AB_1}(\rho_{B_1B_2})\,, \quad \rho_{B_1B_2C} = \Lambda_{B_2\to B_2C}(\rho_{B_1B_2})\,. \tag{4.18}$$

We obtain a global state on ABC as

$$\begin{aligned}\tilde{\rho}_{ABC} &:= \Lambda_{B_2\to B_2C}(\rho_{AB}) &(4.19)\\ &= \Lambda_{B_1\to AB_1}(\rho_{BC}) &(4.20)\\ &= \Lambda_{B_1\to AB_1}\circ\Lambda_{B_2\to B_2C}(\rho_B)\,. &(4.21)\end{aligned}$$

We now show that all bipartite reduced states of $\tilde{\rho}_{ABC}$ are same as ρ_{ABC}. Since Tr_C commutes with $\Lambda_{B_1\to AB_1}$, it immediately follows that

$$\tilde{\rho}_{AB} = \Lambda_{B_1\to AB_1}\circ\mathrm{Tr}_C(\rho_{BC}) = \Lambda_{B_1\to AB_1}(\rho_{B_1B_2}) = \rho_{AB}\,. \tag{4.22}$$

Similarly, $\tilde{\rho}_{BC} = \rho_{BC}$. From assumption (I), we also have $I(A:B_2)_\rho = 0$ and therefore

$$\begin{aligned}\tilde{\rho}_{AB_2C} &= \mathrm{Tr}_{B_1}\circ\Lambda_{B_2\to B_2C}(\rho_{AB_1B_2}) &(4.23)\\ &= \Lambda_{B_2\to B_2C}(\rho_A\otimes\rho_{B_2}) &(4.24)\\ &= \Lambda_{B_2\to B_2C}\left(\rho_A\otimes\mathrm{Tr}_{B_1}\rho_{B_1B_2}\right) &(4.25)\\ &= \rho_A\otimes\mathrm{Tr}_{B_1}\circ\Lambda_{B_2\to B_2C}(\rho_{B_1B_2}) &(4.26)\\ &= \rho_A\otimes\rho_{B_2C}\,. &(4.27)\end{aligned}$$

Thus, $\tilde{\rho}_{AC} = \rho_A\otimes\rho_C = \rho_{AC}$ and $\tilde{\rho}_{ABC}\in\mathcal{C}_\rho^{(2)}$.

$\tilde{\rho}_{ABC}$ is a quantum Markov chain conditioned on B, since $\tilde{\rho}_{ABC} = \Lambda_{B_2\to B_2C}(\rho_{AB}) = \Lambda_{B_2\to B_2C}(\rho_{AB})$ (see Sec. 2.2.3). Quantum Markov chains saturate strong subadditivity (2.35), therefore

$$\begin{aligned}S(ABC)_{\tilde{\rho}} &= S(AB)_{\tilde{\rho}} + S(BC)_{\tilde{\rho}} - S(B)_{\tilde{\rho}} &(4.28)\\ &= S(AB)_\rho + S(BC)_\rho - S(B)_\rho\,. &(4.29)\end{aligned}$$

Due to strong subadditivity of an arbitrary state $\sigma_{ABC}\in\mathcal{C}_\rho^{(2)}$ implies

$$\begin{aligned}S(ABC)_\sigma &\le S(AB)_\sigma + S(BC)_\sigma - S(B)_\sigma &(4.30)\\ &= S(AB)_\rho + S(BC)_\rho - S(B)_\rho &(4.31)\\ &= S(ABC)_{\tilde{\rho}}\,. &(4.32)\end{aligned}$$

Therefore, $\tilde{\rho}_{ABC}$ is the maximum entropy state in $\mathcal{C}_\rho^{(2)}$. The 3rd-order irreducible correlation of ρ_{ABC} is calculated as

$$\begin{aligned}C^{(3)}(\rho_{ABC}) &= S(ABC)_{\tilde{\rho}} - S(ABC)_\rho &(4.33)\\ &= I(A:C|B)_\rho &(4.34)\\ &= S_{topo}\,, &(4.35)\end{aligned}$$

where we used Eq. (4.29) in the second line. Thus completes the proof. □

For later convenience, we show that there exist decompositions $\mathcal{H}_{B_1} = \bigoplus_i \mathcal{H}_{B_{1i}^L} \otimes \mathcal{H}_{B_{1i}^R}$ and $\mathcal{H}_{B_2} = \bigoplus_j \mathcal{H}_{B_{2j}^L} \otimes \mathcal{H}_{B_{2j}^R}$ such that $\tilde{\rho}_{ABC}$ can be decomposed as

$$\tilde{\rho}_{ABC} = \bigoplus_{i,j} p_{B_1}(i) p_{B_2}(j|i) \rho_{AB_{1i}^L} \otimes \rho_{B_{1i}^R B_{2j}^L} \otimes \rho_{B_{2j}^R C} \,. \tag{4.36}$$

Since $I(B_1 : C|B_2)_\rho = 0$, there exists a decomposition

$$\rho_{BC} = \bigoplus_{j} p_{B_2}(j) \rho_{B_1 B_{2j}^L} \otimes \rho_{B_{2j}^R C} \,. \tag{4.37}$$

We denote the orthogonal projector on $\mathcal{H}_{B_{1i}^L} \otimes \mathcal{H}_{B_{1i}^R}$ by $\Pi^i_{B_1}$ and on $\mathcal{H}_{B_{2j}^L} \otimes \mathcal{H}_{B_{2j}^R}$ by $\Pi^j_{B_2}$. From $I(A : B_2|B_1)_\rho = 0$, ρ_{AB} is decomposed as

$$\rho_{AB} = \bigoplus_{i} p_{B_1}(i) \rho_{AB_{1i}^L} \otimes \rho_{B_{1i}^R B_2} \,, \tag{4.38}$$

and therefore

$$\rho_{B} = \bigoplus_{i} p_{B_1}(i) \rho_{AB_{1i}^L} \otimes \rho_{B_{1i}^R B_2} \,. \tag{4.39}$$

Let us consider a CPTP-map $\mathcal{P} : \mathcal{S}(\mathcal{H}_B) \to \mathcal{S}(\mathcal{H}_B)$ defined as

$$\mathcal{P}(\sigma_B) = \bigoplus_{j} \mathrm{Tr}_{B_{2j}^R C} \left[\Pi^j_{B_2} \sigma_{B_1} \Pi^j_{B_2} \right] \otimes \rho_{B_{2j}^R C} \,. \tag{4.40}$$

Since $\mathcal{P}(\rho_{AB}) = \rho_{AB}$ by construction, it holds that

$$\begin{aligned} \rho_{AB} &= \mathcal{P}(\rho_{AB}) && (4.41)\\ &= \bigoplus_{i,j} p_{B_1}(i) \rho_{AB_{1i}^L} \otimes \mathrm{Tr}_{B_{2j}^R} \left[\Pi^j_{B_2} \rho_{B_{1i}^R B_2} \Pi^j_{B_2} \right] \otimes \rho_{B_{2j}^R} \,. && (4.42) \end{aligned}$$

For any i and j,

$$\begin{aligned} \mathrm{Tr} \left[\Pi^j_{B_2} \rho_{B_{1i}^R B_2} \Pi^j_{B_2} \right] &= \mathrm{Tr} \left[\rho_{B_{1i}^L} \otimes \Pi^j_{B_2} \rho_{B_{1i}^R B_2} \Pi^j_{B_2} \right] && (4.43)\\ &= \mathrm{Tr} \left[\Pi^i_{B_1} \Pi^j_{B_2} \rho_{B_1 B_2} \Pi^j_{B_2} \Pi^i_{B_1} \right] / p_{B_1}(i) && (4.44)\\ &= p_{B_2}(j|i) \,. && (4.45) \end{aligned}$$

By setting

$$p_{B_2}(j|i) \rho_{B_{1i}^R B_{2j}^L} := \mathrm{Tr}_{B_{2j}^R} \left[\Pi^j_{B_2} \rho_{B_{1i}^R B_2} \Pi^j_{B_2} \right] \,, \tag{4.46}$$

we see that

$$\rho_{AB} = \bigoplus_{i,j} p_{B_1}(i) p_{B_2}(j|i) \rho_{AB_{1i}^L} \otimes \rho_{B_{1i}^R B_{2j}^L} \otimes \rho_{B_{2j}^R} \,. \tag{4.47}$$

We have

$$\rho_{BC} = \bigoplus_j p_{B_2}(j)\rho_{B_1 B_{2j}^L} \otimes \Lambda_{B_2 \to B_2 C}(\rho_{B_{2j}^R}) \,, \tag{4.48}$$

since $\Lambda_{B_2 \to B_2 C}$ only acts on B_2^R. Thus

$$\Lambda_{B_2 \to B_2 C}(\rho_{B_{2j}^R}) = \rho_{B_{2j}^R C} \tag{4.49}$$

follows from Eq. (4.37). Finally, we obtain Eq. (4.36) by performing $\Lambda_{B_2 \to B_2 C}$ to ρ_{AB} in Eq. (4.47).

4.2.2 Kitaev-Preskill Type Partitions

The proof for Kitaev-Preskill type partitions is more involved since it contains a cyclic structure. In this case, we only use the direct sum structure of quantum Markov chains rather than the recovery maps used in the case of Levin-Wen type partitions.

Proof. We divide all three subregions into two halves as in Fig. 4.4(a'). It is convenient to denote $A_1, A_2, B_1, ...C_2$ by $X_1, X_2, X_3, ...X_6$ interchangeably in this case. Due to the periodicity of these regions, we also set $X_7 \equiv X_1$. For any neighboring regions $X_{i-1}X_iX_{i+1}$, the reduced state is a quantum Markov chain conditioned on X_i. Therefore, for each i, there exists a decomposition $\mathcal{H}_{X_i} = \bigoplus_{j_i} \mathcal{H}^L_{X_i(j_i)} \otimes \mathcal{H}^R_{X_i(j_i)}$ such that the reduced state $\rho_{X_{i-1}X_iX_{i+1}}$ can be decomposed as

$$\rho_{X_{i-1}X_iX_{i+1}} = \bigoplus_{j_i} p_i(j_i)\rho_{X_{i-1}X^L_{i(j_i)}} \otimes \rho_{X^R_{i(j_i)}X_{i+1}} \,, \tag{4.50}$$

where $\{p_i(j_i)\}_{j_i}$ is a probability distribution. We denote the orthogonal projector on $\mathcal{H}^L_{X_i(j_i)} \otimes \mathcal{H}^R_{X_i(j_i)}$ by $\Pi^{(i)}_{j_i}$.

Our goal is to show that the maximum entropy state in $\mathcal{C}^{(2)}_\rho$ can be written as

$$\begin{aligned}\tilde{\rho}_{ABC} = \bigoplus_{i_1,\ldots,i_6} & p_1(i_1|i_6)p_2(i_2|i_1)\cdots p_6(i_6|i_5)\times \\ & \rho_{A^R_{1(i_1)}A^L_{2(i_2)}} \otimes \rho_{A^R_{2(i_2)}B^L_{1(i_3)}} \otimes \cdots \otimes \rho_{C^R_{2(i_6)}A^L_{1(i_1)}} \,,\end{aligned} \tag{4.51}$$

where $p_j(i_j|i_{j-1}) = \mathrm{Tr}(\Pi^{(j)}_{i_j}\Pi^{(j-1)}_{i_{j-1}}\rho_{ABC})/\mathrm{Tr}(\Pi^{(j-1)}_{i_{j-1}}\rho_{ABC})$. As long as it is clear from the context, we omit the lower index for the probabilities $p_j(i_j, i_{j-1})$ and simply write $p(i_j, i_{j-1})$.

We first show that the cyclic products $p(i_1|i_6)\cdots p(i_6|i_5)$ form a probability distribution (i.e., $\tilde{\rho}_{ABC}$ is a quantum state) under assumption (I) and (II). The non-negativity is clear because each conditional probability is non-negative. We

further restrict the domain indices so that the whole distribution is strictly positive. The normalization condition can be shown by the following calculation:

$$\sum_{i_1,\ldots,i_6} p(i_1|i_6)p(i_2|i_1)\cdots p(i_6|i_5) \tag{4.52}$$

$$= \sum_{i_2,\ldots,i_6} \left(\frac{\sum_{i_1} p(i_6|i_1)p(i_2|i_1)p(i_1)}{p(i_6)} \right) p(i_3|i_2)\cdots p(i_6|i_5) \tag{4.53}$$

$$= \sum_{i_2,\ldots,i_6} \left(\sum_{i_1} \frac{p(i_6,i_1,i_2)}{p(i_6)} \right) p(i_3|i_2)\cdots p(i_6|i_5) \tag{4.54}$$

$$= \sum_{i_2,\ldots,i_6} \frac{p(i_6,i_2)}{p(i_6)} p(i_3|i_2)\cdots p(i_6|i_5) \tag{4.55}$$

$$= \sum_{i_2,\ldots,i_6} p(i_2)p(i_3|i_2)\cdots p(i_6|i_5) \tag{4.56}$$

$$= \sum_{i_3,\ldots,i_6} p(i_3)p(i_4|i_3)p(i_5|i_4)p(i_6|i_5) = \cdots = 1\,. \tag{4.57}$$

Here we used the Bayes rule $p(i|j) = p(j|i)p(i)/p(j)$ in the first line. In the second line, we used that $p(i_6, i_1, i_2) = p(i_1)p(i_6|i_1)p(i_2|i_1)$, which follows since $\rho_{C_2A_1A_2}$ is a quantum Markov chain with a decomposition:

$$\rho_{C_2A_1A_2} = \bigoplus_{i_1} p(i_1)\rho_{C_2A^L_{1(i_1)}} \otimes \rho_{A^R_{1(i_1)}A_2} \tag{4.58}$$

and therefore a measurement of i_6 is conditionally independent of i_2. The fourth equality follows from $p(i_6, i_2) = p(i_6)p(i_2)$, which holds since $I(C_2 : A_2)_\rho = 0$.

Next, we show that $\tilde{\rho}_{ABC} \in \mathcal{C}^{(2)}_\rho$. To do so, we first decompose ρ_{AB} as

$$\rho_{AB} = \bigoplus_{i_2,i_3} p(i_2)p(i_3|i_2)\rho_{A_1A^L_{2(i_2)}} \otimes \rho_{A^R_{2(i_2)}B^L_{1(i_3)}} \otimes \rho_{B^R_{1(i_3)}B_2}\,. \tag{4.59}$$

This follows from the same argument deriving Eq. (4.36) and the fact that ρ_{AB} is the maximum entropy state on a tripartite system $A_1(A_2B_1)B_2$ since it is a quantum Markov chain satisfying $I(A_1 : B_2|A_2B_1)_\rho = 0$. In a similar way, we have decompositions

$$\rho_{BC} = \bigoplus_{i_4,i_5} p(i_4)p(i_5|i_4)\rho_{B_1B^L_{2(i_4)}} \otimes \rho_{B^R_{2(i_4)}C^L_{1(i_5)}} \otimes \rho_{C^R_{1(i_5)}C_2} \tag{4.60}$$

and

$$\rho_{AC} = \bigoplus_{i_6,i_1} p(i_6)p(i_1|i_6)\rho_{C_1C^L_{2(i_6)}} \otimes \rho_{C^R_{2(i_6)}A^L_{1(i_1)}} \otimes \rho_{A^R_{1(i_1)}A_2}\,. \tag{4.61}$$

Without loss of generality, we focus on proving $\tilde{\rho}_{AB} = \rho_{AB}$ in the following, since the same arguments can be applied to systems BC and CA owing to the symmetry of the problem. We can obtain a finer decomposition by considering a measurement corresponding to projections $\Pi_{j_i}^{(i)}$. For instance, by decomposing C_1 into C_1^L and C_1^R, Eq. (4.61) can be written as

$$\rho_{AC} = \bigoplus_{i_5,i_6,i_1} p(i_6)p(i_1|i_6)p(i_5|i_6)\rho_{C^L_{1(i_5)}} \otimes \rho_{C^R_{1(i_5)}C^L_{2(i_6)}} \otimes \rho_{C^R_{2(i_6)}A^L_{1(i_1)}} \otimes \rho_{A^R_{1(i_1)}A_2} \,. \tag{4.62}$$

By tracing out $C_1^R C_2 A_2$ from Eq. (4.61) [1], we obtain that

$$\rho_{A_1C} = \bigoplus_{i_5,i_6} p(i_5,i_6)\rho_{C^L_{1(i_5)}} \otimes \rho^{i_6}_{A_1} \,, \tag{4.63}$$

where $\rho^{i_6}_{A^L_{1(i_1)}} := \mathrm{Tr}_{C_2^R}\rho_{C^R_{2(i_6)}A^L_{1(i_1)}}$ and $\rho^{i_6}_{A_1} := \bigoplus_{i_1} p(i_1|i_6)\rho^{i_6}_{A^L_{1(i_1)}} \otimes \rho_{A^R_{1(i_1)}}$. Since $I(C_1 : A_1)_\rho = 0$, it must be that $p(i_5,i_6) = p(i_5)p(i_6)$ or $\rho^{i_6}_{A_1} = \rho_{A_1}$. Similarly, $I(B_2 : C_2)_\rho = 0$ leads to either $p(i_5,i_6) = p(i_5)p(i_6)$ holds or $\rho^{i_5}_{B_2} := \bigoplus_{i_4} p(i_4|i_5)\rho_{B^L_{2(i_4)}} \otimes \rho^{i_5}_{B^R_{2(i_4)}}$ is independent of i_5, where $\rho^{i_5}_{B^R_{2(i_4)}} := \mathrm{Tr}_{C_1^L}\rho_{B^R_{2(i_4)}C^L_{1(i_5)}}$.

If $p(i_5,i_6) = p(i_5)p(i_6)$ holds, then

$$\rho_A = \bigoplus_{i_6,i_1,i_2} p(i_6)p(i_1|i_6)p(i_2|i_1)\rho^{i_6}_{A^L_{1(i_1)}} \otimes \rho_{A^R_{1(i_1)}A^L_{2(i_2)}} \otimes \rho_{A^R_{2(i_2)}} \tag{4.64}$$

$$= \bigoplus_{i_2} p(i_2)\rho_{A_1A^L_{2(i_2)}} \otimes \rho_{A^R_{2(i_2)}} \,. \tag{4.65}$$

ρ_B has a similar decomposition and therefore

$$\rho_B = \bigoplus_{i_3,i_4,i_5} p(i_5)p(i_4|i_5)p(i_3|i_4)\rho_{B^L_{1(i_3)}} \otimes \rho_{B^R_{1(i_3)}B^L_{2(i_4)}} \otimes \rho^{i_5}_{B^R_{2(i_4)}} \tag{4.66}$$

$$= \bigoplus_{i_3} p(i_3)\rho_{B^L_{1(i_3)}} \otimes \rho_{B^R_{1(i_3)}B_2} \,. \tag{4.67}$$

[1] The partial trace over C_1^R from $\mathcal{H}_{C_1} = \bigoplus_{i_5} \mathcal{H}_{C^L_{1(i_5)}} \otimes \mathcal{H}_{C^R_{1(i_5)}}$ is performed by tracing out $\mathcal{H}_{C^R_{1(i_5)}}$ from each direct sum component.

Since $p(i_6|i_5) = p(i_6)$, from Eq. (4.64) and Eq. (4.2.2) it holds that

$$\begin{aligned}
\tilde{\rho}_{AB} &= \bigoplus_{i_1,\ldots,i_6} p(i_1|i_6)p(i_2|i_1)\cdots p(i_5|i_4)p(i_6) \\
&\qquad \times \rho^{i_6}_{A^L_{1(i_1)}} \otimes \rho_{A^R_{1(i_1)}A^L_{2(i_2)}} \otimes \rho_{A^R_{2(i_2)}B^L_{1(i_3)}} \otimes \rho_{B^R_{1(i_3)}B^L_{2(i_4)}} \otimes \rho^{i_5}_{B^R_{2(i_4)}} && (4.68)\\
&= \bigoplus_{i_2,\ldots,i_5} p(i_2)p(i_3|i_2)p(i_4|i_3)p(i_5|i_4)\rho_{A_1A^L_{2(i_2)}} \otimes \rho_{A^R_{2(i_2)}B^L_{1(i_3)}} \otimes \rho_{B^R_{1(i_3)}B^L_{2(i_4)}} \otimes \rho^{i_5}_{B^R_{2(i_4)}} && (4.69)\\
&= \bigoplus_{i_2,\ldots,i_5} p(i_5)p(i_4|i_5)p(i_3|i_4)p(i_2|i_3)\rho_{A_1A^L_{2(i_2)}} \otimes \rho_{A^R_{2(i_2)}B^L_{1(i_3)}} \otimes \rho_{B^R_{1(i_3)}B^L_{2(i_4)}} \otimes \rho^{i_5}_{B^R_{2(i_4)}} && (4.70)\\
&= \bigoplus_{i_2,i_3} p(i_3)p(i_2|i_3)\rho_{A_1A^L_{2(i_2)}} \otimes \rho_{A^R_{2(i_2)}B^L_{1(i_3)}} \otimes \rho_{B^R_{1(i_3)}B_2} && (4.71)\\
&= \rho_{AB}\,. && (4.72)
\end{aligned}$$

Here the third line follows from the Bayes rule $p(i)p(j|i) = p(j)p(i|j)$.

If $\rho^{i_6}_{A_1} = \rho_{A_1}$ and $\rho^{i_5}_{B_2} = \rho_{B_2}$ hold, a simple calculation shows

$$\begin{aligned}
\tilde{\rho}_{AB} &= \bigoplus_{i_1,\ldots,i_4} p(i_1)p(i_2|i_1)\cdots p(i_4|i_3)\rho_{A^L_{1(i_1)}} \otimes \rho_{A^R_{1(i_1)}A^L_{2(i_2)}} \\
&\qquad \otimes \rho_{A^R_{2(i_2)}B^L_{1(i_3)}} \otimes \rho_{B^R_{1(i_3)}B^L_{2(i_4)}} \otimes \rho_{B^R_{2(i_4)}} && (4.73)\\
&= \bigoplus_{i_2,\ldots,i_4} p(i_2)p(i_3|i_2)p(i_4|i_3)\rho_{A_1A^L_{2(i_2)}} \otimes \rho_{A^R_{2(i_2)}B^L_{1(i_3)}} \\
&\qquad \otimes \rho_{B^R_{1(i_3)}B^L_{2(i_4)}} \otimes \rho_{B^R_{2(i_4)}} && (4.74)\\
&= \bigoplus_{i_2,\ldots,i_3} p(i_2)p(i_3|i_2)\rho_{A_1A^L_{2(i_2)}} \otimes \rho_{A^R_{2(i_2)}B^L_{1(i_3)}} \otimes \rho_{B^R_{1(i_3)}B_2} && (4.75)\\
&= \rho_{AB}\,. && (4.76)
\end{aligned}$$

Note that in the first equality we used

$$\begin{aligned}
&\sum_{i_5,i_6} p(i_1|i_6)p(i_2|i_1)p(i_3|i_2)p(i_4|i_3)p(i_5|i_4)p(i_6|i_5) \\
&\quad = p(i_2|i_1)p(i_3|i_2)p(i_4|i_3)\sum_{i_5,i_6} p(i_1|i_6)p(i_5|i_4)p(i_6|i_5) && (4.77)\\
&\quad = p(i_2|i_1)p(i_3|i_2)p(i_4|i_3)\sum_{i_6} p(i_1|i_6)/p(i_4)\sum_{i_5} p(i_5)p(i_4|i_5)p(i_6|i_5) && (4.78)\\
&\quad = p(i_2|i_1)p(i_3|i_2)p(i_4|i_3)\sum_{i_6} p(i_1|i_6)/p(i_4)\sum_{i_5} p(i_4,i_5,i_6) && (4.79)
\end{aligned}$$

$$= p(i_2|i_1)p(i_3|i_2)p(i_4|i_3)\sum_{i_6} p(i_1|i_6)p(i_6) \tag{4.80}$$

$$= p(i_2|i_1)p(i_3|i_2)p(i_4|i_3)\sum_{i_6} p(i_1,i_6) \tag{4.81}$$

$$= p(i_1)p(i_2|i_1)p(i_3|i_2)p(i_4|i_3)\,. \tag{4.82}$$

Here, the fourth line follows from $p(i_4,i_6) = p(i_4)p(i_6)$.

Finally, it remains to show that $\tilde{\rho}_{ABC}$ is the maximum entropy state. We express $\tilde{\rho}_{ABC}$ in a more convenient form by defining new indices $a = (i_1,i_2), b = (i_3,i_4)$ and $c = (i_5,i_6)$ so that

$$\tilde{\rho}_{ABC} = \bigoplus_{a,b,c} p(a|c)p(b|a)p(c|b)\rho_{A_a^R B_b^L} \otimes \rho_{B_b^R C_c^L} \otimes \rho_{C_c^R A_a^L}\,. \tag{4.83}$$

We define the Hamiltonian $H_{ABC} := H_{AB} + H_{BC} + H_{CA}$, where

$$H_{AB} = -\sum_{a,b} \ln[p(b|a)\rho_{A_a^R B_b^L}]\,, \tag{4.84}$$

$$H_{BC} = -\sum_{b,c} \ln[p(c|b)\rho_{B_b^R C_c^L}]\,, \tag{4.85}$$

$$H_{AB} = -\sum_{a,c} \ln[p(a|c)\rho_{C_c^R A_a^L}]\,. \tag{4.86}$$

It is clear that $\tilde{\rho}_{ABC} = e^{-H_{ABC}}$ and hence $\tilde{\rho}_{ABC} \in \overline{\mathcal{E}}_2^{rI}$. This guarantees that $\tilde{\rho}_{ABC}$ is the maximum entropy state (see Sec. 2.3.1).

The entropy of $\tilde{\rho}_{ABC}$ is calculated through the formula

$$S(ABC)_{\tilde{\rho}} = H(ABC)_p + \sum_{a,b,c} p(a,b,c)\left(S\left(\rho_{A_a^R B_b^L}\right) + S\left(\rho_{B_b^R C_c^L}\right) + S\left(\rho_{C_c^R A_a^L}\right)\right) \tag{4.87}$$

and decompositions

$$\tilde{\rho}_{AB} = \rho_{AB} = \bigoplus_{a,b} p(a,b)\rho_{A_a^L} \otimes \rho_{A_a^R B_b^L} \otimes \rho_{B_b^R}\,, \tag{4.88}$$

which follows from Eq. (4.83) and independence of the reduced states from traced-out labels. As result, we obtain

$$S(ABC)_{\tilde{\rho}} = S(AB)_\rho + S_(BC)_\rho + S(CA)_\rho - S(A)_\rho - S(B)_\rho - S(C)_\rho\,, \tag{4.89}$$

which completes the proof. □

4.2.3 Proof of Theorem 8

Theorem 8 is proved in a similar way to the proof for the Kitaev-Preskill type partition. We divide each region X_i into two halves and relabel the $2m$ regions by $\tilde{X}_i$, where $i = 1, ..., 2m$. In the same way to Eq. (4.51), we can show that $\tilde{\rho}^{(m-1)}$ can be written as

$$\tilde{\rho}_X^{(m-1)} = \bigoplus_{i_1,...,i_{2m}} p_1(i_1|i_{2m})p_2(i_2|i_1)\cdots p_{2m}(i_{2m}|i_{2m-1})\times \rho_{\tilde{X}^R_{1(i_1)}\tilde{X}^L_{2(i_2)}} \otimes \cdots \otimes \rho_{\tilde{X}^R_{2m(i_{2m})}\tilde{X}^L_{1(i_1)}}. \tag{4.90}$$

$\tilde{\rho}_X^{(m-1)}$ is equal to $\tilde{\rho}_X^{(2)}$, since it can be written as the Gibbs state of a nearest-neighbor Hamiltonian. Therefore, from the 3nd to the $(m-1)$th order irreducible correlation are zero. The value can be calculated through the direct sum decomposition of $\tilde{\rho}^{(m-1)}$ in the same way as in Eq. (4.89).

4.3 Equivalence to The Optimal Rate of A Secret Sharing Protocol

In this section, we provide an operational meaning to the TEE as the asymptotic optimal rate of a secret sharing protocol, by using the explicit structures of the maximum entropy states revealed in Sec. 4.2. Another analysis of a relation between the TEE and secret sharing has been investigated in the thermodynamic limit of the toric code by using operator algebras [137]. Note that in this work, the authors consider abelian models and different geometry of regions.

We first review what secret sharing protocols are in Sec. 4.3.1, and then introduce our setting and the main result of this section. In Sec. 4.3.4, we show an explicit encoding scheme for the toric code model by using loop operators, which can be generalized to all abelian quantum double models [50].

4.3.1 Secret Sharing Protocol

A secret sharing protocol is a protocol for distributing a secret message amongst players such that only allowed groups of players can read out the shared secret. It is first proposed by Shamir [138] and Blakley [139] in classical information theory, and quantum secret sharing was introduced by Hillery, Buzek and Berthiaume [140]. In quantum secret sharing, a secret is given by either a classical message or quantum state being distributed to multiple quantum systems.

There are mainly two classes of secret sharing protocols, called threshold schemes and ramp schemes. In a threshold scheme, groups which are forbidden to read secret cannot obtain *any* information about the secret, while in a ramp scheme they

can obtain *partial* information. A simple example of a threshold secret sharing is constructed by using a k-degree polynomial function $f(x)$. For some values $\{x_i\}_{i=1}^n$, one can distribute a secret, for example the value of $f(0)$, by sending $f(x_i \neq 0)$ to n players. They can determine the function (and therefore the secret as well) if and only if more than k players collaborate together, otherwise they cannot obtain any information about $f(0)$.

Traditionally, it is important to find particular distributions or quantum states to implement secret sharing protocols satisfying certain properties. In this thesis, in contrast, we employ a secret sharing protocol to *characterize* multipartite correlations contained in a *given* state. In this scenario, multipartite correlations are regarded as a *resource* to encode secrets.

For so-called stabilizer states [133], the irreducible correlation is shown to be written in terms of the asymptotic optimal rate of a particular type of secret sharing protocols [134, 70]. A key idea behind this equivalence is that if a state contains information which cannot be determined from information of $(k-1)$-partite reduced states, we can use such information as a secret hiding from $(k-1)$ parties. Actually, the total correlation is shown to coincide to the optimal rate of a secret sharing protocol, where secrets are encoded by local unitaries and hidden from any player who only has access to the reduced state on one subsystem [143].

Example 10. *Let us consider the GHZ-state $|GHZ_3\rangle$ for example. It is easy to show that $|GHZ_3\rangle$ and $(Z \otimes I \otimes I)|GHZ_3\rangle$ are orthogonal each other, but have exactly the same bipartite reduced states. Therefore, we can encode a secret bit $a \in \{0,1\}$ to the GHZ state by performing Z^a, where a is hidden from all players who can only access to any two of three subsystems. Moreover, we can encode secret bits b, c, which are hidden from one subsystem but not two, by applying $I \otimes X^b \otimes I$ and $I \otimes I \otimes X^c$. This is consistent to the irreducible correlation of the GHZ state:*

$$C^{(3)}(GHZ_3) = 1\,,\ C^{(2)}(GHZ_3) = 2\,. \tag{4.91}$$

4.3.2 Setting and Main Result

We quantify the maximal asymptotic rate R of secret bits that can be encoded and shared by using an infinite number of copies of a given resource state ρ_{ABC}. Let us fix the number of copies $N > 0$. The sender chooses a secret message m from $\mathcal{M}_N = \{1, ..., |\mathcal{M}_N|\}$ and encodes it in the N copies of the tripartite state according to a code-book $\{\rho_m^N\}$. Here, we regard ρ_m^N is a state on a tripartite system in which each subsystem contains N copies of original A, B or C. Each code state ρ_m^N is given by a state of the form $\rho_m^N = U_m \rho_{ABC}^{\otimes N} U_m^\dagger$, which satisfies $\rho_m^N \in \mathcal{C}^{(2)}_{\rho^{\otimes N}}$. The sender then distributes the encoded state ρ_m^N to three receivers associated to (N copies of) A, B and C.

Since the bipartite reduced states of all code states are equal to the one of $\rho_{ABC}^{\otimes N}$, the encoded secret m can be read out only when all three receivers cooperate together. To read the secret, they need to perform a global POVM measurement $\{\Lambda_m^{(N)}\}$. The probability to falsely decode the message m is given by $p^N(m) = \mathrm{Tr}\{(\mathbb{1} - \Lambda_m^{(N)})\rho_m^N\}$, and we denote the maximum error probability by $p_{\max}^N = \max_m p^N(m)$.

We say a secret sharing rate $r(\rho_{ABC})$ for ρ_{ABC} is *achievable* if there exist an appropriate encoding method and a POVM such that $|\mathcal{M}_N| = 2^{N(r(\rho_{ABC})-\delta)}$ and $p_{\max}^N \leq \epsilon$ for any $\delta, \epsilon > 0$ and sufficiently large $N > 0$. The optimal secret sharing rate $r(\rho_{ABC})$ is obtained via the Holevo-Schumacher-Westmoreland theorem [141, 142]:

$$r(\rho_{ABC}) = \lim_{N\to\infty} \frac{1}{N} \left[\max_{\overline{\rho}^N \in \mathcal{C}^{(2)}_{\rho^{\otimes N}}} S\left(\overline{\rho}_{ABC}^N\right) - S(\rho_{ABC}^{\otimes N}) \right], \tag{4.92}$$

where the maximum is over all uniformly distributed ensembles

$$\overline{\rho}_{ABC}^N = \sum_m \frac{1}{\mathcal{M}_N} U_m \rho_{ABC}^{\otimes N} U_m^\dagger \tag{4.93}$$

satisfying $U_m \rho_{ABC}^{\otimes N} U_m^\dagger \in \mathcal{C}^{(2)}_{\rho^{\otimes N}}$ for all $m = 1, ...\mathcal{M}_N$. Here, the uniform distribution is needed to avoid a bias in the choice of the secret message.

Our main result in this section is the equivalence of the 3rd-order irreducible correlation to the optimal secret sharing rate $r(\rho_{ABC})$:

Theorem 11. *Under the setting of Theorem 6, the equality*

$$r(\rho_{ABC}) = C^{(3)}(\rho_{ABC}) \tag{4.94}$$

holds. From Theorem 6, it implies

$$S_{topo} = r(\rho_{ABC})\,. \tag{4.95}$$

4.3.3 Proof: The Equivalence Between TEE and Optimal Secret Sharing Rate

The techniques of the proof are generalization from the proof for the bipartite case in Ref. [143], in which the irreducible correlation is shown to match the mutual information.

Proof. We first prove Theorem 11 for the case of a region with the Levin-Wen type partition. We will then generalize the proof to the other cases. By assumption

and the proof of Theorem 6 in Sec. 4.2, the maximal entropy state $\tilde{\rho}^{(2)}_{ABC}$ is equal to a quantum Markov chain which can be decomposed as

$$\tilde{\rho}^{(2)}_{ABC} = \bigoplus_i p_i \rho_{AB_i^L} \otimes \rho_{B_i^R C} \,, \tag{4.96}$$

associated to decompositions of $\mathcal{H}_{B_1}$ and $\mathcal{H}_{B_2}$.

Consider the spectral decomposition of $\rho_{AB_i^L}$,

$$\rho_{AB_i^L} = \sum_{K_i} \lambda_{K_i} \Pi^{K_i}_{AB_i^L} \,, \tag{4.97}$$

where $\Pi^{K_i}_{AB_i^L}$ is the projector on the eigenspace corresponding to eigenvalue λ_{K_i}. More explicitly, $\Pi^{K_i}_{AB_i^L}$ can written as

$$\Pi^{K_i}_{AB_i^L} = \sum_{m_{K_i}=1}^{d_{K_i}} |K_i, m_{K_i}\rangle\langle K_i, m_{K_i}|_{AB_i^L} \,, \tag{4.98}$$

where $|K_i, m_{K_i}\rangle$ are an orthonormal basis of the eigenspace of λ_{K_i} and d_{K_i} denotes the dimension of the eigenspace. Then we write the state ρ_{ABC} in terms of eigenvectors of $\rho_{AB_i^L}$ to obtain

$$\rho_{ABC} = \sum_{i,K_i,m_{K_i}} \sum_{j,L_j,n_{L_j}} |K_i, m_{K_i}\rangle\langle L_j, n_{L_j}|_{AB^L} \otimes w_{B^R C}^{i,K_i,m_{K_i},j,L_j,n_{L_j}} \,, \tag{4.99}$$

where $\mathcal{H}_{B^L} = \bigoplus_i \mathcal{H}_{B_i^L}$ and $\mathcal{H}_{B^R} = \bigoplus_i \mathcal{H}_{B_i^R}$.

Next, we apply a random unitary $U_{AB^L} \in \mathcal{U}$ of the form

$$U_{AB^L} = \bigoplus_{i,K_i} U^{K_i}_{AB_i^L} \,, \tag{4.100}$$

where for every i and K_i, $U^{K_i}_{AB_i^L}$ are drawn from an exact 1-design of the Haar measure on the eigenspace corresponding to the eigenvalue λ. Note that the cardinality of $\mathcal{U}$ is finite[2]. Clearly, this randomization process does not change the reduced state on AB. According to Schur's lemma, this random unitary operation transforms the state given by Eq. (4.99) to

$$\bar{\rho}_{ABC} = \bigoplus_{i,K_i,m_{K_i}} \Pi^{K_i}_{AB_i^L} \otimes w_{B_i^R C}^{i,K_i,m_{K_i}} \,, \tag{4.101}$$

[2] $\{U^{K_i}_{AB_i^L}\}$ consists of unitaries associated to (i) all possible relative phase flips among the basis states $\{|K_i, m_{K_i}\rangle\}$, (ii) all permutations of the basis states, and (iii) all combinations of (i) and (ii).

where $w_{B_i^R C}^{i,K_i,m_{K_i}} \equiv w_{B^R C}^{i,K_i,m_{K_i},i,K_i,m_{K_i}}$. Since $\tilde{\rho}^{(2)}$ and $\bar{\rho}$ are in $\mathcal{C}_\rho^{(2)}$, we obtain

$$\tilde{\rho}_{AB^L}^{(2)} = \bigoplus_{i,K_i} p_i \lambda_{K_i} \Pi_{AB_i^L}^{K_i} \tag{4.102}$$

$$= \bar{\rho}_{AB^L} \tag{4.103}$$

$$= \bigoplus_{i,K_i} \mathrm{Tr}\left(\sum_{m_{K_i}} w_{B_i^R C}^{i,K_i,m_{K_i}} \right) \Pi_{AB_i^L}^{K_i} . \tag{4.104}$$

Thus, it holds that

$$\mathrm{Tr}\left(\sum_{m_{K_i}} w_{B_i^R C}^{i,K_i,m_{K_i}} \right) = p_i \lambda_{K_i} . \tag{4.105}$$

We denote the normalized operator $\frac{1}{p_i \lambda_{K_i}} \sum_{m_{K_i}} w_{B_i^R C}^{i,K_i,m_{K_i}}$ by $\rho_{B_i^R C}^{K_i}$. Note that $\rho_{B_i^R C} = \sum_{K_i} q_{K_i} \rho_{B_i^R C}^{K_i}$, where $q_{K_i} = \lambda_{K_i} d_{K_i}$, but the states in $\{\rho_{B_i^R C}^{K_i}\}$ are not necessarily orthogonal to each other. Then, $\bar{\rho}_{ABC}$ can be written as

$$\bar{\rho}_{ABC} = \bigoplus_{i,K_i} p_i \lambda_{K_i} \Pi_{AB_i^L}^{K_i} \otimes \rho_{B_i^R C}^{K_i} . \tag{4.106}$$

The difference between $\bar{\rho}_{ABC}$ and $\tilde{\rho}_{ABC}^{(2)}$ is that $\bar{\rho}_{ABC}$ has additional correlations between AB_i^L and $B_i^R C$ via the index K_i.

Summarizing the above calculations, we obtain an ensemble of states

$$\left\{ \frac{1}{|\mathcal{U}|}, U_i \rho_{ABC} U_i^\dagger \in \mathcal{C}_\rho^{(2)} \right\} , \tag{4.107}$$

where the entropy of the averaged state $\bar{\rho}_{ABC}$ is given by

$$S(\bar{\rho}_{ABC}) = H(\{p_i\}) + \sum_i p_i H(\{q_{K_i}\}) + \sum_{i,K_i} p_i q_{K_i} \left(\log d_{K_i} + S(\rho_{B_i^R C}^{K_i}) \right) . \tag{4.108}$$

From Eqs. (4.96) and (4.97), the entropy of $\tilde{\rho}^{(2)}$ is given by

$$\begin{aligned} S(\tilde{\rho}_{ABC}^{(2)}) =& H(\{p_i\}) + \sum_i p_i H(\{q_{K_i}\}) \\ &+ \sum_{i,K_i} p_i q_{K_i} \log d_{K_i} + \sum_i p_i S(\rho_{B_i^R C}) . \end{aligned} \tag{4.109}$$

By taking the difference between Eqs. (4.108) and (4.109), the 3rd-order irreducible correlation of $\bar{\rho}_{ABC}$ can be bounded by

$$C^{(3)}(\bar{\rho}_{ABC}) = S(\tilde{\rho}^{(2)}_{ABC}) - S(\bar{\rho}_{ABC}) \tag{4.110}$$

$$= \sum_i p_i \left[S(\rho_{B_i^R C}) - \sum_{K_i} q_{K_i} S(\rho^{K_i}_{B_i^R C}) \right] \tag{4.111}$$

$$= \sum_i p_i \left[S\left(\sum_{K_i} q_{K_i} \rho^{K_i}_{B_i^R C} \right) - \sum_{K_i} q_{K_i} S(\rho^{K_i}_{B_i^R C}) \right] \tag{4.112}$$

$$\leq \sum_i p_i H(\{q_{K_i}\}) \leq \max_i \log D_i \tag{4.113}$$

$$\leq \log D\,, \tag{4.114}$$

where D_i and D denote the number of different eigenvalues of $\rho_{AB^L_i}$ and ρ_{AB^L}, respectively. If we consider N copies of ρ_{ABC}, D grows only polynomially in N, whereas the total dimension of the Hilbert space grows exponentially. If the dimension of the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_{B^L}$ is denoted by d_{AB^L}, the number of eigenvalues D^N of the N-copy state $\rho^{\otimes N}_{ABC}$ is bounded by [144],

$$D^N \leq (N+1)^{d_{AB^L}}\,. \tag{4.115}$$

Given the expression for the rate

$$r(\rho_{ABC}) = \lim_{N\to\infty} \frac{1}{N} \left[\max_{\bar{\rho}^N \in \mathcal{C}^{(2)}_{\rho^{\otimes N}}} S\left(\bar{\rho}^N_{ABC}\right) - S(\rho^{\otimes N}_{ABC}) \right]\,, \tag{4.116}$$

and using that the irreducible correlation is additive (see Sec. 2.3.1), we obtain

$$r(\rho_{ABC}) = \lim_{N\to\infty} \frac{1}{N} \left[\max_{\bar{\rho}^N \in \mathcal{C}^{(2)}_{\rho^{\otimes N}}} S\left(\bar{\rho}^N_{ABC}\right) - S(\rho^{\otimes N}_{ABC}) \right] \tag{4.117}$$

$$= \lim_{N\to\infty} \frac{1}{N} \left[\max_{\bar{\rho}^N \in \mathcal{C}^{(2)}_{\rho^{\otimes N}}} S\left(\bar{\rho}^N_{ABC}\right) - S(\tilde{\rho}^{(2)\otimes N}) \right] \tag{4.118}$$

$$+ S(\tilde{\rho}^{(2)}_{ABC}) - S(\rho_{ABC}) \tag{4.119}$$

$$= C^{(3)}(\rho_{ABC}) - \lim_{N\to\infty} \frac{1}{N} \max_{\bar{\rho}^N \in \mathcal{C}^{(2)}_{\rho^{\otimes N}}} C^{(3)}\left(\bar{\rho}^N_{ABC}\right) \tag{4.120}$$

$$\geq C^{(3)}(\rho_{ABC}) - \lim_{N\to\infty} \frac{1}{N} \log(N+1)^{d_{AB^L}} \tag{4.121}$$

$$= C^{(3)}(\rho_{ABC})\,. \tag{4.122}$$

This establishes a lower bound on the optimal rate R by $C^{(3)}$. However, the upper bound $r(\rho_{ABC}) \leq C^{(3)}(\rho_{ABC})$ follows directly from Eq. (4.116), and the definition of $C^{(3)}(\rho_{ABC})$. This completes the proof for the Levin-Wen type partition.

In the case of the Kitaev-Preskill type partition, i.e., when the maximum entropy state can be written as Eq. (4.83), we iteratively perform random unitary operations as discussed in the previous case to systems AB and AC. Let us rewrite $\tilde{\rho}^{(2)}_{ABC}$ as

$$\tilde{\rho}^{(2)}_{ABC} = \bigoplus_{a,b,c} p(a,b)p(c|a,b)\rho_{A^R_a B^L_b} \otimes \rho_{B^R_b C^L_c} \otimes \rho_{C^R_c A^L_a}\,, \tag{4.123}$$

where $p(c|a,b) = p(c|a)p(c|b)/p(c)$. We then introduce the spectral decomposition $\rho_{A^R_a B^L_b} = \sum_{K_{ab}} \lambda_{K_{ab}} \Pi^{K_{ab}}_{A^R_a B^L_b}$. Let us define a set of unitaries $\{U_{A^R B^L}\}$ in the same way as in the previous case. Consequently, the averaged state becomes

$$\bar{\rho}_{ABC} = \bigoplus_{a,b,K_{ab}} p(a,b)\lambda_{K_{ab}} \Pi^{K_{ab}}_{A^R_a B^L_b} \otimes \rho^{K_{ab}}_{A^L_a B^R_b C} \tag{4.124}$$

for some state $\rho^{K_{ab}}_{A^L_a B^R_b C}$. We further introduce the spectral decomposition $\rho_{C^R_c A^L_a} = \sum_{L_{ac}} \mu_{L_{ac}} \Pi^{L_{ac}}_{C^R_c A^L_a}$ and a set of unitaries $\{U_{C^R A^L}\}$ similar to $\{U_{A^R B^L}\}$. After performing the second average over the unitaries $\{U_{C^R A^L}\}$, the state can be written as

$$\bar{\bar{\rho}}_{ABC} = \bigoplus_{a,b,c,K_{ab},L_{ac}} p(a,b)p(c|a,b)\lambda_{K_{ab}}\mu_{L_{ac}} \Pi^{K_{ab}}_{A^R_a B^L_b} \otimes \Pi^{L_{ac}}_{C^R_c A^L_a} \otimes \rho^{K_{ab},L_{ac}}_{B^R_b C^L_c}\,. \tag{4.125}$$

Since the remaining correlation in $\bar{\bar{\rho}}_{ABC}$ is also bounded by the logarithm of the number K_{ab}, L_{ac} of different eigenvalues, we can use the same argument as in the case of the Levin-Wen type partition. Therefore, Theorem 11 holds for all situations in which Theorem 6 holds. □

4.3.4 Explicit Encoding for Abelian Models

As mentioned, our assumptions are satisfied for ground states of exactly solvable models, e.g., the toric code model. For the toric code model, we can explicitly demonstrate how to encode secrets to the reduced state on a subregion. It may help to understand the meaning of our result intuitively.

There are two types of string operators, $Z-$string and X-string, in the toric code model. These string operators create anyonic excitations at their endpoints. Performing string operators does not change the reduced states on a region unless the endpoints are outside of the region. Recall that the actions of two Z-string operators on lines C and C' on a ground state are equivalent if one can deform C

to C' by using vertex or face operators. From this property, we can deform a line C overlapping with X to another line C' passing outside of X, namely,

$$\rho_X := \mathrm{Tr}_{X^c}\left[|\psi\rangle\langle\psi|\right] \tag{4.126}$$
$$= \mathrm{Tr}_{X^c}\left[W_{X(Z)}(C')|\psi\rangle\langle\psi|W_{X(Z)}(C')^\dagger\right] \tag{4.127}$$
$$= \mathrm{Tr}_{X^c}\left[W_{X(Z)}(C)|\psi\rangle\langle\psi|W_{X(Z)}(C)^\dagger\right] \tag{4.128}$$

for a state $|\psi\rangle$ on the lattice (Fig. 4.5).

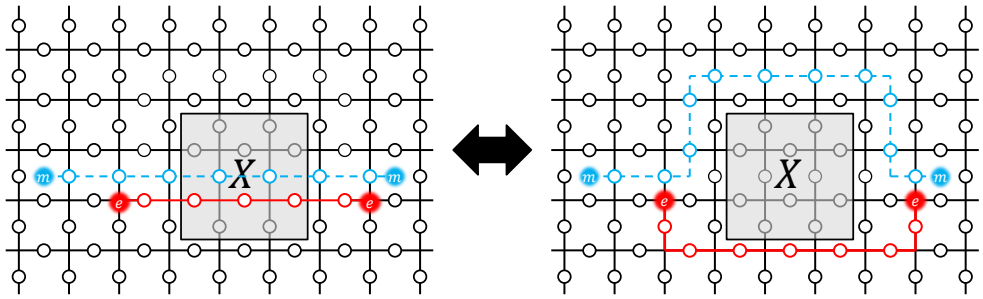

Figure 4.5: Two equivalent states with a different string configurations. Performing Z- and X-string operators do not change ρ_X unless the anyonic excitations are outside of X. By using the deformation property of the string operators, we see that the actions of two string operators in both sides are the same for a ground state of the toric code model.

Let us first consider a disc-like region with the Kitaev-Preskill type partition in the toric code model. For such a region, the TEE is given by

$$S_{topo} = \log_2 \sqrt{4} = 1\,. \tag{4.129}$$

Therefore, we can encode 1 bits of classical information in the region secretly hidden from players knowing reduced states on any two of three subsystems. This is done as follows. For any two dimensional lattice, the three subsystems share either a vertex operator A_v or a face operator Z_p, not the both. When the former holds, we encode a secret by performing Z-string operator $W_Z(C_1)$ as in Fig. 4.6. The string operator does not change the reduced states on any two subsystems, but does change the eigenvalue of the loop operator $W_X(C_2)$ from $+1$ to -1. Therefore, the reduced states on ABC before and after performing $W_Z(C_1)$ are orthogonal to each other. It means that it allows encoding 1 bit of secret. For an annular region, the TEE is given by

$$S_{topo} = 2\,. \tag{4.130}$$

This is performed by encoding secrets by using both X and Z-string operators. Associated to 2 bits of classical secrets, there are four orthogonal states sharing the

same bipartite marginals, where each corresponds to one of four types of anyons (e, m, ϵ and the vacuum). The above discussion can be generalized to all quantum

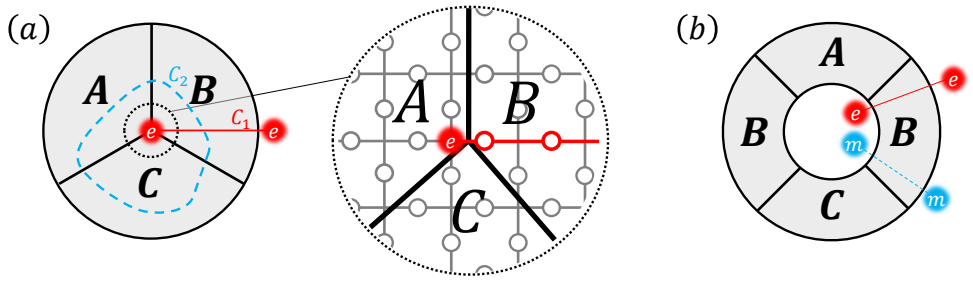

Figure 4.6: A schematic picture of the encoding method for the toric code model. (a): for the Kitaev-Preskill type partition, three subsystem A, B and C share either a vertex (this figure) or a face. We can encode a secret bit $a \in \{0, 1\}$ by performing $(W_Z(C_1))^a$. a can be read out only if we perform a global measurement on ABC, such as measure $W_X(C_2)$. (b): for the Levin-Wen partition, we can use both X-string and Z-string simultaneously, and therefore we can encode 2 bits.

double models exhibiting only abelian anyons. An important property of these models is that string operators are written as tensor products of one-site unitary operators. In other exactly solvable models such as a quantum double model with non-abelian anyons, this property does not hold and we do not know any explicit encoding method achieving the optimal rate. While in abelian models one-copy of the reduced state is sufficient to achieve the optimal rate, in general we need to consider encoding with multiple copies of the reduced state. This is the case for non-abelian models, since the topological entanglement entropy is strictly larger than the logarithm of the number of anyon types. In addition, for Levin-Wen models, we do not know how to create an anyon pair, therefore the explicit encoding method is also unclear.

4.4 Correlation Analysis of Gapped Ground States with Finite Correlation Length

In general, gapped ground states have a non-zero correlation length and the area law is written as in Eq. (3.10). Thus, our assumptions (I) and (II) do not hold exactly. In this section, we extend some of results obtained in Sec. 4.1 to more general cases in which assumptions (I) and (II) are satisfied only approximately. We consider the following modified assumptions,

(I') If two regions A and B are separated, then $I(A : B)_\rho \leq \varepsilon$.

(II') For a simply connected region ABC such that B shields A from C, $I(A : C|B)_\rho \leq \varepsilon$.

Here $\varepsilon > 0$ represents the effect of finite correlation length and vanishes when the characteristic size of regions, which we denote by l, is infinitely large. For simplicity, we further assume that $\varepsilon = e^{-\Theta(l)}$, i.e., the correction term decays exponentially fast with respect to the characteristic scale of regions [3]. We expect that this assumption holds when the system is described by a local commuting Hamiltonian with sufficiently weak perturbations, but no rigorous proof has so far been obtained.

In the case of zero correlation length, we have used two equivalent conditions of Markov chains, a particular decomposition of the state (2.46) and the existence of a recovery map (2.47). However, the first property is known to be fragile, i.e., there exists a tripartite state with small conditional mutual information but is far from any Markov chains in the trace distance [77]. Such a tripartite state cannot be well-approximated by states with the particular decomposition. As discussed in Sec. 2.2.3, the second property survives in the following form for $\varepsilon > 0$,

$$I(A : C|B)_\rho \geq \min_{\Lambda_{B\to BC}} \frac{1}{4\ln(2)} \|\rho_{ABC} - (\mathbb{1}_A \otimes \Lambda_{B\to BC})(\rho_{AB})\|_1^2 . \tag{4.131}$$

We will use this inequality to extend the equivalence between the TEE and the irreducible correlation in Sec. 4.1. This extension is then applied to extend the result in Sec. 4.1.1.

4.4.1 A Smoothed Version of The Irreducible Correlation

To deal with situations with small uncertainty or errors, a "smoothed" version of entropies are used in quantum information theory [145]. We want to define a smoothed version of the irreducible correlation in analogy with this approach.

Let us consider a set of multipartite state where their k-partite reduced states are δ-close to $\rho \in \mathcal{S}(\mathcal{H})$ defined as

$$\mathcal{C}_\rho^{(k),\delta} := \{\sigma \in \mathcal{S}(\mathcal{H}) | \, \|\sigma_{S_k} - \rho_{S_k}\|_1 \leq \delta, \, \forall S_k \subset [n]\} . \tag{4.132}$$

We define the δ variation of the maximum entropy state as

$$\tilde{\rho}^{(k),\delta} := \arg\max_{\sigma \in \mathcal{C}_\rho^{(k,\delta)}} S(\sigma) . \tag{4.133}$$

[3] A function $f(x)$ is said to be in $\Theta(g(x))$ if there exist constants $c, C > 0$ such that $cg(x) \leq f(x) \leq Cg(x)$ for all sufficiently large x.

Note that $\tilde{\rho}^{(k),\delta}$ is uniquely determined since $\mathcal{C}_\rho^{(k),\delta}$ is a closed convex set. The kth-order δ-irreducible correlation is then defined as

$$C_\delta^{(k)}(\rho) := S(\tilde{\rho}^{(k-1),\delta}) - S(\tilde{\rho}^{(k),\delta})\,. \tag{4.134}$$

Recall that the irreducible correlation has another representation in terms of distances from the sets of Gibbs states $\overline{\mathcal{E}}_k^{rI}$. This representation of the irreducible correlation is used to obtain Theorem 9 . To extend Theorem 9 to finite correlation length cases, we consider a modification of the distance-like measure $D^{(k)}(\rho)$ defined in Eq. (2.68) by inserting a cut-off on the set of Hamiltonians $\mathscr{H}_{nn}$. Let us define the set of nearest-neighbor Hamiltonians with the interaction strength less than K by

$$\mathscr{H}_{nn}^K := \left\{ H = \sum_{S_k \subset [n]} h_{S_k} \otimes \mathbb{1}_{S_k^c},\ \|h_{S_k}\| \leq K \right\}\,. \tag{4.135}$$

We are interested in the distance from all Gibbs states of Hamiltonians in $\mathscr{H}_{nn}^K$ by

$$D^{nn,K}(\rho) := \min_{H \in \mathscr{H}_k^K} S(\rho \| e^{-H})\,. \tag{4.136}$$

(Here, we include the normalization factor in the Hamiltonian).

4.4.2 Extensions of Results for States with Finite Correlation Length

In Sec. 4.1, we prove that the TEE is equivalent to the 3rd-order irreducible correlation for a certain tripartition of an annular region (Theorem 6). Moreover, when the number of partitions increases, the TEE is equivalent to the highest-order irreducible correlation (Theorem 4.6). We consider a similar setting with the Levin-Wen type partitions, in which the TEE S_{topo} is approximately the conditional mutual information of certain regions. A key point is that the original proof in Sec. 4.2 only exploits the existence of recovery maps of Markov chains, not the direct sum decomposition of Markov chains.

Let us consider an annular region in a 2D spin lattice as depicted in Fig. 4.7(a). For the δ-irreducible correlation, we have the following theorem.

Theorem 12. *If assumptions (I') and (II') are satisfied, for a region as depicted in Fig. 4.7(a), it holds that*

$$\left| C_\delta^{(3)}(\rho) - S_{topo} \right| \leq e^{-\Theta(l)}\,, \tag{4.137}$$

for $\delta = e^{-\Theta(l)}$.

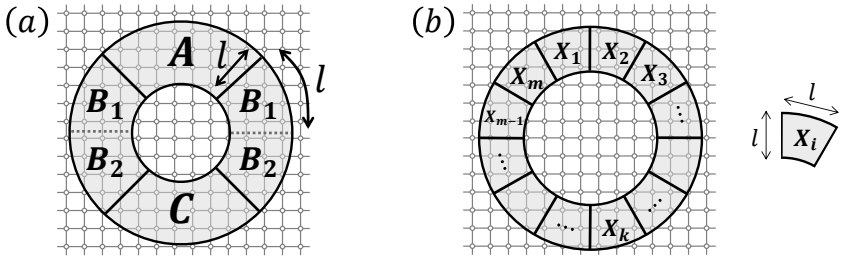

Figure 4.7: Annular regions with appropriate partitions for the proof. (a): we consider a region with the Levin-Wen type partition. The size of each subsystem is characterized by l. (b): a finer partition with sufficiently large subsystems. Regions as in Fig. 3.5(a) is not considered here, since short-range correlations around the center point makes the problem more involved.

Although $C_\delta^{(3)}(\rho)$ has no geometrical meaning yet, we can still assign a distance-like meaning to the TEE. Let us consider a region depicted as in Fig. 4.7(b) and denote the number of spins in the whole region by n. We divide the region into m pieces, where each piece has a sufficiently large width and height characterized by $l = \mathcal{O}(\ln n)$. The following theorem states that the distance from the Gibbs states of Hamiltonians in $\mathscr{H}_{nn}^K$ is approximately characterized by the TEE.

Theorem 13. *If assumptions (I') and (II') are satisfied,*

$$\left|D^{nn,K}(\rho) - S_{topo}\right| \leq e^{-\Theta(l)} \tag{4.138}$$

for $K = \Theta(n)$.

Theorem 13 implies that when the correction terms in the area law decays sufficiently fast, the non-zero TEE implies the entanglement Hamiltonian $-\ln \rho_X$ contains a non-local interaction. Next, we investigate which types of many-body interactions are contained in $-\ln \rho_X$. To address this question, let us set $A \equiv X_1$, $B \equiv X_2X_3X_{m-1}X_m$ and C as the remaining subsystems. In a similar way to Theorem 13, we can show that

$$\left|D^{(2),K}(\rho) - S_{topo}\right| \leq e^{-\Theta(l)}\,, \tag{4.139}$$

where $D^{2,K}(\rho)$ is the set of all Gibbs states of 2-local Hamiltonians with the interaction length less than K. 2-local Hamiltonians in ABC contains at most $(m-1)$ interaction in terms of $X_1 ... X_m$. Therefore, Eq. (4.139) implies that adding $(m-1)$ interactions cannot improve the minimization in Eq. (4.136). This fact suggests that the non-local part in the entanglement Hamiltonian is dominated by genuine m-body interactions as in the case of the zero correlation length.

Finally, we consider an extension of Theorem 9. We introduce a cut-off on the spectrum by

$$\lambda^{\Lambda}(A) := \{\lambda \in \lambda(A) |\, \lambda \leq \ln \Lambda\} \ . \tag{4.140}$$

Let us consider a pure state on a 2D cylinder as depicted by Fig. 4.3. Then, the extended theorem is stated as follows:

Theorem 14. *Suppose that a pure state $\rho = |\psi\rangle\langle\psi|_{YXY'}$ defined on the cylinder in Fig. 4.3 satisfies assumption (I') and (II'). We also assume the reflection symmetry*

$$\rho_Y = \rho_{Y'} \ . \tag{4.141}$$

Then, there exists a Hamiltonian H_X on $X = X_1X_2...X_m$ such that for any $\Lambda > 0$,

$$\left\| \lambda^{\Lambda}\left(H^{(2)}_{\rho_Y}\right) - \lambda^{\Lambda}\left(H_X\right) \right\|_1 \leq \Lambda e^{-\Theta(l)} \ . \tag{4.142}$$

Moreover, H_X satisfies

$$\min_{H'_X \in \mathscr{H}^K_{nn}} |\|H_X - H'_X\| - S_{topo}| \leq e^{-\Theta(l)} \ . \tag{4.143}$$

4.4.3 Proof of Theorem 12

The strategy of the proof is as follows. We first construct a global state on ABC from the reduced state of ρ_{ABC} by using recovery maps. Owing to the Fawzi-Renner bound [78], the constructed state has bipartite reduced states which are δ-close to the marginals of ρ_{ABC}. We then show that this state has a sufficiently high entropy with which we bound the entropy of the maximum entropy state in $\mathcal{C}^{(2),\delta}_{\rho}$ and $C^{(3)}_{\delta}(\rho)$.

By assumption (I') and (II'), it holds that

$$I(A : B_2C)_{\rho} \leq \varepsilon \tag{4.144}$$

and

$$I(A : B_2|B_1)_{\rho} \leq \varepsilon, \tag{4.145}$$

$$I(B_1 : C|B_2)_{\rho} \leq \varepsilon \ . \tag{4.146}$$

By using the Pinsker inequality (2.26), the first assumption (4.144) implies that

$$\|\rho_{AB_2C} - \rho_A \otimes \rho_{B_2C}\|_1 \leq 2\sqrt{\varepsilon} \ . \tag{4.147}$$

By using the monotonicity of the trace norm, this also implies that $\|\rho_{AB_2} - \rho_A \otimes \rho_{B_2}\|_1 \leq 2\sqrt{\varepsilon}$ and $\|\rho_{AC} - \rho_A \otimes \rho_C\|_1 \leq 2\sqrt{\varepsilon}$.

From the Fawzi-Renner bound (4.131), the assumptions represented by (4.145) and (4.146) imply that there exist CPTP maps $\Lambda_{B_1\to AB_1}$ and $\Lambda_{B_2\to B_2C}$ such that

$$\|\rho_{AB_1B_2} - \Lambda_{B_1\to AB_1}(\rho_{B_1B_2})\|_1 \leq 2\sqrt{\varepsilon}\,, \tag{4.148}$$

$$\|\rho_{B_1B_2C} - \Lambda_{B_2\to B_2C}(\rho_{B_1B_2})\|_1 \leq 2\sqrt{\varepsilon}\,, \tag{4.149}$$

where we omitted the identity maps for simplicity and used $\sqrt{\ln(2)} \leq 1$. Similar to the proof of Theorem 6 in Sec. 4.2.1, we define a global state $\tilde{\rho}_{ABC}$ by

$$\tilde{\rho}_{ABC} := \Lambda_{B_2\to B_2C} \circ \Lambda_{B_1\to AB_1}(\rho_{B_1B_2})\,. \tag{4.150}$$

We will show that the bipartite reduced states of $\tilde{\rho}_{ABC}$ are close to the reduced states of ρ. By tracing out system A in Eq. (4.150), we obtain

$$\tilde{\rho}_{B_1B_2C} = \Lambda_{B_2\to B_2C} \circ \mathrm{Tr}_A\left(\Lambda_{B_1\to AB_1}(\rho_{B_1B_2})\right)\,. \tag{4.151}$$

From Eq. (4.148) and the monotonicity of the trace norm under CPTP-maps, we have

$$\begin{aligned}
&\|\Lambda_{B_2\to B_2C} \circ \mathrm{Tr}_A\left(\Lambda_{B_1\to AB_1}(\rho_{B_1B_2})\right) - \Lambda_{B_2\to B_2C}(\rho_{B_1B_2})\|_1 \\
&\leq \|\Lambda_{B_1\to AB_1}(\rho_{B_1B_2}) - \rho_{B_1B_2C}\|_1 &(4.152)\\
&\leq 2\sqrt{\varepsilon}\,. &(4.153)
\end{aligned}$$

We used $\rho_{B_1B_2} = \mathrm{Tr}_A\rho_{AB_1B_2}$ in the second line. Therefore, from Eq. (4.149), we obtain

$$\|\rho_{B_1B_2C} - \tilde{\rho}_{B_1B_2C}\|_1 \leq \|\rho_{B_1B_2C} - \Lambda_{B_2\to B_2C}(\rho_{B_1B_2})\|_1 + \|\Lambda_{B_2\to B_2C}(\rho_{B_1B_2}) - \tilde{\rho}_{B_1B_2C}\|_1 \tag{4.154}$$

$$\leq 4\sqrt{\varepsilon}\,. \tag{4.155}$$

Similarly, we have

$$\|\rho_{AB_1B_2} - \tilde{\rho}_{AB_1B_2}\|_1 \leq 4\sqrt{\varepsilon}\,. \tag{4.156}$$

Finally, we show that $\tilde{\rho}_{AC}$ is close to ρ_{AC}. From Eq. (4.148), it holds that

$$\|\tilde{\rho}_{AC} - \rho_{AC}\|_1 \leq \|\tilde{\rho}_{AC} - \mathrm{Tr}_B\Lambda_{B_2\to B_2C}(\rho_{AB_1B_2})\|_1 + \|\mathrm{Tr}_B\Lambda_{B_2\to B_2C}(\rho_{AB_1B_2}) - \rho_{AC}\|_1 \tag{4.157}$$

$$\leq 2\sqrt{\varepsilon} + \|\mathrm{Tr}_{B_2}\Lambda_{B_2\to B_2C}(\rho_{AB_2}) - \rho_{AC}\|_1\,. \tag{4.158}$$

Equations (4.147) and (4.148) yield

$$\begin{aligned}
\|\mathrm{Tr}_{B_2}\Lambda_{B_2\to B_2C}(\rho_{AB_2}) - \rho_{AC}\|_1 &\leq \|\rho_A \otimes \mathrm{Tr}_{B_2}\Lambda_{B_2\to B_2C}(\rho_{B_2}) - \rho_A\otimes\rho_C\|_1 \\
&\quad + 4\sqrt{\varepsilon} &(4.159)\\
&\leq \|\mathrm{Tr}_B\Lambda_{B_2\to B_2C}(\rho_B) - \mathrm{Tr}_B\rho_{BC}\|_1 \\
&\quad + 4\sqrt{\varepsilon} &(4.160)\\
&\leq 6\sqrt{\varepsilon}\,, &(4.161)
\end{aligned}$$

where the last line follows from Eq. (4.149) and the monotonicity of the trace norm under Tr_B. Inserting Eq. (4.161) to Eq. (4.158) implies

$$\|\rho_{AC} - \tilde{\rho}_{AC}\| \leq 8\sqrt{\varepsilon}\,. \tag{4.162}$$

and therefore $\tilde{\rho}_{ABC} \in \mathcal{C}_\rho^{(2),\delta}$ for $\delta = 8\sqrt{\varepsilon}$.

Although $\tilde{\rho}_{ABC}$ is not the maximum entropy state in $\mathcal{C}_\rho^{(2),\delta}$, we can obtain appropriate bounds of $C_\delta^{(3)}$ from this state. To do so, we first estimate the conditional mutual information of $\tilde{\rho}_{ABC}$. By definition,

$$\begin{aligned}
\|\tilde{\rho}_{ABC} - \Lambda_{B_2\to B_2C}(\tilde{\rho}_{AB})\|_1 &\leq \|\Lambda_{B_1\to B_1A}(\rho_{B_1B_2}) - \tilde{\rho}_{AB}\|_1 && (4.163)\\
&\leq \|\rho_{B_1B_2} - \mathrm{Tr}_C\Lambda_{B_2\to B_2C}(\rho_{B_1B_2})\|_1 && (4.164)\\
&\leq 2\sqrt{\varepsilon}\,, && (4.165)
\end{aligned}$$

thus $\tilde{\rho}_{ABC}$ can be approximately recovered from $\tilde{\rho}_{AB}$ by $\Lambda_{B_2\to B_2C}$. For $\tilde{\rho}_{ABC}$ and $\tilde{\rho}' := \Lambda_{B_2\to B_2C}(\tilde{\rho}_{AB})$, we have

$$I(A:C|B)_{\tilde{\rho}} \leq S(A|B)_{\tilde{\rho}} - S(A|BC)_{\tilde{\rho}} \leq S(A|BC)_{\tilde{\rho}} - S(A|BC)_{\tilde{\rho}'}\,. \tag{4.166}$$

By using the Fannes inequality (2.23), it implies that

$$I(A:C|B)_{\tilde{\rho}} \leq 6\sqrt{2}|A|\varepsilon^{\frac{1}{4}} = 3|A|\sqrt{\delta} \tag{4.167}$$

for $2\sqrt{\varepsilon} \leq \frac{1}{2}$, where $|A| = \log_2(\dim \mathcal{H}_A)$.

From the triangle inequality, $\tilde{\rho}_{ABC} \in \mathcal{C}_\sigma^{(2),2\delta}$ holds for any state (including the maximum entropy state) $\sigma_{ABC} \in \mathcal{C}_\rho^{(2),\delta}$. Therefore, from the Fannes inequality, we have

$$\begin{aligned}
S(ABC)_\sigma &\leq S(A|B)_\sigma + S(BC)_\sigma && (4.168)\\
&\leq S(A|B)_{\tilde{\rho}} + S(BC)_{\tilde{\rho}} + 6|A|\sqrt{2\delta} + 6|BC|\sqrt{2\delta} && (4.169)\\
&= S(A|B)_{\tilde{\rho}} + S(BC)_{\tilde{\rho}} + 6|ABC|\sqrt{2\delta} && (4.170)\\
&= S(ABC)_{\tilde{\rho}} + 6|ABC|\sqrt{2\delta} + I(A:C|B)_{\tilde{\rho}} && (4.171)\\
&\leq S(ABC)_{\tilde{\rho}} + 6|ABC|\sqrt{2\delta} + 3|A|\sqrt{\delta} && (4.172)
\end{aligned}$$

for $\sqrt{2\delta} \leq 1$. It implies that

$$\begin{aligned}
C_\delta^{(3)}(\rho) &\leq S(A|B)_{\tilde{\rho}} + S(BC)_{\tilde{\rho}} + 6|ABC|\sqrt{2\delta} - S(ABC)_\rho && (4.173)\\
&\leq S(A|B)_{\tilde{\rho}} - S(A|B)_\rho + S(BC)_{\tilde{\rho}} - S(BC)_\rho\\
&\quad + I(A:C|B)_\rho + 6|ABC|\sqrt{2\delta} && (4.174)\\
&\leq I(A:C|B)_\rho + 6|ABC|\sqrt{\delta} + 6|ABC|\sqrt{2\delta} && (4.175)\\
&= I(A:C|B)_\rho + 12|ABC|\sqrt{2\delta}\,. && (4.176)
\end{aligned}$$

where we used Eq. (4.171) for $\sigma = \tilde{\rho}^{(2),\delta}$ in the first line and again used the Fannes inequality in the third line.

Eq. (4.168)≤Eq. (4.172) for $\sigma = \rho$ yields

$$S(ABC)_{\tilde{\rho}} \geq S(A|B)_\rho + S(BC)_\rho - 9\sqrt{2}|ABC|\sqrt{\delta}\,. \tag{4.177}$$

Since the maximum entropy state in $\mathcal{C}_\rho^{(2),\delta}$ has a entropy greater than or equal to that of $\tilde{\rho}_{ABC}$, we have

$$\begin{aligned} C_\delta^{(3)}(\rho) &\geq S(ABC)_{\tilde{\rho}} - S(ABC)_\rho && (4.178)\\ &\geq S(A|B)_\rho + S(BC)_\rho - S(ABC)_\rho - 9|ABC|\sqrt{2\delta} && (4.179)\\ &= I(A:C|B)_\rho - 9|ABC|\sqrt{2\delta}\,. && (4.180) \end{aligned}$$

where we used strong subadditivity in the second line. Hence, we conclude that

$$\left| C_\delta^{(3)}(\rho) - I(A:C|B)_\rho \right| \leq 12|ABC|\sqrt{2\delta} \tag{4.181}$$

for sufficiently small δ. By assumption, $|ABC| = n\ln(d)$ and $\sqrt{\delta} = e^{-\Theta(l)}$. Therefore, we can set the right hand side of Eq. (4.181) in $e^{-\Theta(l)}$ by choosing $l = \Theta(\ln(n))$. By assumption (I), we have

$$S_{topo} - I(A:C|B)_\rho = I(A:C)_\rho \leq \varepsilon \tag{4.182}$$

and thus complete the proof.

4.4.4 Proof of Theorem 13

The strategy of the proof is as follows. We first consider $\tilde{\rho}_{ABC}$ introduced in the proof of Theorem 12. We then introduce a modification on $\tilde{\rho}_{ABC}$ and then make a Gibbs state of a nearest-neighbor Hamiltonian from its reduced states. This Gibbs state gives a bound on the distance from the set of Gibbs state of bounded nearest-neighbor Hamiltonians.

Let $A \equiv X_1$, $B_1 \equiv X_m X_2$, $B_2 \equiv X_{m-1}X_3$ and $C \equiv X_4 X_5 ... X_{m-2}$. We also set $X_1 \equiv X_{m+1}$. Without loss of generality, we assume that $|X_1| = \max_i |X_i|$. In the following, we use both notations $X_1....X_m$ and ABC interchangeably. Define a full-rank modification of $\tilde{\rho}_{ABC}$ introduced in Eq. (4.150), that is,

$$\tilde{\rho}_{ABC}^f := \left(1 - \frac{1}{2^{n-1}}\right)\tilde{\rho}_{ABC} + \frac{1}{2^{n-1}}\tau_{ABC}\,, \tag{4.183}$$

where τ_{ABC} is the completely mixed state on ABC. Since by definition it holds that

$$\|\tilde{\rho}_{ABC}^f - \tilde{\rho}_{ABC}\|_1 \leq 2^{-n}\,, \tag{4.184}$$

the Fannes inequality implies

$$I(A:C|B)_{\tilde{\rho}^f} \leq I(A:C|B)_{\tilde{\rho}} + \frac{12|A|}{\sqrt{2^n}} \tag{4.185}$$

$$\leq 3|A|\sqrt{\delta} + \frac{12|A|}{\sqrt{2^n}}\,. \tag{4.186}$$

We consider a Gibbs state $\tilde{\pi}_X$ defined as

$$\tilde{\pi}_X := \frac{1}{Z} e^{-H_X^{\tilde{\rho}}}\,, \tag{4.187}$$

where Z is the normalizer and

$$H_X^{\tilde{\rho}} := \sum_{i=1}^{m} \left(\ln \tilde{\rho}^f_{X_i} - \ln \tilde{\rho}^f_{X_i X_{i+1}} \right)\,. \tag{4.188}$$

By definition, $\tilde{\pi}_X$ is an element of $\mathscr{H}_{nn}^K$ with $K = \Theta(n)$.

By definition of $H_X^{\tilde{\rho}}$and the definition of the relative entropy, it holds that

$$S\left(\tilde{\rho}_X \left\| e^{-H_X^{\tilde{\rho}}} \right.\right) = \sum_{i=1}^{m} S(X_{i+1}|X_i)_{\tilde{\rho}^f} - S(X_1 X_2 ... X_m)_{\tilde{\rho}^f}\,. \tag{4.189}$$

We use an iterative calculation:

$$\sum_{i=1}^{m} S(X_{i+1}|X_i)_{\tilde{\rho}^f} = \sum_{i=1}^{m} S(X_i|X_{i+1})_{\tilde{\rho}^f} \tag{4.190}$$

$$= S(X_1|X_2)_{\tilde{\rho}^f} + S(X_2|X_3)_{\tilde{\rho}^f} + S(X_3|X_4)_{\tilde{\rho}^f} + \sum_{i=4}^{m} S(X_i|X_{i+1})_{\tilde{\rho}^f} \tag{4.191}$$

$$= S(X_1|X_2)_{\tilde{\rho}^f} + I(X_2:X_4|X_3)_{\tilde{\rho}^f} + S(X_2X_3|X_4)_{\tilde{\rho}^f} + \sum_{i=4}^{m} S(X_i|X_{i+1})_{\tilde{\rho}^f} \tag{4.192}$$

$$= S(X_1|X_2)_{\tilde{\rho}^f} + I(X_2:X_4|X_3)_{\tilde{\rho}^f} + S(X_2X_3|X_4)_{\tilde{\rho}^f} + S(X_4|X_5)_{\tilde{\rho}^f} + \sum_{i=5}^{m} S(X_i|X_{i+1})_{\tilde{\rho}^f} \tag{4.193}$$

$$= S(X_1|X_2)_{\tilde{\rho}^f} + I(X_2:X_4|X_3)_{\tilde{\rho}^f} + I(X_2X_3:X_5|X_4)_{\tilde{\rho}^f} + S(X_2X_3X_4|X_5)_{\tilde{\rho}^f} + \sum_{i=5}^{m} S(X_i|X_{i+1})_{\tilde{\rho}^f} \tag{4.194}$$

$$
\begin{aligned}
&\vdots \\
&= S(X_1|X_2)_{\tilde{\rho}^f} + S(X_2 \dots X_{m-1}|X_m)_{\tilde{\rho}^f} + S(X_m|X_1)_{\tilde{\rho}^f} \\
&\quad + \sum_{i=3}^{m-1} I(X_2 ... X_{i-1} : X_{i+1}|X_i)_{\tilde{\rho}^f} \qquad (4.195) \\
&= S(X_2 \dots X_{m-1}|X_m)_{\tilde{\rho}^f} - S(X_2)_{\tilde{\rho}^f} + S(X_m X_1 X_2)_{\tilde{\rho}^f} \\
&\quad + I(X_m : X_2|X_1)_{\tilde{\rho}^f} + \sum_{i=3}^{m-1} I(X_2 ... X_{i-1} : X_{i+1}|X_i)_{\tilde{\rho}^f} \qquad (4.196)
\end{aligned}
$$

By using the subadditivity $S(X_2X_m) \leq S(X_2) + S(X_m)$, we obtain that

$$
(4.196) \leq I(X_1 : X_3 \dots X_{m-1}|X_2X_m)_{\tilde{\rho}^f} + I(X_m : X_2|X_1)_{\tilde{\rho}^f} + \sum_{i=3}^{m-1} I(X_2 ... X_{i-1} : X_{i+1}|X_i)_{\tilde{\rho}^f} . \qquad (4.197)
$$

Therefore, we have

$$
\begin{aligned}
S\left(\tilde{\rho}_X \middle\| e^{-H_X^{\tilde{\rho}}}\right) \leq\ & I(X_m : X_2|X_1)_{\tilde{\rho}^f} + \sum_{i=2}^{m-2} I(X_2 ... X_i : X_{i+2}|X_{i+1})_{\tilde{\rho}^f} \\
& + I(A : B_2|B_1)_{\tilde{\rho}^f} + I(A : C|B_1B_2)_{\tilde{\rho}^f} , \qquad (4.198)
\end{aligned}
$$

where we used the chain rule $I(A : B_2C|B_1)_{\tilde{\rho}^f} = I(A : B_2|B_1)_{\tilde{\rho}^f} + I(A : C|B_1B_2)_{\tilde{\rho}^f}$. Recall that $\tilde{\rho} \in \mathcal{C}_\rho^{(2),\delta}$, where $\delta = 8\sqrt{\varepsilon}$. From Eq. (4.184), $\tilde{\rho}^f \in \mathcal{C}_\rho^{(2),\delta'}$ for $\delta' = \delta + 2^{-n}$. Therefore, $\tilde{\rho}^f$'s reduced states on $AB_1 = X_mX_1X_2$ and $BC = X_2 \dots X_m$ are δ'-close to ρ's reduced states. Combining the Fannes inequality, Eq. (4.186) and the δ'-closeness imply that

$$
\begin{aligned}
S\left(\tilde{\rho}_X \middle\| e^{-H_X^{\tilde{\rho}}}\right) &\leq 12 \left(|X_2| + \sum_{i=2}^{m-2} |X_{i+2}| \right) \sqrt{\delta'} + (m-1)\varepsilon + I(A : C|B_1B_2)_{\tilde{\rho}^f} \qquad (4.199) \\
&\leq 12(m-2)|X_1|\sqrt{\delta'} + (m-1)\varepsilon + 3|X_1|\sqrt{\delta} + \frac{12|X_1|}{\sqrt{2^n}} \qquad (4.200) \\
&= \mathcal{O}\left(n\sqrt{\delta'}\right) , \qquad (4.201)
\end{aligned}
$$

where we used $|X_i| \leq |X_1|$, $m|X_1| = \Theta(n)$ and the asymptotic notation $\mathcal{O}(n\sqrt{\delta'})$ as $n \to \infty$ and $\varepsilon \to 0$ (and therefore $\delta' \to 0$).

Let us estimate the partition function $Z = \mathrm{Tr} e^{-H_X^{\tilde{\rho}}}$. By the Pinsker inequality, the above upper bound (4.201) implies

$$\left\| \tilde{\rho}_X - e^{-H_X^{\tilde{\rho}}} \right\|_1 \leq \sqrt{2S\left(\tilde{\rho}_X \left\| e^{-H_X^{\tilde{\rho}}} \right.\right)} \tag{4.202}$$

$$\leq \mathcal{O}\left(n^{\frac{1}{2}} \delta'^{\frac{1}{4}}\right) . \tag{4.203}$$

We obtain

$$\left| \mathrm{Tr} e^{-H_X^{\tilde{\rho}}} - 1 \right| = \left| \left\| e^{-H_X^{\tilde{\rho}}} \right\|_1 - \|\tilde{\rho}_X\|_1 \right| \tag{4.204}$$

$$\leq \mathcal{O}\left(n^{\frac{1}{2}} \delta'^{\frac{1}{4}}\right) \tag{4.205}$$

and therefore

$$|\log_2 Z| \leq \mathcal{O}\left(n^{\frac{1}{2}} \delta'^{\frac{1}{4}}\right) , \tag{4.206}$$

where we used the inequality $|\|A\|_1 - \|B\|_1| \leq \|A - B\|_1$. Thus, the difference between $\tilde{\rho}_X^f$ and $\tilde{\pi}_X$ is bounded as

$$\left\| \tilde{\rho}_X^f - \tilde{\pi}_X \right\|_1 \leq \sqrt{2S(\tilde{\rho}_X \| \tilde{\pi}_X)} \tag{4.207}$$

$$= \sqrt{2\left(S\left(\tilde{\rho}_X \left\| e^{-H_X^{\tilde{\rho}}} \right.\right) - \log_2 Z\right)} \tag{4.208}$$

$$\leq \mathcal{O}\left(n^{\frac{1}{2}} \delta'^{\frac{1}{8}}\right) . \tag{4.209}$$

The conditional mutual information of $\tilde{\pi}$ is calculated as

$$I(A:C|B)_{\tilde{\pi}} \leq \mathcal{O}\left(|A| n^{\frac{1}{4}} \delta'^{\frac{1}{16}}\right) . \tag{4.210}$$

So far, we have shown that the Gibbs state $\tilde{\pi}_X$ is close to $\tilde{\rho}_X$. Since $\tilde{\rho}^f \in \mathcal{C}_\rho^{(2),\delta'}$, we obtain that

$$\|\rho_{AB} - \tilde{\pi}_{AB}\|_1 \leq \|\rho_{AB} - \tilde{\rho}_{AB}^f\|_1 + \left\| \tilde{\rho}_{AB}^f - \tilde{\pi}_{AB} \right\|_1 \tag{4.211}$$

$$\leq \mathcal{O}\left(n^{\frac{1}{2}} \delta'^{\frac{1}{8}}\right) . \tag{4.212}$$

In a similar way, it can be shown that

$$\|\rho_{BC} - \tilde{\pi}_{BC}\|_1 \leq \mathcal{O}\left(n^{\frac{1}{2}} \delta'^{\frac{1}{8}}\right) . \tag{4.213}$$

Next, we show that $I(A:C|B)_\rho$ approximates the distance between ρ and $\tilde{\pi}$ in terms of the relative entropy. Since $\tilde{\pi}$ and ρ are close on AB and BC, the Fannes inequality yields

$$|S(A|B)_\rho - S(A|B)_{\tilde{\pi}}| \leq \mathcal{O}\left(|A| n^{\frac{1}{4}} \delta'^{\frac{1}{16}}\right) \tag{4.214}$$

and also

$$|S(BC)_\rho - S(BC)_{\tilde{\pi}}| \leq \mathcal{O}\left(|BC| n^{\frac{1}{4}} \delta'^{\frac{1}{16}}\right) . \tag{4.215}$$

These inequalities imply

$$\begin{aligned}
&|I(A:C|B)_\rho - (S(\tilde{\pi}_{ABC}) - S(\rho_{ABC}))| \\
&\quad = |S(A|B)_\rho + S(BC)_\rho - S(ABC)_{\tilde{\pi}}| && (4.216)\\
&\quad \leq |S(A|B)_\rho - S(A|B)_{\tilde{\pi}}| + |S(BC)_\rho - S(BC)_{\tilde{\pi}}| + I(A:C|B)_{\tilde{\pi}} && (4.217)\\
&\quad \leq \mathcal{O}\left(n^{\frac{5}{4}} \delta'^{\frac{1}{16}}\right) , && (4.218)
\end{aligned}$$

where we used the triangle inequality in the second line, and used Eq. (4.210) in the last line. By definition of the distance measure $\mathcal{D}^{nn,K}$, it holds that

$$\begin{aligned}
\mathcal{D}^{nn,K}(\rho) &= \min_{\mu \in \mathcal{E}^K_{nn}} S\left(\rho_{ABC} \,\|\mu_{ABC}\right) && (4.219)\\
&= \min_{\mu \in \mathcal{E}^K_{nn}} S(\rho_{ABC}) - \mathrm{Tr}(\rho \log_2 \mu) . && (4.220)
\end{aligned}$$

Here, $\log_2 \mu \propto H_{AB} + H_{BC}$ for some bounded Hermitian operators H_{AB} and H_{BC} satisfying $\|H_{AB}\| + \|H_{BC}\| \leq \mathcal{O}(mK)$. Equations (4.212) and (4.213) yield that

$$\begin{aligned}
|\mathrm{Tr}(\rho_Y O_Y) - \mathrm{Tr}(\tilde{\pi}_Y O_Y)| &\leq \|O\| \|\rho_Y - \tilde{\pi}_Y\| && (4.221)\\
&\leq \mathcal{O}\left(\|O\| n^{\frac{1}{2}} \delta'^{\frac{1}{8}}\right) && (4.222)
\end{aligned}$$

for $Y = AB, BC$. Therefore, we have

$$\begin{aligned}
\mathcal{D}^{nn,K}(\rho) &= \min_{\mu \in \mathcal{E}^K_{nn}} S(ABC)_\rho - \mathrm{Tr}(\tilde{\pi} \log_2 \mu) + \mathcal{O}\left(mK n^{\frac{1}{2}} \delta^{\frac{1}{8}}\right) && (4.223)\\
&= \min_{\mu \in \mathcal{E}^K_{nn}} S(ABC)_{\tilde{\pi}} - \mathrm{Tr}(\tilde{\pi} \log_2 \mu) + I(A:C|B)_\rho \\
&\quad + \mathcal{O}\left(n^{\frac{5}{4}} \delta'^{\frac{1}{16}}\right) + \mathcal{O}\left(n^{\frac{5}{2}} \delta^{\frac{1}{8}}\right) && (4.224)\\
&= \min_{\mu \in \mathcal{E}^K_{nn}} S\left(\tilde{\pi}_{ABC} \,\|\mu_{ABC}\right) + I(A:C|B)_\rho + \mathcal{O}\left(n^{\frac{5}{2}} \delta'^{\frac{1}{16}}\right) && (4.225)\\
&= I(A:C|B)_\rho + \mathcal{O}\left(n^{\frac{5}{2}} \delta'^{\frac{1}{16}}\right) , && (4.226)
\end{aligned}$$

where we used Eq. (4.218) and $mK = \mathcal{O}(n^2)$ in the second line. The third equality follows from Eq. (4.218) and the last line follows from $\pi_{ABC} \in \mathcal{E}^K_{nn}$. We know that $I(A:C|B)_\rho = S_{topo} + e^{-\Theta(l)}$ by choosing suitable $l = \Theta(\log_2 n)$ so that $n^{\frac{5}{2}} e^{-\Theta(l)} = e^{-\Theta(l)}$. Thus, we complete the proof by using that δ' is $e^{-\Theta(l)}$.

4.4.5 Proof of Theorem 14

We prove Theorem 14 in this subsection.

Proof. Since $|\psi_{YXY'}\rangle$ is pure, it holds that

$$\lambda(\rho_{YY'}) = \lambda(\rho_X)\,. \tag{4.227}$$

Since we choose $\varepsilon = e^{-\Theta(l)}$, assumption (I') implies that

$$I(Y:Y')_\rho \leq e^{-\Theta(l)}\,. \tag{4.228}$$

Therefore, by Pinsker inequality and the reflection symmetry, we obtain that

$$\|\rho_{YY'} - \rho_Y^{\otimes 2}\|_1 \leq e^{-\Theta(l)}\,. \tag{4.229}$$

For bounded Hermitian operators A and B, the difference of their spectrum is bounded as

$$\|\lambda(A) - \lambda(B)\|_1 \leq \|A - B\|_1\,. \tag{4.230}$$

Therefore, we obtain

$$\left\|\lambda(\rho_Y^{\otimes 2}) - \lambda(\rho_X)\right\|_1 \leq e^{-\Theta(l)}\,. \tag{4.231}$$

From Eq. (4.230), Theorem 13 implies that there exists a Hamiltonian H_X such that

$$\left\|\lambda(\rho_Y) - \lambda\left(e^{-H_X}\right)\right\|_1 \leq e^{-\Theta(l)}\,. \tag{4.232}$$

By Theorem 13, this Hamiltonian is short-ranged if $S_{topo} = 0$ and otherwise contains non-local interactions. By using the triangle inequality, we obtain that

$$\left\|\lambda(\rho_Y^{\otimes 2}) - \lambda\left(e^{-H_X}\right)\right\|_1 \leq e^{-\Theta(l)}\,. \tag{4.233}$$

Let us introduce another cut-off to the spectrum

$$\lambda_\Lambda(A) := \left\{\lambda \in \lambda(A) \,\middle|\, \lambda \geq \frac{1}{\Lambda}\right\} \tag{4.234}$$

which bounds possible values of the spectrum from below. Clearly, it implies that

$$\left\|\lambda_\Lambda(\rho_Y^{\otimes 2}) - \lambda_\Lambda\left(e^{-H_X}\right)\right\|_1 \leq e^{-\Theta(l)}\,. \tag{4.235}$$

Using the Lipschitz continuity of the logarithm in $[1/\Lambda, \infty)$, we conclude that

$$\left\|\lambda^\Lambda(-\ln \rho_Y^{\otimes 2}) - \lambda^\Lambda\left(H_X\right)\right\|_1 \leq \Lambda e^{-\Theta(l)}\,. \tag{4.236}$$

Since $H_{\rho_Y}^{(2)} = -\ln \rho_Y^{\otimes 2}$, we complete the proof. □

4.5 Concluding Remarks

We have shown that the TEE is equivalent to the irreducible correlation and the optimal rate of a secret sharing protocol under assumption (I) and (II), and thus provide characterizations of the multipartite correlations in topologically ordered phases by information-theoretically well-defined functions. The equivalence between the TEE and the irreducible correlation implies that the TEE is directly related the structure of the entanglement Hamiltonian on a region, and also the ES of a half cylinder.

A weakness of the TEE as an indicator of topologically ordered phases is that a non-zero value of the TEE for a gapped ground state does not necessary imply that the states are in a topologically ordered phase as shown by a counterexample constructed by Bravyi [65]. A natural question is to ask whether a complete characterization of multipartite correlations in topologically ordered phases excluding Bravyi's counterexample is possible or not by considering the irreducible correlation (or the optimal rate of a secret sharing protocol). Unfortunately, the answer is probably no, since Bravyi's example satisfies our assumptions and therefore the TEE is equivalent to the irreducible correlation. Therefore, the highest-order irreducible correlation and the optimal rate of a secret sharing protocol only provide a partial characterization in this sense. So far, we do not know if there is a better way to characterize correlations in topologically ordered phases or there is no purely "characteristic" multipartite correlations in topologically ordered phases which distinguish them from the trivial phase.

One possible way to solve this problem is to decompose the irreducible correlation into a quantum part and a classical part. The value of the irreducible correlation represents the amount of both quantum and classical correlations, and there is no way to distinguish whether the correlation has a quantum or classical origin. In the case of Bravyi's counterexample, the nontrivial irreducible correlation can be understood as classical, since the reduced state is diagonal in a product basis (and therefore separable). However, one would expect that the multipartite correlations in topologically ordered phases should be somehow "quantum", as sometimes referred as long-range entanglement [109, 48, 123]. Recall that the irreducible correlation is defined by using the set of Gibbs states $\mathcal{E}_k$. It may be possible to consider a convex hull of $\mathcal{E}_k$ and define a quantum version of the irreducible correlation in terms of the distance from these sets of states [4]. Another possibility is to consider a quantum secret sharing of quantum information instead of classical bits. Investigating these extensions remains as a future problem.

Another remaining problem is the relation between the irreducible correlation

[4]The convex hull of $\mathcal{E}_k$ has been already considered in Ref. [146] to define a "witness" of the irreducible correlation.

and the optimal rate of the secret sharing protocol. They are known to coincide for stabilizer states [134, 70], and we have shown that the equivalence also holds for reduced states of gapped ground states with exactly zero correlation length. However, we have not showed that the equivalence holds for the case of the finite-correlation length even approximately, since the proof relies on a property of Markov state which is fragile against perturbations. Further, in the secret sharing protocol, the encoding method is very restricted: it must be performed by unitaries which do not change reduced states. Relaxing these conditions may help to solve this problem. It is also interesting to check whether the fragile property still approximately holds by requiring reasonable constraints which gapped ground states are expected to satisfy. Moreover, whether these two functions, the irreducible correlation and the optimal rate of the secret sharing protocol, are exactly same for all quantum states is also interesting problem to investigate from a viewpoint of quantum information theory.

Chapter 5

1D Quantum Gibbs States and Approximate Markov Chains

Gibbs states which describe systems in thermal equilibrium are a main topic of analysis when studying quantum states at finite temperature. Typical Hamiltonians in many-body physics only contain short-range interactions. For short-range Hamiltonians, the corresponding Gibbs states are known to have several aspects of "locality". For example, it has been proven that such a Gibbs state has finite correlation length in 1D systems at any temperature, and also in higher-dimensional systems above a universal temperature (i.e., independent of the system size and details of the Hamiltonian) [147]. In addition, an area law in terms of the mutual information for any spatial dimension has been proven [44]. These results provide quantitative representations of an intuition that correlations in the Gibbs states of short-range Hamiltonians are short-ranged.

In classical information theory, the Hammersley-Clifford theorem [79] is a fundamental result establishing another type of locality of Gibbs states, the Markov property. A state or probability distribution on a many-body system satisfies the Markov property[1] if $I(A : C|B) = 0$ for any tripartition ABC of the whole system, such that B separates C from A. A distribution or state satisfying the Markov property is called a Markov network, a generalization of Markov chains. The definition states that correlations between A and C are always mediated by B and there is no direct correlation between A and C. The Hammersley-Clifford theorem [79] states that any positive[2] Markov network is equivalent to a Gibbs state of a short-range classical Hamiltonian. When we take the logarithm of a positive distribution, we can always consider the distribution as a Gibbs state of some "Hamiltonian". The Hammersley-Clifford theorem implies that the Markov

[1] It is more precisely called the *global* Markov property.

[2] A probability distribution is called positive if it has only non-zero entries. In a similar manner, a quantum state is called positive if all eigenvalues are strictly positive.

property strictly restricts the possible form of such Hamiltonians, and also implies Gibbs states of short-range Hamiltonians satisfy the Markov property. A quantum analog of the Hammersley-Clifford theorem has been investigated [80, 46, 81]. As a result, it has been shown that the quantum version of the positive Markov networks is equivalent to a Gibbs state of a commuting short-range Hamiltonian.

Since general Hamiltonians of quantum systems contain non-commuting interaction terms, it is natural to ask whether the Gibbs states of general short-range Hamiltonians obey a similar property. Indeed, one can show that the area law for mutual information [44] implies that such Gibbs states obey the Markov property, *approximately*. However, the area law does not show the accuracy of the approximation.

In this chapter, we go one step further and show the very first explicit bound on the accuracy of the Markov property of a Gibbs state of any 1D short-range Hamiltonian. Moreover, we show that in 1D any state approximately satisfying the Markov property, which we call an *approximate Markov chain*, can be well-approximated by a Gibbs state of some short-range Hamiltonian. These statements form a generalization of the quantum Hammersley-Clifford theorem for 1D systems: Gibbs states of 1D short-range Hamiltonians and approximate Markov chains are approximately equivalent.

The proof of the upper bound exploits the fact that the conditional mutual information of a state is bounded from above by its distance to a reconstructed state. The reconstructed state is a global state which is created from the reduced state of the given state by applying a recovery map. In the proof, we explicitly construct a recovery map which approximately creates the Gibbs state. The recovery map consists of operators which remove/add local interactions in the Hamiltonian associating to the Gibbs state. Although such operations only succeed with a constant probability, employing a repeat-until-success method provides a solution to make a deterministic recovery map. For the converse, i.e., the approximate Markov property implies that the corresponding Hamiltonian is short-ranged, we use the equivalence between states having maximum entropy under linear constraints and Gibbs states (see Sec. 2.3 for details). The approximate Markov property offers a new characterization the locality of general 1D Gibbs states. Due to a close relationship between states with small conditional mutual information and local recoverability of the state, this result has an application for preparing of 1D Gibbs states on a quantum computer. The decay rate of our upper bound on the conditional mutual information explains how fast the mutual information of a region saturates the bound from the area law for mutual information. Conversely, our results also provide a new structural characterization of approximate Markov chains in terms of Gibbs states of short-range Hamiltonians.

This chapter is organized as follows. In Sec. 5.1, we review the Hammersley-

Clifford theorem. We then represent our main results and applications in Sec. 5.2. Proofs are given in Sec. 5.3 and Sec. 5.4. We discuss our results in Sec. 5.5.

5.1 The Hammersley-Clifford Theorem

In this section, we review the Hammersley-Clifford theorem and its quantum extension which has been presented in Refs. [80, 81].

Consider random variables $X_{v_1}, ..., X_{v_k}$ associated to vertices $V = (v_1, ..., v_k)$ of a graph $G = (V, E)$. We denote the set of random variables on a subset of vertices $A \subset V$ by X_A. We call a set of random variables on a graph a *Markov network* if the associating distribution $p(X_{v_1} = x_1, ..., X_{v_k} = x_k)$ satisfies

$$I\left(X_A : X_B | X_C\right)_p = 0 \tag{5.1}$$

for any (disjoint) tripartition ABC of V such that any path connecting A and C through B (Fig. 5.1). We say such a situation B shields A from C.

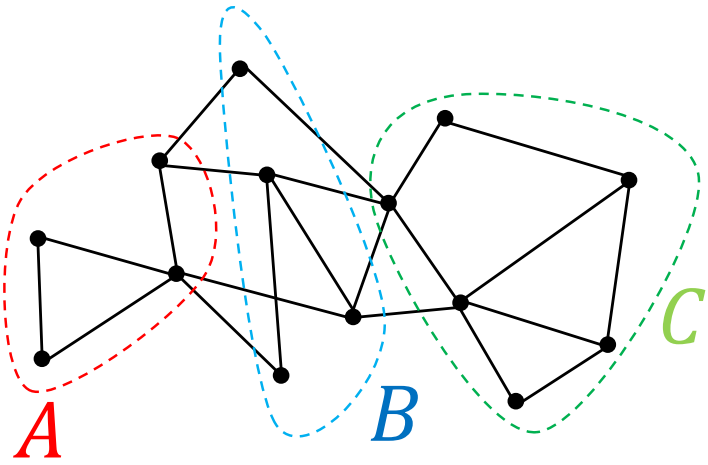

Figure 5.1: An example of graphs with appropriate tripartition ABC of vertices. All paths connecting A and C pass B.

The Hammersley-Clifford theorem states that (positive) Markov networks are equivalent to Gibbs states of short-range Hamiltonians defined on the graph.

Theorem 15. *(The Hammersley-Clifford Theorem [79]) A positive probability distribution on a graph $G = (V, E)$ is a Markov network if and only if it can be written as*

$$p(X_{v_1}, ..., X_{v_k}) = \frac{1}{Z} \exp\left(-\sum_c h_c(X_{V_c})\right), \tag{5.2}$$

where the sum is over all cliques [3] $\{c\}$*,* X_{V_c} *is the set of random variables corresponding to vertices of the clique* c *and* h_c *is a real function on it.*

Note that when the graph G does not contains any triangles, all interactions $h(X_{V_{c_i}})$ are nearest-neighbor interactions. Therefore, any Markov network on the graph is a Gibbs state of a nearest-neighbor Hamiltonian $H \equiv \sum_{c_i} h(X_{V_{c_i}})$ and vice versa.

In a similar manner, we define a quantum Markov network as a state $\rho \in \mathcal{S}(\mathcal{H}_{V_1} \otimes \cdots \otimes \mathcal{H}_{V_k})$ on a Hilbert space $\mathcal{H}_{V_1} \otimes \cdots \otimes \mathcal{H}_{V_k}$ associated to vertices of a graph $G = (V, E)$, such that

$$I\left(A : C|\, B\right)_\rho = 0\,. \tag{5.3}$$

A Markov network ρ is called positive if $\rho > 0$. In Ref. [80], one direction of the equivalence has shown.

Theorem 16. *[80] For any positive Markov network* ρ*, there exists a Hamiltonian* $H = \sum_c h_c$ *such that*

$$\rho = e^{-H}\,, \tag{5.4}$$

where the sum is over all cliques of G*.*

For the reverse direction, we encounter the problem of non-commutativity of quantum Hamiltonians. However, for graphs only containing two-body cliques, positive Markov networks are equivalent to Gibbs states of local and commuting Hamiltonians.

Theorem 17. *[81] If a system defined on a graph* G *contains only two-body cliques, a positive state* ρ *on the system is a Markov network if and only if* $\rho = e^{-H}$*, where*

$$H = \sum_C h_C, \quad [h_C, h_{C'}] = 0 \tag{5.5}$$

for any cliques C*,* C'*.*

For more general short-range commuting Hamiltonians, the Markov property holds when A and C are sufficiently separated. For any Gibbs state ρ of a Hamiltonian $H = \sum_{v \in V} h_v$ on a graph $G = (V, E)$, where each interaction h_v acts on spins within distance r and $[h_v, h_{v'}] = 0$ for any $v, v' \in V$, it holds that

$$I(A : C|B)_\rho = 0 \tag{5.6}$$

for any tripartition ABC of V such that B shields A from C and $d(A, C) \geq r$. The converse also holds, i.e., any positive state on G satisfying above condition can be written as a Gibbs state of a Hamiltonian $H = \sum_v h_v$, where each h_v acts on spins within distance r from v. The proofs are straightforward extensions of the original proofs of the quantum Hammersley-Clifford theorem given in Ref. [81].

[3] A clique of a graph is a subgraph where all two vertices are connected by an edge.

5.2 An Approximate Quantum Hammersley-Clifford Theorem for 1D systems

We summarize our results of this chapter and discuss its implications to the area law for mutual information, preparation algorithm for 1D Gibbs states and an upper bound on the squashed entanglement.

5.2.1 Settings and Notations

We consider a 1D quantum spin system $\Lambda = \{1, ..., n\}$ with n spins. When we consider the periodic boundary condition, we set $n+1 \equiv 1$. Each spin i corresponds to a Hilbert space $\mathcal{H}_i$ with dimension $d < \infty$. We denote $\mathcal{H}_X$ as the Hilbert space corresponding to a subsystem defined by $X \subset \Lambda$. We say that the *support* of an operator V is $X \subset \Lambda$ if V can be written as

$$V = V_X \otimes \mathbb{1}_{X^c}\,, \tag{5.7}$$

i.e., the product of some operator V_X on $\mathcal{H}_X$ and the identity operator acting on the complement of X, which is denoted by X^c. We will denote the support of an operator V by $\mathrm{supp}(V)$.

We consider a short-range Hamiltonian on Λ given by $H_\Lambda = \sum_i h_i$, where each h_i is bounded and $\mathrm{supp}(h_i)$ contains $r < \infty$ neighboring spins around the spin i. We will denote a restricted Hamiltonian on a region X by H_X, which is defined as

$$H_X = \sum_{\mathrm{supp}(h_i) \subset X} h_i\,. \tag{5.8}$$

The Gibbs state of a Hamiltonian H_X at an inverse temperature β is defined as

$$\rho^{H_X} = \frac{e^{-\beta H_X}}{Z}\,, \tag{5.9}$$

where $Z_X = \mathrm{Tr}[e^{-\beta H_X}]$ is the partition function. Note that we will omit β in the notations and simply denote ρ^{H_X} and Z_X, while they depend on β. The reduced state of this Gibbs state of a subsystem $Y \subset \Lambda$ is denoted by $\rho_Y^{H_X}$.

5.2.2 Main Results

We summarize our main results. Our first result is the following theorem:

Theorem 18. *Let H be a short-range 1D Hamiltonian defined on a 1D spin system Λ. For any $\beta > 0$, there exists a constant $l_0 > 0$ such that for any tripartition ABC*

of Λ *with* $d(A,C) \geq l_0$*, there exists a CPTP-map* $\Lambda_{B\to BC} : \mathcal{S}(\mathcal{H}_B) \to \mathcal{S}(\mathcal{H}_{BC})$ *satisfying*

$$\left\|\rho^{H_{ABC}} - (\mathrm{id_A} \otimes \Lambda_{\mathrm{B}\to\mathrm{BC}})\left(\rho_{\mathrm{AB}}^{\mathrm{H_{ABC}}}\right)\right\|_1 \leq e^{-q(\beta)\sqrt{d(A,C)}}\,, \tag{5.10}$$

where $q(\beta) = Ce^{-c\beta}$ *for some constants* $C, c > 0$*.*

This theorem means that if A and C are sufficiently separated, the Gibbs state can be approximately recovered from the reduced state on AB by performing a recovery map acting on B. Moreover, the approximation accuracy is exponentially good with respect to the square root of $d(A,C)$. It turns out that the corresponding conditional mutual information decays in a similar way:

Corollary 19. *Under the setting of Theorem 18, it holds that*

$$I(A:C|B)_{\rho^{H_{ABC}}} \leq \left(d(A,C) + 3q(\beta)^{-2}d(A,C)^{-\frac{1}{2}}\right) e^{-q(\beta)\sqrt{d(A,C)}}\,. \tag{5.11}$$

The proof is given in Sec. 5.3.3. The RHS of Eq. (5.11) is in $e^{-\Theta\left(\sqrt{d(A,C)}\right)}$, and moreover it is independent of the dimensions of subsystems. This is the very first explicit bound on the conditional mutual information of general Gibbs states. However, our bound may not be tight and it may be possible to obtain a better decaying rate for the same class of Gibbs states.

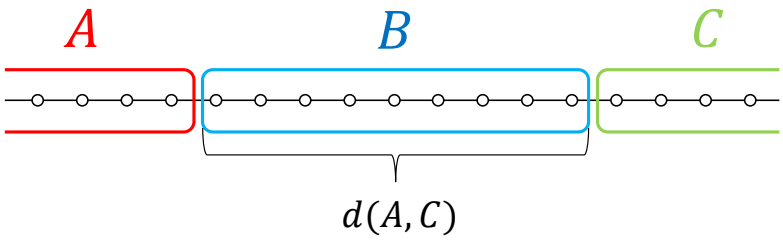

Figure 5.2: A 1D chain with tripartition ABC where B shields A from C.

Our second main result is a kind of converse to Theorem 18.

Theorem 20. *Let* $\rho_{A_1\ldots A_n} \in \mathcal{S}(\mathcal{H}_{A_1} \otimes \cdots \mathcal{H}_{A_n})$ *be an* ε*-approximate Markov chain. Then, there exists a short-range Hamiltonian* $H = \sum_i h_{A_iA_{i+1}}$ *with* $\mathrm{supp}(h_{A_iA_{i+1}}) = A_iA_{i+1}$ *such that*

$$S\left(\rho \middle\| \frac{e^{-H}}{\mathrm{Tr}e^{-H}}\right) \leq \varepsilon n\,. \tag{5.12}$$

Corollary 19 claims that any 1D Gibbs state of a short-range Hamiltonian is an approximate Markov chain with subexponentially good accuracy. Conversely,

Theorem 20 shows that any 1D approximate Markov chain is close to a Gibbs state of a 1D short-range Hamiltonian. The combination of these two results consists a variant of the Hammersley-Clifford theorem for general 1D quantum systems. Next, we discuss three applications of our results.

Saturation Rate of The Area Law for The Mutual Information

A Gibbs state of a short-range Hamiltonian obeys the area law in terms of the mutual information [44]. For a Gibbs state ρ^H of a short-range Hamiltonian H defined on a lattice, it holds that

$$I(A:A^c)_{\rho^H} \leq 2\beta J|\partial A|\,, \tag{5.13}$$

where J is a universal constant depends on the strength of the interactions in H. Note that in 1D systems, $|\partial A| = 2$ and therefore the upper bound is a constant.

As we have discussed before, the area law represented by Eq. (5.13) implies a decay of the conditional mutual information. Let us assume that A is a simply connected region on a lattice and divide A^c into $\{B_i\}$ so that B_1 shields B_2 from A, B_2 shields B_3 from AB_1 and so on. By the monotonicity of the mutual information under the partial trace, we have

$$I(A:B_1B_2...B_l)_{\rho^H} \leq I(A:B_1B_2...B_lB_{l+1})_{\rho^H} \leq I(A:A^c)_{\rho^H} \leq C\beta|\partial A|\,, \tag{5.14}$$

where $C > 0$ is a constant. Since the upper bound is independent of l, for any $\varepsilon > 0$, there exists l such that

$$I(A:B_1...B_{l+1})_{\rho^H} - I(A:B_1...B_l)_{\rho^H} = I(A:B_{l+1}|B_1...B_l)_{\rho^H} \leq \varepsilon\,. \tag{5.15}$$

Our result provides how fast the conditional mutual information decays in 1D Gibbs states. Suppose that each size of B_i is some sufficiently large constant $w > 0$. From Corollary 19, we have

$$I(A:B_1...B_{l+1})_{\rho^H} - I(A:B_1...B_l)_{\rho^H} \leq e^{-\Theta(\sqrt{wl})}\,. \tag{5.16}$$

Therefore, 1D Gibbs states saturate the upper bound of the area law subexponentially fast in l.

Efficient Preparation of 1D Gibbs states

Theorem 18 ensures that there exist local CPTP-maps which approximately recover the state from tracing out operations on local regions. This statement can be rephrased that the Gibbs state can be prepared by locally "patching" reduced states by using recovery maps.

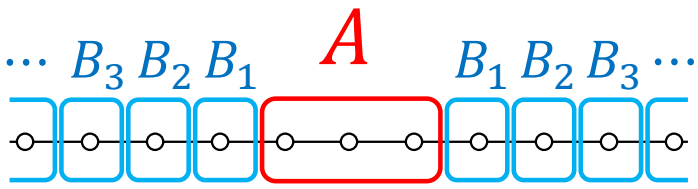

Figure 5.3: An example of regions AB. When the number of regions $\{B_1, B_2, ..., B_l\}$ increases, or equivalently, when the size of $B_1...B_l$ grows, the mutual information $I(A : B_1...B_l)$ of the Gibbs state rapidly saturates the upper bound given by the area law.

Corollary 21. *A Gibbs state of any 1D short-range Hamiltonian can be well-approximated by a depth-two (mixed) circuit with each gate acting on $O(\log^2(n))$ qubits.*

In more detail, there is a CPTP-map of the form

$$\Delta = \left(\bigotimes_i \Delta_{2,i}\right)\left(\bigotimes_i \Delta_{1,i}\right), \tag{5.17}$$

with each CPTP-map $\Delta_{k,i}$ acting on $O(e^{O(\beta)}\log^2(n/\varepsilon))$ sites, with $\Delta_{k,i}$ and $\Delta_{k,j}$ acting on non-overlapping sites for $i \neq j$, such that

$$\left\| \Delta(\tau) - \frac{e^{-\beta H}}{\mathrm{Tr}(e^{-\beta H})} \right\|_1 \leq \varepsilon, \tag{5.18}$$

with the maximally mixed state τ.

The proof is given in Sec. 5.3.4.

Bounds on The Squashed Entanglement of 1D Gibbs states

Entanglement contained in 1D Gibbs states has been investigated for specific models and different entanglement measures [148, 149, 150]. Intuitively, it is natural that one expects any 1D Gibbs state to contain only "short-range" entanglement. Although any 1D Gibbs state ρ^H of a short-range Hamiltonian H exhibits an exponential decay of the correlation functions [151], it only provides a dimension-dependent upper bound on correlation measures.

However, as we have discussed in Sec. 2.2.2, the conditional mutual information provides an upper bound on the amount of entanglement in terms of the squashed entanglement. Therefore, our result implies

$$E_{sq}(A : B)_{\rho^H} \leq Ce^{-c\sqrt{d(A,B)}}. \tag{5.19}$$

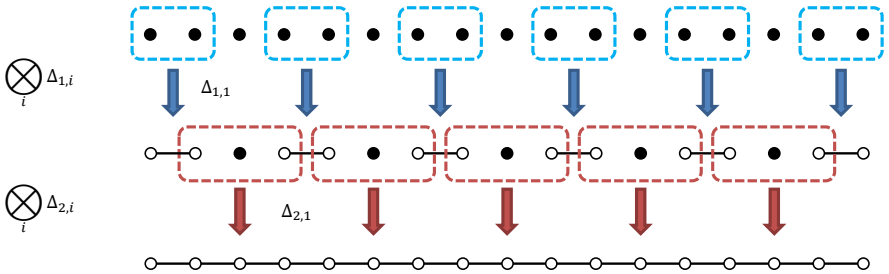

Figure 5.4: A schematic picture of the preparation algorithm for 1D Gibbs states. At the first step (blue arrow), we perform CPTP-map $\bigotimes_i \Delta_{1,i}$ on a product state (black dots). Each $\Delta_{1,i}$ acts on a small set of spins (dotted circle) and produces the reduced state of the target Gibbs state on the set. At the second step (red arrow), we perform another CPTP-map to concatenate these reduced states locally. Due to the approximate Markov property of the Gibbs state, the output state is close to the Gibbs state.

This bound holds for any 1D short-range Hamiltonian, and moreover the bound is independent of the size of regions A and B. Such a general bound had not been known.

5.3 Proof: 1D Gibbs States are Approximate Markov Chains

In this section, we prove Theorem 18 and Corollary 19. A key technical tool in the proof of the main theorem is the quantum belief propagation equations which have been studied in Refs. [152, 47]. By combining with the Lieb-Robinson bounds [153], the equations implies that if a short-range Hamiltonian is perturbed locally, the corresponding Gibbs state also changes (almost) locally. This fact is also derived by Araki in Ref. [151] by using analysis of imaginary time evolutions of operators.

5.3.1 Quantum Belief Propagation Equations

Consider an one-parameter family of a Hamiltonian H on a spin lattice with a perturbation operator V

$$H(s) = H + sV\,, \tag{5.20}$$

where $s \in [0, 1]$. The change of the Gibbs state of $H(s)$ under a small change of s can be computed by a quantum belief propagation equation [152, 47]:

$$\frac{d}{ds}e^{-\beta H(s)} = -\frac{\beta}{2}\left\{e^{-\beta H(s)}, \Phi_\beta^{H(s)}(V)\right\}\,, \tag{5.21}$$

where the operator $\Phi_\beta^{H(a)}(V)$ is given by [47]

$$\Phi_\beta^{H(s)}(V)_{ij} = V_{ij}\tilde{f}_\beta(E_i(s) - E_j(s)) \tag{5.22}$$

in the energy eigenbasis of $H(s) = \sum_i E_i(s)|i\rangle\langle i|$ [4], with $\tilde{f}_\beta(\omega) = \frac{\tanh(\beta\omega/2)}{\beta\omega/2}$. Using the Fourier transform $f_\beta(t) = \frac{1}{2\pi}\int d\omega \tilde{f}_\beta(\omega)e^{i\omega t}$, $\Phi_\beta^{H(s)}(V)$ can be written as the integral form:

$$\Phi_\beta^{H(s)}(V) = \int_{-\infty}^{\infty} dt f_\beta(t) e^{-iH(s)t} V e^{iH(s)t}\,. \tag{5.23}$$

Taking the formal integration of Eq. (5.21), we obtain

$$e^{-\beta H(1)} = O e^{-\beta H(0)} O^\dagger\,, \tag{5.24}$$

where the operator O is defined as

$$O := \mathcal{T}\exp\left(-\frac{\beta}{2}\int_0^1 ds' \Phi_\beta^{H(s')}(V)\right) \tag{5.25}$$

$$:= \sum_{n=0}^{\infty}\left(-\frac{\beta}{2}\right)^n \int_0^1 ds_1' \int_0^{s_1'} ds_2' ... \int_0^{s_{n-1}'} ds_n' \Phi_\beta^{H(s_n')}(V)\cdots\Phi_\beta^{H(s_1')}(V)\,. \tag{5.26}$$

The operator norm of $\Phi_\beta^{H(s)}(V)$ can be evaluated as

$$\left\|\Phi_\beta^{H(s)}(V)\right\| = \left\|\int_{-\infty}^{\infty} dt f_\beta(t) e^{-iH(s)t} V e^{iH(s)t}\right\| \tag{5.27}$$

$$= \left\|\int_{-\infty}^{\infty} dt' f_1(t') e^{-i\beta H(s)t'} V e^{i\beta H(s)t'}\right\| \tag{5.28}$$

$$\leq \|V\|\left|\int_{-\infty}^{\infty} dt' f_1(t')\right| \tag{5.29}$$

$$= \|V\|. \tag{5.30}$$

The second line follows from $dt f_\beta(t) = \frac{dt'}{\beta} f_1(\frac{t'}{\beta})$ and the integral in the last equality can be calculated through the series expansion:

$$\frac{\tanh(x)}{x} = \sum_{k=0}^{\infty}\frac{2}{x^2 + (k+\frac{1}{2})^2\pi}\,. \tag{5.31}$$

The upper bound (5.27) provides an upper bound of $\|O\|$:

$$\|O\| \leq e^{\frac{\beta}{2}\|V\|}\,. \tag{5.32}$$

[4] Each $|i\rangle$ also depends on s as well as the eigenvalues.

When the Hamiltonian H is short-ranged, time-evolutions of a local operator by H is restricted by the Lieb-Robinson bounds [153]. Let us consider two operators O_A and O_B supported on separated local regions A and B. Then, the Lieb-Robinson bound for these local operators is formulated as

$$\left\| \left[O_A, e^{-iHt} O_B e^{iHt} \right] \right\| \leq c \|O_A\| \|O_B\| \min(|A|, |B|) e^{c'(vt - d(A,B))} , \tag{5.33}$$

where $c, v \geq 0, c' > 0$ are constants. Suppose further that $\mathrm{supp}(V)$ is a simply-connected local region. Then, $H(s)$ obeys the Lieb-Robinson bounds for all $s \in [0,1]$. Since $f_\beta(t)$ in Eq. (5.23) decays fast in $|t|$, the Lieb-Robinson bound implies that $\mathrm{supp}(\Phi_\beta^{H(s)}(V))$ is approximated by some local region $\mathcal{V}_l$, which contains all spins within the distance l from $\mathrm{supp}(V)$. Let us introduce a restricted operator

$$\mathrm{Tr}_{\mathcal{V}_l^c} \left[\Phi_\beta^{H(s)}(V) \right] \otimes \mathbb{1}_{\mathcal{V}_l} \tag{5.34}$$

supported on $\mathcal{V}_l$. Then, from Eq. (5.33), we have [47]

$$\left\| \Phi_\beta^{H(s)}(V) - \mathrm{Tr}_{\mathcal{V}_l^c} \left[\Phi_\beta^{H(s)}(V) \right] \otimes \mathbb{1}_{\mathcal{V}_l} \right\| \leq c' \|V\| e^{-\frac{c'l}{1+c'v\beta/\pi}} . \tag{5.35}$$

We also define the integral of (5.34) as

$$O_{\mathcal{V}_l} := \mathcal{T} \exp \left(-\frac{\beta}{2} \int_0^1 ds' \mathrm{Tr}_{\mathcal{V}_l^c} \left[\Phi_\beta^{H(s')}(V) \right] \otimes \mathbb{1}_{\mathcal{V}_l} \right) . \tag{5.36}$$

Let us choose the coefficients c', v so that Eq. (5.35) holds for all $s \in [0,1]$. Then, we obtain that

$$\|O - O_{\mathcal{V}_l}\| \leq \frac{c'\beta\|V\|}{2} e^{\frac{(1+c')\beta\|V\|}{2}} e^{-\frac{c'l}{1+c'v\beta/\pi}} . \tag{5.37}$$

To see this, consider some parametrized operators $Q(s)$ and $\tilde{Q}(s)$ satisfying $\|Q(s)\|$, $\|\tilde{Q}(s)\| \leq C$ and $\|Q(s) - \tilde{Q}(s)\| \leq \Delta$ for all $s \in [0,1]$. From the simple calculation, we obtain

$$\begin{aligned} Q(s_n)Q(s_{n-1}) \cdots Q(s_1) = \tilde{Q}(s_n)\tilde{Q}(s_{n-1}) \cdots \tilde{Q}(s_1) \\ + \sum_{j=1}^{n} Q(s_n) \cdots Q(s_{j+1}) \Delta_j \tilde{Q}(s_{j-1}) \cdots \tilde{Q}(s_1) , \end{aligned} \tag{5.38}$$

where $\Delta_j = Q(s_j) - \tilde{Q}(s_j)$. This implies

$$\|Q(s_n)Q(s_{n-1}) \cdots Q(s_1) - \tilde{Q}(s_n)\tilde{Q}(s_{n-1}) \cdots \tilde{Q}(s_1)\| \leq nC^{n-1}\Delta . \tag{5.39}$$

In our case, $Q(s) = \Phi_\beta^{H(s)}(V)$ and $\tilde{Q}(s) = \mathrm{Tr}_{\mathcal{V}_l^c}\left[\Phi_\beta^{H(s)}(V)\right] \otimes \mathbb{1}_{\mathcal{V}_l}$. The operator norms of these operators can be bounded as

$$\|Q(s)\|, \|\tilde{Q}(s)\| \leq \|Q(s)\| + \|Q(s) - \tilde{Q}(s)\| \tag{5.40}$$

$$\leq \left(1 + c' e^{-\frac{c' l}{1+c' v\beta/\pi}}\right) \|V\| \tag{5.41}$$

$$\leq (1 + c')\|V\| \,. \tag{5.42}$$

Therefore, the assumption for norms holds when we choose $C = (1 + c')\|V\|$ and $\Delta = c' e^{-\frac{c' l}{1+c' v\beta/\pi}}$. By inserting Eq. (5.39) into the definition of O (5.25), we obtain the bound (5.37).

5.3.2 The proof of Theorem 18

We restate Theorem 18: **Theorem 18.** *Let H be a short-range 1D Hamiltonian defined on Λ. For any $\beta > 0$, there exists a constant $l_0 > 0$ such that for any tripartition ABC of Λ with $d(A, C) \geq l_0$, there exists a CPTP-map $\Lambda_{B\to BC} : \mathcal{S}(\mathcal{H}_B) \to \mathcal{S}(\mathcal{H}_{BC})$ satisfying*

$$\left\|\rho^{H_{ABC}} - (\mathrm{id_A} \otimes \Lambda_{\mathrm{B}\to\mathrm{BC}})\left(\rho_{\mathrm{AB}}^{\mathrm{H_{ABC}}}\right)\right\|_1 \leq e^{-q(\beta)\sqrt{d(A,C)}} \,, \tag{5.43}$$

where $q(\beta) = Ce^{-c\beta}$ for some constants $C, c > 0$.

For simplicity, we shall omit $\mathrm{id_A}$ in the following. The proof of Theorem 18 consists of three steps. First, we show that there exists a CP-map which approximately recovers the Gibbs state $\rho^{H_{ABC}}$ from the reduced state on AB. In the second, we normalize the CP-map to be trace non-increasing. The normalized map can be regarded as a probabilistic quantum operation which succeed to recover the Gibbs state with a constant probability. Finally, in the third step, we construct a CPTP recovery map from the probabilistic operation by employing a repeat-until-success strategy.

Let us begin with the following lemma in which the techniques based on the quantum belief propagation equations are used.

Lemma 22. *For any 1D Gibbs state $\rho^{H_{ABC}}$ of a short-range Hamiltonian H_{ABC} on a 1D system with a tripartition ABC with $l := d(A, C)/2 > r$, there exists a CP-map $\kappa_{B\to BC}$, such that*

$$\left\|\rho^{H_{ABC}} - \kappa_{B\to BC}\left(\rho_{AB}^{H_{ABC}}\right)\right\|_1 \leq C_1(\beta) e^{-q_1(\beta) l} \,. \tag{5.44}$$

where $c', v, C_1(\beta)$ and $q_1(\beta)$ are non-negative constants.

Proof. Let us consider a local Hamiltonian $H = \sum_i h_i$ with the range r. We assume that $\| \sum_{j:i\in \mathrm{supp}(h_j)} h_j \| \leq J$ for all i. Without loss of generality, we introduce a tripartition ABC of a 1D system so that each subsystem is simply connected and $d(A, C)$ is chosen to be $2l$ for some integer $l > r$. We then split region B into the left half B^L, which touches A, and the right half B^R which touches C (Fig. 5.5). We denote the sum of the all interactions h_i acting on both B^L and B^R by H_{B^M}.

Remark 23. *When B consists of a fixed number of simply connected regions, each connected component neighboring both A and C is divided into two halves as in the same way. Then, H_{B^M} is the sum of all interaction terms acting on both such divided regions.*

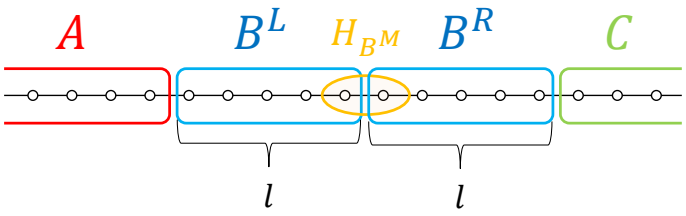

Figure 5.5: A schematic picture of the definition of H_{B^M}. We divide B into two halves B^L and B^R. In the case of a nearest-neighbor Hamiltonian, H_{B^M} is the interaction term acting on both B^L and B^R (the orange circle).

We apply the technical tools discussed in the previous section to an one-parameter family of Hamiltonians $H_{ABC}(s) = H_{AB^L} + H_{B^RC} + sH_{B^M}$. Here, H_{B^M} corresponds to the perturbation operator V in the previous section. From Eq. (5.24), it holds that

$$e^{-\beta H_{ABC}} = O_{ABC} e^{-\beta(H_{AB^L}+H_{B^RC})} O_{ABC}^\dagger , \tag{5.45}$$

where O_{ABC} is defined as

$$O_{ABC} := \mathcal{T} \exp\left(-\frac{\beta}{2} \int_0^1 ds' \Phi_\beta^{H(s')}(H_{B^M}) \right) . \tag{5.46}$$

The inverse of O_{ABC} is given by

$$O_{ABC}^{-1} := \bar{\mathcal{T}} \exp\left(\frac{\beta}{2} \int_0^1 ds' \Phi_\beta^{H(s')}(H_{B^M}) \right) , \tag{5.47}$$

where $\tilde{\mathcal{T}}$ is the inverse time-ordering operator. We shall denote O_{ABC}^{-1} by $\tilde{O}_{ABC}$ as well. By definition, it holds that

$$O_{ABC}\tilde{O}_{ABC} = \tilde{O}_{ABC}O_{ABC} = \mathbb{1}_{ABC} . \tag{5.48}$$

From Eq. (5.32), we have an upper bound of the operator norms:

$$\|O_{ABC}\|, \|\tilde{O}_{ABC}\| \leq e^{\frac{\beta}{2}J} . \tag{5.49}$$

We note that the bound is independent of the size of subsystems, although these operators are non-trivially acting on the whole system ABC.

Since Eq. (5.45) holds, it is not difficult to see that

$$\rho^{H_{ABC}} = O_{ABC} \left[\mathrm{Tr}_{B^R C} \left(\tilde{O}_{ABC} \rho^{H_{ABC}} \tilde{O}_{ABC}^{\dagger} \right) \otimes \rho_{B^R C}^{H_{ABC}} \right] \tag{5.50}$$

holds. Since the "perturbation" H_{B^M} has a local support, there exist a local operator which approximates O_{ABC} as shown in Eq. (5.37). Let us choose the support of the approximating operator to be B and denote it by O_B. For simplicity, we denote

$$K(\beta) = \frac{c'\beta J}{2} e^{\frac{(1+c')\beta J}{2}} . \tag{5.51}$$

Then, there exist constants c' and v such that

$$\|O_{ABC} - O_B\| \leq K(\beta) e^{-q_1(\beta) l} , \tag{5.52}$$

where $q_1(\beta) = \frac{c'}{1+c' v \beta/\pi}$. The same relation holds for $\tilde{O}_{ABC}$ and its local approximation $\tilde{O}_B$. Their operator norms are bounded as

$$\|O_B\|, \|\tilde{O}_B\| \leq \|O_{ABC}\| + \|O_{ABC} - O_B\| \leq e^{\frac{\beta J}{2}} + K(\beta) . \tag{5.53}$$

Let us denote the non-trivial part of $\tilde{O}_B$ by $\tilde{O}_{B|B}$, i.e.,

$$\tilde{O}_B = \tilde{O}_{B|B} \otimes \mathbb{1}_{AC} . \tag{5.54}$$

By using this notation, we define a CP-map $\kappa_{B \to BC}$ by replacing the operators in Eq. (5.50) by their local approximations:

$$\kappa_{B \to BC}(\sigma_B) := O_B \left[\mathrm{Tr}_{B^R} \left(\tilde{O}_{B|B} \sigma_B \tilde{O}_{B|B}^{\dagger} \right) \otimes \rho_{B^R C}^{H_{B^R C}} \right] O_B^{\dagger} . \tag{5.55}$$

Note that the partial trace over C is removed from Eq. (5.50).

In the following, we show that $\kappa_{B \to BC}(\rho_{AB}^{H_{ABC}})$ is close to $\rho^{H_{ABC}}$. Let us denote

$$X_1 := \mathrm{Tr}_{B^R C} \left(\tilde{O}_{ABC} \rho^{H_{ABC}} \tilde{O}_{ABC}^{\dagger} \right) \otimes \rho_{B^R C}^{H_{B^R C}} \tag{5.56}$$

and

$$X_2 := \mathrm{Tr}_{B^R} \left(\tilde{O}_B \rho_{AB}^{H_{ABC}} \tilde{O}_B^{\dagger} \right) \otimes \rho_{B^R C}^{H_{B^R C}} . \tag{5.57}$$

For any state σ_{ABC}, we have

$$\begin{aligned}
\|X_1 - X_2\|_1 &= \left\| \mathrm{Tr}_{B^R C}\left(\tilde{O}_{ABC}\sigma_{ABC}\tilde{O}_{ABC}^\dagger\right) \otimes \rho_{B^R C}^{H_{B^R C}} \right. \\
&\qquad \left. - \mathrm{Tr}_{B^R}\left(\tilde{O}_{B|B}\sigma_{AB}\tilde{O}_{B|B}^\dagger\right) \otimes \rho_{B^R C}^{H_{B^R C}} \right\|_1 && (5.58)\\
&= \left\| \mathrm{Tr}_{B^R C}\left(\tilde{O}_{ABC}\sigma_{ABC}\tilde{O}_{ABC}^\dagger\right) - \mathrm{Tr}_{B^R C}\left(\tilde{O}_B\sigma_{ABC}\tilde{O}_B^\dagger\right) \right\|_1 && (5.59)\\
&\leq \left\| \tilde{O}_{ABC}\sigma_{ABC}\tilde{O}_{ABC}^\dagger - \tilde{O}_B\sigma_{ABC}\tilde{O}_B^\dagger \right\|_1 && (5.60)\\
&\leq \left\| (\tilde{O}_{ABC} - \tilde{O}_B)\sigma_{ABC}\tilde{O}_{ABC}^\dagger \right\|_1 + \left\| \tilde{O}_B\sigma_{ABC}(\tilde{O}_{ABC}^\dagger - \tilde{O}_B^\dagger) \right\|_1 && (5.61)
\end{aligned}$$

We used the monotonicity of the trace-norm in the last inequality. To address the calculation, we use the following spacial case of the Hölder's inequality:

$$\|AB\|_1 \leq \|A\|_1 \|B\| \,. \tag{5.62}$$

It implies that

$$\begin{aligned}
(5.61) &\leq \left\| (\tilde{O}_{ABC} - \tilde{O}_B) \right\| \left\| \sigma_{ABC}\tilde{O}_{ABC}^\dagger \right\|_1 + \left\| \tilde{O}_B\sigma_{ABC} \right\|_1 \left\| (\tilde{O}_{ABC}^\dagger - \tilde{O}_B^\dagger) \right\| && (5.63)\\
&\leq \left\| (\tilde{O}_{ABC} - \tilde{O}_B) \right\| \left\| \tilde{O}_{ABC}^\dagger \right\| + \left\| \tilde{O}_B \right\| \left\| (\tilde{O}_{ABC}^\dagger - \tilde{O}_B^\dagger) \right\| && (5.64)\\
&\leq 2K(\beta)\left(e^{\frac{\beta J}{2}} + K(\beta)\right) e^{-q_1(\beta)l} \,. && (5.65)
\end{aligned}$$

The first and second lines follow from Eq. (5.62) and $\|\sigma_{ABC}\|_1 = 1$. In the last line, we used Eq. (5.52) and Eq. (5.53).

By using the above bound, we bound the difference between the original Gibbs state $\rho^{H_{ABC}}$ and $\kappa_{B\to BC}(\rho_{AB}^{H_{ABC}})$ as

$$\begin{aligned}
&\left\| \rho^{H_{ABC}} - \kappa_{B\to BC}(\rho_{AB}^{H_{ABC}}) \right\|_1 \\
&= \left\| O_{ABC}X_1O_{ABC}^\dagger - O_BX_2O_B^\dagger \right\|_1 && (5.66)\\
&\leq \left\| O_{ABC}X_1O_{ABC}^\dagger - O_BX_1O_B^\dagger \right\|_1 + \left\| O_B(X_1 - X_2)O_B^\dagger \right\|_1 && (5.67)\\
&\leq \left\| (O_{ABC} - O_B)X_1O_{ABC}^\dagger \right\|_1 + \left\| O_BX_1(O_{ABC}^\dagger - O_B^\dagger) \right\|_1 + \|(X_1 - X_2)\|_1 \left\| O_B^\dagger \right\|^2 && (5.68)\\
&\leq \|O_{ABC} - O_B\| \|\tilde{O}_{ABC}\|^2 \|O_{ABC}\| + \|O_{ABC}^\dagger - O_B^\dagger\| \|\tilde{O}_{ABC}\|^2 \|O_B\| \\
&\quad + \|(X_1 - X_2)\|_1 \left\| O_B^\dagger \right\|^2 && (5.69)\\
&\leq K(\beta)\left(e^{\frac{3\beta J}{2}} + e^{\beta J}\left(e^{\frac{\beta J}{2}} + K(\beta)\right) + 2\left(e^{\frac{\beta J}{2}} + K(\beta)\right)^3\right) e^{-q_1(\beta)l} && (5.70)\\
&\leq 4K(\beta)\left(e^{\frac{\beta J}{2}} + K(\beta)\right)^3 e^{-q_1(\beta)l} \,. && (5.71)
\end{aligned}$$

Here we used the fact that $\|X_1\|_1 \leq \|\rho^{H_{ABC}}\|_1 \|\tilde{O}_{ABC}\|^2 = \|\tilde{O}_{ABC}\|^2$ in the fourth line. Choosing $C_1(\beta) = 4K(\beta)\left(e^{\frac{\beta J}{2}} + K(\beta)\right)^3$ completes the proof. □

Unfortunately, the CP-map $\kappa_{B\to BC}$ is not a trace non-increasing map in general. Next, we normalize $\kappa_{B\to BC}$ to obtain a CP and trace non-increasing map, which corresponds to a probabilistic process described by an instrument (see Sec. 2.1.2 for the definition of instruments).

Lemma 24. *Under the setting of Theorem 22, there exists a CP and trace non-increasing map $\tilde{\Lambda}_{B\to BC}$ for any $l \geq l_0 = \frac{\log C_1(\beta)+1}{q_1(\beta)}$ such that*

$$\left\| \rho^{H_{ABC}} - \frac{\tilde{\Lambda}_{B\to BC}(\rho_{AB}^{H_{ABC}})}{\mathrm{Tr}[\tilde{\Lambda}_{B\to BC}(\rho_{AB}^{H_{ABC}})]} \right\| \leq C_2(\beta) e^{-q_1(\beta) l}\,, \tag{5.72}$$

where $C_2(\beta) = \frac{2C_1(\beta)}{1-e^{-1}}$. Moreover, $p := \mathrm{Tr}[\tilde{\Lambda}_{B\to BC}(\rho_{AB}^{H_{ABC}})] > 0$ is a constant independent of the size of subsystems A, B and C.

Proof. We denote the maximum eigenvalue of $O_B^\dagger O_B$ ($\tilde{O}_B^\dagger \tilde{O}_B$) by $\lambda_{\max}^{O_B}$ ($\lambda_{\max}^{\tilde{O}_B}$). From Eq. (5.53) and inequality $\|A^\dagger A\| \leq \|A\|^2$, these eigenvalues are bounded as

$$\lambda_{\max}^{O_B}, \lambda_{\max}^{\tilde{O}_B} \leq \left(e^{\frac{\beta J}{2}} + K(\beta)\right)^2 . \tag{5.73}$$

Define $\lambda_{\max} := \lambda_{\max}^{O_B} \lambda_{\max}^{\tilde{O}_B}$. Then, we define the normalized map $\tilde{\Lambda}_{B\to BC}$ as

$$\tilde{\Lambda}_{B\to BC}(\sigma_B) := \frac{1}{\lambda_{\max}} \kappa_{B\to BC}(\sigma_B)\,. \tag{5.74}$$

By definition, $\tilde{\Lambda}_{B\to BC}$ is CP and trace non-increasing. Let us define a complementary map $\tau_{B\to BC}$ as

$$\tau_{B\to BC}(\sigma_B) := \sigma_B \otimes \tau_C\,, \tag{5.75}$$

for some fixed state $\tau_C \in \mathcal{S}(\mathcal{H}_C)$. We denote $\tau_{B\to BC} - \tilde{\Lambda}_{B\to BC}$ by $\tilde{E}_{B\to BC}$. Then, $\{\tilde{\Lambda}_{B\to BC}, \tilde{E}_{B\to BC}\}$ is an instrument. We refer the output corresponding to $\tilde{\Lambda}_{B\to BC}$ as the "success". The successful output state for the input $\rho_{AB}^{H_{ABC}}$ is

$$\frac{\tilde{\Lambda}(\rho_{AB}^{H_{ABC}})}{\mathrm{Tr}[\tilde{\Lambda}(\rho_{AB}^{H_{ABC}})]}\,. \tag{5.76}$$

Let us introduce $l_0(\beta) = \frac{\ln C_1(\beta)+1}{q_1(\beta)}$. For any $l \geq l_0(\beta)$, $C_1(\beta)e^{-q_1(\beta)l} \leq e^{-1} < 1$. For such l, the success probability p of the instrument for the input $\rho_{AB}^{H_{ABC}}$ is then

estimated as

$$
\begin{aligned}
p &= \mathrm{Tr}[\tilde{\Lambda}(\rho_{AB}^{H_{ABC}})] && (5.77)\\
&= \frac{1}{\lambda_{\max}} \left\| \kappa_{B\to BC}(\rho_{AB}^{H_{ABC}}) \right\|_1 && (5.78)\\
&\geq \frac{1}{\lambda_{\max}} \left| \|\rho^{H_{ABC}}\|_1 - \left\| \rho^{H_{ABC}} - \kappa_{B\to BC}(\rho_{AB}^{H_{ABC}}) \right\|_1 \right| && (5.79)\\
&\geq \frac{1}{\lambda_{\max}} \left(1 - C_1(\beta) e^{-q_1(\beta) l}\right) && (5.80)\\
&\geq \frac{1}{\lambda_{\max}} \left(1 - \frac{1}{e}\right) && (5.81)\\
&\geq \frac{1 - e^{-1}}{\left(e^{\frac{\beta J}{2}} + K(\beta)\right)^4} && (5.82)\\
&> 0\,, && (5.83)
\end{aligned}
$$

where we used Eq. (5.73) in the line before the last.

The approximation error of the succeeded output is then estimated as

$$
\begin{aligned}
&\left\| \rho^{H_{ABC}} - \frac{\tilde{\Lambda}_{B\to BC}(\rho_{AB}^{H_{ABC}})}{\mathrm{Tr}[\tilde{\Lambda}_{B\to BC}(\rho_{AB}^{H_{ABC}})]} \right\|_1 = \left\| \rho^{H_{ABC}} - \frac{\kappa_{B\to BC}(\rho_{AB}^{H_{ABC}})}{\|\kappa_{B\to BC}(\rho_{AB}^{H_{ABC}})\|_1} \right\|_1 && (5.84)\\
&\leq \left| 1 - \frac{1}{\|\kappa_{B\to BC}(\rho_{AB}^{H_{ABC}})\|_1} \right| \|\rho^{H_{ABC}}\|_1 \\
&\qquad\qquad + \frac{1}{\|\kappa_{B\to BC}(\rho_{AB}^{H_{ABC}})\|_1} \left\| \rho^{H_{ABC}} - \kappa_{B\to BC}(\rho_{AB}^{H_{ABC}}) \right\|_1 && (5.85)\\
&\leq \frac{C_1(\beta) e^{-q_1(\beta) l}}{1 - e^{-1}} + \frac{1}{1 - e^{-1}} C_1(\beta) e^{-q_1(\beta) l} && (5.86)\\
&\leq \frac{2 C_1(\beta)}{1 - e^{-1}} e^{-q_1(\beta) l}\,. && (5.87)
\end{aligned}
$$

In the third line, we used the fact

$$
\left| \|\kappa_{B\to BC}(\rho_{AB}^{H_{ABC}})\|_1 - 1 \right| \leq C_1(\beta) e^{-q_1(\beta) l}\,, \tag{5.88}
$$

which follows from

$$
1 - C_1(\beta) e^{-q_1(\beta) l} \leq \|\kappa_{B\to BC}(\rho_{AB}^{H_{ABC}})\|_1 \leq 1 + C_1(\beta) e^{-q_1(\beta) l}\,. \tag{5.89}
$$

Thus, we conclude that Lemma 24 holds by choosing $C_2(\beta) = \frac{2C_1(\beta)}{1-e^{-1}}$. □

We are now in position to prove Theorem 18. Without loss of generality, let us assume that $d(A,C) = |B| = 3l^2 - l$ for $l \in \mathbb{N}$. We divide B into $B = B_l \bar{B}_{l-1} B_{l-1} ... \bar{B}_1 B_1$ as shown in Fig. 5.6, where for each i, $|B_i| = 2l$ and $|\bar{B}_i| = l$. From Lemma 24, there exists a CP and trace non-increasing map $\tilde{\Lambda}_{B_i \to B_i \bar{B}_{i-1} ... B_1 C}$ for each i which approximately recovers $\rho^{H_{ABC}}$ from the reduced state on $AB_l ... \bar{B}_i B_i$. Let us simply denote $\tilde{\Lambda}_{B_i \to B_i \bar{B}_{i-1} ... B_1 C}$ by $\tilde{\Lambda}_i$ and its complementary map $\tilde{E}_{B_i \to B_i \bar{B}_{i-1} ... B_1 C}$ by $\tilde{E}_i$. We also denote the tracing operation over $\bar{B}_i B_i ... B_1 C$ by Tr_i. We define a CPTP-map $\Lambda_{B \to BC}$ as

$$\Lambda_{B \to BC}(\sigma_B) = \\ \tilde{\Lambda}_1(\sigma_B) + \left(\tilde{\Lambda}_2 + \cdots \left(\tilde{\Lambda}_{l-1} + \left(\tilde{\Lambda}_l + \tilde{E}_l \right) \mathrm{Tr}_{l-1} \tilde{E}_{l-1} \right) \cdots \mathrm{Tr}_2 \tilde{E}_2 \right) \mathrm{Tr}_1 \tilde{E}_1(\sigma_B) \quad (5.90)$$

based on the repeat-until-success method (Fig. 5.6).

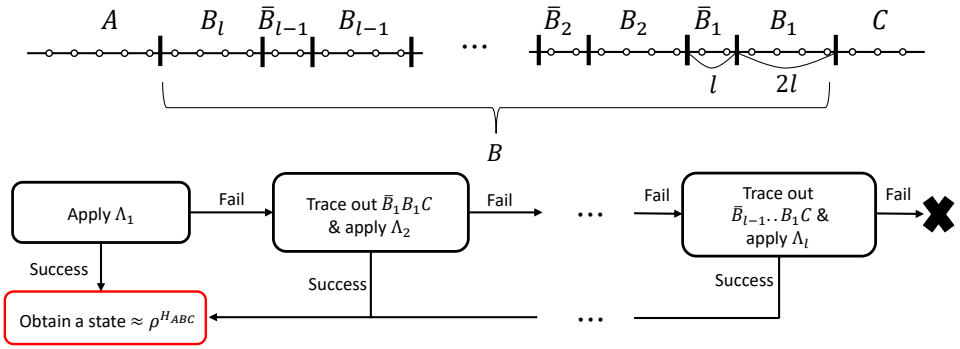

Figure 5.6: A schematic picture of the repeat-until-success method. We introduced buffer systems $\{\bar{B}_i\}$ to suppress the effect of failure. The "failure" output at the end corresponds to the CP-map $\tilde{E}_l \circ \mathrm{Tr}_{l-1} \tilde{E}_{l-1} \cdots \circ \mathrm{Tr}_1 \tilde{E}_1$.

When we input $\rho_{AB}^{H_{ABC}}$ to $\Lambda_{B \to BC}$, the output of each map $\tilde{\Lambda}_i$ corresponds to the success of the recovery process atthe ith step (with probability p_i) and $\tilde{E}_i$ corresponds to the failure of the recovery process (with probability $1-p_i$). If it fails, we trace out both the recovered system and the "buffer" system $\bar{B}_i$, and then, the effect of the failure can be almost neglected. Thus, we can repeat the probabilistic process to obtain the success outcome. The effect of the failure is estimated by the following lemma, which utilizes the exponential decay of correlation of 1D Gibbs states [151].

Lemma 25. *Under the setting of Theorem 22, there exists a constant $\xi \geq 0$ such that*

$$(1-p_i) \left\| \mathrm{Tr}_i \left(\rho^{H_{ABC}} \right) - \frac{\mathrm{Tr}_i \tilde{E} \left(\rho_{AB_l ... B_i}^{H_{ABC}} \right)}{1-p_i} \right\|_1 \leq e^{-\frac{l}{\xi}} . \quad (5.91)$$

Proof. Define a correlation function $\mathrm{Cor}(X:Y)_\rho$ of regions X and Y by

$$\mathrm{Cor}(X:Y)_\rho = \max_{\|M\|,\|N\|\leq 1} \left|\mathrm{Tr}\left[(M\otimes N)(\rho_{XY}-\rho_X\otimes\rho_Y)\right]\right| . \tag{5.92}$$

Consider some CP and trace-decreasing map $(\mathbb{1}_X\otimes\mathcal{E}_Y)(\rho_{XY})=\sum_i E_i\rho_{XY}E_i^\dagger$. By lemma 9 of Ref. [154], it holds that

$$\mathrm{Cor}(X:Y)_\rho \geq \mathrm{Tr}[M_Y\rho_{XY}]\|\rho_X-\sigma_X\|_1 , \tag{5.93}$$

where $M_Y=\sum_i E_i^\dagger E_i$ and

$$\sigma_X = \frac{1}{\mathrm{Tr}[M_Y\rho_{XY}]}\mathrm{Tr}_Y\left[M_Y\rho_{XY}\right] = \mathrm{Tr}_Y\frac{(\mathrm{id_X}\otimes\mathcal{E}_\mathrm{Y})(\rho_\mathrm{XY})}{\mathrm{Tr}(\mathrm{id_X}\otimes\mathcal{E}_\mathrm{Y})(\rho_\mathrm{XY})} . \tag{5.94}$$

It has been shown that any 1D Gibbs states with a short-range Hamiltonian have exponentially decaying $\mathrm{Cor}(X:Y)_\rho$ [151], i.e., there exist constants $c\geq 0$ and $\xi>0$ such that

$$\mathrm{Cor}(X:Y)_\rho \leq ce^{-d(X:Y)/\xi} . \tag{5.95}$$

Choosing $X=AB_l...B_{i+1}$, $Y=B_i$ and $\mathcal{E}_Y=\tilde{E}_i$ prove Lemma 25. □

Without loss of generality, let us assume

$$p_i = \mathrm{Tr}[\tilde{\Lambda}_i(\rho^{H_{ABC}}_{AB_l...B_i})] = p > 0 \tag{5.96}$$

for all i. Lemma 25 allows an iterative calculation. First we have

$$\begin{aligned}\Big\|\Lambda_{B\to BC}(\rho^{H_{ABC}}_{AB}) - \tilde{\Lambda}_1(\rho^{H_{ABC}}_{AB}) + (1-p)\tilde{\Lambda}_2(\rho^{H_{ABC}}_{AB_l...B_2}) \\ +(1-p)\left(\tilde{\Lambda}_3+\left(\cdots\left(\tilde{\Lambda}_l+\tilde{E}_l\right)\mathrm{Tr}_{l-1}\tilde{E}_{l-1}\right)\cdots\right)\mathrm{Tr}_2\tilde{E}_2(\rho^{H_{ABC}}_{AB_l...B_2})\Big\|_1 \leq e^{-l/\xi}\end{aligned} \tag{5.97}$$

Here, we used Lemma 25 for $\mathrm{Tr}_1\tilde{E}_1(\rho^{H_{ABC}}_{AB})$ and $\rho^{H_{ABC}}_{AB_l...B_2}$. Then we can obtain

$$\begin{aligned}\Big\|\Lambda_{B\to BC}(\rho^{H_{ABC}}_{AB}) - \tilde{\Lambda}_1(\rho^{H_{ABC}}_{AB}) + (1-p)\tilde{\Lambda}_2(\rho^{H_{ABC}}_{AB_l...B_2}) \\ +(1-p)^2\tilde{\Lambda}_3(\rho^{H_{ABC}}_{AB_l...B_3}) + (1-p)^2\left(\tilde{\Lambda}_4+\cdots\mathrm{Tr}_4\tilde{E}_4\right)\mathrm{Tr}_3\tilde{E}_3(\rho^{H_{ABC}}_{AB_l...B_3})\Big\|_1 \leq 2e^{-l/\xi} .\end{aligned} \tag{5.98}$$

We can proceed in a similar way, where at each ith step, we replace $\mathrm{Tr}_i\tilde{E}_i(\rho^{H_{ABC}}_{AB})$ by $\mathrm{Tr}_i(\rho^{H_{ABC}})$ by using the triangle inequality. After iterating $l-1$ steps, we obtain

$$\left\|\Lambda_{B\to BC}(\rho^{H_{ABC}}_{AB}) - \sum_{i=1}^{l}(1-p)^{i-1}\tilde{\Lambda}_i(\rho^{H_{ABC}}_{AB}) + (1-p)^{l-1}\tilde{E}_l(\rho^{H_{ABC}}_{AB_l})\right\|_1 \leq (l-1)e^{-l/\xi} . \tag{5.99}$$

Since $\left(\sum_{i=1}^{l} p(1-p)^{i-1}\right) + (1-p)^l = 1$, it follows

$$\rho^{H_{ABC}} = \sum_{i=1}^{l} p(1-p)^{i-1}\rho^{H_{ABC}} + (1-p)^l \rho^{H_{ABC}} \tag{5.100}$$

and thus

$$\left\| \rho^{H_{ABC}} - \sum_{i=1}^{l} (1-p)^{i-1} \tilde{\Lambda}_i(\rho_{AB}^{H_{ABC}}) + (1-p)^{l-1} \tilde{E}_l(\rho_{AB_l}^{H_{ABC}}) \right\|_1 \tag{5.101}$$

$$\leq \sum_{i=1}^{l} p(1-p)^{i-1} \left\| \rho^{H_{ABC}} - \frac{1}{p}\tilde{\Lambda}_i(\rho_{AB_l\ldots B_i}^{H_{ABC}}) \right\|_1 + (1-p)^l \left\| \rho^{H_{ABC}} - \frac{1}{1-p}\tilde{E}_l(\rho_{AB_l}^{H_{ABC}}) \right\|_1 \tag{5.102}$$

$$\leq \{1-(1-p)^l\} C_2(\beta) e^{-q_1(\beta)l} + 2(1-p)^l \,. \tag{5.103}$$

Therefore, by combining Eq. (5.99) and Eq. (5.103), we conclude

$$\|\rho^{H_{ABC}} - \Lambda_{B\to BC}(\rho_{AB}^{H_{ABC}})\|_1 \leq \{1-(1-p)^l\} C_2(\beta) e^{-q_1(\beta)l} + 2e^{-|\ln(1-p)|l} + (l-1)e^{-l/\xi} \tag{5.104}$$

$$\leq C_2(\beta) e^{-q_1(\beta)l} + 2e^{-|\ln(1-p)|l} + l e^{-l/\xi} \,. \tag{5.105}$$

Here, the probability p can be bounded as in Eq. (5.82), and thus we have

$$|\ln(1-p)| \geq \left| \ln\left(1 - \frac{(1-e^{-1})e^{-2\beta J}}{\left(1+e^{-\frac{\beta J}{2}}K(\beta)\right)^4}\right)\right| \geq \frac{(1-e^{-1})e^{-2\beta J}}{\left(1+e^{-\frac{\beta J}{2}}K(\beta)\right)^4} = e^{-\Theta(\beta)} \,, \tag{5.106}$$

where the last inequality follows from $\log(1-x) \leq -x$ for any $x \in [0,1]$. If $\xi = e^{\mathcal{O}(\beta)}$, Eq. (5.105) can be bounded by

$$2C_2(\beta) l e^{-q'(\beta)l} = e^{-q'(\beta)l + \ln(2C_2(\beta)l)} \,, \tag{5.107}$$

where $q'(\beta) = \Omega(e^{-\Theta(\beta)})$. Since $d(A,C) = 3l^2 - l$, Eq. (5.107) is $\Omega(e^{-\Theta(\sqrt{d(A,C)})})$. Therefore, for sufficiently large l, there exists a constant $q(\beta) = \Omega(e^{-\Theta(\beta)})$ such that $e^{-q'(\beta)l + \ln(2C_2(\beta)l)} \leq e^{-q(\beta)\sqrt{d(A,C)}}$.

Note that the all arguments are symmetric under the exchange of A and C. Thus, there exists a recovery map from B to AB with a similar accuracy as well.

5.3.3 The proof of Corollary 19

Let us first consider a 1D open spin chain with a tripartition ABC so that a simply connected region B shields A from C. Then, $d(A,C) = |B|$. Divide C into $C = C_1 \cup C_2 \cup ... \cup C_m$, where m is the maximum number such that $|C_i| = |B|$ for $1 \leq i < m$ and each C_i shields C_{i-1} from C_{i+1} (here, $C_0 \equiv B$).

Theorem 18 and the Fannes inequality imply

$$I(A : C_i|BC_1 \dots C_{i-1})_{\rho^{H_{ABC}}} \leq |B| e^{-q(\beta)\sqrt{i|B|}} \tag{5.108}$$

for any $i \in [1, m]$. By the chain rule

$$I(A : C|B) = I(A : C_1|B) + I(A : C_2|BC_1) + \ldots + I(A : C_m|BC_1 \dots C_{m-1}), \tag{5.109}$$

we have

$$\begin{aligned} I(A : C|B)_{\rho^{H_{ABC}}} &\leq \sum_{i=1}^{m} |B| e^{-q(\beta)\sqrt{i|B|}} \\ &\leq \left(|B| e^{-q(\beta)\sqrt{|B|}} + |B| \sum_{i=1}^{m-1} e^{-q(\beta)\sqrt{|B|(i+1)}} \right) \\ &\leq \left(|B| e^{-c\sqrt{|B|}} + \int_1^{\infty} e^{-q(\beta)\sqrt{|B|x}} dx \right) \\ &= \left(1 + \frac{2(1 + q(\beta)\sqrt{|B|})}{q(\beta)^2 |B|} \right) |B| e^{-q(\beta)\sqrt{|B|}}. \end{aligned} \tag{5.110}$$

Again, the upper bound is $e^{-\Theta(\sqrt{d(A,C)})}$. The same strategy works for a more complicated tripartition ABC of both 1D open chains and closed chains.

5.3.4 The proof of Corollary 21

The original proof of this corollary is first given by Brandao in Ref. [155]. We first divide the whole 1D spin chain into consecutive regions $A_1B_1C_1A_2B_2C_2...A_kB_kC_k$, where we choose $|A_i| = |B_j| = l \geq l_0$ and $|C_i| = 5\xi(\ln d)l$ for all i, where l_0 is the constant given in Theorem 18 and the correlation length ξ is given in Eq. (5.95). Let us consider region $(A_1B_1C_1...C_{i-1}A_iB_{i+1})(B_iA_{i+1})C_i$, where B_iA_{i+1} shields $A_1B_1...B_{i+1}$ from C_i. From Theorem 18, there exists a CPTP-map $\delta_i : \mathcal{S}(\mathcal{H}_{B_iA_{i+1}}) \to \mathcal{S}(\mathcal{H}_{B_iC_iA_{i+1}})$ such that

$$\left\| \Delta_i \left(\rho^H_{A_1B_1...A_iB_iA_{i+1}B_{i+1}C_{i+1}} \right) - \rho^H_{A_1B_1...A_iB_iC_iA_{i+1}B_{i+1}C_{i+1}} \right\|_1 \leq C e^{-q(\beta)\sqrt{l}}. \tag{5.111}$$

Since the Gibbs state has exponentially decaying correlations, after tracing out C_i, the two remained connected components are almost uncorrelated. By using Lemma 20 of Ref. [154], it follows that

$$\begin{aligned}
&\left\|\rho_{A_1B_1C_1\ldots B_{i-1}A_iB_i} - \rho_{A_1B_1C_1\ldots B_{i-1}} \otimes \rho_{A_iB_i}\right\|_1 \\
&\quad \leq (\dim \mathcal{H}_{A_iB_i})^2\, Cor(A_1B_1C_1\ldots B_{i-1} : A_iB_i)_{\rho^H} && (5.112)\\
&\quad \leq d^{4l} e^{-5\xi(\ln d)l/\xi} && (5.113)\\
&\quad = e^{-(\ln d)l}\,. && (5.114)
\end{aligned}$$

Each Δ_i acts on different sets of spins and therefore does not overlap. Then we have

$$\begin{aligned}
&\|\Delta_1 \otimes \cdots \otimes \Delta_k(\rho_{A_1B_1} \otimes \cdots \otimes \rho_{A_kB_k}) - \rho_{A_1B_1C_1\cdots A_kB_kC_k}\|_1 \\
&\leq \|\Delta_1 \otimes \cdots \otimes \Delta_k(\rho_{A_1B_1} \otimes \cdots \otimes \rho_{A_kB_k}) \\
&\quad -\Delta_2 \otimes \cdots \otimes \Delta_k(\rho_{A_1B_1C_1A_2B_2} \otimes \rho_{A_3B_3} \otimes \cdots \otimes \rho_{A_kB_k})\|_1 \\
&\quad + \|\Delta_2 \otimes \cdots \otimes \Delta_k(\rho_{A_1B_1C_1A_2B_2} \otimes \rho_{A_3B_3} \otimes \cdots \otimes \rho_{A_kB_k}) - \rho_{A_1B_1C_1\ldots A_kB_kC_k}\|_1 && (5.115)\\
&\leq \|\Delta_1(\rho_{A_1B_1} \otimes \rho_{A_2B_2}) - \rho_{A_1B_1C_1A_2B_2}\|_1 \\
&\quad + \|\Delta_2 \otimes \cdots \otimes \Delta_k(\rho_{A_1B_1C_1A_2B_2} \otimes \rho_{A_3B_3} \otimes \cdots \otimes \rho_{A_kB_k}) - \rho_{A_1B_1C_1\cdots A_kB_kC_k}\|_1 && (5.116)\\
&\leq \|\Delta_1(\rho_{A_1B_1A_2B_2}) - \rho_{A_1B_1C_1A_2B_2}\|_1 + e^{-(\ln d)l} \\
&\quad + \|\Delta_2 \otimes \cdots \otimes \Delta_k(\rho_{A_1B_1C_1A_2B_2} \otimes \rho_{A_3B_3} \otimes \cdots \otimes \rho_{A_kB_k}) - \rho_{A_1B_1C_1\cdots A_kB_kC_k}\|_1 && (5.117)\\
&\leq 2Ce^{-q(\beta)\sqrt{l}} + \|\Delta_2 \otimes \cdots \otimes \Delta_k(\rho_{A_1B_1C_1A_2B_2} \otimes \cdots \otimes \rho_{A_kB_k}) - \rho_{A_1B_1C_1\cdots A_kB_kC_k}\|_1. && (5.118)
\end{aligned}$$

The first inequality follows from the triangle inequality, the second from the monotonicity of the trace norm under quantum operations, the third from Eq. (5.114), and the fourth from Eq. (5.111) and $e^{-(\ln d)l} \leq Ce^{-q(\beta)\sqrt{l}}$.

Iterating the argument above, we find

$$\|\Delta_1 \otimes \cdots \otimes \Delta_k(\rho_{A_1B_1} \otimes \ldots \otimes \rho_{A_kB_k}) - \rho_{A_1B_1C_1\ldots A_kB_kC_k}\|_1 \leq 2kCe^{-q(\beta)\sqrt{l}}\,. \tag{5.119}$$

Since $k \leq n$, choosing $l = \mathcal{O}(\log^2(n/\varepsilon))$ gives an error bounded by ε. We denote a CPTP-map which construct $\rho_{A_iB_i}$ by $\Delta_{1,i}$ and relabel Δ_i in the above by $\Delta_{2,i}$. The CPTP-map $\bigotimes_i \Delta_{1,i}$ creates product state of the form of $\rho_{A_1B_1} \otimes \rho_{A_2B_2} \otimes \cdots \otimes \rho_{A_kB_k}$, and then $\bigotimes_i \Delta_{2,i}$ approximately creates the target state from this product state.

5.3.5 Extension to more general graphs

Our proof of Theorem 18 for 1D spin chains can be generalized to more general graphs with appropriate partitions. For instance, let us consider a tree graph $G = (E, V)$ with a partition ABC as depicted in Fig. 5.7. Since G is a tree, it is guaranteed that there is a unique path connecting A and C. Then, all spins in B are classified as (i) spins belonging to the unique path (ii) the descendants of spins on the path (iii) the rest spins which are separated from the path. From this property of tree graphs, we can obtain a coarse-grained 1D chain by regarding each spin on the path and its descendants as a single system, and removing all spins in (iii). We can then apply the proof presented in the previous section to this situation. Note that the norm of an interaction term connecting spins on the path is irrelevant to the size of the coarse grained spins.

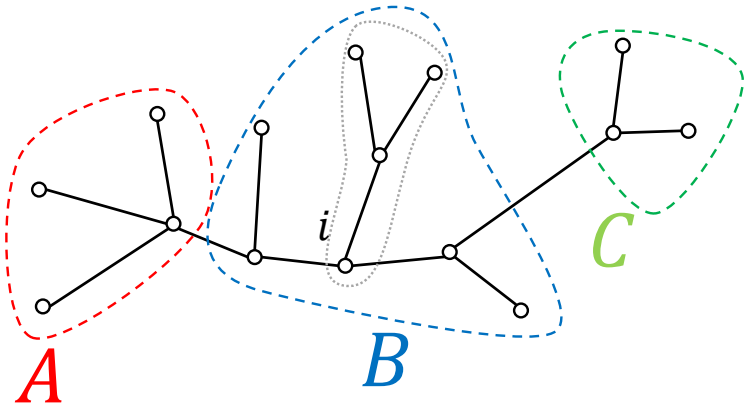

Figure 5.7: An example of the region with a partition ABC for a tree graph. In the coarse-grain procedure, spin i and its descendants (the spins in the gray dotted region) can be regarded as one large system.

An important point of this argument is that the success probability of the recovery map in Lemma 24 is also bounded by a constant of $d(A, C)$ in the case of tree graphs. In general cases, we can consider a partition ABC which cannot be reduced to 1D systems, such as depicted in Fig. 5.8. Remember that the success probability p is in the order $\Omega(e^{\beta\|H_{BM}\|})$. In the case of the situation of Fig. 5.8, H_{BM} is the sum of all interactions along the perimeter of B^L and the strength $\|H_{BM}\|$ is proportional to l. When considering the repeat-until-success method, the success probability decays too rapidly, and therefore our strategy does not work. Because of this reason, we need an alternative way to proof the approximate Markov property of Gibbs states on 2D or higher-dimensional lattice systems.

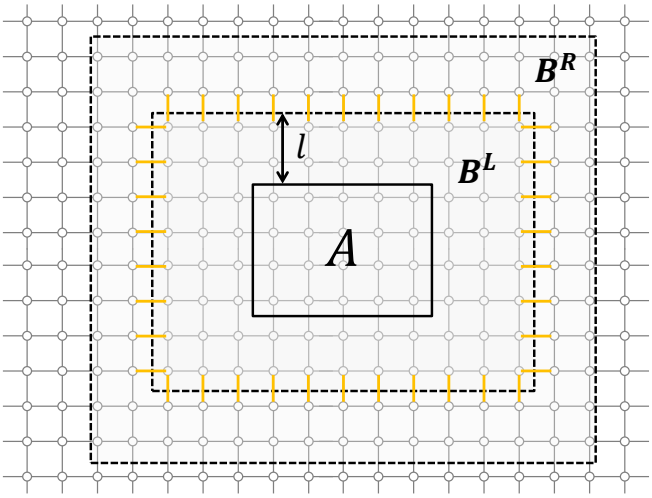

Figure 5.8: An example of the systems of which the partition is not reducible to 1D. When the system is defined on a 2D lattice, $\|H_{B^M}\|$ is proportional to l, since it contains all interactions between B^L and B^R (the interactions corresponding to the orange lines in the case of a nearest-neighbor Hamiltonian).

5.4 Proof: Approximate Markov Chains are 1D Gibbs States

We restate Theorem 20:

Theorem 20 *Let $\rho_{A_1,\ldots,A_n}$ be an ε-approximate Markov chain. Then there exists a short-range Hamiltonian $H = \sum_i h_{A_i,A_{i+1}}$, with $h_{A_i,A_{i+1}}$ only acting on A_iA_{i+1}, such that*

$$S\left(\rho \middle\| \frac{e^{-H}}{\mathrm{Tr}e^{-H}}\right) \leq \varepsilon n.$$

Proof. We first consider a 1D open spin chain. Let $\sigma_{A_1\ldots A_n}$ be the maximum entropy state such that

$$\sigma_{A_iA_{i+1}} = \rho_{A_iA_{i+1}} \tag{5.120}$$

for all i. $\sigma_{A_1\ldots A_n}$ is an element of $\overline{\mathcal{E}}^{rI}(\mathcal{A})$, where $\mathcal{A}$ is the set of all observables or all elements of the orthonormal basis of observables on $\{A_iA_{i+1}\}$. By definition, $\mathcal{E}(\mathcal{A})$ is the set of Gibbs states with Hamiltonians in the form $H = \sum_i h_{A_iA_{i+1}}$. We will show

$$S(\sigma) \leq S(\rho) + \varepsilon n. \tag{5.121}$$

The result then follows from Eq. (2.62) and the fact that $\inf_{\omega\in\mathcal{E}(\mathcal{A})} S(\sigma\|\omega) = 0$.

By strong subadditivity we find

$$\begin{aligned} S(A_1 \dots A_n)_\sigma &\leq S(A_1A_2)_\sigma - S(A_2)_\sigma + S(A_2 \dots A_n)_\sigma \\ &\leq S(A_1A_2)_\sigma - S(A_2)_\sigma + S(A_2A_3)_\sigma - S(A_3)_\sigma + S(A_3 \dots A_n)_\sigma \\ &\dots \\ &\leq \sum_{i=1}^{n-1} S(A_iA_{i+1})_\sigma - S(A_{i+1})_\sigma \\ &= \sum_{i=1}^{n-1} S(A_iA_{i+1})_\rho - S(A_{i+1})_\rho. \end{aligned} \tag{5.122}$$

The last equality follows from Eq. (5.120). Since ρ is an ε-approximate Markov chain, for every $i \in [1, n-1]$,

$$I(A_i : A_{i+2}...A_n|A_{i+1})_\rho \leq \varepsilon, \tag{5.123}$$

which can be written as

$$S(A_i....A_n) \geq S(A_iA_{i+1}) - S(A_{i+1}) + S(A_{i+1}...A_n) - \varepsilon, \tag{5.124}$$

i.e., the strong subadditivity is saturated up to error ε. Therefore we obtain

$$\begin{aligned} \sum_{i=1}^{n-1} S(A_iA_{i+1})_\rho - S(A_{i+1})_\rho &\leq \sum_{i=1}^{n-2} S(A_iA_{i+1})_\rho - S(A_{i+1})_\rho + S(A_{n-2}A_{n-1}A_n) + \varepsilon \\ &\vdots \\ &\leq S(A_1A_2)_\rho - S(A_2)_\rho + S(A_2 \dots A_n)_\rho + (n-2)\varepsilon \\ &\leq S(A_1...A_n)_\rho - (n-1)\varepsilon. \end{aligned} \tag{5.125}$$

Combining Eq. (5.122) and Eq. (5.125) we have

$$S(\rho\|\sigma) = S(\sigma) - S(\rho) \leq \varepsilon(n-1)\,. \tag{5.126}$$

By the discussion considered in Sec. 2.3, there exists a Gibbs state

$$\omega = \frac{1}{Z} \exp\left(- \sum_i H_{A_iA_{i+1}} \right) \in \mathcal{E}(\mathcal{A}) \tag{5.127}$$

which satisfies

$$S(\sigma\|\omega) \leq \varepsilon\,. \tag{5.128}$$

Using the Pythagorean theorem, we obtain

$$S(\rho\|\omega) = S(\rho\|\sigma) + S(\sigma\|\omega) \leq n\varepsilon\,, \tag{5.129}$$

which completes the proof. The above proof can be applied to the case of periodic boundary condition as well. In this case, we set $\mathcal{A}$ as the set of all observables on $\{A_nA_1A_2, A_2A_3, ..., A_{n-1}A_n\}$ and the Hamiltonian H contains a next-nearest-neighbor interaction. □

5.5 Concluding Remarks

The approximate Markov property of 1D Gibbs states implies that the Gibbs state can be constructed by first preparing the reduced states on distinct blocks, and then patching them by locally performing quantum operations. It has been shown that there exists an efficient quantum Gibbs sampling algorithm if the target Gibbs state satisfies the approximate Markov property and another property called uniform clustering [82]. It is known that any 1D Gibbs state satisfies the later condition [151] and our preparation method can be understood as a special case of this algorithm for 1D systems.

The area law of the mutual information guarantees that any Gibbs state of a short-range Hamiltonian fullfills the approximate Markov property in any dimension (even there are various phases in two or higher dimensions). However, the accuracy of the approximate Markov property of Gibbs states in systems with dimension greater than 1 has not been derived. Unfortunately, as we have discussed in Sec. 5.3.5, our techniques cannot be generalized to higher-dimensional systems. Moreover, our bound on the conditional mutual information may not be tight. New mathematical tools are required to address these issues.

One difficulty in higher-dimension originates from the existence of another type of the Markov property. In 1D systems, it is sufficient to consider the approximate Markov property for a tripartition ABC of the whole system, as it ensures the conditional mutual information of a reduced state to be small as well (Fig. 5.9(a)). This is because of that a possible patterns of tripartitions are very restricted in 1D systems. However, this is not the case in higher-dimensional systems (Fig. 5.9(b)). This difficulty also arises when we consider characterization of states on a general graph satisfying the approximate Markov property in terms of Gibbs states.

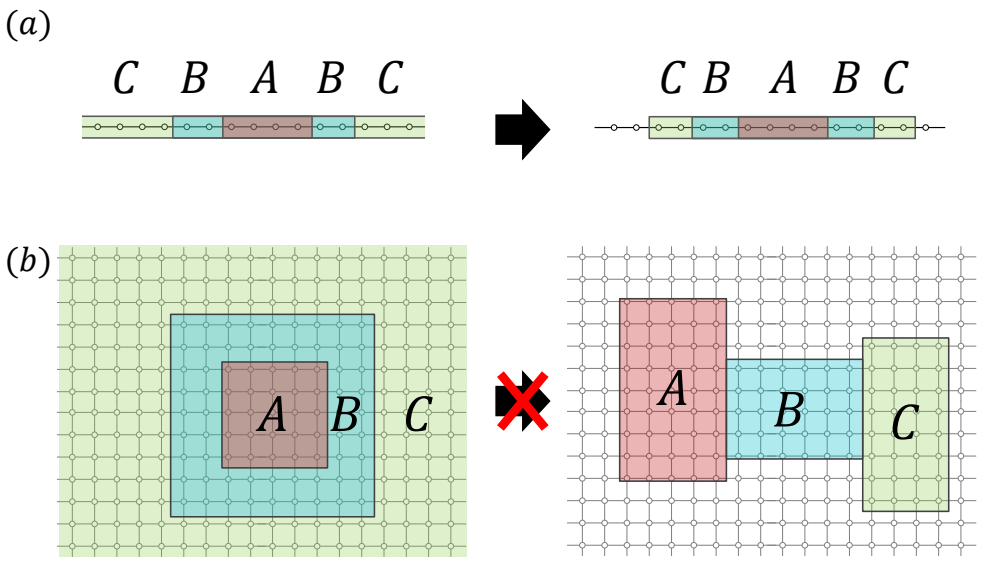

Figure 5.9: Correspondence of the Markov property of the tripartite partitions for the whole system (the left hand side) and a reduced system (the right hand side). (a): In 1D systems, the approximate Markov property for the partition in the left hand side implies that $I(A:C|B)$ is also small for the partition in the right hand side. (b): In higher-dimensional systems, the approximate Markov property for the partition in the left hand side does not imply that $I(A:C|B)$ is small for the partition in the right hand side.

Chapter 6

Conclusion

In this thesis, we have studied generic properties of multipartite correlations in 2D gapped ground states (Chapter 4) and 1D Gibbs states (Chapter 5). We summarize our results obtained in each chapter and present concluding remarks.

6.1 Summary of Results

In Chapter 4, we have studied multipartite correlations in gapped ground states in 2D systems. To characterize multipartite correlations in topologically ordered phases, we have focused on two measures of multipartite correlation: the irreducible correlation which has a geometrical interpretation, and the optimal rate of a secret sharing protocol which provides an operational significance. We have considered gapped ground states with the zero correlation length, and also analyzed the case of finite correlation length.

First, we have calculated the irreducible correlation of each order for regions with certain partitions and have derived that the highest-order irreducible correlation coincides with the TEE of the state. This statement has been conjectured in Refs. [74, 75] for gapped ground states, and we have provided a rigorous proof in the limit of zero correlation length. The second-order irreducible correlation can be written as the sum of the mutual information of bipartite subsystems, and the other orders the irreducible correlations are zero. That is, a non-zero TEE implies that the entanglement Hamiltonian of the reduced state on the region contains genuinely multipartite interactions in addition to bipartite interactions. As an application of this result, we have shown that the value of the TEE provides a restriction on the possible form of the ES on a cylinder. A similar restriction on the ES had been found in the PEPS formalism [131, 132], but our result generalizes previous observations and provides a connection to the value of the TEE.

Second, we have considered a particular type of secret sharing protocols by

regarding the reduced state of a ground state as a resource. We have shown that the optimal rate of the secret sharing protocol coincides with the TEE of the state and thus we have discovered an operational interpretation of the TEE. We have also explicitly demonstrated how to encode secrets in the optimal way for the toric code.

Our results indicate that, as long as gapped ground states satisfy the conditions of zero correlation length, the TEE is an information-theoretically meaningful measure of multipartite correlations. Moreover, if the correction due to the finite correlation length decays sufficiently fast, we can assert a geometrical interpretation to the TEE beyond the ideal situation. We emphasize that the two measures are defined for general quantum states and thus applicable for more general settings in contrast to the TEE, such as many-body states at finite temperature or ground states of long-range Hamiltonians.

In Chapter 5, we have studied the Markov property of 1D Gibbs states and their relation to approximate Markov chains. We have derived that any 1D Gibbs state of a short-range Hamiltonian has sub-exponentially decaying conditional mutual information $I(A:C|B)$ with respect to the distance between regions A and C, and therefore they are approximate Markov chains with a good accuracy. The proof is based on the technique of quantum belief propagation [152, 47], which states that a local perturbation on a Hamiltonian disturbs the corresponding Gibbs states only locally. The decay rate of the conditional mutual information describes how fast the mutual information of a 1D Gibbs state saturates the bound given by the area law. We have also shown that the approximate Markov property can be utilized to show the existence of an efficient preparation method of 1D Gibbs state in a constant-time. In addition, we have shown that any approximate Markov chain can be well-approximated by a Gibbs state of a short-range Hamiltonian. This result provides a characterization of approximate Markov chain in terms Gibbs states, along with the previous result by using recovery maps [78]. The result has been derived by the equivalence between Gibbs states and the maximum entropy states which is also used in Chapter 4.

By combining our results, we have shown that 1D Gibbs states of short-range Hamiltonians and approximate Markov chains are approximately equivalent. This statement forms a generalization of the quantum Hammersley-Clifford theorem [80, 81] for 1D systems.

6.2 Concluding Remarks of This Thesis

In this thesis, we have derived generic properties of multipartite correlations in quantum many-body states obeying an area law. These generic properties provide characterizations of classes of states or phases beyond bipartite correlations. An

important idea behind our results on generic properties of multipartite correlations is that when a state obeys an area law of entanglement or the mutual information, the reduced states on certain regions form (approximate) Markov chains. We have demonstrated properties of Markov chains and approximate Markov chains, which have been intensively investigated in quantum and classical information theory, are useful theoretical tools to derive and understand generic properties of quantum many-body systems independent of details of interactions or physical models.

To analyze complex multipartite correlations, we have used the classification based on the interaction patterns of Hamiltonians in the expression of quantum states as effective Gibbs states. Our results affirm that this classification method, starting from Amari's work presented in Ref. [66], has an affinity with analysis of entanglement Hamiltonian and Gibbs states. We believe that our results contribute to understanding of properties of multipartite correlations in quantum systems by bridging condensed matter physics and quantum information theory.

Bibliography

[1] P. W. Shor. Algorithms for quantum computation: discrete logarithms and factoring. *Proc. 35th Annu. Symp. Found. Comput. Sci.*, 124–134, 1994.

[2] C. H. Bennett and G. Brassard. Quantum cryptography: Public key distribution and coin tossing. In *Proc. IEEE Int. Conf. Comput. Syst. Signal Process.*, 175–179, 1984.

[3] C. H. Bennett. Quantum cryptography: Uncertainty in the service of privacy. *Science*, 257(5071):752–753, 1992.

[4] A. K. Ekert. Quantum cryptography based on Bell's theorem. *Phys. Rev. Lett.*, 67(6):661–663, 1991.

[5] R. L. Rivest, A. Shamir, and L. Adleman. A Method for Obtaining Digital Signatures and Public-key Cryptosystems. *Commun. ACM*, 21(2):120–126, 1978.

[6] S. L. Braunstein and C. M. Caves. Statistical distance and the geometry of quantum states. *Phys. Rev. Lett.*, 72(22):3439–3443, 1994.

[7] P. Kok, S. L. Braunstein, and Jonathan P Dowling. Quantum lithography, entanglement and Heisenberg-limited parameter estimation. *J. Opt. B*, 6(8):S811, 2004.

[8] C. H. Bennett, G. Brassard, R. Jozsa, A. Peres, and W. K. Wootters. Teleporting an Unkown Quantum State via Dual Classical and Einstein-Podolsky-Rosen Channels. *Phys. Rev. Lett.*, 70(12):1895–1899, 1993.

[9] S. Popescu and D. Rohrlich. Quantum nonlocality as an axiom. *Found. Phys.*, 24(3):379–385, 1994.

[10] M. Pawłowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Zukowski. Information causality as a physical principle. *Nature*, 461(7267):1101–1104, 2009.

[11] S. Ryu and T. Takayanagi. Holographic derivation of entanglement entropy from the anti-de sitter space/conformal field theory correspondence. *Phys. Rev. Lett.*, 96(18):181602, 2006.

[12] Y. Sekino and L. Susskind. Fast scramblers. *J. High Energy Phys.*, 2008(10):065–065, 2008.

[13] F. Pastawski, B. Yoshida, D. Harlow, and J. Preskill. Holographic quantum error-correcting codes: toy models for the bulk/boundary correspondence. *J. High Energy Phys.*, 2015(6):149–, 2015.

[14] M. Sarovar, A. Ishizaki, G. R. Fleming, and K. B. Whaley. Quantum entanglement in photosynthetic light-harvesting complexes. *Nat. Phys.*, 6(6):462–467, 2010.

[15] I. Kassal, J. D. Whitfield, A. Perdomo-Ortiz, M.-H. Yung, and A. Aspuru-Guzik. Simulating Chemistry Using Quantum Computers. *Annu. Rev. Phys. Chem.*, 62(1):185–207, 2011.

[16] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghì. Canonical typicality. *Phys. Rev. Lett.*, 96(5):050403, 2006.

[17] S. Popescu, A. J. Short, and A. Winter. Entanglement and the foundations of statistical mechanics. *Nat. Phys.*, 2(11):754–758, 2006.

[18] G. D. L. Cuevas. A quantum information approach to statistical mechanics. *J. Phys. B At. Mol. Opt. Phys.*, 46(24):243001, 2013.

[19] L. Amico, R. Fazio, A. Osterloh, and V. Vedral. Entanglement in many-body systems. *Rev. Mod. Phys.*, 80(2):517–576, 2008.

[20] N. Laflorencie. Quantum entanglement in condensed matter systems. *Physics Report*, 643:1–59, 2016.

[21] J. D. Bekenstein. Black holes and entropy. *Phys. Rev. D*, 7(8):2333–2346, 1973.

[22] P. Hayden, D. W. Leung, and A. Winter. Aspects of generic entanglement. *Commun. Math. Phys.*, 265(1):95–117, 2006.

[23] M. B. Hastings. An Area Law for One Dimensional Quantum Systems. *J. Stat. Mech. Theory Exp.*, 08024(2):9, 2007.

[24] D. Aharonov, I. Arad, Z. Landau, and U. Vazirani. The 1D area law and the complexity of quantum states: A combinatorial approach. *Proc. - Annu. IEEE Symp. Found. Comput. Sci. FOCS*, 324–333, 2011.

[25] G. Vidal. Efficient simulation of one-dimensional quantum many-body systems. *Phys. Rev. Lett.*, 93(4):040502, 2004.

[26] I. Affleck, T. Kennedy, E. H. Lieb, and H. Tasaki. Rigorous results on valence-bond ground states in antiferromagnets. *Phys. Rev. Lett.*, 59(7):799–802, 1987.

[27] S. R. White. Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.*, 69(19):2863–2866, 1992.

[28] U. Schollwöck. The density-matrix renormalization group in the age of matrix product states. *Ann. Phys. (N. Y).*, 326(1):96–192, 2011.

[29] G. Vidal, J. I. Latorre, E. Rico, and A. Kitaev. Entanglement in quantum critical phenomena. *Phys. Rev. Lett.*, 90(22):227902, 2002.

[30] B. Q. Jin and V. E. Korepin. Quantum spin chain, Toeplitz determinants and the Fisher-Hartwig conjecture. *J. Stat. Phys.*, 116(1-4):79–95, 2004.

[31] J. P. Keating and F. Mezzadri. Entanglement in quantum spin chains, symmetry classes of random matrices, and conformal field theory. *Phys. Rev. Lett.*, 94(5):050501, 2005.

[32] C. Callan and F. Wilczek. On geometric entropy. *Phys. Lett. B*, 333(1-2):55–61, 1994.

[33] C. Holzhey, F. Larsen, and F. Wilczek. Geometric and renormalized entropy in conformal field theory. *Nucl. Physics, Sect. B*, 424(3):443–467, 1994.

[34] G. Vidal. Class of Quantum Many-Body States That Can Be Efficiently Simulated. *Phys. Rev. Lett.*, 101(11):110501, 2008.

[35] M. B. Plenio, J. Eisert, J. Dreißig, and M. Cramer. Entropy, entanglement, and area: Analytical results for harmonic lattice systems. *Phys. Rev. Lett.*, 94(6):060503, 2005.

[36] S. Michalakis and J. P. Zwolak. Stability of Frustration-Free Hamiltonians. *Commun. Math. Phys.*, 322(2):277–302, 2013.

[37] B. Swingle. A simple model of many-body localization. *e-print*, arXiv:1307.0507, 2013.

[38] F. G. S. L. Brandão and M. Cramer. Entanglement area law from specific heat capacity. *Phys. Rev. B*, 92(11):115134, 2015.

[39] F. Verstraete and J. I. Cirac. Renormalization algorithms for Quantum-Many Body Systems in two and higher dimensions. *e-print*, arXiv:cond-mat/0407066, 2004.

[40] F. Verstraete, M. M. Wolf, D. Perez-Garcia, and J. I. Cirac. Criticality, the area law, and the computational power of projected entangled pair states. *Phys. Rev. Lett.*, 96(22):220601, 2006.

[41] T. Barthel, M. C. Chung, and U. Schollwöck. Entanglement scaling in critical two-dimensional fermionic and bosonic systems. *Phys. Rev. A*, 74(2):022329, 2006.

[42] D. Gioev and I. Klich. Entanglement entropy of fermions in any dimension and the widom conjecture. *Phys. Rev. Lett.*, 96(10):100503, 2006.

[43] E. Fradkin and J. E. Moore. Entanglement entropy of 2D conformal quantum critical points: Hearing the shape of a quantum drum. *Phys. Rev. Lett.*, 97(5):050404, 2006.

[44] M. M. Wolf, F. Verstraete, M. B. Hastings, and J. I. Cirac. Area laws in quantum systems: Mutual information and correlations. *Phys. Rev. Lett.*, 100(7):070502, 2008.

[45] P. Hayden, R. Jozsa, D. Petz, and A. Winter. Structure of states which satisfy strong subadditivity of quantum entropy with equality. *Commun. Math. Phys.*, 246(2):359–374, 2004.

[46] D. Poulin and M. B. Hastings. Markov entropy decomposition: A variational dual for quantum belief propagation. *Phys. Rev. Lett.*, 106(8):080403, 2011.

[47] I. H. Kim. Perturbative analysis of topological entanglement entropy from conditional independence. *Phys. Rev. B*, 86(24):245116, 2012.

[48] I. H. Kim. Long-range entanglement is necessary for a topological storage of quantum information. *Phys. Rev. Lett.*, 111(8):080503, 2013.

[49] I. H. Kim and B. J. Brown. Ground-state entanglement constrains low-energy excitations. *Phys. Rev. B*, 92(11):115139, 2015.

[50] A. Y. Kitaev. Fault-tolerant quantum computation by anyons. *Ann. Phys. (N. Y).*, 303(1):2–30, 2003.

[51] M. H. Freedman, M. Larsen, and Z. Wang. A Modular Functor Which is Universalfor Quantum Computation. *Commun. Math. Phys.*, 227(3):605–622, 2002.

[52] R. Orús, T. C. Wei, O. Buerschaper, and M. V. Den Nest. Geometric entanglement in topologically ordered states. *New J. Phys.*, 16:013015, 2014.

[53] A. Hamma, R. Ionicioiu, and P. Zanardi. Ground state entanglement and geometric entropy in the Kitaev model. *Phys. Lett. A*, 337(1-2):22–28, 2005.

[54] M. Levin and X. G. Wen. Detecting topological order in a ground state wave function. *Phys. Rev. Lett.*, 96(11):110405, 2006.

[55] A. Kitaev and J. Preskill. Topological entanglement entropy. *Phys. Rev. Lett.*, 96(11):110404, 2006.

[56] S. Furukawa and G. Misguich. Topological entanglement entropy in the quantum dimer model on the triangular lattice. *Phys. Rev. B*, 75(21):214407, 2007.

[57] M. Haque, O. Zozulya, and K. Schoutens. Entanglement entropy in fermionic laughlin states. *Phys. Rev. Lett.*, 98(6):060401, 2007.

[58] S. V. Isakov, M. B. Hastings, and R. G. Melko. Topological Entanglement Entropy of a Bose-Hubbard Spin Liquid. *Nat. Phys.*, 7(10):772–775, 2011.

[59] S. Depenbrock, I. P. McCulloch, and U. Schollwöck. Nature of the spin-liquid ground state of the S=1/2 Heisenberg model on the kagome lattice. *Phys. Rev. Lett.*, 109(6):067201, 2012.

[60] W. Li, A. Weichselbaum, and J. V. Delft. Identifying symmetry-protected topological order by entanglement entropy. *Phys. Rev. B*, 88(24):902–905, 2013.

[61] W. J. McGill. Multivariate information transmission. *Psychometrika*, 19(2):97–116, 1954.

[62] K. Krippendorff. Information of interactions in complex systems. *Int. J. Gen. Syst.*, 38(6):669–680, 2009.

[63] L. Leydesdorff. Redundancy in systems which entertain a model of themselves: interaction information and the self-organization of anticipation. *Entropy*, 12(1):63–79, 2010.

[64] C. Castelnovo and C. Chamon. Entanglement and topological entropy of the toric code at finite temperature. *Phys. Rev. B*, 76(18):184442, 2007.

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

[66] S. I. Amari. Information geometry on hierarchy of probability distributions. *IEEE Trans. Inf. Theory*, 47(5):1701–1711, 2001.

[67] E. Schneidman, S. Still, M. J. Berry, and W. Bialek. Network information and connected correlations. *Phys. Rev. Lett.*, 91(23):238701, 2003.

[68] E. T. Jaynes. Information theory and statistical mechanics. *Phys. Rev.*, 106(4):620–630, 1957.

[69] N. Linden, S. Popescu, and W. K. Wootters. Almost Every Pure State of Three Qubits Is Completely Determined by Its Two-Particle Reduced Density Matrices. *Phys. Rev. Lett.*, 89(20):207901, 2002.

[70] D. L. Zhou. Irreducible multiparty correlations in quantum states without maximal rank. *Phys. Rev. Lett.*, 101(18):180505, 2008.

[71] S. Weis and A. Knauf. Entropy distance: New quantum phenomena. *J. Math. Phys.*, 53(10), 2012.

[72] S. Weis. Continuity of the Maximum-Entropy Inference. *Commun. Math. Phys.*, 330(3):1263–1292, 2014.

[73] S. Weis, A. Knauf, N. Ay, and M.-J. Zhao. Maximizing the Divergence from a Hierarchical Model of Quantum States. *Open Syst. & Inf. Dyn.*, 22(01):1550006, 2015.

[74] J. Chen, Z. Ji, C. K. Li, Y. T. Poon, Y. Shen, N. Yu, B. Zeng, and D. Zhou. Discontinuity of maximum entropy inference and quantum phase transitions. *New J. Phys.*, 17(8):083019, 2015.

[75] Y. Liu, B. Zeng, and D. L. Zhou. Irreducible many-body correlations in topologically ordered systems. *New J. Phys.*, 18(2):023024, 2016.

[76] M. A. Levin and X.-G. Wen. String-net condensation: A physical mechanism for topological phases. *Phys. Rev. A*, 71(4),045110 2005.

[77] B. Ibinson, N. Linden, and A. Winter. Robustness of quantum Markov chains. *Commun. Math. Phys.*, 277(2):289–304, 2008.

[78] O. Fawzi and R. Renner. Quantum Conditional Mutual Information and Approximate Markov Chains. *Commun. Math. Phys.*, 340(2):575–611, 2015.

[79] J. M. Hammersley and P. Clifford. Markov Fields on Finite Graphs and Lattices, available at http://www.statslab.cam.ac.uk/˜grg/books/hammfest/hamm-cliff.pdf, 1971.

[80] M. S. Leifer and D. Poulin. Quantum Graphical Models and Belief Propagation. *Ann. Phys. (N. Y).*, 323(8):1899–1946, 2008.

[81] W. Brown and D. Poulin. Quantum Markov Networks and Commuting Hamiltonians. *e-print*, arXiv:1206.0755, 2012.

[82] F. G. S. L. Brandao and M. J. Kastoryano. Finite correlation length implies efficient preparation of quantum thermal states. *e-print*, arXiv:1609.07877, 2016.

[83] S. L. Braunstein and P. van Loock. Quantum information with continuous variables. *Rev. Mod. Phys.*, 77(2):513, 2005.

[84] W F. Stinespring. Positive functions on C*-algebras. *Proc. Am. Math. Soc.*, 6(2):211–216, 1955.

[85] B. Schumacher. Quantum coding. *Phys. Rev. A*, 51(4):2738–2747, 1995.

[86] M. S. Leifer and R. W. Spekkens. Towards a formulation of quantum theory as a causally neutral theory of Bayesian inference. *Phys. Rev. A*, 88(5):052130, 2013.

[87] M. Horodecki, J. Oppenheim, and A. Winter. Quantum state merging and negative information. *Commun. Math. Phys*, 269(1):107–136, 2005.

[88] F. Hiai, and D. Petz. The proper formula for relative entropy and its asymptotics in quantum probability. *Commun. Math. Phys*, 143(1):99–114, 1991.

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

[90] D. P. DiVincenzo, P. Hayden, and B. M. Terhal. Hiding quantum data. *Found. Phys.*, 33(11):1629–1647, 2003.

[91] S. Watanabe. Information theoretical analysis of multivariate correlation. *IBM J. Res. Dev.*, 4:66–82, 1960.

[92] M. J. Donald, M. Horodecki, and O. Rudolph. The uniqueness theorem for entanglement measures. *J. Math. Phys.*, 43(9):4252–4272, 2002.

[93] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher. Concentrating partial entanglement by local operations. *Phys. Rev. A*, 53(4):2046–2052, 1996.

[94] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters. Mixed-state entanglement and quantum error correction. *Phys. Rev. A*, 54(5):3824–3851, 1996.

[95] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight. Quantifying Entanglement. *Phys. Rev. Lett.*, 78(12):2275–2279, 1997.

[96] M. Christandl and A. Winter. "Squashed entanglement": An additive entanglement measure. *J. Math. Phys.*, 45(3):829–840, 2004.

[97] Y. Huang. Computing quantum discord is NP-complete. *New J. Phys.*, 16:033027, 2014.

[98] D. L. Zhou. Irreducible multiparty correlations can be created by local operations. *Phys. Rev. A*, 80(2):022113, 2009.

[99] D. L. Zhou. Efficient Numerical Algorithm on Irreducible Multiparty Correlations. *Commun. Theor. Phys.*, 61(2):187, 2014.

[100] T. Galla and O. Gühne. Complexity measures, emergence, and multiparticle correlations. *Phys. Rev. E*, 85(4):046209, 2012.

[101] A. J. Bell. Co-information lattice. *4th Int. Symp. Indep. Compon. Anal. Blind Source Sep.*, 921–926, 2003.

[102] P. L. Williams and R. D. Beer. Nonnegative Decomposition of Multivariate Information. *e-print*, arXiv:1004.2515, 2010.

[103] L. D. Landau. On the theory of phase transitions. *Zh. Eks. Teor. Fiz.*, 7(1937):19–32, 1937.

[104] D. C. Tsui, H. L. Stormer, and A. C. Gossard. Two-Dimensional Magnetotransport in the Extreme Quantum Limit. *Phys. Rev. Lett.*, 48(22):1559–1562, 1982.

[105] R. B. Laughlin. Anomalous Quantum Hall Effect: An Incompressible Quantum Fluid with Fractionally Charged Excitations. *Phys. Rev. Lett.*, 50(18):1395–1398, 1983.

[106] D. Arovas, J. R. Schrieffer, and F. Wilczek. Fractional Statistics and the Quantum Hall Effect. *Phys. Rev. Lett.*, 53(7):722–723, 1984.

[107] X.-G. Wen. Vacuum degeneracy of chiral spin states in compactified space. *Phys. Rev. B*, 40(10):7387–7390, 1989.

[108] X.-G. Wen. Topological orders and edge excitations in fractional quantum Hall states. *Adv. Phys.*, 44(5):405–473, 1995.

[109] X. Chen, Z. C. Gu, and X.-G. Wen. Local unitary transformation, long-range quantum entanglement, wave function renormalization, and topological order. *Phys. Rev. B*, 82(15):155138, 2010.

[110] B. Zeng and X.-G. Wen. Gapped quantum liquids and topological order, stochastic local transformations and emergence of unitarity. *Phys. Rev. B*, 91(12):125121, 2015.

[111] S.-Bravyi and M. B. Hastings. A Short Proof of Stability of Topological Order under Local Perturbations. *Commun. Math. Phys.*, 307(3):609–627, 2011.

[112] C. L. Kane and E. J. Mele. Quantum Spin Hall Effect in Graphene. *Phys. Rev. Lett.*, 95(22):226801, 2005.

[113] A. Y. Kitaev. Unpaired Majorana fermions in quantum wires. *Physics-Uspekhi*, 44(10S):131, 2001.

[114] A. Miyake. Quantum computational capability of a 2D valence bond solid phase. *Ann. Phys. (N. Y).*, 326(7):1656–1671, 2011.

[115] P. W. Shor. Scheme for reducing decoherence in quantum computer memory. *Phys. Rev. A*, 52(4):2493–2496, 1995.

[116] B. Yoshida and I. L. Chuang. Framework for classifying logical operators in stabilizer codes. *Phys. Rev. A*, 81(5):052302, 2010.

[117] A. Kómár, O. L.-Cardinal, and K. Temme. Necessity of an energy barrier for self-correction of Abelian quantum doubles. *Phys. Rev. A*, 93(5):052337, 2016.

[118] R. Alicki, M. Horodecki, P. Horodecki, and R. Horodecki. On Thermal Stability of Topological Qubit in Kitaev's 4D Model. *Open Syst. & Inf. Dyn.*, 17(01):1–20, 2010.

[119] H. Li and F. D. M. Haldane. Entanglement spectrum as a generalization of entanglement entropy: Identification of topological order in non-Abelian fractional quantum hall effect states. *Phys. Rev. Lett.*, 101(1):010504, 2008.

[120] S. Dong, E. Fradkin, R. G. Leigh, and S. Nowling. Topological entanglement entropy in Chern-Simons theories and quantum Hall fluids. *J. High Energy Phys.*, 2008(05):016, 2008.

[121] K. Kato, F. Furrer, and M. Murao. Information-theoretical formulation of anyonic entanglement. *Phys. Rev. A*, 90(6),062325 2014.

[122] M. B. Hastings and X.-G. Wen. Quasiadiabatic continuation of quantum states: The stability of topological ground-state degeneracy and emergent gauge invariance. *Phys. Rev. B*, 72(4):045141, 2005.

[123] L. Zou and J. Haah. Spurious Long-range Entanglement and Replica Correlation Length. *Phys. Rev. B*, (94):075151, 2016.

[124] N. B.-Ali, L. Ding, and S. Haas. Topological order in paired states of fermions in two dimensions with breaking of parity and time-reversal symmetries. *Phys. Rev. B*, 80(18):180504, 2009.

[125] A. M. Läuchli, E. J. Bergholtz, J. Suorsa, and M. Haque. Disentangling Entanglement Spectra of Fractional Quantum Hall States on Torus Geometries. *Phys. Rev. Lett.*, 104(15):156404, 2010.

[126] R. Thomale, A. Sterdyniak, N. Regnault, and B. A. Bernevig. Entanglement Gap and a New Principle of Adiabatic Continuity. *Phys. Rev. Lett.*, 104(18):180502, 2010.

[127] H. Yao and X.-L. Qi. Entanglement Entropy and Entanglement Spectrum of the Kitaev Model. *Phys. Rev. Lett.*, 105(8):080501, 2010.

[128] S. Bravyi and A. Y. Kitaev. Quantum codes on a lattice with boundary. *e-print*, arXiv:quantu-ph/9811052, 1998.

[129] W. W. Ho, L. Cincio, H. Moradi, D. Gaiotto, and G. Vidal. Edge-entanglement spectrum correspondence in a nonchiral topological phase and Kramers-Wannier duality. *Phys. Rev. B*, 91(12):125119, 2015.

[130] J. R. Wootton. Towards unambiguous calculation of the topological entropy for mixed states. *J. Phys. A Math. Theor.*, 45(21):215309, 2012.

[131] J. I. Cirac, D. Poilblanc, N. Schuch, and F. Verstraete. Entanglement spectrum and boundary theories with projected entangled-pair states. *Phys. Rev. B.*, 83(24):245134, 2011.

[132] N. Schuch, D. Poilblanc, J. I. Cirac, and D. Pérez-García. Topological order in the projected entangled-pair states formalism: Transfer operator and boundary Hamiltonians. *Phys. Rev. Lett.*, 111(9):090501, 2013.

[133] D. Gottesman. Stabilizer codes and quantum error correction. *Ph.D. thesis*, available at arXiv:quant-ph/9705052, 1997.

[134] D. L. Zhou and L. You. Characterizing the complete hierarchy of correlations in an n-party system. *e-print*, arXiv:quant-ph/0701029, 2007.

[135] I. H. Kim. Determining the structure of the real-space entanglement spectrum from approximate conditional independence. *Phys. Rev. B*, 87(15):155120, 2013.

[136] Y. Zhang, T. Grover, A. Turner, M. Oshikawa, and A. Vishwanath. Quasiparticle statistics and braiding from ground-state entanglement. *Phys. Rev. B*, 85(23):235151, 2012.

[137] L. Fiedler, P. Naaijkens, and T. J. Osborne. Jones index, data hiding and total quantum dimension. *e-print*, arXiv:1608.02618, 2016.

[138] A. Shamir. How to Share a Secret. *Commun. ACM*, 22(11):612–613, 1979.

[139] G. R. Blakley. Safeguarding cryptographic keys. *Proc. 1979 AFIPS Nat. Comp. Conf.*, 313–317, 1979.

[140] M. Hillery, V. Buzek, and A. Berthiaume. Quantum secret sharing. *Phys. Rev. A*, 59(3):1829–1834, 1999.

[141] A. S. Holevo. The Capacity of the Quantum Channel with General Signal States. *IEEE Trans. Info. Theo.*, 44(1):269–273, 1998.

[142] B. Schumacher and M. D. Westmoreland. Sending classical information via noisy quantum channels. *Phys. Rev. A*, 56(1):131–138, 1997.

[143] B. Schumacher and M. D. Westmoreland. Quantum mutual information and the one-time pad. *Phys. Rev. A*, 74(4):042305, 2006.

[144] M. M. Wilde. From Classical to Quantum Shannon Theory. *Cambridge Univ. Press*, 2013.

[145] R. Renner. Security of Quantum Key Distribution. *PhD thesis*, available at arXiv:quant-ph/0512258, 2005.

[146] F. Huber and O. Guhne. Characterizing Ground and Thermal States of Few-Body Hamiltonians. *Phys. Rev. Lett.*, 117(1):010403, 2016.

[147] M. Kliesch, C. Gogolin, M. J. Kastoryano, A. Riera, and J. Eisert. Locality of Temperature. *Phys. Rev. X*, 4(3):031019, 2014.

[148] M. C. Arnesen, S. Bose, and V. Vedral. Natural Thermal and Magnetic Entanglement in the 1D Heisenberg Model. *Phys. Rev. Lett.*, 87(1):017901, 2001.

[149] D. Gunlycke, V. M. Kendon, V. Vedral, and S. Bose. Thermal concurrence mixing in a one-dimensional Ising model. *Phys. Rev. A.*, 64(4):042302, 2001.

[150] X. Wang. Thermal and ground-state entanglement in Heisenberg <math display XX qubit rings. *Phys. Rev. A*, 66(3):034302, 2002.

[151] H. Araki. Gibbs states of a one dimensional quantum lattice. *Commun. Math. Phys.*, 14(2):120–157, 1969.

[152] M. B. Hastings. Quantum belief propagation: An algorithm for thermal quantum systems. *Phys. Rev. B*, 76(20):201102, 2007.

[153] E. H. Lieb and D. W. Robinson. The finite group velocity of quantum spin systems. *Commun. Math. Phys.*, 28(3):251–257, 1972.

[154] F. G. S. L. Brandão and M. Horodecki. Exponential Decay of Correlations Implies Area Law. *Commun. Math. Phys.*, 333(2):761–798, 2014.

[155] K. Kato and F. G. S. L. Brandao. Quantum Approximate Markov Chains are Thermal. *e-print*, arXiv:1609.06636, 2016.

Parts of results in this thesis are based on manuscripts contained in the following papers:

- K. Kato, F. Furrer, and M. Murao. Information-theoretical analysis of topological entanglement entropy and multipartite correlations. *Phys. Rev. A*, 93(2),022317 2016. (main results in Chapter 4 except Theorem 9, Theorem 13, Theorem 14)

- K. Kato and F. G. S. L. Brandao. Quantum Approximate Markov Chains are Thermal. *e-print*, arXiv:1609.06636, 2016. (main results in Chapter 5)

- K. Kato and F. G. S. L. Brandao. Locality of Entanglement Spectrum and Edge States in Two-Dimensional Systems. *in preparation*. (Theorem 9, Theorem 13, Theorem 14 in Chapter 4)