Symmetry and Randomness in Quantum Information Theory: Several Applications

by

Wei Xie

A dissertation submitted for the degree of

Doctor of Philosophy

Centre for Quantum Software and Information
Faculty of Engineering and Information Technology
University of Technology Sydney, Australia

Certificate of Original Authorship

I, Wei Xie, declare that this thesis is submitted in fulfilment of the requirements for the award of PhD in the School of Computer Science at the University of Technology Sydney. This thesis is wholly my own work unless otherwise reference of acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis. This document has not been submitted for qualifications at any other academic institution. This research is supported by the Australian Government Research Training Program.

Production Note:
Signature removed prior to publication.

Abstract

This thesis studies four topics in quantum information theory using tools from representation theory and (high-dimensional) probability theory.

First, we study the nonadditivity of minimum output von Neumann and Rényi entropy of quantum channels. A sketch of the proof by Aubrun, Szarek and Werner for nonadditivity of minimum output entropy is presented, and a slight simplification is given. We show that asymptotically the minimum output entropy of the random channel $\mathcal{E} \otimes \mathcal{E} \otimes \mathcal{E}^*$ is achieved not by a tripartite genuinely entangled state, but by a tensor product of two states. We also study another model of random channel, and our estimation of the minimum output Rényi entropies fails to show the usefulness of genuine multipartite entanglement for the multiple nonadditivity.

Second, we study the generic entanglement in the random near-invariant tensors under the action of $\mathfrak{su}(2)$, and random symmetric invariant tensors under the action of $\mathfrak{su}(d)$ for any d, serving as an extension of the random invariant tensors under $\mathfrak{su}(2)$. We show that both the random tensors are asymptotically close to a maximally entangled state with respect to any bipartite cut.

Third, we study efficient quantum certification for states and unitaries. We present an algorithm that uses $O(\varepsilon^{-4} \ln |\mathcal{P}|)$ copies of an unknown state to distinguish whether the unknown state is contained in or ε-far from a finite set $\mathcal{P}$ of known states with respect to the trace distance. This algorithm is more sample-efficient in some settings. The previous study showed that one can distinguish whether an unknown unitary U is equal to or ε-far from a known or unknown unitary V in fixed dimension with $O(\varepsilon^{-2})$ uses of the unitary, in which an ancilla system should be used. We give an algorithm that distinguishes the two cases with $O(\varepsilon^{-1})$ uses of the unitary, without using ancilla system or using ancilla system of much smaller dimension.

Finally, we study the parallel repetition of extended nonlocal game motivated by its connection with multipartite steering and entanglement detection. We show that the probability of winning an n-fold parallel repetition of commuting nonsignaling extended nonlocal game G decreases exponentially in n, provided that the game value of G is strictly less than 1, following the approach used by Lancien and Winter based on de Finetti reduction.

Contents

Notation

Here is a list of some notation frequently used in this thesis, along with its description unless otherwise noted.

ln	Natural logarithm
log	Binary logarithm
$\mathbb{N}$	The set of all nonnegative integers
$\mathbb{Z}$	The set of all integers
$\mathbb{R}$	The set of all real numbers
$\mathbb{C}$	The set of all complex numbers
$[m]$	The set $\{1, 2, \ldots, m\}$
$\Pr(E)$	Probability of event E
$\mathbb{E}X$	Expectation of random variable X
$\mathrm{Var}X$	Variance of random variable X
$M^\intercal$	Transpose of matrix M
M^*	Complex conjugate of matrix M
$M^\dagger$	Transpose conjugate of matrix M
$M \geq N$	$M - N$ is semidefinite positive for Hermitian M, N
$\lVert M \rVert_p$	Schatten p-norm of matrix M
$\langle u, v \rangle$	Equal to $\sum_i u_i^* v_i$ for vectors u, v
$\langle M, N \rangle$	Equal to $\mathrm{tr}(M^\dagger N)$ for matrices M, N
$\mathcal{H}, \mathcal{K}$	Typical (finite-dimensional) Hilbert spaces
$\mathcal{L}(\mathcal{H}, \mathcal{K})$	The set of all linear operators from $\mathcal{H}$ to $\mathcal{K}$
$\mathcal{L}(\mathcal{H})$	$\mathcal{L}(\mathcal{H}, \mathcal{H})$
$f \lesssim g$	For positive functions f, g of n, $f(n) \leq cg(n)$ holds for some positive constant c and any sufficiently large n, also written as $g \gtrsim f$, or $f = O(g)$, or $g = \Omega(f)$
$f \simeq g$	Both $f \lesssim g$ and $f \gtrsim g$ hold, also written as $f = \Theta(g)$
$f \sim g$	$f(n)/g(n) \to 1$ as $n \to \infty$

$\mathrm{U}(d)$	Unitary group of degree d
$\mathrm{SU}(d)$	Special unitary group of degree d
$\mathrm{SL}(d)$	Special linear group of degree d over $\mathbb{C}$
$\mathrm{GL}(d)$	General linear group of degree d over $\mathbb{C}$
$\mathfrak{su}(d)$	Lie algebra of $\mathrm{SU}(d)$
$\mathfrak{sl}(d)$	Lie algebra of $\mathrm{SL}(d)$ over $\mathbb{C}$
$\mathrm{Par}(n)$	The set of all partitions of n
$\lambda \vdash n$	$\lambda \in \mathrm{Par}(n)$
$\mathrm{Par}(n, d)$	The set of all partitions of n with length at most d
$\mathrm{Type}(n, d)$	The set of all types of strings in $\{1, 2, \ldots, d\}^n$
R	The function that maps a group or an algebra to the set of its representation matrices
S_n	Symmetric group of degree n
W_π	The operator that permutates n tensor factors according to $\pi \in S_n$
W	Typical name for the swap operator
V_λ^L	The irrep of $\mathrm{GL}(d)$ of highest weight λ, sometimes written as V_λ or $\mathcal{H}_\lambda$
V_λ^S	The irrep of S_n labeled by partition λ, sometimes written as $\mathcal{K}_\lambda$
$\vee^n \mathbb{C}^d$	The symmetric subspace of $(\mathbb{C}^d)^{\otimes n}$
$\wedge^n \mathbb{C}^d$	The antisymmetric subspace of $(\mathbb{C}^d)^{\otimes n}$
χ	Character of a representation, or Holevo information, or Holevo capacity
S^{d-1}	Unit sphere in Euclidean space $\mathbb{R}^d$
$\mathcal{S}(\mathcal{H})$	Unit sphere in Hilbert space $\mathcal{H}$
$\mathcal{D}(\mathcal{H})$	The set of all density operators on $\mathcal{H}$
ρ, σ	Typical density operators
ψ, φ	$\psi = \lvert\psi\rangle\langle\psi\rvert$, $\varphi = \lvert\varphi\rangle\langle\varphi\rvert$ (applied for all pure states)
ϕ, ω	Typical maximally entangled state and maximally mixed state respectively
$H_p(\cdot)$	(Quantum or classical) p-Rényi entropy
$H(\cdot)$	Shannon entropy, or von Neumann entropy
$\mathrm{D}(\rho, \sigma)$	Trace distance between two quantum states ρ, σ
$\mathrm{F}(\rho, \sigma)$	Fidelity between two quantum states ρ, σ

Chapter 1

Introduction

1.1 Background and overview

The history of information theory can be traced back to 1948, when Shannon published a seminal paper 'A mathematical theory of communication' [Sha48] laying the foundations for the modern theory of information, in which he proposed a powerful mathematical framework to study information processing tasks such as data compression and data transmission. In the early 1940s, people thought it is impossible to send information at a positive rate with negligible error. Shannon challenged this opinion by proving that the error probability of communication could be made nearly zero for appropriate communication rate. Using this novel framework, Shannon's theory answers two fundamental questions in communication theory: What is the ultimate data compression rate, and what is the ultimate data transmition rate. The entropy and channel capacity are the respective answers.

Information theory, however, should not be viewed as merely a subfield of communication theory. Over the past decades the ideas used in information theory have made indispensable contributions to statistical physics, computer science, probability and statistics, and even to economics. Landauer's 'information is physical' claim [Lan91] and Wheeler's 'it from bit' concept [Whe90] summarize the deep connection between physics and information. Information theory has been playing a more and more important role in the study of physics and other science branches. The readers are referred to [CT12] for an introduction to information theory.

Prior to the invention of information theory, the first half of the 20th century saw the magnificent development of quantum mechanics. The counter-intuitive nature of quantum mechanics puzzled physicists at first but then drastically changed the way people understood the world. Our universe is governed fundamentally by the laws of quantum mechanics, rendering the necessity to study the effect of quantum mechanics upon information processing. It turns out that under the laws of quantum mechanics, information behaves quite differently from classical world.

During the past decades, quantum information science has rapidly developed and has become a rather rich science branch, with subfields including quantum computing, quantum cryptography, quantum communication, quantum complexity theory, quantum entanglement theory, etc. Over recent years it has been offering a new perspective for the study of physics, including thermodynamics, condensed matter and quantum gravity. The field of quantum computing was initiated in the first half of 1980s by the works of Benioff [Ben80] and Manin [Man80], Feynman [Fey82], and Deutsch [Deu85]. Many remarkable quantum information processing techniques and quantum algorithms were proposed in the 1990s [DJ92, BW92, BBC+93, Sho94, Sho96, Gro96, Sho99]. The recent development of quantum information science is also drived by other factors. Due to the decrease in the size of computing components over the past decades, quantum mechanics will inevitably become more relevant to the construction of computer chips, since it offers a remarkably accurate description of microscopic physical systems. The technique of quantum parallelism and interference enables a quantum computer to perform a calculation upon a superposition of quantum states as input and then to extract desired information via quantum measurement. In this way a quantum computer can outperform its classical counterpart when dealing with certain computational tasks. The principle of quantum mechanics is also introduced to the field of communication and cryptography and makes the transmission of information more efficient and secure.

Mathematical foundation of quantum information theory has been based mainly on the viewpoint of operator algebra. Nowadays many mathematical ideas have been widely employed in the study of quantum information. Group symmetry is the foundation for many fundamental theories of modern physics, and is also placed in the heart of many mathematical theories. The basic idea of group symmetry has been applied to various fields in physics, partical physics, quantum field theory and condensed matter physics. As the theory of quantum information is an area utilizing the properties of quantum world, group symmetry is undoubtedly a significant idea in quantum information as well. Representation theory is a useful tool in the study of symmetry in that it reduces problems in abstract algebra to that in linear algebra. The applications of group-theoretic method can be found in almost every corner in quantum information theory such as quantum compression and estimation; the interested reader is referred to [Har05, Chr06, Hay17] and the references therein for more introduction.

Another theme of this thesis is the applications of probabilistic methods in quantum information theory. One striking trait of quantum theory is that the quantum object in question usually lives in a high-dimensional vector space. It is often intractable or impractical to give an analytical or numerical study, hence *the curse of dimensionality.* The probabilistic methods, however, allow one to choose appropriate free parameters to simplify the analysis, and thus transform the curse of dimensionality into *the blessing of dimensionality.* Many tools from random matrix theory have been used successfully in quantum information theory. In high dimension the phenomenon of *concentration of measure* [Led05] could yield many surprising and powerful results. Roughly speaking, concentration of measure refers to the phenomenon that a Lipschitz function of many independent random variables is essentially constant. In some sense, the subject of concentration of measure lies at the core of modern probability theory.

It is a curious but elementary fact that the uniform measure on the Euclidean unit sphere S^{n-1} concentrates strongly about any equator as n gets large. This observation can be rigorously formulated by Levy's lemma. Note that the phenomenon of measure concentration happens not only for the uniform measure on S_n or the Haar measure on compact group but for many general measures. As a quantum state is represented by a unit vector, Levy's lemma offers a natural tool for studying the properties of random quantum state and related concepts such as random subspace, random unitary and random channel in high dimension; see [HLSW04, HLW06, AS17] for more. Among many important applications of concentration of measure to quantum information theory is the disproof of the additivity conjecture of Holevo capacity, which was arguably the most significant conjecture in quantum information theory. Hayden and Winter [Win07, Hay07, HW08] employed random unitary channel and random subspace in high dimension to yield counterexamples to the additivity conjecture of minimum output p-Rényi entropy for all $p > 1$. Hastings [Has09] then showed that the minimum output von Neumann entropy is not additive, and equivalently, the Holevo capacity is not additive, also using a probabilistic method. The reasoning in [Has09] was made clearer in [BH10], and a more concise proof was provided in [ASW11]. All these counterexamples, explicitly or implicitly, use the idea of measure concentration exhibited in high dimensional random quantum state, random subspace or random quantum operation, although no explicit construction is known up to now.

The probabilistic method has close connection to the group-theoretic approach. Indeed, the uniform measure on compact Lie group, widely used in the probabilistic method, can be calculated and studied using group symmetry. The probabilistic method and the group-theoretic method would continue to play an important role in the study of quantum information theory.

The structure and main results of this thesis are summarized as follows. Chapter 1 introduces research background, notation and terminology, and presents basics of linear algebra in Section 1.2, quantum information theory in Section 1.3, and group representation theory and (high-dimensional) probability theory in Section 1.4.

Chapter 2 is devoted to the study of the usefulness of multipartite quantum entanglement in violating the additivity of minimum output entropy. Section 2.1 introduces background of the additivity problem of minimum output entropy, sketches the proof in [ASW11] using Dvoretzky's theorem, and then slightly simplifies part of their argument by using a tighter bound on the Lipschitz constant of a function of bipartite states. Section 2.2 uses the graphical Weingarten calculus to study the asymptotic minimum output entropy of the random channel $\mathcal{E} \otimes \mathcal{E} \otimes \mathcal{E}^*$, and shows that asymptotically it is the tensor product of a bipartite maximally entangled state and a pure state, instead of any genuine tripartite state, that achieves the minimum output von Neumann entropy of the triple channels in high dimension. Section 2.3 analyzes another model of high-dimensional random quantum channels, and no evidence is found to support the usefulness of genuine multipartite entanglement for nonadditivity of minimum output Rényi entropy. So the multiple nonadditivity problem is still left intact.

Chapter 3 studies the generic entanglement in the random near-invariant tensors under the action of $\mathfrak{su}(2)$ and random symmetric invariant tensors under the action of $\mathfrak{su}(d)$ for any dimension d. The symmetric invariant tensor is so named since it lies in the tensor power of symmetric subspace. Section 3.1 introduces the representation theory for special unitary group, focusing on the Clebsch-Gordan transform and the Gelfand-Tsetlin patterns. Section 3.2 shows that a high-dimensional random near-invariant multipartite state is asymptotically perfect, i.e., close to a maximally entangled state with respect to any bipartite cut. Section 3.3 shows that similarly a high-dimensional random symmetric invariant multipartite state is also asymptotically perfect.

Chapter 4 studies the quantum certification for states and unitaries, and gives some efficient certification algorithms. Section 4.1 introduces the background and literature review for the certification problem. Section 4.2 shows that in order to distinguish whether an unknown state is contained in or ε-far from a finite set $\mathcal{P}$ of known states with respect to the trace distance, $O(\varepsilon^{-4} \ln |\mathcal{P}|)$ copies of the unknown state are sufficient. In previous study, $O(\max\{\varepsilon^{-2}, \delta^{-2}\} \ln |\mathcal{P}|)$ copies are needed, where δ is the minimum distance between distinct states in $\mathcal{P}$. Our algorithm uses less copies of states when $\delta \lesssim \varepsilon$. Section 4.3 shows that to order to distinguish whether an unknown unitary U is equal to or far from a known or unknown unitary V in fixed dimension, $O(\varepsilon^{-1})$ uses of the unitary suffice, while $O(\varepsilon^{-2})$ uses of unitary are needed when using previous method based on Choi state. Another advantage of our approach is that we do not need to introduce extra ancilla system or need an ancilla system of much smaller dimension in the certification.

Chapter 5 studies the parallel repetition of extended nonlocal game using nonsignaling strategy. Section 5.1 introduces background, notational convention and prior work for the parallel repetition of nonlocal games. Section 5.2 presents several technical lemmas on multipartite states and operator assemblages, focusing on symmetry-related properties of them. Section 5.3 applies these technical lemmas to show that the probability of winning an n-fold parallel repetition of commuting nonsignaling extended nonlocal game G decreases exponentially in n provided that the game value of G is strictly less than 1, following the approach in [LW16] based on de Finetti reduction. The question that whether this result can be extended to multiplayer case, or general nonsignaling strategy case, is left for future work.

1.2 Linear algebra and notation

This section explains some concepts and basics in linear algebra.

We write $\mathbb{C}$ and $\mathbb{R}$ for the set of complex numbers and the set of reals respectively. For complex number α, we write α^* for its complex conjugate. For a complex matrix M, we write M^* and $M^\dagger$ for its complex conjugate and conjugate transpose respectively.[1] A complex vector space is a nonempty set $\mathcal{H}$ together with two operations, i.e., vector addition $+ : \mathcal{H} \times \mathcal{H} \to \mathcal{H}$ and scalar multiplication $\cdot : \mathbb{C} \times \mathcal{H} \to \mathcal{H}$, satisfying the following conditions: $(\mathcal{H}, +)$ is an abelian group, $1v = v$, $(\alpha + \beta)v = \alpha v + \beta v$, $\alpha(\beta v) = (\alpha\beta)v$, $\alpha(v + u) = \alpha v + \alpha u$ for any $\alpha, \beta \in \mathbb{C}, v, u \in \mathcal{H}$.

An *inner product* over a vector space $\mathcal{H}$ is a mapping $\langle \cdot, \cdot \rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ fulfilling the following three properties. For any $\alpha_1, \alpha_2 \in \mathbb{C}, u, v, v_1, v_2 \in \mathcal{H}$, (i) Sesquilinearity: $\langle u, \alpha_1 v_1 + \alpha_2 v_2 \rangle = \alpha_1 \langle u, v_1 \rangle + \alpha_2 \langle u, v_2 \rangle$, (ii) Conjugate symmetry: $\langle u, v \rangle = (\langle v, u \rangle)^*$, (iii) Positive-definiteness: $\langle v, v \rangle \geq 0$ with equality if and only if $v = 0$. A vector v is called a unit vector if $\langle v, v \rangle = 1$. We adopt the Dirac notation, in which a unit vector is written as $|\psi\rangle$, the adjoint vector to $|\psi\rangle$ is written as $\langle \psi|$, and the inner product is written as $\langle \varphi | \psi \rangle$.

[1] Some literature use overline to denote the complex conjugate and use asterisk to denote the transpose conjugate.

A collection $\{|\psi_i\rangle\}_{i\in I}$ of unit vectors is called an *orthonormal basis* of $\mathcal{H}$ if they are orthogonal to each other and if each vector in $\mathcal{H}$ can be written as a linear combination of vectors in this collection. The cardinality of I is called the dimension of $\mathcal{H}$. Throughout this thesis we consider only one type of Hilbert space: the finite-dimensional complex vector spaces equipped with the Euclidean inner product, that is, $\mathcal{H} \cong \mathbb{C}^d$ for some $d < \infty$, and $\langle u, v\rangle \equiv \sum_{i=1}^d u_i^* v_i$ for $u = \sum_{i=1}^d u_i|i\rangle$ and $v = \sum_{i=1}^d v_i|i\rangle$ where $\{|i\rangle\}$ is an orthonormal basis. We write $\|v\| := \sqrt{\langle v, v\rangle}$ for the length of $v \in \mathcal{H}$.

A *(linear) operator* from Hilbert space $\mathcal{H}_1$ to Hilbert space $\mathcal{H}_2$ is a mapping $M : \mathcal{H}_1 \to \mathcal{H}_2$ satisfying $M(u + v) = Mu + Mv$ and $M(\alpha v) = \alpha(Mv)$ for any $\alpha \in \mathbb{C}, u, v \in \mathcal{H}$. The set of all operators from $\mathcal{H}_1$ to $\mathcal{H}_2$ is denoted by $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$. A linear operator on $\mathcal{H}$ is simply an operator from $\mathcal{H}$ to itself, and the set of operators on $\mathcal{H}$ is denoted by $\mathcal{L}(\mathcal{H})$. The identity operator on a linear space $\mathcal{H}$ is usually written as $\mathbb{1}_{\mathcal{H}}$. An *associative algebra* is a vector space equipped with an associative bilinear multiplication, and a basic example of associative algebra is $\mathcal{L}(\mathcal{H})$.

With respect to any orthonormal bases $(|j\rangle)$ and $(|i\rangle)$ of $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively, an operator M can be written in the matrix form $M = \sum_{ij} M_{ij}|i\rangle\langle j|$, its *transposition* is $M^\intercal = \sum_{ij} M_{ij}|j\rangle\langle i|$, and its *adjoint* is $M^\dagger = \sum_{ij} M_{ij}^*|j\rangle\langle i|$. The *trace* of an operator $M \in \mathcal{L}(\mathcal{H})$ is $\mathrm{tr}(M) = \sum_i M_{ii}$, which is independent of the basis chosen. We write $\mathrm{tr}^k(M)$ for $(\mathrm{tr}\, M)^k$. The Hilbert-Schmidt inner product on $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ is denoted as $\langle X, Y\rangle \equiv \mathrm{tr}(X^\dagger Y)$ for $X, Y \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$, making $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ a Hilbert space.

The concept of tensor is a generalization of that of vector and matrix. For any nonnegative integers p and q, a (p, q) tensor is a linear mapping from $\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_p$ to $\mathcal{K}_1 \otimes \cdots \otimes \mathcal{K}_q$. For instance, a usual vector is a $(0, 1)$ tensor, and $|w\rangle\langle uv|$ is a $(2, 1)$ tensor.

An operator $M \in \mathcal{L}(\mathcal{H})$ is said to be *Hermitian* if $M^\dagger = M$, and the set of Hermitian operators on $\mathcal{H}$ is denoted by $\mathrm{Herm}(\mathcal{H})$. An operator $M \in \mathcal{L}(\mathcal{H})$ is said to be *positive semidefinite*, written $M \geq 0$, if $\langle\psi|M|\psi\rangle \geq 0$ for all $|\psi\rangle \in \mathcal{H}$, in which case M should be Hermitian. The partial order '$\geq$' over the set of Hermitian operators on $\mathcal{H}$ is defined by that $M \geq N$ if and only if $M - N \geq 0$. An operator $M \in \mathcal{L}(\mathcal{H})$ is an *orthogonal projection operator* (or *projector* for short) if $M \geq 0$ and $M^2 = M$. An operator $M \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ is a *(linear) isometry* if $M^\dagger M = \mathbb{1}$. If an isometry maps a space $\mathcal{H}$ to itself, it is called a *unitary operator*. The set of unitary operators on $\mathcal{H}$ is denoted by $\mathrm{U}(\mathcal{H})$.

The *singular value theorem* states that for any operator $M \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$, there exist real numbers $s_1(M) \geq \cdots \geq s_r(M) > 0$ and orthonormal sets $\{|\varphi_1\rangle, \ldots, |\varphi_r\rangle\} \subset \mathcal{H}_1$

and $\{|\psi_1\rangle, \ldots, |\psi_r\rangle\} \subset \mathcal{H}_2$ such that

$$M = \sum_{i=1}^{r} s_i(M)|\psi_i\rangle\langle\varphi_i| \,.$$

The numbers $s_1(M), \ldots, s_r(M)$ are referred to as the singular values of M, and the vector of singular values of M is written as $s(M) := (s_1(M), \ldots, s_r(M))^\intercal$. Sometimes it is convenient to define $s_k(M) = 0$ for $k > \text{rank}(M)$. The *spectral theorem* states that an operator $M \in \mathcal{L}(\mathcal{H})$ is Hermitian if and only if there exist real numbers $\lambda_1, \ldots, \lambda_r$ and an orthonomal set $\{|\psi_1\rangle, \ldots, |\psi_r\rangle\} \subset \mathcal{H}$ such that

$$M = \sum_{i=1}^{r} \lambda_i|\psi_i\rangle\langle\psi_i| \,.$$

The *support* of a Hermitian operator M, denoted $\text{supp}(M)$, is the vector space spanned by the eigenvectors of M with nonzero eigenvalue.

Let $\{|a\rangle\}$ and $\{|b\rangle\}$ be orthonormal bases of $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively. There exists an isomorphism between the Hilbert space $\mathcal{L}(\mathcal{H}_A, \mathcal{H}_B)$ of operators and the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ of bipartite vectors:

$$\begin{aligned} \text{vec} : \mathcal{L}(\mathcal{H}_A, \mathcal{H}_B) &\to \mathcal{H}_A \otimes \mathcal{H}_B \\ |b\rangle\langle a| &\mapsto |a\rangle|b\rangle \,. \end{aligned} \tag{1.1}$$

By linearity, it holds that $\text{vec}(|u\rangle\langle v|) = |v\rangle^*\langle u|$ for $|v\rangle \in \mathcal{H}_A$ and $|u\rangle \in \mathcal{H}_B$, where the complex conjugate is with respect to the basis $\{|a\rangle\}$. Obviously the definition of vec depends on the choice of orthonormal bases. This operator-vector correspondence is widely used in quantum information theory to enable some convenient calculation. For example, via this correspondence the *Schmidt decomposition* can be derived from the singular value theorem: any bipartite pure state $|\psi\rangle_{AB}$ can be written as $|\psi\rangle_{AB} = \mu_i|\psi_i\rangle_A|\varphi_i\rangle_B$ for some orthonomal bases $(|\psi_i\rangle_A)_i$ and $(|\varphi_i\rangle_B)_i$, where μ_i are known as *Schmidt coefficients*.

A *norm* on a vector space is a real-valued function on the space with three properties: positive definiteness, positive scalibility, and the triangle inequality. The *p-norm* (or *ℓ_p-norm*) of a vector $x = (x_1, \ldots, x_d)^\intercal \in \mathbb{C}^d$ is

$$\|x\|_p = \Big(\sum\nolimits_i |x_i|^p\Big)^{1/p} \,.$$

For any operator $M \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ and any real number $p \geq 1$, the *Schatten p-norm* of

M is defined as

$$\|M\|_p = \left(\operatorname{tr}(|M|^p)\right)^{1/p},$$

where $|M| := \sqrt{M^\dagger M}$. The Schatten 1-norm and 2-norm are also known as the *trace norm* and the *Frobenius norm* respectively. The Schatten ∞-norm, also known as the *spectral norm* or *operator norm*, is defined as

$$\|M\|_\infty = \max\{\|Mv\|_2 : v \in \mathcal{H}_1, \|v\|_2 = 1\},$$

which coincides with $\lim_{p\to\infty} \|M\|_p$. It is easy to see that the Schatten p-norm of an operator M is equal to the p-norm of the vector of singular values of M:

$$\|M\|_p = \|s(M)\|_p .$$

We fix several miscellaneous notations as follows. We write S_n for the symmetric group of degree n, while S^{d-1} and $\mathcal{S}(\mathbb{C}^d)$ denote the Euclidean unit spheres in $\mathbb{R}^d$ and $\mathbb{C}^d$ respectively. For sets R and T, we sometimes use T^R to denote the set of all functions from R to T. For positive functions $f(n)$ and $g(n)$ of $n \in \mathbb{N}$, we write $f(n) \gtrsim g(n)$ or $g(n) \lesssim f(n)$ iff $f(n) \geq cg(n)$ for constant c and any large n, write $f(n) \simeq g(n)$ iff $f(n) \lesssim g(n)$ and $f(n) \gtrsim g(n)$, and write $f(n) \sim g(n)$ iff $\frac{f(n)}{g(n)} \to 1$ as $n \to \infty$. Throughout this thesis the letters c, c', C, C' usually denote absolute positive constants, independent of the dimensions involved, whose values may change from occurrence to occurrence. Some notation used in this thesis is summarized in the table on pages 3–4.

1.3 Quantum information theory

This section, largely based on [NC11, Wat18], explains some basic concepts and principles in quantum mechanics and quantum information theory.

The first postulate of quantum mechanics provides the arena in which quantum mechanics takes place. It asserts that associated to any closed physical system is a complex Hilbert space, known as the state space of the system, and that the system is completely described by its state vector, which is a unit vector in this space. We denote by $|A|$ or d_A the dimension of the Hilbert space of system A (we also use $|S|$ to denote the cardinality of a set S when no confusion arises).

The system we are most concerned with and also the simplest one is the two-dimensional system, known as a qubit. Supposing $|0\rangle$ and $|1\rangle$ form an orthonormal basis for this space, any state vector in this space can be written as

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$$

with $\alpha, \beta \in \mathbb{C}$ satisfying $|\alpha|^2 + |\beta|^2 = 1$.

The continuous-time dynamics of a closed quantum system is described by the Schrödinger equation, i.e., $\mathrm{i}\hbar\frac{\mathrm{d}|\psi\rangle}{\mathrm{d}t} = H|\psi\rangle$ where $\hbar$ is Planck's constant and H is the Hamiltonian of the system. In quantum computation, we usually consider the discrete-time evolution of a system. Suppose the states of a closed system at times t_0 and t_1 are $|\psi_0\rangle$ and $|\psi_1\rangle$, respectively. Then the two states are related to each other by a unitary operator U which depends only on the times t_0 and t_1:

$$|\psi_1\rangle = U|\psi_0\rangle\,.$$

This is the second postulate of quantum mechanics. The Hadamard gate H and Pauli matrices $\sigma_x, \sigma_y, \sigma_z$ are among the most frequently used unitary operators acting on a qubit.

A closed quantum system evolves according to unitary operator, but there are times when people interact with quantum system to extract useful information from it. The third postulate is introduced to explain what happens in this process of quantum measurement. Quantum measurement is described by a collection $\{M_i\}_{i\in I}$ of *measurement operators* M_i which satisfy that $\sum_{i\in I} M_i^\dagger M_i = \mathbb{1}$. The index i refers to possible measurement outcome that may occur in the experiment. Suppose the state of quantum system is $|\psi\rangle$ before measurement, then for each $i \in I$ the result i occurs with probability $p_i = \langle\psi|M_i^\dagger M_i|\psi\rangle$. If the measurement outcome is i, the state of system after measurement is $M_i|\psi\rangle$ up to a normalization coefficient.

The general measurement described above has two variations. If the main item of interest is the probabilities of the measurement outcomes, but not the post-measurement states, the *positive operator-valued measure (POVM)* formalism is used to deal with this case. A POVM is a set of positive semidefinite operators that sum up to an identity operator. In other words, a POVM is any $\{E_i\}_{i\in I}$ satisfying $E_i \geq 0$ for each $i \in I$ and $\sum_{i\in I} E_i = \mathbb{1}$. A *projective measurement* is a special measurement in which each measurement operator is a projector. A projective measurement is also described by an *observable* which is simply a Hermitian operator, of which each eigenvalue represents a possible measurement outcome and the projector onto each eigenspace is a measurement operator. Any POVM can be realized as projective measurement by introducing ancilla system.

We are always concerned with a composite quantum system made up of two or more different physical systems. The fourth postulate describes how the state space of a composite system is built up from the state spaces of the component systems. The state space of a composite physical system is the tensor product of the state spaces of the component physical systems. When the states of n systems A_1 through A_n are $|\psi_1\rangle, \ldots, |\psi_n\rangle$ respectively, the joint state of the total system is $|\psi_1\rangle \otimes \cdots \otimes |\psi_n\rangle$.

Now we focus on a component system in a composite system. Suppose the composite system is in the singlet state $\frac{1}{\sqrt{2}}(|0\rangle_A|1\rangle_B - |1\rangle_A|0\rangle_B)$, then the state of the system A can be equivalently viewed as $\frac{1}{2}\mathbb{1}_A$ from the perspective of local observers. If the joint state is $|\psi\rangle_{AB}$, then the state of local system A is $\rho_A := (\mathbb{1} \otimes \mathrm{tr})(|\psi\rangle\langle\psi|_{AB})$. An operator in $\mathcal{L}(\mathcal{H})$ is called a *density operator* if it is positive semidefinite with trace one. It can be shown that an operator in $\mathcal{L}(\mathcal{H})$ represents a valid quantum state of some component system if and only if ρ is a density operator. Throughout this thesis a pure state is usually labeled by a Greek letter, and the density operator of a pure state is simply denoted by the letter, for example, $\psi := |\psi\rangle\langle\psi|$ and $\varphi := |\varphi\rangle\langle\varphi|$. We denote by $\mathcal{D}(\mathcal{H})$ the set of density operators on space $\mathcal{H}$. For a joint state ρ_{AB} on AB, the reduced state on A is $\rho_A := (\mathbb{1} \otimes \mathrm{tr})\rho_{AB}$, where $\mathbb{1} \otimes \mathrm{tr} =: \mathrm{tr}_B$ is called *partial trace* and ρ_A is called the *reduced state* of ρ_{AB} on A. For an n-component composite system $\mathcal{H} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$, any state of form $\rho = \rho_1 \otimes \cdots \otimes \rho_n$ for $\rho_i \in \mathcal{D}(\mathcal{H}_i)$ is called a *product state*. Furthermore, if $\rho \in \mathcal{D}(\mathcal{H})$ can be written as a convex combination of product states, it is called a *separable state*; otherwise, it is called an *entangled state*.

The density operator describes not only the local physical state but also the statistical mixture of an ensemble of states. If a system is in state ψ_i with probability p_i for each $i \in I$, then statistically the system can be viewed in the state $\sum_i p_i|\psi_i\rangle\langle\psi_i|$, which is a density operator.

We use $H(\sigma) := -\mathrm{tr}(\sigma \ln \sigma)$ to denote the von Neumann entropy [vN32] of a quantum state σ, $H(A|B)_\rho := H(AB)_\rho - H(B)_\rho$ the conditional quantum entropy of ρ_{AB}, and $I(A:B)_\rho := H(A)_\rho + H(B)_\rho - H(AB)_\rho$ the quantum mutual information of ρ_{AB}. For $p \geq 0$ and $p \neq 1$, the *p-Rényi entropy* of a density operator ρ is defined as

$$H_p(\rho) = \frac{1}{1-p} \ln \mathrm{tr}(\rho^p)\,.$$

Throughout this thesis we use the natural logarithm ln to define various entropies and relevant concepts, while many texts use the binary logarithm instead. The cases of p-Rényi entropy for $p = 1, p = \infty$ should be understood as the limits $p \to 1, p \to \infty$ respectively. It turns out $H_0(\rho) = \ln \operatorname{rank} \rho$, $H_\infty(\rho) = -\ln \|\rho\|_\infty$, and the 1-Rényi entropy is exactly the von Neumann entropy. When $p > 1$, it holds that $H_p(\rho) = \frac{p}{1-p} \ln \|\rho\|_p$.

The two most natural measures of distance between quantum states are the *fidelity* and *trace distance*. The fidelity between states ρ and σ is defined as $\mathrm{F}(\rho, \sigma) := \|\sqrt{\rho}\sqrt{\sigma}\|_1$, and in particular, $\mathrm{F}(\psi, \varphi) = \sqrt{\langle \psi, \varphi \rangle}$. The trace distance between ρ and σ is defined as $\mathrm{D}(\rho, \sigma) := \frac{1}{2}\|\rho - \sigma\|_1$, and in particular, $\mathrm{D}(\psi, \varphi) = \sqrt{1 - \langle \psi, \varphi \rangle}$.

Any element in $\mathcal{L}(\mathcal{L}(\mathcal{H}_A), \mathcal{L}(\mathcal{H}_B))$ is called a *super-operator*. A super-opertator $\mathcal{E} \in \mathcal{L}(\mathcal{L}(\mathcal{H}_A), \mathcal{L}(\mathcal{H}_B))$ is called *positive* if $\mathcal{E}(M_A) \geq 0$ for any $M_A \geq 0$; $\mathcal{E}$ is called *completely positive* if $\mathcal{E} \otimes \mathbb{1}$ is positive for any identity super-operator $\mathbb{1}$; $\mathcal{E}$ is called *trace preserving* if $\mathrm{tr}(\mathcal{E}(M)) = \mathrm{tr}(M)$ for any $M \in \mathcal{L}(\mathcal{H}_A)$. A *quantum operation* (or *quantum channel*) $\mathcal{E}_{A\to B}$ with input system A and output system B is a completely positive (CP), trace preserving (TP) linear map, mapping the linear operators on $\mathcal{H}_A$ to those on $\mathcal{H}_B$. Since the subscript of an operator or operation specifies its input and output systems, sometimes we may write an operator or operation omitting the identity operator or operation $\mathbb{1}$, e.g., $X_{AB}Y_{BC} \equiv (X_{AB} \otimes \mathbb{1}_C)(\mathbb{1}_A \otimes Y_{BC})$ and $\mathcal{E}_{B\to C}(X_{AB}) \equiv (\mathbb{1}_A \otimes \mathcal{E}_{B\to C})X_{AB}$, and we may also write a partial trace by simply omitting some subscript, e.g., $\rho_A \equiv \mathrm{tr}_B\, \rho_{AB}$ and $\mathcal{E}_{A\to B} \equiv \mathrm{tr}_C \circ \mathcal{E}_{A\to BC}$.

There are several equivalent and convenient representations of quantum channel, including Choi-Jamiołkowski matrix [Jam72, Cho75], Kraus decomposition [HK69, HK70], and Stinespring representation [Sti55], which are introduced as follows.

In the same spirit as the operator-vector correspondence (1.1), the *Choi-Jamiołkowski isomorphism* between $\mathcal{L}(\mathcal{L}(\mathcal{H}_A), \mathcal{L}(\mathcal{H}_B))$ and $\mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is as follows. The Choi matrix of a super-operator $\mathcal{E}_{A\to B}$ is $J_{\mathcal{E}} = (\mathbb{1}_{\tilde{A}\to A} \otimes \mathcal{E}_{A\to B})\phi_{\tilde{A}A}$, where $\phi_{\tilde{A}A} = \sum_{ij} |ii\rangle\langle jj|$ is a fixed unnormalized maximally entangled state. We call $J_{\mathcal{E}}/\dim(\mathcal{H}_A)$ the Choi state of $\mathcal{E}$. The output of the channel $\mathcal{E}_{A\to B}$ with input ρ_A can be recovered from $J_{\mathcal{E}}$ as $\mathcal{E}_{A\to B}(\rho_A) = \mathrm{tr}_A\left(\mathsf{T}_A(J_{\mathcal{E}})\rho_A\right)$, where $\mathsf{T}_A(J_{\mathcal{E}})$ is the partial transpose on A of $J_{\mathcal{E}}$. It can be verified that $\mathcal{E}$ is completely positive iff $J_{\mathcal{E}}$ is positive semidefinite, and that $\mathcal{E}$ is trace preserving iff $\mathrm{tr}_B\, J_{\mathcal{E}} = \mathbb{1}_A$.

Any $\mathcal{E} \in \mathcal{L}(\mathcal{L}(\mathcal{H}_A), \mathcal{L}(\mathcal{H}_B))$ is completely positive iff there exist linear operators $K_i \in \mathcal{L}(\mathcal{H}_A, \mathcal{H}_B)$ such that

$$\mathcal{E}(X) = \sum_i K_i X K_i^\dagger\,.$$

This is called *Kraus decomposition* and the K_i are known as *Kraus operators*. Further, a completely positive $\mathcal{E}$ is trace preserving iff

$$\sum_i K_i^\dagger K_i = \mathbb{1}_A\,.$$

Moreover, $\mathcal{E} \in \mathcal{L}(\mathcal{L}(\mathcal{H}_A), \mathcal{L}(\mathcal{H}_B))$ is completely positive iff there exists an operator $L \in \mathcal{L}(\mathcal{H}_A, \mathcal{H}_B \otimes \mathcal{H}_R)$ for some choice of $\mathcal{H}_R$ such that

$$\mathcal{E}(X) = \mathrm{tr}_R(LXL^\dagger)\,,$$

which is known as *Stinespring dilation*. Further, a completely positive $\mathcal{E}$ is trace preserving iff L is an isometry.

1.4 Symmetry and randomness

In this section we briefly review some concepts and basics in probability theory and representation theory.

A *σ-algebra* (or *σ-field*) on a set Ω is a collection $\mathcal{F}$ of subsets of Ω that includes the empty subset and is closed under the operations of complement, countable unions and countable intersections. The pair $(\Omega, \mathcal{F})$ is called a *measurable space*. A function μ from $\mathcal{F}$ to the extended real number line is called a *measure* if it satisfies that, $\mu(A) \geq 0$ for all $A \in \mathcal{F}$, $\mu(\varnothing) = 0$, and $\mu(\cup_{k=1}^\infty A_k) = \sum_{k=1}^\infty \mu(A_k)$ for any countable collections $\{A_k\}_{k=1}^\infty$ of pairwise disjoint sets in $\mathcal{F}$. A triple $(\Omega, \mathcal{F}, \mu)$ is called a *measure space*. A probability measure, usually written Pr, is a measure such that $\Pr(\Omega) = 1$. A *probability space* is a measure space with a probability measure. Let $(\Omega, \mathcal{F})$ and $(\Omega', \mathcal{F}')$ be two measurable spaces, a function $f : \Omega \to \Omega'$ is called a *measurable function* if $f^{-1}(A') \in \mathcal{F}$ for all $A' \in \mathcal{F}'$.

Let $(\Omega, \mathcal{F}, \Pr)$ be a probability space and $(E, \mathcal{E})$ a measurable space. Then an *$(E, \mathcal{E})$-valued random variable* is a $(\mathcal{F}, \mathcal{E})$-measurable function $X : \Omega \to E$, that is, $X^{-1}(B) \in \mathcal{F}$ for any $B \in \mathcal{E}$. This allows to define a probability measure μ on $\mathcal{E}$, that is, $\mu(B) = \Pr(X^{-1}(B))$. When $E = \mathbb{R}$, X is called a *real-valued random variable*.

A *topological space* is set X, along with a collection $\mathcal{T}$ of subsets of X satisfying that $\mathcal{T}$ contains the empty set and that $\mathcal{T}$ is closed under the operations of arbitrary union and finite intersection. The elements of $\mathcal{T}$ are called open sets. The *Borel algebra* on a topological space X is the smallest σ-algebra containing all open sets. Thus one may define a measure on a topological space.

The Markov inequality and its derived inequalities will be used repeatedly in this thesis. The power of probabilistic method also partially relies on the phenomenon of concentration of measure. It is often an effect related to high dimension or a large number of variables, for which functions with small local oscillation are nearly constant [Led05]. A well-known and illustrative example is the Euclidean sphere S^{n-1} in $\mathbb{R}^n$

for large dimension n, on which the uniform measure is highly concentrated near any equator. This observation may be formulated equivalently on functions as below, known as Levy's lemma [LPH51].

Let $n > 2$ and let $f : S^{n-1} \to \mathbb{R}$ be an L-Lipschitz function, that is, $|f(x) - f(y)| \leq L\,\mathrm{geo}(x, y)$ for any $x, y \in S^{n-1}$ where geo denotes the geodesic metric. Let M be a central value of f, that is, $\Pr(f \geq M) \geq \frac{1}{4}$ and $\Pr(f \leq M) \geq \frac{1}{4}$ for Pr denoting the uniform probability measure on S^{n-1}. Then for any $\varepsilon > 0$,

$$\Pr(f - M \geq \varepsilon) \leq \mathrm{e}^{-\frac{1}{4}n\varepsilon^2/L^2}\,. \tag{1.2}$$

Another strong result is the Dvoretzky's theorem which was conjectured by Grothendieck [Gro56], first proved by Dvoretzky [Dvo61], and then refined by Milman [Mil71]. Roughly speaking, it states, in terms of function instead of convex body, that an L-Lipschitz function f on a unit sphere S^{n-1} with uniform probability measure is at most ε-far from some central value on the intersection of S^{n-1} and a random $O(n\varepsilon^2/L^2)$-dimensional subspace with probability $1 - \exp(-\Omega(n\varepsilon^2/L^2))$. The details will be introduced in later chapter.

When there is difficulty in explicitly constructing some object of interest, one may consider random choices of these objects and then show that the random object possesses the property with a nonzero probability. Thus various concentration inequalities may be useful for this technique. This probabilistic existence argument has been successfully ultilized in many fields to prove the existence of some desired object.

We now turn to introduction to representation theory. A *group* is a set together with an associative binary operation such that an identity element exists and every element has an inverse. A *representation* of a group G is a vector space V together with a homomorphism from G to $\mathrm{End}(V)$, i.e. a map $\mathsf{R} : G \to \mathrm{End}(V)$ such that $\mathsf{R}(g_1g_2) = \mathsf{R}(g_1)\mathsf{R}(g_2)$ for any $g_1, g_2 \in G$. Here, V and R are called *representation space* and *representation map* respectively. When clear from the context, we denote a representation (R, V) simply by the representation space V. Sometimes we write $g \cdot v$ or gv to denote $\mathsf{R}(g)v$ for $g \in G$ and $v \in V$. Thoughout this thesis we consider only representations on finite-dimensional complex vector spaces. The *character* for a representation V is a map $\chi : G \to \mathbb{C}$ given by $\chi(g) = \mathrm{tr}(\mathsf{R}(g))$.

For any two vector spaces V_1 and V_2, define $\mathrm{hom}(V_1, V_2)$ to be the set of linear transformations from V_1 to V_2. If V_1 and V_2 are representations of G with representation maps R_1 and R_2 respectively, then define $\mathrm{hom}_G(V_1, V_2) := \{M \in \mathrm{hom}(V_1, V_2) : M\mathsf{R}_1(g) = \mathsf{R}_2(g)M \text{ for any } g \in G\}$. If $\mathrm{hom}_G(V_1, V_2)$ contains an invertible map, we say V_1 and V_2 are equivalent representation and always identify them.

For a representation (R, V) of G, a subspace W is called an *invariant subspace* under G if and only if $\mathsf{R}(g)v \subset W$ for any $g \in G, v \in W$. We say a representation (R, V) is *irreducible* (and call it an irreducible representation, or *irrep* for short) if the only invariant subspaces of V of G are the empty subspace and the entire space V.

In some situations it is convenient to work with the *group algebra* (also known as *group ring*) rather than the group itself. For a group G, its group algebra $\mathbb{C}[G]$ is the algebra consisting of all complex-valued functions $x : G \to \mathbb{C}, g \mapsto x(g)$. Equivalently, any element x in $\mathbb{C}[G]$ can be written as $x = \sum_{g\in G} x(g)g$ for $x(g) \in \mathbb{C}$. For two elements $x = \sum_{g\in G} x(g)g$ and $y = \sum_{g\in G} y(g)g$, one has

$$xy = \sum_{g\in G} \left(\sum_{h\in G} x(h)y(h^{-1}g) \right) g \,,$$

and equivalently,

$$(xy)(g) = \sum_{h\in G} x(h)y(h^{-1}g) \text{ for any } g \in G \,.$$

The irreps of symmetric group and some Lie groups are briefly introduced as follows. Let $\mathrm{Type}(n, d)$ denote the set of all d-tuples of nonnegative integers that sum to n. A *partition* of n is a nonincreasing tuple of nonnegative integers that sum to n, and let $\mathrm{Par}(n)$ denote the set of all partitions of n. Let $\mathrm{Par}(n, d)$ denote the set of all nonincreasing d-tuples of nonnegative integers that sum to n, i.e., $\mathrm{Par}(n, d) = \{(\lambda_1, \ldots, \lambda_d) : \lambda_1 \geq \cdots \geq \lambda_d \geq 0, \sum \lambda_i = n\}$. Two partitions $(\lambda_1, \ldots, \lambda_k)$ and $(\lambda_1, \ldots, \lambda_k, 0, \ldots, 0)$ are considered equivalent if they differ only in padded zeros. A partition λ can be identified with a *Young diagram* in which there are λ_i empty boxes in the i-th row. A *Young tableau of shape λ and alphabet A* is a result of filling each box in Young diagram λ with a number in A. A *standard Young tableau* is a Young tableau filled with numbers 1 through n which strictly increase from left to right and from top to bottom. Denote by $\mathrm{STab}(\lambda)$ the set of standard Young tableaux of shape λ. A *semistandard Young tableau* is a Young tableau filled with numbers which weekly increase from left to right and strictly increase from top to bottom. Denote by $\mathrm{SSTab}(\lambda, d)$ the set of semistandard Young tableaux of shape λ and alphabet $\{1, \ldots, d\}$.

As a special finite group, the symmetric group S_n of degree n consists of all permutations on n symbols. Throughout this thesis a permutation is written as a product of permutation cycles which is unique up to the ordering of the cycles. Given any $\pi \in S_n$, define an isometry W_π such that

$$W_\pi(u_1 \otimes \cdots \otimes u_n) = u_{\pi^{-1}(1)} \otimes \cdots \otimes u_{\pi^{-1}(n)},$$

for any $u_i \in \mathcal{H}_i$. For the sake of brevity, we sometimes write

$$\pi u := W_\pi u \text{ and } \pi \cdot X := W_\pi X W_\pi^\dagger$$

to denote the action of $\pi \in S_n$ on $u \in \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$ and, respectively, on $X \in \mathcal{L}(\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n)$. Thus W_τ and its action are defined by linearity for any $\tau \in \mathbb{C}[S_n]$.

Since the conjugacy classes of S_n are labeled by partitions of n, the number of inequivalent irreps of S_n is equal to the number of partitions of n. Indeed, each partition λ corresponds to an irrep of S_n. Choose a standard Young tableau T of shape $\lambda \in \mathrm{Par}(n)$. Notice that the choice of standard Young tableau for each partition does not matter since different choices for the same partition give rise to equivalent irreps. Denote by P_T the set of permutations π that permute numbers within each row in T, i.e., i and $\pi(i)$ are in the same row for each $i \in [n]$, and by Q_T the set of permutations π that permute numbers within each column in T. The *Young symmetrizer* c_T is thus defined as $c_T := \sum_{p\in P_T, q\in Q_T} pq\,\mathrm{sgn}(q)$. The left-modules $V_\lambda^S := \mathbb{C}[S_n]c_T$ are called the *Specht modules* of S_n. The character for the irrep V_λ^S is denoted by χ_λ^S. Now we give a basis for V_λ^S. Let $\{|1\rangle, \ldots, |d\rangle\}$ be standard orthonomal basis for $\mathbb{C}^d$. For each $T \in \mathrm{STab}(\lambda)$, define v_T to be the basis vector with i at the positions in row i of T, e.g., $v_T = |12121\rangle$ for $T = \begin{array}{|c|c|c|}\hline 1 & 3 & 5 \\ \hline 2 & 4 \\ \cline{1-2}\end{array}$. Thus $\{c_T v_T : T \in \mathrm{STab}(\lambda)\}$ is a non-orthogonal basis of V_λ^S. By decomposing the irrep of S_k as direct sum of irreps of S_{k-1} using the branching rule, one can obtain an orthogonal basis known as *Young-Yamanouchi basis* [Jam06].

The dimension of V_λ^S, for $\lambda \in \mathrm{Par}(n)$, is given by

$$\dim V_\lambda^S = \frac{n!}{\widetilde{\lambda}_1! \cdots \widetilde{\lambda}_d!} \prod_{1\le i<j\le d} \left(\widetilde{\lambda}_i - \widetilde{\lambda}_j\right), \tag{1.3}$$

where $\widetilde{\lambda}_i := \lambda_i + d - i$. It can be bounded by [Hay02]

$$\binom{n}{\lambda}(n+d)^{-d(d-1)/2} \le \dim V_\lambda^S \le \binom{n}{\lambda} \tag{1.4}$$

and furthermore

$$\exp(nH(\bar{\lambda}))(n+d)^{-d(d+1)/2} \leq \dim V_\lambda^S \leq \exp(nH(\bar{\lambda}))\,, \tag{1.5}$$

where $\bar{\lambda} := \lambda/n$.

The representation theory of compact Lie group is always studied via that of Lie algebra. A *Lie algebra* is a vector space $\mathfrak{g}$ together with a skew-symmetric bilinear map $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ satisfying the Jacobi identity $[[a,b],c] + [[b,c],a] + [[c,a],b] = 0$ for any $a,b,c \in \mathfrak{g}$. This map is called Lie bracket. A representation of Lie algebra $\mathfrak{g}$ is a vector space together with a homomorphism $\mathsf{f} : \mathfrak{g} \to \mathrm{End}V$ such that the Lie bracket is preserved, i.e., $\mathsf{f}([a,b]) = \mathsf{f}(a)\mathsf{f}(b) - \mathsf{f}(b)\mathsf{f}(a)$ for any $a,b \in \mathfrak{g}$. As the Lie algebra can be thought of as the tangent space at the identity of a Lie group, the representation theories of Lie algebras and Lie groups parallel each other in some sense. For an introduction to Lie group and Lie algebra and their representation theory see standard literature, e.g. [FH13, Hal15, GW09].

Define the *torus* $\mathrm{U}(1)^{\times d} = \mathrm{U}(1) \times \cdots \times \mathrm{U}(1)$ as the subgroup of diagonal matrices. Let (R, V) be a rational irrep of $\mathrm{U}(d)$. Since the rational irrep of $\mathrm{U}(1)$ is given by $u \mapsto u^k$ for $k \in \mathbb{Z}$ and $u \in \mathrm{U}(1)$, V as a representation of the torus $\mathrm{U}(1)^{\times d}$ is decomposed as a direct sum of orthogonal subspaces labeled by $\mu \in \mathbb{Z}^d$. The subspace corresponding to μ is $V(\mu) = \{v \in V : \mathsf{R}(\mathrm{diag}(x_1, \ldots, x_d))v = x_1^{\mu_1} \cdots x_d^{\mu_d} v \ \forall x_i \neq 0\}$, called the *$\mu$-weight space*. Any $v \in V(\mu)$ is called a *weight vector* with *weight* μ. We say μ' majorizes μ if $\sum_{i=1}^d \mu_i' = \sum_{i=1}^d \mu_i$ and $\sum_{i=1}^k \mu_i' \geq \sum_{i=1}^k \mu_i$ for each k, denoted $\mu' \succ \mu$. It turns out that every irrep has a *highest weight* μ^{h}, which majorizes any other weights in this irrep, and that any d-tuple λ of integers as a highest weight gives rise to an irrep denoted V_λ^L. The tuples that differ only in order of numbers yield isomorphic representations. Among these irreps there are two special ones: the symmetric subspace $\vee^n \mathbb{C}^d := \{v \in (\mathbb{C}^d)^{\otimes n} : W_\pi v = v \ \forall \pi \in S_n\}$ corresponding to partition (n), and the antisymmetric subspace $\wedge^n \mathbb{C}^d := \{v \in (\mathbb{C}^d)^{\otimes n} : W_\pi v = \mathrm{sgn}(\pi) v \ \forall \pi \in S_n\}$ corresponding to partition $(1, \ldots, 1)$ for $d \geq n$.

By Weyl's tensorial construction, V_λ^L is isomorphic to $c_T(\mathbb{C}^d)^{\otimes n}$ for any $\lambda \in \mathrm{Par}(n,d)$ and for any $T \in \mathrm{STab}(\lambda)$. Since $\mathfrak{gl}(d)$ (resp. $\mathfrak{sl}(d)$) is the complexification of the real Lie algebra $\mathfrak{u}(d)$ (resp. $\mathfrak{su}(d)$) respectively, a representation of $\mathfrak{gl}(d)$ (resp. $\mathfrak{sl}(d)$) is irreducible iff so is the corresponding representation of $\mathfrak{u}(d)$ (resp. $\mathfrak{su}(d)$). The irrep of $\mathrm{U}(d)$ (resp. $\mathrm{SU}(d)$) can be extended to $\mathrm{GL}(d)$ (resp. $\mathrm{SL}(d)$). As the representation matrices for $(\lambda_1, \ldots, \lambda_d)$ and that for $(\lambda_1 + k, \ldots, \lambda_d + k)$ are isomorphic up to a factor $(\det g)^k$ for $g \in \mathrm{U}(d)$, the irreps of $\mathrm{SU}(d)$ and $\mathrm{SL}(d)$ can be labeled by d-tuples with the last number being zero.

For each $\lambda \in \mathrm{Par}(n,d)$, $R \in \mathrm{STab}(\lambda)$, $T \in \mathrm{SSTab}(\lambda,d)$ and $i \in [n]$, let m_i be the number such that the box containing i in R and the box containing m_i in T are in the same place, and define e_T to be the vector $|m_1,\ldots,m_n\rangle$, e.g., $e_T = |1323\rangle$ for $R = \begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\ \hline 3 \\ \cline{1-1}\end{array}$ and $T = \begin{array}{|c|c|c|}\hline 1 & 3 & 3 \\ \hline 2 \\ \cline{1-1}\end{array}$. Thus for any choice of R, $\{c_R e_T : T \in \mathrm{SSTab}(\lambda,d)\}$ is a basis of V_λ^L, which is not orthogonal in the Euclidean space $(\mathbb{C}^d)^{\otimes n}$. By using the branching rule, one can obtain an orthogonal basis known as Gelfand-Tsetlin basis [GT50] which will be detailed and used in Chapter 3. In this basis the irrep of $\mathrm{U}(d)$ is unitary, that is, each representation matrix of U is unitary for $U \in \mathrm{U}(d)$.

The character χ_λ^L for the irrep V_λ^L is

$$\chi_\lambda^L(\mathrm{diag}(x_1,\ldots,x_d)) = \frac{\det(x_i^{\lambda_j+d-j})_{i,j=1}^d}{\det(x_i^{d-j})_{i,j=1}^d}\,. \tag{1.6}$$

The dimension of V_λ^L is given by

$$\dim V_\lambda^L = \prod_{1\le i<j\le d} \frac{\lambda_i - \lambda_j + j - i}{j-i}\,, \tag{1.7}$$

which is bounded above by $(n+1)^{d(d-1)/2}$ [CM06].

The following Schur-Weyl duality, relating finite-dimensional irreps of the general linear and symmetric groups, was first presented in Schur's thesis [Sch] and developed further by Weyl [Wey39]; see [FH13, EGH+11, GW09] for more details.

Theorem 1 (Schur-Weyl duality)**.** *Let G be either $\mathrm{GL}(d)$, $\mathrm{SL}(d)$, $\mathrm{U}(d)$ or $\mathrm{SU}(d)$. Let $\mathcal{A}$ denote the algebra in $\mathrm{End}(\mathbb{C}^d)^{\otimes n}$ generated by W_π for all $\pi \in S_n$, and let $\mathcal{B}$ denote the algebra in $\mathrm{End}(\mathbb{C}^d)^{\otimes n}$ generated by $U^{\otimes n}$ for all $U \in G$. Then $\mathcal{A}$ and $\mathcal{B}$ are commutants of each other, and this leads to the following decomposition*

$$(\mathbb{C}^d)^{\otimes n} \cong \bigoplus_{\lambda\in\mathrm{Par}(n,d)} V_\lambda^L \otimes V_\lambda^S\,, \tag{1.8}$$

where V_λ^L and V_λ^S are irreps of the Lie group G and the symmetric group S_n respectively.

The Haar measure μ on a compact group G is the unique normalized invariant Borel probability measure over G. It can be intuitively thought of as the uniform probability distribution over a group. More specifically, the measure μ is invariant with respect to both left and right multiplication by any group element:

$$\mu(gS) = \mu(S) = \mu(Sg)$$

for any Borel subset $S \subset G$ and group element g.

As an application of Schur-Weyl duality, the twirling operation stated as follows has been widely used in quantum information theory.

Lemma 2. *Given an operator $X \in \mathcal{L}((\mathbb{C}^d)^{\otimes n})$, denote*

$$\widetilde{X} := \int U^{\otimes n} X U^{\dagger,\otimes n}\, \mathrm{d}U\,, \tag{1.9}$$

where $\mathrm{d}U$ is the normalized Haar measure on $\mathrm{U}(\mathbb{C}^d)$ or $\mathrm{SU}(\mathbb{C}^d)$.

(i) When $n = 2$,

$$\widetilde{X} = \alpha \mathbb{1} + \beta W\,, \tag{1.10}$$

where

$$\alpha = \frac{d\,\mathrm{tr}(X) - \mathrm{tr}(XW)}{d^3 - d},\ \beta = \frac{d\,\mathrm{tr}(XW) - \mathrm{tr}(X)}{d^3 - d}\,,$$

and W is the swap operator satisfying $W(|u\rangle \otimes |v\rangle) = |v\rangle \otimes |u\rangle$ for any $|u\rangle, |v\rangle \in \mathbb{C}^d$.

(ii) When $X = |\psi\rangle\langle\psi|^{\otimes n}$,

$$\widetilde{X} = \frac{\Pi_{\mathrm{sym}}}{\mathrm{tr}\,\Pi_{\mathrm{sym}}} = \frac{1}{d^{\uparrow n}} \sum_{\pi \in S_n} W_\pi\,, \tag{1.11}$$

where Π_{sym} is the projector onto the symmetric subspace $\vee^n \mathbb{C}^d$ and $d^{\uparrow n} := d(d+1)\cdots(d+n-1)$ denotes the rising factorial.

Proof. (i) The integration is invariant under unitary conjugation, i.e., $(U \otimes U)\widetilde{X}(U \otimes U)^\dagger = \widetilde{X}$ for any $U \in \mathrm{U}(\mathcal{H})$. Thus $\widetilde{X}$ is a linear combination of operators $\mathbb{1}$ and W due to Schur-Weyl duality. The coefficients α and β are determined using the conditions $\mathrm{tr}(\widetilde{X}) = \mathrm{tr}\, X$ and $\mathrm{tr}(W\widetilde{X}) = \mathrm{tr}(WX)$.

(ii) Since $\widetilde{X}$ commutes with each $U^{\otimes n}$ and maps the irrep $\vee^n \mathbb{C}^d$ to itself, by Schur's lemma, $\widetilde{X}$ equals Π_{sym} up to a scalar. Since $W_\pi \widetilde{X} = \widetilde{X}$ for each $\pi \in S_n$, $\widetilde{X} = \frac{1}{f(n)} \sum_{\pi \in S_n} W_\pi$ for some $f(n)$. Thus $f(n) = \sum_{\pi \in S_n} \mathrm{tr}\, W_\pi = \sum_{\pi \in S_n} d^{\#\pi} =: \sum_{j=1}^n h(n,j) d^j$, where $\#\pi$ denotes the number of cycles in π. Denote by $\Theta_{n,j}$ the set of permutations in S_n having j cycles, and then $h(n,j) = |\Theta_{n,j}|$. For $\pi \in \Theta_{n,j}$, if $\pi(1) = 1$, there are $h(n-1, j-1)$ such permutations; if $\pi(1) = t \neq 1$, by identifying numbers 1 and t, there are $(n-1)h(n-1,j)$ such permutations. Thus

$h(n, j) = h(n-1, j-1) + (n-1)h(n-1, j)$, and $f(n) = (d+n-1)f(n-1)$ follows. $\square$

Chapter 2

On multiple nonadditivity of minimum output Rényi entropy

Many quantities in quantum information theory have regularization form, such as various forms of capacity of a given quantum channel. A natural and important question is whether the quantities, such as the entanglement of formation, Holevo capacity, and minimum output entropy, are additive under tensor product. In this chapter we study the nonadditivity of Holevo capacity, or equivalently that of minimum output entropy. Several additivity results for the minimum output entropy were established for particular classes of channels, including unital qubit channels [Kin02], entanglement breaking channels [Sho02], and depolarizing channel [Kin03].

The maximum output p-norm of a channel $\mathcal{E}$ is $\nu_p(\mathcal{E}) := \max_\rho \|\mathcal{E}(\rho)\|_p$ for $p > 1$ and the minimum output Rényi entropy is $H_p(\mathcal{E}) = \frac{p}{1-p} \ln \nu_p(\mathcal{E})$. Thus the additivity of minimum output Rényi entropy can be turned into multiplicativity of the maximum output p-norm. Werner and Holevo [WH02] first constructed a counterexample $\mathcal{E}_{\mathrm{wh}}(\rho) = \frac{1}{d-1}(\mathrm{tr}(\rho)\mathbb{1} - \rho^\intercal)$ for $\rho \in \mathcal{D}(d)$ which satisfies that $\nu_p(\mathcal{E}_{\mathrm{wh}}^{\otimes 2}) > \nu_p(\mathcal{E}_{\mathrm{wh}})^2$ for $d = 3$ and $p > 4.79$. This Werner-Holevo channel (or called antisymmetric channel) $\mathcal{E}_{\mathrm{wh}}$ violates the additivity conjecture also for other choices of dimension d and parameter p. Hayden and Winter [Win07, Hay07, HW08] employed random unitary channel and random subspace in high dimension to yield counterexamples to the additivity conjecture of minimum output p-Rényi entropy for all $p > 1$. Note that the case $p = 1$ corresponds to the von Neumann entropy. Hastings [Has09] then showed that the minimum output von Neumann entropy is not additive, and equivalently, the Holevo capacity is not additive, settling a longstanding open problem in quantum information theory. It applied a probabilistic method to prove the existence of counterexample, using an unusual measure on quantum states. A detailed elucidation of Hastings' proof was given in [FKM10]. The reasoning in [Has09] was made clearer in [BH10] by explicitly using a measure concentration argument. By using the Dvoretzky's theorem, a more concise disproof to the conjecture was provided in [ASW11]. In high dimension, almost every channel and its conjugate violate the additivity conjecture, but explicit construction of such channel is difficult. A constructive counterexample for any $p > 2$ was obtained in [GHP10] using the antisymmetric subspace, and no constructive counterexample to the minimum output von Neumann entropy is known up to now.

The works mentioned above studied the additivity property all using a bipartite state as input of a pair of channels. An important question is thus whether multipartite entangled state is also useful for the subadditivity of the minimum output entropy of quantum channels. If only bipartite entangled state is useful, then the regularized form of the channel capacity may be formulated as a simpler expression. To deal with this question, it was shown in [FN14] that the tensor product of maximally entangled states asymptotically yields the minimum output entropy of random channel $(\mathcal{N} \otimes \mathcal{N}^*)^{\otimes r}$. This phenomenon also occurs for the random channel $\mathcal{N}^{\otimes 2r}$ where $\mathcal{N}$ is induced by random orthogonal matrices via Stinespring dilation [FN17]. It thus seems that only the bipartite entanglement could be the state achieving the minimum output entropy. In order to further study the multiple additivity of minimum output entropy, we consider alternative construction of random quantum channels. In order to conveniently do the calculation for complicated random quantum channels, a graphical calculus was developed in [CN10], in which the asymptotic spectrum of output state of a pair of high-dimensional random quantum channels with input being maximally entangled state was also studied. We will introduce in Section 2.2 the graphical calculus and use it to study the spectrum of output state of high-dimensional random channels, and study in Section 2.3 the multiple nonadditivity problem using concentration of measure argument.

2.1 Introduction and subadditivity of a pair of channels

It is a fundamental problem in quantum information theory to determine the capacity of a quantum-mechanical communication channel for conveying quantum, classical, or private information. For the case of classical information transmission, the first major result is due to Holevo, who proved in 1973 [Hol73] that the maximum amount of information that can be extracted from an ensemble of states $\eta := \{(p_i, \rho_i)\}$ with p_i

denoting probabilities is bounded above by

$$\chi(\eta) = H\Big(\sum\nolimits_i p_i\rho_i\Big) - \sum\nolimits_i p_i H(\rho_i)\,, \tag{2.1}$$

which is termed the *Holevo χ-quantity* (or *Holevo information*) of this ensemble. Here $H(\sigma) = -\operatorname{tr}(\sigma \ln \sigma)$ is the von Neumann entropy of density operator σ.

For any ensemble $\eta = \{(p_i, \rho_i)\}$ and any channel $\mathcal{N}$, one naturally defines an ensemble $\mathcal{N}(\eta) := \{(p_i, \mathcal{N}(\rho_i))\}$. Thus the *Holevo capacity* [Hol98, SW97] of channel $\mathcal{N}$ is defined as

$$\chi(\mathcal{N}) = \sup_\eta \chi(\mathcal{N}(\eta))\,, \tag{2.2}$$

where the supremum is over all ensembles of input states.

The classical capacity of a given channel $\mathcal{N}$ is defined as the regularization of the Holevo capacity

$$C(\mathcal{N}) = \lim_{n\to\infty} \frac{1}{n}\chi(\mathcal{N}^{\otimes n})\,. \tag{2.3}$$

It is always rather difficult to evaluate or even to estimate the classical capacity C of a given channel, unless its Holevo capacity is additive. If the Holevo capacity is additive, then the classical capacity has a simple single-letter characterization. Otherwise one can use the quantum channel collectively with entangled state as input to gain more efficiency of communication.

Shor [Sho04] showed that the additivity conjecture of Holevo capacity and that of minimum output entropy are equivalent. The minimum output entropy is defined as $H^{\min}(\mathcal{E}) := \min_\psi H(\mathcal{E}(\psi))$ with the minimum over all pure states ψ. Based on the works [Win07, Hay07, HW08], Hastings [Has09] showed that there exists channel $\mathcal{N}$ such that $H^{\min}(\mathcal{N}\otimes\mathcal{N}^*) < H^{\min}(\mathcal{N}) + H^{\min}(\mathcal{N}^*)$, finally falsifying the additivity conjecture. The proof was simplified in [ASW11] using the Doretzky'e theorem.

In the rest of this section we present a sketch of the proof in [ASW11], and then further simplify the proof. Let S, R, A, B be quantum systems of respective dimensions s, r, a, b such that $\mathcal{H}_S \otimes \mathcal{H}_R = \mathcal{H}_A \otimes \mathcal{H}_B$ (and obviously $sr = ab$). We assume $a \le b$. Consider the channel

$$\mathcal{E}_{S\to A}(\rho_S) = \operatorname{tr}_B(U(\rho_S \otimes \xi_r)U^\dagger) \tag{2.4}$$

with ξ_r being a fixed pure state on $\mathcal{H}_R$ in the following.

Let ϕ_q denote the maximally entangled state of rank q. Using the maximally entangled state as the input of $\mathcal{E}\otimes\mathcal{E}^*$, for any U we have

$$\begin{aligned}\langle(\mathcal{E}\otimes\mathcal{E}^*)\phi_s,\phi_a\rangle &= \left\langle (U\otimes U^*)(\phi_s\otimes\xi_r^{\otimes 2})(U\otimes U^*)^\dagger,\phi_a\otimes\mathbb{1}_{BB'}\right\rangle\\ &\geq \left\langle (U\otimes U^*)(\phi_s\otimes\xi_r^{\otimes 2})(U\otimes U^*)^\dagger,\phi_a\otimes\phi_b\right\rangle\\ &= \langle\phi_s\otimes\xi_r^{\otimes 2},\phi_s\otimes\phi_r\rangle = \frac{1}{r}=\frac{s}{ab}\,,\end{aligned}$$

where we used $(U\otimes U^*)|\phi_s\phi_r\rangle = |\phi_a\phi_b\rangle$ for any U. Thus $H^{\min}(\mathcal{E}\otimes\mathcal{E}^*)\leq H(\lambda)\leq 2\ln a-\frac{1}{r}(\ln\frac{a^2}{r}-2)$ for large r, where $\lambda:=(\frac{1}{r},\frac{1-1/r}{a^2-1},\ldots,\frac{1-1/r}{a^2-1})\in\mathbb{R}^{a^2}$ and the first inequality is due to Schur-concavity of H. When $r=ca$ for some constant c, there is a constant C such that for any channel $\mathcal{E}_{S\to A}$,

$$H^{\min}(\mathcal{E}\otimes\mathcal{E}^*)\leq 2\ln a - C\frac{\ln a}{a}\,. \tag{2.5}$$

In order to obtain a lower bound on $H^{\min}(\mathcal{E})=H^{\min}(\mathcal{E}^*)$, we consider the random channel defined in (2.4) where $U_{SR\to AB}$ is a Haar-random unitary. Based on previous works [Has09, BH10], Aubrun, Szarek and Werner [ASW11] applied the Dvoretzky's theorem to show that with high probability a random channel $\mathcal{E}$ satisfies

$$H^{\min}(\mathcal{E})=H^{\min}(\mathcal{E}^*)\geq \ln a-\frac{4}{a}\,. \tag{2.6}$$

Combining Eqs. (2.5) and (2.6), the subadditivity of minimum output entropy, $H^{\min}(\mathcal{E}\otimes\mathcal{E}^*)<H^{\min}(\mathcal{E})+H^{\min}(\mathcal{E}^*)$, is thus established.

Theorem 3 (Dvoretzky's theorem for Lipschitz functions)**.** *There are constants c,c' such that the following holds. Let $f:S_{\mathbb{C}^n}\to\mathbb{R}$ be an L-Lipschitz and phase-invariant function, and μ_f be any central value of f. Let $0<\varepsilon<1$, $k\leq cn\varepsilon^2/L^2$, and E be a k-dimensional random subspace of $\mathbb{C}^n$ with respect to Haar measure. Then* $\Pr(\mathrm{osc}(f,S_E,\mu_f)>\varepsilon)\leq \mathrm{e}^{-c'n\varepsilon^2/L^2}$, *where $S_E:=S_{\mathbb{C}^n}\cap E$.*

In the above theorem, for a function $f:\Omega\to\mathbb{R}$, a subset $\Gamma\subset\Omega$ and a scalar $\alpha\in\mathbb{R}$ we denote $\mathrm{osc}(f,\Gamma,\alpha):=\sup_{x\in\Gamma}|f(x)-\alpha|$, the expression '$L$-Lipschitz' means that $|f(x)-f(y)|\leq\|x-y\|_2$ for any $x,y\in S_{\mathbb{C}^n}$, and 'phase-invariant' means that $f(x)=f(\mathrm{e}^{\mathrm{i}\theta}x)$ for any $x\in S_{\mathbb{C}^n}$ and any $\theta\in\mathbb{R}$.

For $\sigma\in\mathcal{D}(\mathbb{C}^a)$ with eigenvalues p_i, it holds that $H(\sigma)=-\sum_{i=1}^a p_i\ln p_i\geq -\ln\sum p_i^2=$

$\ln a - \ln(a \sum p_i^2) \geq \ln a - a \operatorname{tr} \sigma^2 + 1 = \ln a - a\|\sigma - \omega\|_2^2$. It follows that

$$\begin{aligned} H^{\min}(\mathcal{E}) &= \min_{\psi_S} H(\mathcal{E}(\psi_S)) \\ &\geq \min_{\psi_S}(\ln a - a\|\mathcal{E}(\psi_S) - \omega_A\|_2^2) \\ &= \ln a - a \max_{|\psi\rangle_{AB} \in \mathcal{S}(\mathcal{H}_S)} \|\psi_A - \omega_A\|_2^2 \,. \end{aligned} \tag{2.7}$$

Considering $\mathcal{H}_S$ as an s-dimensional random subspace of $\mathcal{H}_{AB}$, it suffices to show that

$$\mathscr{P} := \Pr_{\mathcal{H}_S}\Big(\max_{|\psi\rangle_{AB} \in \mathcal{S}(\mathcal{H}_S)} \|\psi_A - \omega_A\|_2 \leq 2/a\Big) > 0\,. \tag{2.8}$$

When this is the case, there exists a subspace $\mathcal{H}_S$ such that $H^{\min}(\mathcal{E}) \geq \ln a - \frac{4}{a}$ and hence $H^{\min}(\mathcal{E} \otimes \mathcal{E}^*) < H^{\min}(\mathcal{E}) + H^{\min}(\mathcal{E}^*)$. We now present a sketch of the proof in [ASW11] of Eq. (2.8) as below.

Define functions $g : \mathcal{S}(\mathcal{H}_{AB}) \to \mathbb{R}, |\psi\rangle_{AB} \mapsto \|\psi_A - \omega_A\|_2$ and $h : \mathcal{S}(\mathcal{H}_{AB}) \to \mathbb{R}, |\psi\rangle_{AB} \mapsto \sqrt{\|\psi_A\|_\infty}$. Denote $\Omega_t := \{|\psi\rangle \in \mathcal{S}(\mathcal{H}_{AB}) : \sqrt{\|\psi_A\|_\infty} \leq t\}$ where $t := 4a^{-1/2}$, and let g' denote the restriction of g on Ω_t. The functions g' and h are $2t$-Lipschitz and 1-Lipschitz respectively. Define $\tilde{g} : \mathcal{S}(\mathcal{H}_{AB}) \to \mathbb{R}, |\psi\rangle_{AB} \mapsto \inf_{|\varphi\rangle \in \Omega_t}(g(|\varphi\rangle) + 2t\||\psi\rangle - |\varphi\rangle\|_2)$ which is $2t$-Lipschitz on $\mathcal{S}(\mathcal{H}_{AB})$ and coincides with g' on Ω_t.

The bound $\sqrt{\|\psi_A\|_\infty} \leq \frac{1}{\sqrt{a}} + \frac{1+\varepsilon}{\sqrt{b}}$ holds with probability at least $1 - \mathrm{e}^{-a\varepsilon^2}$ for $a \leq b$ [AS17, Prop. 6.36]. It follows that $g \leq 3b^{-1/2}$ with high probability. Thus h has a central value $\alpha \leq 3a^{-1/2}$ and both functions g and $\tilde{g}$ have a central value $\beta \leq 3b^{-1/2}$ for large a. Take $b = 9a^2$. Then due to Dvoretzky's theorem, we have $\mathscr{P}_1 := \Pr_{\mathcal{H}_S}(\operatorname{osc}(\tilde{g}, \mathcal{S}(\mathcal{H}_S), \beta) \leq 1/a) \geq 1 - \mathrm{e}^{-cb}$ and $\mathscr{P}_2 := \Pr_{\mathcal{H}_S}(\operatorname{osc}(h, \mathcal{S}(\mathcal{H}_S), \alpha) \leq a^{-1/2}) \geq 1 - \mathrm{e}^{-cb}$ for some c. Therefore,

$$\begin{aligned} \mathscr{P} &= \Pr_{\mathcal{H}_S}(\operatorname{osc}(g, \mathcal{S}(\mathcal{H}_S), 0) \leq 2/a) \\ &\geq \Pr_{\mathcal{H}_S}\big(\operatorname{osc}(\tilde{g}, \mathcal{S}(\mathcal{H}_S), 0) \leq 2/a\big) + \Pr_{\mathcal{H}_S}\big(\operatorname{osc}(h, \mathcal{S}(\mathcal{H}_S), 0) \leq t\big) - 1 \geq 1 - 2\mathrm{e}^{-cb}\,, \end{aligned}$$

completing the proof of (2.8).

The proof in [ASW11] used the fact that the Lipschitz constant of the restriction of g on Ω_t is upper-bounded by $2t$. Indeed, the function g is 2-Lipschitz on the sphere $\mathcal{S}(\mathcal{H}_{AB})$; see [HLW06, Lemma III.8], or Lemma 7 in this thesis for general case. In what follows we directly apply Dvoretzky's theorem once, instead of twice, to the function g, further simplifying the argument.

Simpler proof for bounding (2.8). We now show that $\mathscr{P}$ is close to 1 in high dimension. Consider the function $g : |\psi\rangle_{AB} \mapsto \|\psi_A - \omega_A\|_2$ which is 2-Lipschitz on $\mathcal{S}(\mathcal{H}_{AB})$. Since $\sqrt{\|\psi_A\|_\infty} \leq \frac{1}{\sqrt{a}} + \frac{c'}{\sqrt{b}}$ holds with high probability for some constant $c' > 1$ and $\|\psi_A - \omega_A\|_2 \leq \sqrt{a}\|\psi_A - \omega_A\|_\infty \leq \frac{3}{\sqrt{b}}$, the function g has a central value $\beta \leq \frac{3}{\sqrt{b}}$. Using Dvoretzky's theorem, $\Pr_{\mathcal{H}_S}(\operatorname{osc}(g, \mathcal{S}(\mathcal{H}_S), \frac{3}{\sqrt{b}}) \leq \varepsilon) \geq 1 - \mathrm{e}^{-cab\varepsilon^2}$ for some constant c. Taking $b = 9a^2$ and $\varepsilon = \frac{1}{a}$, we thus have

$$\begin{aligned}
\mathscr{P} &= \Pr_{\mathcal{H}_S}(\operatorname{osc}(g, \mathcal{S}(\mathcal{H}_S), 0) \leq 2/a) \\
&\geq \Pr_{\mathcal{H}_S}(\operatorname{osc}(g, \mathcal{S}(\mathcal{H}_S), 1/a) \leq 1/a) \\
&\geq 1 - \mathrm{e}^{-ca}\,.
\end{aligned}$$

□

2.2 Minimum output entropy of a triple of random channels

In this section we investigate the asymptotic behavior of the output state of a triple of quantum channels using tools from [FN14]. A graphical representation of quantum state and quantum channel in the context of quantum information theory was proposed in [CN10]. This graphical calculus can be well suited for the computation of integration over unitary group, orthogonal group and symplectic group. We now review first the result in [CŚ06] on calculation of integration over unitary group, and then the graphical calculus for tensors [CN10].

The (unitary) Weingarten function $\mathrm{Wg}_{n,d}$, also written Wg_d or Wg, is the inverse of $\sum_{\pi\in S_n} \pi \operatorname{tr}(W_\pi)$ in the symmetric group algebra $\mathbb{C}[S_n]$, namely

$$\sum_{\pi} (\operatorname{tr} W_{\pi^{-1}})\mathrm{Wg}(\pi\sigma) = \delta_{\sigma,\mathrm{id}}\,, \tag{2.9}$$

where δ denotes Kronecker delta. The Weingarten function is a class function on symmetric group since so is its inverse.

Proposition 4 ([CŚ06, Corollary 2.4])**.** *Let U be a random unitary or random special unitary according to the Haar measure, and U_{kl} be its elements. Then for any*

$i_t, j_t, i'_t, j'_t \in \{1,2,\ldots,d\}$ *with* $t \in \{1,2,\ldots,n\}$,

$$\int_{\mathrm{U}(d)} U_{i_1 j_1}\cdots U_{i_n j_n} U^*_{i'_1 j'_1}\cdots U^*_{i'_n j'_n}\,\mathrm{d}U = \sum_{\pi,\sigma\in S_n} \delta_{i_{\sigma(1)} i'_1}\cdots\delta_{i_{\sigma(n)} i'_n}\delta_{j_1 j'_{\pi(1)}}\cdots\delta_{j_n j'_{\pi(n)}}\mathrm{Wg}(\pi\sigma)\,. \tag{2.10}$$

We give a concise proof as follows.

Proof. The character of S_n acting on $(\mathbb{C}^d)^{\otimes n}$ is written as $\chi^S := \sum_{\pi\in S_n}\pi\,\mathrm{tr}(W_\pi) \in \mathbb{C}[S_n]$. By the Schur-Weyl decomposition $(\mathbb{C}^d)^{\otimes n} \cong \bigoplus_{\lambda\in\mathrm{Par}(n,d)} V^L_\lambda\otimes V^S_\lambda$, it can be also written as $\chi^S = \sum_{\lambda\in\mathrm{Par}(n,d)}\dim(V^L_\lambda)\chi^S_\lambda$, where χ^S_λ is the character for irrep V^S_λ of S_n. Notice that $\sum_{\lambda\in\mathrm{Par}(n,d)} p_\lambda = \mathrm{id}$ where $p_\lambda := \frac{\dim V^S_\lambda}{n!}\chi^S_\lambda$, and $\chi^S_\lambda\chi^S_{\lambda'} = \frac{n!}{\dim V^S_\lambda}\delta_{\lambda,\lambda'}\chi^S_\lambda$. Take

$$\mathrm{Wg} = \frac{1}{n!^2}\sum_\lambda \frac{(\dim V^S_\lambda)^2}{\dim V^L_\lambda}\chi^S_\lambda\,,$$

where the sum is over $\mathrm{Par}(n)$ or $\mathrm{Par}(n,d)$, in either case it can be verified that the equality $\chi^S\mathrm{Wg} = \mathrm{id}$ holds.

Denote $\tilde{X} := \mathbb{E}_U U^{\otimes n} X U^{\dagger,\otimes n}$ where U is a Haar-random unitary or special unitary. Since $\tilde{X}$ commutes with each $U^{\otimes n}$, $\tilde{X} = \sum_\sigma c_\sigma W_\sigma$ for some numbers c_σ. Thus $\mathrm{tr}(XW_\pi) = \mathrm{tr}(\tilde{X}W_\pi) = \sum_\sigma c_\sigma\,\mathrm{tr}\,W_{\sigma\pi}$. It follows that $\sum_\pi \mathrm{tr}(XW_\pi)\mathrm{Wg}(\sigma'\pi) = \sum_\sigma c_\sigma\sum_\pi \mathrm{tr}(W_{\sigma\pi})\mathrm{Wg}(\sigma'\pi) = c_\sigma\delta_{\sigma,\sigma'}$, where the latter equality used (2.9), hence $c_\sigma = \sum_\pi \mathrm{tr}(XW_\pi)\mathrm{Wg}(\sigma\pi)$. Therefore $\tilde{X} = \sum_{\pi,\sigma}\mathrm{tr}(XW_\pi)\mathrm{Wg}(\pi\sigma)W_\sigma$. Then

$$\mathrm{tr}(\tilde{X}Y) = \mathbb{E}_U\,\mathrm{tr}(U^{\otimes n}XU^{\dagger,\otimes n}Y) = \sum_{\pi,\sigma\in S_n}\mathrm{tr}(XW_\pi)\,\mathrm{tr}(YW_\sigma)\mathrm{Wg}(\pi\sigma)\,. \tag{2.11}$$

Taking $X = |j_1\rangle\langle j'_1|\otimes\cdots\otimes|j_n\rangle\langle j'_n|$ and $Y = |i'_1\rangle\langle i_1|\otimes\cdots\otimes|i'_n\rangle\langle i_n|$, we have $\mathrm{tr}(\tilde{X}Y) = \mathbb{E}U_{i_1j_1}\cdots U_{i_nj_n}U^*_{i'_1j'_1}\cdots U^*_{i'_nj'_n}$. The result follows by noticing that $\mathrm{tr}(XW_\pi) = \delta_{j_1,j'_{\pi(1)}}\cdots\delta_{j_n,j'_{\pi(n)}}$ and $\mathrm{tr}(YW_\sigma) = \delta_{i_{\sigma(1)},i'_1}\cdots\delta_{i_{\sigma(n)},i'_n}$. □

Since (2.11) holds for any Y, we have

$$\mathbb{E}_U(U^{\otimes n}XU^{\dagger,\otimes n}) = \sum_{\pi,\sigma\in S_n}\mathrm{Wg}(\pi\sigma)\,\mathrm{tr}(XW_\pi)W_\sigma\,. \tag{2.12}$$

The following is some examples of Weingarten function:

$$\begin{aligned}
\mathrm{Wg}_d(1) &= \frac{1}{d} \\
\mathrm{Wg}_d(1^2) &= \frac{1}{d^2-1} \\
\mathrm{Wg}_d(2) &= \frac{-1}{d(d^2-1)} \\
\mathrm{Wg}_d(1^3) &= \frac{d^2-2}{d(d^2-1)(d^2-4)} \\
\mathrm{Wg}_d(12) &= \frac{-1}{(d^2-1)(d^2-4)}\,,
\end{aligned}$$

where the permutations are described by their cycle shapes, that is, $(1^{\mu_1}\cdots n^{\mu_n})$ denotes the conjugacy class in S_n which has μ_j cycles of length j for $j=1,\ldots,n$.

Let $\pi\in S_n$ be written as product of disjoint cycles: $\pi = C_1C_2\cdots C_{\#\pi}$. Denote $|\pi| := n-\#\pi$ and $\mathrm{Mob}(\pi) := \prod_i (-1)^{|C_i|}\mathrm{Cat}_{|C_i|}$, where $|C_i|$ is the number of elements in C_i minus 1, and the Catalan number is $\mathrm{Cat}_m = \frac{1}{m+1}\binom{2m}{m}$. The Weingarten function has the following asymptotics:

$$\mathrm{Wg}_{n,d}(\pi) = d^{-n-|\pi|}(\mathrm{Mob}(\pi)+O(d^{-2}))\,. \tag{2.13}$$

In particular, $\mathrm{Wg}(\mathrm{id}) \sim d^{-n}(1+O(d^{-2}))$.

We now introduce the graphic presentation of the unitary integration, which makes the calculation more convenient in some complicated situations; see [CN10] for more details (but notice some minor difference from here). Each tensor is represented by a box attached labels of different shapes corresponding to vector spaces. The label can be empty (white color) or filled (black color). A (p,q)-tensor is represented by a box with p white labels and q black labels. Besides boxes the diagram contains wires connecting the labels, which represents the contraction on tensor. One can compute expectation values of such diagrams containing Haar-random unitary box. There is a trick on dealing with the transpose of unitary. One can simply turn the color of each label of the unitaries $U^\dagger, U^\intercal$ to its opposite while replacing $U^\dagger, U^\intercal$ by U^*, U respectively to obtain a diagram without $U^\dagger$ or $U^\intercal$. First we index the boxes by positive integers. Each pair of permutations, (π,σ), in Eq. (2.10) will be used to eliminate the boxes of U and U^* as follows. Wires are added to connect the white label of the U box with index i from inside to the black label of the U^* box with index $\pi(i)$, or connect the black label of the U^* box with index i from inside to the white label of the U box with index $\sigma(i)$. After this action the unitary boxed are eliminated, and the remaining diagram is called an (π, σ)-*removal*.

See Fig. 2.1 for examples of representation of a matrix and a quantum channel. The integration formula (2.12) thus can be represented by the graphics in Fig. 2.2.

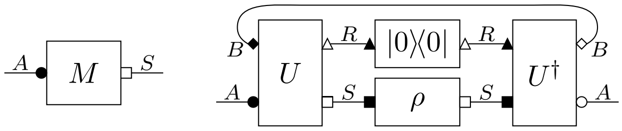

Figure 2.1: Diagrams for a matrix $M_{S\to A}$ and a quantum channel $\mathcal{E}_{S\to A}(\rho) = \mathrm{tr}_B(U(\rho \otimes |0\rangle\langle 0|)U^\dagger)$

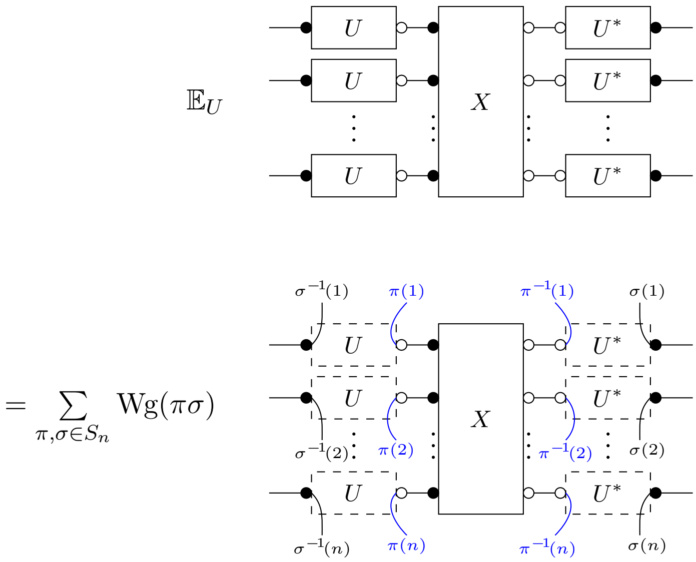

Figure 2.2: Graphical presentation for the integration $\mathbb{E}_U(U^{\otimes n} X U^{\dagger,\otimes n})$ in (2.12). The boxes for unitaries U and U^* have been removed after integration, and the lines out from the interior of these boxes are joined according to permutation π or σ.

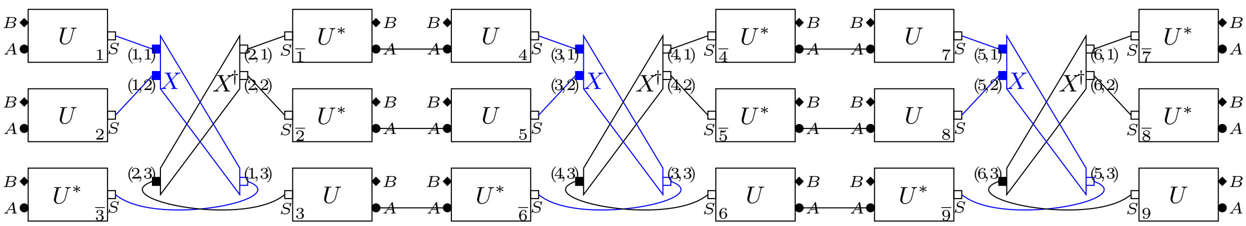

Figure 2.3: Diagram for calculating $\mathbb{E}\,\mathrm{tr}(Z^3)$. The pure state on system R is omitted.

Let S, R, A, B be quantum systems of respective dimensions s, r, a, b such that $\mathcal{H}_S \otimes \mathcal{H}_R = \mathcal{H}_A \otimes \mathcal{H}_B$ and $s = b$. For any unitary $U_{SR\to AB}$, define a quantum channel $\mathcal{E}_{S\to A}$ as

$$\mathcal{E}_{S\to A}(\rho) = \mathrm{tr}_B(U(\rho \otimes \xi_R)U^\dagger)\,, \tag{2.14}$$

where $\rho \in \mathcal{D}(\mathcal{H}_S)$ and ξ is a fixed pure state. Similarly the quantum channel $\mathcal{E}^*$ can be defined for the unitary U^*. A random quantum channel $\mathcal{E}$ is thus obtained by choosing a Haar-random unitary U. We now investigate the minimal ouput entropy of the random channel $\mathcal{E} \otimes \mathcal{E} \otimes \mathcal{E}^*$ when the pure input state is fixed following the method in [FN14].

Let $|\psi\rangle$ be a tripartite state on systems 1, 2 and 3. Using the operator-vector correspondence (1.1), the state $|\psi\rangle$ can be represented by an operator X from systems 1 and 2 to system 3. Denote by Z the output state of quantum channel $\mathcal{E}^U \otimes \mathcal{E}^U \otimes \mathcal{E}^{U^*}$ with input state ψ. We first calculate $\mathbb{E}\,\mathrm{tr}(Z^p)$ to study the asymptotic property of Z using the moment method. See Fig. 2.3 for the diagram of $\mathbb{E}\,\mathrm{tr}(Z^p)$ for the case $p = 3$.

The unitaries U in $\mathrm{tr}(Z^p)$ are labeled by integers $1, 2, \ldots, 3p$, and the unitaries U^* are labeled by $\overline{1}, \overline{2}, \ldots, \overline{3p}$. Define $\gamma \in S_{3p}$ as $\gamma(k) = \overline{k-3}$ if $k \mod 3 = 1$ or 2, $\gamma(k) = \overline{k+3}$ if $k \mod 3 = 0$ except that $\gamma(1) = \overline{3p-2}, \gamma(2) = \overline{3p-1}, \gamma(3p) = \overline{3}$. Using the integration formula (2.12), we write $\mathbb{E}\,\mathrm{tr}(Z^p)$ as

$$\mathbb{E}\,\mathrm{tr}(Z^p) = \sum_{\pi,\sigma\in S_{3p}} K_{\pi,\sigma} \text{ with } K_{\pi,\sigma} := a^{\#\sigma\gamma} b^{\#\sigma} f_X(\pi)\mathrm{Wg}(\pi\sigma)\,,$$

where $f_X(\pi) := \mathrm{tr}(W_\pi(X \otimes X^\dagger)^{\otimes 2p})$.

We relabel the $6p$ systems S of $2p$ boxes X, X^* as $[i, x]$ where $i \in [2p]$ denotes the position of box and $x \in [3]$ denotes the label attached on it. Consider the subset Θ of S_{3p}:

$$\Theta := \{\pi \in S_{3p} : \exists E \sqcup F = [2p] \text{ s.t. } |E| = |F| = 2p \text{ and } \forall i \in E, \exists j \in F, \pi([i, x]) = [j, y]\}\,. \tag{2.15}$$

Using technique in [FN14, Theorem A.1], denoting $t := \min_{\theta\in\Theta} |\pi^{-1}\theta|$, we can write

$$f_X(\pi) = \mathrm{tr}(Y_1 W_{\pi_1} Y_2 W_{\pi_2})\,. \tag{2.16}$$

where Y_1, Y_2 both are tensor products of X, X^* and $\mathbb{1}_{b^t}$. Thus, using Cauchy inequality,

$$|f_X(\pi)| \le \|Y_1\|_2 \|Y_2\|_2 = b^t\,. \tag{2.17}$$

Thus $|K_{\pi,\sigma}| \le a^{\#\sigma\gamma} b^{\#\sigma} b^t \mathrm{Wg}(\pi\sigma) \sim a^{\#\sigma\gamma-3p-|\pi\sigma|} b^{\#\sigma+t-3p-|\pi\sigma|}\mathrm{Mob}(\pi\sigma)$, where we have used the asymptotics of Wg function, (2.13). We are concerned with the case that b goes to infinity and a is fixed and finite.

It is known that the distance on S_n defined as $|\pi^{-1}\sigma| \equiv n - \#(\pi^{-1}\sigma)$ for $\pi, \sigma \in S_n$ satisfies the triangle inequality $|\pi_1^{-1}\pi_2| + |\pi_2^{-1}\pi_3| \ge |\pi_1^{-1}\pi_3|$ for any $\pi_1, \pi_2, \pi_3 \in S_n$. Using this triangle inequality, the power of b can be bounded as

$$\begin{aligned} &\#\sigma + \min_{\theta\in\Theta} |\pi^{-1}\theta| - 3p - |\pi\sigma| \\ &\quad\le \min_{\theta\in\Theta} |\pi^{-1}\theta| - |\pi| \\ &\quad\le \min_{\theta\in\Theta} |\theta| = 0\,, \end{aligned} \tag{2.18}$$

where the first inequality is saturated iff $|\sigma| + |\pi\sigma| = |\pi|$ (in this case we write $\sigma \le \pi^{-1}$) and the second is saturated iff $|\pi^{-1}\theta| \ge |\pi|$ for any $\theta \in \Theta$. Denoting $J_1 = \{3k - l : k \in [p], l \in [2]\}$ and $J_2 = \{3k : k \in [p]\}$, the set $\{(i,j) \in S_{3p} : i, j \in J_1 \text{ or } i, j \in J_2 \text{ or } \lceil i/3 \rceil \ne \lceil j/3 \rceil\}$ is contained in Θ. Following the reasoning in [FN14], the inequalities in (2.18) are saturated iff

$$(\pi, \sigma) \in \Xi := \{(\mathrm{id}, \mathrm{id}), ((13), \mathrm{id}), ((13), (13)), ((23), \mathrm{id}), ((23), (23))\}\,.$$

Denote $\Xi' := \{\mathrm{id}, (13), (23)\}$.

Using the integration formula, we have

$$\begin{aligned} \mathbb{E}Z &= \sum_{\pi,\sigma\in S_3} \mathrm{Wg}(\pi\sigma) f_X(\pi) b^{\#\sigma} W^A_\sigma \\ &\sim \sum_{(\pi,\sigma)\in\Xi} a^{-3-|\pi\sigma|}(-1)^{|\pi\sigma|} \frac{f_X(\pi)}{b^t} W^A_\sigma \\ &= a^{-3} \sum_{(\pi,\sigma)\in\Xi} \frac{f_X(\pi)}{b^t} R_\pi\,, \end{aligned}$$

where $t := \min_{\theta\in\Theta} |\pi^{-1}\theta|$ and $R_\pi := \sum_{\sigma:(\pi,\sigma)\in\Xi} (-a^{-1})^{|\pi\sigma|} W^A_\sigma$. Following [FN14] we call the sequence of X (or equivalently $|\psi\rangle$) well-behaved if $\frac{f_X(\pi)}{b^t} \to \alpha_\pi$ as $b \to \infty$. Then $\mathbb{E}Z$ converges almost surely to $Y = a^{-3}\sum_{\pi\in\Xi'} \alpha_\pi R_\pi$. Define $\rho_{\mathrm{id}} = a^{-3}\mathbb{1}$, $\rho_{(13)} = a^{-3}W^A_{(13)} + (a^{-3} - a^{-4})\mathbb{1}$, $\rho_{(23)} = a^{-3}W^A_{(23)} + (a^{-3} - a^{-4})\mathbb{1}$, all of which act on $(\mathbb{C}^a)^{\otimes 3}$. Thus Y can be rewritten as $Y = \sum_\sigma p_\sigma \rho_\sigma$ with $p_\sigma := \sum_{(\pi,\sigma)\in\Xi} (-1)^{|\pi\sigma|}\alpha_\pi$. Due to the Schur-concavity of von Neumann entropy, $H(\rho_{(13)}) = H(\rho_{(23)}) < H(\rho_{\mathrm{id}})$. Thus the largest $p_{(13)}$ or $p_{(23)}$ achieves least $H(Y)$. Since $p_{(13)} = \alpha_{(13)}$ achieves maximal value 1 when $|\psi\rangle$ is tensor of a maximally entangled state and a pure state. Thus we have the following:

Theorem 5. *Among all sequences of well-behaved input states, the tensor product of a bipartite maximally entangled state and a pure state achieves asymptotically the minimum output entropy of the random channel $\mathcal{E} \otimes \mathcal{E} \otimes \mathcal{E}^*$.*

2.3 On multiple nonadditivity of minimum output p-Rényi entropy

It was shown in [FW07] that weak additivity of minimum output entropy (for two identical channels) implies its strong additivity (for different ones). It can be easily extended to the case of triple channels following similar construction. We claim that for $1 \leq p \leq \infty$, if $H_p^{\min}(\mathcal{E}_1 \otimes \mathcal{E}_2 \otimes \mathcal{E}_3) < H_p^{\min}(\mathcal{E}_1) + H_p^{\min}(\mathcal{E}_2) + H_p^{\min}(\mathcal{E}_3)$ for some channels $\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3$, then there exists a channel $\mathcal{N}$ such that $H_p^{\min}(\mathcal{N}^{\otimes 3}) < 3H_p^{\min}(\mathcal{N})$. Indeed, let ρ_i be the states that minimize $H_p^{\min}(\mathcal{E}_i)$ for $i = 1, 2, 3$ respectively, and define $\mathcal{N}_1(\cdot) = \mathcal{E}_1(\cdot) \otimes \mathcal{E}_2(\rho_2) \otimes \mathcal{E}_3(\rho_3)$, $\mathcal{N}_2(\cdot) = \mathcal{E}_1(\rho_1) \otimes \mathcal{E}_2(\cdot) \otimes \mathcal{E}_3(\rho_3)$, $\mathcal{N}_3(\cdot) = \mathcal{E}_1(\rho_1) \otimes \mathcal{E}_2(\rho_2) \otimes \mathcal{E}_3(\cdot)$ and $\mathcal{N} = \mathcal{N}_1 \oplus \mathcal{N}_2 \oplus \mathcal{N}_3$. Thus $H_p^{\min}(\mathcal{N}) = \min_i H_p^{\min}(\mathcal{N}_i)$ and $H_p^{\min}(\mathcal{N}_1) = H_p^{\min}(\mathcal{N}_2) = H_p^{\min}(\mathcal{N}_3) = H_p^{\min}(\mathcal{E}_1) + H_p^{\min}(\mathcal{E}_2) + H_p^{\min}(\mathcal{E}_3)$. It follows that $H_p^{\min}(\mathcal{N}^{\otimes 3}) \leq H_p^{\min}(\mathcal{N}_1 \otimes \mathcal{N}_2 \otimes \mathcal{N}_3) = 2(H_p^{\min}(\mathcal{E}_1) + H_p^{\min}(\mathcal{E}_2) + H_p^{\min}(\mathcal{E}_3)) + H_p^{\min}(\mathcal{E}_1 \otimes \mathcal{E}_2 \otimes \mathcal{E}_3) < 3(H_p^{\min}(\mathcal{E}_1) + H_p^{\min}(\mathcal{E}_2) + H_p^{\min}(\mathcal{E}_3)) = 3H_p^{\min}(\mathcal{N})$, obtaining the claim.

Now consider a random unitary $U : \mathbb{C}^s \otimes \mathbb{C}^r \to \mathbb{C}^a \otimes \mathbb{C}^b$ with $sr = ab$. Define the channel $\mathcal{E}_1 = \mathcal{E}_2 \in \mathcal{L}(\mathcal{L}(\mathbb{C}^s), \mathcal{L}(\mathbb{C}^a))$ by $\rho_S \mapsto \operatorname{tr}_B(U(\rho_S \otimes \xi_R)U^\dagger)$. Similarly, define the channel $\mathcal{E}_3 \in \mathcal{L}(\mathcal{L}(\mathbb{C}^{s^2}), \mathcal{L}(\mathbb{C}^{a^2}))$ which is induced by $U^* \otimes U^*$ via Stinespring dilation.

The following two lemmas will be used to estimate $\nu_p(\mathcal{E}_3)$.

Lemma 6. *Consider the random state $|\psi\rangle_{AB} = U^{\otimes 2}|\xi\rangle$, where U is a Haar-random unitary in $\mathrm{SU}(\mathbb{C}^a \otimes \mathbb{C}^b)$, $|\xi\rangle \in (\mathbb{C}^d)^{\otimes 2}$ is a fixed state with $d = ab$, and the systems A and B have dimensions a^2 and b^2 respectively with $a \leq b$. Let $1 \leq p < \infty$ be an integer. Then $\lim_{d\to\infty} \frac{\mathbb{E}\|\psi_A\|_p}{((2p)!)^{1/p} a^{-2+2/p}} \leq 1$.*

Proof. We have

$$\begin{aligned}
(\mathbb{E}\|\psi_A\|_p)^p &\leq \mathbb{E}\|\psi_A\|_p^p \\
&= \mathbb{E}\operatorname{tr}\psi_A^p \\
&= \mathbb{E}\langle \psi_A^{\otimes p}, W_\gamma^A \rangle \\
&= \mathbb{E}\langle \psi_{AB}^{\otimes p}, W_\gamma^A \rangle,
\end{aligned}$$

where the first inequality is due to the convexity of x^p for $p \geq 1$, and $\gamma = (135 \cdots (2p-1))(246 \cdots (2p)) \in S_{2p}$.

By Schur-Weyl duality, $\mathbb{E}\psi_{AB}^{\otimes p} = \mathbb{E}U^{\otimes 2p}\xi^{\otimes p}U^{\dagger,\otimes 2p} = \sum_{\pi \in S_{2p}} c_\pi W_\pi$ where c_π's are coefficients. Thus for each $\sigma \in S_{2p}$, $\langle \mathbb{E}\psi^{\otimes p}, W_\sigma \rangle = \langle \sum_{\pi \in S_{2p}} c_\pi W_\pi, W_\sigma \rangle$, i.e., $\sum_{\pi \in S_{2p}} c_\pi d^{\#\pi\sigma} = \langle \xi^{\otimes p}, W_\sigma \rangle$, which constitute a system of $(2p)!$ linear equations in variables c_π. Consider the $(2p)! \times (2p)!$ matrix $M := (d^{\#\pi\sigma})_{\pi,\sigma \in S_{2p}}$, and the matrix M_π formed by replacing the column $(d^{\#\pi\sigma})_{\sigma \in S_{2p}}$ in M by the vector $(\langle \xi^{\otimes p}, W_\sigma \rangle)_{\sigma \in S_{2p}}$. Then by Cramer's rule, $c_\pi = \frac{\det M_\pi}{\det M}$. Since each element of M equals d^k with $k \in [2p]$ and d^{2p} appears exactly once in every column and row of M, $\lim_{d\to\infty} \frac{|\det M|}{d^{2p \cdot (2p)!}} = 1$. Since $|\langle \xi^{\otimes p}, W_\sigma \rangle| \leq 1$, we have $\lim_{d\to\infty} \frac{|\det M_\sigma|}{d^{2p\cdot((2p)!-1)}} \leq 1$. Thus for each π,

$$\lim_{d\to\infty} |c_\pi| d^{2p} \leq 1 \,. \tag{2.19}$$

Since $\langle W_\pi, W_\gamma^A \rangle = a^{\#\pi\gamma} b^{\#\pi}$, we have

$$\begin{aligned} d^{-2p}\langle W_\pi, W_\gamma^A \rangle &= a^{\#\pi\gamma - 2p} b^{\#\pi - 2p} \\ &\leq a^{\#\pi\gamma + \#\pi - 4p} \\ &= a^{-|\pi\gamma| - |\pi|} \\ &\leq a^{-|\gamma|} \\ &= a^{-2p+2} \,, \end{aligned}$$

where $|\pi| := 2p - \#\pi$ for $\pi \in S_{2p}$ is a metric, the first inequality used that $a \leq b$, and the second inequality used the triangle inequality.

It then follows that $(\mathbb{E}\|\psi_A\|_p)^p \leq \sum_{\pi \in S_{2p}} c_\pi \langle W_\pi, W_\gamma^A \rangle \leq \sum_{\pi \in S_{2p}} c_\pi d^{2p} a^{-2p+2}$. Thus by (2.19), $\lim_{d\to\infty} \left(\frac{\mathbb{E}\|\psi_A\|_p}{((2p)!)^{1/p} a^{-2+2/p}}\right)^p \leq \lim_{d\to\infty} \frac{1}{(2p)!} \sum_{\pi \in S_{2p}} |c_\pi| d^{2p} \leq 1$, completing the proof. ☐

The Lipschitz constants of von Neumann entropy and 2-norm of reduced state are upper-bounded by $\sqrt{8}\log|A|$ and 2 respectively [HLW06]. Following their approach, we now give an upper-bound for the case of general p-norm.

Lemma 7. *The Lipschitz constant of function $f : (\mathcal{S}(\mathcal{H}_{AB}), \|\cdot\|_2) \to \mathbb{R}, |\psi\rangle_{AB} \mapsto \|\psi_A\|_p$ with $p \geq 1$, is upper-bounded by 2.*

Proof. For fixed orthonormal bases $\{|j\rangle\}$ and $\{|k\rangle\}$, we write $|\psi\rangle_{AB} = \sum_{jk} \psi_{jk} |j\rangle_A |k\rangle_B = \sum_{jk}(t_{jk0} + \mathrm{i}t_{jk1})|j\rangle|k\rangle$ where t_{jkl}'s are reals. Thus $\psi_A = \sum_{j,j',k} \psi_{jk}\psi_{j'k}^* |j\rangle\langle j'|$. Consider first the commutative case. For diagonal $\psi_A = \sum_{jk} \psi_{jk}\psi_{jk}^* |j\rangle\langle j| = \sum_j x_j |j\rangle\langle j|$, where

$x_j := \sum_{kl} t_{jkl}^2$ with $l \in \{0,1\}$, we have $\|\psi_A\|_p = (\sum_j x_j^p)^{1/p}$. Then $\frac{\partial}{\partial t_{j'k'l'}}\|\psi_A\|_p = \frac{1}{p}(\sum_j x_j^p)^{1/p-1} p x_{j'}^{p-1} 2t_{j'k'l'}$. It follows that

$$\begin{aligned}
\|\nabla f(|\psi\rangle)\|_2^2 &= \sum_{j'k'l'} 4\big(\sum\nolimits_j x_j^p\big)^{2/p-2} x_{j'}^{2p-2} t_{j'k'l'}^2 \\
&= 4\big(\sum\nolimits_j x_j^p\big)^{2/p-2} \sum\nolimits_{j'} x_{j'}^{2p-1} \\
&\le 4\big(\sum\nolimits_j x_j^p\big)^{2/p-2} \big(\sum\nolimits_{j'} x_{j'}^p\big)^{2-1/p} \\
&= 4\|\psi_A\|_p \\
&\le 4\,,
\end{aligned}$$

where the first inequality is due to the convexity of $x^{2-1/p}$ for $p \ge 1$. Therefore, for any $|\psi\rangle_{AB}, |\varphi\rangle_{AB}$ with diagonal ψ_A, φ_A, we have $|\|\psi_A\|_p - \|\varphi_A\|_p| \le 2\||\psi\rangle - |\varphi\rangle\|_2$.

For general states $|\psi\rangle, |\varphi\rangle$, choosing the eigenvectors of φ_A as fixed basis, i.e., $\varphi_A = \mathrm{diag}(\varphi_A)$, we have $\|\psi_A\|_p \ge \|\mathrm{diag}(\psi_A)\|_p$ due to the Schur-convexity of p-norm and the Schur-Horn theorem. Thus $\|\psi_A\|_p - \|\varphi_A\|_p \le \|\psi_A\|_p - \|\mathrm{diag}(\varphi_A)\|_p \le 2\||\psi\rangle - |\varphi\rangle\|_2$, where we have assumed w.l.o.g. that $\|\psi_A\|_p \ge \|\varphi_A\|_p$, completing the proof. □

The following two lemmas will be used to estimate $\nu_p(\mathcal{E}_1)$.

Lemma 8. *Consider random state $|\psi\rangle = U|\xi\rangle$, where $|\xi\rangle \in \mathbb{C}^d$ is a fixed state with $d = ab$ and $a \le b$, $U \in \mathrm{SU}(\mathbb{C}^d)$ is distributed Haar-randomly, and the systems A and B have dimensions a and b respectively. Let $1 \le p < \infty$ be an integer. Then $\mathbb{E}\|\psi_A\|_p \le (p!)^{1/p} a^{1/p-1}$.*

Proof. Using similar method as in Lemma 6, we have $(\mathbb{E}\|\psi_A\|_p)^p \le \mathbb{E}\langle \psi_{AB}^{\otimes p}, W_\gamma^A\rangle$ for

$\gamma = (12\cdots p) \in S_p$. It holds that $\mathbb{E}\psi_{AB}^{\otimes p} = (d^{\uparrow p})^{-1}\sum_{\pi\in S_p} W_\pi$. Thus

$$\begin{aligned}(\mathbb{E}\|\psi_A\|_p)^p &\le (d^{\uparrow p})^{-1}\sum_{\pi\in S_p}\langle W_\pi, W_\gamma^A\rangle\\ &= (d^{\uparrow p})^{-1}\sum_{\pi\in S_p} a^{\#\pi\gamma}b^{\#\pi}\\ &\le \sum_{\pi\in S_p} a^{\#\pi\gamma-p}b^{\#\pi-p}\\ &\le \sum_{\pi\in S_p} a^{-|\pi\gamma|-|\pi|}\\ &\le \sum_{\pi\in S_p} a^{-|\gamma|}\\ &= p!a^{1-p}\,,\end{aligned}$$

where the second inequality is due to that $d^{\uparrow p} > d^p$. □

Lemma 9. *Let $s < \frac{ab\varepsilon^2}{16\ln(2/\delta)}$ be an integer for positive ε, δ. There exists an s-dimensional subspace of $\mathbb{C}^a \otimes \mathbb{C}^b$ that contains only states $|\psi\rangle$ such that $\|\psi_A\|_p < (p!)^{1/p}a^{-1+1/p} + \varepsilon + 2\delta$.*

Proof. Using Levy's lemma (1.2) on a sphere and Lemmas 7 and 8, we have

$$\Pr(\|\psi_A\|_p \ge (p!)^{1/p}a^{1/p-1} + \varepsilon) \le \mathrm{e}^{-\frac{1}{4}ab\varepsilon^2}\,.$$

Since any s-dimensional subspace $\mathcal{H}_S$ with Euclidean metric has a δ-net N that has $(\frac{2}{\delta})^{2s}$ elements, we have $\Pr_{\mathcal{H}_S}(\exists|\varphi\rangle \in N \text{ s.t. } \|\psi_A\|_p \ge (p!)^{1/p}a^{1/p-1}+\varepsilon) \le (\frac{2}{\delta})^{2s}\mathrm{e}^{-\frac{1}{4}ab\varepsilon^2}$. When s takes the required value, this probability is less than 1, and there exists a subspace $\mathcal{H}_S$ such that any state $|\varphi\rangle$ in N satisfies $\|\varphi_A\|_p < (p!)^{1/p}a^{1/p-1} + \varepsilon$. Then $\forall|\psi\rangle \in \mathcal{H}_S$, $\exists|\varphi\rangle \in N$ such that $\||\psi\rangle - |\varphi\rangle\|_2 \le \delta$ and $\|\psi_A\|_p \le \|\varphi_A\|_p + 2\||\psi\rangle - |\varphi\rangle\|_2 < (p!)^{1/p}a^{1/p-1} + \varepsilon + 2\delta$. □

It follows from the above lemma that

$$\nu_p(\mathcal{E}_1) := \max_{|\psi\rangle}\|\mathcal{E}_1(\psi)\|_p < (p!)^{1/p}a^{1/p-1} + \varepsilon + 2\delta\,. \tag{2.20}$$

Lemma 10. *Let s be an integer such that $s^2 < \frac{ab\varepsilon^2}{16\ln\frac{\pi}{\delta}}$. Then $\nu_p(\mathcal{E}_3) < 2((2p)!)^{1/p}a^{-2+2/p} + \varepsilon + 2\delta$.*

Proof. Consider the function $f : U \mapsto \| \operatorname{tr}_B(U^{\otimes 2}\xi U^{\dagger,\otimes 2})\|_p$. Due to Lemma 7,

$$\begin{aligned}|f(U)-f(V)| &\leq 2\|U^{\otimes 2}|\xi\rangle - V^{\otimes 2}|\xi\rangle\|_2\\ &\leq 2g_2(U^{\otimes 2}|\xi\rangle, V^{\otimes 2}|\xi\rangle)\\ &\leq 2g_2(U^{\otimes 2}, V^{\otimes 2})\\ &= 2g_2(U,V)^2\\ &\leq 2g_2(U,V)\,,\end{aligned}$$

where the first inequality is due to Lemma 7 and the last inequality is valid when $g_2(U,V) \leq 1$ for g_2 denoting the geodesic metric induced by 2-norm. For arbitrary U, V, there are unitaries $U_1, \ldots, U_n$ such that $U_1 \equiv U, U_n \equiv V$ and $g_2(U,V) = g_2(U_1,U_2) + \cdots + g_2(U_{n-1},U_n)$ with each summand less than 1. Then $\frac{|f(U)-f(V)|}{g_2(U,V)} \leq \frac{|f(U_1)-f(U_2)|+\cdots+|f(U_{n-1})-f(U_n)|}{g_2(U_1,U_2)+\cdots+g_2(U_{n-1},U_n)} \leq \max_k \frac{|f(U_k)-f(U_{k+1})|}{g_2(U_k,U_{k+1})} \leq 2$, that is, the Lipschitz constant of f is bounded by 2. By Lemma 6, for large d, $\mathbb{E}\|\psi_A\|_p \leq 2((2p)!)^{1/p}a^{-2+2/p}$. Due to Levy's lemma for $\mathrm{SU}(d)$ [AS17, Eq. 5.22], we have $\Pr(\|\psi_A\|_p \geq 2((2p)!)^{1/p}a^{-2+2/p} + \varepsilon) \leq \mathrm{e}^{-\frac{1}{8}ab\varepsilon^2}$.

Since any s^2-dimensional subspace $(\mathbb{C}^s)^{\otimes 2}$ equipped with geodesic metric g_2 induced by 2-norm has a δ-net N that has $(\frac{\pi}{\delta})^{2s^2}$ elements, we have $\Pr(\exists|\psi\rangle \in N \text{ s.t. } \|\psi_A\|_p \geq 2((2p)!)^{1/p}a^{-2+2/p}+\varepsilon)$ is less than $(\frac{\pi}{\delta})^{2s^2}\mathrm{e}^{-\frac{1}{8}ab\varepsilon^2}$. When s takes the required value, the probability is less than 1, and there exists a subspace of dimension s^2 such that any state in N satisfies $\|\psi_A\|_p < 2((2p)!)^{1/p}a^{-2+2/p} + \varepsilon$. Then $\forall|\psi\rangle \in \mathcal{H}_S$, $\exists|\varphi\rangle \in N$ such that $g_2(|\psi\rangle, |\varphi\rangle) \leq \delta$ and thus $\|\psi_A\|_p \leq \|\varphi_A\|_p + 2g_2(|\psi\rangle,|\varphi\rangle) < 2((2p)!)^{1/p}a^{-2+2/p} + \varepsilon + 2\delta$, completing the proof. ∎

Lemma 11. *For $s^2 = \frac{ab\varepsilon^2}{17\ln\frac{\pi}{\delta}}$, when $a = b$ gets large enough, we have $\nu_p(\mathcal{E}_1\otimes\mathcal{E}_2\otimes\mathcal{E}_3) \geq a^{-2}\frac{\varepsilon^2}{17\ln\frac{\pi}{\delta}}$.*

Proof. Denoting by ϕ_q the maximally entangled state of rank q, we have

$$\begin{aligned}\nu_p(\mathcal{E}_1\otimes\mathcal{E}_2\otimes\mathcal{E}_3) &\geq \|(\mathcal{E}_1\otimes\mathcal{E}_2\otimes\mathcal{E}_3)(\phi_s^{\otimes 2})\|_p\\ &\geq \langle(\mathcal{E}_1\otimes\mathcal{E}_2\otimes\mathcal{E}_3)\phi_s^{\otimes 2}, \phi_a^{\otimes 2}\rangle\\ &\geq \langle U^{\otimes 4}\phi_s^{\otimes 2}U^{\dagger,\otimes 4}, \phi_{ab}^{\otimes 2}\rangle\\ &= \langle \phi_s^{\otimes 2}, \phi_{ab}^{\otimes 2}\rangle\\ &= \frac{s^2}{a^2b^2} = a^{-2}\frac{\varepsilon^2}{17\ln(\pi/\delta)}\,.\end{aligned}$$

∎

Unfortunately, it turns out that the above construction and calculation in this section fail to show the multiple nonadditivity of minimum output p-entropy. The bounds on the minimum output p-entropy obtained in the lemmas above may be too loose. However, the multiple nonadditivity of minimum output p-entropy or von Neumann entropy using an alternative random channels deserves further study. For example, one may consider $\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3$ which are induced respectively via Stinespring dilation by $U, V, U \otimes V$ for independent random unitaries U and V, or random unitary channels $\mathcal{F}_1 = \frac{1}{n}\sum_{i=1}^n U_i(\cdot)U_i^\dagger, \mathcal{F}_2 = \frac{1}{n}\sum_{j=1}^n V_j(\cdot)V_j^\dagger, \mathcal{F}_3 = \frac{1}{n^2}\sum_{i,j=1}^n (U_i \otimes V_j)(\cdot)(U_i \otimes V_j)^\dagger$ where U_i, V_j are all independent. In order to find multiple nonadditivity, if exists, alternative models or methods may be needed.

Chapter 3

Generic entanglement in random invariant tensors

Quantum entanglement as a mysterious and striking phenomenon exhibits fundamental nonclassical manifestation of quantum world. Since the celebrated Einstein-Podolsky-Rosen (EPR) state as an entanglement was first discussed in 1935 [EPR35], this counterintuitive feature has been a central theme in many fields of physics. Nowadays quantum entanglement is a key resource in quantum information theory with applications in quantum cryptography [BB84], quantum dense coding [BW92], quantum teleportation [BBC+93] and distributed computation [FGM01]. Entanglement is also used in many quantum algorithms although its role is not quite clear yet.

An n-partite tensor on $\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$ is called *perfect* if it is proportional to an isometric tensor from A to B for any bipartite cut $A : B$ of the n systems such that the dimension of party A is no larger than that of B [PYHP15]. This class of entangled states has been studied in quantum information theory, condensed matter theory and quantum gravity [HQRY16, PYHP15, ADH15]. Such states are also known as absolutely maximally entangled states [HCL+12, Hel13]. It is obvious that the Greenberger-Horne-Zeilinger (GHZ) state $\frac{1}{\sqrt{2}}(|000\rangle + |111\rangle)$ is perfect while the W state $\frac{1}{\sqrt{3}}(|001\rangle + |010\rangle + |100\rangle)$ is not.

An invariant state is in the tensor power of an irreducible representation of SU(2), and it is invariant under global SU(2) action. Let V be an irrep of SU(2), then $|\psi\rangle \in V^{\otimes n}$ is an invariant state if $U^{\otimes n} \cdot |\psi\rangle = |\psi\rangle$ for any $U \in \mathrm{SU}(2)$. Invariant tensor is a significant notion in the theory of loop quantum gravity [Thi08, HMH07, AL04, RV14].

Because of the importance of invariant state and perfect state in quantum gravity, the concept of invariant perfect state was introduced and studied in [LHGZ17, LHRZ18] which also showed that random invariant state is asymptotically perfect as the representation dimension goes to infinity. This is similar to the well-known fact that a high-dimensional random pure bipartite state $|\psi\rangle_{AB}$ is nearly maximally entangled, that is, for large dimensions $d_A \leq d_B$, the entanglement entropy of ψ_{AB} is highly concentrated around its expectation $\mathbb{E}H(\psi_A)$ which is at least $\ln(d_A) - \frac{d_A}{2d_B}$ [HLW06]. The entanglement in bipartite $\mathfrak{su}(2)$-invariant state was also studied in [Sch03, Sch05] using various entanglement criteria.

In this chapter we study two classes of states: the states that could be slightly disturbed by the global SU(2) action, and the states that are kept invariant by SU(d) instead of SU(2). As extension of the concept of invariant state, the two classes of states will be shown still generically perfect. The SU(d)-invariant state will find important application in Chapter 2. The random invariant state serves as an alternative model for random quantum state, and it is tempting to apply the model into other scenarios.

In the calculation of the entanglement entropy, one needs to decompose tensor product of two irreps into a direct sum of irreps which is known as the *Clebsch-Gordan transform*, and the coefficients arising in this transform are called *Clebsch-Gordan coefficients* (CGCs). The Clebsch-Gordan transform is a basic tool in the analysis of invariant state. In physics, the SU(2) CGCs arise in the context of angular momentum coupling, and the SU(3) CGCs arise for example, in the context of quantum chromodynamics. The SU(d) CGCs are useful for some new standard model [Sla81] and for the numerical calculation for various models with SU(d) symmetry in which the application of Wigner-Eckart theorem needs knowledge of CGCs [AKHvD11]. The Clebsch-Gordan transform is also used to construct the Schur-Weyl transform [BCH06].

3.1 Representation theory of special unitary group

The Lie algebra $\mathfrak{su}(d)$ consists of all traceless skew-Hermitian matrices with Lie bracket being the commutator. Since SU(d) is a simply connected compact group, there is a one-to-one correspondence between the representations of SU(d) and $\mathfrak{su}(d)$. We introduce here the irreps of $\mathfrak{su}(d)$ from which the irreps of SU(d) can be obtained via the exponential mapping. We first briefly review the representation theory of $\mathfrak{su}(2)$, which can be found in standard textbook of quantum mechanics, and then that of $\mathfrak{su}(d)$. For detailed introduction to representation theory of Lie groups, we refer to [FH13, Hal15, GW09].

Angular momentum operators are Hermitian operators J_x, J_y, J_z satisfying $[J_k, J_l] = \mathrm{i}\epsilon_{klm}J_m$ where $k, l, m \in \{x, y, z\}$ and ϵ_{klm} is Levi-Civita symbol (we take $\hbar = 1$ for simplicity). The three operators $\mathrm{i}J_x, \mathrm{i}J_y, \mathrm{i}J_z$ constitute a basis of Lie algebra $\mathfrak{su}(2)$. The raising and lowering operators are $J_\pm := J_x \pm \mathrm{i}J_y$.

For any positive integer k, there exists a k-dimensional irrep of $\mathfrak{su}(2)$. Let j be such that $k = 2j+1$ and let $V_{(j)} := \mathrm{span}\{|j,j\rangle, |j,j-1\rangle, \dots, |j,-j\rangle\}$ denote this k-dimensional irrep. Any k-dimensional irrep of $\mathfrak{su}(2)$ is isomorphic to $V_{(j)}$. The action of $\mathfrak{su}(2)$ on $V_{(j)}$ can be written as

$$\begin{aligned} J_z|j,m\rangle &= m|j,m\rangle \\ J_\pm|j,m\rangle &= \sqrt{(j\mp m)(j\pm m+1)}|j,m\pm 1\rangle\,. \end{aligned} \tag{3.1}$$

The tensor product of irreps of $\mathfrak{su}(2)$ is decomposed into a direct sum of irreps:

$$V_{(j_1)} \otimes V_{(j_2)} = \bigoplus_j V_{(j)}\,, \tag{3.2}$$

which is *Clebsch-Gordan transform*. The basis state of $\bigoplus_j V_{(j)}$ can be written as

$$|(j_1j_2)jm\rangle = \sum_{m_1,m_2} |j_1m_1j_2m_2\rangle\langle j_1m_1j_2m_2|(j_1j_2)jm\rangle\,, \tag{3.3}$$

where $\langle j_1m_1j_2m_2|(j_1j_2)jm\rangle =: C^{j_1,j_2,j}_{m_1,m_2,m}$ are *Clebsch-Gordan coefficients*.

The Littlewood-Richardson rule is a combinatorical method to decompose a product of two Schur polynomials into a linear combination of Schur polynomials. It follows from the Littlewood-Richardson rule that $V_{(j)}$ has multiplicity one if and only if j is at least $|j_1-j_2|$ and at most j_1+j_2. The fact that $|j_1-j_2| \le j \le j_1+j_2$ can be also seen from the selection rule:

$$\text{if } m \neq m_1+m_2 \text{ then } C^{j_1j_2j}_{m_1m_2m} = 0\,, \tag{3.4}$$

which is obtained by applying J_z to both sides of (3.3).

The action of $J_\pm$ on both sides of (3.3) gives a recurrence relation for CGCs. By the phase convention $C^{j_1,j_2,j}_{j_1,j-j_1,j} > 0$, it is clear that all CGCs are real, which constitute an orthogonal matrix. It follows from the recurrence relation and the normalization condition that $C^{j_1,j_2,0}_{m_1,m_2,0} = \delta_{j_1,j_2}\delta_{m_1,-m_2}\frac{(-1)^{j_1-m_1}}{\sqrt{2j_1+1}}$.

A closed-form expression for the CGCs of $\mathfrak{su}(2)$, known as Racah's formula [Rac42], is

$$C^{j_1,j_2,j}_{m_1,m_2,m} = \delta_{m,m_1+m_2}\sqrt{\frac{(2j+1)(j+j_1-j_2)!(j-j_1+j_2)!(j_1+j_2-j)!}{(j_1+j_2+j+1)!}}$$
$$\sqrt{(j+m)!(j-m)!(j_1+m_1)!(j_1-m_1)!(j_2+m_2)!(j_2-m_2)!}\sum_k(-1)^k f_k^{-1}\,,$$

where $f_k = k!(j_1+j_2-j-k)!(j_1-m_1-k)!(j_2+m_2-k)!(j-j_2+m_1+k)!(j-j_1-m_2+k)!$.

We now introduce the representation theory for $\mathfrak{su}(d)$. Since $\mathfrak{sl}(d)$ is the complexification of the Lie algebra $\mathfrak{su}(d)$, there is a one-to-one correspondence of their irreps. It is convenient for our purpose to work with $\mathfrak{sl}(d)$ and $\mathfrak{gl}(d)$.

An irrep of $\mathfrak{gl}(d)$ is labeled by its highest weight, i.e., a sequence λ of d nonincreasing integers. Let V_λ denote the irrep of highest weight $\lambda = (\lambda^d_1, \ldots, \lambda^d_d)$. As irreps of $\mathfrak{sl}(d)$ and $\mathfrak{su}(d)$, $V_{(\lambda^d_1,\ldots,\lambda^d_d)}$ and $V_{(\lambda^d_1+c,\ldots,\lambda^d_d+c)}$ are equivalent for any integer c. Thus irreps of $\mathfrak{sl}(d)$ or $\mathfrak{su}(d)$ are labeled by sequences of d nonincreasing integers with the last integer being zero. The basis states of the irrep V_λ can be labeled by the *Gelfand-Tsetlin (GT) patterns*, which are arrays of integers, of the following form

$$\boldsymbol{\lambda} = \begin{pmatrix} \lambda^d_1 & & \lambda^d_2 & & \cdots & & \lambda^d_d \\ & \lambda^{d-1}_1 & & \cdots & & \lambda^{d-1}_{d-1} & \\ & & \ddots & & \iddots & & \\ & & & \lambda^1_1 & & & \end{pmatrix}.$$

Each GT pattern $\boldsymbol{\lambda}$ corresponds to a basis state, denoted $|\boldsymbol{\lambda}\rangle$, and via the branching rule, these states constitute an orthonormal basis of V_λ, i.e., $V_\lambda = \mathrm{span}\{|\boldsymbol{\lambda}\rangle : \boldsymbol{\lambda} \in \mathrm{GT}(\lambda)\}$.

For each k and l, the k-th element in row l of $\boldsymbol{\lambda}$ is denoted by $\boldsymbol{\lambda}^l_k$ or λ^l_k. Each row $\lambda^l := (\lambda^l_1, \ldots, \lambda^l_l)$, also written $\boldsymbol{\lambda}^l$, for $l = 1, \ldots, d$, is a nonincreasing sequence of integers. Here the superscript is used to distinguish different sequences of numbers. The top row of $\boldsymbol{\lambda}$ is written $\lambda := \lambda^d$ for short. Any two adjacent rows, say λ^l and λ^{l-1}, satisfy the interlacing condition:

$$\lambda^l_1 \geq \lambda^{l-1}_1 \geq \lambda^l_2 \geq \lambda^{l-1}_2 \geq \cdots \geq \lambda^{l-1}_{l-1} \geq \lambda^l_l\,. \tag{3.5}$$

Let $\mathrm{GT}(\lambda)$ denote the set of GT patterns with the top row being λ, and let $\mathrm{GT}(\lambda;\lambda^{d-1})$ denote the set of GT patterns with the top two rows being λ and λ^{d-1} respectively. In this chapter, the bold letters $\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu}$ denote GT patterns with the top row being partitions λ, μ, ν respectively.

In the same spirit as that a partition can be represented by a Young diagram, a GT pattern $\boldsymbol{\lambda}$ can be represented by a semistandard Young tableau of shape λ and alphabet $\{1,\dots,d\}$, e.g.,

$$\begin{pmatrix} 5 & & 3 & & 2 \\ & 4 & & 2 & \\ & & 3 & & \end{pmatrix} \iff \begin{array}{|c|c|c|c|c|} \hline 1 & 1 & 1 & 2 & 3 \\ \hline 2 & 2 & 3 \\ \cline{1-3} 3 & 3 \\ \cline{1-2} \end{array} \, .$$

Let $\boldsymbol{\lambda}+1^{K,L}$ denote the array of integers such that $\boldsymbol{\lambda}+1^{K,L}$ and $\boldsymbol{\lambda}$ have the same elements except that the K-th element in row L of $\boldsymbol{\lambda}+1^{K,L}$ is λ^L_K+1. The array $\boldsymbol{\lambda}-1^{K,L}$ is defined in the same fashion. Notice that $\boldsymbol{\lambda}\pm 1^{K,L}$ may not be a valid GT pattern.

Denote by $E^{i,j}$ the matrix with a one in the j-th entry in row i and with zero elsewhere. For $1\le l\le d-1$, a basis state $|\boldsymbol{\lambda}\rangle$ is acted by $\mathfrak{gl}(d)$ as

$$\begin{aligned} E^{l,l}|\boldsymbol{\lambda}\rangle &= (r^{\boldsymbol{\lambda}}_l - r^{\boldsymbol{\lambda}}_{l-1})|\boldsymbol{\lambda}\rangle \, , \\ E^{l,l+1}|\boldsymbol{\lambda}\rangle &= \sum_{k=1}^{l} a^{\boldsymbol{\lambda}}_{k,l}|\boldsymbol{\lambda}+1^{k,l}\rangle \, , \\ E^{l+1,l}|\boldsymbol{\lambda}\rangle &= \sum_{k=1}^{l} b^{\boldsymbol{\lambda}}_{k,l}|\boldsymbol{\lambda}-1^{k,l}\rangle \, , \end{aligned} \tag{3.6}$$

where

$$\begin{aligned} r^{\boldsymbol{\lambda}}_l &= \textstyle\sum_{k=1}^{l} \lambda^l_k \text{ and } r^{\boldsymbol{\lambda}}_0 = 0 \, , \\ a^{\boldsymbol{\lambda}}_{k,l} &= \left(- \frac{\prod_{i=1}^{l+1}(\hat{\lambda}^{l+1}_i - \hat{\lambda}^l_k)\prod_{i=1}^{l-1}(\hat{\lambda}^{l-1}_i - \hat{\lambda}^l_k - 1)}{\prod_{i=1,i\neq k}^{l}(\hat{\lambda}^l_i - \hat{\lambda}^l_k)(\hat{\lambda}^l_i - \hat{\lambda}^l_k - 1)} \right)^{1/2} , \\ b^{\boldsymbol{\lambda}}_{k,l} &= a^{\boldsymbol{\lambda}-1^{k,l}}_{k,l} \, , \end{aligned} \tag{3.7}$$

and $\hat{\lambda}^l_i := \lambda^l_i - i$ for $i\le l$, $\hat{\lambda}^l_i := 0$ for $i>l$, and zero factors in the products are skipped. The coefficient $a^{\boldsymbol{\lambda}}_{k,l}$ vanishes if $\boldsymbol{\lambda}+1^{k,l}$ is not a valid pattern, and $b^{\boldsymbol{\lambda}}_{k,l}$ vanishes if $\boldsymbol{\lambda}-1^{k,l}$ is not a valid pattern. It can be seen that the expressions (3.6) subsume (3.1) as a special case for $\mathfrak{su}(2)$ by observing that $|j,m\rangle = \left|\left(\begin{smallmatrix} 2j & & 0 \\ & j+m & \end{smallmatrix}\right)\right\rangle$.

The direct product of two irreps of $\mathfrak{su}(d)$, V_λ and V_μ, acted by $X\otimes \mathbb{1} + \mathbb{1}\otimes X$ for $X\in\mathfrak{su}(d)$, is still a representation, and is in general reducible. Consider the following decomposition

$$V_\lambda \otimes V_\mu = \bigoplus_\nu V_\nu^{\oplus N^\nu_{\lambda,\mu}} \, , \tag{3.8}$$

where $N^{\nu}_{\lambda,\mu}$ is the multiplicity of V_ν in the decomposition of $V_\lambda \otimes V_\mu$, or in short, the multiplicity of ν in $\lambda \otimes \mu$. The multiplicity of V_ν can be determined by the Littlewood-Richardson rule.

The basis state of the space on the right-hand side of Eq. (3.8) is written as

$$|\boldsymbol{\nu}, \gamma\rangle = \sum_{\boldsymbol{\lambda},\boldsymbol{\mu}} C^{\boldsymbol{\nu},\gamma}_{\boldsymbol{\lambda},\boldsymbol{\mu}} |\boldsymbol{\lambda}\rangle|\boldsymbol{\mu}\rangle \,, \tag{3.9}$$

where $\gamma = 1, \ldots, N^{\nu}_{\lambda,\mu}$.

Denote the weight $w^{\boldsymbol{\lambda}} := (w^{\boldsymbol{\lambda}}_1, \ldots, w^{\boldsymbol{\lambda}}_{d-1})$ where $w^{\boldsymbol{\lambda}}_l = r^{\boldsymbol{\lambda}}_l - \frac{1}{2}(r^{\boldsymbol{\lambda}}_{l+1} + r^{\boldsymbol{\lambda}}_{l-1})$. Applying $J_{z,l} := \frac{1}{2}(E^{l,l} - E^{l+1,l+1})$ to both sides of (3.9) yields $w^{\boldsymbol{\nu}}_l|\boldsymbol{\nu},\gamma\rangle = \sum_{\boldsymbol{\lambda},\boldsymbol{\mu}} C^{\boldsymbol{\nu},\gamma}_{\boldsymbol{\lambda},\boldsymbol{\mu}}(w^{\boldsymbol{\lambda}}_l + w^{\boldsymbol{\mu}}_l)|\boldsymbol{\lambda}\rangle|\boldsymbol{\mu}\rangle$. It follows that $C^{\boldsymbol{\nu},\gamma}_{\boldsymbol{\lambda},\boldsymbol{\mu}} \neq 0$ only if $w^{\boldsymbol{\nu}} = w^{\boldsymbol{\lambda}} + w^{\boldsymbol{\mu}}$.

When $|\boldsymbol{\nu}\rangle$ is a highest-weight state, applying $E^{l,l+1}$ to both sides yields its CGCs. An outer multiplicity larger than one leads to ambiguous choices of CGCs, which can be resolved by adopting some fixed rules. Applying $J_{-,l}$ to both sides of the expression for highest-weight states just obtained, one gets the CGCs for lower-weight states. See [AKHvD11] for an algorithmatic description of the Clebsch-Gordan transform.

3.2 Random near-invariant tensors

Throughout this section we denote by $V_{(s)}$ the irrep of $\mathfrak{su}(2)$ of dimension $2s+1$ for any nonnegative integer or half-integer s, and we consider the states in $V_{(s)}^{\otimes n}$ which are kept invariant under the action of $\mathfrak{su}(2)$. Then $V_{(s)}^{\otimes n}$, as a new representation of $\mathfrak{su}(2)$, has the following decomposition based on Clebsch-Gordan transform

$$V_{(s)}^{\otimes n} = \bigoplus_{j \leq ns} V_{(j)}^{\oplus N(n,j)} \,, \tag{3.10}$$

where $N(n,j)$ is the multiplicity of $V_{(j)}$ in decomposition of $V_{(s)}^{\otimes n}$.

The *coupled* basis of $V_{(s)}^{\otimes n}$, i.e. the basis of the right-hand side of the above equation, consists of $|j, m, \gamma\rangle$ with $j \in \{0, 1, \ldots, ns\}$ (or $j \in \{\frac{1}{2}, \frac{3}{2}, \ldots, ns\}$ when s is a half-integer), $\gamma \in \{1, 2, \ldots, N(n,j)\}$ and $m \in \{j, j-1, \ldots, -j\}$. A state $|\psi\rangle \in V_{(s)}^{\otimes n}$ is called an *invariant tensor* (or *invariant state*) if $X \cdot |\psi\rangle := (X \otimes \mathbb{1}^{\otimes(n-1)} + \mathbb{1} \otimes X \otimes \mathbb{1}^{\otimes(n-2)} + \cdots + \mathbb{1}^{\otimes(n-1)} \otimes X)|\psi\rangle = 0$ for any $X \in \mathfrak{su}(2)$, or equivalently, $U^{\otimes n} \cdot |\psi\rangle = |\psi\rangle$ for any $U \in \mathrm{SU}(2)$. Obviously, $|\psi\rangle$ is invariant if and only if it is a superposition of states $|j = 0, m = 0, \gamma\rangle$.

Denote the basis states of $V_{(s)}$ by $|s, m\rangle =: |s - m + 1\rangle$ for $m = s, s-1, \ldots, -s$. For $n = 2$, the invariant tensor $|\psi\rangle$ satisfies $(J_k \otimes \mathbb{1} + \mathbb{1} \otimes J_k)|\psi\rangle = 0$ for each $k =$

x, y, z. It is readily verified that the unique invariant tensor for the case $n = 2$ is $|\psi_s\rangle = (|1, 2s+1\rangle - |2, 2s\rangle + \cdots + (-1)^{2s}|2s+1, 1\rangle)/\sqrt{2s+1}$, e.g., $|\psi_1\rangle = (|13\rangle - |22\rangle + |31\rangle)/\sqrt{3}$ for $s = 1$. For the case $n = 3$ and $s = 1$, the unique invariant tensor is $(|123\rangle + |231\rangle + |312\rangle - |132\rangle - |213\rangle - |321\rangle)/\sqrt{6} \in \wedge^3\mathbb{C}^3$.

It is shown in [LHGZ17] that invariant perfect tensor exists for $n = 2$ or $n = 3$ but not for $n = 4$ and that a random four-partite invariant tensor is asymptotically perfect as the local dimension $(2s+1)$ goes to infinity. This result is then extended in [LHRZ18] which shows a random invariant tensor is asymptotically perfect for any n.

A random SU(2) invariant state, corresponding to $j = 0$, has asymptotic perfectness [LHGZ17, LHRZ18]. Does a random state corresponding to bounded j is also asymptotically perfect? Besides, when s is a half-integer and n is odd, the invariant tensor does not exist, since for this case j cannot be zero in the decomposition (3.10). Thus it is necessary to consider a generalized notion of invariant tensor for this setting. We introduce the notion of *near-invariant state*, which is kept invariant or slightly disturbed by the action of SU(2). To be specific, for the decomposition (3.10), we define in this section the *near-invariant state space* as $\mathcal{H}^{\text{inv}}_{s,n,j_0} := \text{span}\{|j, m, \gamma\rangle : j \leq j_0, -j \leq m \leq j, 1 \leq \gamma \leq N(n, j)\}$ where j_0 is finite and fixed. Any state in this space is called a near-invariant state.

Consider a bipartition of the n Hilbert spaces such that Alice holds p systems and Bob holds $q := n - p$ systems, where $p \leq q$. Thus

$$V_{(s)}^{\otimes n} = \left(\bigoplus_{j_1 \leq ps} V_{(j_1)}^{\oplus N(p,j_1)}\right) \bigotimes \left(\bigoplus_{j_2 \leq qs} V_{(j_2)}^{\oplus N(q,j_2)}\right). \tag{3.11}$$

The space $V_{(s)}^{\otimes n}$ on the left-hand side of (3.11) has a basis $\{|j, m, \gamma\rangle\}$, while the space on the right-hand side has a basis $\{|j_1, m_1, \alpha_{j_1}\rangle \otimes |j_2, m_2, \beta_{j_2}\rangle\}$ where $1 \leq \alpha_{j_1} \leq N(p, j_1)$ and $1 \leq \alpha_{j_2} \leq N(q, j_2)$. The near-invariant state space $\mathcal{H}^{\text{inv}}_{s,n,j_0}$ has a basis $\{|jmj_1j_2\alpha_{j_1}\beta_{j_2}\rangle : j \leq j_0, -j \leq m \leq j, j_1 \leq ps, j_2 \leq qs, 1 \leq \alpha_{j_1} \leq N(p, j_1), 1 \leq \beta_{j_2} \leq N(q, j_2)\}$, where

$$|jmj_1j_2\alpha_{j_1}\beta_{j_2}\rangle = \sum_{m_1,m_2:m_1+m_2=m} |j_1m_1\alpha_{j_1}\rangle_A |j_2m_2\beta_{j_2}\rangle_B C^{j_1,j_2,j}_{m_1,m_2,m}. \tag{3.12}$$

Let $\varphi = \varphi_{AB}$ be a uniformly random (normalized) state in $\mathcal{H}^{\text{inv}}_{s,n,j_0}$, and φ_A be its reduced state on system A. The 2-Rényi entropy of φ_A is

$$\begin{aligned} H_2(\varphi_A) &= -\ln \operatorname{tr}(\varphi_A^2) = -\ln \operatorname{tr}(W_{AA'}(\varphi_A \otimes \varphi_{A'})) \\ &= -\ln \operatorname{tr}(W_{AA'}(\varphi_{AB} \otimes \varphi_{A'B'}))\,, \end{aligned} \tag{3.13}$$

where A', B' denote isomorphic systems of A, B respectively, the second equality uses the swap trick, $W_{AA'}$ is the swap operator, and φ_{AB} and $\varphi_{A'B'}$ are the same state.

Note that the 2-Rényi entanglement entropy of a maximally entangled state on AB is $H_{\max} := \ln(2s+1)^p$. We are concerned with how close the random invariant state φ_{AB} is to a maximally entangled state, measured by the ratio

$$\eta(\varphi_{AB}) := \frac{H_2(\varphi_A)}{H_{\max}}\,. \tag{3.14}$$

As it is difficult to calculate the average of $H_2(\varphi_A)$ directly, we will estimate the average of $\mathrm{e}^{-H_2(\varphi_A)}$, and then show that the random variable $H_2(\varphi_A)$ is highly concentrated around $-\ln \mathbb{E}_\varphi \mathrm{e}^{-H_2(\varphi_A)}$ with small fluctuation for large dimension.

It follows from Eq. (3.13) that

$$\mathbb{E}_\varphi \mathrm{e}^{-H_2(\varphi_A)} = \mathbb{E}_\varphi \operatorname{tr}(\varphi_A^2) = \operatorname{tr}(W_{AA'}\mathbb{E}_\varphi(\varphi_{AB} \otimes \varphi_{A'B'}))\,. \tag{3.15}$$

Using the Werner twirling formula (1.10) in Lemma 2, where the unitary U in the integration is over $\mathrm{U}(\mathcal{H}^{\mathrm{inv}}_{s,n,j_0})$, we have

$$\mathbb{E}_\varphi(\varphi_{AB} \otimes \varphi_{A'B'}) = \frac{\mathbb{1}_{\mathrm{inv}}^{\otimes 2} + W_{\mathrm{inv}}}{d_{\mathrm{inv}}^2 + d_{\mathrm{inv}}}\,, \tag{3.16}$$

where $\mathbb{1}_{\mathrm{inv}}$ is the projector onto the near-invariant space $\mathcal{H}^{\mathrm{inv}}_{s,n,j_0}$, $d_{\mathrm{inv}} = \operatorname{tr}(\mathbb{1}_{\mathrm{inv}})$ is its dimension, and W_{inv} is the swap operator on $(\mathcal{H}^{\mathrm{inv}}_{s,n,j_0})^{\otimes 2}$.

Theorem 12. *Asymptotically, a random near-invariant state is perfect (i.e. maximally entangled with respect to any bipartite cut). To be specific, for any p, q such that $q \geq p \geq 2$ and $q \geq 3$, and for any fixed $\delta > 0$, a random near-invariant state $\varphi_{AB} \in \mathcal{H}^{\mathrm{inv}}_{s,n,j_0} \subset V_{(s)}^{\otimes p} \otimes V_{(s)}^{\otimes q}$ satisfies $\Pr(|\eta(\varphi_{AB}) - 1| \geq \delta) \to 0$ as $s \to \infty$.*

Indeed, it can be seen from the proof of the theorem above that δ can be chosen arbitrarily small as long as $\delta \gtrsim s^{-1/2-\varepsilon}$ for any $\varepsilon > 0$. The theorem above follows from the two propositions as below.

Proposition 13. *Let $q \geq p \geq 2$. For random near-invariant state φ_{AB}, as $s \to \infty$,*

$$s^{-p} \lesssim \mathbb{E}_{\varphi} \mathrm{e}^{-H_2(\varphi_A)} \lesssim s^{-p} \ln s\,.$$

Proposition 14. *Let $q \geq p \geq 2$ and $q \geq 3$. For random near-invariant state φ_{AB}, as $s \to \infty$,*

$$\frac{\mathbb{E}_{\varphi} \operatorname{tr}^2 \varphi_A^2}{(\mathbb{E}_{\varphi} \operatorname{tr} \varphi_A^2)^2} - 1 \lesssim s^{-1}\,.$$

Proof of Theorem 12. Denoting $K := -\ln \mathbb{E}\mathrm{e}^{-H_2(\varphi_A)}$, by Proposition 13, there exists constants c, C such that $p \ln s - \ln\ln s - \ln C \leq K \leq p \ln s - \ln c$. Thus $\frac{K}{H_{\max}} \to 1$ as $s \to \infty$. So for any $\delta > 0$ and s large enough,

$$\left|\frac{K}{H_{\max}} - 1\right| \leq \frac{\delta}{2}\,. \tag{3.17}$$

We have

$$\begin{aligned}
\Pr\left(\left|\frac{H_2(\varphi_A)}{H_{\max}} - 1\right| \geq \delta\right) &= \Pr\left(\left|\frac{H_2(\varphi_A) - K}{H_{\max}} + \frac{K}{H_{\max}} - 1\right| \geq \delta\right) \\
&\leq \Pr(|H_2(\varphi_A) - K| \geq H_{\max}\delta/2) \\
&\leq \Pr\left(\left|\frac{\operatorname{tr}\varphi_A^2}{\mathbb{E}\operatorname{tr}\varphi_A^2} - 1\right| \geq \varepsilon\right) \\
&\leq \frac{1}{\varepsilon^2}\left(\frac{\mathbb{E}\operatorname{tr}^2\varphi_A^2}{(\mathbb{E}\operatorname{tr}\varphi_A^2)^2} - 1\right) \\
&\lesssim \frac{1}{s\varepsilon^2} \to 0\,,
\end{aligned}$$

where $\varepsilon := 1 - \mathrm{e}^{-H_{\max}\delta/2}$, the second line uses Eq. (3.17), the third line uses the fact that if $|\ln x| \geq t$ then $|x - 1| \geq 1 - \mathrm{e}^{-t}$, the fourth line uses Markov inequality, and the last line is due to Proposition 14. □

Before proving Propositions 13 and 14, we give several lemmas.

Lemma 15. *Let $N(n,k)$ denote the multiplicity of irrep $V_{(k)}$ in $V_{(s)}^{\otimes n}$ for $n \geq 2$. Then, as $s \to \infty$,*

$$N(n,k) \lesssim s^{n-2} \text{ for each } k\,, \tag{3.18}$$

and

$$N(n,k) \simeq s^{n-2} \text{ for each } k \in [s, (n-1)s]\,. \tag{3.19}$$

Proof. Consider $V_{(s)}^{\otimes n} = \bigoplus_k V_{(k)}^{\oplus N(n,k)}$ for $s \in \frac{1}{2}\mathbb{N}$. When $k > ns$ or $k < 0$, set $N(n,k) = 0$. Note that $V_{(j)} \otimes V_{(s)}$ contains $V_{(k)}$ iff $|s - j| \leq k \leq s + j$. Thus $N(n,k) =$

$N(n-1,k-s)+\cdots+N(n-1,k+s)$. It follows that by induction for $0\leq k\leq ns$, $N(n,k)\leq(2s+1)\max_{k'}N(n-1,k')\leq\cdots\leq(2s+1)^{n-2}\max_{k''}N(2,k'')=(2s+1)^{n-2}$. Conversely, for $s\leq k\leq(n-1)s$, $N(n,k)\geq s\min_{s\leq k'\leq(n-2)s}N(n-1,k')\geq\cdots\geq s^{n-2}N(2,s)=s^{n-2}$. □

Lemma 16. *Let j,Δ,m be finite and fixed, and m_1 may depends on j_1. As $j_1\to\infty$,*

$$|C^{j_1,j_1+\Delta,j}_{m_1,j-m_1,j}|\simeq j_1^{-j-\frac{1}{2}}(j_1-m_1)^{\frac{1}{2}(j+\Delta)}(j_1+m_1)^{\frac{1}{2}(j-\Delta)}\,, \tag{3.20}$$

and

$$|C^{j_1,j_1+\Delta,j}_{m_1,m-m_1,m}|\lesssim j_1^{-j-\frac{1}{2}}\sum_{k=0}^{j-\Delta}(j_1-m_1)^{j-k+\frac{1}{2}(m-\Delta)}(j_1+m_1)^{k+\frac{1}{2}(\Delta-m)}\,. \tag{3.21}$$

Proof. Eq. (3.20) simply follows from that

$$\begin{aligned}(C^{j_1,j_1+\Delta,j}_{m_1,j-m_1,j})^2&=\frac{(2j+1)!(2j_1+\Delta-j)!}{(2j_1+\Delta+j+1)!(j+\Delta)!(j-\Delta)!}\frac{(j_1-m_1+j+\Delta)!(j_1+m_1)!}{(j_1+m_1-j+\Delta)!(j_1-m_1)!}\\&\simeq j_1^{-2j-1}(j_1-m_1)^{j+\Delta}(j_1+m_1)^{j-\Delta}\,.\end{aligned}$$

Due to Racah's formula for the CGCs, $|C^{j_1,j_1+\Delta,j}_{m_1,m-m_1,m}|\leq\sqrt{C_1}\sqrt{C_2}C_3$, where

$$\begin{aligned}C_1&=\frac{(2j+1)(j-\Delta)!(j+\Delta)!(2j_1+\Delta-j)!}{(2j_1+\Delta+j+1)!}\simeq j_1^{-2j-1}\,,\\C_2&=(j+m)!(j-m)!(j_1+m_1)!(j_1-m_1)!(j_1-m_1+\Delta+m)!(j_1+m_1+\Delta-m)!\,,\\C_3&=\sum\nolimits_{k=0}^{j-\Delta}f_k^{-1}\end{aligned}$$

with $f_k=(j_1-m_1-j+\Delta+k)!(j_1+m_1-k)!(j-\Delta-k)!(j+m-k)!k!(\Delta-m+k)!$. The summation in C_3 is extended over those k such that the argument of every factorial is nonnegative (thus $0\leq k\leq j-\Delta$). Note that $C_3\leq\sum_k((j_1-m_1-j+\Delta+k)!(j_1+m_1-k)!)^{-1}$. It follows that

$$\begin{aligned}\sqrt{C_2}C_3&\lesssim\sum_k\Big(\frac{(j_1+m_1)!(j_1-m_1)!(j_1-m_1+\Delta+m)!(j_1+m_1+\Delta-m)!}{(j_1-m_1-j+\Delta+k)!^2(j_1+m_1-k)!^2}\Big)^{\frac{1}{2}}\\&\simeq\sum_k(j_1-m_1)^{j-k+\frac{1}{2}(m-\Delta)}(j_1+m_1)^{k+\frac{1}{2}(\Delta-m)}\,,\end{aligned}$$

completing the proof of Eq. (3.21). □

Lemma 17. *For any nonnegative integers t and r, $\sum_{i=1}^{n-1} i^t(n-i)^r \simeq n^{t+r+1}$ as $n \to \infty$.*

Proof. It is known $\sum_{k=1}^{n} k^t \sim \frac{n^{t+1}}{t+1}$ as $n \to \infty$ by Faulhaber's formula. The case $rt = 0$ is readily verified, and we now prove the case $r, t > 0$. We have

$$\begin{aligned}\sum_{i=1}^{n} i^t(n-i)^r &= \sum_{k=0}^{r}(-1)^k\binom{r}{k}n^{r-k}\sum_{i=1}^{n} i^{t+k} \\ &\sim \sum_{k=0}^{r}(-1)^k\binom{r}{k}n^{r-k}\frac{n^{t+k+1}}{t+k+1} \\ &=: F(t+1,r)n^{t+r+1}\,,\end{aligned}$$

where the second line holds if the coefficient $F(t+1,r)$ is nonzero. To complete the proof of this lemma, it suffices to show $F(t+1,r) \neq 0$. We now prove that

$$F(t,r) := \sum_{k=0}^{r}(-1)^k\binom{r}{k}\frac{1}{k+t} = \frac{1}{t\binom{r+t}{t}} \tag{3.22}$$

for positive integers t, r.

For any positive integer m, denote

$$f(m) := \sum_{j=0}^{r}(-1)^j\binom{r+m}{j+m}\,. \tag{3.23}$$

By calculating $f(m) + f(m+1)$, we have $f(m) = \binom{r+m-1}{m-1}$.

Now we calculate $F(1,r)$ and $F(2,r)$. It holds that $F(1,r) = \sum_{k=0}^{r}(-1)^k\binom{r}{k}\frac{1}{k+1} = \frac{1}{r+1}\sum_{k=0}^{r}(-1)^k\binom{r+1}{k+1} = \frac{1}{r+1}$ and that $F(2,r) = \frac{1}{r+1}\sum_{k=0}^{r}(-1)^k\binom{r+1}{k+1} - \frac{1}{(r+1)(r+2)}\sum_{k=0}^{r}(-1)^k\binom{r+2}{k+2} = \frac{1}{(r+1)(r+2)}$, where we have used Eq. (3.23).

Since $F(t,r+1) = \sum_{k=0}^{r+1}(-1)^k\binom{r+1}{k}\frac{1}{k+t}$, we have

$$\begin{aligned}F(t,r) - F(t,r+1) &= \sum_{k=1}^{r}(-1)^{k+1}\binom{r}{k-1}\frac{1}{k+t} - (-1)^{r+1}\frac{1}{r+1+t} \\ &= \sum_{k=0}^{r-1}(-1)^k\binom{r}{k}\frac{1}{k+t+1} + (-1)^r\frac{1}{r+1+t} \\ &= F(t+1,r)\end{aligned}$$

Consequently, by induction on t, using the recurrence relation $F(t+1,r) = F(t,r) - F(t,r+1)$ and the expressions of $F(1,r)$ and $F(2,r)$, we achieve the desired result

(3.22). □

Proof of Proposition 13. Since $\{|jmj_1j_2\alpha_{j_1}\beta_{j_2}\rangle\}$ is a basis of the near-invariant space,

$$\mathbb{1}_{\text{inv}}^{\otimes 2} + W_{\text{inv}} = \sum \left(|jmj_1j_2\alpha_{j_1}\beta_{j_2}\rangle|j'm'j_1'j_2'\alpha'_{j_1'}\beta'_{j_2'}\rangle + |j'm'j_1'j_2'\alpha'_{j_1'}\beta'_{j_2'}\rangle|jmj_1j_2\alpha_{j_1}\beta_{j_2}\rangle\right) \left(\langle jmj_1j_2\alpha_{j_1}\beta_{j_2}|\langle j'm'j_1'j_2'\alpha'_{j_1'}\beta'_{j_2'}|\right), \tag{3.24}$$

the sum over $j, m, j_1, j_2, \alpha_{j_1}, \beta_{j_2}, j', m', j_1', j_2', \alpha'_{j_1'}, \beta'_{j_2'}$.

Using Eq. (3.12), we have

$$\begin{aligned}
&\operatorname{tr}\left(W_{AA'}(|jmj_1j_2\alpha_{j_1}\beta_{j_2}\rangle|j'm'j_1'j_2'\alpha'_{j_1'}\beta'_{j_2'}\rangle)(\langle jmj_1j_2\alpha_{j_1}\beta_{j_2}|\langle j'm'j_1'j_2'\alpha'_{j_1'}\beta'_{j_2'}|)\right) \\
&= \sum_{\substack{m_1+m_2=m \\ m_1'+m_2'=m'}} |j_1m_1\alpha_{j_1}\rangle|j_2m_2\beta_{j_2}\rangle C^{j_1j_2j}_{m_1,m_2,m}|j_1'm_1'\alpha'_{j_1'}\rangle|j_2'm_2'\beta'_{j_2'}\rangle C^{j_1'j_2'j'}_{m_1',m_2',m'} \\
&\quad \sum_{\substack{\hat{m}_1+\hat{m}_2=m \\ \hat{m}_1'+\hat{m}_2'=m'}} \langle j_1\hat{m}_1\alpha_{j_1}|\langle j_2\hat{m}_2\beta_{j_2}|C^{j_1j_2j}_{\hat{m}_1,\hat{m}_2,m}\langle j_1'\hat{m}_1'\alpha'_{j_1'}|\langle j_2'\hat{m}_2'\beta'_{j_2'}|C^{j_1'j_2'j'}_{\hat{m}_1',\hat{m}_2',m'} \\
&= \sum_{m_1} \delta_{j_1,j_1'}\delta_{\alpha_{j_1},\alpha'_{j_1}}\left(C^{j_1,j_2,j}_{m_1,m-m_1,m}\right)^2\left(C^{j_1,j_2',j'}_{m_1,m'-m_1,m'}\right)^2,
\end{aligned} \tag{3.25}$$

and similarly,

$$\begin{aligned}
&\operatorname{tr}\left(W_{AA'}(|j'm'j_1'j_2'\alpha'_{j_1'}\beta'_{j_2'}\rangle|jmj_1j_2\alpha_{j_1}\beta_{j_2}\rangle)(\langle jmj_1j_2\alpha_{j_1}\beta_{j_2}|\langle j'm'j_1'j_2'\alpha'_{j_1'}\beta'_{j_2'}|)\right) \\
&= \sum_{m_2} \delta_{j_2,j_2'}\delta_{\beta_{j_2},\beta'_{j_2}}\left(C^{j_1,j_2,j}_{m-m_2,m_2,m}\right)^2\left(C^{j_1',j_2,j'}_{m'-m_2,m_2,m'}\right)^2.
\end{aligned} \tag{3.26}$$

It follows that

$$\begin{aligned}
\operatorname{tr}(W_{AA'}\mathbb{1}_{\text{inv}}^{\otimes 2}) &= \sum_{j,m,j_1,j_2,\alpha_{j_1},\beta_{j_2},j',m',j_2',\beta'_{j_2'}} \sum_{m_1} \left(C^{j_1,j_2,j}_{m_1,m-m_1,m}\right)^2\left(C^{j_1,j_2',j'}_{m_1,m'-m_1,m'}\right)^2, \\
\operatorname{tr}(W_{AA'}W_{\text{inv}}) &= \sum_{j,m,j_1,j_2,\alpha_{j_1},\beta_{j_2},j',m',j_1',\alpha'_{j_1'}} \sum_{m_2} \left(C^{j_1,j_2,j}_{m-m_2,m_2,m}\right)^2\left(C^{j_1',j_2,j'}_{m'-m_2,m_2,m'}\right)^2.
\end{aligned} \tag{3.27}$$

Denoting $j_2 := j_1 + \Delta$ and $j_2' := j_1 + \Delta'$, noticing $|\Delta| \le j$ and $|\Delta'| \le j'$, we have

$$\begin{aligned}
\operatorname{tr}(W_{AA'}\mathbb{1}_{\text{inv}}^{\otimes 2}) &\ge \sum_{j,j_1,\Delta,\alpha_{j_1},\beta_{j_1+\Delta},j',\Delta',\beta'_{j_1+\Delta'}} \sum_{|m_1|<j_1} \left(C^{j_1,j_1+\Delta,j}_{m_1,j-m_1,j}\right)^2 \left(C^{j_1,j_1+\Delta',j'}_{m_1,j'-m_1,j'}\right)^2 \\
&\simeq \sum_{j,j_1,\Delta,\alpha_{j_1},\beta_{j_1+\Delta},j',\Delta',\beta'_{j_1+\Delta'}} \sum_{|m_1|<j_1} j_1^{-2j-2j'-2}(j_1-m_1)^{j+j'+\Delta+\Delta'}(j_1+m_1)^{j+j'-\Delta-\Delta'} \\
&\simeq \sum_{j,j_1,\Delta,\alpha_{j_1},\beta_{j_1+\Delta},j',\Delta',\beta'_{j_1+\Delta'}} j_1^{-1} \\
&\gtrsim \sum_{j_1} N(p,j_1)N(q,j_1)^2 j_1^{-1} \\
&\simeq \sum_{j_1} s^{p-2}(s^{q-2})^2 j_1^{-1} \\
&\sim s^{p+2q-6}\ln(p-1)\,,
\end{aligned} \tag{3.28}$$

and thus $\operatorname{tr}(W_{AA'}\mathbb{1}_{\text{inv}}^{\otimes 2}) \gtrsim s^{p+2q-6}$. In the above derivation, the first line uses the special cases $m = j$ and $m' = j'$, the second line uses Eq. (3.20), the third line uses Lemma 17, the fourth line, where the sum is over $j_1 \in [s,(p-1)s]$, uses the special cases $\Delta = \Delta' = 0$, the fifth line uses Lemma 15, and the last line uses the asymptotics of the harmonic series, i.e. $\sum_{n=1}^{k}\frac{1}{n} \sim \ln k$ as $k \to \infty$.

In the other direction,

$$\begin{aligned}
\operatorname{tr}(W_{AA'}\mathbb{1}_{\text{inv}}^{\otimes 2}) &\le \sum_{j,j_1,\Delta,\alpha_{j_1},\beta_{j_1+\Delta},j',\Delta',\beta'_{j_1+\Delta'}} \sum_{m_1} j_1^{-2j-1}\sum_{k=0}^{j-\Delta}(j_1-m_1)^{2j-2k-\Delta+m}(j_1+m_1)^{2k+\Delta-m} \\
&\qquad \cdot j_1^{-2j'-1}\sum_{k'=0}^{j'-\Delta'}(j_1-m_1)^{2j'-2k'-\Delta'+m'}(j_1+m_1)^{2k'+\Delta'-m'} \\
&\simeq \sum_{j,j_1,\Delta,\alpha_{j_1},\beta_{j_1+\Delta},j',\Delta',\beta'_{j_1+\Delta'}} j_1^{-1} \\
&\lesssim \max_{j_2}\sum_{j_1} N(p,j_1)N(q,j_2)^2 j_1^{-1} \\
&\lesssim s^{p+2q-6}\ln(ps)\,,
\end{aligned}$$

where the first line uses Eq. (3.21) and that $(\sum_{k=1}^{n} a_k)^2 \le n\sum_{k=1}^{n} a_k^2$, the '$\simeq$' is due to Lemma 17, and the last line uses Lemma 15.

Consequently,

$$s^{p+2q-6} \lesssim \operatorname{tr}(W_{AA'}\mathbb{1}_{\text{inv}}^{\otimes 2}) \lesssim s^{p+2q-6}\ln s\,.$$

In the same way, we have

$$s^{2p+q-6} \lesssim \operatorname{tr}(W_{AA'}W_{\mathrm{inv}}) \lesssim s^{2p+q-6}\ln s\,.$$

Since $\{|jmj_1j_2\alpha_{j_1}\beta_{j_2}\rangle\}$ is a basis of the near-invariant space, its dimension is

$$d_{\mathrm{inv}} = \sum_{j,m,\Delta}\sum_{j_1} N(p,j_1)N(q,j_1+\Delta) \simeq \sum_{j_1} s^{p-2}s^{q-2} \simeq s^{n-3}\,. \tag{3.29}$$

Thus,

$$\frac{s^{p+2q-6}+s^{2p+q-6}}{(s^{n-3})^2} \lesssim \mathbb{E}_\varphi \operatorname{tr}(\varphi_A^2) \lesssim \frac{(s^{p+2q-6}+s^{2p+q-6})\ln s}{(s^{n-3})^2}\,, \tag{3.30}$$

completing the proof by noticing $p \le q$. $\square$

Proof of Proposition 14. By Lemma 2,

$$\mathbb{E}(\operatorname{tr}\varphi_A^2)^2 = \frac{1}{(d_{\mathrm{inv}})^{\uparrow 4}} \operatorname{tr}\left((W^A_{(12)}W^A_{(34)})\sum\nolimits_{\pi\in S_4} W_\pi\right) \tag{3.31}$$

and

$$(\mathbb{E}\operatorname{tr}\varphi_A^2)^2 = \frac{1}{d_{\mathrm{inv}}^2(d_{\mathrm{inv}}+1)^2} \operatorname{tr}\left((W^A_{(12)}W^A_{(34)})\sum\nolimits_{\pi\in S_2\times S_2} W_\pi\right), \tag{3.32}$$

where $W^A_{(12)}$ swaps the first and second system A, and $W^A_{(34)}$ swaps the third and fourth system A. Thus

$$\frac{\mathbb{E}(\operatorname{tr}\varphi_A^2)^2}{(\mathbb{E}\operatorname{tr}\varphi_A^2)^2} - 1 = \frac{-4d_{\mathrm{inv}}-6}{(d_{\mathrm{inv}}+2)(d_{\mathrm{inv}}+3)} + \frac{(d_{\mathrm{inv}}-1)!}{(d_{\mathrm{inv}}+3)!}\frac{1}{(\mathbb{E}\operatorname{tr}\varphi_A^2)^2}\sum\nolimits_{\pi\in S_4\backslash S_2\times S_2} \operatorname{tr}\left((W^A_{(12)}W^A_{(34)})W_\pi\right). \tag{3.33}$$

Together with Eq. (3.29) and Eq. (3.32), it follows that

$$\frac{\mathbb{E}\operatorname{tr}^2\varphi_A^2}{(\mathbb{E}\operatorname{tr}\varphi_A^2)^2} - 1 \lesssim s^{-2p-4q+12}\sum\nolimits_{\pi\in S_4\backslash S_2\times S_2} \operatorname{tr}\left((W^A_{(12)}W^A_{(34)})W_\pi\right). \tag{3.34}$$

In the following, for $\pi \in S_4\backslash S_2\times S_2$, denote $t_i := \pi^{-1}(i)$ for $i\in\{1,2,3,4\}$. We have

$$\begin{aligned}W_\pi = \sum \big|j^{t_1}m^{t_1}j_1^{t_1}j_2^{t_1}\alpha^{t_1}_{j_1^{t_1}}\beta^{t_1}_{j_2^{t_1}}\big\rangle\big|j^{t_2}m^{t_2}j_1^{t_2}j_2^{t_2}\alpha^{t_2}_{j_1^{t_2}}\beta^{t_2}_{j_2^{t_2}}\big\rangle\big|j^{t_3}m^{t_3}j_1^{t_3}j_2^{t_3}\alpha^{t_3}_{j_1^{t_3}}\beta^{t_3}_{j_2^{t_3}}\big\rangle\big|j^{t_4}m^{t_4}j_1^{t_4}j_2^{t_4}\alpha^{t_4}_{j_1^{t_4}}\beta^{t_4}_{j_2^{t_4}}\big\rangle \\ \langle j^1m^1j_1^1j_2^1\alpha^1_{j_1^1}\beta^1_{j_2^1}|\langle j^2m^2j_1^2j_2^2\alpha^2_{j_1^2}\beta^2_{j_2^2}|\langle j^3m^3j_1^3j_2^3\alpha^3_{j_1^3}\beta^3_{j_2^3}|\langle j^4m^4j_1^4j_2^4\alpha^4_{j_1^4}\beta^4_{j_2^4}|\,,\end{aligned}$$

where the sum is over $j^k, m^k, j_1^k, j_2^k, \alpha_{j_1^k}^k, \beta_{j_2^k}^k$ for $k \in \{1,2,3,4\}$, and the superscripts indicate different systems. In the following, k in each summation also ranges over $\{1,2,3,4\}$.

Using

$$
\begin{aligned}
\left|j^k m^k j_1^k j_2^k \alpha_{j_1^k}^k \beta_{j_2^k}^k\right\rangle &= \sum_{m_1^k + m_2^k = m^k} \left|j_1^k m_1^k \alpha_{j_1^k}^k\right\rangle \left|j_2^k m_2^k \beta_{j_2^k}^k\right\rangle C_{m_1^k m_2^k m^k}^{j_1^k j_2^k j^k}, \\
\left\langle j^k m^k j_1^k j_2^k \alpha_{j_1^k}^k \beta_{j_2^k}^k\right| &= \sum_{\hat{m}_1^k + \hat{m}_2^k = m^k} \left\langle j_1^k \hat{m}_1^k \alpha_{j_1^k}^k\right| \left\langle j_2^k \hat{m}_2^k \beta_{j_2^k}^k\right| C_{\hat{m}_1^k \hat{m}_2^k m^k}^{j_1^k j_2^k j^k},
\end{aligned}
\tag{3.35}
$$

we have

$$
\operatorname{tr}\left((W_{(12)}^A W_{(34)}^A) W_\pi\right) = \sum_{k, j^k, m^k, j_1^k, j_2^k} T_{\text{cgc}} T_{\alpha\beta},
$$

where

$$
T_{\alpha\beta} = \sum_{k, \alpha_{j_1^k}^k, \beta_{j_2^k}^k} 1,
$$

and

$$
\begin{aligned}
T_{\text{cgc}} = \sum_{k, m_1^k + m_2^k = m^k} & C_{m_1^1, m_2^1, m^1}^{j_1^1, j_2^1, j^1} C_{m_1^2, m_2^2, m^2}^{j_1^2, j_2^2, j^2} C_{m_1^3, m_2^3, m^3}^{j_1^3, j_2^3, j^3} C_{m_1^4, m_2^4, m^4}^{j_1^4, j_2^4, j^4} \\
& C_{m_1^{t_2}, m_2^{t_1}, m^1}^{j_1^{t_2}, j_2^1, j^1} C_{m_1^{t_1}, m_2^{t_2}, m^2}^{j_1^{t_1}, j_2^2, j^2} C_{m_1^{t_4}, m_2^{t_3}, m^3}^{j_1^{t_4}, j_2^3, j^3} C_{m_1^{t_3}, m_2^{t_4}, m^4}^{j_1^{t_3}, j_2^4, j^4}.
\end{aligned}
$$

In the following, we consider $\pi \in S_4 \backslash S_2 \times S_2$, for which the pair $(\#\pi, \#((12)(34)\pi))$ is in the set $\{(3,1),(2,2),(1,3),(1,1)\}$. Using Lemma 17 and Eq. (3.21), we have $T_{\text{cgc}} \lesssim 1$ for each such π.

When $\#\pi = 1$ or 2,

$$
\operatorname{tr}(W_{(12)}^A W_{(34)}^A W_\pi) \lesssim \sum_{j_1, \Delta} N(p, j_1)^{\#((12)(34)\pi)} N(q, j_1 + \Delta)^{\#\pi}.
$$

Thus, $\operatorname{tr}(W_{(12)}^A W_{(34)}^A W_\pi) \lesssim s^{3p+q-7}$ for $\#\pi = 1$, and $\operatorname{tr}(W_{(12)}^A W_{(34)}^A W_\pi) \lesssim s^{2p+2q-7}$ for $\#\pi = 2$.

Specially, when $\#\pi = 3$, we have $T_{\text{cgc}} \simeq (j_1^1)^{-2}$, and thus $\sum_{k,j_1^k} T_{\text{cgc}} \lesssim 1$, hence

$$\operatorname{tr}(W^A_{(12)} W^A_{(34)} W_\pi) \lesssim \max_{j_1} N(p,j_1)^{\#((12)(34)\pi)} N(q, j_1+\Delta)^{\#\pi} \lesssim s^{p+3q-8}.$$

Consequently, using Eq. (3.34), it follows that

$$\frac{\mathbb{E}\operatorname{tr}^2 \varphi_A^2}{(\mathbb{E}\operatorname{tr}\varphi_A^2)^2} - 1 \lesssim \max\{s^{p-3q+5}, s^{-2q+5}, s^{-p-q+4}\},$$

of which the right-hand side vanishes as $s \to \infty$ provided $q \geq 3$ and $q \geq p \geq 2$. □

We studied the generalized case of the near-invariant states, subsuming the results for invariant states. The idea of our proof of Theorem 12 is similar to that used in [HNQ+16] and [LHRZ18]. Our proof, however, does not involve the uncoupled basis, simplifies the calculation, and gives explicit asymptotic estimate for some variables.

3.3 Symmetric invariant tensors of higher degree

Although the basic idea of representation of $\mathfrak{su}(d)$ for $d \geq 3$ is similar to that of $\mathfrak{su}(2)$, substantive difference exists between the two cases as the former case involves complex structure of roots and weights [Hal15]. In this section we study the random $\mathfrak{su}(d)$-invariant states to investigate whether the entanglement property still holds for this new symmetry.

Consider the Schur-Weyl decomposition $(\mathbb{C}^d)^{\otimes n} = \bigoplus_{\lambda \in \mathrm{Par}(n,d)} V_\lambda \otimes \mathcal{K}_\lambda$, where V_λ and $\mathcal{K}_\lambda$ are the Weyl module of $\mathfrak{su}(d)$ and Specht module S_n respectively. Let $\mathbb{C}^d$ be the defining irrep of $\mathfrak{su}(d)$, then when n is divisible by d, the invariant state space in $(\mathbb{C}^d)^{\otimes n}$ always exists uniquely, which is $V_\lambda \otimes \mathcal{K}_\lambda$ for λ being $(k, k, \ldots, k)$. This section studies the invariant state space in $V_{(s)}^{\otimes n}$ where $V_{(s)}$ is the symmetric subspace of $(\mathbb{C}^d)^{\otimes s}$. Still, the invariant state is vanished by the action of $\mathfrak{su}(d)$. As $V_{(s)}$ is the symmetric subspace, this invariant state is called symmetric invariant state (tensor) in this section.

We say two partitions $\lambda := (\lambda_1, \ldots, \lambda_d)$ and $\mu := (\mu_1, \ldots, \mu_d)$ are dual to each other if $\lambda_k + \mu_{d+1-k} = \lambda_1 + \mu_d$ for any $k \in [d]$. Similarly, we say two GT patterns $\boldsymbol{\lambda} := (\lambda_k^l)_{l\in[d],k\in[l]}$ and $\boldsymbol{\mu} := (\mu_k^l)_{l\in[d],k\in[l]}$ are dual to each other if $\lambda_k^l + \mu_{l+1-k}^l = \lambda_1^l + \mu_l^l$ for any $l \in [d]$ and $k \in [l]$.

Lemma 18. *Let $\lambda := (\lambda_1, \ldots, \lambda_d)$ and $\mu := (\mu_1, \ldots, \mu_d)$ be two partitions. As representation of $\mathfrak{su}(d)$, $V_\lambda \otimes V_\mu$ contains invariant state space if and only if λ and μ are* *dual to each other. Further, for $\boldsymbol{\lambda} \in \mathrm{GT}(\lambda)$ and $\boldsymbol{\mu} \in \mathrm{GT}(\mu)$, $C^0_{\boldsymbol{\lambda},\boldsymbol{\mu}} \neq 0$ if and only if $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ are dual to each other.*

Proof. $V_\lambda \otimes V_\mu$ contains invariant state space iff $V_\lambda \otimes V_\mu$ contains V_ν for ν being $(\lambda_1 + \mu_d, \ldots, \lambda_1 + \mu_d)$. Due to the Littlewood-Richardson rule, we first add μ_d boxes filled with integer 1 in the first row, and then add μ_d boxes filled with integer 2 and $\lambda_2 - \lambda_1$ boxes filled with integer 1. Continuing so, we need to add $\mu_d + \lambda_1 - \lambda_{d+1-k}$ boxes filled with k for each k. Thus $\lambda_k + \mu_{d+1-k} = \lambda_1 + \mu_d$ for each k, that is, λ is dual to μ.

Given a partition λ^d, V_{λ^d} is an irrep of $\mathrm{SU}(d)$. By the branching rule, as irrep of $\mathfrak{su}(d-1)$, V_{λ^d} decomposes as $V_{\lambda^d} = \bigoplus_{\lambda^{d-1}} V_{(\lambda^d;\lambda^{d-1})}$, where the direct sum is over all partitions λ^{d-1} interlacing λ^d, and $V_{(\lambda^d;\lambda^{d-1})} := \mathrm{span}\{|\boldsymbol{\lambda}\rangle : \boldsymbol{\lambda} \in \mathrm{GT}(\lambda^d;\lambda^{d-1})\}$. The irrep V_{μ^d} decomposes in the same fashion. It follows that

$$V_{\lambda^d} \otimes V_{\mu^d} = \bigoplus_{\lambda^{d-1},\mu^{d-1}} V_{(\lambda^d;\lambda^{d-1})} \otimes V_{(\mu^d;\mu^{d-1})}\,. \tag{3.36}$$

Due to Littlewood-Richardson rule, in the orthogonal direct sum the term $V_{(\lambda^d;\lambda^{d-1})} \otimes V_{(\mu^d;\mu^{d-1})}$ contains $V_{(0)}$ iff λ^{d-1} and μ^{d-1} are dual to each other. Continuing so, we have $C^0_{\boldsymbol{\lambda},\boldsymbol{\mu}} \neq 0$ iff $\boldsymbol{\lambda}$ is dual to $\boldsymbol{\mu}$. □

Lemma 19. *For each dual pair of partitions λ and μ, and for each dual pair of GT patterns $\boldsymbol{\lambda} \in \mathrm{GT}(\lambda)$ and $\boldsymbol{\mu} \in \mathrm{GT}(\mu)$, $|C^0_{\boldsymbol{\lambda},\boldsymbol{\mu}}| = (\dim V_\lambda)^{-1/2}$.*

Proof. The irrep $V_{(0)}$ is contained in decomposition of $V_\lambda \otimes V_\mu$, hence

$$|0\rangle = \sum_{\boldsymbol{\lambda}} C^0_{\boldsymbol{\lambda},\boldsymbol{\mu}} |\boldsymbol{\lambda}\rangle|\boldsymbol{\mu}\rangle \tag{3.37}$$

for $|0\rangle \in V_{(0)}$.

For $l \in [d-1]$, applying $E^{l,l+1}$ to Eq. (3.37), we have

$$\begin{aligned} 0 &= \sum_{k\in[l],\boldsymbol{\lambda}} C^0_{\boldsymbol{\lambda},\boldsymbol{\mu}} \left(a^{\boldsymbol{\lambda}}_{k,l} |\boldsymbol{\lambda} + 1^{k,l}\rangle|\boldsymbol{\mu}\rangle + a^{\boldsymbol{\mu}}_{k,l} |\boldsymbol{\lambda}\rangle|\boldsymbol{\mu} + 1^{k,l}\rangle\right) \\ &= \sum_{k\in[l],\boldsymbol{\lambda}} C^0_{\boldsymbol{\lambda},\boldsymbol{\mu}} \left(a^{\boldsymbol{\lambda}}_{l+1-k,l} |\boldsymbol{\lambda} + 1^{l+1-k,l}\rangle|\boldsymbol{\mu}\rangle + a^{\boldsymbol{\mu}}_{k,l} |\boldsymbol{\lambda}\rangle|\boldsymbol{\mu} + 1^{k,l}\rangle\right). \end{aligned}$$

The coefficient of $|\boldsymbol{\lambda}\rangle|\boldsymbol{\mu} + 1^{k,l}\rangle$ is

$$C^0_{\boldsymbol{\lambda}-1^{l+1-k,l},\boldsymbol{\mu}+1^{k,l}} a^{\boldsymbol{\lambda}-1^{l+1-k,l}}_{l+1-k,l} + C^0_{\boldsymbol{\lambda},\boldsymbol{\mu}} a^{\boldsymbol{\mu}}_{k,l} = 0\,.$$

It can be verified from Eq. (3.7) that

$$a_{l+1-k,l}^{\boldsymbol{\lambda}-1^{l+1-k,l}} = a_{k,l}^{\boldsymbol{\mu}} .$$

It follows that $C^0_{\boldsymbol{\lambda}-1^{l+1-k,l},\boldsymbol{\mu}+1^{k,l}} + C^0_{\boldsymbol{\lambda},\boldsymbol{\mu}} = 0$ for each pair $(\boldsymbol{\lambda},\boldsymbol{\mu})$ of dual patterns. Thus for any dual pairs $(\boldsymbol{\lambda},\boldsymbol{\mu})$ and $(\boldsymbol{\lambda'},\boldsymbol{\mu'})$, either $C^0_{\boldsymbol{\lambda},\boldsymbol{\mu}} = C^0_{\boldsymbol{\lambda'},\boldsymbol{\mu'}}$ or $C^0_{\boldsymbol{\lambda},\boldsymbol{\mu}} = -C^0_{\boldsymbol{\lambda'},\boldsymbol{\mu'}}$. Using the normalization condition of (3.37) and the fact that $\mathrm{GT}(\lambda)$ has cardinality $\dim(V_\lambda)$, the result follows. $\square$

We now restrict our attention to the dual pairs of partitions and patterns in $V_{(s)}^{\otimes p} \otimes V_{(s)}^{\otimes q}$. In the rest of this section we assume that $\frac{p+q}{d}s$ is an integer so that $V_{(s)}^{\otimes p} \otimes V_{(s)}^{\otimes q}$ contains invariant states. Given $\lambda = (\lambda_1, \ldots, \lambda_d) \in \mathrm{Par}(ps, d)$, the dual of λ in $\mathrm{Par}(qs, d)$ is unique and denoted by λ_*, and moreover, for given $\boldsymbol{\lambda} \in \mathrm{GT}(\lambda)$, the dual of $\boldsymbol{\lambda}$ in $\mathrm{GT}(\mu)$ with $\mu \in \mathrm{Par}(qs, d)$ is also unique and denoted by $\boldsymbol{\lambda}_*$. For a state φ_{AB}, denote

$$\eta(\varphi_{AB}) := \frac{H_2(\varphi_A)}{H_{\max}} , \tag{3.38}$$

where $H_{\max} := \ln(\dim V_{(s)}^{\otimes p}) = \ln \binom{s+d-1}{d-1}^p \simeq \ln s^{(d-1)p}$.

For any integer k and $\lambda \in \mathrm{Par}(ks)$, denote by $N(\lambda)$ the multiplicity of V_λ in decomposition of $V_{(s)}^{\otimes k}$, or in short, the multiplicity of λ in $(s)^{\otimes k}$. Consider the decomposition

$$V_{(s)}^{\otimes n} = V_{(s)}^{\otimes p} \otimes V_{(s)}^{\otimes q} = \Big(\bigoplus\nolimits_{\lambda \vdash ps} V_\lambda^{\oplus N(\lambda)}\Big) \bigotimes \Big(\bigoplus\nolimits_{\mu \vdash qs} V_\mu^{\oplus N(\mu)}\Big) . \tag{3.39}$$

For any dual pair of partitions $\lambda \in \mathrm{Par}(ps, d)$ and $\mu \in \mathrm{Par}(qs, d)$ and for $1 \le \alpha_\lambda \le N(\lambda)$ and $1 \le \beta_\mu \le N(\mu)$,

$$|\lambda, \alpha_\lambda, \beta_\mu\rangle = \sum C^0_{\boldsymbol{\lambda},\boldsymbol{\mu}} |\boldsymbol{\lambda}, \alpha_\lambda\rangle_A |\boldsymbol{\mu}, \beta_\mu\rangle_B , \tag{3.40}$$

where the sum is over dual pairs of patterns $\boldsymbol{\lambda} \in \mathrm{GT}(\lambda)$ and $\boldsymbol{\mu} \in \mathrm{GT}(\mu)$, is an invariant state.

Using the same approach used in last section, in what follows we first estimate the expectation

$$\mathbb{E}\mathrm{e}^{-H_2(\varphi_A)} = \mathbb{E}_\varphi \operatorname{tr} \varphi_A^2 = \frac{\operatorname{tr}(W_{AA'}(\mathbb{1}_{\mathrm{inv}}^{\otimes 2} + W_{\mathrm{inv}}))}{d_{\mathrm{inv}}^2 + d_{\mathrm{inv}}} , \tag{3.41}$$

where

$$d_{\mathrm{inv}} = \sum_{\lambda \vdash ps, \lambda_* \vdash qs} N(\lambda) N(\lambda_*) \tag{3.42}$$

is the dimension of invariant state space, and then give a bound on its variance for using Markov inequality.

The states of form (3.40) constitute an orthonormal basis of the invariant state space, so

$$\mathbb{1}_{\mathrm{inv}}^{\otimes 2} = \sum_{\substack{\lambda,\alpha_\lambda,\beta_\mu \\ \lambda',\alpha'_{\lambda'},\beta'_{\mu'}}} \left(|\lambda,\alpha_\lambda,\beta_\mu\rangle|\lambda',\alpha'_{\lambda'},\beta'_{\mu'}\rangle\right)\left(\langle\lambda,\alpha_\lambda,\beta_\mu|\langle\lambda',\alpha'_{\lambda'},\beta'_{\mu'}|\right), \tag{3.43}$$

where $\lambda \vdash ps$ is dual to $\mu \vdash qs$. Inserting (3.40) into (3.43), we have

$$\mathrm{tr}(W_{AA'}\mathbb{1}_{\mathrm{inv}}^{\otimes 2}) = \sum_{\lambda,\alpha_\lambda,\beta_{\lambda_*},\beta'_{\lambda_*}} \sum_{\boldsymbol{\lambda}} (C^0_{\boldsymbol{\lambda},\boldsymbol{\lambda}_*})^4 = \sum_{\lambda} N(\lambda)N(\lambda_*)^2 \frac{1}{\dim V_\lambda}, \tag{3.44}$$

since $(C^0_{\boldsymbol{\lambda},\boldsymbol{\lambda}_*})^2 = (\dim V_\lambda)^{-1}$.

Similarly,

$$\mathrm{tr}(W_{AA'}W_{\mathrm{inv}}) = \sum_{\lambda} N(\lambda)^2 N(\lambda_*) \frac{1}{\dim V_\lambda}. \tag{3.45}$$

The multiplicities of irreps in decomposition of $V_{(s)}^{\otimes p} \otimes V_{(s)}^{\otimes q}$ with $q \geq p$ are given in the following two lemmas.

Lemma 20. *Let $\mu = \mu^p = (\mu_1, \mu_2, \ldots, \mu_{\min\{p,d\}})$ be any strictly decreasing sequence of positive numbers that sum to p. Then, as $s \to \infty$,*

(i) $N(\mu s) \simeq s^{(p-1)(p-2)/2}$ when $p \leq d$,

(ii) $N(\mu s) \simeq s^{(d-1)(d-2)/2+(d-1)(p-d)}$ when $p \geq d$,

(iii) $N(\lambda^p) \lesssim N(\mu s)$ for any $\lambda^p \in \mathrm{Par}(ps, \min\{p,d\})$.

Proof. Without loss of generality we here consider the closest partition in $\mathrm{Par}(ps, \min\{p,d\})$ to μs if μs itself is not a valid partition, since the distance between the two vectors vanishes as $s \to \infty$.

For (i), let $\mu^{p-1} = (\mu_1^{p-1}, \ldots, \mu_{p-1}^{p-1})$ be a strictly decreasing sequence of sum $p-1$. By Littlewood-Richardson rule, the decomposition of $\mu^{p-1}s \otimes (s)$ contains μs iff μ^{p-1} interlaces μ, i.e. $\mu_1 \geq \mu_1^{p-1} \geq \mu_2 \geq \mu_2^{p-1} \geq \cdots \geq \mu_{p-1}^{p-1} \geq \mu_p$. For given μ, such a μ^{p-1} exists iff $\mu_1 > 1 > \mu_p$. Continuing so, there exists a strictly decreasing sequence $\mu^{p-2} = (\mu_1^{p-2}, \ldots, \mu_{p-2}^{p-2})$ such that $\mu^{p-2}s \otimes (s)$ contains $\mu^{p-1}s$, iff $\mu_1 + \mu_2 > 2 > \mu_{p-1} + \mu_p$. Consequently, μs has nonzero multiplicity in $(s)^{\otimes p}$ iff $\sum_{i=1}^k \mu_i > k$ for each k. Indeed, the condition $\sum_{i=1}^k \mu_i > k$ holds for each k, since μ is a strictly decreasing sequence of sum p. Now we have obtained $\mu^k \in \mathbb{R}^k$ of sum k for each $k \in \{2, \ldots, p\}$, such that $\mu^k s$ is contained in $\mu^{k-1}s \otimes (s)$ for each k. Since $N(\mu^2 s) = 1$, using induction it suffices to show $N(\mu^k s) \simeq s^{k-2}N(\mu^{k-1}s)$ for each k.

Let $\theta = (\theta_1, \ldots, \theta_k)$ be such that $\theta_1 \in (0,1)$, $\theta_i \in \left(0, 1 - \sum_{j=1}^{i-1} \theta_j\right)$ for $2 \le i \le k-1$, and $\theta_k = -\sum_{j=1}^{k-1} \theta_j$, and let $\theta' = (\theta'_1, \ldots, \theta'_k)$ be defined similarly. For $\varepsilon > 0$ small and fixed (independent of s), denote $\mu^k_\theta = \mu^k + \theta\varepsilon$. Set ε so small that μ^{k-1}_θ interlaces $\mu^k_{\theta'}$ for each θ, θ'. Since $\mu^{k-1}_\theta s$ has $\simeq s^{k-2}$ choices as θ varies for given μ^{k-1}, and $N(\mu^k_{\theta'} s) \ge \sum_\theta N(\mu^{k-1}_\theta s)$, we have $N(\mu^k s) \gtrsim s^{k-2} N(\mu^{k-1} s)$. Due to the upcoming proof of (iii) (for $p \le d$), it holds that $N(\mu^k s) \lesssim s^{k-2} N(\mu^{k-1} s)$, completing the proof of (i).

We now prove (iii) for the case $p \le d$. Since $N(\lambda^2) = 1$ for $\lambda^2 \in \mathrm{Par}(2s, 2)$, it suffices to show that $N(\lambda^k) \lesssim s^{k-2} \max_{\lambda^{k-1}} N(\lambda^{k-1})$ for any $2 \le k \le p$ and $\lambda^k \in \mathrm{Par}(ks, k)$, where the maximum is over $\lambda^{k-1} \in \mathrm{Par}((k-1)s, k-1)$. λ^k is contained in decomposition of $\lambda^{k-1} \otimes (s)$ for some λ^{k-1} iff λ^{k-1} interlaces λ^k. Since $\lambda^k_i - \lambda^k_{i-1} \lesssim s$ for each i, and λ^{k-1} has sum $(k-1)s$, the number of λ^{k-1}'s that interlace λ^k is $\lesssim s^{k-2}$. Thus $N(\lambda^k) \lesssim s^{k-2} \max_{\lambda^{k-1}} N(\lambda^{k-1})$.

The case $p > d$ in (ii) and (iii) is proved in the same way. For $p \ge d+1$, $N(\mu s) \simeq s^{d-1} N(\mu^{p-1} s) \simeq s^{(d-1)(p-d)} N(\mu^d s) \simeq s^{(d-1)(d-2)/2 + (d-1)(p-d)}$. □

Lemma 21. *Let $p \le q$, and denote $b := \frac{p+q}{d}$ and $t := p+q-d$. Let $\mu = \mu^p = (\mu_1, \mu_2, \ldots, \mu_{\min\{p,d\}})$, where $\mu_1 < b$, be any strictly decreasing sequence of positive numbers that sum to p. If $t \le 0$, $N(\lambda_*) = 0$ for any $\lambda \in \mathrm{Par}(ps, \min\{p,d\})$. If $t > 0$,*

(i) $N((\mu s)_*) \simeq s^{(d-1)(d-2)/2 + (d-1)(q-d)}$ when $p \ge d$,

(ii) $N((\mu s)_*) \simeq s^{(d-1)(d-2)/2 + (d-1)(t-d) + (p+d-1)(d-p)/2}$ when $p \le d$ and $t \ge d$,

(iii) $N((\mu s)_*) \simeq s^{(p+t-1)(q-d)/2 + (t-1)(d-t) + (t-1)(t-2)/2}$ when $p \le d$, $t \le d$ and $q \ge d$,

(iv) $N((\mu s)_*) \simeq s^{(p-1)(d-p) + (t-1)(t-2)/2}$ when $p \le d$, $t \le d$ and $q \le d$,

(v) $N((\lambda^p)_*) \lesssim N((\mu s)_*)$ for any $\lambda^p \in \mathrm{Par}(ps, \min\{p,d\})$.

Proof. Notice that $\lambda_* \in \mathrm{Par}(qs)$ for $\lambda \in \mathrm{Par}(ps)$. Since the proof idea is similar to that of Lemma 20, a brief calculation is given as follows.

For (i), when $p \ge d$, for any $\lambda = (\mu_1, \ldots, \mu_d)s \in \mathrm{Par}(ps, d)$, $\lambda_* = (b - \mu_d, \ldots, b - \mu_1)s$ is a strictly decreasing sequence of sum q. Since $q \ge d$, and $\mu_1 < b$, by Lemma 20, $N(\lambda_*) \simeq s^{(d-1)(d-2)/2 + (d-1)(q-d)}$.

When $p < d$, for any $\lambda = \mu^p s = (\mu_1, \ldots, \mu_p)s$ with $\mu_1 < b$, we have $\lambda_* = \nu^q s$, where $\nu^q := (b, \ldots, b, b - \mu_p, \ldots, b - \mu_1)$ has sum q. Let ν^q be interlaced by ν^{q-1}, ν^{q-1} be interlaced by ν^{q-2}, and so on until ν^2 be interlaced by $\nu^1 = (1)$, where ν^k has sum k for each k. Three subcases are calculated as follows.

For (ii), when $t \ge d$, $\nu^q s$ has multiplicity $N(\nu^q s) \simeq s^p N(\nu^{q-1} s) \simeq s^{p + (p+1) + \cdots + (d-1)} N(\nu^t s) = s^{(p+d-1)(d-p)/2} N(\nu^t s)$, $\nu^t s$ has multiplicity $N(\nu^t s) \simeq s^{(d-1)(p+q-2d)} N(\nu^d s)$, and $\nu^d s$ has multiplicity $N(\nu^d s) \simeq s^{(d-1)(d-2)/2}$. Thus $N((\mu s)_*) = N(\nu^q s)$ is obtained.

For (iii), when $t < d$ and $d \leq q$, $\nu^q s$ has multiplicity $N(\nu^q s) \simeq s^p N(\nu^{q-1} s) \simeq s^{p+(p+1)+\cdots+(p+q-d-1)} N(\nu^d s) = s^{(2p+q-d-1)(q-d)/2} N(\nu^d s)$, $\nu^d s$ has multiplicity $N(\nu^d s) \simeq s^{(p+q-d-1)(2d-p-q)} N(\nu^t s)$, and $\nu^t s$ has multiplicity $N(\nu^t s) \simeq s^{(p+q-d-1)(p+q-d-2)/2}$. Thus $N(\nu^q s)$ is obtained.

For (iv), when $t < d$ and $d \geq q$, $\nu^q s$ has multiplicity $N(\nu^q s) \simeq s^{(p-1)(d-p)} N(\nu^t s)$, and $\nu^t s$ has multiplicity $N(\nu^t s) \simeq s^{(t-1)(t-2)/2}$. □

For large n and constant k, we have $|\mathrm{Par}(n,k)| \simeq n^{k-1}$ since $\frac{1}{k!}|\mathrm{Type}(n,k)| \leq \mathrm{Par}(n,k) \leq |\mathrm{Type}(n,k)|$ and $|\mathrm{Type}(n,k)| = \binom{n+k-1}{k-1}$. By Weyl dimension formula, $\dim V_\lambda \simeq s^{h(h-1)/2}$ for $h := \min\{p,d\}$.

We now estimate the value of η.

Theorem 22. *For $p \geq d \geq 2$, a random symmetric invariant state is asymptotically maximally entangled with respect to any bipartite cut. To be specific, for any $p \leq q$, and for any fixed $\delta > 0$, a random invariant state $\varphi_{AB} \in \mathcal{H}^{\mathrm{inv}}_{s,n} \subset V_{(s)}^{\otimes p} \otimes V_{(s)}^{\otimes q}$ satisfies $\Pr(|\eta(\varphi_{AB}) - 1| \geq \delta) \to 0$ as $s \to \infty$.*

Proof. Using the proof idea of Theorem 12 we only need to estimate $\mathbb{E}\,\mathrm{tr}\,\psi_A^2$ and $\frac{\mathbb{E}\,\mathrm{tr}^2\,\psi_A^2}{(\mathbb{E}\,\mathrm{tr}\,\psi_A^2)^2}$.

For $p \geq d$, λ has s^{d-1} choices. By Weyl dimension formula, $\dim V_\lambda \simeq s^{d(d-1)/2}$. Let $\mu = \mu^p = (\mu_1, \mu_2, \ldots, \mu_h)$, where $\mu_1 < b$, be any strictly decreasing sequence of positive numbers that sum to p, and let $\lambda = \mu s$. Then by Lemmas 20 and 21,

$$
\begin{aligned}
d_{\mathrm{inv}} &= \sum_{\lambda \vdash ps, \lambda_* \vdash qs} N(\lambda) N(\lambda_*) \\
&\simeq s^{d-1} s^{(d-1)(d-2)/2+(d-1)(p-d)} s^{(d-1)(d-2)/2+(d-1)(q-d)} \\
&= s^{(d-1)(t-1)}\,,
\end{aligned}
\tag{3.46}
$$

and

$$
\begin{aligned}
\mathrm{tr}(W_{AA'} \mathbb{1}^{\otimes 2}_{\mathrm{inv}}) + \mathrm{tr}(W_{AA'} W_{\mathrm{inv}}) &\simeq s^{d-1} s^{(d-1)(d-2)/2+(d-1)(p-d)} \left(s^{(d-1)(d-2)/2+(d-1)(q-d)}\right)^2 s^{-d(d-1)/2} \\
&= s^{(d-1)(p+2q-2d-2)}\,,
\end{aligned}
$$

from which it follows that $\mathbb{E}\,\mathrm{tr}\,\varphi_A^2 \simeq s^{-p(d-1)}$ by (3.41).

We now give a bound on the variance of $\mathrm{tr}\,\varphi_A^2$. Using Eq. (3.46), and Eq. (3.33) which still holds for the case $d \geq 3$, we have

$$
\frac{\mathbb{E}(\mathrm{tr}\,\varphi_A^2)^2}{(\mathbb{E}\,\mathrm{tr}\,\varphi_A^2)^2} - 1 \lesssim s^{-2(d-1)(p+2q-2d-2)} \sum\nolimits_{\pi \in S_4 \setminus S_2 \times S_2} \mathrm{tr}\left((W^A_{(12)} W^A_{(34)}) W_\pi\right). \tag{3.47}
$$

$$W_\pi = \sum_{k,\lambda^k,\alpha^k_{\lambda^k},\beta^k_{\mu^k}} |\lambda^{t_1}\alpha^{t_1}_{\lambda^{t_1}}\beta^{t_1}_{\mu^{t_1}}\rangle|\lambda^{t_2}\alpha^{t_2}_{\lambda^{t_2}}\beta^{t_2}_{\mu^{t_2}}\rangle|\lambda^{t_3}\alpha^{t_3}_{\lambda^{t_3}}\beta^{t_3}_{\mu^{t_3}}\rangle|\lambda^{t_4}\alpha^{t_4}_{\lambda^{t_4}}\beta^{t_4}_{\mu^{t_4}}\rangle$$
$$\langle\lambda^1\alpha^1_{\lambda^1}\beta^1_{\mu^1}|\langle\lambda^2\alpha^2_{\lambda^2}\beta^2_{\mu^2}|\langle\lambda^3\alpha^3_{\lambda^3}\beta^3_{\mu^3}|\langle\lambda^4\alpha^4_{\lambda^4}\beta^4_{\mu^4}|$$

$$= \sum_{k,\lambda^k,\alpha^k_{\lambda^k},\beta^k_{\mu^k}} \sum_{k,\boldsymbol{\lambda}^k} |\boldsymbol{\lambda}^{t_1}\alpha^{t_1}_{\lambda^{t_1}}\rangle|\boldsymbol{\mu}^{t_1}\beta^{t_1}_{\mu^{t_1}}\rangle|\boldsymbol{\lambda}^{t_2}\alpha^{t_2}_{\lambda^{t_2}}\rangle|\boldsymbol{\mu}^{t_2}\beta^{t_2}_{\mu^{t_2}}\rangle|\boldsymbol{\lambda}^{t_3}\alpha^{t_3}_{\lambda^{t_3}}\rangle|\boldsymbol{\mu}^{t_3}\beta^{t_3}_{\mu^{t_3}}\rangle|\boldsymbol{\lambda}^{t_4}\alpha^{t_4}_{\lambda^{t_4}}\rangle|\boldsymbol{\mu}^{t_4}\beta^{t_4}_{\mu^{t_4}}\rangle$$
$$\langle\boldsymbol{\lambda}^{t_2}\alpha^1_{\lambda^1}|\langle\boldsymbol{\mu}^{t_1}\beta^1_{\mu^1}|\langle\boldsymbol{\lambda}^{t_1}\alpha^2_{\lambda^2}|\langle\boldsymbol{\mu}^{t_2}\beta^2_{\mu^2}|\langle\boldsymbol{\lambda}^{t_4}\alpha^3_{\lambda^3}|\langle\boldsymbol{\mu}^{t_3}\beta^3_{\mu^3}|\langle\boldsymbol{\lambda}^{t_3}\alpha^4_{\lambda^4}|\langle\boldsymbol{\mu}^{t_4}\beta^4_{\mu^4}|$$
$$\cdot C^0_{\boldsymbol{\lambda}^{t_1},\boldsymbol{\mu}^{t_1}}C^0_{\boldsymbol{\lambda}^{t_2},\boldsymbol{\mu}^{t_2}}C^0_{\boldsymbol{\lambda}^{t_3},\boldsymbol{\mu}^{t_3}}C^0_{\boldsymbol{\lambda}^{t_4},\boldsymbol{\mu}^{t_4}} \cdot C^0_{\boldsymbol{\lambda}^{t_2},\boldsymbol{\mu}^{t_1}}C^0_{\boldsymbol{\lambda}^{t_1},\boldsymbol{\mu}^{t_2}}C^0_{\boldsymbol{\lambda}^{t_4},\boldsymbol{\mu}^{t_3}}C^0_{\boldsymbol{\lambda}^{t_3},\boldsymbol{\mu}^{t_4}} \tag{3.48}$$

where $\boldsymbol{\lambda}^k, \hat{\boldsymbol{\lambda}}^k \in \mathrm{GT}(\lambda^k)$, and $\boldsymbol{\mu}^k, \hat{\boldsymbol{\mu}}^k \in \mathrm{GT}(\mu^k)$.

In order to calculate $\mathrm{tr}(W^A_{(12)}W^A_{(34)}W_\pi)$, assume w.l.o.g. that $\boldsymbol{\lambda}^{t_1} = \hat{\boldsymbol{\lambda}}^2$, $\boldsymbol{\lambda}^{t_2} = \hat{\boldsymbol{\lambda}}^1$, $\boldsymbol{\lambda}^{t_3} = \hat{\boldsymbol{\lambda}}^4$, $\boldsymbol{\lambda}^{t_4} = \hat{\boldsymbol{\lambda}}^3$, $\boldsymbol{\mu}^{t_1} = \hat{\boldsymbol{\mu}}^1$, $\boldsymbol{\mu}^{t_2} = \hat{\boldsymbol{\mu}}^2$, $\boldsymbol{\mu}^{t_3} = \hat{\boldsymbol{\mu}}^3$, $\boldsymbol{\mu}^{t_4} = \hat{\boldsymbol{\mu}}^4$, $\alpha^1_{\lambda^1} = \alpha^{t_2}_{\lambda^{t_2}}$, $\alpha^2_{\lambda^2} = \alpha^{t_1}_{\lambda^{t_1}}$, $\alpha^3_{\lambda^3} = \alpha^{t_4}_{\lambda^{t_4}}$, $\alpha^4_{\lambda^4} = \alpha^{t_3}_{\lambda^{t_3}}$, $\beta^1_{\mu^1} = \beta^{t_1}_{\mu^{t_1}}$, $\beta^2_{\mu^2} = \beta^{t_2}_{\mu^{t_2}}$, $\beta^3_{\mu^3} = \beta^{t_3}_{\mu^{t_3}}$, $\beta^4_{\mu^4} = \beta^{t_4}_{\mu^{t_4}}$.

Thus

$$\mathrm{tr}(W^A_{(12)}W^A_{(34)}W_\pi) = \sum_{\lambda^1,\lambda^2,\lambda^3,\lambda^4} T_{\mathrm{cgc}}T_{\alpha\beta}\,, \tag{3.49}$$

where

$$T_{\alpha\beta} = \sum_{k,\alpha^k_{\lambda^k},\beta^k_{\mu^k}} 1 \simeq N(\lambda)^{\#(W\pi)}N(\lambda_*)^{\#\pi}\,,$$

and

$$T_{\mathrm{cgc}} = \sum_{\boldsymbol{\lambda}^1,\boldsymbol{\lambda}^2,\boldsymbol{\lambda}^3,\boldsymbol{\lambda}^4} C^0_{\boldsymbol{\lambda}^{t_1},\boldsymbol{\mu}^{t_1}}C^0_{\boldsymbol{\lambda}^{t_2},\boldsymbol{\mu}^{t_2}}C^0_{\boldsymbol{\lambda}^{t_3},\boldsymbol{\mu}^{t_3}}C^0_{\boldsymbol{\lambda}^{t_4},\boldsymbol{\mu}^{t_4}}$$
$$C^0_{\boldsymbol{\lambda}^{t_2},\boldsymbol{\mu}^{t_1}}C^0_{\boldsymbol{\lambda}^{t_1},\boldsymbol{\mu}^{t_2}}C^0_{\boldsymbol{\lambda}^{t_4},\boldsymbol{\mu}^{t_3}}C^0_{\boldsymbol{\lambda}^{t_3},\boldsymbol{\mu}^{t_4}} \tag{3.50}$$

Since the summand in right-hand side of Eq. (3.50) is nonzero only if $\boldsymbol{\lambda}^{t_1} = \boldsymbol{\lambda}^{t_2}$ and $\boldsymbol{\lambda}^{t_3} = \boldsymbol{\lambda}^{t_4}$. It follows that $T_{\mathrm{cgc}} \neq 0$ only if $\lambda^{t_1} = \lambda^{t_2}$ and $\lambda^{t_3} = \lambda^{t_4}$, in which case,

$$T_{\mathrm{cgc}} = \sum\nolimits_{\boldsymbol{\lambda}^{t_1},\boldsymbol{\lambda}^{t_3}} (\dim V_{\lambda^{t_1}})^{-2}(\dim V_{\lambda^{t_3}})^{-2} = (\dim V_{\lambda^{t_1}})^{-1}(\dim V_{\lambda^{t_3}})^{-1} \simeq s^{-d(d-1)}\,. \tag{3.51}$$

Thus

$$\operatorname{tr}(W^A_{(12)}W^A_{(34)}W_\pi) \simeq \sum_{\lambda^{t_1},\lambda^{t_3}} N(\lambda)^{\#(W\pi)} N(\lambda_*)^{\#\pi} s^{-d(d-1)} . \tag{3.52}$$

When $\#(W\pi) = 1$ and $\#\pi = 3$,

$$\begin{aligned}\operatorname{tr}(W^A_{(12)}W^A_{(34)}W_\pi) &\simeq s^{2(d-1)} s^{(d-1)(d-2)/2+(d-1)(p-d)} \big(s^{(d-1)(d-2)/2+(d-1)(q-d)}\big)^3 s^{-d(d-1)} \\ &= s^{(d-1)(p+3q-3d-2)} .\end{aligned}$$

Using Eq. (3.47), we have

$$\frac{\mathbb{E}(\operatorname{tr}\varphi_A^2)^2}{(\mathbb{E}\operatorname{tr}\varphi_A^2)^2} - 1 \lesssim s^{(d-1)(-p-q+d+2)} , \tag{3.53}$$

which vanishes as $s \to \infty$ if $p+q > d+2$. Under the condition $q \geq p \geq d$, $p+q > d+2$ iff $q \geq 3$.

When $(\#\pi, \#((12)(34)\pi))$ equals $(2,2)$, $(1,3)$, or $(1,1)$, the estimate similar to Eq. (3.53) can be obtained. Combining these estimates, we have

$$\begin{aligned}\frac{\mathbb{E}(\operatorname{tr}\varphi_A^2)^2}{(\mathbb{E}\operatorname{tr}\varphi_A^2)^2} - 1 &\lesssim \max\left\{s^{(d-1)(-p-q+d+2)}, s^{(d-1)(-2q+d+2)}, s^{(d-1)(p-3q+d+2)}, s^{(d-1)(-p-3q+2d+4)}\right\} \\ &= s^{(d-1)(-p-q+d+2)} .\end{aligned}$$

□

Using the same calculation method, when $p < d$, we have that $\mathbb{E}\operatorname{tr}\varphi_A^2 \gtrsim s^{-p(p-1)}$, and thus $\frac{-\ln \mathbb{E}\operatorname{tr}\varphi_A^2}{H_{\max}} \leq \frac{p-1}{d-1} < 1$ for large s. The calculation in this chapter is not valid for the case where the numbers p, q of systems Alice and Bob holds are relatively small compared with the local dimension d, in that $\operatorname{tr}\varphi_A^2$ is so small that $-\ln \mathbb{E}\operatorname{tr}\varphi_A^2$ is far from $-\mathbb{E}\ln\operatorname{tr}\varphi_A^2$, as the second-order derivative of $-\ln x$ is large for small positive x.

Chapter 4

Certification of quantum states and unitaries

4.1 Introduction and previous work

In the emerging quantum information technology, a fundamental task in building reliable quantum information processing devices is to obtain parameters of an unknown quantum state or device. The process is called *tomography* if all parameters are required to be known. In many scenarios, however, we are concerned only with whether the unknown state or operation satisfies specific property. For example, to assess the quality of a quantum chip after production, one needs to check whether the circuit is close to a given unitary transformation, and it is unnecessary to get all parameters about this chip. This process is called *certification*, which usually saves samples and storage space compared with the quantum tomography. See [MdW16] for a survey on quantum certification.

The abstract setting for certification can be described as follows. Given a known set $\mathcal{P}$ and an unknown element x, a tester (or an algorithm) $\mathcal{T}$ either accepts (i.e. reports $x \in \mathcal{P}$) or rejects (i.e. reports $x \notin \mathcal{P}_\varepsilon$) with some probability after measuring x, where $\mathcal{P}_\varepsilon := \{y : \mathrm{dist}(y, \mathcal{P}) \leq \varepsilon\}$ with dist denoting some metric. The tester $\mathcal{T}$ is eligible if the following conditions hold:

(1) (Completeness) If $x \in \mathcal{P}$, $\mathcal{T}$ accepts with probability at least $2/3$;

(2) (Soundness) If $x \notin \mathcal{P}_\varepsilon$, $\mathcal{T}$ accepts with probability at most $1/3$.

The numbers $2/3, 1/3$ have no special meaning and can be replaced by any constants $c > 1/2, s < 1/2$ respectively due to the confidence amplification by using repeating the test. If the tester accepts with certainty when $x \in \mathcal{P}$, we say the tester has perfect completeness.

We first recall that the trace distance between two quantum states is $\mathrm{D}(\rho, \sigma) := \frac{1}{2}\|\rho - \sigma\|_1$, and their fidelity is $\mathrm{F}(\rho, \sigma) := \|\sqrt{\rho}\sqrt{\sigma}\|_1$. Given a known state $|\varphi\rangle$ and an unknown state $|\psi\rangle$, the aim is to decide whether $\psi = \varphi$, i.e., $|\psi\rangle = \mathrm{e}^{\mathrm{i}\theta}|\varphi\rangle$ for some real θ, or $\mathrm{D}(\psi, \varphi) \geq \varepsilon$. In this simple task of certification, $\mathcal{P}$ is $\{\mathrm{e}^{\mathrm{i}\theta}|\varphi\rangle : \theta \in \mathbb{R}\}$ and x is $|\psi\rangle$. It turns out $O(\varepsilon^{-2})$ copies of given states are sufficient for this task. The test is simply to perform the POVM $\{|\varphi\rangle\langle\varphi|, \mathbb{1} - |\varphi\rangle\langle\varphi|\}$ independently n times, and accept if and only if every outcome is the first one. If $\mathrm{D}(\psi, \varphi) = \sqrt{1 - \langle\psi, \varphi\rangle} \geq \varepsilon$, then $\langle\psi, \varphi\rangle^n \leq (1 - \varepsilon^2)^n \leq \frac{1}{3}$ and one can take $n = \frac{\ln \frac{1}{3}}{\ln(1-\varepsilon^2)} = O(\varepsilon^{-2})$ for small ε.

In order to test whether a pair of unknown states ρ and σ on $\mathcal{H}$ are equal, the swap test [BCWdW01] is usually used. The swap test is simply a two-outcome measurement $\{P, Q\}$, where P and Q are projectors onto the symmetric subspace and the antisymmetric subspace of $\mathcal{H}^{\otimes 2}$ respectively. The first outcome occurs with probability $\frac{1}{2}(1 + \mathrm{tr}(\rho\sigma))$, and when this is the case, we say the swap test accepts or passes. By applying group representation theory, it is shown in [BOW19] that $O(d/\varepsilon^2)$ copies of quantum states suffices to distinguish whether an unknown ρ is equal to some known σ or ε-far from σ in trace distance. This method is efficient since the tomography needs $\Omega(d^2)$ copies of quantum states.

Let $\mathcal{P}$ be a finite set of pure states such that $\min_{\varphi \neq \varphi' \in \mathcal{P}} \mathrm{D}(\varphi, \varphi') =: \delta$, and let ψ be an unkown pure state. Then $O(\max\{\varepsilon^{-2}, \delta^{-2}\} \ln |\mathcal{P}|)$ copies suffice to distinguish whether $\psi \in \mathcal{P}$ or $\mathrm{D}(\psi, \mathcal{P}) \geq \varepsilon$ [Wan11]. By using the method in [BOW19] one can use $O(\varepsilon^{-2}|\mathcal{P}|)$ copies of the unknown state ρ to distinguish whether $\rho \in \mathcal{P}$ or $\min_{\sigma \in \mathcal{P}} \|\rho - \sigma\|_2 \geq \varepsilon$. We also show that $O(\varepsilon^{-4} \ln |\mathcal{P}|)$ copies of unknown state ψ suffice to distinguish whether $\psi \in \mathcal{P}$ or $\mathrm{D}(\psi, \mathcal{P}) \geq \varepsilon$. When δ is small compared with ε, our method exhibits advantage by using notably less samples.

The certification of unitaries is quite different from that of quantum states in that the quantum unitary certification requires a double optimization, one is the input state for the quantum unitary and the other is the choice of measurement after the action of quantum unitary. The works in [dSLCP11, SdSF$^+$12, RGK13] used methods based on Monte Carlo sampling to estimate the fidelity of an unknown gate to a fixed one, and also studied the optimality of estimation strategy. Given an unknown unitary U and a known or unknown unitary V, by using the Choi correspondence between quantum unitaries and states, there exists a tester that distinguishes whether their distance is zero or larger than ε with $O(\varepsilon^{-2})$ uses of unitaries. By using the Schur-Weyl decomposition, we show that for fixed dimension of the unitary, only $O(\varepsilon^{-1})$ uses of the unitaries suffice to achieve the same goal. Another advantage of our algorithm is that we do not need to introduce extra ancilla system as in the method using Choi state.

The productness property can be tested with $O(\varepsilon^{-2})$ copies [HM13] using the procedure first discussed in [MKB05] which applies the swap test across each corresponding pair of subsystems of two copies of $|\psi\rangle$. In [HM13], using the Choi-Jamiołkowski isomorphism, testing product unitaries is reduced to testing product pure states. The tasks of testing Pauli matrices and Clifford gates are also studied in [MO10] and [Low09, Wan11] respectively. The method based on group representation theory may be useful in testing other properties of quantum states and unitaries, which deserves further study.

4.2 Testing membership of a finite set of states

In the study of testing whether an unknown state is close to or far from a given state and whether two unknown states are close to each other, Bădescu *et al.* [BOW19] used the Chebyshev inequality to bound the deviation of a random variable from its expectation so that the completeness and soundness conditions can be satisfied. To be specific, for a real-valued random variable X and a scalar $0 < \gamma < \frac{1}{2}$, consider a tester which reports $\mathbb{E}X \le (1-2\gamma)\theta$ when $X \le (1-\gamma)\theta$ is observed and reports $\mathbb{E}X \ge \theta$ when $X > (1-\gamma)\theta$ is observed. When $\mathbb{E}X \le (1-2\gamma)\theta$, the probability that the tester reports correctly is

$$\Pr(X \le (1-\gamma)\theta) \ge 1 - \frac{\mathrm{Var}X}{((1-\gamma)\theta - \mathbb{E}X)^2} \ge 1 - \frac{\mathrm{Var}X}{(\gamma\theta)^2}\,. \tag{4.1}$$

When $\mathbb{E}X \ge \theta$, the probability that the tester reports correctly is

$$\Pr(X > (1-\gamma)\theta) \ge 1 - \frac{\mathrm{Var}X}{(\mathbb{E}X - (1-\gamma)\theta)^2} \ge 1 - \frac{\mathrm{Var}X}{(\gamma\theta)^2}\,. \tag{4.2}$$

If $\mathrm{Var}X$ is small enough as compared with $\gamma\theta$, the tester is eligible. In our settings, let X denote the random variable of measurement outcome of M on a state ϱ. Obviously the expectation of X is $\mathbb{E}X = \mathrm{tr}(\varrho M)$ and the variance of X is $\mathrm{Var}X = \mathrm{tr}(\varrho M^2) - (\mathrm{tr}\,\varrho M)^2$. A quantum algorithm was proposed in [BOW19] to test whether an unknown state is close to or far from a fixed state:

Proposition 23 ([BOW19])**.** *Given access to an unknown state ρ, for a fixed state σ and a constant $\alpha \in (0,1]$, there exists a tester that distinguishes whether $\|\rho-\sigma\|_2 \le (1-\alpha)\varepsilon$ or $\|\rho-\sigma\|_2 \ge \varepsilon$ using $O(\varepsilon^{-2})$ copies of ρ.*

Proof sketch [BOW19]. Choose an orthonormal basis such that $\sigma = \mathrm{diag}(\beta_1, \ldots, \beta_d)$. Let Π_t denote the projector onto the subspace of $(\mathbb{C}^d)^{\otimes n}$ spanned by basis states of type t. Define observables $M = \sum_{t \in \mathrm{Type}(n,d)} \frac{\langle \beta, t\rangle}{n} \Pi_t$ and $N = \sum_{1 \le i < j \le n} W_{(ij)} / \binom{n}{2}$, where $W_{(ij)}$ is the operator swapping the i-th and j-th systems. Denote $L := N + \mathrm{tr}(\sigma^2)\mathbb{1} - 2M$, and denote by X the measurement outcome of the observable L on state $\rho^{\otimes n}$. It can be calculated that $\mathbb{E}X = \|\rho - \sigma\|_2^2$, and $\mathrm{Var}X = O\big(\frac{1}{n^2} + \frac{\|\rho-\sigma\|_2^2}{n}\big)$. Finally taking $\gamma = \alpha - \frac{1}{2}\alpha^2$ and $\theta = \varepsilon^2$, and using Eqs. (4.1) and (4.2) we get that $O(\varepsilon^{-2})$ copies of ρ are sufficient. $\square$

In order to test whether an unknown state is contained in or far from a given finite set of states, we need a revised tester. For the random variable $Y = \min\{X_1, \ldots, X_m\}$ where X_i's are independent random variables, consider the tester which accepts (i.e. reports $\min_i \mathbb{E}X_i \le (1-2\gamma)\theta$) when $\min_i X_i \le (1-\gamma)\theta$ and rejects (i.e. reports $\min_i \mathbb{E}X_i \ge \theta$) otherwise.

When $\min \mathbb{E}X_i \le (1-2\gamma)\theta$, we have

$$\begin{aligned}\Pr(\min X_i \le (1-\gamma)\theta) &= 1 - \prod\nolimits_i \Pr(X_i > (1-\gamma)\theta) \\ &\ge 1 - \Pr(X_k > (1-\gamma)\theta) \\ &\ge 1 - \frac{\mathrm{Var}X_k}{(\gamma\theta)^2}\,,\end{aligned} \tag{4.3}$$

where $\mathbb{E}X_k = \min_i \mathbb{E}X_i$.

When $\min \mathbb{E}X_i \ge \theta$, we have

$$\begin{aligned}\Pr(\min X_i > (1-\gamma)\theta) &= \prod\nolimits_i \Pr(X_i > (1-\gamma)\theta) \\ &\ge \prod\nolimits_i \Big(1 - \frac{\mathrm{Var}X_i}{(\mathbb{E}X_i - (1-\gamma)\theta)^2}\Big)\,.\end{aligned} \tag{4.4}$$

Using the above bounds, Proposition 23 yields the following:

Corollary 24. *Given a finite set $\mathcal{P}$ of states and an unknown state ρ, there exists a tester that distinguishes whether $\rho \in \mathcal{P}$ or $\min_{\sigma\in\mathcal{P}} \|\rho - \sigma\|_2 \ge \varepsilon$ using $O(\varepsilon^{-2}|\mathcal{P}|)$ copies of the unknown state.*

Proof. Take $\gamma = \frac{1}{2}$, $\theta = \varepsilon^2$ and denote $m = |\mathcal{P}|$. Using Eq. (4.3) and the estimate of variance of X_k, $O(\varepsilon^{-2})$ copies of ρ suffice to ensure that $\Pr(\min X_i \le (1-\gamma)\theta) \ge 2/3$. For Eq. (4.4), $\Pr(\min X_i > (1-\gamma)\theta) \ge \big(1 - c\varepsilon^{-4}(\frac{1}{n^2} + \frac{\varepsilon^2}{n})\big)^m$ for some constant c and large n. So n should satisfy that $1 - c\varepsilon^{-4}(\frac{1}{n^2} + \frac{\varepsilon^2}{n}) \ge (\frac{2}{3})^{1/m}$. In order for this inequality to hold, take $n = c\varepsilon^{-2}\big(1 - (\frac{2}{3})^{1/m}\big)^{-1} = O(\varepsilon^{-2}m)$. $\square$

We now apply the exponential Markov inequality to bound the deviation of X from its expectation, yielding an alternative sample complexity for testing the property of a finite set of pure states. When $\mathbb{E}X \leq (1-2\gamma)\theta$, the tester accepts with probability

$$\Pr(X \leq (1-\gamma)\theta) = 1 - \Pr(X > (1-\gamma)\theta) \geq 1 - \frac{\mathbb{E}\mathrm{e}^{sX}}{\mathrm{e}^{s(1-\gamma)\theta}}\,, \tag{4.5}$$

for any $s > 0$. When $\mathbb{E}X \geq \theta$, then the tester rejects with probability

$$\Pr(X > (1-\gamma)\theta) = 1 - \Pr(X \leq (1-\gamma)\theta) \geq 1 - \frac{\mathbb{E}\mathrm{e}^{sX}}{\mathrm{e}^{s(1-\gamma)\theta}}\,, \tag{4.6}$$

for any $s < 0$.

Theorem 25. *Given a finite set $\mathcal{P}$ of pure states and an unknown pure state ψ, there exists a tester with perfect completeness that distinguishes whether $\psi \in \mathcal{P}$ or $\mathrm{D}(\psi, \mathcal{P}) \geq \varepsilon$ using $O(\varepsilon^{-4} \ln |\mathcal{P}|)$ copies of the unknown state.*

Proof. For any $\varphi \in \mathcal{P}$, choose an orthonormal basis such that $\varphi = \mathrm{diag}(1, 0, \ldots, 0)$ in this basis, and denote the diagonal elements of ψ by $x_1, x_2, \ldots, x_d$ in this basis. Let Π_t denote the projector onto the subspace of $(\mathbb{C}^d)^{\otimes n}$ spanned by states of type t. Consider the observable $M = \sum_{t \in \mathrm{Type}(n,d)} \frac{t_1}{n} \Pi_t$ which was first used in [BOW19]. Let X denote the measurement outcome of M on $\psi^{\otimes n}$, then $\mathbb{E}X = \mathrm{tr}(\psi^{\otimes n} M) = \langle \psi, \varphi \rangle = x_1$. The moment-generating function of X is

$$\begin{aligned}
\mathbb{E}\mathrm{e}^{sX} = \mathrm{tr}(\psi^{\otimes n} \mathrm{e}^{sM}) &= \sum_{t_1 + \cdots + t_d = n} \mathrm{e}^{st_1/n}\, \mathrm{tr}(\psi^{\otimes n} \Pi_t) \\
&= \sum_{t_1=0}^{n} \mathrm{e}^{st_1/n} \sum_{t_2 + \cdots + t_d = n - t_1} \mathrm{tr}(\psi^{\otimes n} \Pi_t) \\
&= \sum_{t_1=0}^{n} \mathrm{e}^{st_1/n} \sum_{t_2 + \cdots + t_d = n - t_1} \binom{n}{t} x_1^{t_1} \cdots x_d^{t_d} \\
&= \sum_{t_1=0}^{n} \mathrm{e}^{st_1/n} \binom{n}{t_1} x_1^{t_1} (1 - x_1)^{n - t_1} \\
&= \left(1 + (\mathrm{e}^{s/n} - 1) x_1\right)^n ,
\end{aligned}$$

where $\binom{n}{t} := \frac{n!}{t_1! \cdots t_d!}$ and $\binom{n}{t_1} := \frac{n!}{t_1!(n-t_1)!}$.

When $\mathbb{E}X = x_1 \leq 1 - \varepsilon^2$, take $s = 2n\varepsilon^2$. Since $x_1 = \langle \psi, \varphi \rangle \leq 1 - \varepsilon^2$, we have

$$\begin{aligned}\mathbb{E}e^{sX} &\leq \left(1 + (e^{2\varepsilon^2} - 1)(1 - \varepsilon^2)\right)^n \\ &= (1 + 2\varepsilon^2 + O(\varepsilon^6))^n \,.\end{aligned}$$

Using Eqs. (4.5) and (4.6) with $\gamma = \varepsilon^2/2$ and $\theta = 1$, it follows that

$$\begin{aligned}\Pr(X \leq 1 - \varepsilon^2/2) &\geq 1 - \frac{\mathbb{E}e^{sX}}{e^{s(1-\varepsilon^2/2)}} \\ &\geq 1 - \frac{(1 + 2\varepsilon^2 + O(\varepsilon^6))^n}{e^{n(2\varepsilon^2 - \varepsilon^4)}} \\ &= 1 - (1 - 2\varepsilon^4 + O(\varepsilon^6))^n \,.\end{aligned}$$

When $\mathbb{E}X = x_1 = 1$, the measurement outcome X of M is 1 with certainty.

Now we consider the task of testing whether an unknown state ψ is contained in or far from the set $\mathcal{P}$. For each $\varphi \in \mathcal{P}$, one can define an corresponding observable M_φ. Measure the observable M_φ on the state $\psi^{\otimes n}$, and denote the outcome by X_φ. Consider the tester that reports $\mathrm{D}(\psi, \mathcal{P}) \geq \varepsilon$ i.e. $\max_\varphi \mathbb{E}X_\varphi \leq 1 - \varepsilon^2$ when $\max_\varphi X_\varphi \leq 1 - \frac{\varepsilon^2}{2}$, and reports $\psi \in \mathcal{P}$ otherwise. For the soundness condition, in order for

$$\Pr\left(\max_\varphi X_\varphi \leq 1 - \varepsilon^2/2\right) = \Pi_\varphi \Pr(X_\varphi \leq 1 - \varepsilon^2/2) \geq \left(1 - (1 - 2\varepsilon^4 + O(\varepsilon^6))^n\right)^{|\mathcal{P}|} \geq \frac{2}{3}$$

to hold, it suffices to take

$$n = \frac{\ln\left(1 - (\frac{2}{3})^{1/|\mathcal{P}|}\right)}{\ln(1 - 2\varepsilon^4 + O(\varepsilon^6))} = O(\varepsilon^{-4} \ln |\mathcal{P}|) \,.$$

As for the completeness condition, when $\psi \in \mathcal{P}$, $\max_\varphi X_\varphi = 1$. Thus this tester has perfect completeness. Therefore using $O(\varepsilon^{-4} \ln |\mathcal{P}|)$ copies of ψ the tester fulfills both completeness and soundness conditions. □

Denote $\delta := \{\mathrm{D}(\varphi_1, \varphi_2) : \varphi_1 \neq \varphi_2 \in \mathcal{P}\}$. It was shown in [Wan11] that $O(\max\{\varepsilon^{-2}, \delta^{-2}\} \ln |\mathcal{P}|)$ copies of states suffice to distinguish whether $\psi \in \mathcal{P}$ or $\mathrm{D}(\psi, \mathcal{P}) \geq \varepsilon$. When $\delta = O(\varepsilon^2)$, our method is more sample-efficient than that in [Wan11]. The method based on Schur-Weyl decomposition used in this chapter may be useful in the certification of bipartite or multipartite states.

4.3 Testing equality of unitaries

Following [MdW16], define the distance between two unitaries $U, V \in \mathrm{U}(d)$ as

$$\operatorname{dist}(U,V) = \sqrt{1 - \left|\frac{1}{d}\langle U, V\rangle\right|^2}, \tag{4.7}$$

where $\langle U, V\rangle := \operatorname{tr}(U^\dagger V)$. The distance of two unitaries is no larger than 1.

Notice that the normalized inner product of unitaries is equal to the inner product of their Choi states, i.e.,

$$\frac{1}{d}\langle U, V\rangle = \langle \phi^{\mathrm{m}}|(U^\dagger \otimes \mathbb{1})(V \otimes \mathbb{1})|\phi^{\mathrm{m}}\rangle\,. \tag{4.8}$$

Here and in the following we use $|\phi^{\mathrm{m}}\rangle$ to denote the (normalized) maximally entangled state. Consider the Schur-Weyl decomposition

$$(\mathbb{C}^d)^{\otimes n} = \bigoplus_{\lambda \vdash (n,d)} \mathcal{H}_\lambda \otimes \mathcal{K}_\lambda\,,$$

where $\mathcal{H}_\lambda$ are irreducible representation spaces of $\mathrm{U}(d)$ and $\mathcal{K}_\lambda$ are multiplicity spaces. The entanglement between the representation space and multiplicity space was used to estimate the group transformation [CDS05]. It turns out that it can be also used in the certification of unitaries as shown in the following.

Theorem 26. *Given access to an unknown single-qubit unitary U, there exists a tester with perfect completeness that distinguishes whether U is equal to a fixed and known V up to a phase or $\operatorname{dist}(U,V) \geq \varepsilon$ with $O(\varepsilon^{-1})$ uses of U, without using any ancilla system.*

Proof. By Schur-Weyl duality, $(\mathbb{C}^2)^{\otimes n} = \bigoplus_j \mathcal{H}_j \otimes \mathcal{K}_j$ where $\mathcal{H}_j$ is the irrep of $\mathrm{U}(2)$ of dimension $2j+1$ and $\mathcal{K}_j$ is the corresponding irrep of S_n. In this decomposition we write $U^{\otimes n} = \bigoplus_j U_j \otimes \mathbb{1}_j$ for any $U \in \mathrm{U}(2)$ since $U^{\otimes n}$ acts as identity operator on each $\mathcal{K}_j$.

Notice that $\mathcal{H}_{\frac{n}{2}-1}$ and $\mathcal{K}_{\frac{n}{2}-1}$ have the same dimension $n-1$. Let $|\phi^{\mathrm{m}}\rangle$ be the maximally entangled state in $\mathcal{H}_{\frac{n}{2}-1} \otimes \mathcal{K}_{\frac{n}{2}-1}$, i.e. $|\phi^{\mathrm{m}}\rangle = \frac{1}{\sqrt{n-1}}\sum_{i=1}^{n-1} |\alpha_i\rangle|\beta_i\rangle$ where $\{|\alpha_i\rangle\}$ and $\{|\beta_i\rangle\}$ are orthonormal bases of $\mathcal{H}_{\frac{n}{2}-1}$ and $\mathcal{K}_{\frac{n}{2}-1}$ respectively.

Apply the unitary $U^{\otimes n}$ to $|\phi^{\mathrm{m}}\rangle$, and then perform the POVM $\{\phi^{\mathrm{m}}_V, \mathbb{1} - \phi^{\mathrm{m}}_V\}$ where $|\phi^{\mathrm{m}}_V\rangle := (V_{\frac{n}{2}-1} \otimes \mathbb{1}_{\frac{n}{2}-1})|\phi^{\mathrm{m}}\rangle$. The tester reports that U equals V up to a phase if the first outcome occurs, and reports $\operatorname{dist}(U,V) \geq \varepsilon$ otherwise. Obviously the tester has perfect completeness irrespective of the choice of n. Next step is to derive the requirement for n to ensure the soundness condition $\langle U^{\otimes n}\phi^{\mathrm{m}}U^{\dagger,\otimes n}, \phi^{\mathrm{m}}_V\rangle \leq 1/3$ when $\mathrm{dist}(U,V) \geq \varepsilon$.

Denote the eigenvalues of $U^\dagger V$ by $\mathrm{e}^{\mathrm{i}\alpha}$ and $\mathrm{e}^{\mathrm{i}\beta}$ for $\alpha,\beta \in [0,2\pi)$. When $\mathrm{dist}(U,V) \geq \varepsilon$, then by the definition (4.7), $|\langle U,V\rangle|^2 \leq 4(1-\varepsilon^2)$, i.e. $|\mathrm{e}^{\mathrm{i}\alpha}+\mathrm{e}^{\mathrm{i}\beta}|^2 \leq 4(1-\varepsilon^2)$. It follows that $|\sin\frac{1}{2}(\alpha-\beta)| \geq \varepsilon$. We thus have

$$\begin{aligned}
\sqrt{\langle U^{\otimes n}\phi^{\mathrm{m}}U^{\dagger,\otimes n}, \phi^{\mathrm{m}}_V\rangle} &= \Big|\frac{1}{n-1}\big\langle U_{\frac{n}{2}-1}, V_{\frac{n}{2}-1}\big\rangle\Big| \\
&= \Big|\frac{1}{n-1}\frac{\mathrm{e}^{\mathrm{i}(n\alpha+\beta)}-\mathrm{e}^{\mathrm{i}(\alpha+n\beta)}}{\mathrm{e}^{\mathrm{i}\alpha}-\mathrm{e}^{\mathrm{i}\beta}}\Big| \\
&= \Big|\frac{1}{n-1}\frac{\sin\frac{n-1}{2}(\alpha-\beta)}{\sin\frac{1}{2}(\alpha-\beta)}\Big| \\
&\leq \Big|\frac{1}{n-1}\frac{1}{\sin\frac{1}{2}(\alpha-\beta)}\Big| \\
&\leq \frac{1}{(n-1)\varepsilon}\,,
\end{aligned} \tag{4.9}$$

where the second equality used the Weyl character formula (1.6) for irrep $\mathcal{H}_{\frac{n}{2}-1}$. We now take $n = \frac{\sqrt{3}}{\varepsilon}+1 = O(\varepsilon^{-1})$ so that the soundness condition $\langle U^{\otimes n}\phi^{\mathrm{m}}U^{\dagger,\otimes n}, \phi^{\mathrm{m}}_V\rangle \leq 1/3$ for $\mathrm{dist}(U,V) \geq \varepsilon$ is satisfied. □

When both single-qubit unitaries are unknown, we have the following.

Theorem 27. *Given access to two unknown single-qubit unitaries U and V, there exists a tester with perfect completeness that distinguishes whether U equals V up to a phase or $\mathrm{dist}(U,V) \geq \varepsilon$ with $O(\varepsilon^{-1})$ uses of U and V, without using any ancilla system.*

Proof. Similar to the proof of Theorem 26, we first decompose the space as $(\mathbb{C}^2)^{\otimes n} = \bigoplus_j \mathcal{H}_j \otimes \mathcal{K}_j$ where $\mathcal{H}_j$ and $\mathcal{K}_j$ are irreps of $\mathrm{U}(2)$ and S_n respectively. Denote by $|\phi^{\mathrm{m}}\rangle$ the maximally entangled state in $\mathcal{H}_{\frac{n}{2}-1}\otimes\mathcal{K}_{\frac{n}{2}-1}$. Then we apply the swap test to $U^{\otimes n}|\phi^{\mathrm{m}}\rangle$ and $V^{\otimes n}|\phi^{\mathrm{m}}\rangle$. Repeat the above procedure, and report that the two unitaries are equal up to a phase if and only if the swap test accepts twice. When U and V are equal, the tester reports correctly with certain. When $\mathrm{dist}(U,V) \geq \varepsilon$, since $\big|\langle\phi^{\mathrm{m}}|U^{\dagger,\otimes n}V^{\otimes n}|\phi^{\mathrm{m}}\rangle\big| \leq \frac{1}{(n-1)\varepsilon}$ by (4.9), the tester reports incorrectly with probability less than $\big(\frac{1}{2}(1+\frac{1}{(n-1)^2\varepsilon^2})\big)^2$, which is in turn less than $1/3$ if we take $n = \frac{3}{\varepsilon}+1$. □

In order to deal with the higher dimensional case, an upper bound on $|\operatorname{tr} U_\lambda|/d_\lambda$

is needed given $|\operatorname{tr} U|/d \le 1-\varepsilon^2$, where U_λ is representation matrix of U on irrep $\mathcal{H}_\lambda$ for appropriate $\lambda \in \mathrm{Par}(n)$ and d_λ is the dimension of $\mathcal{H}_\lambda$.

Lemma 28. *Let $\lambda = (d-1, d-2, \ldots, 0)n/\binom{d}{2} \in \mathrm{Par}(n,d)$ and $s = n/\binom{d}{2} + 1$. If $\frac{1}{d}|\langle U, V\rangle| \le 1-\varepsilon^2$ for $U, V \in \mathrm{U}(d)$, then*

$$\frac{1}{\dim \mathcal{H}_\lambda}|\langle U_\lambda, V_\lambda\rangle| \le \left(\frac{2}{s\varepsilon}\right)^m$$

for some positive m.

Proof. The dimension of the irrep $\mathcal{H}_\lambda$ is

$$\dim \mathcal{H}_\lambda = s^{d(d-1)/2}\,. \tag{4.10}$$

The character for $\mathcal{H}_\lambda$ is

$$\chi_\lambda^L(\mathrm{diag}(x_1, \ldots, x_d)) = \frac{\prod_{1\le j<k\le d}(x_k^s - x_j^s)}{\prod_{1\le j<k\le d}(x_k - x_j)}\,.$$

Since $\mathsf{R}_\lambda : U \mapsto U_\lambda$ is a unitary representation of $\mathrm{U}(d)$, we have $U_\lambda V_\lambda = (UV)_\lambda$ and $(U_\lambda)^\dagger = (U^\dagger)_\lambda$. Thus $\langle U_\lambda, V_\lambda\rangle = \langle (U^\dagger V)_\lambda, \mathbb{1}\rangle$. It suffices to consider the case $V = \mathbb{1}$.

Consider a unitary U having eigenvalues $\mathrm{e}^{\mathrm{i}\theta_k}$ with $0 \le \theta_k < 2\pi$ for each $k \in [d]$. We have

$$\begin{aligned}|\operatorname{tr} U|^2 = \left|\sum_{k=1}^d \mathrm{e}^{\mathrm{i}\theta_k}\right|^2 &= d + 2\sum\nolimits_{1\le j<k\le d}\cos(\theta_k - \theta_j) \\ &= d^2 - 4\sum\nolimits_{1\le j<k\le d}\sin^2\frac{1}{2}(\theta_k - \theta_j)\,,\end{aligned} \tag{4.11}$$

while

$$\begin{aligned}|\operatorname{tr} U_\lambda| &= \prod_{1\le j<k\le d}\left|\frac{\mathrm{e}^{\mathrm{i}s\theta_k} - \mathrm{e}^{\mathrm{i}s\theta_j}}{\mathrm{e}^{\mathrm{i}\theta_k} - \mathrm{e}^{\mathrm{i}\theta_j}}\right| \\ &= \prod_{1\le j<k\le d}\left|\frac{\sin\frac{s}{2}(\theta_k - \theta_j)}{\sin\frac{1}{2}(\theta_k - \theta_j)}\right|.\end{aligned} \tag{4.12}$$

Before proceeding with the proof we now show that

$$\left|\frac{\sin sx}{\sin x}\right| \le s \tag{4.13}$$

holds for any odd positive integer s and any $|x| \le \pi$. Indeed, it suffices to consider the case $0 \leq \sin x \leq \frac{1}{s}$, and (4.13) follows by noticing that the function $\frac{\sin sx}{s \sin x}$ is even and is symmetric about the line $x = \frac{\pi}{2}$. When $0 \leq \sin x \leq \frac{1}{s}$, since $\sin x \geq \frac{2}{\pi}x$, we have $sx \leq \frac{\pi}{2}$. It follows that for any $s \geq 1$, we have $\cos x \geq \cos sx$, and thus $s \sin x \geq \sin sx$, completing the proof of (4.13).

As for the soundness condition, when $\frac{1}{d}|\operatorname{tr} U| \leq 1 - \varepsilon^2$, by (4.11) we have

$$\sum_{1 \leq j < k \leq d} \sin^2 \frac{1}{2}(\theta_k - \theta_j) \geq \frac{1}{4} d^2 (2\varepsilon^2 - \varepsilon^4) \,.$$

It follows that there exists (j, k) satisfying

$$\sin^2 \frac{1}{2}(\theta_k - \theta_j) \geq \frac{d(2\varepsilon^2 - \varepsilon^4)}{2(d-1)} \,,$$

and let m be the number of such pairs in totally $\binom{d}{2}$ pairs. For any such pair,

$$\frac{|\sin \frac{s}{2}(\theta_k - \theta_j)|}{|\sin \frac{1}{2}(\theta_k - \theta_j)|} \leq \frac{1}{|\sin \frac{1}{2}(\theta_k - \theta_j)|} \leq \sqrt{\frac{2(d-1)}{d(2\varepsilon^2 - \varepsilon^4)}} \leq \frac{2}{\varepsilon} \tag{4.14}$$

for $\varepsilon \leq 1$. On the other hand, there are $\binom{d}{2} - m$ pairs of (j, k) satisfying $\sin^2 \frac{1}{2}(\theta_k - \theta_j) < \frac{d(2\varepsilon^2 - \varepsilon^4)}{2(d-1)}$. For any such pair, since $|\frac{1}{2}(\theta_k - \theta_j)| \leq \pi$, by (4.13) we have

$$\frac{|\sin \frac{s}{2}(\theta_k - \theta_j)|}{|\sin \frac{1}{2}(\theta_k - \theta_j)|} \leq s \,. \tag{4.15}$$

Therefore, combining (4.10), (4.12), (4.14) and (4.15),

$$\frac{1}{\dim \mathcal{H}_\lambda} |\operatorname{tr} U_\lambda| \leq s^{-\binom{d}{2}} (2/\varepsilon)^m s^{-m + \binom{d}{2}} = \left(\frac{2}{s\varepsilon}\right)^m \,.$$

□

Theorem 29. *Given access to an unknown unitary $U \in \mathrm{U}(d)$ and a known or unknown unitary $V \in \mathrm{U}(d)$, there exists a tester with perfect completeness that distinguishes whether $\operatorname{dist}(U, V) = 0$ or $\operatorname{dist}(U, V) \geq \varepsilon$ with $O(d^2/\varepsilon)$ uses of unitaries.*

Proof. Consider the partition $\lambda = (d-1, d-2, \ldots, 0) n / \binom{d}{2}$, and denote $s := n / \binom{d}{2} + 1$. The proof follows similar approach used in Theorems 26 and 27.

By noticing that λ / n is independent of n, it follows from (1.5) and (1.7) that when ε is so small that n is much larger than d, the dimension of $\mathcal{K}_\lambda$ is exponential in n while the dimension of $\mathcal{H}_\lambda$ is polynomial in n, thus $\dim \mathcal{K}_\lambda$ is larger than $\dim \mathcal{H}_\lambda$.

Denote by $|\phi^{\mathrm{m}}\rangle$ the maximally entangled state in $\mathcal{H}_\lambda \otimes \mathcal{K}_\lambda$, and denote $|\phi^{\mathrm{m}}_U\rangle := U^{\otimes n}|\phi^{\mathrm{m}}\rangle$ and $|\phi^{\mathrm{m}}_V\rangle = V^{\otimes n}|\phi^{\mathrm{m}}\rangle$. When U is unknown and V is given, perform a measurement $\{\phi^{\mathrm{m}}_V, \mathbb{1} - \phi^{\mathrm{m}}_V\}$ on ϕ^{m}_U, and repeat, and accept iff the first outcome occurs. When U and V are both unknown, perform a swap test for ϕ^{m}_U and ϕ^{m}_V, and repeat.

The completeness condition is obvious. When $\mathrm{dist}(U,V) \geq \varepsilon$, using Lemma 28, the soundness condition is satisfied by taking s to be the smallest odd integer larger than $6/\varepsilon$, for which $n \leq \binom{d}{2}(\frac{6}{\varepsilon}+1) = O(d^2/\varepsilon)$. □

For the asymptotic case where the value of ε is small, n will be so large that the irrep $\mathcal{H}_\lambda$ has smaller dimension than its multiplicity space $\mathcal{K}_\lambda$ does. Thus there exists a maximally entangled state ϕ^{m} on $\mathcal{H}_\lambda \otimes \mathcal{H}_\lambda$ which is a subspace of $\mathcal{H}_\lambda \otimes \mathcal{K}_\lambda$. If ε is not small enough and thus n is not large enough, one can introduce a reference system of dimension $\frac{\dim \mathcal{H}_\lambda}{\dim \mathcal{K}_\lambda}$ to make them have equal dimension [CDS05]. The dimension of the ancilla system, however, has been exponentially decreased compared with the method using Choi states directly. Since it is common to deal with low-dimensional quantum systems in quantum computing under current technology, our method exhibits advantage in practice.

For the certification of identity of qudit unitaries, we have considered the irrep corresponding partition $\lambda = (d-1, d-2, \ldots, 0)n/\binom{d}{2}$. One issue that would be worth investigating is whether the irreps of unitary group used in this chapter are optimal compared with other irreps.

Using Theorem 29, one can also efficiently test whether one unitary is identity, and whether two unitaries are inverse to each other. It is known that many properties of quantum channel can be reduced to testing properties of quantum state via the Choi-Jamiołkowski isomorphism, but by employing this approach one usually needs to use quantum channels too many times and also to introduce extra ancilla system. In this section we used group representation to test properties of unitaries more sample-efficiently without using ancilla system. Our approach may be promising in the property testing of noisy quantum operation and measurement to achieve better performance.

Chapter 5

Parallel repetition for extended nonlocal games

5.1 Introduction and overview

The model of nonlocal games has been extensively studied in theoretical physics and computational complexity theory. In physics, the celebrated Bell inequality experiments, proposed by Bell [Bel64] and then studied by Clauser *et al.* [CHSH69], are built on nonlocal game as theoretical basis. The experiments provided substantial evidence for that the world is not locally realistic, and made the theory of quantum mechanics more convincing. In computational complexity theory, nonlocal game is employed as a simple model via which interactive proof systems can be analyzed [GMR85, Bab85].

In a two-player one-round nonlocal game G, the referee selects a pair of questions (x_1, x_2) at random from a set according to a probability distribution μ which is fixed and publicly known, then sends x_1 to one player and x_2 to the other, who then respond with answers a and b respectively. Upon receiving the answers a_1 and a_2, the referee evaluates a publicly known predicate $V(a_1, a_2, x_1, x_2)$. The two players jointly win the game if the predicate $V(a_1, a_2, x_1, x_2)$ is satisfied, and they lose otherwise. Before the game starts, the players can corroborate on a strategy based on μ and V to maximize their winning probability. During the game, each player is unware of the questions received by other players. For example, in the CHSH game, the question pair (x, y) is randomly selected uniformly from the set $\{0, 1\} \times \{0, 1\}$, the answer from each player is a single bit, and the predicate is $a \oplus b = x \wedge y$. The maximum winning probability for the players is $\frac{3}{4}$, but using entanglement shared by the two players their winning probability can be $\cos^2 \frac{\pi}{8} > 0.85$.

We first fix some notation for discussion. Given finite alphabets $\mathcal{A}$ and $\mathcal{B}$, for a function $P_{\mathcal{AB}}$ on $\mathcal{A} \times \mathcal{B}$, denote $P_{\mathcal{A}}(a) := \sum_{b \in \mathcal{B}} P_{\mathcal{AB}}(a, b)$ for each $a \in \mathcal{A}$. For a space $\mathcal{H}$ which may be trivial and a function $V : \mathcal{B} \to \mathcal{L}(\mathcal{H})$, we define $V^{\wedge n} : \mathcal{B}^n \to \mathcal{L}(\mathcal{H}^{\otimes n}), b^n \mapsto V(b^{(1)}) \otimes \cdots \otimes V(b^{(n)})$ for $b^n := (b^{(1)}, \ldots, b^{(n)}) \in \mathcal{B}^n$.

In the rest of this chapter we denote $\mathcal{A} := \mathcal{A}_1 \times \mathcal{A}_2$ and $\mathcal{X} := \mathcal{X}_1 \times \mathcal{X}_2$ with each $\mathcal{A}_i, \mathcal{X}_i$ being finite alphabet, and correspondingly we write $a := (a_1, a_2) \in \mathcal{A}$, $x := (x_1, x_2) \in \mathcal{X}$. Let μ be a probability distribution (or called distribution for short) on a set $\mathcal{X}$, and let $V : \mathcal{A} \times \mathcal{X} \to \{0, 1\}$ be a function. The game value $\mathrm{val}(G)$ of the two-player game G specified by μ and V is defined as the maximum winning probability

$$\max_{P} \mathop{\mathbb{E}}_{x \sim \mu} \sum_{a \in \mathcal{A}} P(a|x) V(a, x) \,,$$

where the maximum is over some class of strategy P which is a conditional probability distribution on $\mathcal{A}|\mathcal{X}$. Based on the strategy class, the nonlocal games can be divided into several categories. For classical strategy, $P(a|x)$ can be written as $P(a|x) = P_1(a_1|x_1)P_2(a_2|x_2)$ for some conditional distributions P_i. For quantum strategy, or called entangled strategy, it can be written as $P(a|x) = \langle \rho, E_1(a_1|x_1) \otimes E_2(a_2|x_2) \rangle$ where ρ is a quantum state shared by the two players and $\{E_i(a_i|x_i)\}_{a_i}$ is a POVM for each x_i. Nonsignaling strategies are those strategies which do not imply communication.

The parallel repetition question is the following natural question: what is the value of the game G^n, in which n independent instances of G are played in parallel? In the game G^n, the questions $x^n := (x^{(1)}, \ldots, x^{(n)})$ are chosen independently, the answers $a^n := (a^{(1)}, \ldots, a^{(n)})$ are decided collectively, and the players win the game iff they win all the n instances. The players may use the information of $x^{(i)}$ to give $a^{(j)}$, so the value $\mathrm{val}(G^n)$ may be larger than $\mathrm{val}(G)^n$. For the classical strategy, Raz's parallel repetition theorem [Raz98] states that $\mathrm{val}(G^n) \leq (1 - (1 - \mathrm{val}(G))^3)^{cn}$ where c is a constant on G. The proof has been simplified and improved upon in [Hol09] and [BG16]. For the multiplayer setting, the exponential decay of the nonsignaling value of a game G was established in [BFS14], and similar results were achieved in [LW16] and [AFRV16] using a different technique based on de Finetti reduction.

The extended nonlocal game as a generalization of (ordinary) nonlocal game discussed above was studied in [JMRW16]. In this new type of games, the predicate becomes operator-valued, and the result of the game is determined by the outcome of measurement performed by the referee. The monogamy-of-entanglement games [TFKW13] and the steering game [Fri12] are examples of entended nonlocal game.

For two functions $P_{\mathcal{X}}, Q_{\mathcal{X}} : \mathcal{X} \to \mathrm{Pos}$, where Pos denotes the set of positive semidefinite operators on some Hilbert space, the trace distance D between $P_{\mathcal{X}}$ and $Q_{\mathcal{X}}$ is defined as $\mathrm{D}(P_{\mathcal{X}}, Q_{\mathcal{X}}) := \sum_{x\in\mathcal{X}} \frac{1}{2}\|P_{\mathcal{X}}(x) - Q_{\mathcal{X}}(x)\|_1$. An *operator assemblage* on $\mathcal{A}|\mathcal{X}$ is a function $P_{\mathcal{A}|\mathcal{X}} : \mathcal{A} \times \mathcal{X} \to \mathrm{Pos}$ such that $\sum_{a\in\mathcal{A}} P_{\mathcal{A}|\mathcal{X}}(a|x)$ has unit trace for each $x \in \mathcal{X}$. An operator assemblage $P_{\mathcal{A}|\mathcal{X}}$ is called *nonsignaling* [Rus17], denoted NS, if for each $i = 1, 2$ there exists a conditional distribution P'_i on $\mathcal{A}_i|\mathcal{X}_i$ such that $P(a_i|x) = P'_i(a_i|x_i)$. It is also required that $\sum_{a_1} P'_1(a_1|x_1) = \sum_{a_2} P'_2(a_2|x_2)$.

In a two-player extended game, the referee holds a system $\mathcal{R} \cong \mathbb{C}^d$ and the i-th player holds a quantum system $\mathcal{H}_i$. Let μ be a distribution on the set $\mathcal{X}$ of questions, and let $V : \mathcal{A} \times \mathcal{X} \to [0, \mathbb{1}_{\mathcal{R}}]$ be an operator-valued function. The game value of the extended nonlocal game specified by μ and V is defined as

$$\max_P \mathop{\mathbb{E}}_{x\sim\mu} \sum_{a\in\mathcal{A}} \langle P(a|x), V(a,x)\rangle \,, \tag{5.1}$$

where the maximum is over some class of operator assemblage P on $\mathcal{A}|\mathcal{X}$.

For quantum strategy in an extended game, there is a state shared by the referee and all players such that $P(a|x) = \mathrm{tr}_{\mathcal{H}}(\rho(E_1(a_1|x_1) \otimes E_2(a_2|x_2) \otimes \mathbb{1}_{\mathcal{R}}))$ where $\mathcal{H} := \mathcal{H}_1 \otimes \mathcal{H}_2$, and $\{E_i(a_i|x_i)\}_{a_i}$ is POVM for each x_i. Equivalently, $\langle P(a|x), V(a,x)\rangle = \langle \rho, E_1(a_1|x_1) \otimes E_2(a_2|x_2) \otimes V(a,x)\rangle$. The quantum and nonsignaling value of the extended game G thus can be defined as the game value using quantum and nonsignaling strategies respectively.

In an n-fold parallel repetition of a two-player game G, n pairs of questions are selected independently according to $\mu^{\wedge n}$. The questions $(x_1, \ldots, x_n)$ are sent to one player and $(y_1, \ldots, y_n)$ to the other. The players respond with values $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$ respectively. They win this game if the predicate $Q(x_i, y_i, a_i, b_i)$ is satisfied for all $i \in [n]$. Thus the value of n-fold parallel repetition of an extended nonlocal game G is

$$\max_{Q_{\mathcal{A}^n|\mathcal{X}^n}} \sum_{a^n, x^n} \mu_{\mathcal{X}}^{\wedge n}(x^n) \big\langle Q_{\mathcal{A}^n|\mathcal{X}^n}(a^n|x^n), V_{\mathcal{A}\mathcal{X}}^{\wedge n}(a^n, x^n)\big\rangle \,, \tag{5.2}$$

where $a^n := (a^{(1)}, \ldots, a^{(n)}) \in \mathcal{A}^n$, $x^n := (x^{(1)}, \ldots, x^{(n)}) \in \mathcal{X}^n$, and the maximum is over some class of strategies.

The extended games may be viewed as being equivalent to multipartite steering inequalities, in a similar way to the equivalence between nonlocal games and Bell inequalities [JMRW16]. The study of parallel repetition of extended nonlocal game is partly motivated by its connection with multipartite quantum steering. Similarly to the fact that the violation of Bell inequalities certifies the presence of entanglement in device-independent scenario, the violation of steering inequalities certifies the presence of entanglement in semi-device-independent scenario [CSA+15]. In order to distinguish two possible cases corresponding to two game values in this task, the parallel repetition is always used to amplify the gap between the two values. Thus it is necessary to study the behavior of extended game under parallel repetition.

In this chapter we consider the nonsignaling value of parallel repetition of a two-player game G, denoted $\mathrm{val}(G^n)$. When elements in $\{V(a,x) : a \in \mathcal{A}, x \in \mathcal{X}\}$ commute with each other, it suffices to consider the *commuting operator assemblage* $P_{\mathcal{A}|\mathcal{X}}$, for which elements in $\{P_{\mathcal{A}|\mathcal{X}}(a|x) : a \in \mathcal{A}, x \in \mathcal{X}\}$ commute with each other. Similarly to the nonsignaling correlations used in the ordinary games studied in [Hol09], the commuting nonsignaling operator assemblage enjoys a nice property, denoted $\mathcal{P}$. Namely, for any distribution $Q_{\mathcal{A}\mathcal{X}}$ with $\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2, \mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2$, if $\mathrm{D}(Q_{\mathcal{X}} R_{\mathcal{A}_1|\mathcal{X}_1}, Q_{\mathcal{A}_1\mathcal{X}})$ and $\mathrm{D}(Q_{\mathcal{X}} T_{\mathcal{A}_2|\mathcal{X}_2}, Q_{\mathcal{A}_2\mathcal{X}})$ are both small for some conditional distributions $R_{\mathcal{A}_1|\mathcal{X}_1}$ and $T_{\mathcal{A}_2|\mathcal{X}_2}$, then there exists a nonsignaling conditional distribution $P_{\mathcal{A}|\mathcal{X}}$ such that $\mathrm{D}(Q_{\mathcal{X}} P_{\mathcal{A}|\mathcal{X}}, Q_{\mathcal{A}\mathcal{X}})$ is also small.

Following the approach in [LW16] we show that the value of n-fold parallel repetition of a two-player game G with commuting nonsignaling strategy is no larger than $(1-c\delta)^n$ for constant c, provided that the game value of G is no larger than $1-\delta$ with $\delta > 0$. It can be seen from the proof that this parallel repetition theorem still holds for general nonsignaling strategy if any nonsignaling operator assemblage satisfies the property $\mathcal{P}$.

In order to extend the parallel repetition theorem for nonsignaling games to the extended games using the de Finetti reduction in [LW16], we need to further study properties of the operator assemblages. For the parallel repetition problem, we need to deal with the symmetry in $Q_{\mathcal{A}^n|\mathcal{X}^n}(a^n|x^n)$ which is an operator on $(\mathbb{C}^d)^{\otimes n}$, while in the ordinary game it is simply a scalar. Thus the calculation is not straightforward in the study of the parallel repetition of extended games.

5.2 Technical lemmas

In this section we will give some necessary lemmas for the proof in next section. The fidelity $\mathrm{F}(X,Y) := \left\|\sqrt{X}\sqrt{Y}\right\|_1$ and the trace distance $\mathrm{D}(X,Y) := \frac{1}{2}\|X-Y\|_1$ are two widely used functions to measure the similarity of two quantum states. Some properties of the functions are collected in the following.

Lemma 30 (See e.g. [Wat18]). *For any positive semidefinite operators X, X', Y, Y', Z and any quantum states ρ, σ,* *(1) Additivity under direct sum:* $\mathrm{F}(X \oplus X', Y \oplus Y') = \mathrm{F}(X, Y) + \mathrm{F}(X', Y')$*;*

(2) Multiplicativity under tensor product: $\mathrm{F}(X \otimes X', Y \otimes Y') = \mathrm{F}(X, Y)\mathrm{F}(X', Y')$*;*

(3) Superadditivity under addition: $\mathrm{F}(X + X', Y + Y') \geq \mathrm{F}(X, Y) + \mathrm{F}(X', Y')$*, and in particular,* $\mathrm{F}(X, Y) \leq \mathrm{F}(X, Z)$ *if* $Y \leq Z$*;*

(4) Monotonicity: for any CPTP $\mathcal{E}$*,* $\mathrm{F}(X, Y) \leq \mathrm{F}(\mathcal{E}(X), \mathcal{E}(Y))$ *and* $\mathrm{D}(X, Y) \geq \mathrm{D}(\mathcal{E}(X), \mathcal{E}(Y))$*;*

(5) Fuchs-van de Graaf inequalities [FVDG99]: $\mathrm{F}(\rho, \sigma)^2 + \mathrm{D}(\rho, \sigma)^2 \leq 1 \leq \mathrm{F}(\rho, \sigma) + \mathrm{D}(\rho, \sigma)$*.*

Lemma 31. *For each* $k \in [p]$*, denote by* ρ_k *the reduced state on* $\mathcal{H}_k$ *of a state* $\rho \in \mathcal{D}(\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_p)$*, and denote by* Π_k *the projector onto the support of* ρ_k*. It holds that*

$$\rho \leq \Pi_1 \otimes \cdots \otimes \Pi_p \,. \tag{5.3}$$

Proof. For any state σ acting on $\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_p$, let Π_k^σ denote the projector onto the support of the reduced state σ_k of σ on $\mathcal{H}_k$. We first consider the pure state case that $\rho = |\psi\rangle\langle\psi|$. Let $\{|i_k\rangle\}_{i_k=1}^{d_k}$ be an orthonormal basis of the support of ψ_k, and it can be extended to $\{|i_k\rangle\}_{i_k=1}^{D_k}$ which is an orthonormal basis of $\mathcal{H}_k$. Let the general form of $|\psi\rangle$ be $|\psi\rangle = \sum_{i_1,\ldots,i_p=1}^{D_1,\ldots,D_p} \lambda_{i_1,\ldots,i_p} |i_1 \cdots i_p\rangle$. Thus

$$\psi = \sum_{i_1,\ldots,i_p=1}^{D_1,\ldots,D_p} \sum_{i'_1,\ldots,i'_p=1}^{D_1,\ldots,D_p} \lambda_{i_1,\ldots,i_p} \lambda^*_{i'_1,\ldots,i'_p} |i_1 \cdots i_p\rangle\langle i'_1 \cdots i'_p|$$

and $\psi_1 = \sum_{i_1,i'_1,i_2,\ldots,i_p} \lambda_{i_1,\ldots,i_p} \lambda^*_{i'_1,i_2,\ldots,i_p} |i_1\rangle\langle i'_1|$. For integer t_1 satisfying $d_1 < t_1 \leq D_1$, one has $\langle t_1|\psi_1|t_1\rangle = 0$. That is, $\lambda_{t_1,i_2,\ldots,i_p} = 0$ for each $t_1 > d_1$ and each $i_2, \ldots, i_p$. Similarly, for k and t_k satisfying $d_k < t_k \leq D_k$, we have $\lambda_{i_1,\ldots,i_{k-1},t_k,i_{k+1},\ldots,i_p} = 0$. So $|\psi\rangle = \sum_{i_1,\ldots,i_p=1}^{d_1,\ldots,d_p} \lambda_{i_1,\ldots,i_p} \, |i_1 \cdots i_p\rangle$ with normalization condition $\sum_{i_1,\ldots,i_p} |\lambda_{i_1,\ldots,i_p}|^2 = 1$. It follows that $\langle\psi|\Pi_1 \otimes \cdots \otimes \Pi_p|\psi\rangle = 1$, and thus $|\psi\rangle$ is a unit vector in the subspace corresponding to the projector $\Pi_1 \otimes \cdots \otimes \Pi_p$, so $\psi \leq \Pi_1^\psi \otimes \cdots \otimes \Pi_p^\psi$.

When ρ is a general state, it can be written as a convex combination of pure states, i.e. $\rho = \sum_\psi c_\psi |\psi\rangle\langle\psi|$. Since $\rho_k \geq c_\psi \psi_k$ for each ψ and each k, we have $\mathrm{supp}(\psi_k) \subset \mathrm{supp}(\rho_k)$ where supp denotes the support space of a Hermitian operator. To see this, notice the fact that $|\varphi\rangle \in \mathrm{supp}(M)$ for a Hermitian operator M if and only if $\langle\varphi|\xi\rangle = 0$ holds for any $|\xi\rangle$ satisfying $\langle\xi|M|\xi\rangle = 0$. We thus have $\Pi_k^\psi \leq \Pi_k^\rho$ and hence $\bigotimes_k \Pi_k^\psi \leq \bigotimes_k \Pi_k^\rho$. It follows that

$$\rho = \sum_\psi c_\psi \psi \leq \sum_\psi c_\psi \bigotimes_k \Pi_k^\psi \leq \bigotimes_k \Pi_k^\rho \,,$$

completing the proof. $\square$

Similarly to [DSW16, Lemma 18], we have the following lemma.

Lemma 32. *Let $\mathcal{B} = [m]$. For each $b^n := \big(b^{(1)}, \ldots, b^{(n)}\big) \in \mathcal{B}^n$, $P_{\mathcal{B}^n}(b^n)$ is a positive semidefinite operator on $(\mathbb{C}^d)^{\otimes n}$ such that if $b^{(i)} = b^{(j)}$ for $i \neq j$, then $W_{(ij)} P_{\mathcal{B}^n}(b^n) W_{(ij)}^\dagger = P_{\mathcal{B}^n}(b^n)$, where $W_{(ij)}$ is the operator swapping the i-th and j-th systems. Then there exists some measure $\mathrm{d}\rho_{\mathcal{B}}$ on $(\mathcal{D}(\mathbb{C}^d))^{\mathcal{B}}$ such that for any $b^n \in \mathcal{B}^n$,*

$$\frac{P_{\mathcal{B}^n}(b^n)}{\operatorname{tr} P_{\mathcal{B}^n}(b^n)} \leq \operatorname{poly}(n) \int \mathrm{F}\Big(\frac{P_{\mathcal{B}^n}(b^n)}{\operatorname{tr} P_{\mathcal{B}^n}(b^n)}, \rho_{\mathcal{B}}^{\wedge n}(b^n)\Big)^2 \rho_{\mathcal{B}}^{\wedge n}(b^n)\, \mathrm{d}\rho_{\mathcal{B}}\,, \tag{5.4}$$

where $\rho_{\mathcal{B}}^{\wedge n}(b^n) := \rho_{\mathcal{B}}(b^{(1)}) \otimes \cdots \otimes \rho_{\mathcal{B}}(b^{(n)})$.

Proof. We first consider the pure state case that $P_{\mathcal{B}^n}(b^n)$ is a pure state $\psi_{\mathcal{B}^n}(b^n)$ on $(\mathbb{C}^{d^2})^{\otimes n}$ for each b^n. Let $\mathrm{d}\varphi_{\mathcal{B}} = \mathrm{d}\varphi_1 \cdots \mathrm{d}\varphi_m$, where $\mathrm{d}\varphi_i$ is the uniform spherical measure on $S_{\mathbb{C}^d}$ for each i. For each b^n, There exists $\pi \in S_n$ such that $\pi b^n = 1^{q_1} 2^{q_2} \cdots m^{q_m}$ for some integers q_i. Since the inequality (5.4) is permutation invariant in the sense that it is implied by $\psi_{\mathcal{B}^n}(\sigma b^n) \leq \operatorname{poly}(n) \int \langle \psi_{\mathcal{B}^n}(\sigma b^n), \varphi_{\mathcal{B}}^{\wedge n}(\sigma b^n)\rangle \varphi_{\mathcal{B}}^{\wedge n}(\sigma b^n)\, \mathrm{d}\varphi_{\mathcal{B}}$ for $\sigma \in S_n$, it suffices to prove the case that $b^n = 1^{q_1} 2^{q_2} \cdots m^{q_m}$. It holds that

$$\psi_{\mathcal{B}^n}(b^n) = (\Pi_1 \otimes \cdots \otimes \Pi_m) \psi_{\mathcal{B}^n}(b^n) (\Pi_1 \otimes \cdots \otimes \Pi_m)\,, \tag{5.5}$$

where Π_i is the projector onto the symmetric subspace $\vee^{q_i} \mathbb{C}^d$. Denote $\beta_{d,N} := \binom{d-1+N}{d-1}$ for any positive integers d and N. Note that $\operatorname{Herm}\big(\vee^{q_i} \mathbb{C}^d\big) \cong \mathbb{R}^{\beta_{d,q_i}^2}$, and thus that $\{|\psi\rangle\langle\psi|^{\otimes q_i} : |\psi\rangle \in S_{\mathbb{C}^d}\}$, subject to the normalization condition, is contained in a real affine space of dimension $\beta_{d,q_i}^2 - 1$. Due to Caratheodory's theorem, for each i there exists an ensemble $\{(p_{k_i}, |\varphi_{k_i}\rangle\langle\varphi_{k_i}|)\}_{k_i}$ such that

$$\int_{|\varphi\rangle \in S_{\mathbb{C}^d}} |\varphi\rangle\langle\varphi|^{\otimes q_i}\, \mathrm{d}\varphi = \sum_{k_i=1}^{\beta_{d,q_i}^2} c_{k_i} |\varphi_{k_i}\rangle\langle\varphi_{k_i}|^{\otimes q_i}\,, \tag{5.6}$$

where $\mathrm{d}\varphi$ stands for the uniform probability measure on the unit sphere $\mathcal{S}(\mathbb{C}^d)$. Thus since $\Pi_i = \beta_{d,q_i} \int_{|\varphi\rangle \in S_{\mathbb{C}^d}} |\varphi\rangle\langle\varphi|^{\otimes q_i}\, \mathrm{d}\varphi$,

$$\Pi_1 \otimes \cdots \otimes \Pi_m = \beta_{d,q_1} \cdots \beta_{d,q_m} \sum_{k_1,\ldots,k_m} c_{k_1} \cdots c_{k_m} \varphi_{k_1}^{\otimes q_1} \otimes \cdots \otimes \varphi_{k_m}^{\otimes q_m} \tag{5.7}$$

where k_i ranges over $\big[\beta_{d,q_i}^2\big]$.

It follows from Eqs. (5.5) and (5.7) that

$$\begin{aligned}\psi_{\mathcal{B}^n}(b^n) &\le \beta_{d,q_1}^4\cdots\beta_{d,q_m}^4 \sum_{k_1,\dots,k_m} c_{k_1}^2\cdots c_{k_m}^2 \left\langle \psi_{\mathcal{B}^n}(b^n), \bigotimes_i \varphi_{k_i}^{\otimes q_i}\right\rangle \bigotimes_i \varphi_{k_i}^{\otimes q_i} \\ &\le \beta_{d,q_1}^3\cdots\beta_{d,q_m}^3 \sum_{k_1,\dots,k_m} c_{k_1}\cdots c_{k_m} \left\langle \psi_{\mathcal{B}^n}(b^n), \bigotimes_i \varphi_{k_i}^{\otimes q_i}\right\rangle \bigotimes_i \varphi_{k_i}^{\otimes q_i}, \end{aligned} \tag{5.8}$$

where the first inequality is due to that $\sum_{i,j=1}^r M_i X M_j^\dagger \le r\sum_{i=1}^r M_i X M_i^\dagger$ for $X \ge 0$ [LW17], and the second follows from that $p_{k_i} \le \beta_{d,q_i}^{-1}$.

Notice that $\beta_{d,q_1}^3\cdots\beta_{d,q_m}^3 = \mathrm{poly}(n)$ for fixed d and m. Since Eq. (5.6) holds when all $|\varphi_{k_i}\rangle$ are replaced by $U|\varphi_{k_i}\rangle$ for any unitary U, one can integrate Eq. (5.8) with respect to all φ_{k_i} to obtain

$$\psi_{\mathcal{B}^n}(b^n) \le \mathrm{poly}(n)\int \left\langle \psi_{\mathcal{B}^n}(b^n), \varphi_{\mathcal{B}}^{\wedge n}(b^n)\right\rangle \varphi_{\mathcal{B}}^{\wedge n}(b^n)\,\mathrm{d}\varphi_{\mathcal{B}}\,. \tag{5.9}$$

We now turn to the general case that $P_{\mathcal{B}^n}(b^n)$ is positive semidefinite for each b^n. Using the operator-vector correspondence (1.1), any permutation invariant operator acting on $\mathcal{H}_1\otimes\cdots\otimes\mathcal{H}_n$ has a symmetric purification on $\mathcal{H}_1\otimes\mathcal{K}_1\otimes\cdots\otimes\mathcal{H}_n\otimes\mathcal{K}_n$. Let $\psi_{\mathcal{B}^n}(b^n)$ be such purification of $P_{\mathcal{B}^n}(b^n)$ for each b^n. Then the desired result follows. □

Lemma 33 ([AFR15])**.** *Let $\mathcal{X} = [m], \mathcal{A} = [l]$, and let $P_{\mathcal{A}^n|\mathcal{X}^n} : \mathcal{A}^n \times \mathcal{X}^n \to [0,1]$ be some permutation invariant conditional distribution in the sense that $P(a^n|x^n) = P(\pi a^n|\pi x^n)$ for any $x^n \in \mathcal{X}^n, a^n \in \mathcal{A}^n$ and $\pi \in S_n$. Then there exists some measure $\mathrm{d}Q_{\mathcal{A}|\mathcal{X}}$ on $[0,1]^{\mathcal{A}\times\mathcal{X}}$ such that*

$$P_{\mathcal{A}^n|\mathcal{X}^n} \le (n+1)^{m(l-1)}\int Q_{\mathcal{A}|\mathcal{X}}^{\wedge n}\,\mathrm{d}Q_{\mathcal{A}|\mathcal{X}}\,. \tag{5.10}$$

Lemma 34. *Let $\mathcal{Y} = [m], \mathcal{B} = [l]$, and let $P_{\mathcal{B}^n|\mathcal{Y}^n} : \mathcal{B}^n \times \mathcal{Y}^n \to \mathrm{Pos}\big((\mathbb{C}^d)^{\otimes n}\big)$ be an operator assemblage satisfying $P(\pi b^n|\pi y^n) = \pi\cdot P(b^n|y^n)$ for any $\pi \in S_n$. Then there exists a measure $\mathrm{d}Q_{\mathcal{B}|\mathcal{Y}}$ on $(\mathbb{C}^d)^{\mathcal{B}\times\mathcal{Y}}$ such that*

$$P_{\mathcal{B}^n|\mathcal{Y}^n} \le \mathrm{poly}(n)\int Q_{\mathcal{B}|\mathcal{Y}}^{\wedge n}\,\mathrm{d}Q_{\mathcal{B}|\mathcal{Y}}\,. \tag{5.11}$$

Proof. Let $\mathrm{d}Q_{\mathcal{B}|\mathcal{Y}} = \frac{\mathrm{d}p_{1,1}}{c_{1,1}}\cdots\frac{\mathrm{d}p_{1,l-1}}{c_{1,l-1}}\cdots\frac{\mathrm{d}p_{m,l-1}}{c_{m,l-1}}\,\mathrm{d}\varphi_{1,1}\cdots\mathrm{d}\varphi_{1,l}\cdots\mathrm{d}\varphi_{m,l}$ where $\mathrm{d}\varphi_{y,b}$ is the uniform spherical measure. Thus

$$\begin{aligned}\int Q_{\mathcal{B}|\mathcal{Y}}^{\wedge n}(b^n|y^n)\,\mathrm{d}Q_{\mathcal{B}|\mathcal{Y}} = &\int_0^{c_{1,1}} \frac{\mathrm{d}p_{1,1}}{c_{1,1}} \cdots \int_0^{c_{m,l-1}} \frac{\mathrm{d}p_{m,l-1}}{c_{m,l-1}} p_{1,1}^{q_{1,1}} \cdots p_{1,l}^{q_{1,l}} \cdots p_{m,l}^{q_{m,l}} \\ &\cdot \frac{\Pi_{q_{1,1}}}{\beta_{d,q_{1,1}}} \otimes \cdots \otimes \frac{\Pi_{q_{m,l}}}{\beta_{d,q_{m,l}}}\,,\end{aligned} \tag{5.12}$$

where $p_{y,b}, q_{y,b}, c_{y,b}$ are defined in the same way as in Lemma 33.

Since $P(\pi b^n|\pi y^n) = \pi \cdot P(b^n|y^n)$, one has $\operatorname{tr} P(b^n|y^n) = \operatorname{tr} P(\pi b^n|\pi y^n)$. Due to Lemma 33,

$$\operatorname{tr} P(b^n|y^n) \le (n+1)^{m(l-1)} \int_0^{c_{1,1}} \frac{\mathrm{d}p_{1,1}}{c_{1,1}} \cdots \int_0^{c_{m,l-1}} \frac{\mathrm{d}p_{m,l-1}}{c_{m,l-1}} p_{1,1}^{q_{1,1}} \cdots p_{1,l}^{q_{1,l}} \cdots p_{m,l}^{q_{m,l}}\,. \tag{5.13}$$

Due to Lemma 31, we have

$$\frac{P(b^n|y^n)}{\operatorname{tr} P(b^n|y^n)} \le \Pi_{q_{1,1}} \otimes \cdots \otimes \Pi_{q_{m,l}}\,. \tag{5.14}$$

Taking together (5.12), (5.13) and (5.14), and using that $\beta_{d,q_{1,1}} \cdots \beta_{d,q_{m,l}} = \mathrm{poly}(n)$, we obtain

$$P_{\mathcal{B}^n|\mathcal{Y}^n} \le \mathrm{poly}(n) \int Q_{\mathcal{B}|\mathcal{Y}}^{\wedge n}\,\mathrm{d}Q_{\mathcal{B}|\mathcal{Y}}\,.$$

□

5.3 Parallel repetition for nonsignaling strategy

In this section we show that for n-fold parallel repetition of an extended nonlocal game G using commuting nonsignaling strategy, the game value $\mathrm{val}(G^n)$ vanishes exponentially as n grows, given that the one-shot game value $\mathrm{val}(G)$ is strictly less than 1. By definition,

$$\mathrm{val}(G^n) = \max_{Q_{\mathcal{A}^n|\mathcal{X}^n} \in \mathrm{NS}} \sum_{a^n, x^n} \mu_{\mathcal{X}}^{\wedge n}(x^n) \big\langle Q_{\mathcal{A}^n|\mathcal{X}^n}(a^n|x^n), V_{\mathcal{A}\mathcal{X}}^{\wedge n}(a^n, x^n)\big\rangle\,. \tag{5.15}$$

We now study the symmetric property in this expression and shall exploit this symmetry in later analysis. For $\pi \in S_n$, $a^n := \big(a^{(1)}, \ldots, a^{(n)}\big) \in \mathcal{A}^n$ and $x^n := \big(x^{(1)}, \ldots, x^{(n)}\big) \in \mathcal{X}^n$, we write $\pi a^n := \pi \cdot a^n := \big(a^{(1')}, \ldots, a^{(n')}\big)$ where $k' := \pi^{-1}(k)$, and write $\pi \cdot (a^n|x^n) := (\pi a^n|\pi x^n)$. We also write $\pi \cdot X := W_\pi X W_\pi^\dagger$ where W_π is the operator permuting the systems according to $\pi \in S_n$.

Notice that $V^{\wedge n}(a^n, x^n) = V(a^{(1)}, x^{(1)}) \otimes \cdots \otimes V(a^{(n)}, x^{(n)})$. For any $\pi \in S_n$ we have that $\pi \cdot V^{\wedge n} = V^{\wedge n}$, and thus

$$\begin{aligned}
&\sum_{a^n, x^n} \mu^{\wedge n}(x^n) \langle Q(a^n|x^n), V^{\wedge n}(a^n, x^n)\rangle \\
&\quad = \sum_{a^n, x^n} \mu^{\wedge n}(x^n) \langle Q(\pi a^n|\pi x^n), V^{\wedge n}(\pi a^n, \pi x^n)\rangle \\
&\quad = \sum_{a^n, x^n} \mu^{\wedge n}(x^n) \langle Q(\pi a^n|\pi x^n), \pi \cdot V^{\wedge n}(a^n, x^n)\rangle \\
&\quad = \sum_{a^n, x^n} \mu^{\wedge n}(x^n) \big\langle (\pi^{-1} \cdot Q \cdot \pi)(a^n|x^n), V^{\wedge n}(a^n, x^n)\big\rangle \,.
\end{aligned}$$

Thus the strategies Q and $\pi^{-1} \cdot Q \cdot \pi$ give the same game value for any $\pi \in S_n$. Therefore by using the symmetrization $\frac{1}{n!}\sum_{\pi \in S_n} \pi^{-1} \cdot Q \cdot \pi$, we can assume w.l.o.g. that the strategy Q in (5.15) satisfies

$$Q = \pi^{-1} \cdot Q \cdot \pi \tag{5.16}$$

for any $\pi \in S_n$.

If $a^{(i)} = a^{(j)}$ and $x^{(i)} = x^{(j)}$ for some $i \neq j$, it follows from (5.16) that $W_{(ij)} Q(a^n|x^n) W_{(ij)}^\dagger = Q(a^n|x^n)$ for this case, where $W_{(ij)}$ is the operator swapping the i-th and j-th systems. Thus, for any a^n, x^n, there exists $\sigma \in S_n$ and integers $q_{i,j}$ with $i \in [|\mathcal{A}|]$ and $j \in [|\mathcal{X}|]$, such that $(\pi_{1,1} \otimes \cdots \otimes \pi_{|\mathcal{A}|,|\mathcal{X}|}) \cdot \sigma \cdot Q(a^n|x^n) = \sigma \cdot Q(a^n|x^n)$ holds for any $\pi_{i,j} \in S_{q_{i,j}}$.

Theorem 35. *Let G be a two-player game such that $\mathrm{val}(G) \leq 1-\delta$ for some $0 < \delta < 1$. Then for any positive integer n, we have*

$$\mathrm{val}(G^n) \leq \left(1 - c\delta^2\right)^n \tag{5.17}$$

for some constant c.

Proof. Due to Lemma 32, by noticing that $\mathrm{F}(X, |u\rangle\langle u|)^2 = \langle X, |u\rangle\langle u|\rangle$ for $X \geq 0$, one has for any a^n and x^n,

$$\mu_{\mathcal{X}}^{\wedge n}(x^n) Q_{\mathcal{A}^n|\mathcal{X}^n}(a^n|x^n) \leq \mathrm{poly}(n) \int \mathrm{F}\Big(\mu_{\mathcal{X}}^{\wedge n}(x^n) Q_{\mathcal{A}^n|\mathcal{X}^n}(a^n|x^n), R_{\mathcal{AX}}^{\wedge n}(a^n, x^n)\Big)^2 R_{\mathcal{AX}}^{\wedge n}(a^n, x^n)\, \mathrm{d}R_{\mathcal{AX}} \,. \tag{5.18}$$

The fidelity F in the expression above can be upper-bounded using Lemma 30 as

$$\begin{aligned}
&\mathrm{F}\Big(\mu_{\mathcal{X}}^{\wedge n}(x^n)Q_{\mathcal{A}^n|\mathcal{X}^n}(a^n|x^n), R_{\mathcal{A}\mathcal{X}}^{\wedge n}(a^n,x^n)\Big)\\
&\quad\le \mathrm{F}\big(\mu_{\mathcal{X}}^{\wedge n}Q_{\mathcal{A}^n|\mathcal{X}^n}, R_{\mathcal{A}\mathcal{X}}^{\wedge n}\big)\\
&\quad\le \mathrm{F}\big(\mu_{\mathcal{X}}^{\wedge n}Q_{\mathcal{A}_i^n|\mathcal{X}^n}, R_{\mathcal{A}_i\mathcal{X}}^{\wedge n}\big)\\
&\quad= \mathrm{F}\big(\mu_{\mathcal{X}}^{\wedge n}Q'_{\mathcal{A}_i^n|\mathcal{X}_i^n}, R_{\mathcal{A}_i\mathcal{X}}^{\wedge n}\big)\\
&\quad\le \mathrm{poly}(n)\mathrm{F}\Big(\mu_{\mathcal{X}}^{\wedge n}\int T_{\mathcal{A}_i|\mathcal{X}_i}^{\wedge n}\,\mathrm{d}T_{\mathcal{A}_i|\mathcal{X}_i}, R_{\mathcal{A}_i\mathcal{X}}^{\wedge n}\Big)\\
&\quad\le \mathrm{poly}(n)\max_{T_{\mathcal{A}_i|\mathcal{X}_i}}\mathrm{F}(\mu_{\mathcal{X}}T_{\mathcal{A}_i|\mathcal{X}_i}, R_{\mathcal{A}_i\mathcal{X}})^n\,,
\end{aligned}\tag{5.19}$$

where the second line uses the additivity of fidelity under direct sum, the third uses the superadditivity of fidelity under addition, the fourth uses definition of the nonsignaling strategy, the fifth uses Lemma 34, and the sixth uses the multiplicativity of fidelity under tensor product.

Since the inequality (5.19) holds for each $i=1,2$, denoting

$$\tilde{\mathrm{F}}(R_{\mathcal{A}\mathcal{X}}) := \min_i \max_{T_{\mathcal{A}_i|\mathcal{X}_i}} \mathrm{F}\big(\mu_{\mathcal{X}}T_{\mathcal{A}_i|\mathcal{X}_i}, R_{\mathcal{A}_i\mathcal{X}}\big)\,.$$

we have that

$$\mathrm{F}\Big(\mu_{\mathcal{X}}^{\wedge n}(x^n)Q_{\mathcal{A}^n|\mathcal{X}^n}(a^n|x^n), R_{\mathcal{A}\mathcal{X}}^{\wedge n}(a^n,x^n)\Big) \le \mathrm{poly}(n)\tilde{\mathrm{F}}(R_{\mathcal{A}\mathcal{X}})^n\,. \tag{5.20}$$

Inserting (5.20) into (5.18), we have

$$\mu_{\mathcal{X}}^{\wedge n}(x^n)Q_{\mathcal{A}^n|\mathcal{X}^n}(a^n|x^n) \le \mathrm{poly}(n)\int \tilde{\mathrm{F}}(R_{\mathcal{A}\mathcal{X}})^{2n}R_{\mathcal{A}\mathcal{X}}^{\wedge n}(a^n,x^n)\,\mathrm{d}R_{\mathcal{A}\mathcal{X}}\,. \tag{5.21}$$

Denoting $\mathcal{Q}_\varepsilon := \big\{R_{\mathcal{A}\mathcal{X}} : \tilde{\mathrm{F}}(R_{\mathcal{A}\mathcal{X}}) \ge 1-\varepsilon\big\}$, we have

$$\begin{aligned}
&\mu_{\mathcal{X}}^{\wedge n}(x^n)Q_{\mathcal{A}^n|\mathcal{X}^n}(a^n|x^n)\\
&\le \mathrm{poly}(n)\Big(\int_{R\in\mathcal{Q}_\varepsilon}\tilde{\mathrm{F}}(R_{\mathcal{A}\mathcal{X}})^{2n}R_{\mathcal{A}\mathcal{X}}^{\wedge n}(a^n,x^n)\,\mathrm{d}R_{\mathcal{A}\mathcal{X}} + \int_{R\notin\mathcal{Q}_\varepsilon}\tilde{\mathrm{F}}(R_{\mathcal{A}\mathcal{X}})^{2n}R_{\mathcal{A}\mathcal{X}}^{\wedge n}(a^n,x^n)\,\mathrm{d}R_{\mathcal{A}\mathcal{X}}\Big)\\
&\le \mathrm{poly}(n)\Big(\int_{R\in\mathcal{Q}_\varepsilon}R_{\mathcal{A}\mathcal{X}}^{\wedge n}(a^n,x^n)\,\mathrm{d}R_{\mathcal{A}\mathcal{X}} + (1-\varepsilon)^{2n}\int_{R\notin\mathcal{Q}_\varepsilon}R_{\mathcal{A}\mathcal{X}}^{\wedge n}(a^n,x^n)\,\mathrm{d}R_{\mathcal{A}\mathcal{X}}\Big)\,.
\end{aligned}$$

For $R_{\mathcal{A}\mathcal{X}} \in \mathcal{Q}_\varepsilon$, applying the Fuchs-van de Graaf inequality, we have

$$\max_i \min_{T_{\mathcal{A}_i|\mathcal{X}_i}} \mathrm{D}(\mu_{\mathcal{X}} T_{\mathcal{A}_i|\mathcal{X}_i}, R_{\mathcal{A}_i\mathcal{X}}) \leq \sqrt{2\varepsilon}\,. \tag{5.22}$$

By the property of the commuting nonsignaling strategies, there exists a nonsignaling operator assemblage $K_{\mathcal{A}|\mathcal{X}}$ such that

$$\mathrm{D}(\mu_{\mathcal{X}} K_{\mathcal{A}|\mathcal{X}}, R_{\mathcal{A}\mathcal{X}}) \leq C\sqrt{\varepsilon} \tag{5.23}$$

for some constant C. From the hypothesis $\mathrm{val}(G) \leq 1-\delta$ we get that

$$\sum_{a,x} \mu(x)\langle P(a|x), V(a,x)\rangle \leq 1-\delta \tag{5.24}$$

for any nonsignaling $P_{\mathcal{A}|\mathcal{X}}$. Combining Eqs. (5.23) and (5.24), we have

$$\sum_{a,x} \langle R(a|x), V(a,x)\rangle \leq 1-\delta+C\sqrt{\varepsilon}\,,$$

since $\|V(a,x)\|_\infty \leq 1$ for any a and x. It follows that

$$\sum_{a^n,x^n} \langle R^{\wedge n}(a^n|x^n), V^{\wedge n}(a^n,x^n)\rangle \leq (1-\delta+C\sqrt{\varepsilon})^n\,.$$

Notice that for $R_{\mathcal{A}\mathcal{X}} \notin \mathcal{Q}_\varepsilon$, $\sum_{a^n,x^n}\langle R^{\wedge n}(a^n|x^n), V^{\wedge n}(a^n,x^n)\rangle \leq 1$. Thus the winning probability $\mathrm{val}(G^n)$ is upper bounded by $\mathrm{poly}(n)\big((1-\delta+C\sqrt{\varepsilon})^n + (1-\varepsilon)^n\big)$. Taking $\varepsilon = (\frac{\delta}{2C})^2$, one has

$$\mathrm{val}(G^n) \leq \mathrm{poly}(n)\big(1-c\delta^2\big)^n \tag{5.25}$$

for some constant c. Notice that $\mathrm{val}(G^n) \geq \mathrm{val}(G)^n$. Using the fact that if a non-negative series t_k satisfy that $t_{mn} \geq t_n^m$ for any positive integers m and n, and if $t_n \leq \mathrm{poly}(n)$, then $t_n \leq 1$ for any n, one can remove the prefactor $\mathrm{poly}(n)$ in (5.25) to conclude the proof. $\square$

Bibliography

[ADH15] A. Almheiri, X. Dong, and D. Harlow. Bulk locality and quantum error correction in ads/cft. *Journal of High Energy Physics*, 2015(4):163, 2015.

[AFR15] R. Arnon-Friedman and R. Renner. de Finetti reductions for correlations. *Journal of Mathematical Physics*, 56(5):052203, 2015.

[AFRV16] R. Arnon-Friedman, R. Renner, and T. Vidick. Non-signaling parallel repetition using de Finetti reductions. *IEEE Transactions on Information Theory*, 62(3):1440–1457, 2016.

[AKHvD11] A. Alex, M. Kalus, A. Huckleberry, and J. von Delft. A numerical algorithm for the explicit calculation of $\mathrm{SU}(N)$ and $\mathrm{SL}(N,\mathbb{C})$ Clebsch-Gordan coefficients. *Journal of Mathematical Physics*, 52(2):023507, 2011.

[AL04] A. Ashtekar and J. Lewandowski. Background independent quantum gravity: A status report. *Classical and Quantum Gravity*, 21(15):R53, 2004.

[AS17] G. Aubrun and S. J. Szarek. *Alice and Bob meet Banach: the interface of asymptotic geometric analysis and quantum information theory*, volume 223. American Mathematical Society, 2017.

[ASW11] G. Aubrun, S. Szarek, and E. Werner. Hastings's additivity counterexample via Dvoretzky's theorem. *Communications in Mathematical Physics*, 305(1):85–97, 2011.

[Bab85] L. Babai. Trading group theory for randomness. In *Proceedings of the 17th annual ACM Symposium on Theory of Computing*, pages 421–429, 1985.

[BB84] C. H. Bennett and G. Brassard. Quantum cryptography. In *Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing*, pages 175–179, 1984.

[BBC+93] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Physical Review Letters*, 70(13):1895, 1993.

[BCH06] D. Bacon, I. L. Chuang, and A. W. Harrow. Efficient quantum circuits for Schur and Clebsch-Gordan transforms. *Physical Review Letters*, 97(17):170502, 2006.

[BCWdW01] H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf. Quantum fingerprinting. *Physical Review Letters*, 87(16):167902, 2001.

[Bel64] J. S. Bell. On the Einstein-Podolsky-Rosen paradox. *Physics*, 1(3):195–200, 1964.
Reprinted in J. S. Bell, *Speakable and unspeakable in quantum mechanics*, Cambridge University Press, Cambridge, 1987.

[Ben80] P. Benioff. The computer as a physical system: a microscopic quantum mechanical Hamiltonian model of computers as represented by Turing machines. *Journal of Statistical Physics*, 22(5):563–591, 1980.

[BFS14] H. Buhrman, S. Fehr, and C. Schaffner. On the parallel repetition of multi-player games: the no-signaling case. *Leibniz International Proceedings in Informatics*, 27:24–35, 2014.

[BG16] M. Braverman and A. Garg. Small value parallel repetition for general games. In *Proceedings of the 47th Annual ACM Symposium on Theory of Computing*, pages 335–340, 2016.

[BH10] F. G. Brandão and M. Horodecki. On Hastings' counterexamples to the minimum output entropy additivity conjecture. *Open Systems & Information Dynamics*, 17(01):31–52, 2010.

[BOW19] C. Bădescu, R. O'Donnell, and J. Wright. Quantum state certification. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, pages 503–514. ACM, 2019.

[BW92] C. H. Bennett and S. J. Wiesner. Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states. *Physical Review Letters*, 69(20):2881, 1992.

[CDS05] G. Chiribella, G. D'Ariano, and M. Sacchi. Optimal estimation of group transformations using entanglement. *Physical Review A*, 72(4):042338, 2005.

[Cho75] M.-D. Choi. Completely positive linear maps on complex matrices. *Linear Algebra and its Applications*, 10(3):285–290, 1975.

[Chr06] M. Christandl. *The structure of bipartite quantum states – insights from group theory and cryptography.* PhD thesis, University of Cambridge, 2006.

[CHSH69] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt. Proposed experiment to test local hidden-variable theories. *Physical Review Letters*, 23(15):880, 1969.

[CM06] M. Christandl and G. Mitchison. The spectra of quantum states and the Kronecker coefficients of the symmetric group. *Communications in Mathematical Physics*, 261(3):789–797, 2006.

[CN10] B. Collins and I. Nechita. Random quantum channels I: graphical calculus and the Bell state phenomenon. *Communications in Mathematical Physics*, 297(2):345–370, 2010.

[CŚ06] B. Collins and P. Śniady. Integration with respect to the Haar measure on unitary, orthogonal and symplectic group. *Communications in Mathematical Physics*, 264(3):773–795, 2006.

[CSA+15] D. Cavalcanti, P. Skrzypczyk, G. Aguilar, R. Nery, P. S. Ribeiro, and S. Walborn. Detection of entanglement in asymmetric quantum networks and multipartite quantum steering. *Nature Communications*, 6(1):1–6, 2015.

[CT12] T. M. Cover and J. A. Thomas. *Elements of information theory.* John Wiley & Sons, 2012.

[Deu85] D. Deutsch. Quantum theory, the Church-Turing principle and the universal quantum computer. *Proceedings of the Royal Society of London A*, 400(1818):97–117, 1985.

[DJ92] D. Deutsch and R. Jozsa. Rapid solution of problems by quantum computation. *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, 439(1907):553–558, 1992.

[dSLCP11] M. P. da Silva, O. Landon-Cardinal, and D. Poulin. Practical characterization of quantum devices without tomography. *Physical Review Letters*, 107(21):210404, 2011.

[DSW16] R. Duan, S. Severini, and A. Winter. On zero-error communication via quantum channels in the presence of noiseless feedback. *IEEE Transactions on Information Theory*, 62(9):5260–5277, 2016.

[Dvo61] A. Dvoretzky. Some results on convex bodies and Banach spaces. *Proc. Internat. Sympos. Linear Spaces (Jerusalem, 1960)*, pages 123–160, 1961.

[EGH+11] P. I. Etingof, O. Golberg, S. Hensel, T. Liu, A. Schwendner, D. Vaintrob, and E. Yudovina. *Introduction to representation theory*, volume 59. American Mathematical Society Providence, RI, 2011.

[EPR35] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Physical Review*, 47(10):777, 1935.

[Fey82] R. P. Feynman. Simulating physics with computers. *International Journal of Theoretical Physics*, 21(6-7):467–488, 1982.

[FGM01] M. Fitzi, N. Gisin, and U. Maurer. Quantum solution to the Byzantine agreement problem. *Physical Review Letters*, 87(21):217901, 2001.

[FH13] W. Fulton and J. Harris. *Representation theory: a first course*, volume 129. Springer Science & Business Media, 2013.

[FKM10] M. Fukuda, C. King, and D. K. Moser. Comments on Hastings' additivity counterexamples. *Communications in Mathematical Physics*, 296(1):111–143, 2010.

[FN14] M. Fukuda and I. Nechita. Asymptotically well-behaved input states do not violate additivity for conjugate pairs of random quantum channels. *Communications in Mathematical Physics*, 328(3):995–1021, 2014.

[FN17] M. Fukuda and I. Nechita. On the minimum output entropy of random orthogonal quantum channels. *IEEE Transactions on Information Theory*, 64(2):1374–1384, 2017.

[Fri12] T. Fritz. Tsirelson's problem and kirchberg's conjecture. *Reviews in Mathematical Physics*, 24(05):1250012, 2012.

[FVDG99] C. A. Fuchs and J. Van De Graaf. Cryptographic distinguishability measures for quantum-mechanical states. *IEEE Transactions on Information Theory*, 45(4):1216–1227, 1999.

[FW07] M. Fukuda and M. M. Wolf. Simplifying additivity problems using direct sum constructions. *Journal of mathematical physics*, 48(7):072101, 2007.

[GHP10] A. Grudka, M. Horodecki, and Ł. Pankowski. Constructive counterexamples to the additivity of the minimum output Rényi entropy of quantum channels for all $p > 2$. *Journal of Physics A: Mathematical and Theoretical*, 43(42):425304, 2010.

[GMR85] S. Goldwasser, S. Micali, and C. Rackoff. The knowledge complexity of interactive proof systems. In *Proceedings of the 17th annual ACM Symposium on Theory of Computing*, pages 291–304, 1985.

[Gro56] A. Grothendieck. Sur certaines classes de suites dans les espaces de banach et le théorème de Dvoretzky-Rogers. *Bol. Soc. Mat. Sao Paulo*, 8(81-110):1953, 1956.

[Gro96] L. K. Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the 28th Annual ACM Symposium on Theory of Computing*, pages 212–219. ACM, 1996.

[GT50] I. M. Gelfand and M. L. Tsetlin. Finite-dimensional representations of the group of unimodular matrices. In *Dokl. Akad. Nauk Ser. Fiz.*, pages 825–828, 1950.

[GW09] R. Goodman and N. R. Wallach. *Symmetry, representations, and invariants*, volume 255. Springer, 2009.

[Hal15] B. Hall. *Lie groups, Lie algebras, and representations: an elementary introduction*, volume 222. Springer, 2015.

[Har05] A. W. Harrow. *Applications of coherent classical communication and the Schur transform to quantum information theory.* PhD thesis, MIT, 2005.

[Has09] M. B. Hastings. Superadditivity of communication capacity using entangled inputs. *Nature Physics*, 5(4):255–257, 2009.

[Hay02] M. Hayashi. Exponents of quantum fixed-length pure-state source coding. *Physical Review A*, 66(3):032321, 2002.

[Hay07] P. Hayden. The maximal p-norm multiplicativity conjecture is false. *arXiv preprint arXiv:0707.3291*, 2007.

[Hay17] M. Hayashi. *A group theoretic approach to quantum information.* Springer, 2017.

[HCL$^+$12] W. Helwig, W. Cui, J. I. Latorre, A. Riera, and H.-K. Lo. Absolute maximal entanglement and quantum secret sharing. *Physical Review A*, 86(5):052335, 2012.

[Hel13] W. Helwig. Absolutely maximally entangled qudit graph states. *arXiv preprint arXiv:1306.2879*, 2013.

[HK69] K.-E. Hellwig and K. Kraus. Pure operations and measurements. *Communications in Mathematical Physics*, 11(3):214–220, 1969.

[HK70] K.-E. Hellwig and K. Kraus. Operations and measurements II. *Communications in Mathematical Physics*, 16(2):142–147, 1970.

[HLSW04] P. Hayden, D. Leung, P. W. Shor, and A. Winter. Randomizing quantum states: constructions and applications. *Communications in Mathematical Physics*, 250(2):371–391, 2004.

[HLW06] P. Hayden, D. W. Leung, and A. Winter. Aspects of generic entanglement. *Communications in Mathematical Physics*, 265(1):95–117, 2006.

[HM13] A. W. Harrow and A. Montanaro. Testing product states, quantum Merlin-Arthur games and tensor optimization. *Journal of the ACM*, 60(1):3, 2013.

[HMH07] M. Han, Y. Ma, and W. Huang. Fundamental structure of loop quantum gravity. *International Journal of Modern Physics D*, 16(09):1397–1474, 2007.

[HNQ$^+$16] P. Hayden, S. Nezami, X.-L. Qi, N. Thomas, M. Walter, and Z. Yang. Holographic duality from random tensor networks. *Journal of High Energy Physics*, 2016(11):9, 2016.

[Hol73] A. S. Holevo. Statistical decision theory for quantum systems. *Journal of Multivariate Analysis*, 3(4):337–394, 1973.

[Hol98] A. S. Holevo. The capacity of the quantum channel with general signal states. *IEEE Transactions on Information Theory*, 44(1):269–273, 1998.

[Hol09] T. Holenstein. Parallel repetition: simplification and the no-signaling case. *Theory of Computing*, 5(1):141–172, 2009.

[HQRY16] P. Hosur, X.-L. Qi, D. A. Roberts, and B. Yoshida. Chaos in quantum channels. *Journal of High Energy Physics*, 2016(2):4, 2016.

[HW08] P. Hayden and A. Winter. Counterexamples to the maximal p-norm multiplicativity conjecture for all $p > 1$. *Communications in Mathematical Physics*, 284(1):263–280, 2008.

[Jam72] A. Jamiołkowski. Linear transformations which preserve trace and positive semidefiniteness of operators. *Reports on Mathematical Physics*, 3(4):275–278, 1972.

[Jam06] G. D. James. *The representation theory of the symmetric group*, volume 682. Springer, 2006.

[JMRW16] N. Johnston, R. Mittal, V. Russo, and J. Watrous. Extended non-local games and monogamy-of-entanglement games. *Proceedings of the Royal Society A*, 472(2189):20160003, 2016.

[Kin02] C. King. Additivity for unital qubit channels. *Journal of Mathematical Physics*, 43(10):4641–4653, 2002.

[Kin03] C. King. The capacity of the quantum depolarizing channel. *IEEE Transactions on Information Theory*, 49(1):221–229, 2003.

[Lan91] R. Landauer. Information is physical. *Physics Today*, 44(5):23–29, 1991.

[Led05] M. Ledoux. *The concentration of measure phenomenon.* Number 89. American Mathematical Society, 2005.

[LHGZ17] Y. Li, M. Han, M. Grassl, and B. Zeng. Invariant perfect tensors. *New Journal of Physics*, 19(6):063029, 2017.

[LHRZ18] Y. Li, M. Han, D. Ruan, and B. Zeng. Random SU(2) invariant tensors. *Journal of Physics A: Mathematical and Theoretical*, 51(17):175303, 2018.

[Low09] R. A. Low. Learning and testing algorithms for the Clifford group. *Physical Review A*, 80(5):052314, 2009.

[LPH51] P. Lévy, F. Pellegrino, and J. Hadamard. *Problèmes concrets d'analyse fonctionnelle*, volume 6. Gauthier-Villars Paris, 1951.

[LW16] C. Lancien and A. Winter. Parallel repetition and concentration for (sub-)no-signalling games via a flexible constrained de Finetti reduction. *Chicago Journal of Theoretical Computer Science*, 2016(11):1–22, 2016.

[LW17] C. Lancien and A. Winter. Flexible constrained de Finetti reductions and applications. *Journal of Mathematical Physics*, 58(9):092203, 2017.

[Man80] Y. I. Manin. *Computable and uncomputable* [in Russian]. Sovetskoye Radio, Moscow, 1980.

[MdW16] A. Montanaro and R. de Wolf. A survey of quantum property testing. *Theory of Computing*, (7):1–81, 2016.

[Mil71] V. D. Milman. A new proof of A. Dvoretzky's theorem on cross-sections of convex bodies. *Funkcional. Anal. i Prilozen*, 5(4):28–37, 1971.

[MKB05] F. Mintert, M. Kuś, and A. Buchleitner. Concurrence of mixed multipartite quantum states. *Physical Review Letters*, 95(26):260502, 2005.

[MO10] A. Montanaro and T. J. Osborne. Quantum boolean functions. *Chicago Journal of Theoretical Computer Science*, 2010(1):1–45, 2010.

[NC11] M. A. Nielsen and I. L. Chuang. *Quantum computation and quantum information*. Cambridge University Press, New York, NY, USA, 10th edition, 2011.

[PYHP15] F. Pastawski, B. Yoshida, D. Harlow, and J. Preskill. Holographic quantum error-correcting codes: toy models for the bulk/boundary correspondence. *Journal of High Energy Physics*, 2015(6):149, 2015.

[Rac42] G. Racah. Theory of complex spectra II. *Physical Review*, 62(9-10):438, 1942.

[Raz98] R. Raz. A parallel repetition theorem. *SIAM Journal on Computing*, 27(3):763–803, 1998.

[RGK13] D. M. Reich, G. Gualdi, and C. P. Koch. Optimal strategies for estimating the average fidelity of quantum gates. *Physical Review Letters*, 111(20):200401, 2013.

[Rus17] V. Russo. *Extended nonlocal games*. PhD thesis, University of Waterloo, 2017.

[RV14] C. Rovelli and F. Vidotto. *Covariant loop quantum gravity: an elementary introduction to quantum gravity and spinfoam theory*. Cambridge University Press, 2014.

[Sch] I. Schur. *Über eine Klasse von Matrizen, die sich einer gegebenen Matrix zuordnen lassen, Inaugural*. PhD thesis, Dissertation, Friedrich-Wilhelms-Universitat zu Berlin 1901.
Reprinted in *Gesammelte Abhandlungen I*, Springer-Verlag, Heidelberg, 1973, pp. 1–72.

[Sch03] J. Schliemann. Entanglement in su(2)-invariant quantum spin systems. *Physical Review A*, 68(1):012309, 2003.

[Sch05] J. Schliemann. Entanglement in su(2)-invariant quantum spin systems: the positive partial transpose criterion and others. *Physical Review A*, 72(1):012307, 2005.

[SdSF$^+$12] L. Steffen, M. P. da Silva, A. Fedorov, M. Baur, and A. Wallraff. Experimental Monte Carlo quantum process certification. *Physical Review Letters*, 108(26):260506, 2012.

[Sha48] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423, 1948.

[Sho94] P. W. Shor. Algorithms for quantum computation: discrete logarithms and factoring. In *Proceedings of the 35th Annual Symposium on Fundamentals of Computer Science*, pages 124–134. IEEE, 1994.

[Sho96] P. W. Shor. Fault-tolerant quantum computation. In *Proceedings of the 37th Annual Symposium on Fundamentals of Computer Science*, pages 56–65. IEEE, 1996.

[Sho99] P. W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Review*, 41(2):303–332, 1999. This is an expanded version of [Sho94].

[Sho02] P. W. Shor. Additivity of the classical capacity of entanglement-breaking quantum channels. *Journal of Mathematical Physics*, 43(9):4334–4340, 2002.

[Sho04] P. W. Shor. Equivalence of additivity questions in quantum information theory. *Communications in Mathematical Physics*, 246(3):453–472, 2004.

[Sla81] R. Slansky. Group theory for unified model building. *Physics reports*, 79(1):1–128, 1981.

[Sti55] W. F. Stinespring. Positive functions on C*-algebras. *Proceedings of the American Mathematical Society*, 6(2):211–216, 1955.

[SW97] B. Schumacher and M. D. Westmoreland. Sending classical information via noisy quantum channels. *Physical Review A*, 56(1):131, 1997.

[TFKW13] M. Tomamichel, S. Fehr, J. Kaniewski, and S. Wehner. A monogamy-of-entanglement game with applications to device-independent quantum cryptography. *New Journal of Physics*, 15(10):103002, 2013.

[Thi08] T. Thiemann. *Modern canonical quantum general relativity.* Cambridge University Press, 2008.

[vN32] J. von Neumann. *Mathematische grundlagen der quantenmechanik.* Springer, 1932.

[Wan11] G. Wang. Property testing of unitary operators. *Physical Review A*, 84(5):052328, 2011.

[Wat18] J. Watrous. *The theory of quantum information.* Cambridge University Press, 2018.

[Wey39] H. Weyl. *The classical groups: their invariants and representations.* Princeton University Press, 1939. Second edition, with supplement, 1946.

[WH02] R. F. Werner and A. S. Holevo. Counterexample to an additivity conjecture for output purity of quantum channels. *Journal of Mathematical Physics*, 43(9):4353–4357, 2002.

[Whe90] J. A. Wheeler. Information, physics, quantum: the search for links. *Complexity, entropy, and the physics of information*, 8, 1990.

[Win07] A. Winter. The maximum output p-norm of quantum channels is not multiplicative for any $p > 2$. *arXiv preprint arXiv:0707.0402*, 2007.