Information-theoretic quantities of quantum channels with partition quantum group symmetries

Master Thesis

Sam van Poelgeest

Delft University of Technology

Information-theoretic quantities of quantum channels with partition quantum group symmetries

by

Sam van Poelgeest

to obtain the degree of

Master of Science in Applied Mathematics & Master of Science in Applied Physics

at the Delft University of Technology

Student number:	4573463	
Project duration:	September, 2021 – July, 2022	
Thesis committee:	Dr. Martijn Caspers,	TU Delft, supervisor
	Dr. David Elkouss Coronas,	TU Delft, supervisor
	Prof. Dr. Dion Gijswijt,	TU Delft
	Dr. Miriam Blaauboer,	TU Delft

Abstract

In the field of *quantum information theory*, it is well-known that the purely quantum phenomenon called *quantum entanglement* can boost the capacity of a quantum channel, which is called the *superadditivity* of the capacity. Shor showed in his breakthrough paper on the equivalence of additivity conjectures that this superadditivity is equivalent to the subadditivity of the *minimum output entropy* of quantum channels, and Hastings gave a high-dimensional and probabilistic counterexample to the minimum output entropy additivity conjecture. Since then, researchers have endeavored to classify which quantum channels have subadditive minimum output entropy and to find deterministically constructed quantum channels with subadditive minimum output entropy. It is well-known that determining the minimum output entropy of an arbitrary quantum channel is a hard problem (in fact, it has been shown to be NP-complete), which in turn makes it difficult to determine whether a quantum channel has subadditive minimum output entropy.

In this thesis, we utilize the representation theory of compact (partition) quantum groups to construct so-called covariant *Clebsch-Gordan quantum channels* and analyze their minimum output entropy. We review the work of Brannan and Collins in the case of the free orthogonal quantum groups O_N^+, and introduce a similar analysis for the quantum permutation group S_N^+. Afterwards, we specialize to the subfamily of *lowest-weight* Clebsch-Gordan channels, and we show that, in the case where one embeds the fundamental representation of O_N^+ in the tensor product of two irreducible representations of O_N^+, the associated lowest-weight quantum channels have sufficient additional structure to analytically compute their minimum output entropy in terms of a recurrence relation.

Lastly, we present an analysis of certain numerical methods that can be utilized to bound the minimum output entropy from below and from above in low dimensions: we use ε-covers for lower bounds, and we use a modified version of the derivative-free optimization method called *Particle Swarm Optimization* over the unit sphere, hybridized with gradient descent, to find upper bounds. We show a proof-of-concept by applying the ε-cover to three S_N^+ channels with small input dimensions, but also note that the exponential scaling of ε-covers with the input dimension makes them intractable with higher input dimensions. We benchmark the Particle Swarm Optimization enriched with gradient-descent on the *highest-weight* Clebsch-Gordan channels for which it can be shown that the minimum output entropy is zero, and we see that the optimization technique performs adequately. We also apply the optimization scheme to the tensor product of two S_N^+-channels, but do not find any violation of the minimum output entropy additivity conjecture.

Acknowledgments

I would like to take this opportunity to express my gratitude for the support and patience of my dear friends and my family, who have supported me tirelessly throughout this journey. In particular, I would like to thank my girlfriend, who has listened to my stories, and offered a helping hand, even when she did not understand a word of what I was saying about compact quantum groups. I would also like to thank Moritz Weber, Alexander Wendel and Adam Skalski for inviting me to their seminars and for our fruitful discussions. To both my supervisors, David and Martijn, thank you for the many meetings, questions and feedback, which I enjoyed greatly. Thank you in particular for facilitating my wish to combine the thesis projects of Applied Physics and Applied Mathematics.

Contents

1

Introduction

Since Feynman's conjecture that quantum computers can be programmed to simulate quantum systems, sizable advancements have been made in the fields of *quantum computation* and *quantum information theory* towards scalable quantum computation and a fully functional quantum internet. Many parties across the globe, both academic and commercial, are working towards this goal. Among them is *QuTech*, a collaborative research institute between The Delft University of Technology and the Netherlands Organisation for Applied Scientifc Research (TNO). Although numerous obstacles remain on the road towards a fully functional quantum internet, it is imperative to consider the theoretical possibilities of future quantum networks now, so that we can utilize these quantum technologies to their theoretical limitations once they are scalable and functional. As these technologies will communicate using quantum information, a *theory of quantum information* is necessary to describe the communication between quantum systems. To develop such a theory, we draw inspiration from the *classical information theory*, where questions such as *at what rate can we send information reliably with negligible risk of losing information?* have been satisfactorily answered: Shannon proved that classical channels can not exceed a certain maximal rate, the so-called *channel capacity*, and gave a method to calculate these capacities. However, when we consider quantum channels, a truly quantum phenomenon invalidates the same analysis: *quantum entanglement*. Suddenly, it might be possible to entangle the quantum information over multiple channel uses (modelled by taking *tensor products* of the quantum channels), which could possibly boost the overall capacity of the quantum channel. In turn, this would have a tangible influence on our choice of quantum internet protocols that we deploy on the quantum communication channels.

In the classical situation, the channel capacity is additive: when we use two channels to send information from A to B, the capacity is the sum of the individual capacities of the channels. In the quantum world, it was first conjectured that the classical capacity of a quantum channel was also additive, which would directly imply that entanglement between channel uses could not boost the channel capacity. After Shor showed in his breakthrough paper [Sho04] that this conjecture was equivalent to three other quantum additivity conjectures, among them the additivity conjecture of the *minimum output entropy* (from now on: *MOE*), proving any of these four conjectures false would prove that the classical capacity of a quantum channel was *not* necessarily additive. Hastings provided definitive proof that all these conjectures were false by constructing a counterexample to the *MOE*-additivity conjecture [Has09]. However, his proof used a random construction of a quantum channel that has a non-zero *probability* of violating the additivity conjecture – to this day, we have not found a deterministic description of any quantum channel that violates this additivity conjecture.

These additivity conjectures have not only resulted in theoretical research such as efficient methods of determining quantum capacities [ON00], but have also attracted the attention of experimental researchers, who have found coherent information additivity violations [YMP+20] and proof of superadditive quantum coding gain for small codes [FTMS03].

In general, computing the minimum output entropy of a quantum channel is a very hard problem (it is, in fact, NP-complete [BS07]), which makes it difficult to investigate whether quantum channels violate MOE additivity. Therefore, we could try to find families of quantum channels that have particular useful properties that allow us to calculate or approximate their MOE. A natural source of such quantum channels is the *representation theory* of mathematical objects such as (compact) groups: these quantum channels retain the symmetries of their underlying group to a certain degree, which allows us to exploit the group structure to calculate information-theoretic quantities of the quantum channels, such as their MOE. This method dates back to at least 2005, when Holevo proved the equivalence of MOE-additivity and χ-additivity for what he calls *irreducibly covariant quantum channels* [Hol05]. For specific groups, these quantum channels have been investigated: Al Nuwairan studied the convex set of $\mathrm{SU}(2)$-irreducibly covariant channels, in particular their extreme points, the so-called *EPOSIC* (*extreme points of* $\mathrm{SU}(2)$*-irreducibly covariant*) channels [Nuw14, Nuw13]. This technique can also be carried out for so-called *compact quantum groups*, an active mathematical research subfield of operator algebras and functional analysis that looks at non-commutative generalizations of the C^*-algebras of continuous functions on compact groups. The use of compact quantum groups allows us to capture the non-commutativity that is a recurring theme in quantum mechanics on a more fundamental level, namely inside the structure from which we build the quantum channels in the first place. Consensus has been reached on the basic concepts of compact quantum groups, and their representation theory is well-developed, which allows us to consider covariant quantum channels. Brannan and Collins introduced the quantum channels associated to the *free orthogonal quantum group* O_N^+ in [BC16b], and they also computed asymptotically sharp bounds on the MOE of these quantum channels [BCLY20]. The lower bounds on the MOE of these quantum channels grow relatively quickly with the dimensions of the input and output Hilbert spaces, which allows us to investigate whether the MOE of the tensor product of these channels could be strictly smaller than the sum of their respective MOE's. Furthermore, although computing the MOE is NP-complete, we will consider the possibilities of utilizing numerical methods to compute the MOE of these covariant quantum channel for low input and output dimensions.

Both the exploration of the symmetries of covariant quantum channels and the investigation of numerical methods for computing their MOE are important building blocks and techniques that we can apply to find deterministic descriptions of quantum channels that are strictly MOE-subadditive. This leads us to pose two questions in this thesis:

1. Can we apply the construction of Brannan and Collins to other compact quantum groups, such as *quantum permutation group* S_N^+, to yield a family of covariant channels? And in this case, how do the MOE bounds of Brannan and Collins generalize to this other quantum group?
2. Can we numerically approximate the MOE of these covariant quantum channel in low dimensions to satisfactory accuracy?

In this thesis, we seek to understand the mathematical construction of Brannan and Collins that is foundational to these quantum channels. Afterwards, we give an affirmative answer to the first question by introducing quantum channels associated to the quantum permutation group S_N^+, and we derive similar lower bounds for the MOE of these quantum channels.

To answer the second question, we propose the use of ε-covers to bound the MOE from below. We show a proof-of-concept for low dimensions applied to the S_N^+-quantum channels, and investigate how their size grows with the input dimension of these quantum channels. To approximate the MOE from above, we must solve a difficult optimization problem (concave minimization over a convex set), and we investigate whether the derivative-free optimization method called *Particle Swarm Optimization* (enriched with gradient-based descent) can be deployed to find good minimums for the MOE.

Lastly, we discuss and implement both the DPS hierarchy (named after its inventors Doherty, Parrilo and Spedalieri, [DPS02]) and its modification, the DPS* criterion, which are

methods that indicate whether a quantum state is separable or entangled. We utilize these methods to investigate whether certain O_N^+-quantum channels are *entanglement-breaking*, as quantum channels with that property cannot violate the MOE additivity.

The numerical tools gathered and developed in this thesis can be publicly accessed at `qittoolbox` on GitHub .

We briefly discuss the outline of this thesis. In Chapters 2 and 3 we will cover the necessary mathematical and physical background. We hope that these chapters can be useful references for the reader whenever further chapters utilize theory that is unfamiliar to the reader. In Chapter 4, we discuss the theory of compact quantum groups, in particular the compact *matrix* quantum groups, together with the Tannaka-Krein duality, which is necessary to understand how a compact matrix quantum group is in a certain sense fully characterized by the category of its representations. In Chapter 5 we specialize to the O_N^+ and S_N^+ quantum groups, and discuss their representation theory in more detail. This chapter also lays the foundation of the S_N^+-quantum channels that we introduce in this thesis. In Chapter 6 we show how one constructs quantum channels out of the representation theory of the compact quantum groups discussed in the previous chapter(s), and we note some results on the MOE of the S_N^+-channel.

Afterwards, we specialize our attention to the so-called "lowest-weight" quantum channels which have the smallest input dimension (without being trivial) in Chapter 7. There, we strengthen some of the results on positive partial transpose and entanglement-breaking properties as found in [BCLY20] for the lowest-weight quantum channels, and we show a surprising result which allows us to calculate the MOE exactly in the smallest case where we embed the fundamental representation in a larger tensor product of representations.

Lastly, we turn our attention to the numerical approximations and implementations to find bounds on the MOE of quantum channels in Chapter 8. We discuss the continuity of the von Neumann entropy, the ε-covers which provide a lower bound on the MOE, the *Particle Swarm Optimization* algorithm which provides an upper bound on the MOE, and the DPS hierarchy and the DPS* criterion to determine whether a quantum channel is entanglement-breaking.

We conclude our findings in Chapter 9.

2

An introduction to universal C^*-algebras and partition theory

As this thesis deals with a topic that heavily relies on both mathematics and quantum informational physics, it is imperative to sketch a broad background of both the physical and mathematical tools involved in the project. We expect at least an undergraduate mathematical maturity level (e.g. , the reader is familiar with vector spaces, groups and some elementary results from group theory, and elementary set theory), but we will discuss relevant mathematical theory (such as C^*-algebras).

Let us commence with some notation and terminology. We denote the natural numbers starting from 1 by $\mathbb{N}$, and will write $\mathbb{N}_0$ if we wish to include 0. Similarly, $\mathbb{Z}$, $\mathbb{R}$ and $\mathbb{C}$ denote the integers, real numbers and complex numbers.By $\mathbb{S}^{n-1}$ we denote the $(n-1)$-sphere embedded in $\mathbb{R}^n$ (or embedded in $\mathbb{C}^n$, which will be clear from context). The sesquilinear forms on a complex vector space V are usually denoted by $\langle\cdot|\cdot\rangle : V \times V \to \mathbb{C}$, and we define them to be *conjugate-linear* in the *first* coordinate, and *linear* in the second – this choice is made to be consistent with the Dirac bra-ket notation, which is omnipresent in quantum information: in a Hilbert space $\mathcal{H}$, vectors $\xi \in \mathcal{H}$ are sometimes also written as $|\xi\rangle \in \mathcal{H}$. In this case, $\eta^* \in \mathcal{H}^*$, i.e. the linear map $\eta^* : \mathcal{H} \to \mathbb{C}$ by $\eta^*(\xi) = \langle\eta|\xi\rangle$, is also denoted by $\langle\eta| := |\eta\rangle^* \in \mathcal{H}^*$. Then, notation like $\langle\eta|\,|\xi\rangle$ corresponds with the inner product $\langle\eta|\xi\rangle$. If $\mathcal{H}$ and $\mathcal{K}$ are Hilbert spaces, $\mathcal{H}\otimes\mathcal{K}$ denotes the Hilbert space tensor product of $\mathcal{H}$ and $\mathcal{K}$. If A and B are C^*-algebras, $A\otimes B$ denotes the minimal C^*-tensor product.

Having fixed the terminology and notation, we can move on to the mathematical tools we will need to develop to understand compact quantum groups, their representations, and quantum channels and their information-theoretic properties. As compact quantum groups, contrary to what the name suggests, are not "groups" but rather C^*-algebras together with a comultiplication and a compatibility axiom, we start with C^*-algebras. Conveniently, the language of C^*-algebras is also the language in which quantum systems and quantum states are described.

2.1. C^*-algebras

As we will exclusively deal with C^*-algebras over the complex field $\mathbb{C}$ in this thesis, for brevity we will restrict our following definitions to the complex field, although one can easily generalize this to arbitrary fields $\mathbb{F}$. As we can view a C^*-algebra as a vector space with a multiplication, norm, involution, and compatibility rules between these structures, in order to show the interplay of mathematical structure we give a C^*-algebra, we first discuss those structures separately. Two key examples of C^*-algebras will be $C[0,1]$, the continuous (complex-valued) functions on the real interval $[0,1]$, and $\mathcal{B}(\mathcal{H})$, the set of all bounded linear operators on a

Hilbert space $\mathcal{H}$. [1] Necessarily, the introduction will be short, and will be far from a complete introduction to the basics of C^*-algebra theory. See [Mur14] for an excellent introduction.

Let us commence with a vector space A over $\mathbb{C}$. We endow A with a $\mathbb{C}$-bilinear *multiplication* $\cdot : A \times A \to A$ that is *associative*, i.e. $(ab)c = a(bc)$ for all $a, b, c \in A$. This turns A into a $\mathbb{C}$-algebra. If we have two such $\mathbb{C}$-algebras A and B, a map $\varphi : A \to B$ is called an *algebra homomorphism* if it is linear and respects products, i.e. $\varphi(ab) = \varphi(a)\varphi(b)$. If A has an element $1_A \in A$ which we call its *unit* such that $1_A a = a 1_A = a$ for all $a \in A$, then we say that A is a *unital algebra*. For example, we can turn $C[0,1]$ into a unital algebra by considering the product $(f \cdot g)(x) = f(x)g(x)$, and we can turn $\mathcal{B}(\mathcal{H})$ into a unital algebra by considering the matrix product $a \cdot b = a \circ b$ (composition of linear maps). Some algebras such as $C[0,1]$ have the property that $f \cdot g = g \cdot f$ for all $f, g \in C[0,1]$, and we call them *commutative*.

Our examples $C[0,1]$ and $\mathcal{B}(\mathcal{H})$ not only have algebraic structure, they also have topological structure induced by a *norm* on the underlying vector space. To make both structures compatible, we demand that the norm is *submultiplicative*, i.e. $\|ab\| \leqslant \|a\| \cdot \|b\|$. If a $\mathbb{C}$-algebra A has such a submultiplicative norm, we call A a *normed algebra*. For example, $C[0,1]$ can be endowed with the sup-norm $\|f\|_\infty := \sup_{x \in [0,1]} |f(x)|$ and $\mathcal{B}(\mathcal{H})$ with the operator norm $\|a\| = \sup_{\xi \neq 0} \|a\xi\|/\|\xi\|$ where the supremum ranges over all nonzero $\xi \in \mathcal{H}$. The submultiplicativity of the sup-norm follows directly, and it is a typical exercise to show that the operator norm is submultiplicative. Hence, $C[0,1]$ and $\mathcal{B}(\mathcal{H})$ are normed algebras.

However, $C[0,1]$ and $\mathcal{B}(\mathcal{H})$ have the additional property that as normed vector spaces, they are *complete* (i.e., a Cauchy sequence in this space always converges to a point in this space). This turns a normed algebra into a *Banach algebra*.

Lastly, we need the structure of an *involution*: for example, a complex-valued function $f \in C[0,1]$ can be conjugated by setting $f^*(x) := \overline{f(x)}$ (where $\overline{z}$ is the complex conjugate of a complex number z), and an operator $a \in \mathcal{B}(\mathcal{H})$ has an *adjoint* $a^* \in \mathcal{B}(\mathcal{H})$ (in Physics, this is also often referred to as the *hermitian conjugate*, and denoted by $a^\dagger$). In general, an involution on an algebra A is a $\mathbb{C}$-antilinear map $* : A \to A$ such that $(a^*)^* = a$ and $(ab)^* = b^* a^*$ for all $a, b \in A$. If A is endowed with an involution, we call A a $*$-algebra. A map φ between two $*$-algebras A and B is called a *$*$-homomorphism* if it is an algebra homomorphism and $\varphi(a^*) = \varphi(a)^*$ for all $a \in A$. If A is a Banach algebra, we additionally demand that $\|a^*\| = \|a\|$, and then call A a Banach $*$-algebra. If we additionally demand that the C^*-identity $\|a^* a\| = \|a\|^2$ holds for all $a \in A$, we call A a C^*-algebra. By Gelfand's spectral radius formula, we in fact have $\|a\|^2 = \|a^* a\| = \sup\{|\lambda| \ : \ a^* a - \lambda 1_A \text{ is not invertible}\}$, hence the C^*-norm is defined by purely algebraic structure, which also directly shows that at most one C^*-norm exists on a $*$-algebra.

Now that we have established the structures that turn $C[0,1]$ and $\mathcal{B}(\mathcal{H})$ into C^*-algebras, we consider the properties of their individual elements. Given a C^*-algebra A and $a \in A$, we say that a is *normal* if $aa^* = a^* a$ (in the commutative case such as $C[0,1]$, any element is normal. In the finite-dimensional case of $\mathcal{B}(\mathcal{H}) \simeq M_n(\mathbb{C})$ of $n \times n$ matrices, normal elements are precisely those matrices that are diagonalizable by a unitary matrix). We say that a is *self-adjoint* or *hermitian* if $a^* = a$. We call a *unitary* if $aa^* = a^* a = 1_A$. We say that a is *positive* if $a = b^* b$ for some $b \in A$, note that positive elements are always self-adjoint [2]. Lastly, a is a *projection* if it is both self-adjoint and idempotent, i.e. $a^* = a = a^2$.

Of course, there is much more to say about C^*-algebras, but for now we discuss two topics that are vital to the constructions and arguments we will later encounter.

[1] In a certain sense, these encompass "all" examples, because deep theorems following from the Gelfand isomorphism and the Gelfand-Naimark theorem revealed that any commutative C^*-algebra is isomorphic to $C_0(X)$ for some locally compact space X (and in the unital case, this is $C(X)$ on some compact space X), and any C^*-algebra is isomorphic to a norm-closed $*$-subalgebra of $\mathcal{B}(\mathcal{H})$ for some Hilbert space $\mathcal{H}$ through the GNS construction.

[2] Note that positive elements are generally defined as those $a \in A$ that are hermitian and whose spectrum $\sigma(a)$ satisfies $\sigma(a) \subseteq \mathbb{R}_{\geqslant 0}$, but we did not define the spectrum (as we will not need it), hence we use an equivalent definition. For further information, please see [Mur14, Thm. 2.2.5.]

2.1.1. QUANTUM MECHANICAL AND C^*-ALGEBRAIC STATES

In quantum mechanics, all relevant information about a quantum system is stored in the quantum "state", where we distinguish between so-called *pure states* and (statistically) *mixed states*. This notion coincides in the finite-dimensional case with the more general notion of a state as a normalized positive functional on a C^*-algebra. Let us first define:

Definition 2.1.1: Faithful, positive and normalized functionals and states

Given a functional (i.e., a linear map) $\varphi : A \to \mathbb{C}$ from a complex algebra A to its field $\mathbb{C}$. We say:

- φ is *faithful* if, for any $a \in A$ with $a \neq 0$, we have $\varphi(aA) \neq \{0\}$ and $\varphi(Aa) \neq \{0\}$.
- φ is *normalized* if A is a unital algebra and $\varphi(1_A) = 1$.
- φ is *positive* if A is a $*$-algebra and for all $a \in A$, we have $\varphi(a^*a) \geqslant 0$. We write $\varphi \geqslant 0$ to say φ is positive. For two positive functionals φ and τ, we write $\varphi \geqslant \tau$ if $\varphi - \tau \geqslant 0$.

A functional φ that is both normalized and positive is called a *state* on the unital $*$-algebra A. The set of all states on a C^*-algebra A is denoted by $\mathcal{S}(A)$. We call such a a state *pure* if for any positive functional $\tau \leqslant \varphi$ one can find a $t \in [0, 1]$ such that $\tau = t\varphi$, and denote the set of all pures states on A by $\mathcal{PS}(A)$.

In the quantum mechanical setting, the following definition is standard:

Definition 2.1.2: Quantum mechanical state

In finite dimensions, a quantum system is described in a finite-dimensional complex Hilbert space $\mathcal{H} = \mathbb{C}^n$, and the quantum mechanical information is modelled either by a *pure state* or a *mixed state*:

- A *pure state* is modelled by a vector $|\psi\rangle \in \mathcal{H}$ with $\langle\psi|\psi\rangle = \|\psi\|^2 = 1$, where the expectation of an observable $x \in M_n(\mathbb{C})$ is calculated by $x \mapsto \langle x\rangle_\psi := \langle\psi|x|\psi\rangle$.
- A *mixed state* is modelled by a *density matrix* $\rho \in \mathcal{B}(\mathcal{H}) = M_n(\mathbb{C})$ which is a positive matrix with $\mathrm{Tr}(\rho) = 1$, where the expectation of an observable $x \in M_n(\mathbb{C})$ is calculated by $x \mapsto \mathrm{Tr}\,(\rho x)$.

When one refers to a *state* (without specifying whether it is pure or mixed), in general one is referring to a *mixed state* (note that pure states are included in the definition of mixed states, by using $\rho = |\psi\rangle\langle\psi|$), and we denote with $\mathcal{S}(\mathcal{H})$ the set of all states on a Hilbert space $\mathcal{H}$.

Note that a quantum mechanical pure state modelled by $|\psi\rangle \in \mathcal{H} = \mathbb{C}^n$ can be associated with a C^*-algebraic state $\varphi_\psi : \mathcal{B}(\mathcal{H}) = M_n(\mathbb{C}) \to \mathbb{C}$ by $x \mapsto \langle\psi|x|\psi\rangle$. Similarly, a *density matrix* $\rho \in M_n(\mathbb{C})$ can be associated with a C^*-algebraic state $\varphi_\rho : M_n(\mathbb{C}) \to \mathbb{C}$ by $x \mapsto \mathrm{Tr}(\rho x)$. In fact, these associations are bijections:

Theorem 2.1.1

In the finite-dimensional case $\mathcal{H} = \mathbb{C}^n$, the quantum mechanical pure states in Definition 2.1.2 are in bijective correspondence with the C^*-algebraic pure states on $\mathcal{B}(\mathcal{H})$ in Definition 2.1.1 by mapping a state $|\psi\rangle$ to $\varphi_\psi : x \mapsto \langle\psi|x|\psi\rangle$, and the quantum mechanical mixed states are in bijective correspondence with the C^*-algebraic states by mapping a state ρ to $\varphi_\rho : x \mapsto \mathrm{Tr}(\rho x)$.

Proof. Murphy's Example 5.1.1 in [Mur14] shows the more general result that pure states of a C^*-algebra $A := \mathcal{K}(\mathcal{H})$ for some Hilbert space $\mathcal{H}$ are precisely of this form [3]. The mixed state case is also a well-known result, see for example [Jan10, Thm. 2.7]. □

[3] Here, $\mathcal{K}(\mathcal{H})$ denotes the space of all *compact* operators on $\mathcal{H}$. In the finite-dimensional case, $\mathcal{B}(\mathcal{H}) = \mathcal{K}(\mathcal{H}) = M_n(\mathbb{C})$

2.1.2. $*$-IDEALS OF AND SEMINORMS ON C^*-ALGEBRAS

A necessary tool that we must develop before discussing universal C^*-algebras is the construction of a C^*-algebra from a $*$-algebra with a C^*-seminorm. The latter is defined as follows:

Definition 2.1.3: C^*-seminorm

Given a $*$-algebra A. A C^*-seminorm is a function $p : A \to [0, \infty)$ such that:

1. $p(\lambda a) = |\lambda| p(a)$ for all $\lambda \in \mathbb{C}$ and $a \in A$
2. $p(a + b) \leqslant p(a) + p(b)$ for all $a, b \in A$
3. $p(ab) \leqslant p(a)p(b)$ for all $a, b \in A$
4. $p(a^*a) = p(a)^2$ for all $a \in A$.

Example 2.1.1. Given a $*$-algebra A and a C^*-algebra B, and any $*$-homomorphism $\varphi : A \to B$. We then see that p given by $p(x) := \|\varphi(x)\|$ is a C^*-seminorm on A.

It is directly clear from the definition that the norm $\|\cdot\|$ of a C^*-algebra A is also a C^*-seminorm on A. However, it is *not* true that a C^*-seminorm p is generally a norm, as $p(a) = 0$ does not generally imply $a = 0$. However, we can look at quotient $*$-algebras, where we mod out the kernel of p, to make p an actual norm. This construction succeeds because the kernel of p is a two-sided $*$-ideal. An ideal is a subspace $B \subseteq A$ of an algebra A such that $ab, ba \in B$ for any $a \in A$ and $b \in B$. If A is a $*$-algebra, B is called a *two-sided $*$-ideal* if additionally $B^* = B$. Given such a two-sided $*$-ideal I, we can "mod out I": we denote by A/I the collection of all cosets $a + I$ for $a \in A$, with $(a + I) + (b + I) = (a + b) + I$, $(a + I)(b + I) = (ab) + I$, $(a + I)^* = a^* + I$ and $\lambda(a + I) = (\lambda a) + I$ for $\lambda \in \mathbb{C}$. This turns A/I into a $*$-algebra.

The thought of modding out the kernel of the C^*-seminorm p on A has merit, because $\ker(p) := \{x \in A : p(x) = 0\}$ is a two-sided $*$-ideal, which is a standard exercise in using the properties of a C^*-seminorm. We have now achieved a $*$-algebra $A/\ker(p)$ on which the induced map $\hat{p}$ given by $a + \ker(p) \mapsto p(a)$ has a trivial kernel. This turns $\hat{p}$ into a norm on $A/\ker(p)$, and by densely embedding $A/\ker(p)$ in its completion $\hat{A} := \overline{A/\ker(p)}^{\hat{p}}$ as a metric space, the space $\hat{A}$ turns into a C^*-algebra with C^*-norm $\hat{p}$.

2.2. UNIVERSAL C^*-ALGEBRAS

In order to construct compact quantum groups, which consist of a C^*-algebra, comultiplication, and some compatibility conditions, it is often useful to describe the C^*-algebra as a "universal C^*-algebra" of elements with certain constraints. For example, $M_n(\mathbb{C})$ is a C^*-algebra (as it is, in fact, a special case of $\mathcal{B}(\mathcal{H})$ with $\mathcal{H} = \mathbb{C}^n$), but we could equivalently describe it as the C^*-algebra brought forth by the distinct "basis elements" E_{ij} for $1 \leqslant i, j \leqslant n$, with the constraints that $E_{ij}^* = E_{ji}$ and $E_{ij}E_{kl} = \delta_{jk}E_{il}$. We make this notion exact in the following section, which is largely based on [LVWB21] and [Bla06]. Let us commence with defining a free $*$-algebra on a set of elements.

Definition 2.2.1: Free (complex) algebra on the generator set X

Let $X := \{x_i \,:\, i \in I\}$ be a set of elements x_i, indexed by I. We define a (non-commutative) *monomial* in X as a word $x_{i_1} \dots x_{i_n}$ where $i_j \in I$ for all $1 \leqslant j \leqslant n$ and $n \in \mathbb{N}$. We define a (non-commutative) *polynomial* in X as a formal $\mathbb{C}$-linear combination of monomials y_k for $1 \leqslant k \leqslant N$ with $N \in \mathbb{N}$, and so we write the polynomial as $\sum_{k=1}^{N} \lambda_k y_k$ for $\lambda_k \in \mathbb{C}$. We define the *addition* of two non-commutative polynomials $\sum_{k=1}^{N} \lambda_k y_k$ and $\sum_{k=1}^{M} \mu_k z_k$ as the polynomial $\sum_{k=1}^{N+M} \nu_k t_k$ where $\nu_k = \lambda_k$ and $t_k = y_k$ for $k \leqslant N$, and $\nu_k = \mu_{k-N}$ and $t_k = z_{k-N}$ for $N < k \leqslant N+M$. We define the *scalar multiplication* on polynomials in X as $\mu \cdot \sum_{k=1}^{N} \lambda_k y_k = \sum_{k=1}^{N} (\mu\lambda_k) y_k$. The associative *multiplication* of two monomials is defined by concatenation, i.e. $(x_{i_1} \cdots x_{i_n})(x_{j_1} \cdots x_{j_m}) := x_{i_1} \cdots x_{i_n} x_{j_1} \cdots x_{j_m}$, and we canonically extend this $\mathbb{C}$-linearly to polynomials in X. We then define the *free algebra Alg(X) on the generator set X* as the set of polynomials in X, together with the addition, scalar multiplication and associative multiplication defined above.

Note that the elements $x_i \in X$ are non-commutative in general, so the monomials $x_1 x_2$ and $x_2 x_1$ are in general distinct in the free algebra on the generator set X. Furthermore, note that if we include the *empty word* (the monomial in X with length 0) in our definition, this empty word is the multiplicative *unit* of Alg(X), turning it into a *unital* free algebra. We will also denote the unital free algebra on X by Alg(X), and will in this case explicitly state that it is unital.

Definition 2.2.2: Free $*$-algebra on the generator set X

Following Definition 2.2.1, we let X be the set of generators x_i, and we now consider another set $X^* := \{x_i^* \,:\, i \in I\}$ which we see as *disjoint* from X. We consider the free algebra Alg$(X \cup X^*)$, and we define the involution on this free algebra as

$$\left(\mu x_{i_1}^{\epsilon_1} \cdots x_{i_n}^{\epsilon_n}\right)^* := \overline{\mu} \cdot x_{i_n}^{\epsilon_n^*} \cdots x_{i_1}^{\epsilon_1^*}, \tag{2.1}$$

where $\mu \in \mathbb{C}$, and $\epsilon_i \in \{1, *\}$ (where we set $x_i^1 := x_i$ for all $i \in I$), and $\epsilon_i^* := *$ if $\epsilon_i = 1$ and $\epsilon_i^* := 1$ if $\epsilon_i = *$, and we extend this linearly to all of Alg$(X \cup X^*)$. The *free $*$-algebra $*$-Alg(X) on the generator set X* is defined as Alg$(X \cup X^*)$ with the involution as defined above.

Definition 2.2.3: Universal $*$-algebra with generators and relations

Let $X := \{x_i \,:\, i \in I\}$ be a set of elements x_i and consider $*$-Alg(X). Assume we have a set of polynomials in X called $R \subseteq *$-Alg(X). Find the two-sided $*$-ideal $J(R)$ in $*$-Alg(X) generated by R. We define the *universal $*$-algebra with generators X and relations R* as the quotient $*$-Alg$(X|R) := *$-Alg$(X)/J(R)$.

Definition 2.2.4: Universal C^*-algebra

Let X be a set of generators, $R \subseteq *$-Alg(X) a set of relations. Let us put

$$\|x\| := \sup\{p(x) \mid p \text{ is a } C^*\text{-seminorm on } *\text{-Alg}(X|R)\}. \tag{2.2}$$

The reader should convince him- or herself that $\|\cdot\|$ is a C^*-seminorm, provided $\|x\| < \infty$ for all $x \in *$-Alg$(X|R)$ (this follows directly from the definition). We know that the kernel K of $\|\cdot\|$ is a two-sided $*$-ideal, so we can define *the universal C^*-algebra of generators X with relations R* as

$$C^*(X|R) := \overline{*\text{-Alg}(X|R)/K}^{\|\cdot\|}, \tag{2.3}$$

where the overline means that we take a completion of the quotient space with respect to the C^*-seminorm $\|\cdot\|$.

Note that it is not directly clear *when* this construction will work, as we assume that $\|x\| < \infty$ for all $x \in *\text{-Alg}(X|R)$. Furthermore, it is entirely possible that $C^*(X|R)$ is the trivial C^*-algebra $\{0\}$. Let us investigate both problems.

Lemma 2.2.1: Sufficient condition for $\|x\| < \infty$

Given a generator set $X = \{x_i \ : \ i \in I\}$ and relations $R \subseteq P(X)$, assume there is a constant $C > 0$ such that $p(x_i) < C$ for all C^*-seminorms p on $*\text{-Alg}(X|R)$ and all $x_i \in X$. Then $\|x\| < \infty$ for all $x \in *\text{-Alg}(X|R)$.

Proof. Note that $p(x_{i_1} \cdots x_{i_n}) \leqslant p(x_{i_1})p(x_{i_2}) \cdots p(x_{i_n}) < C \cdot C \cdots C = C^n$, so the quantity $\|x\|$ for any monomial x of length n is bounded by C^n. By the triangle inequality, any polynomial in X is also $\|\cdot\|$-bounded, hence $\|x\| < \infty$ for all $x \in *\text{-Alg}(X|R)$. □

To investigate the problem where $C^*(X|R)$ might be trivial, we can use the *universality property* of this C^*-algebra – we can show that $C^*(X|R)$ is non-trivial by finding a non-trivial $*$-homomorphism to another C^*-algebra, provided we can find a set of elements that satisfy the same relations R, and provided that $C^*(X|R)$ exists in the first place:

Proposition 2.2.1: Universality property of $C^*(X|R)$

Let $X = \{x_i | i \in I\}$ be a set of elements, $R \subseteq *\text{-Alg}(X)$ be the relations, and assume that $C^*(X|R)$ exists. Given any C^*-algebra B, with a subset $Y := \{y_i | i \in I\} \subseteq B$ that satisfy the relations R, i.e. if we change all x_i for y_i in the polynomial expressions in R, then each expression is identically zero. Then, there exists a unique $*$-homomorphism $\varphi : C^*(X|R) \to B$ mapping x_i to y_i for all $i \in I$.

Proof. First, let $\varphi_0 : *\text{-Alg}(X) \to B$ be the $*$-homomorphism sending x_i to y_i – we can do this, because $*\text{-Alg}(X)$ is a free $*$-algebra without any constraints over the generator set X. Note that the two sided $*$-ideal $J(R)$ generated by R in $*\text{-Alg}(X)$ vanishes in B if we replace all x_i by y_i, by assumption. Hence, we get an induced $*$-homomorphism $\hat{\varphi}_0 : *\text{-Alg}(X|R) \to B$ by $\hat{\varphi}_0 : x + R \mapsto \varphi(x)$. What is left is to show that $\hat{\varphi}_0$ is continuous with respect to the sup-norm $\|\cdot\|$ on $*\text{-Alg}(X|R)$, such that we can uniquely extend it to a $*$-homomorphism $\varphi : C^*(X|R) \to B$. We can use Example 2.1.1: we set $p(x) = \|\hat{\varphi}_0(x)\|_B$, and note that this is a C^*-seminorm on $*\text{-Alg}(X|R)$, hence we see $\|\hat{\varphi}_0(x)\|_B \leqslant \|x\|$, so $\hat{\varphi}_0$ is indeed continuous. Uniqueness follows from the fact that $*$-homomorphisms that agree on the generators of a C^*-algebra must necessarily agree on the entire generated C^*-algebra [4]. □

Proposition 2.2.2: Uniqueness of $C^*(X|R)$

Assume two C^*-algebras A and B are both generated as C^*-algebras by a set of elements $X := \{x_i | i \in I\}$ and $Y := \{y_i | i \in I\}$, respectively, and that these elements satisfy a set of relations $R \subseteq *\text{-Alg}(X)$ (for B, we exchange x_i for y_i). Furthermore, assume that they both satisfy the same universality property: if another C^*-algebra C contains a subset $Z := \{z_i | i \in I\}$ such that they satisfy the relations R (if we exchange x_i with z_i), then there exist unique $*$-homomorphisms $\varphi_A : A \to C$ and $\varphi_B : B \to C$ mapping x_i to z_i and mapping y_i to z_i, respectively. If this is the case, then $A \simeq B$ as C^*-algebras.

Proof. As $Y \subseteq B$ satisfies the relations R, by the universality property of A we can construct a $*$-homomorphism $\varphi_A : A \to B$ mapping x_i to y_i. As $X \subseteq A$ satisfies the relations R, the universality property of B implies that we have a $*$-homorphism $\varphi_B : B \to A$ mapping y_i to x_i.

Clearly, $\varphi_A \circ \varphi_B = \text{id}_B$ and $\varphi_B \circ \varphi_A = \text{id}_A$ as $*$-homomorphisms that agree on the generators of a C^*-algebra must necessarily agree on the entire generated C^*-algebra, hence we have $A \simeq B$ as C^*-algebras. □

[4] See Lemma 3.26 in [LVWB21] for details on this

Let us now provide two situations where the universal C^*-algebra construction "fails": we show that $C^*(X|R)$ need not exist, and we show a situation where $C^*(X|R) = 0$.

Example 2.2.1 (Non-existent universal C^*-algebra). Let $X := \{x\}$ and let $R := \{x - x^*\}$, adopt the more pleasant notation $C^*(X|R) = C^*(x|x = x^*)$. This object does not exist. We show this by using Example 2.1.1 and Proposition 2.2.1. Namely, consider the C^*-algebra $C[0,1]$, and for any $\lambda > 0$, define $f_\lambda := \lambda 1_{[0,1]} \in C[0,1]$. As $f^* = f$, it satisfies the relationship R, hence by Proposition 2.2.1, construct the $*$-homomorphism $\varphi_\lambda : C^*(x|x^* = x) \to C[0,1]$ mapping x to f_λ. This yields a C^*-seminorm p_λ on $C^*(x|x^* = x)$ given by $p_\lambda(y) = \|\varphi_\lambda(y)\|_\infty$. But, as $p_\lambda(x) = \|f\|_\infty = \lambda$, and we must have $p_\lambda(x) \leqslant \|x\|$, we see that $\lambda \leqslant \|x\|$. But $\lambda > 0$ was arbitrary, so $\|x\| = \infty$, hence $C^*(x|x^* = x)$ does not exist!

Example 2.2.2 (Trivial universal C^*-algebra). Given $X = \{x\}$ and $R = \{x^2, xx^*x - x, x - x^*\}$, then we have $C^*(x|x^2 = 0, xx^*x = x, x = x^*) = 0$. Namely, for any C^*-seminorm p on $*$-Alg$(x|x^2 = 0, xx^*x = x, x = x^*)$, we have $p(x)^2 = p(x^*x) = p(x^*xx^*x) = p(x^*x)^2 = p(x)^4$, so $p(x) \in \{0,1\}$, thus by Lemma 2.2.1, $C^*(X|R)$ exists. However, $\|x\|^2 = \|x^*x\| = \|x^2\| = \|0\| = 0$, so $C^*(X|R) = 0$.

In contrast to these examples where the universal C^*-algebra construction fails, we also have situations where it is successful, such as the following construction, see [LVWB21, Prop. 6.11]:

Proposition 2.2.3

For $n \in \mathbb{N}_{\geqslant 2}$, the C^*-algebra $M_n(\mathbb{C})$ is isomorphic as C^* algebra with

$$C^*(E_{ij}, 1 \leqslant i,j \leqslant n \mid E_{ij}^* = E_{ji}, E_{ij}E_{kl} = \delta_{jk}E_{il}) \tag{2.4}$$

Remark 2.2.1. Note that Blackader in [Bla06] uses a slightly different construction, where the generators $X := \{x_i | i \in \mathcal{I}\}$ need to be realized on a Hilbert space and the relations in R are of the form $\|p(x_{i_1}, \ldots, x_{i_n}, x_{i_1}^*, \ldots, x_{i_n}^*)\| \leqslant \eta$, where p is a polynomial of $2n$ generators in X, and $\eta \geqslant 0$. These relations must (either explicitly or implicitly) bound the norm of those generators seen as operators on this Hilbert space. He defines a *representation* of X with relations R as a set of bounded operators $\{T_i | i \in I\}$ on a certain Hilbert space $\mathcal{H}$, such that if one replaces x_i with T_i (and x_i^* with the adjoint T_i^*) in the relations in R, the relations $\|p(T_{i_1}, \ldots, T_{i_n}, T_{i_1}^*, \ldots, T_{i_n}^*)\| \leqslant \eta$ are satisfied. He then sets $\|x\| := \sup\{\|\pi(x)\|_\mathcal{H}\}$, where the supremum is over all $(\pi, \mathcal{H})$ that are representations of X with relations R. Then, $C^*(X|R)$ is defined as the free $*$-algebra $*$-Alg(X), where one divides out all $\{x : \|x\| = 0\}$, and completed with respect to the C^*-seminorm $\|\cdot\|$, assuming that the supremum in the definition of $\|\cdot\|$ is finite for all $x \in *$-Alg(X). The constructions coincide if we choose all constraints in R to have $\eta = 0$, as follows:

Proposition 2.2.4

Given a set of generators $X := \{x_i | i \in I\}$ and a set of relations $R := \{q_j = 0 | j \in J\}$ where q_j are polynomials in the $2n$ non-commuting variables $x_{i_1}, \ldots, x_{i_n}$ and $x_{i_1}^*, \ldots, x_{i_n}^*$. Assuming the universal C^*-algebra $C^*(X|R)$ as defined in Definition 2.2.4 exists, then it coincides with Blackader's universal C^*-algebra, which we shall call $C_B^*(X|R)$ with C^*-norm $\|x\|_B$.

Proof. Let us show that $\|x\|_B \leqslant \|x + J(R)\|$. Given a $*$-representation $(\pi, \mathcal{H})$ of $\mathcal{A} := *$-Alg(X) such that $\{\pi(x_i) \mid i \in \mathcal{I}\} \subset \mathcal{B}(\mathcal{H})$ satisfy R. Let $\mathcal{B} := *$-Alg$(X)/J(R) = *$-Alg$(X|R)$, and abbreviate $J(R)$ by J. We can descend to a $*$-homomorphism $\hat{\pi} : \mathcal{B} \to \mathcal{B}(\mathcal{H})$ by $\hat{\pi}(x + J) = \pi(x)$, because if $x + J = y + J$, then $x - y \in J$, and as $\{\pi(x_i) \mid i \in \mathcal{I}\}$ satisfy R, we must have $\pi(x - y) = 0$. Then, the map $x + J \mapsto \|\hat{\pi}(x + J)\|_\mathcal{H} = \|\pi(x)\|_\mathcal{H}$ defines a C^*-seminorm on $\mathcal{B} = *$-Alg$(X|R)$, hence $\|x + J\| \geqslant \|\pi(x)\|_\mathcal{H}$ for any $*$-representation $(\pi, \mathcal{H})$ on $*$-Alg(X) that satisfies R, hence by taking a supremum over those representations we get $\|x + J\| \geqslant \|x\|_B$.

Let us now show that $\|x\|_B \geqslant \|x+J\|$. Given the C^*-seminorm p on $\mathcal{B} = *\text{-Alg}(X|R)$, consider the $*$-algebra $\mathcal{B}_p := \mathcal{B}/\text{Ker}(p)$, and complete it with respect to the C^*-seminorm p to yield a C^*-algebra $\overline{\mathcal{B}_p} := \overline{\mathcal{B}/\text{Ker}(p)}^p$. Now consider the canonical projection $*$-homomorphism $\pi_p : \mathcal{A} \to \overline{\mathcal{B}_p}$ by $x \mapsto x + J \mapsto (x + J) + K$, where $K := \text{Ker}(p)$. As π_p clearly contains J in its kernel, the elements $\{\pi_p(x_i)|i \in \mathcal{I}\}$ satisfy R. Each C^*-algebra can be realized as a closed $*$-subalgebra of a $\mathcal{B}(\mathcal{H}_p)$ for some Hilbert space $\mathcal{H}_p$, so $(\pi_p, \mathcal{H}_p)$ is a representation of X satisfying R, with the property $\|\pi_p(x)\| = p(x + J)$. Hence, $p(x + J) \leqslant \|x\|_B$, and by taking a supremum over all possible C^*-seminorms p on $\mathcal{B}$, we retrieve $\|x + J\| \leqslant \|x\|_B$.

It follows that both constructions yield the same universal C^*-algebra $C^*(X|R)$. □

Note that Blackader's definition can be used to define expressions such as $C^*(x|x = x^*, \|x\| \leqslant 1)$, namely by constructing $\mathcal{A} := *\text{-Alg}(X)$, $\|x\| := \sup\{\|\pi(x)\|\}$ with the supremum over all $*$-representations $(\pi, \mathcal{H})$ of $\mathcal{A}$ such that $\pi(x) = \pi(x)^*$ and $\|\pi(x)\| \leqslant 1$, and then finally constructing $\overline{\mathcal{A}/K}^{\|\cdot\|}$ with $K := \text{Ker}\|\cdot\|$. This universal C^*-algebra, in contrast with Example 2.2.1, actually *does* exist! It is, up to isomorphism, precisely $C_0(X)$ with $X := [-1, 1] \setminus \{0\}$ (i.e., the C^*-algebra of all continuous functions on $[-1, 1]$ that vanish in 0), with x identified with the identity function $\text{id} : \lambda \mapsto \lambda$. Clearly, x satisfies R, $C_0(X)$ is generated as a C^*-algebra by x, and let us check the universality condition.

Given any C^*-algebra B with an element $y \in B$ satisfying $y = y^*$ and $\|y\| \leqslant 1$. We need to find a $*$-homomorphism φ from $C_0(X)$ to B mapping x to y. We can see this will work through Continuous Functional Calculus: construct the C^*-subalgebra $C^*\langle y\rangle \subseteq B$ that is generated by the element y, which is in general non-unital, and then consider the Gelfand-Naimark isomorphism $\pi : C_0(\sigma(y) \setminus \{0\}) \to C^*\langle y\rangle \subseteq B$. By $\|y\| \leqslant 1$, we know $\sigma(y) \subseteq [-1, 1]$. We also know that $\pi^{-1}(y) = \text{id} : \lambda \mapsto \lambda$. Hence, our $*$-homomorphism φ maps a function in $C_0(X)$ to its restriction in $C_0(\sigma(y) \setminus \{0\})$ and then through π to an element in $C^*\langle y\rangle \subseteq B$. This is clearly a $*$-homomorphism, and maps x to y.

Note that the universal C^*-algebra of generators X with relations R is in general non-unital, but we can add a unit element 1 to the generator set X and the relations $1 = 1^* = 1^2$ and $1x_i = x_i1 = x_i$ for all $x_i \in X$ to make the universal C^*-algebra unital.

2.3. PARTITION THEORY

Having developed the theory of universal C^*-algebras in the previous section, we now turn our attention to the theory of *partitions* – the compact quantum groups we are interested in are so-called *partition* quantum groups (or sometimes, in the terminology of Banica and Speicher, "easy" quantum groups), whose intertwiner spaces are characterised by categories of partitions. Hence, we must first develop the theory of partitions. We follow [FW16].

Definition 2.3.1: Partition

Given $k, l \in \mathbb{Z}_{\geqslant 0}$, consider the set X as the union of the sets $\{1, 2, \ldots, k\}$ and $\{1', 2', \ldots, l'\}$, where we see the primed numbers as distinct objects from the unprimed ones. A *partition* p is then a partition of the set X, i.e. a collection of subsets $V_i \subseteq X$ indexed by some index set I such that each $V_i \neq \emptyset$, $V_i \cap V_j = \emptyset$ for all $i \neq j \in I$, and $\bigcup_{i\in I} V_i = X$. These subsets V_i are called *blocks*, and if such a V_i only contains one point, it is called a *singleton*. The *collection of all partitions with k unprimed and l primed numbers* (for $k, l \in \mathbb{Z}_{\geqslant 0}$) is denoted $\mathcal{P}(k, l)$, and by definition we set $\mathcal{P}(0, 0) := \{\emptyset\}$. The *collection of all partitions* is called $\mathcal{P}$.

We can graphically denote p by a diagram where we draw $\{1, 2, \ldots, k\}$ as separate points in a horizontal line, $\{1', 2', \ldots, l'\}$ on a horizontal line below the previous one, and connect all objects within one V_i for each $i \in I$. This looks like:

Example 2.3.1. An example of a partition with $k = 4$ and $l = 6$ is $\{V_1, V_2, V_3, V_4, V_5\}$ with $V_1 = \{1, 1', 2'\}$, $V_2 = \{5', 6'\}$, $V_3 = \{2, 4'\}$, $V_4 = \{3, 3'\}$ and $V_5 = \{4\}$. Visually, we can draw this as:

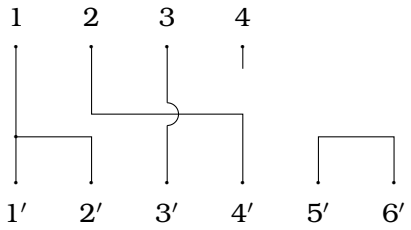


Figure 2.1: The partition $p \in \mathcal{P}(4,6)$ given by $p = \{V_1, V_2, V_3, V_4, V_5\}$ with $V_1 = \{1, 1', 2'\}$, $V_2 = \{5', 6'\}$, $V_3 = \{2, 4'\}$, $V_4 = \{3, 3'\}$ and $V_5 = \{4\}$.

Here, V_i for $1 \leqslant i \leqslant 5$ are called *blocks*. We see that a block can contain more than 2 points (e.g., V_1), it can contain only points from one side (e.g., V_2), it can be a pair of points (e.g., V_2 and V_3), it can be a *singleton* (e.g., V_5), and the blocks are allowed to "intersect" (e.g., V_3 and V_4).

We see that certain blocks may contain both primed and unprimed indices, whilst others only contain primed, or only unprimed integers. We will give them a name: a block V_i containing both a primed and unprimed number is called a *through-block*, and otherwise it is called a *non-through-block*. Denote the number of *blocks* of p by $b(p)$ and the number of *through-blocks* of p by $t(p)$. Finally, we let the $\beta(p)$ denote the number of non-through-blocks of p, so we have the relationship $b(p) = t(p) + \beta(p)$.

Example 2.3.2. Looking at Example 2.3.1, we see that V_1, V_3 and V_4 are *through-blocks*, whilst V_2 and V_5 are *non-through-blocks*. Hence, for this partition p, $b(p) = 5$, $t(p) = 3$ and $\beta(p) = 2$.

We consider two special kinds of partitions: given a partition $p \in \mathcal{P}$, it might happen that all blocks V_i in $p = \{V_1, \ldots, V_r\}$ precisely contain 2 points. In that case, we call p a *pair partition*, and denote the collection of all those pair partitions by $\mathcal{P}_2$, and similar to the partition definition, we set $\mathcal{P}_2(k,l)$ as all pair partitions on k unprimed and l primed numbers. It may happen that, after drawing p as in Example 2.3.1, no strings that belong to different V_i's ever intersect each other. If this is the case, we say that p is a *non-crossing partition*, and we denote the collection of those non-crossing partitions by NC, and similar to above, we define $\mathrm{NC}(k,l)$. Furthermore, we denote the collection of all non-crossing pair partitions by $\mathrm{NC}_2 := \mathrm{NC} \cap \mathcal{P}_2$ and $\mathrm{NC}_2(k,l) := \mathrm{NC}(k,l) \cap \mathcal{P}_2(k,l)$.

Lastly, it will be useful to identify two special partitions:

Definition 2.3.2

Let $| \in P(1,1)$ be the partition $| = \{\{1, 1'\}\}$. We call this the *identity partition*. Let $\uparrow \in \mathcal{P}(0,1)$ be the partition $\uparrow = \{\{1'\}\}$. We call this the *singleton partition*.

2.3.1. OPERATIONS ON PARTITIONS

We define the necessary operations that we can apply to partitions:

Definition 2.3.3: (Diagram) composition of partitions

Define a map $\circ : \mathcal{P}(k,l) \times \mathcal{P}(l,m) \to \mathcal{P}(k,m)$ for $k, l, m \in \mathbb{N}_0$ by $(p,q) \mapsto p \circ q$ as follows: draw the diagram corresponding to the partition p, so with k unprimed numbers on top, and l primed numbers at the bottom. Now draw the diagram corresponding to the partition q by using the l primed numbers from the diagram of p as top layer, and draw the m primed numbers of q at the bottom. Remove the nodes in the centre. Count the number of blocks that are now "floating" in the centre of the diagram without any connection to the upper points of p nor to the lower points of q, and call this number $\mathrm{rl}(q,p)$. Then, remove these blocks. What is left is the diagram of the partition $p \circ q \in P(k,m)$.

Example 2.3.3. An example of the diagram composition of two partitions p and q is as follows:

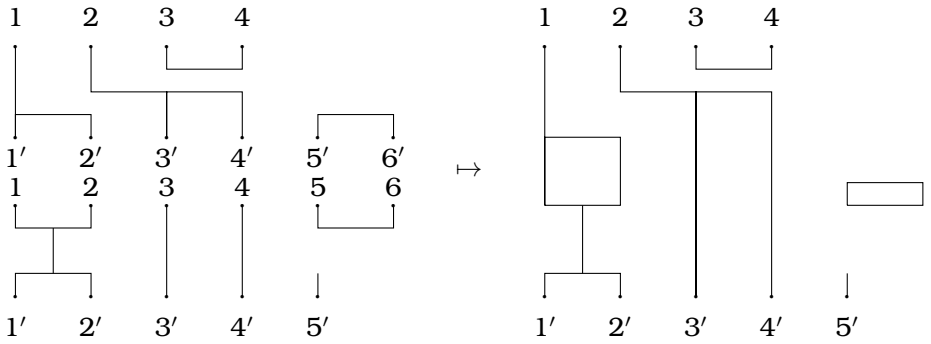


Figure 2.2: The composition of two partitions p and q.

We call the result $p \circ q := \{V_1, V_2, V_3, V_4\}$ with $V_1 = \{1, 1', 2'\}$, $V_2 = \{2, 3', 4'\}$, $V_3 = \{3, 4\}$ and $V_4 = \{5'\}$. We have removed the floating block in the middle, so $\mathrm{rl}(q, p) = 1$.

Definition 2.3.4: Tensor product of partitions

Define a map $\otimes : \mathcal{P}(k, l) \times \mathcal{P}(m, n) \to \mathcal{P}(k + m, l + n)$ for $k, l, m, n \in \mathbb{Z}_{\geqslant 0}$ by $(p, q) \mapsto p \otimes q$ as follows: draw the diagram of p, and then horizontally concatenate it on the right by the diagram of q. Relabel the upper points from 1 to $k+m$, and relabel the lower points from $1'$ to $l' + n'$. This is the diagram of $p \otimes q \in \mathcal{P}(k + m, l + n)$.

Definition 2.3.5: Involution of a partition

Define a map $* : \mathcal{P}(k, l) \to \mathcal{P}(l, k)$ for $k, l \in \mathbb{Z}_{\geqslant 0}$ by $p \mapsto p^*$ as follows: draw the diagram of p, then flip it upside-down and relabel the unprimed numbers to primed numbers and vice-versa. This is the diagram of $p^* \in \mathcal{P}(l, k)$.

Definition 2.3.6: Rotation of a partition

Define maps $\curvearrowright$: $\mathcal{P}(k, l) \to \mathcal{P}(k - 1, l + 1)$ and $\curvearrowleft$: $\mathcal{P}(k, l) \to \mathcal{P}(k - 1, l + 1)$ for $k \in \mathbb{Z}_{>0}$ and $l \in \mathbb{Z}_{\geqslant 0}$ by $p \mapsto p^{\curvearrowright}$ and $p \mapsto p^{\curvearrowleft}$ as follows: draw the diagram of p, and move the left-most upper point labelled 1 to the left of the left-most lower point labelled $1'$, without changing any of the strings in the diagram. Then, relabel the new left-most lower point from 1 to $1'$, and relabel all $x' \to (x + 1)'$ for $x = 1, \ldots, l$. This is the diagram of the partition $p^{\curvearrowleft} \in P(k - 1, l + 1)$. Similarly, moving the right-most upper point labelled k to the right of the right-most lower point labelled l', and relabelling, yields $p^{\curvearrowright} \in P(k - 1, l + 1)$.

Also define maps $\cup : \mathcal{P}(k, l) \to \mathcal{P}(k + 1, l - 1)$ and $\cup : \mathcal{P}(k, l) \to \mathcal{P}(k + 1, l - 1)$ for $k \in \mathbb{Z}_{\geqslant 0}$ and $l \in \mathbb{Z}_{>0}$ by $p \mapsto p\cup$ and $p \mapsto p\cup$ in similar fashion as above: instead of moving the left- or right-most upper point to the left- or right-most lower position (respectively), now move the left-most lower point to the left-most upper point to get the diagram of $p\cup$, or move the right-most lower point to the right-most upper point to get the diagram $p\cup$.

Note that $(p^{\curvearrowright})\cup = p$ and $(p\cup)^{\curvearrowright} = p$ and $(p^{\curvearrowleft})\cup = p$ and $(p\cup)^{\curvearrowleft} = p$ for any $p \in \mathcal{P}(k, l)$ for any $k, l \in \mathbb{Z}_{>0}$, so in this sense one may consider the operations $\curvearrowright$ and $\cup$ to be inverses, and similarly for $\curvearrowleft$ and $\cup$.

Having defined the operations that we can apply to partitions, we can define, analogous to the operator algebraic case, a *symmetric partition* as a partition $p \in \mathcal{P}(k, k)$ for some $k \in \mathbb{N}_0$ that satisfies $p^* = p$. Similarly, a partition $p \in \mathcal{P}(k, k)$ for some $k \in \mathbb{N}_0$ is called *diagrammatically idempotent* if $p \circ p = p$, and we call p *diagrammatically projective* if it is both diagrammatically idempotent and symmetric, i.e. $p^* = p = p \circ p$.

2.3.2. Categories and linear categories of partitions

Definition 2.3.7: Category of partitions

Consider subsets $\mathcal{C}(k,l) \subseteq \mathcal{P}(k,l)$ for each $k,l \in \mathbb{Z}_{\geqslant 0}$, and bundle them into a collection $\mathcal{C} \subseteq \mathcal{P}$. We call $\mathcal{C}$ a *category of partitions* if it is invariant under the *category operations*, which are precisely the *diagram composition, tensor product, involution and four rotations* defined above. Additionally, we demand that it contains the identity partition $| \in \mathcal{C}(1,1) \subseteq \mathcal{P}(1,1)$.

We say that a category of partitions $\mathcal{C}$ is non-crossing if $\mathcal{C}(k,l) \subseteq \mathrm{NC}(k,l)$ for all $k,l \in \mathbb{N}_0$. Furthermore, note that the invariance of a category of partitions $\mathcal{C}$ under the tensor product, together with the demand $| \in \mathcal{C}(1,1)$, directly demands that each $\mathcal{C}(k,k) \subseteq \mathcal{P}(k,k)$ contains the identity $|^{\otimes k} \in \mathcal{C}(k,k)$.

Example 2.3.4. Examples of categories of partitions are $\mathcal{P}$ itself, the category of all pair partitions $\mathcal{P}_2$ (remember that rotations do not alter any strings, so pairs are preserved), NC and NC_2. For the latter two, one may need to convince oneself that the non-crossing property of a partition is invariant under the four rotations defined above.

The *partition quantum groups* studied in this thesis have intertwiner spaces that can be described by such categories of partitions, and it would be beneficial if we can capture linear combinations and compositions of intertwiners in the language of partitions as well. In order to facilitate this, we define the *linear categories of partitions* as introduced by Gromada and Weber [GW19]. Let $\delta > 0$ be a fixed parameter, and let $\mathcal{P}_{\mathrm{lin}}[\delta](k,l)$ as a vector space be defined as $\mathbb{C}[\mathcal{P}(k,l)]$, i.e., the $\mathbb{C}$-vector space whose basis consists of the diagrams of partitions in $\mathcal{P}(k,l)$. Define $\mathcal{P}_{\mathrm{lin}}[\delta] := \cup_{k,l}\mathcal{P}_{\mathrm{lin}}[\delta](k,l)$. We endow $\mathcal{P}_{\mathrm{lin}}[\delta]$ with an algebra structure by defining the composition of $p \in \mathcal{P}(k,l)$ with $q \in \mathcal{P}(l,m)$ by $qp := \delta^{\mathrm{rl}(q,p)}(p \circ q)$ (note the inverted order in qp), where $p \circ q$ is the usual *diagram composition*. Extend this definition linearly. We endow $\mathcal{P}_{\mathrm{lin}}[\delta]$ with a monoidal structure by extending the tensor product $p \otimes q$ linearly, and an involution by extending the involution $p \mapsto p^*$ anti-linearly. In this framework, we define [GW19]:

Definition 2.3.8: Linear category of partitions

Consider for each $k,l \in \mathbb{N}_0$ a *subspace* $\mathcal{C}(k,l) \subseteq \mathcal{P}_{\mathrm{lin}}[\delta](k,l)$, and collect those subspaces in $\mathcal{C} := \cup_{k,l}\mathcal{C}(k,l)$. Assume that the identity partition $| \in \mathcal{C}(1,1)$, the pair partition $\cap \in \mathcal{C}(0,2)$, and assume that $\mathcal{C}$ is closed under the composition, tensor product and involution operations defined on $\mathcal{P}_{\mathrm{lin}}[\delta]$. Then, $\mathcal{C}$ is called a *linear category of partitions*.

Note that a linear category of partitions is automatically closed under rotations [GW19, Lemma 3.1]. Furthermore, a *category of partitions* $\mathcal{C}$ corresponds to what Gromada and Weber call an *easy* linear category of partitions $\mathcal{D}$ if we let $\mathcal{D}(k,l) = \mathrm{span}\ \mathcal{C}(k,l)$, and their corresponding quantum groups are called *easy* (partition) quantum groups.

Example 2.3.5. For each of the categories of partitions $\mathcal{P}_2$, NC and NC_2, we denote by $\mathcal{P}_{2,\mathrm{lin}}[\delta]$, $\mathrm{NC}_{\mathrm{lin}}[\delta]$ and $\mathrm{NC}_{2,\mathrm{lin}}[\delta]$ their respective easy linear categories of partitions with loop parameter δ.

Remark 2.3.1. We call a partition $p \in \mathcal{P}_{\mathrm{lin}}[\delta](k,k)$ for some $k \in \mathbb{N}_0$ *symmetric* if $p^* = p$, *idempotent* if $pp = p$ and *projective* if $p^* = p = pp$. Note that these notions are *not* the same as being *diagrammatically idempotent* or *diagrammatically projective*, where no scalar factor is taken into account! As an example, $p =$  is *diagrammatically projective*, but composing p with itself yields δp, so p is not *projective* in $\mathcal{P}_{\mathrm{lin}}[\delta](2,2)$. Note that we can always normalize a *diagrammatically projective* partition p by $\delta^{-\mathrm{rl}(p,p)}$, such that the partition $\delta^{\mathrm{rl}(p,p)}p$ is projective in $\mathcal{P}_{\mathrm{lin}}[\delta]$.

2.3.3. Intertwiners associated to partitions

In this subsection, we consider a finite-dimensional complex Hilbert space $\mathcal{H} := \mathbb{C}^N$ for some $N \in \mathbb{N}$, with the canonical basis $\{e_1, \ldots, e_N\}$.

Definition 2.3.9: Partition map

Given a partition $p \in \mathcal{P}(k,l)$ for $k,l \in \mathbb{N}_0$. We associate to this partition a linear map $T_p : \mathcal{H}^{\otimes k} \to \mathcal{H}^{\otimes l}$ by

$$T_p\left(e_{i_1} \otimes \cdots \otimes e_{i_k}\right) = \sum_{j_1,\dots,j_l=1}^{n} \delta_p(\boldsymbol{i},\boldsymbol{j}) e_{j_1} \otimes \cdots \otimes e_{j_l}, \tag{2.5}$$

where $\boldsymbol{i} := (i_1,\dots,i_k)$ and $\boldsymbol{j} := (j_1,\dots,j_k)$, and $\delta_p(\boldsymbol{i},\boldsymbol{j}) \in \{0,1\}$ is determined as follows: draw the diagram of p, replace the labels of the upper nodes by $\boldsymbol{i}$, the labels of the lower nodes by $\boldsymbol{j}$, and consider each of the blocks in p. If all blocks in p connect equal indices, then $\delta_p(\boldsymbol{i},\boldsymbol{j}) = 1$. Otherwise, $\delta_p(\boldsymbol{i},\boldsymbol{j}) = 0$. Please note that this does not mean that equal indices must always be in the same block to have $\delta_p(\boldsymbol{i},\boldsymbol{j}) = 1$, as several blocks may be filled with the same indices.

We extend this notion of a partition map linearly to linear combinations of partitions in $\mathcal{P}_{\text{lin}}[N]$, i.e. $\delta_{\alpha p+q}(\boldsymbol{i},\boldsymbol{j}) = \alpha\delta_p(\boldsymbol{i},\boldsymbol{j}) + \delta_q(\boldsymbol{i},\boldsymbol{j})$, and $T_{\alpha p+q} = \alpha T_p + T_q$.

This allows us to see the assignment $p \mapsto T_p$ as a monoidal $*$-homomorphism [GW19, Prop 3.2]:

Theorem 2.3.1

The map $T_\bullet : p \mapsto T_p$ satisfies

- $T_{p\otimes q} = T_p \otimes T_q$,
- $T_{qp} = T_q T_p$ where qp refers to the composition in $\mathcal{P}_{\text{lin}}[N]$ with $\delta := N$,
- $T_{p^*} = T_p^*$.

Note that this notation is *not* consistent with Freslon and Weber [FW16] for two reasons. Firstly, as they do not consider linear categories of partitions, what they call composition is what we have called *diagram composition* (i.e. composition without the scalar factor $\delta^{\text{rl}(q,p)}$) because there is no linear structure on the collection of partitions $\mathcal{P}$. Secondly, they call the partition maps as we have introduced them $\mathring{T}$ (their use of T_p is a normalized version of $\mathring{T}_p$, and this does *not* coincide with our use of T_p). In that case, $\mathring{T}_\bullet$ still satisfies $\mathring{T}_{p\otimes q} = \mathring{T}_p \otimes \mathring{T}_q$ and $\mathring{T}_p^* = \mathring{T}_{p^*}$, but the composition rule becomes $\mathring{T}_{p\circ q} = N^{-\text{rl}(p,q)}\mathring{T}_p\mathring{T}_q$ [FW16, Prop 2.14]. This is an important distinction, because this means that $\mathring{T}_\bullet$ is not a functor, whilst $T_\bullet$ is, and we will invoke category theoretical notions in the following chapters. Of course, all results can be recast into the notation of Freslon and Weber's [FW16] by carefully examining the normalization of the partition maps involved.

Lastly, we cover an important Lemma about the linear independence of these partition maps. Note that this does not depend on the particular normalization of the partition maps involved. We state the lemma slightly differently from [FW16, Lemma 4.16], as a small mistake was discovered in the statement, to which we will return in Lemma 5.1.1.

Lemma 2.3.1: Non-crossing partition maps are linearly independent

[FW16, Lemma 4.16] Consider any category of non-crossing partitions $\mathcal{C}$. If $N \geqslant 4$, then the partition maps $\{T_p : p \in \mathcal{C}(k,l)\}$ are linearly independent for every $k,l \in \mathbb{N}_0$. In the case of $\mathcal{C} = \text{NC}$, the converse also holds.

3

An introduction to quantum Shannon theory

As we investigate information theoretic properties of certain quantum channels arising from the representation theory of compact quantum groups, it is imperative to discuss in detail the information theoretic quantities that we are interested in. This chapter is complementary to the previous chapter. If we are to investigate the properties of certain quantum channels, we should first clearly define what constitutes a *quantum channel*. Afterwards, we translate some classical information theoretical quantities to their quantum counterparts, and introduce the reader to the *minimum output entropy*, which is the quantity that we will give most attention to in this thesis. We follow the work of [NC09, Wil13, Wat18].

3.1. Quantum channels

We will only deal with quantum channels between finite-dimensional Hilbert spaces, as the all the representation spaces that play an integral role in this thesis will be finite-dimensional. This simplifies the theory, and provides a clearer overview.

Let us commence with a short recap of the basics of quantum information theory. In a finite-dimensional Hilbert space $\mathcal{H} = \mathbb{C}^n$, the state of a quantum system is either *pure*, or *mixed* – in the first case, we represent the quantum state by a vector $|\psi\rangle \in \mathcal{H}$ with $\langle\psi|\psi\rangle = \||\psi\|_2^2 = 1$ (actually, we identify *rays* in this complex space, i.e. the unit vectors $|\psi\rangle$ and $|\phi\rangle$ describe the same quantum state if they have the relation $|\phi\rangle = e^{i\varphi}|\psi\rangle$ for some *global phase* $\varphi \in [0, 2\pi)$), and in the latter case, we use a density matrix $\rho \in \mathcal{B}(\mathcal{H}) = M_n(\mathbb{C})$, which is a positive matrix with $\mathrm{Tr}(\rho) = 1$ (see Definition 2.1.2). The set of all mixed states on $\mathcal{H}$ is denoted by $\mathcal{S}(\mathcal{H})$. Note that the set of pure states can be viewed as a subset of the set of mixed states under the identification $\rho_{|\psi\rangle} = |\psi\rangle\langle\psi|$.

To model the transition, manipulation, or evolution of such a quantum system, we use the quantum operation formalism, which is extraordinarily general: it can model *unitary* evolution (which happens in a closed system), but also stochastic changes, and allows for the modelling of a noisy environment to which the quantum system is coupled and with which it therefore exchanges information [NC09, Chap. 8.2]. We start with a quantum state $\rho \in \mathcal{S}(\mathcal{H})$, and after our experiment, end up with a quantum system $\Phi(\rho) \in \mathcal{B}(\mathcal{H}')$ on a finite-dimensional Hilbert space $\mathcal{H}'$. Physically, when a quantum system, originally in an initial state ρ, comes into contact with an environment that is in the state ρ_{E}, the unitary evolution of the *entire* system should be taken into account, i.e. after some time the entire system is described by the state $U(\rho \otimes \rho_{\mathrm{E}})U^*$, where $U \in \mathcal{B}(\mathcal{H} \otimes \mathcal{H}_{\mathrm{E}})$ is a unitary matrix, where $\mathcal{H}_{\mathrm{E}}$ is the (finite-dimensional) Hilbert space of the environment. Afterwards, the quantum system no longer interacts with the environment, and thus, we must trace over the environment to recover the reduced state of our quantum system, $\Phi(\rho) = \mathrm{Tr}_{\mathrm{E}}(U(\rho \otimes \rho_{\mathrm{E}})U^*)$, where Tr_{E} denotes the partial trace over the environment E.

In general, when a linear map $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H}')$ is described by $\Phi(\rho) = \mathrm{Tr}_{\mathrm{E}}(V\rho V^*)$ for some

isometry $V : \mathcal{H} \to \mathcal{H}' \otimes \mathcal{H}_{\mathrm{E}}$, we say that Φ is described in the *Stinespring representation*, and we call V the *Stinespring isometry* or the *Stinespring dilation*.

We can alternatively describe Φ in its *Kraus representation* (also called *operator-sum representation* [NC09, Sect. 8.2.3]), by explicitly tracing out the environment:

$$\Phi(\rho) = \mathrm{Tr}_{\mathrm{E}}(V\rho V^*) = \sum_k (\iota_{\mathcal{H}'} \otimes \langle e_k|)V\rho V^*(\iota_{\mathcal{H}'} \otimes |e_k\rangle) = \sum_k E_k \rho E_k^*, \tag{3.1}$$

where $\{|e_k\rangle\}_{k=1}^{\dim \mathcal{H}_{\mathrm{E}}}$ is an orthonormal basis of $\mathcal{H}_{\mathrm{E}}$, $\iota_{\mathcal{H}'}$ is the identity matrix on $\mathcal{H}'$, and where we put $E_k := (\iota_{\mathcal{H}'} \otimes \langle e_k|)\, V$. The operators E_k are called the *operation elements* or *Kraus operators* associated to the quantum channel Φ.

Finally, we can describe the action of Φ by its *Choi matrix* $J(\Phi)$, defined as $J(\Phi) = \sum_{ij} \Phi(e_{ij}) \otimes e_{ij} \in \mathcal{B}(\mathcal{H}') \otimes \mathcal{B}(\mathcal{H})$, where $\{e_{ij}\}_{i,j=1}^n$ is the canonical basis of $M_n(\mathbb{C}) \simeq \mathcal{B}(\mathcal{H})$ with $n = \dim \mathcal{H}$. Namely, we have [1]

$$\Phi(\rho) = \mathrm{Tr}_{\mathcal{H}} \left(J(\Phi) \left(\iota_{\mathcal{H}'} \otimes \rho^t \right) \right). \tag{3.2}$$

Remarkably, these descriptions are equivalent, and they match the axiomatic description of a *quantum channel*. Given a linear map $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H}')$, we call Φ trace-preserving (abbreviated "TP") if $\mathrm{Tr}(\Phi(\rho)) = \mathrm{Tr}(\rho)$ for any $\rho \in \mathcal{B}(\mathcal{H})$. We call Φ completely positive (abbreviated "CP") if for *any* auxiliary finite-dimensional Hilbert space $\mathcal{H}_{\mathrm{A}}$, the matrix $(\iota_{\mathrm{A}} \otimes \Phi)(\sigma)$ is positive for any positive matrix $\sigma \in \mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}$, where ι_{A} is the identity operator on $\mathcal{H}_{\mathrm{A}}$. We say Φ is CPTP iff it is CP and TP, and in this case we call Φ a *quantum channel*.

Theorem 3.1.1: Equivalent descriptions of quantum channels

Given a linear map $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H}')$. Then the following are equivalent:

1. Φ is completely positive and trace preserving.
2. We can find a finite-dimensional Hilbert space $\mathcal{H}_{\mathrm{E}}$ and a *Stinespring isometry* $V : \mathcal{H} \to \mathcal{H}' \otimes \mathcal{H}_{\mathrm{E}}$ such that we have $\Phi(\rho) = \mathrm{Tr}_{\mathrm{E}}(V\rho V^*)$ for any $\rho \in \mathcal{B}(\mathcal{H})$, i.e. a *Stinespring representation* of Φ exists.
3. We can find a finite set of *Kraus operators* $\{E_i\}_{i=1}^d$ for some $d \in \mathbb{N}$ with matrices $E_i : \mathcal{H} \to \mathcal{H}'$ such that $\Phi(\rho) = \sum_i E_i \rho E_i^*$, and $\sum_i E_i^* E_i = \iota_{\mathcal{H}}$, i.e. a *Kraus representation* of Φ exists.
4. The Choi matrix $J(\Phi) = \sum_{ij} \Phi(e_{ij}) \otimes e_{ij} \in \mathcal{B}(\mathcal{H}' \otimes \mathcal{H})$ is *positive* and we have $\mathrm{Tr}_{\mathcal{H}'} J(\Phi) = \iota_{\mathcal{H}}$.

Furthermore, in statement 2 we can find an environment $\mathcal{H}_{\mathrm{E}}$ with $\dim \mathcal{H}_{\mathrm{E}} = \mathrm{rank}(J(\Phi))$, and in statement 3 we can find a finite set of Kraus operators with $d = \mathrm{rank}(J(\Phi))$.

Proof. See [NC09, Thm 8.1], [Sti55], and for a detailed treatment also see [Wat18, Thm. 2.22 up to Corol. 2.27]. □

Note that the representations of Φ are not unique: for example, two sets of Kraus operators $\{E_i\}_{i=1}^m$ and $\{F_j\}_{j=1}^n$ may satisfy $\sum_i E_i \rho E_i^* = \sum_j F_j \rho F_j^*$ for all $\rho \in \mathcal{B}(\mathcal{H})$. In this case, however, if we append zero operators to whichever list of operators is shorter, we may assume $n = m$, and in this case the operators must be related by an $m \times m$ unitary u, i.e. $E_i = \sum_j u_{ij} F_j$ [NC09, Thm 8.2]. Similarly, if we have two Stinespring representations $(V, \mathcal{H}_{\mathrm{E}})$ and $(\tilde{V}, \mathcal{H}_{\tilde{\mathrm{E}}})$ such that $\mathrm{Tr}_{\mathrm{E}}(V\rho V^*) = \mathrm{Tr}_{\tilde{\mathrm{E}}}(\tilde{V}\rho\tilde{V}^*)$ for all $\rho \in \mathcal{B}(\mathcal{H})$, without loss of generality we can assume $\dim \mathcal{H}_{\mathrm{E}} \leqslant \dim \mathcal{H}_{\tilde{\mathrm{E}}}$, and we can find an isometry $\tilde{U} : \mathcal{H}' \otimes \mathcal{H}_{\mathrm{E}} \to \mathcal{H}' \otimes \mathcal{H}_{\tilde{\mathrm{E}}}$ such that $\tilde{U}V = \tilde{V}$ [KSW08].

[1] This follows from the fact that $J(\Phi)\left(\iota_{\mathcal{H}'} \otimes \rho^t\right) = \sum_{ij} \Phi(e_{ij}) \otimes (e_{ij}\rho^t)$, and then tracing out $\mathcal{H}$ yields $\sum_{ijk} \Phi(e_{ij}) \otimes \langle k|\, e_{ij}\rho^t\, |k\rangle = \sum_{ij} \Phi(e_{ij}) \otimes \langle j|\, \rho^t\, |i\rangle = \sum_{ij} \Phi(e_{ij}) \otimes \rho_{ij} = \Phi(\rho)$.

As a consequence of the previous theorem, we can find a quantum channel that is closely related to our original quantum channel Φ with Stinespring isometry $V : \mathcal{H} \to \mathcal{H}' \otimes \mathcal{H}_E$ if we choose to trace out Bob's subsystem instead of the environment E. This yields the definition of the so-called *complementary channel*:

Definition 3.1.1: Complementary quantum channel

Given a quantum channel $\Phi : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$ with a Stinespring isometry $V : \mathcal{H}_A \to \mathcal{H}_B \otimes \mathcal{H}_E$ for some environment Hilbert space $\mathcal{H}_E$. The *complementary quantum channel* Φ^c is defined by tracing out Bob's subsystem instead of the environment:

$$\Phi^c : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_E) \quad : \quad \rho \mapsto \mathrm{Tr}_B\left(V \rho V^*\right). \tag{3.3}$$

As the Stinespring representation of a quantum channel is not unique, two Stinespring representations $(V, \mathcal{H}_E)$ and $(\tilde{V}, \mathcal{H}_{\tilde{E}})$ of Φ give rise to two complementary quantum channels Φ^c_E and $\Phi^c_{\tilde{E}}$, but they are equivalent in the following sense: a partial isometry $W : \mathcal{H}_E \to \mathcal{H}_{\tilde{E}}$ exists such that $\Phi^c_{\tilde{E}}(\rho) = W \Phi^c_E(\rho) W^*$ and similarly $\Phi^c_E(\rho) = W^* \Phi^c_{\tilde{E}}(\rho) W$ [Hol07].

Having described different variants of representations of quantum channels, we can continue investigating some of their information theoretic properties.

3.2. CLASSICAL INFORMATION THEORY

Let us sketch the background of classical information theory, which was adapted in the last part of the 20th century into *quantum Shannon theory* to accomodate for quantum channels. We briefly discuss the concepts that play an integral role in classical information theory and find their way into quantum Shannon theory. We mainly follow [Wil13].

3.2.1. INFORMATION CONTENT, ENTROPY, AND COMPRESSION

Given a discrete random variable X whose realizations are *letters* x in a finite *alphabet* $\mathcal{X} := \{a_1, \ldots, a_{|\mathcal{X}|}\}$. We denote the probability mass function of X by $p_X(x)$. Say we do an experiment with a random generator that generates an x_j according to the probability distribution p_X, we measure our *surprise* at finding the outcome x by $i(x) = -\log p_X(x)$, which indeed monotonously decreases to 0 as $p_X(x)$ grows to 1. The quantity $i(x)$ is sometimes referred to as the *information content* of the symbol x. Note that if our source produces two symbols, with corresponding random variables X_1 and X_2, and if we assume that the source is *memoryless*, i.e. X_1 and X_2 are independent, then $p_{X_1,X_2}(x_1, x_2) = p_{X_1}(x_1) p_{X_2}(x_2)$, and we have

$$i(x_1, x_2) = -\log(p_{X_1,X_2}(x_1, x_2)) = i(x_1) + i(x_2), \tag{3.4}$$

so the information content is *additive*. The *expected information content* is then $H(X) := \mathbb{E}_X(i(X))$, where $i(X) := -\log p_X(X)$ is a random variable, and we see that

$$H(X) = -\sum_{x \in \mathcal{X}} p_X(x) \log p_X(x), \tag{3.5}$$

which we also call the *entropy* of the information source. Here, the convention is $0 \cdot \log(0) := 0$ as $\lim_{\varepsilon \downarrow 0} \varepsilon \log(\varepsilon) = 0$. We have $0 \leqslant H(X) \leqslant \log |\mathcal{X}|$. [2]

Assume we wish to send a *word* or *block* $x^n := x_1 x_2 \ldots x_n$ from Alice to Bob over a noiseless channel, where each $x_i \in \mathcal{X}$ is a realization of the random variable X_i, which themselves are independent and identically distributed (i.i.d.). We assume Alice and Bob can beforehand agree on a *codebook*. Naively, Alice would need to use $n \log |\mathcal{X}|$ bits to uniquely encode all $|\mathcal{X}|^n$ possible words. Assume that the use of the communication channel is costly, so we wish to minimize the amount of bits Alice needs to send to Bob to get her message across – so we wish

[2] $H(X) \geqslant 0$ follows from the non-negativity of $x \mapsto -x \log x$ on $[0, 1]$, and for the upper bound, use the fact that $H(X)$ is concave and consider the Lagrangian $\mathcal{L} := H(X) + \lambda(\sum_x p_X(x) - 1)$, then differentiate it to the "variables" $p_X(x)$ to find $\partial_{p_X(x)} \mathcal{L} = \lambda - 1 - \log p_X(x)$, equate it to 0 and solve for $p_X(x)$ to find $p_X(x) = \exp(\lambda - 1)$, and we see that we must have a uniform distribution. It is clear that the uniform distribution has entropy $\log |\mathcal{X}|$.

to *compress* her message. Can we do better than sending $n \log |\mathcal{X}|$ bits? Sure, for example, we might consider sending bitstrings of variable length, where the more likely messages are encoded in shorter strings, and the less likely messages are encoded in longer strings – this way, the *expected* length of an encoded message is shorter than $n \log |\mathcal{X}|$. Alternatively, we could allow for schemes that yield the correct message on Bob's side with a very high probability, and trade the certainty of a decoding success of Bob's side with a much smaller expected length of an encoded message. The latter description is the starting point from which Shannon proves his *compression theorem*.

We define the *compression rate* as the ratio between the number of noiseless channel bits we need to send and the number of source symbols that we encode in this message. For a compression rate of R, Alice must encode her message from $x^n \in \mathcal{X}^n$ to a string of bits in $\{0,1\}^{nR}$, and therefore uses the communication channel nR times – let this process be denoted by E. Bob decodes this bitstring in $\{0,1\}^{nR}$ to a message in $\mathcal{X}^n$, let this map be denoted by D. We call such a scheme an (n, R, ε)-*source code* if $p(\text{error}) = \mathbb{P}\left((D \circ E)(X^n) \neq X^n\right) \leqslant \varepsilon$. If such a source code exists *asymptotically*, we call the compression rate R *achievable*: R is called *achievable for* X if for all $\varepsilon \in (0,1)$ and $\delta > 0$, there exists an $n \in \mathbb{N}$ such that an $(n, R+\delta, \varepsilon)$-source code exists. The compression rate $H(X)$ turns out to be achievable. Shannon's *data compression theorem* tells us that this is the best source code with respect to the compression rate:

Theorem 3.2.1: Shannon compression, noiseless channel

Given an information source described by the discrete random variable X, then communication over a noiseless channel has an achievable rate that satisfies

$$\inf\{R \ : \ R \text{ is achievable for } X\} = H(X). \tag{3.6}$$

3.2.2. CLASSICAL CHANNELS, MUTUAL INFORMATION AND CLASSICAL CAPACITY

Although Shannon's compression theorem for noiseless channels definitively answers the question how well one can compress information when one has a *noiseless* channel, in practice, channels are not noiseless.In this framework, Alice and Bob again meet up beforehand to discuss a *channel code* $\mathcal{C} = \{x^n(m)\}_{m \in \mathcal{M}}$, where $x^n(m)$ is a codeword Alice will send to Bob over the channel, using the channel n times, where m is any of the possible messages in $\mathcal{M}$ that Alice can send to Bob. We again assume that the codewords $x^n(m)$ are strings of letters x_i in some finite alphabet $\mathcal{X} = \{a_1, \ldots, a_{|\mathcal{X}|}\}$. The encoding step is modeled by a map $E^n : \mathcal{M} \to \mathcal{X}^n$. We model the input of the channel as the random variable X^n associated with the sequence x^n. As we have a noisy channel in this case, the output of the channel is not simply x^n, but we model it instead by another random variable Y^n associated to the sequence $y^n = y_1 \cdots y_n$, where each y_i is a letter in the alphabet $\mathcal{Y}$, which may be different than the input alphabet $\mathcal{X}$. The decoding step is then modeled by a map $D^n : \mathcal{Y}^n \to \mathcal{M}$. A *classical channel* $\mathcal{N}$ is a function that maps the sequence x^n that Alice presents to the channel to the random variable Y^n that Bob receives. Thus, the entire chain from the message m to the message $\hat{m}$ that Bob decodes is $\hat{m} = (D^n \circ \mathcal{N}^n \circ E^n)(m)$. We discuss the *discrete memoryless channel*, where the random variables are i.i.d., which allows us to write $p_{Y^n|X^n}(y^n|x^n) = \prod_{i=1}^n p_{Y|X}(y_i|x_i)$.

Given such a channel code $\mathcal{C}$, we let $p_e(\mathcal{C}, m)$ be the probability that something goes wrong when Alice attempts to send the message m using the code $\mathcal{C}$. The average error over all messages m is denoted $\overline{p_e}(\mathcal{C})$, and the maximal error over all messages m is denoted $p_e^*(\mathcal{C})$.

As there are $|\mathcal{M}|$ messages in total, we need at least $\log |\mathcal{M}|$ bits to uniquely identify them, and we define the *rate* R as the ratio of the number of bits in the message divided by the number of channel uses – in this case, we have $R = \frac{1}{n} \log |\mathcal{M}|$. We call $\mathcal{C}$ a (n, R, ε)-channel code if its maximal error satisfies $p_e^*(\mathcal{C}) \leqslant \varepsilon$. We say that a communication rate R is *achievable* for a channel $\mathcal{N}$ if, for all $\varepsilon \in (0,1)$ and $\delta > 0$, a sufficiently large $n \in \mathbb{N}$ exists such that a $(n, R-\delta, \varepsilon)$-channel code $\mathcal{C}$ exists. We define the *channel capacity* $C(\mathcal{N})$ as the supremum over all achievable communication rates R for $\mathcal{N}$. To state Shannon's channel capacity theorem, we need to adapt the concept of information content to the context of a noisy channel.

If we assume that Alice posses a random variable X, and Bob a random variable Y that could be correlated with X, the information content of a symbol x from Alice is affected by Bob's knowledge of the symbol y from his random variable Y. To formalize this, we define the *conditional information content* $i(x|y)$ as $i(x|y) = -\log p_{X|Y}(x|y)$, and similarly $H(X|Y = y) = \mathbb{E}_{X|Y=y}(i(X|y))$. The *conditional entropy* is then defined as $H(X|Y) := \mathbb{E}_{X|Y}(i(X|Y)) = \sum_y p_Y(y)H(X|Y = y)$. As expected, we have $H(X) \geqslant H(X|Y) \geqslant 0$, and the first inequality is an equality iff $X \perp Y$. Hence, we define their difference as the *mutual information* $I(X;Y) := H(X) - H(X|Y)$, which measures how much the uncertainty in X is reduced by knowing Y.

We can now state:

Theorem 3.2.2: Shannon's discrete memoryless channel capacity

Given a noisy channel $\mathcal{N}$ modelled by the conditional probability distribution $p_{Y|X}(y|x)$, the capacity $C(\mathcal{N})$ of $\mathcal{N}$ satisfies:

$$C(\mathcal{N}) := \sup\{R \ : \ R \text{ is achievable for } \mathcal{N}\} = I(\mathcal{N}) := \max_{p_X} I(X;Y), \tag{3.7}$$

where we call $I(\mathcal{N})$ the maximum mutual information of channel $\mathcal{N}$, and the maximum is over all possible discrete probability distributions p_X.

3.3. SENDING CLASSICAL INFORMATION OVER QUANTUM CHANNELS

Having discussed some of the fundamentals of classical information theory, we can now attempt to translate these notions to the quantum world. In a sense, the quantum Shannon theory is even richer than the classical information theory, because we simply have more options – for example, we can choose to send classical or quantum information over a quantum channel. Furthermore, we can make different analyses based on what kind of resources are available to Alice and Bob, we could for example assume that they have a (noiseless, or perhaps noisy) classical channel alongside their quantum channel, or they have a classical communication channel and some shared entangled states (i.e. entanglement-assisted classical communication). We restrict ourselves to the case where Alice wants to send classical information to Bob over a quantum channel, for an overview see for example [Wil13].

3.3.1. CLASSICAL INFORMATION OVER A NOISELESS QUANTUM CHANNEL

We remind the reader of the POVM formalism: if an observer is given a density matrix $\rho \in \mathcal{S}(\mathcal{H})$, and wishes to perform a measurement on ρ, and is only interested about the probabilities of different outcomes and not the post-measurement state, we can describe this measurement with a *positive operator-valued measure* (POVM), which is a set of operators $\{F_i\}_i$, where we only demand that each $F_i \in \mathcal{B}(\mathcal{H})$ is a positive semi-definite matrix, and $\sum_i F_i = \iota_{\mathcal{H}}$. The probability of obtaining outcome i is obtained by computing $\text{Tr}(F_i\rho)$. Note that the POVM formalism includes the *projective measurements* (PVMs) if we pick $F_i = \Pi_i$ where the Π_i are mutually orthogonal projections that sum to the identity.

With this formalism, we can see that classical information can be carried by a quantum system, showing that the quantum channel formalism is in this sense broader than the classical channel formalism. In this case, Alice has an ensemble $\mathcal{E}$ of density matrices $\{\rho_x\}_{x\in\mathcal{X}}$ and a discrete probability distribution p_X. On Bob's side, the random variable Y is modelled by the POVM $\{F_y\}_{y\in\mathcal{Y}}$, and we see that $p_{Y|X}(y|x) = \text{Tr}(F_y\rho_x)$. Analogous to the classical case, if Bob wishes to learn as much information about Alice's random variable X as possible, we should try to maximize their mutual information $I(X;Y)$. Note that in this case, Bob can pick which POVM he wishes to use, so we optimize over those, and define this as the *accessible information of the ensemble* $\mathcal{E}$, i.e. $I_{\text{acc}}(\mathcal{E}) = \max_{\{F_y\}} I(X;Y)$.

Let us further investigate the possibility of communicating classical information over a *noiseless* quantum channel. We identify four stages in this scheme:

State preparation: Before we *encode* our message, we must prepare the quantum system we wish to send in the right state – we call this *state preparation*. We can assume that we have an ensemble of quantum states $\{|\psi(x)\rangle\}_{x\in\mathcal{X}}$ and a probability distribution $p_X(x)$ for the discrete random variable X with realizations x in the alphabet $\mathcal{X}$. Assume these quantum states live in the Hilbert space $\mathcal{H}_\mathrm{A}$. The word $x^n = x_1x_2\cdots x_n$ is then encoded in the quantum state

$$|\psi(x^n)\rangle := |\psi(x_1)\rangle \otimes \cdots \otimes |\psi(x_n)\rangle \in \mathcal{H}_\mathrm{A} \otimes \cdots \otimes \mathcal{H}_\mathrm{A} := \mathcal{H}_\mathrm{A}^{\otimes n}. \tag{3.8}$$

Similarly, to an observer who does not know the actual sequence x^n, the quantum state is modeled as

$$\rho^{\otimes n} := \rho \otimes \cdots \otimes \rho \quad \text{, where } \rho := \sum_x p_X(x)\,|\psi(x)\rangle\langle\psi(x)| \in \mathcal{S}(\mathcal{H}_\mathrm{A}). \tag{3.9}$$

Encoding: After successfully preparing our quantum system, we use an *encoding quantum channel* $\mathcal{E}$ that maps $\mathcal{B}(\mathcal{H}_\mathrm{A}^{\otimes n})$ to $\mathcal{B}(\mathcal{H}_\mathrm{C})$, where $\dim\mathcal{H}_\mathrm{C} = 2^{nR}$. We call R the *rate of compression* again, and we note that $R = \frac{1}{n}\dim\mathcal{H}_\mathrm{C}$.

Transmission: After encoding the quantum state to a quantum state living on $\mathcal{H}_\mathrm{C}$, Alice sends this state to Bob using a noiseless qubit channel nR times.

Decoding: Bob then applies the *decoding quantum channel* $\mathcal{D}$ that maps $\mathcal{B}(\mathcal{H}_\mathrm{C})$ to $\mathcal{B}(\mathcal{H}_\mathrm{B}^{\otimes n})$, where $\mathcal{H}_\mathrm{B} \simeq \mathcal{H}_\mathrm{A}$. We compare the input and the output using the normalized trace distance, and say it is ϵ-close if:

$$\tfrac{1}{2}\left\| |\psi(x^n)\rangle\langle\psi(x^n)| - (\mathcal{D}\circ\mathcal{E})\left(|\psi(x^n)\rangle\langle\psi(x^n)|\right)\right\|_1 \leqslant \varepsilon. \tag{3.10}$$

We can summarize this as follows graphically:

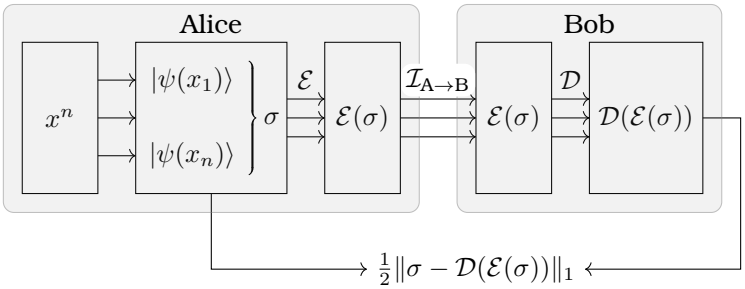


Figure 3.1: A general protocol for sending classical data over a quantum channel. Alice encodes her codeword x^n in $|\psi(x^n)\rangle := |\psi(x_1)\rangle \otimes \cdots \otimes |\psi(x_n)\rangle$, where we abbreviate $\sigma := |\psi(x^n)\rangle\langle\psi(x^n)|$. Then, she *encodes* those states using her *encoding quantum channel* $\mathcal{E}$. She then uses a *noiseless channel* $\mathcal{I}_{\mathrm{A}\to\mathrm{B}}$ to transfer the quantum state $\mathcal{E}(\sigma)$ to Bob. Afterwards, Bob applies his *decoding quantum channel* $\mathcal{D}$. To measure the accuracy of this scheme, the normalized trace distance $\frac{1}{2}\|\sigma - \mathcal{D}(\mathcal{E}(\sigma))\|_1$ is used.

Call a quantum compression rate R *achievable* if for all $\varepsilon \in (0,1)$ and $\delta > 0$, we can find a sufficiently large $n \in \mathbb{N}$ such that an $(n, R+\delta, \varepsilon)$-quantum compression code exists. A similar result to Shannon's noiseless compression theorem exists, but it requires a quantum mechanical version of the *entropy* of the random variable X. We remind the reader that the entropy $H(X)$ of a discrete random variable is defined as $H(X) = -\sum_x p_X(x)\log p_X(x)$. In the quantum framework, instead of an alphabet $\mathcal{X}$ consisting of letters x_i, we have a set of quantum states $\{|\psi_x\rangle\}_x \subset \mathcal{H}$ we can send for some Hilbert space $\mathcal{H}$, each with a probability $p_X(x)$. To an outsider without any other knowledge of the system, this quantum state is modeled by the density matrix $\rho = \sum_x p_X(x)\,|\psi_x\rangle\langle\psi_x|$. Assume for a moment that all states $|\psi_x\rangle$ are mutually orthogonal, then the previous expression for ρ is actually its eigenvalue decomposition, which allows us to define the *von Neumann entropy of the quantum state* ρ:

$$H(\rho) = -\sum_x p_X(x)\log p_X(x) \quad , \quad \text{with eigenval. decomp. } \rho = \sum_x p_X(x)\,|\psi_x\rangle\langle\psi_x|\,. \tag{3.11}$$

So in this case, $H(\rho)$ is equal to the entropy $H(X)$ of the random variable X. By *continuous functional calculus*, we can also write this as $H(\rho) = -\mathrm{Tr}(\rho \log \rho)$, where $\log \rho$ is the matrix logarithm. In literature, one also frequently encounters the notation $H(\mathrm{A})_\rho$ where ρ is the state of the quantum system labelled A, and thus $\rho \in \mathcal{S}(\mathcal{H}_\mathrm{A})$.

Note that one should be careful – if the quantum states in the set $\{|\psi_x\rangle\}_x$ are not mutually orthogonal, $\rho = \sum_x p_X(x) |\psi_x\rangle\langle\psi_x|$ is still a valid density matrix, we still define $H(\rho) := -\mathrm{Tr}(\rho \log \rho)$, but in this case we do *not* have $H(\rho) = H(X)$, as the values $p_X(x)$ no longer have to be the eigenvalues of ρ. In this case, we instead have $H(X) \geqslant H(\rho)$ [Wil13, Ex. 11.9.3].

With this definition in mind, we can state:

Theorem 3.3.1: Schumacher quantum data compression

Given an ensemble $\{|\psi(x)\rangle\}_{x\in\mathcal{X}}$ and a discrete probability distribution $p_X(x)$, then the *quantum data compression limit of* ρ, that is, the infimum over all achievable quantum compression rates for $\rho = \sum_x p_X(x) |\psi(x)\rangle\langle\psi(x)|$, is equal to the von Neumann entropy $H(\rho)$.

Lastly, let us introduce the quantum analogs of the *conditional entropy*. Assume we have two quantum systems, labelled A and B, whose Hilbert spaces are $\mathcal{H}_\mathrm{A}$ and $\mathcal{H}_\mathrm{B}$, respectively. Assume we have a density matrix $\rho_{\mathrm{AB}} \in \mathcal{S}(\mathcal{H}_\mathrm{A} \otimes \mathcal{H}_\mathrm{B})$. We use the notation $H(\mathrm{AB})_\rho$ to denote the von Neumann entropy $H(\rho)$. We define the *marginal entropy* $H(\mathrm{B})_\rho$ as the von Neumann entropy $H(\sigma)$ for $\sigma := \mathrm{Tr}_\mathrm{A}(\rho_{AB})$. The *conditional entropy* is then defined as $H(\mathrm{A}|\mathrm{B})_\rho := H(\mathrm{AB})_\rho - H(\mathrm{B})_\rho$. Just like the classical case, we have the inequality $H(\mathrm{A})_\rho \geqslant H(\mathrm{A}|\mathrm{B})_\rho$. However, in stark contrast to the classical case, the conditional entropy can be negative! Let Alice (A) and Bob (B) share a Bell pair of qubits for example, then the state of the system AB is pure, hence the entropy is 0, but the marginal state that Bob sees, by tracing out Alice's state, is a maximally mixed state, so $H(\mathrm{B})_\rho = 1$, so $H(\mathrm{A}|\mathrm{B})_\rho = -1$. To quantize this phenomenon, we define the *coherent information* $I(\mathrm{A}\rangle\mathrm{B})_\rho := H(\mathrm{B})_\rho - H(\mathrm{AB})_\rho = -H(\mathrm{A}|\mathrm{B})_\rho$. The quantum analog of the *mutual information* is $I(\mathrm{A};\mathrm{B})_\rho := H(\mathrm{A})_\rho - H(\mathrm{A}|\mathrm{B})_\rho$, and just like the classical case, $I(\mathrm{A};\mathrm{B})_\rho = I(\mathrm{B};\mathrm{A})_\rho$.

3.3.2. CLASSICAL INFORMATION OVER A NOISY QUANTUM CHANNEL

In the previous exposition, we assumed that Alice sends her quantum system to Bob without any loss or noise, and we got to Schumacher's quantum data compression theorem, which mirror's Shannon source coding theorem in the classical case. If we wish to model the quantum communication between Alice and Bob as a *noisy* channel, the analysis becomes more involved. An important issue arises from the fact that we cannot reasonably assume that the repeated uses of the same quantum channel are independent of each other, because we can entangle the input states. If we forget about this for a second, we could device a scheme that is analogous to the classical case: Alice has a message $m \in \mathcal{M}$ she wishes to send to Bob. She encodes it in a codeword $x^n(m)$, and now creates the "quantum codeword" $\rho_{x^n(m)} := \rho_{x_1(m)} \otimes \cdots \otimes \rho_{x_n(m)}$, where $\{\rho_x\}_{x\in\mathcal{X}}$ are the possible density matrices she inputs into her channel $\mathcal{N}$. Bob decodes this message using a POVM $\{F_y\}_{y\in\mathcal{Y}}$, and because the input state is a product state, one quickly realizes that $p_{Y^n|X^n}(y^n|x^n) = \prod_{i=1}^n \mathrm{Tr}(F_{y_i}\mathcal{N}(\rho_{x_i}))$, which closely mirrors the classical case where we have many independent and identically distributed uses of a classical channel. The optimal rate that Alice and Bob can communicate is then defined as the *accessible information of* $\mathcal{N}$ defined by $I_\mathrm{acc}(\mathcal{N}) := \max I(X;Y)$, where we optimize over all discrete distributions p_X, the ensemble $\{\rho_x\}_{x\in\mathcal{X}}$, and the POVM $\{F_y\}_{y\in\mathcal{Y}}$.

However, what if Alice uses her quantum channel n times, and *entangles* her input over those n uses? That is, Alice's input is no longer a product of n density matrices, but instead a general $\rho^m_{\mathrm{A}^n} \in \mathcal{B}(\mathcal{H}_\mathrm{A}^{\otimes n})$ for each message $m \in \mathcal{M}$. In this scheme, Bob receives $\mathcal{N}^{\otimes n}(\rho^m_{\mathrm{A}^n})$, and decodes this message with a POVM $\{F_m\}_{m\in\mathcal{M}}$. The probability that he makes an error on message m is then

$$p_e(m) = \mathrm{Tr}\left((\iota_{\mathcal{H}_\mathrm{B}^{\otimes n}} - F_m)\mathcal{N}^{\otimes n}(\rho^m_{\mathrm{A}^n})\right), \tag{3.12}$$

and we define $p_e^* = \max_{m\in\mathcal{M}} p_e(m)$. We say this code has an ϵ error if $p_e^* \leqslant \epsilon$ for some $\epsilon \in [0,1]$. The rate of this code is $C := \frac{1}{n}\log|\mathcal{M}|$, and we call this scheme a (n, C, ε)-code. Analogous to the classical situation, a rate C is *achievable* if for all $\varepsilon \in (0,1)$ and for all $\delta > 0$, there exists a large $n \in \mathbb{N}$ such that an $(n, C-\delta, \varepsilon)$-code exists. We then define the *classical capacity* $C(\mathcal{N})$ as the supremum over all achievable rates.

In this situation, we would need to *regularize* the accessible information $I_{\text{acc}}(\mathcal{N})$, i.e. $I_{\text{reg}}(\mathcal{N}) = \lim_{n\to\infty}\frac{1}{n}I_{\text{acc}}(\mathcal{N}^{\otimes n})$, which is in general intractable. However, we can be upper bound the accessible information by the so-called *Holevo information* $\chi(\mathcal{N}) := \max_{\rho_{\text{XA}}} I(\text{X};\text{B})_\rho$, where $I(\text{X};\text{B})_\rho$ is the mutual information of the classical-quantum state $\rho_{\text{XB}} := \sum_x p_X(x)\,|x\rangle\langle x|\otimes \mathcal{N}(\rho_x)$, which is the state Bob receives if Alice decides to send a classical-quantum state $\sum_x p_X(x)\,|x\rangle\langle x|\otimes\rho_x$. The celebrated Holevo-Schumacher-Westmoreland theorem tells us that the regularized Holevo information is not only an achievable rate, but also the best possible one:

Theorem 3.3.2: Holevo-Schumacher-Westmoreland

Given a quantum channel $\mathcal{N}$,the *classical capacity* $C(\mathcal{N})$ can be computed as:

$$C(\mathcal{N}) = \chi_{\text{reg}}(\mathcal{N}) \quad \text{, where} \quad \chi_{\text{reg}}(\mathcal{N}) := \lim_{n\to\infty}\frac{1}{n}\chi(\mathcal{N}^{\otimes n}). \tag{3.13}$$

We call the quantity χ_{reg} the *regularization* of the Holevo information χ.

Lastly, let us rewrite the Holevo information of a quantum channel $\mathcal{N}$ to a slightly different form. We recall that χ was defined as $\chi(\mathcal{N}) := \max_{\rho_{\text{XA}}} I(\text{X};\text{B})_\rho$. Using $\rho_{\text{XB}} := \sum_x p_X(x)\,|x\rangle\langle x|\otimes \mathcal{N}(\rho_x)$, and recalling that $I(\text{X};\text{B})_\rho := H(\text{X})_\rho - H(\text{X}|\text{B})_\rho$, we can calculate $H(\text{X})_\rho = H(X) = -\sum_x p_X(x)\log p_X(x)$, and

$$H(B)_\rho = H(\text{Tr}_X(\rho_{\text{XB}})) = H(\sum_x p_X(x)\mathcal{N}(\rho_x)) = H(\mathcal{N}(\sum_x p_X(x)\rho_x)). \tag{3.14}$$

Furthermore, we can explicitly calculate $H(\text{XB})_\rho$, see for example [Wil13, Thm. 11.2.2]: we have $H(\text{XB})_\rho = H(X) + \sum_x p_X(x)H(\mathcal{N}(\rho_x))$. Thus, we find:

$$\chi(\mathcal{N}) = \max_{\rho_{\text{XA}}}\left(H(\mathcal{N}(\sum_x p_X(x)\rho_x)) - \sum_x p_X(x)H(\mathcal{N}(\rho_x))\right). \tag{3.15}$$

This expression of the Holevo information is sometimes used to *define* $\chi(\mathcal{N})$, such as in [BCLY20].

3.4. Additivity conjectures for Holevo information and Minimum Output Entropy

Note that the Holevo-Schumacher-Westmoreland theorem features the *regularized* Holevo information χ_{reg}, which is for all practical purposes impossible to calculate directly for a general channel $\mathcal{N}$. To further complicate matters, some channels exhibit *additive* Holevo information, i.e. $\chi(\mathcal{N}^{\otimes n}) = n\chi(\mathcal{N})$ (and in those cases, $\chi_{\text{reg}}(\mathcal{N}) = \chi(\mathcal{N})$), but until recently it was not known whether *all* quantum channels exhibit additive Holevo information. In 2003, Shor published his now celebrated breakthrough paper [Sho04], in which he proved that the additivity conjecture for Holevo information is equivalent to three other additivity conjectures, including the additivity of the so-called *minimum output entropy* $H_{\min}$, defined for a quantum channel $\mathcal{N}: \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H}')$ by

$$H_{\min}(\mathcal{N}) := \min_{\rho\in\mathcal{S}(\mathcal{H})} H(\mathcal{N}(\rho)). \tag{3.16}$$

The additivity conjecture for the minimal output entropy can then be stated as

$$\text{For all quantum channels } \mathcal{N}_1, \mathcal{N}_2 \ : \ H_{\min}(\mathcal{N}_1\otimes\mathcal{N}_2) = H_{\min}(\mathcal{N}_1) + H_{\min}(\mathcal{N}_2). \tag{3.17}$$

Note that we trivially have $H_{\min}(\mathcal{N}_1 \otimes \mathcal{N}_2) \leqslant H_{\min}(\mathcal{N}_1) + H_{\min}(\mathcal{N}_2)$ as $H((\mathcal{N}_1 \otimes \mathcal{N}_2)(\rho \otimes \sigma)) = H(\mathcal{N}_1(\rho)) + H(\mathcal{N}_2(\sigma))$, but a strict inequality *may* occur when we also consider states that are not product states of the form $\rho \otimes \sigma$. The additivity conjecture for the minimum output entropy was initiated (at least in print) by King and Ruskai in [KR01], and was definitively *disproved* by a counterexample by Hastings in 2009 (after Hayden and Winter found counterexamples to the maximal p-norm multiplicativity conjecture for $p > 1$ which would lead to a counterexample for the additivity of minimum output entropy for $p = 1$) in [Has09]. Note that Hastings' counterexample relies on a quantum channel $\mathcal{E}$ described by $\mathcal{E}(\rho) = \sum_{i=1}^{D} p_i U_i^* \rho U_i$ for a sufficiently large dimension N, sufficiently many randomly chosen unitaries $U_i \in M_N(\mathbb{C})$ indexed by $1 \leqslant i \leqslant D$, and a randomly chosen discrete probability distribution $\{p_i\}_{i=1}^{D}$, such that the quantum channel $\mathcal{E}$ has a non-zero *probability* of breaking the additivity conjecture. So far, we have never found a deterministic construction of any pair of quantum channels that break the additivity conjecture for the minimal output entropy [CY19], although Hastings' counterexample has generated quite some research [ASW10, BH10, CFN12, FKM10, BCN16]. We can, however, identify certain classes of quantum channels that must always satisfy the additivity conjecture (and thus cannot yield a deterministic counterexample), which we will do in the following section.

3.5. Special classes of quantum channels

Certain quantum channels have useful properties that allow us to determine whether they have additivity Holevo information, or additive minimum output entropy. In this section, with Φ we denote a quantum channel $\Phi : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$ where $\mathcal{H}_A, \mathcal{H}_B$ are finite-dimensional Hilbert spaces. Recall from the previous section (see Definition 3.1.1) that we also have a *complementary quantum channel* $\Phi^c : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_E)$ given by $\Phi^c(\rho) = \mathrm{Tr}_B(V\rho V^*)$, if the Stinespring representation of Φ is $\Phi(\rho) = \mathrm{Tr}_E(V\rho V^*)$ with a Stinespring isometry $V : \mathcal{H}_A \to \mathcal{H}_B \otimes \mathcal{H}_E$. We then have the following two properties:

Definition 3.5.1: Positive Partial Transpose (PPT)

We say that Φ is *positive partial transpose* (PPT) if $\Theta_B \circ \Phi$ is still a quantum channel, where $\Theta_B : \mathcal{B}(\mathcal{H}_B) \to \mathcal{B}(\mathcal{H}_B)$ is the matrix transpose map if we identify $\mathcal{H}_B \simeq \mathbb{C}^{\dim \mathcal{H}_B}$.

This is equivalent to demanding that $(\Theta_B \otimes \iota_{\mathcal{H}_A})J(\Phi)$ is a positive matrix, where $J(\Phi)$ is the Choi matrix $J(\Phi) = \sum_{ij} \Phi(e_{ij}) \otimes e_{ij}$.

Definition 3.5.2: Entanglement breaking (EBT)

We say that Φ is *entanglement-breaking* ("EBT", the "T" appears because these channels are historically also called *entanglement-breaking trace-preserving*) if any of the following equivalent conditions is met:

1. The Choi matrix $J(\Phi)$ is *separable*, which means that we can find a discrete probability distribution $\{p_i\}_{i=1}^{d}$ for some $d \in \mathbb{N}$ and product states $\rho_i^B \otimes \sigma_i^A$ (where $\rho_i^B \in \mathcal{S}(\mathcal{H}_B)$ and $\sigma_i^A \in \mathcal{S}(\mathcal{H}_A)$ are density matrices) such that

$$J(\Phi) = \sum_{i=1}^{d} p_i \rho_i^B \otimes \sigma_i^A \tag{3.18}$$

2. The matrix $(\iota_C \otimes \Phi)(\rho)$ is separable for any density matrix $\rho \in \mathcal{S}(\mathcal{H}_C \otimes \mathcal{H}_A)$ with $\mathcal{H}_C$ any finite-dimensional Hilbert space, where *separable* again means that we can find a discrete probability distribution $\{p_i\}_{i=1}^{d}$ for some $d \in \mathbb{N}$ and product states $\sigma_i^C \otimes \rho_i^B$ where $\rho_i^B \in \mathcal{S}(\mathcal{H}_B)$ and $\sigma_i^C \in \mathcal{S}(\mathcal{H}_C)$ such that

$$(\iota_C \otimes \Phi)(\rho) = \sum_i p_i \sigma_i^C \otimes \rho_i^B. \tag{3.19}$$

3. The quantum channel Φ can be written as $\Phi(\rho) = \sum_{k=1}^{d} R_k \mathrm{Tr}(F_k \rho)$, where $d \in \mathbb{N}$, R_k are density matrices for $1 \leqslant k \leqslant d$, and $\{F_k\}_{k=1}^{d}$ form a POVM. This is called the "Holevo form" in literature [HSR03].
4. The quantum channel Φ can be written as $\Phi(\rho) = \sum_{k=1}^{d} |\psi_k\rangle\langle\psi_k| \langle\phi_k|\rho|\phi_k\rangle$ where $d \in \mathbb{N}$, $|\psi_k\rangle \in \mathcal{H}_\mathrm{B}$ for $1 \leqslant k \leqslant d$ and $|\psi_k\rangle \in \mathcal{H}_\mathrm{A}$ for $1 \leqslant k \leqslant d$. [HSR03]
5. The quantum channel Φ has a Kraus operator representation using only Kraus operators of rank 1 [HSR03].

Lemma 3.5.1: EBT implies PPT

Given a quantum channel Φ. If Φ is entanglement-breaking, then it is also positive partial transpose.

Proof. Write $J(\Phi) = \sum_{k=1}^{d} p_k \rho_k^\mathrm{B} \otimes \sigma_k^\mathrm{A}$, then we see that $(\Theta_B \otimes \iota_{\mathcal{H}_\mathrm{A}}) J(\Phi) = \sum_{k=1}^{d} p_k \left(\rho_k^\mathrm{B}\right)^t \otimes \sigma_k^\mathrm{A}$, but this is a convex combination of positive operators as the transpose of a matrix does not affect its eigenvalues, thus the convex combination itself is positive, thus Φ is PPT. □

The previous definitions of the PPT and EBT property, and the implication EBT $\implies$ PPT, are important to our consideration of quantum channels in this thesis because EBT channels are *always* MOE-additive:

Theorem 3.5.1: EBT channels are strongly MOE-additive and χ-additive

Given any entanglement-breaking channel Φ. Then, for *any arbitrary* other channel Ψ (note that this channel does *not* have to be entanglement breaking!), we have both

$$H_{\min}(\Phi \otimes \Psi) = H_{\min}(\Phi) + H_{\min}(\Psi), \tag{3.20}$$

and

$$\chi(\Phi \otimes \Psi) = \chi(\Phi) + \chi(\Psi). \tag{3.21}$$

Proof. See Theorem 1 and 2 in [Sho02]. □

As a matter of fact, it is not only important to check whether a quantum channel is entanglement-breaking (as it has to be MOE-additive in that case by the previous theorem), it is also imperative to check the EBT property for the complementary channel:

Theorem 3.5.2: MOE-additivity for complementary channels

For any quantum channels Φ and Ψ, the MOE-additivity holds if and only if the MOE-additivity holds for their complementary channels, i.e.

$$H_{\min}(\Phi \otimes \Psi) = H_{\min}(\Phi) + H_{\min}(\Psi) \iff H_{\min}(\Phi^c \otimes \Psi^c) = H_{\min}(\Phi^c) + H_{\min}(\Psi^c) \tag{3.22}$$

Proof. Theorem 5 in [KMNR05]. □

MOE-additivity results pertaining to unital qubit channels, depolarizing channels and Hadamard channels are also known, and can be found in Appendix B.

4

Compact Quantum Groups

The theory of compact quantum groups dates back to the last decades of the 20$^{\text{th}}$ century, with seminal work by Woronowicz in [Wor80], where he developed the notion of what he called "pseudogroups". The attempt to define group structures on these pseudospaces was already started by Kac in the '60s [Kac63, Kac65], and continued in the language of von Neumann algebras by Takesaki [Tak72] and Enock and Schwartz [ES75], and later on cast into C^*-algebra language by for example Jimbo and Drinfel'd [Dri88]. In [Wor87], and later on in [Wor91], Woronowicz developed a standard theory of what he then called *compact matrix pseudogroups*, which coincide with the theory of *compact matrix quantum groups* which we will develop in this chapter.

The starting point of his exploration was the Gelfand-Naimark isomorphism theorem, which showed that any commutative unital C^*-algebra A is isomorphic as C^*-algebra to the continuous functions $C(X)$ on a certain compact space X. In the non-unital but still commutative case, A is isomorphic as C^*-algebra to $C_0(X)$ for some *locally* compact space X, where $C_0(X)$ are the continuous functions that vanish at infinity.

In the framework of category theory, we consider the category **Cpt** of compact topological spaces (with continuous maps as morphisms). Furthermore, we consider the category $\mathbf{C^*alg}$ of C^*-algebras (with $*$-homomorphisms as morphisms), and its subcategory $\mathbf{C^*com}_1$ of commutative unital C^*-algebras (with unital $*$-homomorphisms as morphisms). Then, we have:

- The functor $C : \mathbf{Cpt}^{\text{op}} \to \mathbf{C^*com}_1$ which maps a compact space X to $C(X)$, and maps a morphism $f^{\text{op}} \in \text{Hom}_{\mathbf{Cpt}^{\text{op}}}(X, Y)$ to $C(f^{\text{op}}) \in \text{Hom}_{\mathbf{C^*com}_1}(C(X), C(Y))$ which takes a function $h \in C(X)$ to $h \circ f \in C(Y)$;
- The functor $\text{sp} : \mathbf{C^*com}_1^{\text{op}} \to \mathbf{Cpt}$ which maps A to the set $\text{sp}(A)$ of all non-trivial $*$-homomorphisms $f : A \to \mathbb{C}$ (called the *characters* on A), which becomes a compact space when endowed with the spectral topology. It maps unital $*$-homomorphisms $\varphi^{\text{op}} \in \text{Hom}_{\mathbf{C^*com}_1}(A, B)$ to $\text{sp}(\varphi^{\text{op}}) \in \text{Hom}_{\mathbf{Cpt}}(\text{sp}(A), \text{sp}(B))$, which takes a non-trivial $*$-homomorphism $f \in \text{sp}(A)$ to $f \circ \varphi \in \text{sp}(B)$.

In this framework, we can state: [Neg71, Del09]

Theorem 4.0.1

The functor $C : \mathbf{Cpt}^{\text{op}} \to \mathbf{C^*com}_1$ and the functor $\text{sp} : \mathbf{C^*com}_1^{\text{op}} \to \mathbf{Cpt}$ are *contravariant equivalences of categories* (also called a *duality*), and they are quasi-inverses.

Although this shows the duality of *commutative* C^*-algebras with compact topological spaces [1], Woronowicz noted that no such theory exists for non-commutative C^*-algebras, and therefore invented what he called the "category of pseudospaces" , which should be

[1] Note that such a duality also exists for the non-unital case, but we will not go into detail here.

constructed such that it is categorically dual to the category of all C^*-algebras $\mathbf{C^*alg}$. Although, as Woronowicz noted himself, this approach is still fundamentally a theory about C^*-algebras, it proved fruitful to reason from this framework, quite like a change of reference frame in physics can yield insights, without actually changing any of the physical laws or objects involved.

An important subclass of these so-called pseudospaces is the class of *pseudo-semigroups*, where one endows a pseudospace $\mathbb{G}$ with a comultiplication $\varphi : \mathbb{G} \to \mathbb{G}\times\mathbb{G}$ such that it satisfies the associativity $(\mathrm{id} \times \varphi) \circ \varphi = (\varphi \times \mathrm{id}) \circ \varphi$, where $\mathrm{id} : \mathbb{G} \to \mathbb{G}$ is the identity morphism. In this context, one can wonder whether generalizations of the Pontrjagin duality exist in the non-abelian case [Wor80]. In the classical case, a locally compact *abelian* group G has a dual group $\hat{G}$: as a set, it consists of all *characters* on G, i.e. all continuous group homomorphisms $G \to \mathbb{T}$, endowed with pointwise multiplication, and the topology of uniform convergence on compact subsets. The Pontrjagin duality theorem then tells us that the map $G \to \hat{\hat{G}}$ given by $x \mapsto \mathrm{ev}_x$ is an isomorphism of topological groups, where $\mathrm{ev}_x(\varphi) = \varphi(x)$ for any $\varphi \in \hat{G}$. The Pontrjagin duality is instrumental in understanding and generalizing the Fourier transform, and the Fourier inversion theorem.

In '88, Woronowicz famously proved that a generalization of the Tannaka-Krein theorem exists for compact quantum groups [Wor88], which, roughly speaking, tells us that compact quantum groups are uniquely determined by the category of their finite-dimensional unitary corepresentations. We will come back to the Tannaka-Krein theorem later in this chapter. According to [MD98], Woronowicz also distributed the preprint *Compact Quantum Groups* [Wor92], which contains the definition for a compact quantum group that is still widely used in the field. Let us commence from this point.

4.1. Definition of a compact quantum group

In this thesis, we will consider particular type of compact quantum groups (which we will often abbreviate with CQG), namely the so-called *partition* (or "easy") quantum groups, which are a special subset of the so-called *compact matrix quantum groups* (CMQG). Although this identification would allow us to focus only on the definition of CMQG's, it is beneficial to generalize the setting slightly to include all compact quantum groups:

Definition 4.1.1

A *compact quantum group* (CQG) $\mathbb{G}$ is a pair $\mathbb{G} := (A, \Delta)$, with A a unital C^*-algebra, and $\Delta : A \to A \otimes A$ a *comultiplication* or *coproduct*, i.e. a unital $*$-homomorphisms with the *coassociativity property* that $(\Delta \otimes \iota) \circ \Delta = (\iota \otimes \Delta) \circ \Delta$, where ι is the identity map on A. Furthermore, we demand that the *cancellation property* holds: $\Delta(A)(1_A \otimes A)$ and $\Delta(A)(A \otimes 1_A)$ lie linearly dense in $A \otimes A$.

Given two compact quantum groups (A, Δ_A) and (B, Δ_B), a map between them is called a *morphism of quantum groups* if it is a non-degenerate unital $*$-homomorphism $f : A \to B$ that satisfies $\Delta_B \circ f = (f \otimes f) \circ \Delta_A$. Here, *non-degenerate* refers to the property that $f(A)B$ lies linearly dense in B.

Example 4.1.1 (Compact groups are compact quantum groups). Given a compact group G, consider the C^*-algebra $A := C(G)$. The multiplication $\cdot : G \times G \to G$ by $(a, b) \mapsto ab$ induces a comultiplication $\Delta_G : C(G) \to C(G) \otimes C(G)$, if we identify $C(G) \otimes C(G) \simeq C(G \times G)$, by the rule $\Delta_G(f)(x, y) = f(xy)$ for all $x, y \in G$. One can check that this turns $(C(G), \Delta_G)$ into a compact quantum group, where the cancellation property follows from the fact that $\Delta(A)(A \otimes 1_A)$ is spanned by functions of the form $(x, y) \mapsto f(xy)g(x)$ (where $f, g \in C(G)$). As a unital $*$-subalgebra of $C(G \times G)$, it clearly separates the points of $G \times G$ [2]. Hence, by Stone-Weierstrass, this must be a dense subspace – the density of $\Delta(A)(1_A \otimes A)$ is similar.

Remark 4.1.1. In literature, one often finds the notation $C(\mathbb{G})$ for A in $\mathbb{G} = (A, \Delta)$. This precisely ties in to the non-commutative generalization of the duality described in Theo-

[2] Let $(x, y) \neq (x', y')$, if $x \neq x'$ put $f = 1_{C(G)}$ and pick g such that $g(x) \neq g(x')$, otherwise $x = x'$ so put $g = 1_{C(G)}$ and pick f such that $f(xy) \neq f(xy')$.

rem 4.0.1: we should read this as the "non-commutative" space of continuous functions on a virtual compact group $\mathbb{G}$ (which Woronowicz called a "pseudogroup"), i.e. the continuous functions in $C(\mathbb{G})$ in general do not commute. This highlights the generalization of compact quantum groups from compact groups, in which case the C^*-algebra $C(G)$ with G a compact group is, in fact, commutative. This also inspires the following:

Proposition 4.1.1: Commutative compact quantum groups

[Tim08, Prop. 5.1.3] Let $\mathbb{G} := (A, \Delta)$ be a compact quantum group, and let A be abelian. Then one can find a compact group G and an isomorphism of compact quantum groups between (A, Δ) and $C(G)$, where the comultiplication on $C(G)$ is constructed as follows: $(\Delta_G(f))(x, y) = f(xy)$ for all $x, y \in G$.

Sketch of proof. As A is a unital commutative C^*-algebra, by the Gelfand-Naimark isomorphism $A \simeq C(G)$ as C^*-algebra, where G is the compact space $\mathrm{sp}(A)$ of all non-trivial characters on A with the spectral topology. The comultiplication $\Delta : C(G) \to C(G) \otimes C(G)$ can be seen as a map $C(G) \to C(G \times G)$ under the identification $C(G) \otimes C(G) \simeq C(G \times G)$. Thus, we find a multiplication $m : G \times G \to G$ by identifying $G \times G$ with $\mathrm{sp}(A \otimes A)$, and then m maps $\varphi \in \mathrm{sp}(A \otimes A)$ to $\varphi \circ \Delta$. It can be shown that m is associative by the coassociativity of Δ, turning G into a compact semigroup. The requirement that $\Delta(A)(1_A \otimes A)$ and $\Delta(A)(A \otimes 1_A)$ lie linearly dense in $A \otimes A$ translate into the *left and right cancellation properties* for G, respectively. Here, left cancellation means that if $f, g \in G$, if $hf = hg$ for all $h \in G$, then $f = g$. Right cancellation is similar. A compact semigroup with the cancellation property is, in fact, a compact group [MD98, Prop 3.2]. See also [Web17a, Ex. 2.5]. □

4.2. States on compact quantum groups

Definition 4.2.1

Given a compact quantum group $\mathbb{G} := (A, \Delta)$ and a state φ on A. We say that the state φ is *left-invariant* if $(\mathrm{id}_A \otimes \varphi)(\Delta(a)) = \varphi(a)1_A$ for all $a \in A$. Similarly, φ is called *right-invariant* if $(\varphi \otimes \mathrm{id}_A)(\Delta(a)) = \varphi(a)1_A$ for all $a \in A$.

Theorem 4.2.1

[Tim08, Thm. 5.1.6] Given a compact quantum group $\mathbb{G} := (A, \Delta)$. This quantum group has a unique state h that is both left-invariant and right-invariant, i.e. any other state that is both left- and right-invariant coincides with h. This unique state h is called the *Haar state*.

4.3. Corepresentations of compact quantum groups

The quantum channels that are analyzed in this thesis arise from looking at intertwiners between (tensor products of) corepresentations of compact quantum groups, which necessitates an introduction to the representation theory of compact quantum groups.

We begin with the definition of a corepresentation:

Definition 4.3.1

Given a CQG $\mathbb{G} = (A, \Delta)$. A *finite-dimensional (unitary) corepresentation operator* of $\mathbb{G}$ on a finite-dimensional Hilbert space $\mathcal{H}$ is an invertible (unitary) operator $X \in \mathcal{B}(\mathcal{H}) \otimes A$, that satisfies the equation

$$X_{[12]}X_{[13]} = (\iota_{\mathcal{H}} \otimes \Delta)(X) \in \mathcal{B}(\mathcal{H}) \otimes A \otimes A. \tag{4.1}$$

Here, $X_{[12]} := X \otimes 1_A \in \mathcal{B}(\mathcal{H}) \otimes A \otimes A$ and $X_{[23]} := (\iota_{\mathcal{H}} \otimes \Sigma)(X_{[12]}) \in \mathcal{B}(\mathcal{H}) \otimes A \otimes A$, where $\Sigma : A \otimes A \to A \otimes A$ is the flip operator $a \otimes b \mapsto b \otimes a$. We call $\dim \mathcal{H}$ the dimension of X.

As finite-dimensional Hilbert spaces can be identified with $\mathbb{C}^n$ with $n = \dim \mathcal{H}$, we can identify $\mathcal{B}(\mathcal{H}) \otimes A$ with $M_n(\mathbb{C}) \otimes A$, which can be identified with $M_n(A)$ (matrices with elements in A as entries), and we find an equivalent view of finite-dimensional corepresentations in terms of matrices:

Definition 4.3.2

Given a CQG $\mathbb{G} = (A, \Delta)$. A *corepresentation matrix* of $\mathbb{G}$ is an invertible matrix $a \in M_n(A)$ that satisfies $\Delta(a_{ij}) = \sum_k a_{ik} \otimes a_{kj}$. We call a *unitary* if for all $1 \leqslant i, j \leqslant n$,

$$\sum_k a_{ki}^* a_{kj} = \delta_{ij} 1_A = \sum_k a_{ik} a_{jk}^* \tag{4.2}$$

Theorem 4.3.1

Given a CQG $\mathbb{G} = (A, \Delta)$ and a finite-dimensional Hilbert space $\mathcal{H}$. The (unitary) corepresentation operators of $\mathbb{G}$ on $\mathcal{H}$ are in bijective correspondence with the (unitary) corepresentation matrices of $\mathbb{G}$ with dimension $n = \dim \mathcal{H}$.

Sketch of proof. Given a corepresentation matrix $a = (a_{ij}) \in M_n(A)$, one can construct the corepresentation operator $X = \sum_{i,j=1}^n e_{ij} \otimes a_{ij}$, where $e_{ij} \in M_n(\mathbb{C})$ is the matrix with all entries zero except the (i, j)-th entry, which is 1. Conversely, given a corepresentation operator X, one can construct a corepresentation matrix $a \in M_n(A)$ by the rule $a_{ij} = (\omega_{e_i, e_j} \otimes \iota_A)(X)$ where $\omega_{\eta,\xi} : \mathcal{B}(\mathcal{H}) \to \mathbb{C}$ is given by $T \mapsto \langle \eta | T\xi \rangle$.

For details, please see [Tim08, Section 5.2.3]. □

Remark 4.3.1. The definition of corepresentations of CQG's in literature, especially when we are concerned with compact *matrix* quantum groups (also see the next section), such as in [FW16, Def. 3.2], corresponds with the corepresentation matrices defined in Definition 4.3.2.

Now that we have defined corepresentations (and the different ways of looking at them), we can consider the following:

Definition 4.3.3

Given a CQG $\mathbb{G} = (A, \Delta)$ and two corepresentation operators X and Y on Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$, respectively. An *intertwiner* between X and Y is an operator $T \in \mathcal{B}(\mathcal{H}, \mathcal{K})$ (i.e. a bounded linear operator $T : \mathcal{H} \to \mathcal{K}$), such that

$$Y(T \otimes \iota_A) = (T \otimes \iota_A) X. \tag{4.3}$$

The collection of all intertwiners between X and Y shall be denote $\mathrm{Hom}_{\mathbb{G}}(X, Y)$, and if the CQG $\mathbb{G}$ is clear from context, sometimes $\mathrm{Hom}(X, Y)$. If there exists an *invertible intertwiner* in $\mathrm{Hom}_{\mathbb{G}}(X, Y)$, we call X and Y *equivalent* and write $X \simeq Y$.

Remark 4.3.2. If we write $X = \sum_{ij} e_{ij} \otimes a_{ij}$ with corepresentation matrix $a = (a_{ij})_{ij} \in M_n(A)$ and $Y = \sum_{ij} f_{ij} \otimes b_{ij}$ with corepresentation matrix $b = (b_{ij})_{ij} \in M_m(A)$, then writing out the

equation $Y(T \otimes \iota_A) = (T \otimes \iota_A)X$ with $T = \sum_{ij} t_{ij} |i\rangle\langle j|$, one finds that T intertwines X and Y iff $\sum_j t_{ij} a_{jl} = \sum_j t_{jl} b_{ij}$ for all $1 \leqslant i \leqslant m$ and $1 \leqslant j \leqslant n$, which is precisely the matrix equation $Ta = bT$. Therefore, we say that T *intertwines corepresentation matrices* a *and* b if $Ta = bT$, write $\mathrm{Hom}_{\mathbb{G}}(a, b)$ for the intertwiner space between a and b, and thus note that $\mathrm{Hom}(a, b) = \mathrm{Hom}(X, Y)$.

Definition 4.3.4

Given a CQG $\mathbb{G} = (A, \Delta)$ and a corepresentation operator X on a Hilbert space $\mathcal{H}$, a subspace $L \subseteq \mathcal{H}$ is called *invariant for* X if the canonical projection $p : \mathcal{H} \twoheadrightarrow L$ satisfies:

$$X\,(p \otimes \iota_A) = (p \otimes \iota_A)\,X\,(p \otimes \iota_A)\,. \tag{4.4}$$

Furthermore, if the only invariant subspaces for X are 0 and $\mathcal{H}$ itself, we call X an *irreducible corepresentation operator*.

Remark 4.3.3. Following the spirit of Remark 4.3.2, if we say $X = \sum_{ij} e_{ij} \otimes a_{ij}$ with corepresentation matrix $a = (a_{ij})_{ij} \in M_n(A)$, and $p = \sum_{ij} p_{ij} |i\rangle\langle j|$, we see that $X(p \otimes \iota_A) = (p \otimes \iota_A)X(p \otimes \iota_A)$ iff $ap = pap$.

Remark 4.3.4. Note that if we have a projection $p : \mathcal{H} \twoheadrightarrow p\mathcal{H}$ such that $X(p \otimes \iota_A) = (p \otimes \iota_A)X(p \otimes \iota_A)$, or equivalently $ap = pap$ for some corepresentation operator $X = \sum_{ij} e_{ij} \otimes a_{ij}$ and corepresentation matrix $a = (a_{ij})_{ij} \in M_n(A)$, then we also have

$$(p \otimes \iota_A)X = X(p \otimes \iota_A) \quad \text{and} \quad pa = ap. \tag{4.5}$$

See [Tim08, Prop. 5.2.8] for a proof.

We can construct new corepresentations on a CQG $\mathbb{G}$ from other corepresentations by considering their direct sum, tensor product, or by conjugating a corepresentation.

Definition 4.3.5: Direct sum

Given a finite family of corepresentation operators $(X_i)_{i \in I}$ on finite-dimensional Hilbert spaces $(\mathcal{H}_i)_{i \in I}$ of a CQG $\mathbb{G} = (A, \Delta)$, we define the *direct sum* $\bigoplus_i X_i \in \mathcal{B}(\oplus_i \mathcal{H}_i) \otimes A$ by demanding that for each $i \in I$, the natural inclusion $\iota_i : \mathcal{H}_i \hookrightarrow \oplus_j \mathcal{H}_j$ is an intertwiner from X_i to $\oplus_i X_i$. This makes the operator unique.

Remark 4.3.5. If we associate the corepresentation matrix $a = (a_{ij})_{ij} \in M_n(A)$ to the corepresentation operator X, and similarly $b = (b_{ij})_{ij} \in M_m(A)$ to corepresentation matrix Y, then the corepresentation matrix associated to $X \oplus Y$, denoted by $a \oplus b$, is the matrix $\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \in M_{m+n}(A)$.

Definition 4.3.6: Tensor product

Given two corepresentation operators X and Y on finite-dimensional Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$, respectively, of a CQG $\mathbb{G} = (A, \Delta)$. We define the *tensor product* $X \otimes Y$ as the operator

$$X \otimes Y := X_{[13]} Y_{[23]} \in \mathcal{B}\,(\mathcal{H} \otimes \mathcal{K}) \otimes A, \tag{4.6}$$

where we have defined $Y_{[23]} = \iota_{\mathcal{H}} \otimes Y$ and $X_{[13]} = (\Sigma \otimes \iota_A)\,(\iota_{\mathcal{K}} \otimes X)$ with the flip operation $\Sigma : \mathcal{B}(\mathcal{K}) \otimes \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H}) \otimes \mathcal{B}(\mathcal{K})$ by $T \otimes R \mapsto R \otimes T$.

Remark 4.3.6. If we express the corepresentation operator X as $X = \sum_{ij}^{n} e_{ij} \otimes a_{ij}$ where $a := (a_{ij})_{ij}$ is its associated corepresentation matrix, and similarly $Y = \sum_{ij}^{m} e'_{ij} \otimes b_{ij}$, where $n = \dim \mathcal{H}$ and $m = \dim \mathcal{K}$ and we identify $\mathcal{H} \simeq \mathbb{C}^n$ and $\mathcal{K} \simeq \mathbb{C}^m$, and $\{e_{ij}\}$ and $\{e'_{ij}\}$ are the

canonical orthonormal bases of $M_n(\mathbb{C})$ and $M_m(\mathbb{C})$, writing out the definition of the tensor product yields:

$$X \otimes Y = \sum_{i,j=1}^{n} \sum_{k,l=1}^{m} e_{ij} \otimes e'_{kl} \otimes a_{ij} b_{kl}, \tag{4.7}$$

hence we can associate with $X \otimes Y$ the corepresentation matrix $a \otimes b := (a_{ij} b_{kl})_{ijkl}$ if we identify $\mathbb{C}^n \otimes \mathbb{C}^m \simeq \mathbb{C}^{nm}$ with canonical basis $\{e_{ij} \otimes e'_{kl} \simeq e_{ijkl}\}_{ijkl}$.

Definition 4.3.7: Conjugation

Given a corepresentation operator X on a finite-dimensional Hilbert space $\mathcal{H}$ of a CQG $\mathbb{G} = (A, \Delta)$. Let $\overline{\mathcal{H}}$ denote the conjugate Hilbert space, which contains vectors $\overline{\xi}$ for $\xi \in \mathcal{H}$, with the scalar multiplication $\lambda \overline{\xi} = \overline{\overline{\lambda} \xi}$ for $\lambda \in \mathbb{C}$ and the inner product $\langle \overline{\eta} | \overline{\xi} \rangle := \langle \xi | \eta \rangle_{\mathcal{H}}$ for $\xi, \eta \in \mathcal{H}$. A bounded operator $\overline{T}$ on this space, i.e. $\overline{T} \in \mathcal{B}(\overline{\mathcal{H}})$, can be described by $\overline{\xi} \mapsto \overline{T\xi}$. This gives a conjugate-linear isomorphism between $\mathcal{H}$ and $\overline{\mathcal{H}}$ by $\xi \mapsto \overline{\xi}$ and also a conjugate-linear isomorphism between $\mathcal{B}(\mathcal{H})$ and $\mathcal{B}(\overline{\mathcal{H}})$ by $T \mapsto \overline{T}$.

Express the corepresentation operator X as $X = \sum_{ij} e_{ij} \otimes a_{ij}$ where $\{e_{ij}\}$ is a basis for $\mathcal{H}$ and $a = (a_{ij}) \in M_n(A)$ is the corresponding corepresentation matrix. Then, we define the *conjugate corepresentation operator* $\overline{X}$ as

$$\overline{X} := \sum_{ij} \overline{e_{ij}} \otimes a_{ij}^*. \tag{4.8}$$

One can prove (see [Tim08, Cor. 5.3.10]) that $\overline{X}$ is again a corepresentation operator. Note that if X is a unitary corepresentation operator, $\overline{X}$ is not necessarily so.

The additional structure of tensor products and conjugation of corepresentations also leads to an important consequence known as Frobenius reciprocity:

Proposition 4.3.1

[Tim08] Let X, Y and Z be corepresentation operators of a compact quantum group $\mathbb{G}$ on Hilbert spaces $\mathcal{H}, \mathcal{K}, \mathcal{L}$, respectively. Then, there exist isomorphisms

$$\mathrm{Hom}_{\mathbb{G}}\left(\overline{X} \otimes Y, Z\right) \simeq \mathrm{Hom}_{\mathbb{G}}\left(Y, X \otimes Z\right) \tag{4.9}$$

$$\mathrm{Hom}_{\mathbb{G}}\left(Y \otimes X, Z\right) \simeq \mathrm{Hom}_{\mathbb{G}}\left(Y, Z \otimes \overline{X}\right). \tag{4.10}$$

We will also need the following identification:

Lemma 4.3.1

Given a CQG $\mathbb{G} = (A, \Delta)$, a corepresentation matrix u of $\mathbb{G}$ on the finite-dimensional Hilbert space $\mathcal{H}$, integers $n, m \in \mathbb{N}_0$, and *subrepresentations* $u_p := pu^{\otimes n}$ and $u_q := qu^{\otimes m}$ for two projections $p : \mathcal{H}^{\otimes n} \twoheadrightarrow \mathcal{H}_p$ and $q : \mathcal{H}^{\otimes m} \twoheadrightarrow \mathcal{H}_q$ that satisfy $u^{\otimes n} p = p u^{\otimes n} p$ and $u^{\otimes m} q = q u^{\otimes m} q$. Then,

$$\mathrm{Hom}(u_p, u_q) = q\mathrm{Hom}(u^{\otimes n}, u^{\otimes m})p, \tag{4.11}$$

if we see both as subspaces of $\mathcal{B}(\mathcal{H}^{\otimes n}, \mathcal{H}^{\otimes m})$.

Proof. We identify $\mathcal{H}_p = p\mathcal{H}^{\otimes n} \subset \mathcal{H}^{\otimes n}$ and $\mathcal{H}_q = q\mathcal{H}^{\otimes m} \subset \mathcal{H}^{\otimes m}$. Following Remark 4.3.4, we know that $pu^{\otimes n} = u^{\otimes n} p$ and $qu^{\otimes m} = u^{\otimes m} q$.

Assume we have an intertwiner $T \in \mathrm{Hom}(u^{\otimes n}, u^{\otimes m})$, then we claim that $qTp \in \mathrm{Hom}(u_p, u_q)$. This follows directly by definition:

$$qTpu_p = qTp(pu^{\otimes n}) = qTpu^{\otimes n} = qTu^{\otimes n} p = qu^{\otimes m} Tp = q^2 u^{\otimes m} Tp = qu^{\otimes m} qTp = u_q qTp. \tag{4.12}$$

Conversely, given a $T \in \mathrm{Hom}(u_p, u_q)$, under the embedding of $\mathrm{Hom}(u_p, u_q)$ into $\mathcal{B}(\mathcal{H}^{\otimes n}, \mathcal{H}^{\otimes m})$, we see

$$qTpu^{\otimes n} = qTp^2u^{\otimes n} = qTpu_p = qTu_pp = qu_qTp = q^2u^{\otimes m}Tp = qu^{\otimes m}qTp = u_qqTp, \tag{4.13}$$

so $qTp \in \mathrm{Hom}(u^{\otimes n}, u^{\otimes m})$ and $q(qTp)p = qTp$ so $qTp \in q\mathrm{Hom}(u^{\otimes n}, u^{\otimes m})p$. □

4.4. THE STRUCTURE OF ALL COREPRESENTATIONS

Using the techniques from the previous subsection, which allow us to make new corepresentations from others through direct sums, tensor products and conjugation, we can prove that this is sufficient to describe "all" the corepresentations:

Theorem 4.4.1

[Tim08, Sect. 5.3] Give a CQG $\mathbb{G}$, the following holds:

1. **Peter-Weyl**: Every corepresentation operator is equivalent to the direct sum of irreducible finite-dimensional unitary corepresentation operators. Furthermore, any irreducible corepresentation operator is finite-dimensional. Moreover, the subspace spanned by the matrix elements of all irreducible corepresentations lies dense in $C(\mathbb{G})$, and they satisfy orthogonality relations with respect to the Haar state on $\mathbb{G}$.
2. **Schur's lemma** : given two irreducible corepresentations X and Y, we either have $\mathrm{Hom}(X, Y) = 0$, or we have $X \simeq Y$ and $\dim \mathrm{Hom}(X, Y) = 1$.
3. **Tannaka-Krein**: The CQG $\mathbb{G}$ can be uniquely recovered from the complete concrete monoidal W^*-category of its finite-dimensional unitary corepresentations.

Remark 4.4.1

Note that the first item of this theorem also concerns corepresentation operators on *infinite-dimensional* Hilbert spaces, and the direct sum may also be a direct sum over an *infinite* family of corepresentations – we have not defined this, as this brings along some technical details concerning compact operators and multiplier algebras, whilst we will not extensively need this generality. See [Tim08] for these generalizations.

4.5. TANNAKA-KREIN DUALITY

In his seminal work in 1988, Woronowicz proved a generalization of the Tannaka-Krein duality for compact quantum groups [Wor88]. The theorem is cast in the language of concrete W^*-categories, whose definition we will repeat here.

Definition 4.5.1: W^*-category

[Web17a, Def. 4.15] Consider a set of objects R, and a binary operation $\cdot : R \times R \to R$. Associate to each object $r \in R$ a finite-dimensional Hilbert space $\mathcal{H}_r$, and associate for any pair $r, s \in R$ a linear subspace of $\mathcal{B}(\mathcal{H}_r, \mathcal{H}_s)$ which we will denote by $\mathrm{Hom}(r, s)$. The tuple $\mathcal{R} := (R, \cdot, \{\mathcal{H}_r\}_{r\in R}, \{\mathrm{Hom}(r, s)\}_{r\in R})$ is called a *concrete monoidal W^*-category* if:

1. The identity operator $\mathrm{id} \in \mathcal{H}_r$ is included in $\mathrm{Hom}(r, r)$ for any $r \in R$
2. The category is closed under composition: if $a \in \mathrm{Hom}(r, s)$ and $b \in \mathrm{Hom}(s, t)$, then $b \cdot a \in \mathrm{Hom}(r, t)$ for any $r, s, t \in R$.
3. The category is closed under the involution: if $a \in \mathrm{Hom}(r, s)$ then $a^* \in \mathrm{Hom}(s, r)$ for any $s, r \in R$.
4. If $\mathcal{H}_r = \mathcal{H}_s$ and $\mathrm{id} \in \mathrm{Hom}(r, s)$, then $r = s$.
5. The category is closed under the tensor product: if $a \in \mathrm{Hom}(r, s)$, and $b \in \mathrm{Hom}(r', s')$, then $a \otimes b \in \mathrm{Hom}(r \cdot r', s \cdot s')$ for any $r, r', s, s' \in R$.

6. The binary operation is associative: $(r \cdot s) \cdot t = r \cdot (s \cdot t)$ for any $r, s, t \in R$.
7. R contains a unit element 1 such that $\mathcal{H}_1 = \mathbb{C}$, and $1 \cdot r = r \cdot 1 = r$ for any $r \in R$.

Furthermore, we call this concrete monoidal W^*-category *complete* if additionally:

8. For any $r \in R$ and any unitary $u : \mathcal{H}_r \to \mathcal{K}$ where $\mathcal{K}$ is a Hilbert space, there exists an $s \in R$ such that $\mathcal{H}_s = \mathcal{K}$ and $u \in \mathrm{Hom}(r, s)$.
9. For any $r \in R$ and any orthogonal projection $p \in \mathrm{Hom}(r, r)$, there exists an $s \in R$ with $\mathcal{H}_s = p\mathcal{H}_r$ and the embedding $\iota : \mathcal{H}_s \hookrightarrow \mathcal{H}_r$ is in $\mathrm{Hom}(s, r)$.
10. We can take direct sums: for any $r, r' \in R$, we can find an $s \in R$ such that $\mathcal{H}_s = \mathcal{H}_r \oplus \mathcal{H}_{r'}$, and the embeddings $\iota_r : \mathcal{H}_r \hookrightarrow \mathcal{H}_s$ and $\iota_{r'} : \mathcal{H}_{r'} \hookrightarrow \mathcal{H}_s$ are in $\mathrm{Hom}(r, s)$ and $\mathrm{Hom}(r', s)$, respectively.

We say that an element $\overline{r} \in R$ is the *complex conjugate* of $r \in R$ if we can find a map $j : \mathcal{H}_r \to \mathcal{H}_{\overline{r}}$ that is invertible and anti-linear, such that the map $\mathbb{C} \to \mathcal{H}_r \otimes \mathcal{H}_{\overline{r}}$ given by $1 \mapsto \sum_i e_i \otimes j(e_i)$ is in $\mathrm{Hom}(1, r \cdot \overline{r})$, and similarly the map $\mathcal{H}_{\overline{r}} \otimes \mathcal{H}_r \to \mathbb{C}$ given by $\overline{e_i} \otimes e_j \mapsto \langle j^{-1}(e_i), e_j \rangle$ is in $\mathrm{Hom}(\overline{r} \cdot r, 1)$.

Given a concrete monoidal W^*-category $\mathcal{R}$, we say that a finite subset $Q \subseteq R$ *generates* $\mathcal{R}$ if for any $s \in R$, we can find morphisms $b_k \in \mathrm{Hom}(q_k, s)$ for $1 \leqslant k \leqslant m$ for some $m \in \mathbb{N}$ such that $\sum_k b_k b_k^* = \mathrm{id} \in \mathrm{Hom}(s, s)$, where q_k are monomials in Q.

We have now established enough theory to understand the compact quantum group setting: the category of finite-dimensional unitary representations Rep $\mathbb{G}$ from a compact quantum group $\mathbb{G}$ form a complete concrete monoidal W^*-category:

Theorem 4.5.1

[Wor88] Let $\mathbb{G}$ be a CQG. The class of all finite-dimensional unitary representations Rep $\mathbb{G}$ is a complete concrete monoidal W^*-category, where the morphisms are intertwiners and the binary operation is the tensor product. For any representation $a \in$ Rep $\mathbb{G}$, there exists the complex conjugation $\overline{a} \in$ Rep $\mathbb{G}$. The fundamental representation u is the *distinguished object* of Rep $\mathbb{G}$: u and $\overline{u}$ together generate Rep $\mathbb{G}$.

One of the most celebrated results in the field of compact quantum groups is the realization that the dual to the previous theorem exists:

Theorem 4.5.2: Tannaka-Krein for CMQGs

[Wor88] Let $\mathcal{R} := (R, \cdot, \{\mathcal{H}_r\}_{r \in R}, \{\mathrm{Hom}(r, s)\}_{r,s \in R})$ be a concrete monoidal $W*$-category, such that $\{f, \overline{f}\}$ generates $\mathcal{R}$. Then, we can find a compact matrix quantum group $\mathbb{G} = (A, u)$ and a natural completion $\overline{\mathcal{R}}$ of $\mathcal{R}$ such that $\overline{\mathcal{R}} =$ Rep $\mathbb{G}$. Moreover, if another compact matrix quantum group $\mathbb{G}' = (B, v)$ exists with the property that $\mathcal{R} \subseteq$ Rep $\mathbb{G}'$, then there exists a morphism of compact quantum groups $A \to B$ such that $u \mapsto v$.

For a sketch of the proof, see for example [Web17a].

4.6. COMPACT MATRIX QUANTUM GROUPS

In this thesis, we will be concerned with compact quantum groups that are generated by a unitary matrix of formal elements, which are called compact *matrix* quantum groups:

Definition 4.6.1: Compact matrix quantum group

[Tim08, Def. 6.1.1] A compact matrix quantum group $\mathbb{G} = (A, \Delta, u)$ is a unital C^*-algebra A together with a coproduct Δ, i.e. a $*$-homomorphism $\Delta : A \to A \otimes A$, together with a unitary $u \in M_n(A)$ for some $n \in \mathbb{Z}_{>0}$ such that

1. $\Delta(u_{ij}) = \sum_{k=1}^{n} u_{ik} \otimes u_{kj}$ for all $1 \leqslant i, j \leqslant n$;
2. $\overline{u}$ is an invertible matrix;
3. the elements in the set $\{u_{ij} \mid 1 \leqslant i, j \leqslant n\}$ generate A as a C^*-algebra.

The matrix u is called the *fundamental corepresentation (matrix)*.

Note that the comultiplication Δ in Definition 4.6.1 is uniquely determined by the first and third condition, hence we may also refer to the CMQG $\mathbb{G}$ as $\mathbb{G} = (A, u)$. [Tim08, Remark 6.1.2]

Proposition 4.6.1: CMQG's are CQG's

[Tim08, Prop. 6.1.4] Given a compact matrix quantum group $\mathbb{G} := (A, \Delta, u)$, then (A, Δ) is a compact quantum group.

Definition 4.6.2

[Tim08, Def. 6.1.3] We say that two CMQG's $\mathbb{G} = (A, \Delta_A, u)$ and $\mathbb{H} = (B, \Delta_B, v)$ are *similar*, denoted as $\mathbb{G} \simeq \mathbb{H}$, if there exists an $n \in \mathbb{N}$ such that $u \in M_n(A)$ and $v \in M_n(B)$, and there exists a $*$-isomorphism $f : A \to B$, and there exists an invertible matrix $T \in \mathrm{GL}_n(\mathbb{C})$ such that $v = T(f(u_{ij})_{ij})T^{-1}$.

An important result for compact matrix quantum groups is the fact that their irreducible corepresentations can be described as follows:

Theorem 4.6.1

[Tim08, Lemma 6.1.5] Given a CMQG $\mathbb{G} = (A, \Delta, u)$, and given any irreducible corepresentation of (A, Δ). This irreducible corepresentation must be contained in an iterated tensor product of u, $\overline{u}$, and the trivial corepresentation.

In this thesis, we will be concerned with two compact matrix quantum groups: the *free orthogonal quantum group* O_N^+, and the *quantum permutation group* S_N^+. Let us define them.

4.6.1. The free orthogonal quantum group O_N^+

Definition 4.6.3

The **orthogonal quantum group**, denoted O_N^+ or $A_o(N)$, with dimension $N \in \mathbb{N}$, is the compact matrix quantum group $O_N^+ = (C(O_N^+), u)$, where $C(O_N^+)$ is the universal unital C^*-algebra generated by the elements $\{u_{ij}\}_{i,j=1}^{N}$ placed in the corepresentation matrix $u := (u_{ij})_{ij} \in M_N(C(O_N^+))$ with the relation that u is *orthogonal*, i.e. we have

$$C(O_N^+) := C^*\left(u_{ij} \ , \ 1 \leqslant i, j \leqslant N \mid u_{ij}^* = u_{ij} \ , \ \sum_k u_{ik}u_{jk} = \sum_k u_{ki}u_{kj} = \delta_{ij} \ \forall i, j \right). \tag{4.14}$$

We can generalize this definition to the free orthogonal quantum group with parameter matrix F, such that the previous definition coincides with this generalization under the choice $F = I_N \in \mathrm{GL}_N(\mathbb{C})$, where I_N is the identity matrix: [vDW96, Tim08]

Definition 4.6.4

For some $N \in \mathbb{N}$, assume we are given a matrix $F \in \mathrm{GL}_N(\mathbb{C})$ with the property $F\overline{F} \in \mathbb{R}I_N$. The *free orthogonal quantum group with parameter matrix* F, denoted by O_F^+ or $A_o(F)$, is the CMQG $(C(O_F^+), u)$, where $C(O_F^+)$ is the universal unital C^*-algebra generated by the elements $\{u_{ij}\}_{i,j=1}^N$ placed in a corepresentation matrix $u := (u_{ij})_{ij} \in M_N(C(O_F^+))$ with the relation that u is *unitary*, i.e. $uu^* = u^*u = 1$, and $u = F\overline{u}F^{-1}$.

Remark 4.6.1. The relation $u = F\overline{u}F^{-1}$ means that the corepresentation matrices u and $\overline{u}$ are similar, and the relation $F\overline{F} \in \mathbb{R}I_N$ is chosen because $u = F\overline{u}F^{-1}$ implies $\overline{u} = \overline{F}u\overline{F}^{-1}$, which implies $u = (F\overline{F})u(F\overline{F})^{-1}$, so $F\overline{F}$ is an intertwiner between u and itself. Hence, if we want u to be an irreducible corepresentation, Schur's lemma forces us to have $F\overline{F} = \lambda I_N$ for some $\lambda \in \mathbb{C}^\times$. As this means $\overline{F}/\lambda$ is the inverse of F, we must also have $\frac{1}{\lambda}\overline{F}F = I_N$, so $\overline{F}F = \lambda I_N$, but conjugating $F\overline{F} = \lambda I_N$ yields $\overline{F}F = \overline{\lambda} I_N$, hence $\lambda = \overline{\lambda}$ and we conclude $F\overline{F} \in \mathbb{R}I_N$. [Ban98]

4.6.2. THE QUANTUM PERMUTATION GROUP S_N^+

Definition 4.6.5

The **permutation quantum group** denoted S_N^+ or $A_s(N)$, depending on the dimension $N \in \mathbb{N}$, is the compact matrix quantum group $S_N^+ = (C(S_N^+), u)$, where $C(S_N^+)$ is the universal unital C^*-algebra generated by the elements $\{u_{ij}\}_{i,j=1}^N$ placed in the corepresentation matrix $u := (u_{ij})_{ij} \in M_N(C(S_N^+))$ with the relation that u is a *magic unitary matrix*, i.e. we have

$$C(S_N^+) := C^*\left(u_{ij}, 1 \leqslant i,j \leqslant N \mid u_{ij}^* = u_{ij} = u_{ij}^2, \sum_k u_{ik}u_{jk} = \sum_k u_{ki}u_{kj} = \delta_{ij}\ \forall i,j\right) \tag{4.15}$$

The relations imposed on the elements $(u_{ij})_{ij}$ in the definition of S_N^+ differ per paper. We identify five key relations that all hold for the generators of S_N^+:

- R1. The generators are self-adjoint, i.e. $u_{ij} = u_{ij}^*$.
- R2. The generators are idempotent, i.e. $u_{ij}^2 = u_{ij}$.
- R3. The matrix u is orthogonal, i.e. $uu^t = u^tu = 1$, i.e. $\sum_k u_{ik}u_{jk} = \delta_{ij} = \sum_k u_{ki}u_{kj}$.
- R4. The elements in the rows and columns of u are orthogonal to each other, i.e. $u_{ik}u_{jk} = u_{ki}u_{kj} = 0$ if $i \neq j$.
- R5. The rows and columns of u sum to 1, i.e. $\sum_k u_{ik} = \sum_k u_{kj} = 1$.

Timmermann [Tim08, Ex. 6.1.13] defines S_N^+ with the relations {R1,R2,R3,R4,R5}, Banica and Speicher [BS09, Def. 3.4] say that u has to be *magic*, which is defined by the relations {R1,R2,R3}, Weber [Web13] says that u has to be *magic* which is defined by the relations {R1,R3,R4,R5}, Freslon and Weber [FW16] define S_N^+ with the relations {R1,R2,R4,R5}, Weber defines S_N^+ in a different paper [Web17b] with relations {R1,R2,R5}, and Banica, Bichon and Collins [BBC06] define S_N^+ with relations {R1,R2,R4,R5}. All these definitions are equivalent:

Proposition 4.6.2

The following sets of relations are equivalent:

1. u_{ij} are projections and u is orthogonal : {R1,R2,R3}.
2. u_{ij} are self-adjoint, u is orthogonal, elements in rows and columns are orthogonal, rows and columns sum to 1 : {R1,R3,R4,R5}.
3. u_{ij} are projections, rows and columns sum to 1 : {R1,R2,R5}.
4. u_{ij} are projections, elements in rows and columns are orthogonal, rows and columns sum to 1 : {R1,R2,R4,R5}.

Proof.

1. {R1,R2,R3} implies {R1,R3,R4,R5} : By the fact that u_{ij} are projections and u is orthogonal, we have $\sum_k u_{ik} = \sum_k u_{ik}^2 = \sum_k u_{ik}u_{ik} = \delta_{ii} = 1$, similarly $\sum_k u_{kj} = 1$, hence {R1,R2,R3} implies R5. But 1 itself is a projection, and if projections sum up to a projection, they must be orthogonal, i.e. $u_{ik}u_{jk} = u_{ki}u_{kj} = 0$ if $i \neq j$, hence R4 follows.
2. {R1,R3,R4,R5} implies {R1,R2,R5} : By the fact that elements in the rows and columns of u are orthogonal, and then by the fact that the rows and columns of u sum to 1, we have $u_{ij}^2 = \sum_k u_{ij}u_{ik} = u_{ij}\sum_k u_{ik} = u_{ij}1 = u_{ij}$, hence R2 follows.
3. {R1,R2,R5} implies {R1,R2,R4,R5} : again, 1 itself is a projection, and if projections sum up to a projection, they must be mutually orthogonal, so R4 follows.
4. {R1,R2,R4,R5} implies {R1,R2,R3}: for the diagonal elements, by the fact that generators are idempotent and the rows and columns of u sum to 1, we have $\sum_k u_{ik}u_{ik} = \sum_k u_{ik}^2 = \sum_k u_{ik} = 1$, and similarly $\sum_k u_{ki}u_{ki} = 1$. For the off-diagonal elements, all the summands are 0 by the orthogonality of elements in the rows and columns. Hence, R3 follows.

□

Remark 4.6.2. Note that $C(S_N^+)$ is non-commutative for $N \geqslant 4$ [Web17a]. We can see this for $N = 4$ by considering the universality property of the universal C^* algebra $C(S_N^+)$. Start with a C^*-algebra A that is generated by two non-commuting projections p and q. In concrete terms, take for example $\mathcal{H} = \mathbb{C}^2$, and consider $p = |0\rangle\langle 0|$ and $q = \frac{1}{2}(|0\rangle + |1\rangle)(\langle 0| + \langle 1|)$. Then, consider $M_4(M_2(\mathbb{C}))$, in which we put the matrix

$$v := \begin{pmatrix} p_1 & 1-p_1 & & \\ 1-p_1 & p_1 & & \\ & & p_2 & 1-p_2 \\ & & 1-p_2 & p_2 \end{pmatrix}. \tag{4.16}$$

Clearly, v satisfies the relations that were imposed on $u = (u_{ij})_{ij}$ if we swap u_{ij} for v_{ij}, so by the universality property of $C(S_N^+)$, we get a $*$-homomorphism φ that sends u_{ij} to v_{ij}. But then, as $p_1p_2 \neq p_2p_1$, $\varphi(u_{11}u_{33} - u_{33}u_{11}) \neq 0$, whence $u_{11}u_{33} \neq u_{33}u_{11}$, and thus $C(S_4^+)$ is not commutative.

Having introduced the relevant compact matrix quantum groups, the next step is understanding their corepresentation theory. By Theorem 4.6.1, all their irreducible corepresentations are contained in the repeated tensor product of their fundamental corepresentation u, its conjugate $\overline{u}$ and the trivial corepresentation. In the case of O_N^+ and S_N^+, as their generators are self-adjoint, this simplifies even further, because $\overline{u} \simeq u$, so we only need to understand what the irreducible subrepresentations are of tensor products of the form $u^{\otimes k}$ for $k \in \mathbb{Z}_{\geqslant 0}$. The CMQG's that we are interested in have a particularly nice structure in this regard: they are examples of so-called *partition* or *easy* compact quantum groups, whose intertwiner spaces can be described by certain categories of partitions. We investigate this in the next section.

4.7. (EASY) PARTITION QUANTUM GROUPS

By the Tannaka-Krein duality theorem, we know that we can view Rep $\mathbb{G}$, the category of all finite-dimensional unitary representations with intertwiners as morphisms, as the dual to a compact quantum group $\mathbb{G}$. For compact matrix quantum groups $\mathbb{G} = (C(\mathbb{G}), u)$ with $\overline{u} \simeq u$ (such as O_N^+ and S_N^+), the fundamental corepresentation matrix u generates Rep $\mathbb{G}$, hence the only data we need to investigate in the complete concrete monoidal W^*-category Rep $\mathbb{G}$ are the intertwiner spaces $\mathrm{Hom}_{\mathbb{G}}(u^{\otimes k}, u^{\otimes l})$ for $k, l \in \mathbb{N}_0$. They are often abbreviated as $\mathrm{Hom}_{\mathbb{G}}(k, l)$, or if $\mathbb{G}$ is clear from context, simply by $\mathrm{Hom}(k, l)$. We can now precisely state:

Definition 4.7.1

[FW16, Def 3.5] An *easy (partition) quantum group*, is a compact matrix quantum group $\mathbb{G} = (C(\mathbb{G}), u)$ whose fundamental corepresentation satisfies $u \simeq \overline{u}$, and is orthogonal in the sense that $uu^t = u^t u = 1$, and whose intertwiner spaces $\mathrm{Hom}_{\mathbb{G}}(k, l)$ are of the form:

$$\mathrm{Hom}_{\mathbb{G}}(k, l) = \mathrm{span}\{T_p \ : \ p \in \mathcal{C}(k, l)\} = \{T_p \ : \ p \in \mathcal{C}_{\mathrm{lin}}[N](k, l)\}, \tag{4.17}$$

where $\mathcal{C}(k, l)$ is a subset of all partitions $\mathcal{P}(k, l)$, such that the collection $\mathcal{C}$ of all $\mathcal{C}(k, l)$ for $k, l \in \mathbb{N}_0$ is a *category of partitions*. Here, $\mathcal{C}_{\mathrm{lin}}[N]$ is its associated easy linear category of partitions. The objects T_p are the associated *partition maps*.

As a brief reminder, the partition maps T_p are defined as

$$T_p : \mathcal{H}^{\otimes k} \to \mathcal{H}^{\otimes l} \quad : \quad e_{i_1} \otimes \cdots \otimes e_{i_k} \mapsto \sum_{j_1, \ldots, j_l = 1}^{n} \delta_p(\boldsymbol{i}, \boldsymbol{j}) e_{j_1} \otimes \cdots \otimes e_{j_l}, \tag{4.18}$$

where $\delta_p(\boldsymbol{i}, \boldsymbol{j}) \in \{0, 1\}$ and $\delta_p(\boldsymbol{i}, \boldsymbol{j}) = 1$ iff the through-blocks in the diagram of p with the upper nodes replaced by $\boldsymbol{i}$ and the lower nodes by $\boldsymbol{j}$ connect equal indices. For details on what constitutes a partition, please see Definition 2.3.1, for (linear) categories of partitions, please see Section 2.3.2, and for the partition maps, please see Section 2.3.3.

Both the *free orthogonal quantum group* O_N^+ and the *quantum permutation group* S_N^+ that we introduced in the previous sections are examples of (easy) partition quantum groups. We will investigate their intertwiner spaces and fusion rules in the next chapter.

5

The Quantum Groups O_N^+ and S_N^+

In this thesis, we are interested in the compact matrix quantum groups O_N^+ and S_N^+ that were introduced in the previous chapter. They are *easy (partition) quantum groups*, see Section 4.7, which means that the intertwiner spaces $\mathrm{Hom}_{\mathbb{G}}(u^{\otimes k}, u^{\otimes l})$ can be described in terms of partition maps (where $\mathbb{G}$ is O_N^+ or S_N^+, u is their fundamental representation, and $k, l \in \mathbb{N}_0$). We discuss the intertwiner spaces and fusion rules of O_N^+ and S_N^+ in detail in this chapter.

5.1. The intertwiner spaces and fusion rules of O_N^+

In the case of O_N^+ with $N \geqslant 2$, the intertwiner spaces are described by the category of *non-crossing pair partitions* NC_2 [Ban99, BC07b], or equivalently in terms of the *linear category of non-crossing pair partitions* $\mathrm{NC}_{2,\mathrm{lin}}[N]$, see Definition 2.3.8:

Theorem 5.1.1: Intertwiner spaces for O_N^+

$$\mathrm{Hom}_{O_N^+}\left(u^{\otimes k}, u^{\otimes l}\right) = \mathrm{span}\left\{T_p \ : \ p \in \mathrm{NC}_2(k,l)\right\} = \left\{T_p \ : \ p \in \mathrm{NC}_{2,\mathrm{lin}}[N](k,l)\right\}. \tag{5.1}$$

Banica also established in [Ban98] a classification of the irreducible representations and their fusion rules for $N \geqslant 2$:

Theorem 5.1.2: Fusion rules for O_N^+

The irreducible corepresentations v_k of O_N^+ are self-adjoint, and can be indexed by $\mathbb{N}_0$. We let $v_0 = 1$ (the trivial corepresentation) and $v_1 = u$ (the fundamental corepresentation). Furthermore, for all $k, l \in \mathbb{N}_0$, we have the *fusion rules*:

$$v_k \otimes v_l \simeq v_{|k-l|} \oplus v_{|k-l|+2} \oplus \cdots \oplus v_{k+l}. \tag{5.2}$$

Hence, the fusion rules are identical to those of $\mathrm{SU}(2)$.

If we label the corepresentation Hilbert spaces by $\mathcal{H}_k$, and their dimension $\dim \mathcal{H}_k$ by d_k, we see the recurrence relation:

$$d_k d_l = d_{|k-l|} + d_{|k-l|+2} \cdots + d_{k+l}, \tag{5.3}$$

where $\mathcal{H}_0 = \mathbb{C}$ so $d_0 = 1$, and $\mathcal{H}_1 = \mathbb{C}^N$ so $d_1 = N$, where N is the dimension of the square matrix F. Then, putting $l = 1$, we find the recurrence

$$d_{k+1} = d_k d_1 - d_{k-1} \implies d_{k+1} = N d_k - d_{k-1}. \tag{5.4}$$

One can prove that the solution to this recurrence equation for $N \geqslant 3$ is: [1]

$$d_k = \frac{q_0^{k+1} - q_0^{-(k+1)}}{q_0 - q_0^{-1}} \quad \text{where} \quad q_0^{-1} = N/2 + \tfrac{1}{2}\sqrt{N^2-4} \text{ is the solution to } q_0 + q_0^{-1} = N. \tag{5.6}$$

The authors in [BC16b] and [BCLY20] denote this by $d_k = [k+1]_{q_0}$, where the *quantum integers* $[n]_{q_0}$ for $n \in \mathbb{N}_0$ are defined by $[n]_\alpha = \frac{\alpha^n - \alpha^{-n}}{\alpha - \alpha^{-1}}$, where $\alpha \in \mathbb{R}$ is called the *quantum parameter*. Note that multiple definitions of quantum numbers exist – for example, q-integers are often defined as $(n)_q := \frac{q^n-1}{q-1} = \sum_{k=0}^{n-1} q^k$ [SQ15, Nat02], although $[n]_q$ is sometimes also referred to as a q-integer [Lus10, Kas95]. Stum and Quirós call $[n]_q$ the *symmetric quantum state*, and note that the two definitions are closely related by $[n]_q = q^{1-n}(n)_{q^2}$ [SQ15].

Note that in the case of $N = 2$, $q_0 = 1$, and we have $d_k = \lim_{q_0 \to 1}[k+1] = k+1$. In the case of $N \geqslant 3$, we instead see that d_k grows approximately as q_0^{-k} and $q_0^{-1} \to N$ for large N and fixed k [BC16b].

5.1.1. THE TEMPERLEY-LIEB CATEGORY AND ALGEBRAS

The concrete monoidal W^* category Rep $\mathbb{G}$ of finite-dimensional unitary representations of O_N^+ is intimately related to the so-called *Temperley-Lieb category* $\mathrm{TL}(\delta)$ with objects $k \in \mathbb{N}_0$, and morphism classes $\mathrm{TL}_{k,l}(\delta)$ for $k, l \in \mathbb{N}_0$. As the fundamental representation u of O_N^+ satisfies $u \simeq \overline{u}$ (by definition, as we demand that all generators u_{ij} are self-adjoint), we know that Rep O_N^+ is generated by the concrete monoidal W^*-category $R_0(O_N^+)$ whose objects are $u^{\otimes k}$ for $k \in \mathbb{N}_0$, and whose morphisms are the intertwiners $\mathrm{Hom}(u^{\otimes k}, u^{\otimes l})$. A fundamental result from Banica in [Ban98] tells us that there exists an isomorphism of monoidal W^*-categories

$$\pi : R_0(O_N^+) \to \mathrm{TL}(N) \quad : \quad \pi(u^{\otimes k}) = k \quad \text{and} \quad \pi(\mathrm{Hom}(u^{\otimes k}, u^{\otimes l})) = \mathrm{TL}_{k,l}(N), \tag{5.7}$$

such that $\iota \in \mathrm{Hom}(u,u) \mapsto \iota \in \mathrm{TL}_{1,1}(N)$ and $\sum_{i=1}^N e_i \otimes e_i \in \mathrm{Hom}(1, u^{\otimes 2}) \mapsto \cap \in \mathrm{TL}_{0,2}(N)$. Let us investigate this Temperley-Lieb category.

Abstractly, for a given *loop parameter* $\delta \in \mathbb{C}^\times$, the Temperley-Lieb category $\mathrm{TL}(\delta)$ is a strict C^*-tensor category, whose self-dual irreducible objects can be labelled as $\{0, 1, \ldots\} = \mathbb{N}_0$, where 0 denotes the unit for the tensor category, and the morphism classes $\mathrm{Hom}(1^{\otimes k}, 1^{\otimes l}) =: \mathrm{TL}_{k,l}(\delta)$ are generated by the two morphisms $\iota \in \mathrm{Hom}(1,1)$ (the identity morphism) and $\cap \in \mathrm{Hom}(0, 1 \otimes 1)$, where we demand $\cup\cap = \delta \in \mathrm{Hom}(0,0) = \mathbb{C}$ (with $\cup = \cap^*$) and $(\iota \otimes \cup)(\cap \otimes \iota) = (\cup \otimes \iota)(\iota \otimes \cap) = \iota$, the so-called *snake equation* [BC20, BCLY20]. Luckily, a diagrammatic interpretation of $\mathrm{TL}(\delta)$ exists – as a matter of fact, we have already encountered it, as we will show an equivalence with the category $\mathrm{NC}_{2,\mathrm{lin}}[\delta]$. Let us commence with the Temperley-Lieb *algebras*:

Definition 5.1.1: Temperley-Lieb algebras

The *Temperley-Lieb algebra* $\mathrm{TL}_k(\delta) := \mathrm{TL}_{k,k}(\delta)$ with *loop parameter* $\delta \in \mathbb{C}^\times$ and parameter $k \in \mathbb{N}$ is the unital (and associative) algebra over $\mathbb{C}$ that is generated by the elements $\{u_i\}_{i=1}^{k-1}$ and the unit 1 such that $u_i u_j = u_j u_i$ whenever $|i-j| \geqslant 2$, and $u_i u_{i+1} u_i = u_i$, and $u_i^2 = \delta u_i$.

[1] The characteristic polynomial associated with this recurrence relation is $p(\lambda) = \lambda^2 - N\lambda + 1$, whose roots are $q_1 = N/2 + \sqrt{N^2-4}/2$ and $q_2 = N/2 - \sqrt{N^2-4}/2$. The general solution to the recurrence equation is then $d_k = c_1 q_1^k + c_2 q_2^k$ for some constants $c_1, c_2 \in \mathbb{R}$. Using $d_0 = 1$ yields $c_1 = 1 - c_2$, and $d_1 = N$ together with $q_2 = N - q_1$ yields $N = q_1 - c_2 q_1 + c_2 N - c_2 q_1$, i.e. $c_2 = (N-q_1)/(N-2q_1)$ and $c_1 = 1 - c_2 = -q_1/(N-2q_1)$. We can then rewrite:

$$d_k = \frac{-q_1}{N-2q_1} q_1^k + \frac{N-q_1}{N-2q_1}(N-q_1)^k = \frac{(N-q_1)^{k+1} - q_1^{k+1}}{N-2q_1} = \frac{q_1^{k+1} - q_1^{-(k+1)}}{q_1 - q_1^{-1}}, \tag{5.5}$$

where in the last step we used the fact that $q_1 + q_1^{-1} = N$ (which follows from $p(\lambda) = 0$) to find both $N - q_1 = q_1^{-1}$ and $2q_1 - N = q_1 - q_1^{-1}$. Letting $q_0^{-1} = q_1$ yields the claim, where q_0 is the quantum parameter q in [BC16b], and the quantum parameter q_0 in [BCLY20].

It is well-known that $\mathrm{TL}_k(\delta)$ is a finite-dimensional $\mathbb{C}$-algebra with $\dim \mathrm{TL}_k(d) = \frac{(2k)!}{k!(k+1)!}$ (the k-th Catalan number), if δ is not twice the real part of a root of unity [Jon83]. Furthermore, $\mathrm{TL}_k(\delta)$ can be seen as an involutive $*$-algebra under the declaration that $u_i^* = u_i$ for all $1 \leqslant i \leqslant k-1$, which has non-trivial $*$-representation into a C^*-algebra if $\delta \in [2, \infty) \cup \{2\cos(\pi/n) \ : \ n \in \mathbb{N}, n \geqslant 3\}$ [BC16a].

The Temperley-Lieb algebras can be realized as a $\mathbb{C}$-algebra over non-crossing pair partitions as follows: recall from Section 2.3 that $\mathrm{NC}_2(k,l)$ denotes the collection of all partitions between $\{1,\dots,k\}$ and $\{1',\dots,l'\}$ such that all *blocks* consist of precisely two points (i.e., pairs), and none of the blocks cross. In literature, partitions $p \in \mathrm{NC}_2(k,l)$ are also referred to as *Temperley-Lieb diagrams* D_p, such as:

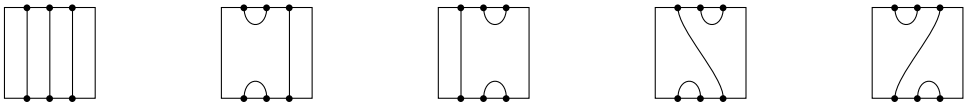


Figure 5.1: The five Temperley-Lieb diagrams in $\mathrm{NC}_2(3,3)$

We know that we can endow NC_2 with a $*$-algebra structure, which we have called $\mathrm{NC}_{2,\mathrm{lin}}[\delta]$ with the loop parameter δ such that $qp = \delta^{\mathrm{rl}(q,p)}(p \circ q)$ (where $p \circ q$ is the usual *diagram composition*, see Definition 2.3.3), see Section 2.3.2. An example of the composition is then:

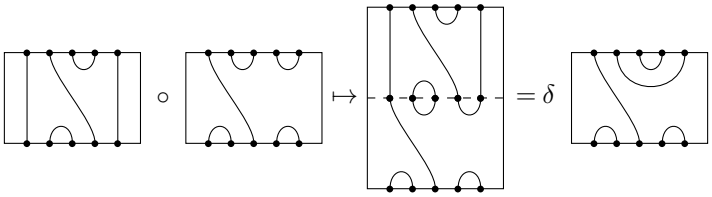


Figure 5.2: Composition of Temperley-Lieb diagrams

Remarkably, the $*$-algebra $\mathrm{NC}_{2,\mathrm{lin}}[\delta](k,k)$ is in fact isomorphic to the Temperley-Lieb algebra $\mathrm{TL}_k(\delta)$ as defined in Definition 5.1.1, under the following isomorphism [BC16a]:

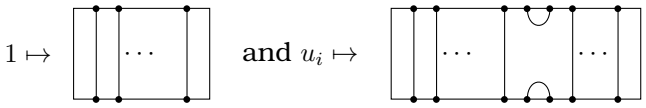


Figure 5.3: Isomorphism mapping 1 to $p = |^{\otimes k}$, and u_i to $p = |^{\otimes(i-1)} \otimes c \otimes |^{\otimes(k-i-1)}$, where the tensor operation $\otimes$ is defined in Definition 2.3.4, and where $c = \cap\cup$ is the "cap-cup" partition $c =$  .

Therefore, from now on we refer to both $\mathrm{TL}_k(\delta)$ and $\mathrm{NC}_{2,\mathrm{lin}}[\delta](k,k)$ as "the k-th Temperley Lieb algebra (with loop parameter δ)". Lastly, the partitions $\cup \in \mathrm{NC}_2(2,0)$ and $\cap \in \mathrm{NC}_2(0,2)$ play an important role in the non-crossing pair partition case: for example, the only pairs of indices in $p \in \mathrm{NC}_2(k,l)$ that contain two upper indices or lower indices must be (nested) cups or caps, respectively. Therefore, we define:

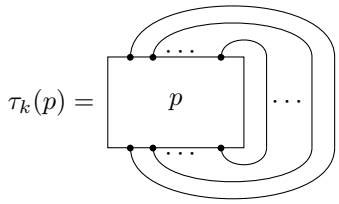
Definition 5.1.2

The cup partition $\cup$ is the unique partition $\cup \in \mathrm{NC}_2(2,0)$, and the cap partition $\cap$ is the unique partition $\cap \in \mathrm{NC}_2(0,2)$. Their nested variants, $\cup^k \in \mathrm{NC}_2(2k,0)$ and $\cap^k \in \mathrm{NC}_2(0,2k)$ are those partitions consisting of $\{i, 2k-i+1\}$ for $1 \leqslant i \leqslant k$, or with primed numbers in the case of $\cap^k$.

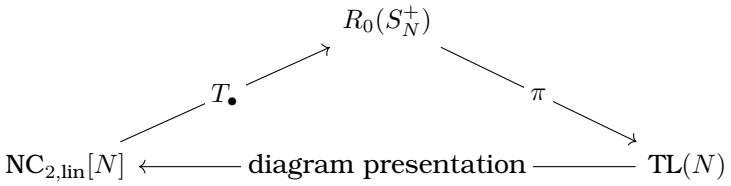
We can then also define:

Definition 5.1.3: Markov trace

The *Markov trace* $\tau_k : \mathrm{TL}_k(\delta) \to \mathbb{C}$ is a tracial linear functional that sends a diagram $p \in \mathrm{TL}_k(\delta)$ to its *tracial closure* $\tau_k(p) \in \mathbb{C}$ by looping all its lower points to its upper points. We can write $\tau_k(p) = \cup^k (p \otimes |^{\otimes k}) \cap^k$, which evaluates to the value $\delta^{\#\text{loops}}$, where #loops is the number of loops (i.e., connected components) in the diagram of $\cup^k (p \otimes |^{\otimes k}) \cap^k$:



In general, the morphism classes $\mathrm{TL}_{k,l}(\delta)$ for general $k, l \in \mathbb{N}_0$ can be bundled into a universal graded C^*-algebra $(\mathrm{TL}_{k,l}(\delta))_{k,l\in\mathbb{N}_0}$, and defined in terms of generators and relations, see [Bra12]. And analogous to the previous exposition, an isomorphism exists between $\mathrm{TL}_{k,l}(\delta)$ and $\mathrm{NC}_{2,\mathrm{lin}}[\delta](k,l)$ that respects this structure, yielding an equivalence of monoidal $*$-categories between $\mathrm{TL}(\delta) = \{\mathbb{N}_0, \{\mathrm{TL}_{k,l}(\delta)\}_{k,l\in\mathbb{N}_0}\}$ and $\mathrm{NC}_{2,\mathrm{lin}}[\delta] = \{\mathbb{N}_0, \{\mathrm{NC}_{2,\mathrm{lin}}[\delta][k,l]\}_{k,l\in\mathbb{N}_0}\}$. As a matter of fact, in literature one sometimes encounters the diagrammatic presentation as the *definition* of the Temperley-Lieb category (e.g. [BSA18]). We remind the reader that $R_0(S_N^+)$ is the concrete monoidal W^*-category with object $u^{\otimes k}$ for $k \in \mathbb{N}_0$ and with morhpism classes that are the intertwiner spaces $\mathrm{Hom}(u^{\otimes k}, u^{\otimes l})$. In summary, we then have:



Here, we used the fact that $\mathrm{Hom}(u^{\otimes k}, u^{\otimes l}) = \{T_p : p \in \mathrm{NC}_{2,\mathrm{lin}}[N](k,l)\}$ for the left-most arrow, which we stated in Theorem 5.1.1. We know that $T_\bullet$ is both full (by definition) and essentially surjective, and by Lemma 2.3.1 it is faithful for $N \geqslant 4$, and in that case the diagram above consists of monoidal $*$-isomorphisms such that their composition is the identity functor on R_0, as they agree on the generators of R_0 . Note that this result *also* holds for $N \geqslant 2$, in contradiction with [FW16, Lemma 4.16]: [2]

Lemma 5.1.1: Independence of partition maps for $N \geqslant 2$ in NC_2

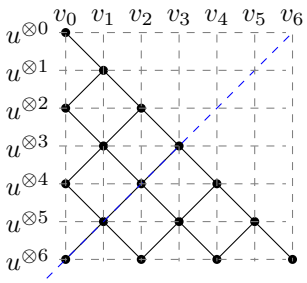
For $N \geqslant 2$, the maps $(T_p)_{p\in\mathrm{NC}_2(k,l)}$ are linearly independent.

Proof. We prove this by dimensionality: we know $\#\mathrm{NC}_2(k,l) = C_{(k+l)/2} = \frac{1}{k+l+1}\binom{2(k+l)}{k+l}$ if $k+l$ is even, else $\#\mathrm{NC}_2(k,l) = 0$. On the other hand, by the identification $\mathrm{span}\{T_p : p \in \mathrm{NC}_2(k,l)\} = \mathrm{Hom}(u^{\otimes k}, u^{\otimes l})$, we can look at the dimension of $\mathrm{Hom}(u^{\otimes k}, u^{\otimes l})$. By rotating the partitions, it

[2] The small error lies in the fact that Freslon and Weber prove that for the category of partitions NC, the maps $(T_p)_{p\in\mathrm{NC}(k,l)}$ are *not* linearly independent for $N < 4$, but this does not imply that for a smaller category, for example NC_2, the maps $(T_p)_{p\in\mathrm{NC}_2(k,l)}$ must also be linearly dependent. The converse, however, *does* hold: if $(T_p)_{p\in\mathrm{NC}(k,l)}$ are linearly independent, then surely picking a *smaller* category still yields linearly independent maps. This is corroborated by the fact that a rich theory has been developed around the *Weingarten calculus*, in which the *Weingarten matrix* W_{kn} is defined as the inverse of the matrix $G_{kn} = [G_{kn}(p,q)]_{p,q\in\mathrm{TL}_k(n)}$, where $G_{kn}(p,q) = N^{\mathrm{rl}(p,q)}$ [BC07a]. Brannan and Collins note in [BC16a] that this matrix is invertible (and give a list of references) for all $k \geqslant 1$ and $N \geqslant 2$, whilst Lemma 4.16 in [FW16] applied to NC_2 would claim that the same Gram matrix G_{kn} is singular for $N < 4$.

suffices to consider $\mathrm{Hom}(u^{\otimes m}, 1)$ and partitions in $\mathrm{NC}_2(m,0)$ with $m = k+l$. Using the fusion rule $v_1 \otimes v_a = v_{a-1} \oplus v_{a+1}$ for $a \in \mathbb{N}$, we repeatedly tensor $v_1 = u$ to $u^{\otimes m}$ and use the fusion rules to decompose $u^{\otimes m+1}$ into irreducibles. Some contemplation yields the table of multiplicities of v_i in $u^{\otimes m}$, and a graphical diagram where each vertex in column v_i and row $u^{\otimes j}$ shows that $v_i \subset u^{\otimes j}$, and each edge from $(u^{\otimes j}, v_i)$ to $(u^{\otimes j+1}, v_{i-1})$ and $(u^{\otimes j+1}, v_{i+1})$ indicates the fusion rules $v_i \otimes v_1 \simeq v_{i-1} \oplus v_{i+1}$.

	v_0	v_1	v_2	v_3	v_4	v_5	v_6
$u^{\otimes 0}$	1						
$u^{\otimes 1}$		1					
$u^{\otimes 2}$	1		1				
$u^{\otimes 3}$		2		1			
$u^{\otimes 4}$	2		3		1		
$u^{\otimes 5}$		5		4		1	
$u^{\otimes 6}$	5		9		5		1



We have purposefully written the decomposition in this form: it becomes clear that the multiplicity of v_0 in the decomposition of $u^{\otimes 2k}$ is the number of unique paths along the graph one can take from vertex $(u^{\otimes 0}, v_0)$ to $(u^{\otimes 2k}, v_0)$, where one is not allowed to cross the diagonal dashed blue line. It is also clear that $v_0 \not\subset u^{\otimes 2k+1}$. It is well-known that this number of unique paths is precisely the k-th Catalan number, see e.g. [SF99]. As all v_i are irreducible representations, we have $\mathrm{Hom}(u^{\otimes m}, 1) \simeq \mathrm{Hom}(v_0^{\oplus t(m)}, v_0)$, where $t(m)$ is the multiplicity of v_0 in $u^{\otimes m}$. But then we recover $\dim \mathrm{Hom}(u^{\otimes m}, 1) = C_{m/2}$ if m is even, and else 0. This precisely corresponds to $\#\mathrm{NC}_2(m,0)$, so by a dimensionality argument all $(T_p)_{p \in \mathrm{NC}_2(m,0)}$ must be linearly independent. Note that this construction is independent of the choice of $N \geqslant 2$.

□

Having firmly established the connection between the non-crossing pair partitions, the Temperley Lieb category, and the intertwiner spaces $\mathrm{Hom}(u^{\otimes k}, u^{\otimes l})$, we can reap the fruits of our labour and utilize some useful technology from the Temperley-Lieb theory: specifically, we can consider the Jones-Wenzl projections.

5.1.2. JONES-WENZL PROJECTION

The *Jones-Wenzl projections* are exceptionally useful projections in the Temperley-Lieb algebras [BC16a]:

Theorem 5.1.3: Jones-Wenzl projections

Given a finite-dimensional Temperley-Lieb algebra $\mathrm{TL}_k(\delta)$ for $k \in \mathbb{N}$ and $\delta \in \mathbb{C}^\times$ but $\delta \notin \{2\cos(\pi/n)\}_{2 \leqslant n \leqslant k+1}$, then there exists a unique non-zero idempotent $p_k \in \mathrm{TL}_k(\delta)$ called the *Jones-Wenzl projection* that has the property that $u_i p_k = p_k u_i = 0$ for all $1 \leqslant i \leqslant k-1$.

For example, the reader can check that $p_1 = |$, and $p_2 = |^{\otimes 2} - \frac{1}{\delta} c$, where $c = \cap\cup$ is the "cap-cup" partition $c =$ [image]. Using $u_1^2 = \delta u_1$, it directly follows that $p_2 u_2 = u_2 p_2 = 0$. The fact that $p_2^2 = p_2$ follows similarly.

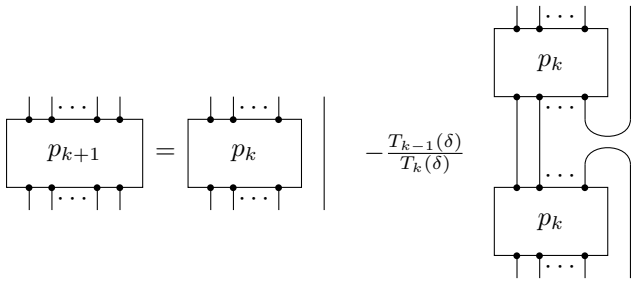
To determine the k-th Jones-Wenzl projection, we can use the *Wenzl recursion formula*. Firstly, we can embed the k-th Jones-Wenzl projection p_k into $\mathrm{TL}_{k+1}(\delta)$ by adding a vertical strand to the right in its diagram, i.e. we map p_k to $p_k \otimes | \in \mathrm{TL}_{k+1}(\delta)$. Then, we have [BC16a, Mor17, BLS19]:

Theorem 5.1.4: Wenzl recurrence relation

The Jones-Wenzl projection p_k satisfies the *Wenzl recursion formula*:

$$p_{k+1} = p_k \otimes | - \tfrac{T_{k-1}(\delta)}{T_k(\delta)} (p_k \otimes |) u_k (p_k \otimes |), \tag{5.8}$$

where T_k is the k-th Type 2 Chebyshev polynomial defined by $T_0(x) = 1, T_1(x) = x$, $xT_k(x) = T_{k+1}(x) + T_{k-1}(x)$. Graphically, this is displayed as:



Lemma 5.1.2

[Mor17, Lemma 2.1] The coefficient of the identity 1_k in the Jones-Wenzl projection $p_k \in \mathrm{TL}_k(\delta) = \langle 1_k, u_i \ (1 \leqslant i \leqslant k-1) \rangle$ is equal to 1, independent of k.

Proof. If we denote the two-sided ideal generated by $\{u_1, \ldots, u_{k-1}\}$ in $\mathrm{TL}_k(\delta)$ by $\mathcal{I}_k$, then we can express $p_k = a1_k + f_k$ for some $f_k \in \mathcal{I}_k$ and $a \in \mathbb{C}$. By the fact that $p_k u_i = 0$ for all $1 \leqslant i \leqslant k-1$, we see that $p_k \mathcal{I}_k = 0$. Hence,

$$p_k = p_k^2 = p_k(a1_k + f_k) = ap_k + 0, \tag{5.9}$$

so $p_k = ap_k$, which implies $a = 1$. □

Lemma 5.1.3: Absorption property of Jones-Wenzl projections

[BLS19, Prop. 2.4] We have the following *absorption properties* in the Temperley-Lieb algebra $\mathrm{TL}_k(\delta)$ with the Jones-Wenzl projection p_k:

$$(p_k \otimes 1_m)p_{k+m} = (1_k \otimes p_m)p_{k+m} = p_{k+m} = p_{k+m}(1_k \otimes p_m) = p_{k+m}(p_k \otimes 1_m). \tag{5.10}$$

Proof. By Lemma 5.1.2, we can express $p_k = 1_k + f_k$ for some $f_k \in \mathcal{I}_k$, where $\mathcal{I}_l$ is the two-sided ideal generated by $\{u_1, \ldots, u_{l-1}\}$ in $\mathrm{TL}_l(\delta)$ for some $l \in \mathbb{N}$. Then $p_k \otimes 1_m = 1_{k+m} + f_k \otimes 1_m$, so $f_k \otimes 1_m \in \mathcal{I}_{k+m}$. But $p_{k+m}\mathcal{I}_{k+m} = 0$, so $p_{k+m}(p_k \otimes 1_m) = p_{k+m}$. The other identities follow with analogous arguments. □

It is well-known that we can identify the irreducible representations v_k with $p_k u^{\otimes k}$, where $p_k \in \mathrm{TL}_k(N)$ (so, we pick $\delta = N$), with representation spaces $\mathcal{H}_k := p_k \mathcal{H}_1^{\otimes k}$, where $\mathcal{H}_1 := \mathbb{C}^N$ is the representation space of the fundamental representation [Bra12, BC16b, BCLY20]. To be pedantic, these equations suppress the monoidal $*$-isomorphism π between R_0 and $\mathrm{TL}(N)$, as we should pull p_k back to $\pi^{-1}(p_k) \in \mathrm{Hom}(u^{\otimes k}, u^{\otimes k})$ to make sense of $p_k u^{\otimes k}$. As $T_\bullet$ is a monoidal $*$-isomorphism for $N \geqslant 2$, we can actually identify p_k with $T_{p_k} \in \{T_q : q \in \mathrm{NC}_{2,\mathrm{lin}}[N](k,k)\}$, which yields the same operator in $\mathrm{Hom}(u^{\otimes k}, u^{\otimes k})$. However, outside this section, we will often forget about these monoidal $*$-isomorphisms and simply write p_k for the operator T_{p_k} that lives in $\mathrm{Hom}(u^{\otimes k}, u^{\otimes k})$.

5.2. THE INTERTWINER SPACES AND FUSION RULES OF S_N^+

The intertwiner spaces of the quantum permutation group S_N^+ for $N \geqslant 4$ are related to the non-crossing partitions, in the following sense [BC07a, BC07b, BBC06]:

Theorem 5.2.1: Intertwiner spaces for S_N^+

$$\mathrm{Hom}_{S_N^+}(k,l) = \mathrm{Hom}_{S_N^+}\left(u^{\otimes k}, u^{\otimes l}\right) = \mathrm{span}\left\{T_p \ : \ p \in \mathrm{NC}(k,l)\right\} \tag{5.11}$$

The fusion rules for S_N^+ are also known [Ban99, Ban02, BBC06]:

Theorem 5.2.2: Fusion rules for S_N^+

For $N \geqslant 4$, the irreducible corepresentations v_k of S_N^+ can be indexed by $\mathbb{N}_0$. We set $v_0 = 1$ (the trivial corepresentation) and $v_1 = u$ (the fundamental corepresentation). For all $k, l \in \mathbb{N}_0$, we have the *fusion rules*:

$$v_k \otimes v_l \simeq v_{|k-l|} \oplus v_{|k-l|+1} \oplus \cdots \oplus v_{k+l}. \tag{5.12}$$

Hence, the fusion rules are identical to those of $\mathrm{SO}(3)$.

We again label the Hilbert spaces of the corepresentations v_k by $\mathcal{H}_k$. their dimension $\dim \mathcal{H}_k$ by d_k, and we then find the recurrence relation

$$d_k d_l = d_{|k-l|} + d_{|k-l|+1} + \cdots + d_{k+l}, \tag{5.13}$$

with $d_0 = 1$ as $\mathcal{H}_0 = \mathbb{C}$ and $d_1 = N - 1$ as $\dim(u) = N = 1 + d_1$. Using $l = 1$, we find

$$d_k d_1 = d_{k-1} + d_k + d_{k+1} \implies d_{k+1} - (d_1 - 1)d_k + d_{k-1} = 0. \tag{5.14}$$

One can prove that the solution to this recurrence relation for $N \geqslant 4$ yields [3]

$$d_k = [2k+1]_{\hat{q}} = \frac{\hat{q}^{2k+1} - \hat{q}^{-(2k+1)}}{\hat{q} - \hat{q}^{-1}} \quad \text{where} \quad \hat{q} + \hat{q}^{-1} = \sqrt{N}. \tag{5.16}$$

As the fundamental representation u of S_N^+ satisfies $u \simeq \overline{u}$ (because all generators u_{ij} are self-adjoint), we know that Rep S_N^+ is generated by the concrete monoidal W^*-categroy $R_0(S_N^+)$ whose objects are $u^{\otimes k}$ for $k \in \mathbb{N}_0$, and whose morphism classes are the intertwiner spaces $\mathrm{Hom}(u^{\otimes k}, u^{\otimes l})$, analogous to the O_N^+ case covered in Section 5.1.1. Furthermore, the morphisms in $R_0(S_N^+)$ are generated by the multiplication $m \in \mathrm{Hom}(u^{\otimes 2}, u)$, the unit $\nu \in \mathrm{Hom}(1, u)$ and the identity $\mathrm{id}_B \in \mathrm{Hom}(u, u)$ [Bra12]. In this case, Banica's fundamental result in [Ban02] tells us that we can identify this concrete W^*-category $R_0(S_N^+)$ with the *two-cabled Temperley-Lieb monoidal W^*-category* $\mathrm{TL}^2(\delta) = \{\mathbb{N}, \{\mathrm{TL}_{2k,2l}(\delta)\}_{k,l\in\mathbb{N}}\}$ *with loop parameter* δ. Here, we see $\mathrm{TL}^2(\delta)$ as a subcategory of *Temperley-Lieb category* $\mathrm{TL}(\delta) := \{\mathbb{N}_0, \{\mathrm{TL}_{k,l}(\delta)\}_{k,l\in\mathbb{N}_0}\}$ introduced in Section 5.1.1, which also showed that we can think of $\mathrm{TL}(\delta)$ as $\mathrm{NC}_{2,\mathrm{lin}}[\delta]$. To

[3] The characteristic polynomial associated with this recurrence relation is $p(\lambda) = \lambda^2 - (d_1 - 1)\lambda + 1$, whose roots are $q_{1,2} = \frac{d_1 - 1}{2} \pm \frac{1}{2}\sqrt{(d_1-1)^2 - 4}$. As $q_1 - q_2 = d_1 - 1$, and we can factor $p(\lambda) = 0$ as $\lambda(\lambda - (d_1 - 1)) = -1$, we see $q_1 q_2 = 1$, whence $q_2 = q_1^{-1}$. The general solution to the recurrence relation is then $d_k = aq_1^k + bq_1^{-k}$. With $d_0 = 1$, we learn $a + b = 1$, so let $b = 1 - a$. Filling in $k = 1$ yields $aq_1 + (1-a)q_1^{-1} = d_1$. This yields $a = (1 + q_1)/(q_1 - q_1^{-1})$, and rewriting d_k yields

$$d_k = \tfrac{1+q_1}{q_1 - q_1^{-1}} q_1^k - \tfrac{1+q_1^{-1}}{q_1 - q_1^{-1}} q_1^{-k} = \tfrac{q_1^k - q_1^{-k}}{q_1 - q_1^{-1}} + \tfrac{q_1^{k+1} - q_1^{-(k+1)}}{q_1 - q_1^{-1}} = [k]_{q_1} + [k+1]_{q_1}. \tag{5.15}$$

Instead of using q_1 that satisfies $q_1 + q_1^{-1} = d_1 - 1 = N - 2$, we can choose $\hat{q}$ satisfying $\hat{q} + \hat{q}^{-1} = \sqrt{N}$. Then, we see $q_1 + q_1^{-1} = \hat{q}^2 + \hat{q}^{-2}$ and $q_1 - q_1^{-1} = \sqrt{(N-2)^2 - 4} = (\hat{q} + \hat{q}^{-1})(\hat{q} - \hat{q}^{-1}) = \hat{q}^2 - \hat{q}^{-2}$, thus $q_1 = \hat{q}^2$. Rewriting the previous expression for d_k and using $1 + \hat{q}^2 = \hat{q}(\hat{q}^{-1} + \hat{q})$ and $1 + \hat{q}^{-2} = \hat{q}^{-1}(\hat{q} + \hat{q}^{-1})$ then yields $d_k = [2k+1]_{\hat{q}}$.

be precise, Banica's result tells us that we have an isomorphism of monoidal W^*-categories, using $\delta^2 = N$ so $\delta = \sqrt{N}$: [Ban99, Prop. 2.2]

$$\pi : R_0(S_N^+) \to \mathrm{TL}^2(\delta) \quad : \quad \pi(u^{\otimes k}) = k \quad \text{and} \quad \pi(\mathrm{Hom}(u^{\otimes k}, u^{\otimes l})) = \mathrm{TL}_{2k,2l}(\delta). \tag{5.17}$$

And concretely, for the generators, we have the result [Bra12]:

$$\pi(\nu) = \delta^{-\frac{1}{2}} \cap \in \mathrm{TL}_{0,2}(\delta) \quad , \quad \pi(m) = \delta^{\frac{1}{2}} \; | \cup | \in \mathrm{TL}_{4,2}(\delta) \quad , \quad \pi(\mathrm{id}_B) = || \in \mathrm{TL}_{2,2}(\delta). \tag{5.18}$$

The fact that we have to pick the loop parameter $\delta = \sqrt{N}$ follows from the more general theory of *quantum automorphism groups* $\mathbb{G}_{\mathrm{aut}}(B, \psi)$, where S_N^+ corresponds to a specific choice of B and ψ. Although we will not define $\mathbb{G}_{\mathrm{aut}}(B, \psi)$ precisely (as we do not need its generality in this thesis), we give a small exposition that allows us to understand the relations imposed on the multiplication m, unit ν and identity id_B morphisms that generate the morphisms of $R_0(\mathbb{G}_{\mathrm{aut}}(B, \psi))$ as described in [Bra12] as we will need those relations in the following subsection to investigate the intertwiners of S_N^+ further.

The quantum permutation group S_N^+ can be considered a special case in the theory of *quantum automorphism groups* $\mathbb{G}_{\mathrm{aut}}(B, \psi)$ where B is a finite-dimensional C^*-algebra and $\psi : B \to \mathbb{C}$ a faithful state, originally investigated by Wang [Wan98]. The choice of the state ψ has a deep influence on the structure of $\mathbb{G}_{\mathrm{aut}}(B, \psi)$, which are generally not isomorphic for different choices of ψ. Banica called ψ a "canonical trace" if it was the restriction to B of the unique trace on $\mathcal{B}(B)$ where we embed B in $\mathcal{B}(B)$ via the left-regular representation [Ban99, Prop. 2.1], which is not the same choice of ψ that Wang originally made. In [Ban02], Banica further generalized this choice to any *δ-form* on B (to which we will return shortly), and proved that for any δ-form ψ on B, as long as $\dim B \geqslant 4$ (here, B is seen as a Hilbert space with inner product induced by ψ), the finite-dimensional representation theory of $\mathbb{G}_{\mathrm{aut}}(B, \psi)$ can be described through Temperley-Lieb algebras. Using this technology, he proved that the irreducible representations of these quantum automorphism groups can always be labeled by $\mathbb{N}_0$, and their fusion rules always follow the fusion rules of $\mathrm{SO}(3)$.

Let us return to the δ-forms that we left undefined: we call a faithful state ψ a *δ-form* for some $\delta > 0$ if the multiplication map $m : B \otimes B \to B$ satisfies $mm^* = \delta^2 \mathrm{id}_B$, where the inner products on B and $B \otimes B$ are induced by ψ and $\psi \otimes \psi$ [Bra12, Def. 3.5]. In this case, if we call $\nu : \mathbb{C} \to B$ the unit map, the relations $\nu^*\nu = 1$, $m(m \otimes \mathrm{id}_B) = m(\mathrm{id}_B \otimes m)$, $m(\mathrm{id}_B \otimes \nu) = m(\nu \otimes \mathrm{id}_B) = \mathrm{id}_B$ and $m^*m = (m \otimes \mathrm{id}_B)(\mathrm{id}_B \otimes m^*)$ also hold. The quantum permutation group S_N^+ corresponds in this framework to the "simplest situation", where we pick $B = C(X_N)$ with $X_N := \{1, \ldots, N\}$, and ψ the uniform probability measure on X_N. This means that ψ is a δ-form with $\delta^2 = \dim B = N$ [Bra12].

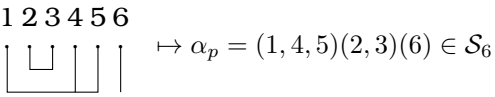
As Brannan notes, we can find an explicit model of the k-th irreducible representation v_k of S_N^+ through π, Banica's isomorphism of monoidal W^*-categories, by considering the Jones-Wenzl projection $p_{2k} \in \mathrm{TL}_{2k,2k}(\sqrt{N})$, and pulling it back through π to find the irreducible representation $v_k = \pi^{-1}(p_{2k})u^{\otimes k}$ and its representation space $\mathcal{H}_k = \pi^{-1}(p_{2k})B^{\otimes k}$ [Bra12]. Note that this approach is not yet explicit enough for our purposes, as we will need an explicit expression for the projection operator $\pi^{-1}(p_{2k}) \in \mathrm{Hom}(u^{\otimes k}, u^{\otimes k})$ that allows us to actually analyze and implement the associated Clebsch-Gordan quantum channels. We can contrast this with the O_N^+-case, where we identify $\pi^{-1}(p_k)$ with T_{p_k}, yielding an explicit operator in $\mathrm{Hom}(u^{\otimes k}, u^{\otimes k})$. In the S_N^+-case, however, this does not work (yet): the underlying partition category for S_N^+ should be $\mathrm{NC}_{\mathrm{lin}}[N]$ (in the sense that $\mathrm{Hom}(u^{\otimes k}, u^{\otimes k}) = \{T_p \; : \; p \in \mathrm{NC}_{\mathrm{lin}}[N](k,k)\}$), *not* $\mathrm{TL}^2(\sqrt{N})$ which contains morphisms classes containing non-crossing *pair* partitions. We therefore need to investigate how the categories $\mathrm{TL}^2(\sqrt{N})$ and $\mathrm{NC}_{\mathrm{lin}}[N]$ are related. The details of this relation lie in a so-called *fattening procedure*, which is a bijection between $\mathrm{NC}(k,l)$ and $\mathrm{NC}_2(2k, 2l)$.

5.2.1. THE FATTENING PROCEDURE

The category of non-crossing partitions, NC, is intimately related to the category of non-crossing *pair* partitions NC_2, as we will show that a bijection exists between $\mathrm{NC}(k)$ and $\mathrm{NC}_2(2k)$. This bijection is sometimes referred to as a *fattening procedure* in literature [CS11, BBC06]), presumably because it "doubles" the strands in the $\mathrm{NC}(k)$-diagram. We follow the exposition of [FN19], which uses [NS06] as a reference.

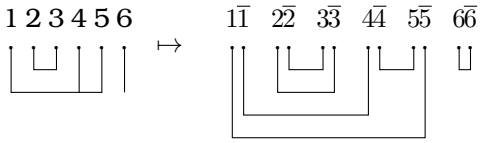
Note that a category of partitions is by definition invariant under the *rotation* of a partition (see Definition 2.3.6), and so, applying the rotation $p\curvearrowright$ precisely l times for a non-crossing partition $p \in \mathrm{NC}(k,l)$, we get a partition $q \in \mathrm{NC}(k+l,0)$. The *fattening procedure* as described in the referenced literature seems to always take place on partitions of the form $q \in \mathrm{NC}(k+l,0)$ (which is denoted in the referenced literature by "$\mathrm{NC}(n)$", leading to a notational clash with our short-hand $\mathrm{NC}(n) = \mathrm{NC}(n,n)$ in the partition framework as described by Freslon and Weber [FW16], thus caution is adviced.). In the general case of $\mathrm{NC}(k,l)$, where $l \neq 0$, we translate this "fattening procedure" by first rotating a partition to a partition in $\mathrm{NC}(k+l,0)$, then applying the procedure, and then rotating it back to $\mathrm{NC}_2(2k,2l)$. As the rotations are invertible operations, and both NC and NC_2 are categories of partitions and thus closed under the rotation operation, this yields a bijection between $\mathrm{NC}(k,l)$ and $\mathrm{NC}_2(2k,2l)$.

Let us consider a non-crossing partition $p \in \mathrm{NC}(n,0)$. We see that p can be naturally identified with a *permutation* $\alpha_p \in \mathcal{S}_n$, which are called *geodesic permutations* (or simply *non-crossing permutations*), see [NS06], as follows: say p contains blocks $V_1, \ldots, V_t$. For each block V_i, order its elements from small to large, put them in a tuple, and call this the *cycle* s_i. The permutation α_p is then $\alpha_p = s_1 s_2 \ldots s_t$. For example:



$$\mapsto \alpha_p = (1,4,5)(2,3)(6) \in \mathcal{S}_6$$

Now that we have found the geodesic permutation α_p associated with p, we "fatten" p to get a partition $\tilde{p} \in \mathrm{NC}_2(2k,0)$ by first doubling the index set from $\{1,\ldots,k\}$ to $\{1,\overline{1},\ldots,k,\overline{k}\}$ (so we "inject" the $\overline{i}$-indices between the original indices). Then, the pairs in $\tilde{p}$ are constructed by the rule: connect $\overline{i}$ with j if and only if $\alpha_p(i) = j$. In our example, this means that $\overline{1}$ gets connected to 4, $\overline{4}$ to 5, and $\overline{5}$ to 1 $\tilde{p}$. Visualized:



Inversely, consider a non-crossing pair partition $q \in \mathrm{NC}_2(2k,0)$. Note that if we label the indices $\{1,\overline{1},\ldots,k,\overline{k}\}$ again, any pair in q *must* be of the form $\{i,\overline{j}\}$ for some $1 \leqslant i,j \leqslant k$: if not, then an odd number of indices exists strictly between these two indices, which cannot be paired to obtain a valid non-crossing *pair* partition. Hence, starting with the pair $\{\overline{1},i_1\}$, we go to $\{\overline{i_1},i_2\}$, then to $\{\overline{i_2},i_3\}$, etcetera. This process terminates when we arrive back at index 1. We have then found the cycle $(1,i_1,i_2,\ldots,i_t)$ for some $t \in \mathbb{N}_0$. Find all such disjoint cycles in the indices of q, making sure to always start with the smallest index that is not yet a member of any other cycle: they make up a geodesic permutation α, and we can translate this back into a partition $q' \in \mathrm{NC}(k)$ by letting each of the cycles be a block of q', the inverse of the assignment $r \mapsto \alpha_r$ for $r \in \mathrm{NC}(2k,0)$.

We wish to determine the inverse image of the Temperley-Lieb generators $u_i \in \mathrm{NC}_2(2k,2k)$ under the fattening procedure such that we can understand the inverse image of the Jones-Wenzl projection under this fattening procedure. As an aid, we translate the fattening procedure to the general $\mathrm{NC}(k,l)$ (with $l \neq 0$) case.

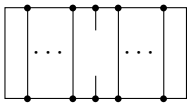
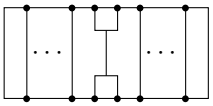
Remark 5.2.1

A direct algorithm for the fattening procedure for $\mathrm{NC}(k,l)$ (with $l \neq 0$) is as follows:

1. for a partition $p \in \mathrm{NC}(k,l)$, label the k upper indices by $\{1,\ldots,k\}$ and the l lower indices by $\{1',\ldots,l'\}$.
2. For each block $V_x = \{i_1,\ldots,i_t,j_1',\ldots,j_s'\}$, sort the unprimed indices from small to large, and concatenate to this the primed indices sorted from large to small (rotation operators invert the order of the lower indices once they become upper indices, hence we need to invert the sorting order), and call this tuple the cycle s_x.
3. From this, make the permutation $\alpha_p = s_1 \ldots s_u$, which is an automorphism on the set $[k] \sqcup [l']$, where $[k] = \{1,\ldots,k\}$ and $[l'] = \{1',\ldots,l'\}$.
4. "Fatten" the upper and lower indices to the sets $\{1,\overline{1},\ldots,k,\overline{k}\}$ and $\{\overline{1}',1'\ldots,\overline{l}',l'\}$. Note that the insertion of $\overline{i}'$ in the lower indices happens to the *left* of i', in contrast to the case for the upper indices.
5. Create the partition $q \in \mathrm{NC}_2(2k,2l)$ whose blocks are precisely those pairs $(\overline{x},y)$ with $x,y \in [k] \sqcup [l']$ for which $\alpha_p(x) = y$ (note that this includes cases in which precisely one of the indices x and y is primed, whilst the other is unprimed).

Lemma 5.2.1: Inverse image of Temperley-Lieb generators

Under the fattening procedure, the inverse image of the Temperley-Lieb generators $u_i \in \mathrm{NC}_2(2k,2k)$ for some $k \in \mathbb{N}$ and $1 \leqslant i \leqslant 2k-1$ is given by $a_{(i+1)/2}$ if i is odd, and $b_{i/2}$ if i is even, where $a_j, b_j \in \mathrm{NC}(k)$ are given by

$a_j =$  and $b_j =$  , (5.19)

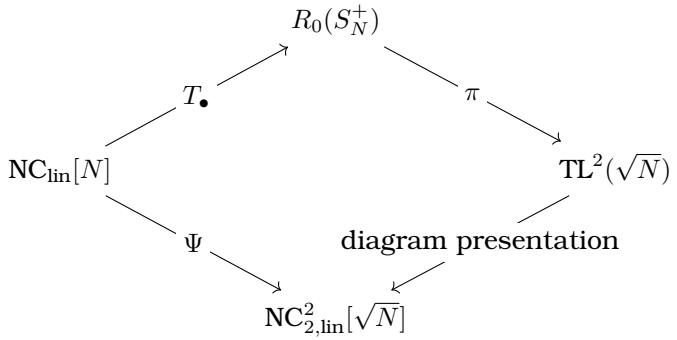
where the j'th upper and lower nodes in a_j are singletons (so a_j is defined for $1 \leqslant j \leqslant k$), and the j'th and $(j+1)$'th upper and lower nodes of b_j are connected in a block of size 4 (so b_j is defined for $1 \leqslant j \leqslant k-1$).

Proof. This follows directly from applying the fattening procedure to a_j, which results in u_{2j-1}, and applying the fattening procedure to b_j, which results in u_{2j}. □

We see that the fattening procedure that we have sketched here is precisely the connection between the underlying category of partitions NC for the intertwiner spaces and the two-cabled Temperley Lieb category $\mathrm{TL}^2(\sqrt{N})$. Let us make this precise. First, note that the functor $T_\bullet$ is faithful for $N \geqslant 4$ by Lemma 2.3.1, whence $T_\bullet$ is a monoidal $*$-isomorphism between $\mathrm{NC}_{\mathrm{lin}}[N]$ and $R_0(S_N^+)$ with objects $u^{\otimes k}$ with $k \in \mathbb{N}_0$, and morphism classes $\mathrm{Hom}(u^{\otimes k}, u^{\otimes l})$ for $k,l \in \mathbb{N}_0$. Together the Banica's monoidal $*$-isomorphism π, and the monoidal $*$-isomorphism between the two-cabled Temperley-Lieb category $\mathrm{TL}^2(\sqrt{N})$ and the linear category of partitions $\mathrm{NC}^2_{2,\mathrm{lin}}[\sqrt{N}] := \{\mathbb{N}_0, \{\mathrm{NC}_{2,\mathrm{lin}}[\sqrt{N}](2k,2l)\}_{k,l\in\mathbb{N}_0}\}$, we have:

Theorem 5.2.3: The fattening isomorphism

The fattening procedure $\psi : \mathrm{NC} \to \mathrm{NC}_2$ coincides with the restriction of the *fattening isomorphism* $\Psi : \mathrm{NC}_{\mathrm{lin}}[N] \to \mathrm{NC}^2_{2,\mathrm{lin}}[\sqrt{N}]$ to NC (so, we forget the linear structure on $\mathrm{NC}_{\mathrm{lin}}[N]$), where Ψ is the monoidal $*$-isomorphism defined by the following commutative diagram:



Here, $R_0(S_N^+)$ is the concrete monoidal W^*-category with objects $u^{\otimes k}$ for $k \in \mathbb{N}_0$ and with the intertwiner spaces $\mathrm{Hom}(u^{\otimes k}, u^{\otimes l})$ as morphisms.

Proof. We use $\delta := \sqrt{N}$. We claim that $\frac{1}{\delta} \uparrow \in \mathrm{NC}_{\mathrm{lin}}[N](0,1)$ is mapped to the unit $\nu \in \mathrm{Hom}(1,u)$ by $T_\bullet$. Clearly, $\mathrm{NC}_{\mathrm{lin}}[N](0,1) = \mathbb{C}\uparrow$, and $\nu^*\nu = 1$ forces $\nu = \frac{1}{\delta}T_\uparrow = T_{\delta^{-1}\uparrow}$. Similarly, the inverse image under $T_\bullet$ of the multiplication $m \in \mathrm{Hom}(u^{\otimes 2}, u)$ must be some linear combination of diagrams in $\mathrm{NC}_{\mathrm{lin}}[N](2,1)$. One can check that, if we scale the diagram $h =$  by δ, the operator $T_{\delta h}$ satisfies the relations from [Bra12, Remark 3.6], whence $m = T_{\delta h}$. Then, if we compose $\pi \circ T_\bullet$ and see the result as diagrams in $\mathrm{NC}^2_{2,\mathrm{lin}}[\sqrt{N}]$, we see that Ψ behaves as:

$$\delta^{-1}\uparrow \quad \mapsto \quad \nu = \delta^{-1}T_\uparrow \quad \mapsto \quad \delta^{-1/2}\cap \quad \in \mathrm{TL}_{0,2}(\delta) \tag{5.20}$$

$$\delta \, \cup \quad \mapsto \quad m = \delta T_{\cup} \quad \mapsto \quad \delta^{1/2}|\cup| \quad \in \mathrm{TL}_{4,2}(\delta) \tag{5.21}$$

$$| \quad \mapsto \quad \mathrm{id}_B = T_| \quad \mapsto \quad || \quad \in \mathrm{TL}_{2,2}(\delta) \tag{5.22}$$

But then, we see that Ψ coincides with the fattening procedure ψ on the generators, as long as we forget about the scalar factors. As Ψ is a monoidal $*$-isomorphism, this fully describes Ψ. From the definition of the fattening procedure, it readily follows that ψ fulfills $\psi(p \otimes q) = \psi(p) \otimes \psi(q)$, $\psi(p^*) = \psi(p)^*$ and $\psi(p \circ q) = \psi(p) \circ \psi(q)$ (note that we are referring to the *diagrammatic composition* here, without the scalar factors, because the domain and codomain of ψ do not have linear structure). Thus, the theorem follows.

□

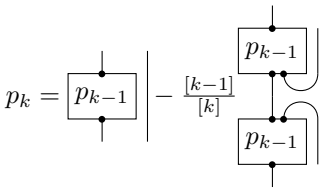
5.2.2. JONES-WENZL PROJECTIONS FOR S_N^+

Having established the fattening isomorphism and its relation to the intertwiner spaces $\mathrm{Hom}(u^{\otimes k}, u^{\otimes l})$, we have an alternative route to finding an expression for $\hat{p}_k := \pi^{-1}(p_{2k}) \in \mathrm{Hom}(u^{\otimes k}, u^{\otimes k})$: when we consider Theorem 5.2.3, we see that pulling p_{2k} back through π^{-1} is equivalent to $T_\bullet \circ \Psi^{-1}$, where Ψ is the fattening isomorphism $\mathrm{NC}_{\mathrm{lin}}[N] \to \mathrm{NC}^2_{2,\mathrm{lin}}[\sqrt{N}]$. As long as we are careful with the scalar factors, we understand the fattening procedure well enough to compute the inverse image of p_{2k}, and we understand $T_\bullet$ enough to yield a concrete description of an operator in $\mathrm{Hom}(u^{\otimes k}, u^{\otimes k})$. Note that the original Jones-Wenzl projections $p_{2k} \in \mathrm{TL}_{2k,2k}(\sqrt{N})$ do not have a direct formula, but are expressed in the Wenzl recursion relation (see Theorem 5.1.4), so a more modest goal is to find a recursion relation for $\hat{p}_k$. When we contemplate this strategy, we see hints to the fact that this calculation will be quite involved: we cannot use *one* recursion step by expressing p_{2k} in terms of p_{2k-1}, because p_{2k-1} does *not*

live in $\mathrm{TL}^2(\sqrt{N})$ as $2k-1$ is odd, which means that we cannot use the fattening isomorphism. We are therefore forced to take another recursion step, and express p_{2k} in terms of p_{2k-2}.

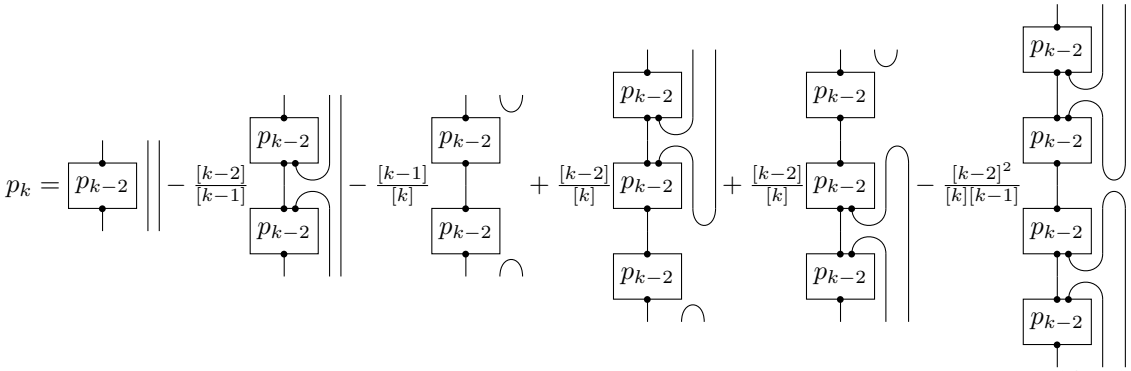
Lemma 5.2.2: Two-step recursion for Jones-Wenzl

The Wenzl-recursion formula reads:



(5.23)

Applying this recursion formula once more, for p_{k-1} in this expression, yields:



(5.24)

Proof. This can be directly shown by filling in the Wenzl recursion formula for p_{k-1} in terms of p_{k-2}. □

This lemma is the basis of:

Theorem 5.2.4: Wenzl recursion for S_N^+

Let Ψ be the fattening isomorphism from Theorem 5.2.3. Let $\hat{p}_{k-1} := \Psi^{-1}(p_{2k-2})$. Let $u_i^{(2k)}$ denote the i'th generator of $\mathrm{TL}_{2k}(\sqrt{N})$, and let $a_j^{(k)}, b_j^{(k)} \in \mathrm{NC}(k,k)$ be the inverse images of those generators as described in Lemma 5.2.1. We claim that we have

$$\begin{aligned}\hat{p}_k = \hat{p}_{k-1}\otimes| &- \delta\tfrac{[2k-2]}{[2k-1]}(\hat{p}_{k-1}\otimes|)b_{k-1}^{(k)}(\hat{p}_{k-1}\otimes|) - \delta^{-1}\tfrac{[2k-1]}{[2k]}(\hat{p}_{k-1}\otimes|)a_k^{(k)}\\ &+ \tfrac{[2k-2]}{[2k]}a_k^{(k)}(\hat{p}_{k-1}\otimes|)b_{k-1}^{(k)}(\hat{p}_{k-1}\otimes|) + \tfrac{[2k-2]}{[2k]}(\hat{p}_{k-1}\otimes|)b_{k-1}^{(k)}(\hat{p}_{k-1}\otimes|)a_k^{(k)}\\ &- \delta\tfrac{[2k-2]^2}{[2k][2k-1]}(\hat{p}_{k-1}\otimes|)b_{k-1}^{(k)}(\hat{p}_{k-1}\otimes|)a_k^{(k)}b_{k-1}^{(k)}(\hat{p}_{k-1}\otimes|).\end{aligned} \tag{5.25}$$

Proof. We can directly calculate the images of the $\mathrm{TL}_{2k}(\sqrt{N})$-generators under Ψ^{-1}:

$$u_{2k-2}^{(2k)} = |^{\otimes 2k-3}\otimes \cup\cap \otimes| \quad\mapsto\quad \delta b_{k-1}^{(k)} \quad=\quad \delta|^{\otimes k-2}\otimes \sqcap\!\sqcup \tag{5.26}$$

$$u_{2k-1}^{(2k)} = |^{\otimes 2k-2}\otimes \cup\cap \quad\mapsto\quad \delta^{-1}a_k^{(k)} \quad=\quad \delta^{-1}|^{\otimes k-1}\otimes | \tag{5.27}$$

Let us call the six diagrams A_k up to and including F_k in Lemma 5.2.2, i.e. $p_k = A_k + \cdots + F_k$.

We then see that the action of Ψ^{-1} on the diagrams A_{2k} up to F_{2k} is:

$$\begin{aligned}
A_{2k} &= p_{2k-2} \otimes |^{\otimes 2} && \mapsto && \hat{p}_{k-1} \otimes |.\\
B_{2k} &= -\tfrac{[2k-2]}{[2k-1]}(p_{2k-2} \otimes |^{\otimes 2}) u_{2k-2}^{(2k)} (p_{2k-2} \otimes |^{\otimes 2}) && \mapsto && -\delta \tfrac{[2k-2]}{[2k-1]} (\hat{p}_{k-1} \otimes |) b_{k-1}^{(k)} (\hat{p}_{k-1} \otimes |).\\
C_{2k} &= -\tfrac{[2k-1]}{[2k]}(p_{2k-2} \otimes |^{\otimes 2}) u_{2k-1}^{(2k)} && \mapsto && -\delta^{-1} \tfrac{[2k-1]}{[2k]} (\hat{p}_{k-1} \otimes |) a_k^{(k)}.\\
D_{2k} &= \tfrac{[2k-2]}{[2k]} u_{2k-1}^{(2k)} (p_{2k-2} \otimes |^{\otimes 2}) u_{2k-2}^{(2k)} (p_{2k-2} \otimes |^{\otimes 2}) && \mapsto && \tfrac{[2k-2]}{[2k]} a_k^{(k)} (\hat{p}_{k-1} \otimes |) b_{k-1}^{(k)} (\hat{p}_{k-1} \otimes |).\\
E_{2k} &= \tfrac{[2k-2]}{[2k]} (p_{2k-2} \otimes |^{\otimes 2}) u_{2k-2}^{(2k)} (p_{2k-2} \otimes |^{\otimes 2}) u_{2k-1}^{(2k)} && \mapsto && \tfrac{[2k-2]}{[2k]} (\hat{p}_{k-1} \otimes |) b_{k-1}^{(k)} (\hat{p}_{k-1} \otimes |) a_k^{(k)},
\end{aligned}$$

and lastly

$$\begin{aligned}
F_{2k} &= -\tfrac{[2k-2]^2}{[2k][2k-1]} (p_{2k-2} \otimes |^{\otimes 2}) u_{2k-2}^{(2k)} (p_{2k-2} \otimes |^{\otimes 2}) u_{2k-1}^{(2k)} u_{2k-2}^{(2k)} (p_{2k-2} \otimes |^{\otimes 2})\\
&\mapsto -\delta \tfrac{[2k-2]^2}{[2k][2k-1]} (\hat{p}_{k-1} \otimes |) b_{k-1}^{(k)} (\hat{p}_{k-1} \otimes |) a_k^{(k)} b_{k-1}^{(k)} (\hat{p}_{k-1} \otimes |). \qquad (5.28)
\end{aligned}$$

This can be shown by a direct calculation, using the fact that Ψ^{-1} respects compositions and tensor products. Note that we have simplified the result somewhat by using $p_{2k-2}^2 = p_{2k-2}$, such as in the definition of C_{2k}. Please also note that $u_{2k-1}^{(2k)}$ and $p_{2k-2} \otimes |^{\otimes 2}$ commute. □

6

Clebsch-Gordan Quantum Channels

When the quantum analogs of classical capacity theories were discovered, such as the HSW-theorem (Theorem 3.3.2), *regularized* information-theoretic quantities such as the regularized Holevo information appeared, which are in general intractable. Hastings' counterexample to the additivity conjecture showed us that there exist quantum channels whose information-theoretic quantities are truly different than the regularized ones. Thus the scientific endeavor appeared to classify which classes of quantum channels are (strongly) additive, and which are not – some of those results have been recorded in Section 3.5. As we can see, if we assume our quantum channel has a *nice* property, such as entanglement-breaking, we can try to leverage that property to figure out whether our channel is (strongly) additive, or to calculate other information-theoretic properties.

Another fruitful idea is to consider channels that have an "underlying symmetry", i.e. channels that are somehow closely related to a mathematical object that we understand well. A natural candidate is to take a mathematical group whose representation theory is well-understood, and use the Hilbert space in which the linear operators of the group representation live to define a quantum channel. This method dates at least back to 2005, where Holevo proved the equivalence of MOE-additivity and χ-additivity for what he calls *irreducibly covariant quantum channels* [Hol05]. For a group G, two representations $\pi_{\mathcal{H}}, \pi_{\mathcal{K}}$ on Hilbert spaces $\mathcal{H}$,$\mathcal{K}$, a quantum channel $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ is called G-covariant if $\Phi(\pi_{\mathcal{H}}(g)\rho\pi_{\mathcal{H}}(g)^*) = \pi_{\mathcal{K}}(g)\Phi(\rho)\pi_{\mathcal{K}}(g)^*$ for all $\rho \in \mathcal{B}(\mathcal{H})$ and for all $g \in G$. The channel Φ is called *irreducibly covariant* if both representations are irreducible.

For specific groups, these quantum channels have been investigated: Al Nuwairan studied the convex set of $\mathrm{SU}(2)$-irreducibly covariant channels, in particular their extreme points, the so-called *EPOSIC* (extreme points of $\mathrm{SU}(2)$-irreducibly covariant) channels [Nuw14, Nuw13]. The starting point of this thesis was the paper by Brannan and Collins, in which they introduce quantum channels that are $\mathbb{G}$-covariant, where $\mathbb{G}$ is the compact *quantum* group O_N^+.

6.1. Channels with compact quantum group symmetry

The theory of compact quantum groups yields an interesting collection of quantum channels: given a compact quantum group $\mathbb{G}$, and three unitary irreducible corepresentations $(u, \mathcal{H}_u)$, $(v, \mathcal{H}_v)$ and $(w, \mathcal{H}_w)$ such that w is a subrepresentation contained in the tensor product $u \otimes v$ on $\mathcal{H}_u \otimes \mathcal{H}_v$. Then, we know an intertwining isometry $\alpha_u^{v,w}$ exists that embeds $\mathcal{H}_w$ in $\mathcal{H}_u \otimes \mathcal{H}_v$. By the Stinespring dilation theorem, we can then naturally consider the quantum channels given by:

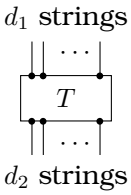
$$\Phi_u^{(v),w} : \mathcal{B}(\mathcal{H}_u) \to \mathcal{B}(\mathcal{H}_w) \ : \ \rho \mapsto \mathrm{Tr}_{\mathcal{H}_v}\left(\alpha_u^{v,w}\rho(\alpha_u^{v,w})^*\right); \tag{6.1}$$

$$\Phi_u^{v,(w)} : \mathcal{B}(\mathcal{H}_u) \to \mathcal{B}(\mathcal{H}_v) \ : \ \rho \mapsto \mathrm{Tr}_{\mathcal{H}_w}\left(\alpha_u^{v,w}\rho(\alpha_u^{v,w})^*\right). \tag{6.2}$$

These channels are called *Clebsch-Gordan quantum channels* in [BCLY20], as the Clebsch-Gordan coefficients are encoded in the isometries $\alpha_u^{v,w}$ in the case where one picks SU(2) as the compact quantum group.

6.2. DIAGRAMMATIC CALCULUS

We wish to broaden our use of string diagrams to depict partitions to include all linear operators, such that we could for example depict $\rho \mapsto \alpha_u^{v,w}\rho(\alpha_u^{v,w})^*$ diagrammatically if the isometry $\alpha_u^{v,w}$ is (a linear combination of) a partition diagram. This technique is called *string diagram calculus* in [BCLY20]. A linear transformation $T \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$, where $\dim \mathcal{H}_i = d_i$, is depicted graphically as



Notice that we choose the convention that the diagram is read top-to-bottom, which breaks with the convention of [BCLY20]! We wish to display the composition T_2T_1 of two operators $T_1 \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$ and $T_2 \in \mathcal{B}(\mathcal{H}_2, \mathcal{H}_3)$ by stacking the diagram of T_1 on top of T_2 and connecting the lower nodes of T_1 with the upper nodes of T_2, to be consistent with the diagram composition introduced in the partition theory section (Section 2.3), and especially with the introduction of the partition maps T_p in Definition 2.3.9. For example, the partition $\cup \in \mathrm{NC}_2(2,0)$ yields a partition map $T_\cup : \mathcal{H}^{\otimes 2} \to \mathbb{C}$ (where $\mathcal{H} = \mathbb{C}^N$) that maps $|i_1 i_2\rangle \to \delta_{i_1,i_2}$, so if we look at the diagram $\boxed{T_\cup}$, the flow of the diagram must be from top to bottom. Again, the diagram of the tensor product of $T_1 \otimes T_2$ is drawn by horizontal concatenation. We will use this diagrammatic string calculus for linear transformations extensively in the following sections.

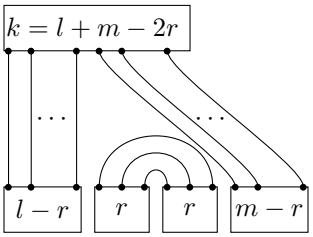
6.3. QUANTUM CHANNELS ASSOCIATED TO O_N^+

In the case of the compact quantum group O_N^+, we know that its irreducible corepresentations v_k can be labelled by $\mathbb{N}_0$, where $v_0 := 1$, $v_1 := u$ (the fundamental representation), and the fusion rules $v_k \otimes v_l \simeq v_{|k-l|} \oplus v_{|k-l|+2} \cdots \oplus v_{k+l}$, see Theorem 5.1.2. We know that we can identify v_k by $p_k u^{\otimes k}$, where p_k is the k-th Jones-Wenzl projection. We label the corepresentation Hilbert spaces by $\mathcal{H}_k$, with $\dim \mathcal{H}_k = [k+1]_q$, see Section 5.1. The main results in this section are due to [BC16b, BC20, BCLY20], but we prove a small uniqueness result that we could not directly find in literature, using the partition theory tools from [FW16].

In order to explicitly describe the Clebsch-Gordan quantum channels associated with O_N^+ (as described in Section 6.1), we will need to find explicit descriptions of the intertwining isometries $\alpha_k^{l,m}$ that embed $\mathcal{H}_k$ in $\mathcal{H}_l \otimes \mathcal{H}_m$. By studying the fusion rules, we see that $v_k \subset v_l \otimes v_m$ iff $k = l + m - 2r$ for some $0 \leqslant r \leqslant \min\{l, m\}$, and in this case we call the triple $(k, l, m) \in \mathbb{N}_0^3$ *admissible*. In Section 5.1, we developed theory that showed a bijection between the intertwiner space $\mathrm{Hom}_{O_N^+}(k, l)$ and the non-crossing pair partitions $\mathrm{NC}_2(k, l)$. By Lemma 4.3.1, we know that the intertwiner space $\mathrm{Hom}(v_k, v_l \otimes v_m)$ is equal to $(p_l \otimes p_m)\mathrm{Hom}(u^{\otimes k}, u^{\otimes l} \otimes u^{\otimes m})p_k$ where $p_k : \mathcal{H}_1^{\otimes k} \twoheadrightarrow \mathcal{H}_k$ is the canonical projection such that $v_k = p_k u^{\otimes k}$, and similarly for p_l and p_m. From the fusion rules, we see that the multiplicity of v_k in $v_l \otimes v_m$ is 1, so $\dim \mathrm{Hom}(v_k, v_l \otimes v_m) = 1$. Hence, we simply need to find any intertwiner T in $\mathrm{Hom}(u^{\otimes k}, u^{\otimes(l+m)}) = \mathrm{span}\{T_p : p \in \mathrm{NC}_2(k, l+m)\}$ such that $(p_l \otimes p_m)Tp_k \neq 0$. As the linear assignment $T_p \mapsto (p_l \otimes p_m)T_p p_k$ maps onto a one-dimensional space, we know there must exist a $p \in \mathrm{NC}_2(k, l+m)$ such that $(p_l \otimes p_m)T_p p_k \neq 0$, i.e. we do not need to consider linear combinations of T_p. We have the following characterisation:

Theorem 6.3.1

Given an admissible triple $(k,l,m) \in \mathbb{N}_0^3$, let $k = l+m-2r$ with $0 \leqslant r \leqslant \min\{l,m\}$. Precisely *one* T_p exists for $p \in \mathrm{NC}_2(k,l+m)$ such that $(p_l \otimes p_m)T_p p_k \neq 0$, and this p is precisely $|^{\otimes(l-r)} \otimes \cap^r \otimes |^{\otimes(m-r)}$. Diagrammatically:



Proof. As $p \in \mathrm{NC}_2(k,l+m)$, if the partition contains a pair $X = \{i,j\}$ of upper nodes, it *must* contain a pair $Y = \{i,i+1\}$ of adjacent nodes (and the other pairs of non-adjacent upper nodes are "nested caps" around such a pair of adjacent nodes). This means that $p = \frac{1}{N} p u_i^{(k)}$ where $u_i^{(k)} \in \mathrm{TL}_k(N)$ is the i'th generator, but the Jones-Wenzl projection p_k kills u_i, so $T_p p_k = 0$.

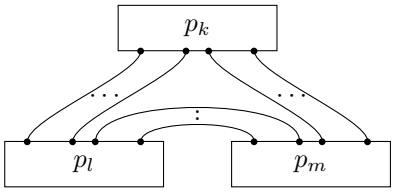
Hence, p contains k through-blocks (connecting upper and lower nodes). Similarly, if p contains a pair of two of the first l lower nodes, it must contain a pair $Z = \{i, i+1\}$, so $p = \frac{1}{N}(u_i^{(l)} \otimes |^{\otimes m})p$, but $p_l \otimes p_k$ kills $u_i^{(l)} \otimes |^{\otimes m}$, so $(p_l \otimes p_m)T_p = 0$. A similar argument shows that p cannot contains pairs of two of the last m lower nodes.

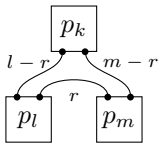
But then, as p is a pair partition, k of the $l+m$ lower nodes are connected to the upper k nodes, leaving $(l+m-k)/2 = r$ pairs of lower nodes that need to be connected. By the previous argument, none of the pairs can be exclusively in the first l or the last m nodes, so we must connect pairs of elements of the form $\{i,j'\}$. By the fact that p is non-crossing, we again find that these pairs must form "nested cups" around the pair $\{l,1'\}$. This fully describes p and shows that it is precisely of the form as stated in the theorem. □

We denote the unique intertwiner that we found in Theorem 6.3.1 by $A_k^{l,m} \in \mathrm{Hom}(v_k, v_l \otimes v_m)$. We see that the unique partition map T_p with $p \in \mathrm{NC}_2(k,l+m)$ such that $(p_l \otimes p_m)T_p p_k$ is given by $\iota^{\otimes l-r} \otimes T_{\cap^r} \otimes \iota^{\otimes m-r}$, where ι is the identity operator on $\mathcal{H}_1$ and where the intertwiner $T_{\cap^r} \in \mathrm{Hom}(1, u^{\otimes 2r})$ is given by the equation

$$T_{\cap^r}(1) = \sum_{i:[r]\to[N]} |i_1 \cdots i_{r-1} i_r\rangle \, |i_r i_{r-1} \cdots i_1\rangle \,, \tag{6.3}$$

where $[r] := \{1,\ldots,r\}$, similarly for $[N]$, and the summation is over all functions i from the finite set $[r]$ to $[N]$, and $\{|j\rangle\}_{j=1}^N$ is the canonical orthonormal basis of $\mathcal{H}_1 = \mathbb{C}^N$. Graphically, this is displayed as

$A_k^{l,m} =$ 

displayed as 

(6.4)

The diagram on the right-hand side has been simplified (repeating strings between objects have been removed), and the labels next to the string indicate how many repeated strands are encoded in one string. In some contexts, the labels might be superfluous (and are left out), as long as only one possible configuration is possible.

This object $A_k^{l,m}$ is called a *three-vertex* in the Temperley-Lieb recoupling theory as described in [KL94]. We can normalize $A_k^{l,m}$ to an isometry $\alpha_k^{l,m}$ given by [BCLY20]

$$\alpha_k^{l,m} = \left(\frac{\tau_k(\iota_k)}{\tau_k\big((A_k^{l,m})^* A_k^{l,m}\big)}\right)^{1/2} A_k^{l,m} = \left(\frac{[k+1]_q}{\theta_q(k,l,m)}\right)^{1/2} A_k^{l,m}, \tag{6.5}$$

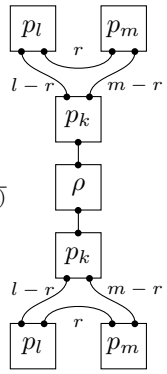
where the *theta-net* $\theta_q(k,l,m)$ is given by

$$\theta_q(k,l,m) := \tau_k\left((A_k^{l,m})^* A_k^{l,m}\right) = \frac{[r]_q![l-r]_q![m-r]_q![k+r+1]_q!}{[l]_q![m]_q![k]_q!}, \tag{6.6}$$

where the *quantum factorial* is defined as $[x]_q! := [x]_q[x-1]_q\cdots[2]_q[1]_q$ for $x \in \mathbb{N}$.

This leads us to diagrammatically display the action $\rho \mapsto \alpha_k^{l,m}\rho(\alpha_k^{l,m})^*$ as:

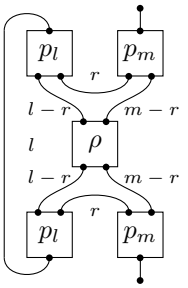
$$\alpha_k^{l,m}\rho(\alpha_k^{l,m})^* = \frac{[k+1]_q}{\theta_q(k,l,m)}$$



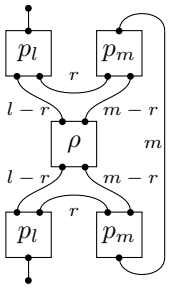
(6.7)

This analysis allows us to build the Clebsch-Gordan quantum channels as discussed in Section 6.1: the input space is $\mathcal{B}(\mathcal{H}_k)$, and the output space is either $\mathcal{B}(\mathcal{H}_l)$ or $\mathcal{B}(\mathcal{H}_m)$, depending on whether we trace out the first or the second subsystem, respectively. As $\mathcal{H}_k = p_k\mathcal{H}_1^{\otimes k}$, we see that $p_k\rho p_k = \rho$ for $\rho \in \mathcal{H}_k$, so we can simplify the diagram above slightly. Taking the partial trace over $\mathcal{H}_l$ or $\mathcal{H}_m$ corresponds to connecting the left-most p_l blocks with each other, or connecting the right-most p_m blocks with each other, respectively:

$$\Phi_k^{(l),m}(\rho) = \frac{[k+1]_q}{\theta_q(k,l,m)} \cdot$$



$$, \quad \Phi_k^{l,(m)}(\rho) = \frac{[k+1]_q}{\theta_q(k,l,m)} \cdot$$



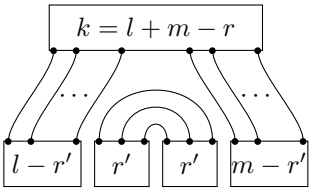
(6.8)

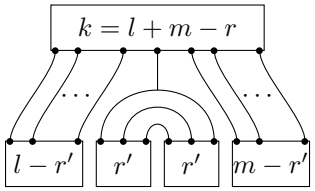
6.4. QUANTUM CHANNELS ASSOCIATED TO S_N^+

In the case of the compact quantum group S_N^+, we know its irreducible corepresentations v_k can be labelled by $\mathbb{N}_0$, with $v_0 := 1$, and $u = v_0 \oplus v_1$ where u is the fundamental representation, and we have the fusion rules $v_k \otimes v_l \simeq v_{|k-l|} \oplus v_{|k-l|+1} \oplus \cdots \oplus v_{k+l}$, see Theorem 5.2.2. We again label the Hilbert spaces of the corepresentations v_k by $\mathcal{H}_k$. To study the Clebsch-Gordan channels as described in Section 6.1 for S_N^+, we look at the intertwining isometries $\alpha_k^{l,m}$ embedding $\mathcal{H}_k$ in $\mathcal{H}_l \otimes \mathcal{H}_m$. With the fusion rules in mind, we call a tuple $(k,l,m) \in \mathbb{N}_0^3$ *admissible* if $k = l+m-r$ for some $0 \leqslant r \leqslant 2\min\{l,m\}$. The main results of this section can be found in [Bra12], but we adapt the results to match the framework in which we described the O_N^+-channels. Analogous to the O_N^+ case covered in Theorem 6.3.1, we can state a uniqueness result:

Theorem 6.4.1

Given an admissible triple $(k,l,m) \in \mathbb{N}_0^3$, let $k = l+m-r$ with $0 \leqslant r \leqslant 2\min\{l,m\}$. Precisely one T_p exists for $p \in \mathrm{NC}(k, l+m)$ such that $(\hat{p}_l \otimes \hat{p}_m)T_p\hat{p}_k \neq 0$. Furthermore, if we let $r' := r/2$ if r is even, and $r' := (r+1)/2$ if r is odd, p is precisely:

$p =$ 

if r is even, else $p =$ 

(6.9)

Proof. Firstly, note that p cannot contain any singleton with an upper index, say $\{i_1\}$: in this case, by Lemma 5.2.1, we know that p is equal to (a scalar multiple of) pa_{i_1}, but $a_{i_1}\hat{p}_k = 0$. Secondly, note that p cannot contain any block X that contains two *adjacent* upper indices, say $\{i_2, i_2 + 1\} \subseteq X$. In this case, using the same Lemma 5.2.1, p is equal to pb_{i_2} up to a non-zero scalar, but $b_{i_2}\hat{p}_k = 0$ as well.

Then, note that the only way for a non-crossing p to neither have an upper index singleton nor any block with two adjacent upper indices, is for all k upper indices to be in separate through-blocks. An analogous argument reveals that the leftmost l lower indices must all be in different blocks, and the rightmost m lower indices as well.

The r additional indices we have at the bottom must be connected in caps containing precisely one lower index i'_1 with $i'_1 \leqslant l$ and one lower index i'_2 with $i'_2 > l$, or in $\{i'_1, i'_2, j\}$ (of the latter, only one is possible for a non-crossing partition!). They must be nested caps and they must be symmetrically distributed around the center of the lower axis, as the partition is non-crossing. All other lower indices must then be in a paired with precisely one upper index. Hence, only one partition is possible, which is precisely the partition p that we drew. □

Let us call the resulting intertwiner $\hat{A}_k^{l,m} := (\hat{p}_l \otimes \hat{p}_m)T_p\hat{p}_k \in \mathrm{Hom}(v^k, v^l \otimes v^m)$. Note that as an alternative proof to the previous theorem, one can apply the inverse of the fattening isomorphism as described in Section 5.2.1 to the partition found in the O_N^+ case in Theorem 6.3.1, which is a valid procedure as the Jones-Wenzl projections $\hat{p}_t$ (for $t \in \{k, l, m\}$) are also inverses under the fattening isomorphism of the usual Jones-Wenzl projections $p_t \in \mathrm{TL}_t(\delta)$. This approach also reveals the proper normalization constant to turn $\hat{A}_k^{l,m}$ into an isometry. Let us carefully examine the normalization constant.

We have previously seen in Theorem 5.2.3 that under the fattening isomorphism, $a := \uparrow \in \mathrm{NC}_{\mathrm{lin}}[N](0,1)$ is mapped to $\Psi(a) := \delta^{1/2}\cap \in \mathrm{TL}_{0,2}(\delta)$, $b :=$  is mapped to $\Psi(b) := \delta^{-1/2}|\cup| \in \mathrm{TL}_{4,2}(\delta)$, and $| \in \mathrm{NC}_{\mathrm{lin}}[N](1,1)$ is mapped to $|| \in \mathrm{TL}_{2,2}(\delta)$, where $\delta = \sqrt{N}$.

In the case where r is even, we claim that the fattening isomorphism takes the intertwining $p = |^{\otimes l-r'} \otimes \cap^{r'} \otimes |^{\otimes m-r'}$ to $\Psi(p) = |^{\otimes 2l-r} \otimes \cap^{r} \otimes |^{\otimes 2m-r}$, where $r' := r/2$. Namely, a cap $\cap \in \mathrm{NC}_{\mathrm{lin}}[N]$ can be realized as b^*a, and is thus mapped to $\cap^2 \in \mathrm{TL}_{0,4}(\delta)$. Similarly, $\cap^2 \in \mathrm{NC}_{\mathrm{lin}}[N]$ can be realized as $(|\otimes b^* \otimes |)(|\otimes a \otimes |)b^*a$ and is thus mapped to $\cap^4 \in \mathrm{TL}_{0,8}(\delta)$. In general, $\cap^t$ is mapped to $\cap^{2t}$ for $t \in \mathbb{N}$. In the case where r is odd, we define $r' := (r+1)/2$, and then a similar check shows that $\Psi(p) = \delta^{-1/2} \cdot |^{\otimes 2l-r} \otimes \cap^{r+1} \otimes |^{\otimes 2m-r}$ (note that, in this case, building p in terms of a, b^* and $|$ always includes one more b^* than a, which causes the scalar factor in $\Psi(p)$).

We conclude that $\Psi(\hat{A}_k^{l,m}) = A_{2k}^{2l,2m}$ in the case where r is even, and $\Psi(\hat{A}_k^{l,m}) = \delta^{-1/2}A_{2k}^{2l,2m}$ in the case where r is odd. Thus, we have

$$\hat{\alpha}_k^{l,m} = \left(\tfrac{[2k+1]_q}{\theta_q(2k,2l,2m)}\right)^{1/2} \hat{A}_k^{l,m} \text{ if } r \text{ is even, else } \hat{\alpha}_k^{l,m} = \left(\tfrac{\delta[2k+1]_q}{\theta_q(2k,2l,2m)}\right)^{1/2} \hat{A}_k^{l,m} \tag{6.10}$$

Akin to the O_N^+-channels, we therefore find a set of Clebsch-Gordan quantum channels for S_N^+ labelled by $\hat{\Phi}_k^{(l),m}$ and $\hat{\Phi}_k^{l,(m)}$, given by:

$$\hat{\Phi}_k^{(l),m} : \mathcal{B}(\mathcal{H}_k) \to \mathcal{B}(\mathcal{H}_m) : \rho \mapsto \mathrm{Tr}_{\mathcal{H}_l}\left(\hat{\alpha}_k^{l,m}\rho(\hat{\alpha}_k^{l,m})^*\right) \tag{6.11}$$

$$\hat{\Phi}_k^{l,(m)} : \mathcal{B}(\mathcal{H}_k) \to \mathcal{B}(\mathcal{H}_l) : \rho \mapsto \mathrm{Tr}_{\mathcal{H}_m}\left(\hat{\alpha}_k^{l,m}\rho(\hat{\alpha}_k^{l,m})^*\right). \tag{6.12}$$

6.5. MINIMUM OUTPUT ENTROPY FOR O_N^+ AND S_N^+ CHANNELS

Proposition 6.5.1: Highly entangled intertwiner for O_N^+

[BC16b, Prop. 3.1] For $N \geqslant 3$, fix an O_N^+-admissible triple $(k,l,m) \in \mathbb{N}_0^3$. For any unit vectors $\xi \in \mathcal{H}_k, \eta \in \mathcal{H}_l, \zeta \in \mathcal{H}_m$, we have:

$$|\langle \alpha_k^{l,m}(\xi) \mid \eta \otimes \zeta \rangle| \leqslant \left(\frac{[k+1]_q}{\theta_q(k,l,m)}\right)^{1/2}. \tag{6.13}$$

Proof. Note that $\mathcal{H}_t := p_t \mathcal{H}_1^{\otimes t}$ for $t \in \mathbb{N}$, and by the absorption property of the Jones-Wenzl projections we have the useful identity:

$$\mathcal{H}_m = p_m \mathcal{H}_1^{\otimes m} = p_m(\iota_r \otimes p_{m-r})\mathcal{H}_1^{\otimes m} = p_m(\mathcal{H}_1^{\otimes r} \otimes \mathcal{H}_{m-r}). \tag{6.14}$$

Thus, we can see $\mathcal{H}_m$ as a subspace of $\mathcal{H}_1^{\otimes r} \otimes \mathcal{H}_{m-r}$, and similarly $\mathcal{H}_l$ as a subspace of $\mathcal{H}_{l-r} \otimes \mathcal{H}^{\otimes r}$, which allows us to express $\eta = \sum_{i:[r]\to[N]} \eta_i \otimes e_i$ and $\zeta = \sum_{i:[r]\to[N]} e_i \otimes \zeta_i$ for vectors $\eta_i \in \mathcal{H}_{l-r}$, $\zeta_i \in \mathcal{H}_{m-r}$, and where we let $\{e_i := |i_1 \dots i_r\rangle\}_{i:[r]\to[N]}$ be the canonical orthonormal basis for $\mathcal{H}_1^{\otimes r}$. If we apply the adjoint of $A_k^{l,m}$ to $\eta \otimes \zeta$, we can use the fact that the adjoint of the intertwiner $T_{\cap r} \in \mathrm{Hom}(1, u^{\otimes 2r})$ is given by

$$T_{\cap r}^* = T_{\cup r} = \sum_{i:[r]\to[N]} \langle i_1 \cdots i_r | \, \langle i_r \cdots i_1 | \,. \tag{6.15}$$

Thus, we see:

$$\begin{aligned}
(A_k^{l,m})^*(\eta \otimes \zeta) &= p_k(\iota_{l-r} \otimes T_{\cup r} \otimes \iota_{m-r})(p_l \otimes p_m)(\eta \otimes \zeta) && (6.16)\\
&= p_k(\iota_{l-r} \otimes T_{\cup r} \otimes \iota_{m-r}) \sum_{i,j:[r]\to[N]} (\eta_i \otimes e_i \otimes e_j \otimes \zeta_j) && (6.17)\\
&= p_k \sum_{i,j:[r]\to[N]} T_{\cup r}(e_i \otimes e_j)\eta_i \otimes \zeta_j = p_k \sum_{i:[r]\to[N]} \eta_i \otimes \zeta_{\check{i}}, && (6.18)
\end{aligned}$$

where $\check{i} : [r] \to [N]$ has $\check{i}_t := i_{N-t+1}$. Using the fact that $\| \sum_i a_i \otimes b_i \|^2 \leqslant (\sum_i \|a_i\|^2)(\sum_i \|b_i\|^2)$ for arbitrary vectors a_i and b_i [1], we see

$$|\langle \alpha_k^{l,m}\xi, \eta \otimes \zeta \rangle| \leqslant \|\xi\| \cdot \|(\alpha_k^{l,m})^*(\eta \otimes \zeta)\| \leqslant \left(\frac{[k+1]_q}{\theta_q(k,l,m)}\right)^{1/2} (\sum_i \|\eta_i\|^2)^{1/2} (\sum_i \|\zeta_{\check{i}}\|^2)^{1/2}. \tag{6.19}$$

Finally, we use the fact that $\sum_i \|\eta_i\|^2 = \|\eta\|^2 = 1$ and similarly for ζ. □

Proposition 6.5.2: Highly entangled intertwiner for S_N^+

For $N \geqslant 4$, fix an S_N^+-admissible triple $(k,l,m) \in \mathbb{N}_0^3$. For any unit vectors $\xi \in \mathcal{H}_k$, $\eta \in \mathcal{H}_l$ and $\zeta \in \mathcal{H}_m$, we have:

$$|\langle \hat{\alpha}_k^{l,m}(\xi) | \eta \otimes \zeta \rangle| \leqslant \begin{cases} \left(\frac{[2k+1]_q}{\theta_q(2k,2l,2m)}\right)^{1/2} & r := l+m-k \text{ is even} \\ \left(\frac{[2k+1]_q}{\theta_q(2k,2l,2m)}\right)^{1/2} \cdot N^{1/4} & r := l+m-k \text{ is odd} \end{cases}. \tag{6.20}$$

Proof. In the case $r := l+m-k$ is even, this follows directly from Proposition 6.5.1, where we can use the same proof, *mutatis mutandis*. In the case r is odd, however, we need to be careful: the intertwiner $T_p \in \mathrm{Hom}(u^{\otimes k}, u^{\otimes l+m})$ is now given by $T_p = \iota_{l-r'} \otimes T_q \otimes \iota_{m-r'}$, where

[1] Expand $\| \sum_i a_i \otimes b_i \|^2 = \sum_{ij} \langle a_i \otimes b_i, a_j \otimes b_j \rangle = \sum_{ij} \langle a_i, a_j \rangle \langle b_i, b_j \rangle$, let $\gamma_{ij} := \langle a_i, a_j \rangle$ and $\delta_{ij} := \langle b_i, b_j \rangle$ be components of vectors γ and δ, apply Cauchy-Schwartz to conclude $\langle \gamma, \delta \rangle \leqslant \|\gamma\| \cdot \|\delta\|$, again use Cauchy-Schwartz on $\|\gamma\|$ to conclude $\|\gamma\|^2 = \sum_{ij} |\langle a_i, a_j \rangle|^2 \leqslant \sum_{ij} \|a_i\|^2 \|a_j\|^2 = (\sum_i \|a_i\|^2)^2$ and similarly for $\|\delta\|^2$. Combining these results yields the claim.

$r' = (r-1)/2$ and q is the diagram in $\mathrm{NC}(1, 2r')$ given by: $q :=$  , where the double lines indicate $r'-1$ nested caps, see Theorem 6.4.1. We repeat the first part of Proposition 6.5.1: write $\mathcal{H}_l = \hat{p}_l(\mathcal{H}_{l-r'} \otimes \mathcal{H}^{\otimes r'})$ and $\mathcal{H}_m = \hat{p}_m(\mathcal{H}^{\otimes r'} \otimes \mathcal{H}_{m-r'})$ (note that we use $\mathcal{H} := \mathbb{C}^N$ here, whilst $\mathcal{H}_1 := \hat{p}_1\mathcal{H} \simeq \mathbb{C}^{N-1}$, in contrast to the O_N^+-case where p_1 is the identity!). Then, express $\eta = \sum_{i:[r']\to[N]} \eta_i \otimes e_i$ and $\zeta = \sum_{i:[r']\to[N]} e_i \otimes \zeta_i$, where $\eta_i \in \mathcal{H}_{l-r'}$, $\zeta_i \in \mathcal{H}_{m-r'}$, and where we let $\{e_i := |i_1 \ldots i_{r'}\rangle\}_{i:[r']\to[N]}$ be the canonical orthonormal basis for $\mathcal{H}^{\otimes r'}$. We again see

$$(\hat{A}_k^{l,m})^*(\eta \otimes \zeta) = \hat{p}_k \sum_{i,j:[r']\to[N]} \eta_i \otimes T_q^*(e_i \otimes e_j) \otimes \zeta_j, \tag{6.21}$$

but this time we have $T_q^*(e_i \otimes e_j) = \delta_{\check{i},j} |i_1\rangle$, where $\delta_{\check{i},j}$ is zero unless j is equal to $\check{i}$. Thus,

$$(\hat{A}_k^{l,m})^*(\eta \otimes \zeta) = \hat{p}_k \sum_{i:[r']\to[N]} \eta_i \otimes |i_1\rangle \otimes \zeta_{\check{i}}. \tag{6.22}$$

We still use the inequality $\| \sum_i a_i \otimes b_i \|^2 \leqslant (\sum_i \|a_i\|^2)(\sum_i \|b_i\|^2)$, but let $a_i := \eta_i \otimes |i_1\rangle$ in this case, and then note that

$$\sum_i \|a_i\|^2 = \sum_i \|\eta_i \otimes |i_1\rangle\|^2 = \sum_i \|\eta_i\|^2 \cdot \| |i_1\rangle \|^2 = \sum_i \|\eta_i\|^2, \tag{6.23}$$

so the argument still holds. Using the correct normalization to pass from $\hat{A}_k^{l,m}$ to $\hat{\alpha}_k^{l,m}$ yields the proposition. □

Theorem 6.5.1: Minimum output entropy for O_N^+

[BC16b, Cor. 4.2] For $N \geqslant 3$, fix an O_N^+-admissible triple $(k, l, m) \in \mathbb{N}_0^3$. We have:

$$H_{\min}(\Phi_k^{(l),m}) = H_{\min}(\Phi_k^{l,(m)}) \geqslant \log \tfrac{\theta_q(k,l,m)}{[k+1]_q}. \tag{6.24}$$

Proof. We know that the value $H_{\min}(\Phi_k^{(l),m})$ is achieved on a pure input state, call this state $\rho = |\psi\rangle\langle\psi| \in \mathcal{B}(\mathcal{H}_k)$. This input state is first mapped to the pure state $|\phi\rangle := \alpha_k^{l,m} |\psi\rangle \in \mathcal{B}(\mathcal{H}_l \otimes \mathcal{H}_m)$ before the partial trace is applied. We make a Schmidt decomposition of $|\phi\rangle$ in the form $|\phi\rangle = \sum_i \sqrt{\lambda_i} |\eta_i\rangle \otimes |\zeta_i\rangle$, where $\lambda_i \geqslant 0$ with $\sum_i \lambda_i = 1$, $\{\eta_i\}$ is an orthonormal system in $\mathcal{H}_l$ and $\{\zeta_i\}$ is an orthonormal system in $\mathcal{H}_m$. By Proposition 6.5.1 applied to $\xi := |\phi\rangle$, $\eta := |\eta_i\rangle$ and $\zeta := |\zeta_i\rangle$, we see that $\lambda_i \leqslant [k+1]_q / \theta_q(k,l,m)$. But then, applying the partial trace over $\mathcal{H}_l$ yields $\Phi_k^{(l),m}(\rho) = \sum_i \lambda_i |\eta_i\rangle$. Using the fact that $-\sum_i p_i \log p_i \geqslant -\log \max_i p_i$ for any discrete probability distribution $\{p_i\}$, we see

$$H_{\min}(\Phi_k^{(l),m}) = H(\Phi_k^{(l),m}(\rho)) = -\sum_i \lambda_i \log \lambda_i \geqslant -\log \max_i \lambda_i \geqslant -\log \tfrac{[k+1]_q}{\theta_q(k,l,m)} = \log \tfrac{\theta_q(k,l,m)}{[k+1]_q}. \tag{6.25}$$

□

Theorem 6.5.2: Minimum output entropy for S_N^+

For $N \geqslant 4$, fix an S_N^+-admissible triple $(k, l, m) \in \mathbb{N}_0^3$. We have:

$$H_{\min}(\hat{\Phi}_k^{(l),m}) = H_{\min}(\hat{\Phi}_k^{l,(m)}) \geqslant \begin{cases} \log \frac{\theta_q(2k,2l,2m)}{[2k+1]_q} & r := l+m-k \text{ is even} \\ \log \frac{\theta_q(2k,2l,2m)}{[2k+1]_q} - \frac{1}{2}\log(N) & r := l+m-k \text{ is odd.} \end{cases} \tag{6.26}$$

Proof. The proof is entirely analogous to the proof of Theorem 6.5.1. □

Remark 6.5.1. Note that in the limit of $N \to \infty$, the quantum parameter q for O_N^+ defined by $q+q^{-1} = N$ and $q \leqslant 1$ approaches $1/N$, and in that case the quantum integers $[x+1]_q$ approach N^x for a fixed $x \in \mathbb{N}$. In that case, the quantum factorial $[x+1]_q!$ approaches $N^{x(x+1)/2}$, and an algebraic manipulation of the expression for the θ-net then shows that $\theta_q(k,l,m)$ approaches $N^{(k+l+m)/2}$. Together with the fact that $[k+1]_q$ approaches N^k, this shows that $\frac{[k+1]_q}{\theta_q(k,l,m)}$ is approximately $N^{(k-l-m)/2} = N^{-r}$ where $k = l+m-2r$. In more precise terms: according to [BCLY20, Lem. 4.2], for any O_N^+-admissible $(k,l,m) \in \mathbb{N}_0^3$, $N^r[k+1]_q/\theta_q(k,l,m) = 1 + \mathcal{O}(N^{-2})$. This show that $H_{\min}(\Phi_k^{(l),m})$ grows at least as fast as $r \log N$, which has the same N-dependency as the entropy of the maximally mixed output state, which must be a strict upper bound: this state has entropy $\log[m+1]_q \approx m \log N$, or if we trace out $\mathcal{H}_m$, it has entropy $\log[l+1]_q \approx l \log N$ (remember that $0 \leqslant r \leqslant \min\{l,m\}$).

Similarly, the quantum parameter $\hat{q}$ for S_N^+ given by $\hat{q} + \hat{q}^{-1} = \sqrt{N}$ and $\hat{q} \leqslant 1$ approaches $1/\sqrt{N}$ in the limit $N \to \infty$, and $[2x+1]_{\hat{q}}$ approaches N^x for a fixed $x \in \mathbb{N}_0$. In this case, we have $\frac{[2k+1]_{\hat{q}}}{\theta_{\hat{q}}(2k,2l,2m)} \approx N^{-r/2}$ with r defined by $k = l+m-r$. Thus, $H_{\min}(\hat{\Phi}_k^{(l),m})$ grows at least as fast as $\frac{r}{2} \log N$ in the case where r is even (here, $0 \leqslant r \leqslant 2\min\{l,m\}$), and it grows at least as fast as $\frac{r-1}{2} \log N$ in the case where r is odd.

Remark 6.5.2. The previous theorems show a direct link between the quantum channels associated to O_N^+ and S_N^+: in the case where $N = M^2$ for some $M \in \mathbb{N}$ with $M \geqslant 3$, and we pick an S_N^+-admissible triple $(k,l,m) \in \mathbb{N}_0^3$, the lower bound for the MOE of the S_N^+-channel $\hat{\Phi}_k^{(l),m}$ is identical to the lower bound for the MOE of the O_M^+-channel $\Phi_{2k}^{(2l),2m}$, as long as r defined by $k = l+m-r$ is even. This is a direct consequence of the fattening isomorphism, but it is nonetheless surprising that the lower bound for the MOE shifts by a factor $-\log(M)$ in the case where r is odd.

Theorem 6.5.3

[BC16b, Thm. 3.2] For $N \geqslant 3$, fix a highest-weight O_N^+-admissible triple $(k = l+m, l, m) \in \mathbb{N}_0^3$ (where $l, m \in \mathbb{N}_0$). We have $H_{\min}(\Phi_{l+m}^{(l),m}) = H_{\min}(\Phi_{l+m}^{l,(m)}) = 0$.

Proof. Let $|\eta_l\rangle = |0101\cdots\rangle \in \mathcal{H}_1^{\otimes l}$. We claim $|\eta_l\rangle \in \mathcal{H}_l$, i.e. $p_l |\eta_l\rangle = |\eta_l\rangle$. Consider *any* diagram p in $\mathrm{NC}_2(l,l)$ except the identity: this diagram must have at least one cup and cap. Such a cup connects an odd and an even index, otherwise it encloses an odd number of points which cannot be paired. But then $T_p |\eta_l\rangle = 0$ as this cup connects a $|0\rangle$ with a $|1\rangle$. Hence, by Lemma 5.1.2, $p_l = 1_l + f_l$ with $f_l \in \mathcal{I}_l = \langle u_1, \cdots, u_{l-1}\rangle$, we see $p_l |\eta_l\rangle = |\eta_l\rangle$.

Similarly, we let $|\eta_m\rangle = |0101\cdots\rangle \in \mathcal{H}_1^{\otimes m}$ if the pattern in $|\eta_l\rangle$ ends with a 1, otherwise $|\eta_m\rangle = |1010\cdots\rangle \in \mathcal{H}_1^{\otimes m}$. Then, by a similar argument as before, $|\eta_m\rangle \in \mathcal{H}_m$, and additionally $|\eta_l\rangle \otimes |\eta_m\rangle = |0101\cdots\rangle \in \mathcal{H}_1^{\otimes l+m}$ also has $|\eta_l\rangle \otimes |\eta_m\rangle \in \mathcal{H}_{l+m}$. But then, as $\alpha_{l+m}^{l,m} = (p_l \otimes p_m) p_{l+m}$ by Theorem 6.3.1, and by the absorption property $\alpha_{l+m}^{l,m} = p_{l+m}$ (see Lemma 5.1.3), we see that $\alpha_{l+m}^{l,m} |\eta_l\rangle \otimes |\eta_m\rangle = |\eta_l\rangle \otimes |\eta_m\rangle$. But this is an elementary tensor product, so taking the partial trace over either the first or the second subsystem yields a pure state, in turn we see that $H_{\min}(\Phi_{l+m}^{(l),m}) = H_{\min}(\Phi_{l+m}^{l,(m)}) = 0$. □

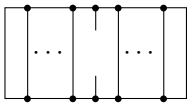
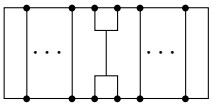
Theorem 6.5.4

For $N \geqslant 4$, fix a highest-weight S_N^+-admissible triple $(k = l+m, l, m) \in \mathbb{N}_0^3$ (with $l, m \in \mathbb{N}_0$). We have $H_{\min}(\hat{\Phi}_{l+m}^{(l),m}) = H_{\min}(\hat{\Phi}_{l+m}^{l,(m)}) = 0$.

Proof. Let

$$|\eta_l\rangle = \tfrac{1}{2^{l/2}}(|0\rangle - |1\rangle) \otimes (|2\rangle - |3\rangle) \otimes (|0\rangle - |1\rangle) \otimes \cdots \in \mathcal{H}^{\otimes l}. \tag{6.27}$$

We claim that $\hat{p}_l |\eta_l\rangle = |\eta_l\rangle$. Through the inverse fattening procedure, the Temperley-Lieb generators u_i are transformed into a_j and b_j, see Lemma 5.2.1, for the reader's convenience:

$a_j =$  and $b_j =$  . (6.28)

Clearly, $T_{b_j} |\eta_l\rangle = 0$ for all $1 \leqslant j \leqslant l-1$, as b_j connects two adjacent nodes whilst two adjacent tensor legs in $|\eta_l\rangle$ never contain the same basis elements. Furthermore, $T_{a_j} |\eta_l\rangle = 0$ for all $1 \leqslant j \leqslant l$ as each individual tensor leg contains an alternating vector. As $\hat{p}_l = 1_l + \hat{f}_l$, where $\hat{f}_l \in \hat{\mathcal{I}}_l = \langle a_1, \ldots, a_l, b_1, \ldots, b_{l-1} \rangle \subseteq \mathrm{NC}(l,l)$ by the fattening procedure, and we just showed that $|\eta_l\rangle$ is in the kernel of all generators except the identity, we see that $\hat{p}_l |\eta_l\rangle = |\eta_l\rangle$.

Analogously, let

$$|\eta_m\rangle = \tfrac{1}{2^{m/2}} (|0\rangle - |1\rangle) \otimes (|2\rangle - |3\rangle) \otimes (|0\rangle - |1\rangle) \otimes \cdots \in \mathcal{H}^{\otimes m} \tag{6.29}$$

if the last tensor leg in $|\eta_l\rangle$ is $|2\rangle - |3\rangle$, otherwise we start $|\eta_m\rangle$ with $|2\rangle - |3\rangle$ in the first tensor leg, and $|0\rangle - |1\rangle$ in the second, etcetera. *Mutatis mutandis*, we see $\hat{p}_m |\eta_m\rangle = |\eta_m\rangle$.

Lastly, we claim that $\hat{p}_{l+m} (|\eta_l\rangle \otimes |\eta_m\rangle) = |\eta_l\rangle \otimes |\eta_m\rangle$. All generators a_j and all but one b_j in $\mathrm{NC}(l+m, l+m)$ are obtained from generators a_i and b_i of $\mathrm{NC}(l,l)$ or $\mathrm{NC}(m,m)$ by tensoring on $|^{\otimes m}$ from the right, or $|^{\otimes l}$ from the left, respectively. It is straightforward to see that these generators still kill $|\eta_l\rangle \otimes |\eta_m\rangle$. We are left with one truly new generator: b_l, connecting the last tensor leg from $|\eta_l\rangle$ with the first tensor leg of $|\eta_m\rangle$. However, we precisely chose those such that they do not contain the same basis elements, so we see $T_{b_l}(|\eta_l\rangle \otimes |\eta_m\rangle) = 0$ as well, and we conclude by a similar argument as before that $\hat{p}_{l+m}(|\eta_l\rangle \otimes |\eta_m\rangle) = |\eta_l\rangle \otimes |\eta_m\rangle$. But then, again using $\hat{\alpha}_{l+m}^{l,m} = \hat{p}_{l+m}$, the input state $|\eta_l\rangle \otimes |\eta_m\rangle$ is a fixed point of $\hat{\alpha}_{l+m}^{l,m}$, whence tracing out either $\hat{H}_l$ or $\hat{H}_m$ yields a pure state, yielding the theorem. □

7

Lowest Weight Clebsch-Gordan Quantum Channels

Having established estimates for the minimum output entropy of the O_N^+-channels and S_N^+-channels in the previous chapter, we wish to investigate their sharpness. Brannan and Collins conjectured in [BC16b, Remark 7] that the lower bound for the MOE of the O_N^+-channel was asymptotically sharp for $N \to \infty$, and indeed in [BCLY20, Thm 4.3] they proved together with Lee and Youn that for any O_N^+-admissible triple $(k, l, m) \in \mathbb{N}_0^3$ with $k = l + m - 2r$:

$$r \log N - C(N) \leqslant H_{\min}(\Phi_k^{(l),m}) = H_{\min}(\Phi_k^{l,(m)}) \leqslant r \log N + D(N), \tag{7.1}$$

with $C(N), D(N) \to 0$ as $N \to \infty$. In this chapter, we will analytically investigate the "lowest-weight" channels, for which we pick $k = |l - m|$, where we will in particular focus on the case of $k = 1$ (thus, $l = m \pm 1$), where we will show that the minimum output entropy can be analytically calculated. This chapter also provides the necessary tools for the numerical approximations to the minimum output entropy for small N that we will describe in following chapters.

7.1. Quantum channels characterized by a projection

Let us commence with a small lemma we will need in the following sections – in the "lowest weight" case $k = |l-m|$ (where (k, l, m) is an admissible triple), we can rewrite the description of the channels $\Phi_k^{(l),m}$ and $\hat{\Phi}_k^{(l),m}$ to an easier form, and this lemma shows us that this description is a valid channel description (i.e., they are completely positive and trace-preserving).

Lemma 7.1.1

Given the finite-dimensional Hilbert spaces $\mathcal{H}_1$, $\mathcal{K}$ and $\mathcal{H}_2 = \mathcal{H}_1 \otimes \mathcal{K}$, and given a linear map $\mathcal{N} : \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$ of the form

$$\mathcal{N}(\rho) = p\,(\rho \otimes \iota_{\mathcal{K}})\,p, \tag{7.2}$$

where $\rho \in \mathcal{B}(\mathcal{H}_1)$, $\iota_{\mathcal{K}}$ is the identity map on $\mathcal{K}$, and $p \in \mathcal{B}(\mathcal{H}_2)$ is an orthogonal projection. The following are equivalent:

1. It is possible to normalize $\mathcal{N}$ to make it trace-preserving on a restricted domain $\mathcal{B}(\mathcal{H}')$, where $\mathcal{H}' \subseteq \mathcal{H}_1$ is a non-trivial subspace;
2. There exists a non-trivial subspace $\mathcal{H}' \subseteq \mathcal{H}_1$ such that the partial trace of p, denoted $\mathrm{Tr}_{\mathcal{K}}(p)$, restricted to this subspace $\mathcal{H}'$, can be expressed as $\mathrm{Tr}_{\mathcal{K}}(p)|_{\mathcal{H}'} = \lambda 1_{\mathcal{H}'}$ where $\lambda \in \mathbb{C}$ and $1_{\mathcal{H}'}$ is the identity on $\mathcal{H}'$.

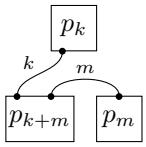
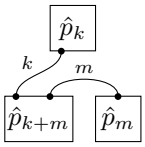
In this case, the normalization of $\mathcal{N}$ is $1/\lambda$.

Proof. By cyclicity of the trace and the fact that p is an orthogonal projection, we consider $\mathrm{Tr}(\mathcal{N}(\rho)) = \mathrm{Tr}(p(\rho \otimes \iota_{\mathcal{K}}))$. From this form, it becomes apparent that $\mathrm{Tr}(\mathcal{N}(\rho)) = \mathrm{Tr}\left(\mathrm{Tr}_{\mathcal{K}}(p)\rho\right)$. Hence, if we assume that $\mathcal{N}$ can be normalized by $1/\lambda$ on a subspace $\mathcal{H}' \subseteq \mathcal{H}_1$ such that it is a quantum channel, we must have $\mathrm{Tr}_{\mathcal{K}}(p)|_{\mathcal{H}'} = \lambda 1_{\mathcal{H}'}$. Conversely, if $\mathrm{Tr}_{\mathcal{K}}(p)|_{\mathcal{H}'} = \lambda 1_{\mathcal{H}'}$ for some subspace $\mathcal{H}' \subseteq \mathcal{H}_1$ and $\lambda > 0$, we can restrict ourselves to $\rho \in \mathcal{B}(\mathcal{H}')$, and we see that $\lambda^{-1}\mathcal{N} : \mathcal{B}(\mathcal{H}') \to \mathcal{B}(\mathcal{H}_2)$ is trace-preserving, that is, $\lambda^{-1}\mathcal{N}$ is a quantum channel. □

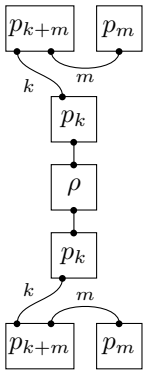
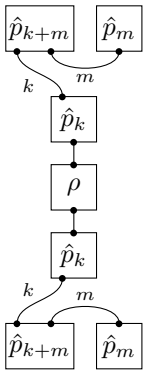
Remark 7.1.1. Note that, if the normalization constant λ in the previous lemma is strictly positive, the map $\lambda^{-1}\mathcal{N}$ is a quantum channel – complete positivity follows directly from the description $\mathcal{N}(\rho) = p(\rho \otimes 1_{\mathcal{K}})p$.

7.2. DESCRIPTION OF THE LOWEST WEIGHT QUANTUM CHANNELS

In the case of O_N^+, a triplet of indices $(k, l, m) \in \mathbb{N}_0^3$ is admissible if there exists an $r \in \mathbb{N}_0$ with $0 \leqslant r \leqslant \min\{l, m\}$ such that $k = l + m - 2r$, corresponding with the fusion rules that tell us that $v_k \subset v_l \otimes v_m$ precisely when $k = l + m - 2r$. We can consider the subrepresentation $v_{|l-m|}$, which we will call the "lowest weight representation". In the S_N^+ case, $k = |l - m|$ is also the lowest-weight representation. Equivalently, we can parametrize the quantum channels by the admissible triplets $(k, l = k + m, m)$ for free parameters $k, m \in \mathbb{N}_0$. In this case, the quantum channels $\Phi_k^{k+m,(m)}$ and $\Phi_k^{(k+m),m}$ (and their S_N^+ counterparts $\hat{\Phi}_k^{k+m,(m)}$ and $\hat{\Phi}_k^{(k+m),m}$) allow for a much easier description than the Stinespring representation, by manipulating their diagrammatic representations and using the absorption property of the Jones-Wenzl projection. Specifically, let us first consider the intertwiners $A_k^{k+m,m}$ and $\hat{A}_k^{k+m,m}$ in $\mathrm{Hom}(v_k, v_{k+m} \otimes v_m)$: using Theorem 6.3.1, we find:

$A_k^{k+m,m} =$  ; $\hat{A}_k^{k+m,m} =$  (7.3)

Hence, the action $\rho \mapsto A_k^{k+m,m}\rho(A_k^{k+m,m})^*$ is:

$A_k^{k+m,m}\rho(A_k^{k+m,m})^* =$  ; $\hat{A}_k^{k+m,m}\rho(\hat{A}_k^{k+m,m})^* =$  (7.4)

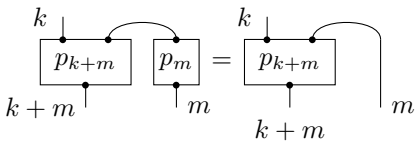
Intuitively, the outer p_m's should not contribute anything, as they are fully connected to p_{k+m}, and the projections p_{k+m} are dominated by $1_k \otimes p_m$. Indeed, we can state a modified absorption property, for which we did not find a direct reference, but it is for example implicitly used by Brannan in [BC20, Section 5] to compute Choi matrices of O_N^+-quantum channels:

Lemma 7.2.1

The modified absorption property of the Jones-Wenzl projection holds:

$$(p_{k+m} \otimes p_m)(\iota_k \otimes T_{\cap^m}) = (p_{k+m} \otimes \iota_m)(\iota_k \otimes T_{\cap^m}), \tag{7.5}$$

or graphically:



The exact same relations also hold for the generalized Jones-Wenzl projections $\hat{p}_k$ and $\hat{p}_{k+m}$.

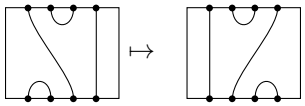
To facilitate this proof, it is beneficial to describe some of the "capping" of diagrams that we see here in more detail. From this discussion in the next subsection, the proof of the lemma will then directly follow.

7.2.1. CAPFORMS AND DIAGRAM MANIPULATIONS

The following section seems related to Morrison's work in [Mor17, Sect. 5], where he manipulates diagram by "folding them down to the right", similar to the expressions $(\iota_k \otimes p)T_{\cap^m}$ below. This is also discussed in [Kau01, Appendix, Sect. III], where the structure of folding a diagram is called a "capform". Note that such capforms actually also fit into the framework of partition theory as described by Freslon and Weber in [FW16]: a capform $(\iota_k \otimes p)T_{\cap^m}$ or $(p \otimes \iota_k)T_{\cap^m}$ is obtained by performing the rotation maps $\curvearrowright$ and $\curvearrowleft$ precisely k times on the partition p, for their definition see Definition 2.3.6.

Definition 7.2.1

Given a partition $p \in \mathcal{P}(k,l)$ for some $k,l \in \mathbb{N}_0$ (note that we allow both crossing and non-pair partitions). We define its *horizontal flip* $p_\leftrightarrow$ as follows. Say the partition p consists of blocks V_i, and has upper nodes $(1,\dots,k)$ and lower nodes $(1',\dots,l')$. Then, $p_\leftrightarrow$ has the same blocks V_i, but we switch its upper nodes around to $(k,\dots,1)$ and its lower nodes to $(l',\dots,1')$. For example:



Remark 7.2.1. Note that we can equivalently define $p_\leftrightarrow = x_k p x_l$, where we will call x_t for $t \in \mathbb{N}$ the "reflection partition" in $\mathcal{P}(t,t)$ that only contain through-blocks connecting the i-th upper node to the $(t-i+1)$'th lower node. This is a useful identity as this allows us to define $a_\leftrightarrow := T_{x_l} a T_{x_k}$ for any *operator* $a \in \mathcal{B}(\mathcal{H}_1^{\otimes k}, \mathcal{H}_1^{\otimes l})$. Also note that $(pq)_\leftrightarrow = p_\leftrightarrow q_\leftrightarrow$, which follows from the fact that $x_k^2 = \iota_k$.

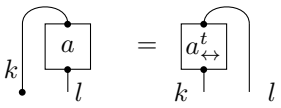
Remark 7.2.2. A moment's reflection will convince the reader that for any $p \in \mathcal{P}(k,l)$, $(p^*)_\leftrightarrow = (p_\leftrightarrow)^*$, as this corresponds to the fact that the reflections along the canonical axes in two dimensions commute. We will therefore freely write $p^*_\leftrightarrow$.

Lemma 7.2.2: Pulling an operator through a cap

For any operator $a \in \mathcal{B}(\mathcal{H}_1^{\otimes k}, \mathcal{H}_1^{\otimes l})$ for some $k,l \in \mathbb{N}$, we have

$$(\iota_k \otimes a)T_{\cap^k} = (a^t_\leftrightarrow \otimes \iota_k)T_{\cap^k}, \tag{7.6}$$

where $a^t_\leftrightarrow := T_{x_k} a^t T_{x_l}$, where x_i are the reflection partitions from Remark 7.2.1. Diagrammatically:



Proof. Writing out both sides yields:

$$(\iota_k \otimes a)T_{\cap^k} = \sum_{\boldsymbol{i}} |i_1 \cdots i_k\rangle \otimes (a\,|i_k \cdots i_1\rangle)\,, \tag{7.7}$$

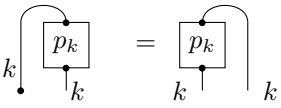
$$(a^t_{\leftrightarrow} \otimes \iota_k)T_{\cap^k} = \sum_{\boldsymbol{i}} \left(a^t_{\leftrightarrow}\,|i_1 \cdots i_l\rangle\right) \otimes |i_l \cdots i_1\rangle\,. \tag{7.8}$$

Taking the inner product with the basis element $\langle j_1, \ldots, j_k, j'_1, \ldots, j'_l|$ then yields the expressions $\langle j'_1, \ldots, j'_l|\, a\, |j_k, \ldots, j_1\rangle$ and $\langle j_1, \ldots, j_k|\, T_{x_k} a^t T_{x_l}\, |j'_l, \ldots, j'_1\rangle$, respectively, which are equal to each other. □

Note that the previous lemma holds in particular for partitions, so we can prove the useful corollary:

Lemma 7.2.3

The Jones-Wenzl projection $p_k \in \mathrm{TL}_k(d)$ satisfies $p_k^t = p_k$ and $(p_k)_{\leftrightarrow} = p_k$, hence we can "pull it through" a cap or cup:



The generalized version $\hat{p}_k \in \mathrm{NC}_{\mathrm{lin}}[d]$ satisfies the same relation.

Proof. Note that $p_k^* = p_k$ by the fact that p_k is an orthogonal projection, and $p_k^t = p_k$ follows from the Wenzl recursion formula, which shows that p_k contains only real coefficients. So, we are left with proving $(p_k)_{\leftrightarrow} = p_k$. Note first that $(p_k)_{\leftrightarrow}$ itself is again a projection in $\mathrm{TL}_k(d)$, which follows from Remark 7.2.1 and Remark 7.2.2, as

$$(p_k)_{\leftrightarrow}(p_k)_{\leftrightarrow} = (p_k p_k)_{\leftrightarrow} = (p_k)_{\leftrightarrow} \quad , \quad (p_k)^t_{\leftrightarrow} = (p_k^t)_{\leftrightarrow} = (p_k)_{\leftrightarrow}. \tag{7.9}$$

Now, our strategy is to show that $(p_k)_{\leftrightarrow} u_i = 0$ for all generators u_i of $\mathrm{TL}_k(d)$, as this is the unique property of the Jones-Wenzl projection, and thus implies $(p_k)_{\leftrightarrow} = p_k$. A quick look at the diagram of u_i shows that $(u_i)_{\leftrightarrow} = u_{k-i}$. But then, $(p_k)_{\leftrightarrow} u_i = (p_k)_{\leftrightarrow}(u_{k-i})_{\leftrightarrow} = (p_k u_{k-i})_{\leftrightarrow} = 0$. Hence, $(p_k)_{\leftrightarrow} \in \mathrm{TL}_k(d)$ is a projection that satisfies the unique relations of the Jones-Wenzl projection, so the lemma follows. □

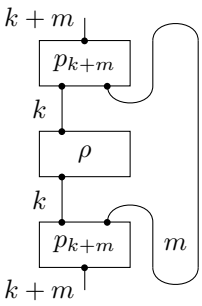
Proof of Lemma 7.2.1. Pull p_m through the cap to the right by using Lemma 7.2.3, then we get

$$(p_{k+m} \otimes \iota_m)(\iota_k \otimes p_m \otimes \iota_m)(\iota_k \otimes T_{\cap^m}), \tag{7.10}$$

but then we can use the usual adsorption property $p_{k+m}(\iota_k \otimes p_m) = p_{k+m}$, and the lemma follows. □

7.2.2. LOWEST WEIGHT CHANNELS: PARTIAL TRACE OVER $\mathcal{H}_m$

We consider the quantum channels $\Phi_k^{k+m,(m)}$. By the absorption property of the Jones-Wenzl projection, the p_m's are absorbed by p_{k+m}, so we have, up to normalization:



That is, up to normalization, we see that the quantum channel is described by the action

$$\rho \mapsto p_{k+m}\,(\rho \otimes 1_m)\,p_{k+m}, \tag{7.11}$$

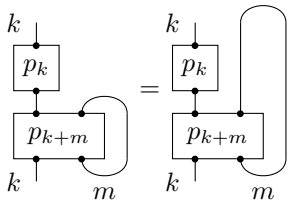
where 1_m is the identity on $\mathcal{H}_1^{\otimes m}$. It is not *a priori* clear that this description is normalizable to make the map trace-preserving (and for general projections, it truly is not normalizable, as we might be able to find a density matrix ρ such that $\rho \otimes 1_m$ is in the kernel of p_{k+m}), but according to Lemma 7.1.1, this mapping can be normalized iff the partial trace of p_{k+m} over $\mathcal{H}_m$ yields a positive scalar multiple of an orthogonal projection. Furthermore, as we already know that the domain of $\Phi_k^{k+m,(m)}$ is $\mathcal{B}(\mathcal{H}_k)$, this partial trace would have to be a positive scalar multiple of the Jones-Wenzl projection p_k. Luckily, we have:

Lemma 7.2.4

For any $k, m \in \mathbb{N}_0$, the partial trace over $\mathcal{H}_m$ of the Jones-Wenzl projection $p_{k+m} \in \mathcal{B}(\mathcal{H}_k \otimes \mathcal{H}_m)$ yields a positive scalar multiple λ of the Jones-Wenzl projection p_k, and this scalar is:

$$\lambda = \frac{[k+m+1]_q}{[k+1]_q}. \tag{7.12}$$

Proof. Note that the diagram of the partial trace of p_{k+m} over $\mathcal{H}_m$ lives in $\mathrm{TL}_k(d)$, so it must be expressible as $q := \mathrm{Tr}_{\mathcal{H}_m}(p_{k+m}) = \lambda 1_k + f_k$, where f_k is an element of the two-sided ideal $\mathcal{I}_k$ generated by the generators $\{u_1, \ldots, u_{k-1}\} \subseteq \mathrm{TL}_k(d)$, and $\lambda \in \mathbb{C}$. However, if we compute $p_k q$, we can see from the diagrams:



Now, from the absorption property $(p_k \otimes \iota^{\otimes m})p_{k+m} = p_{k+m}$, see Lemma 5.1.3, it is clear that $p_k q = q$. However, then:

$$q = p_k q = p_k(\lambda 1_k + f_k) = \lambda p_k + 0, \tag{7.13}$$

so indeed $\mathrm{Tr}_{\mathcal{H}_m}(p_{k+m}) = \lambda p_k$. We are left with calculating the value of λ.

By Lemma 7.1.1, $\rho \mapsto \lambda^{-1} p_{k+m}(\rho \otimes 1_m)p_{k+m}$ is trace-preserving on the domain $\mathcal{B}(\mathcal{H}_k)$. Hence, if we input $\rho = p_k$, we find:

$$p_k \mapsto \lambda^{-1} p_{k+m}(p_k \otimes 1_m)p_{k+m} = \lambda^{-1} p_{k+m}, \tag{7.14}$$

which again follows from the absorption property of p_{k+m}, see Lemma 5.1.3. But then, we compute:

$$[k+1]_q = \dim \mathcal{H}_k = \dim(p_k H_1^{\otimes k}) = \mathrm{Tr}(p_k) = \lambda^{-1}\mathrm{Tr}(p_{k+m}) = \lambda^{-1} \dim \mathcal{H}_{k+m} = \lambda^{-1}[k+m+1]_q, \tag{7.15}$$

whence $\lambda = [k+m+1]_q / [k+1]_q$. □

From the previous lemma and Lemma 7.1.1, we directly recover:

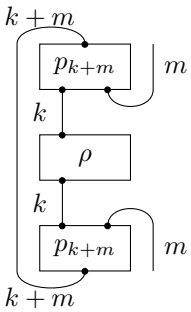
Lemma 7.2.5

We have

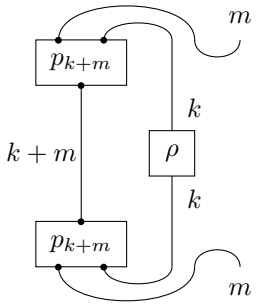
$$\Phi_k^{k+m,(m)} : \mathcal{B}(\mathcal{H}_k) \to \mathcal{B}(\mathcal{H}_{k+m}) \quad : \quad \rho \mapsto \frac{[k+1]_q}{[k+m+1]_q} p_{k+m}(\rho \otimes 1_m) p_{k+m}. \tag{7.16}$$

7.2.3. LOWEST WEIGHT CHANNELS: PARTIAL TRACE OVER $\mathcal{H}_{k+m}$

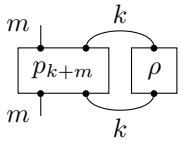
We now consider the quantum channels $\Phi_k^{(k+m),m}$, so we have up to normalization the following diagram:



By Lemma 7.2.3, we can pull the Jones-Wenzl projection p_{k+m} through the caps and cups on the left-hand side, which transforms this diagram into:



But now, as $p_{k+m}^2 = p_{k+m}$, this simplifies to



Hence, up to normalization, our quantum channel can be described by

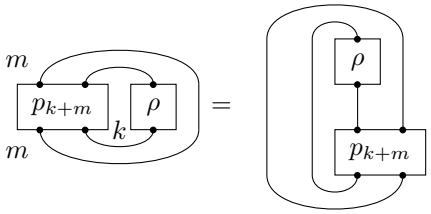
$$\rho \mapsto (\iota_m \otimes T_{\cup^k})(p_{k+m} \otimes \rho)(\iota_m \otimes T_{\cap^k}). \tag{7.17}$$

Alternatively, pulling ρ through $\cap^k$ and taking the partial Markov trace over the last k strands yields

$$\rho \mapsto \mathrm{Tr}_{\mathcal{H}_k}\left(p_{k+m}(\iota_m \otimes T_{x_k} \rho^t T_{x_k})\right), \tag{7.18}$$

where T_{x_k} is the partition map associated to the partition $x_k \in \mathcal{P}_2(k,k)$ that contains pairs $\{i, (k-i+1)'\}$.

Again, linearity and complete positivity are obvious, but whether this channel can be normalized to be trace-preserving is not immediately clear. However, again using Lemma 7.2.3, we see that when we take the trace, our diagrams can be manipulated as:



And this trace can be evaluated as $\mathrm{Tr}(p_{k+m}(\rho \otimes 1_m))$, and we see the familiar expression

$$\mathrm{Tr}(p_{k+m}(\rho \otimes 1_m)) = \mathrm{Tr}(\mathrm{Tr}_{\mathcal{H}_m}(p_{k+m})\rho) = \frac{[k+m+1]_q}{[k+1]_q}\mathrm{Tr}(p_k\rho) = \frac{[k+m+1]_q}{[k+1]_q}\mathrm{Tr}(\rho), \tag{7.19}$$

which gives us the proof to the following lemma.

Lemma 7.2.6

We have

$$\Phi_k^{(k+m),m} : \mathcal{B}(\mathcal{H}_k) \to \mathcal{B}(\mathcal{H}_m) \quad : \quad \rho \mapsto \frac{[k+1]_q}{[k+m+1]_q}(\iota_m \otimes T_{\cup^k})(p_{k+m} \otimes \rho)(\iota_m \otimes T_{\cap^k}). \tag{7.20}$$

7.3. PPT AND EBT PROPERTIES FOR LOWEST WEIGHT CHANNELS

Having established the descriptions of $\Phi_k^{k+m,(m)}$ and $\Phi_k^{(k+m),m}$, we can investigate their information-theoretic properties such as whether they are positive-partial transpose or entanglement-breaking. Before stating the next theorem, we need a small lemma.

Lemma 7.3.1

Consider any partition $p \in \mathrm{NC}_2(k)$ for any $k \in \mathbb{N}$. If a block X in p contains two indices from either the upper row or the lower row, then one of these indices is odd and the other is even. If a block Y in p contains an index from the upper row and an index from the lower row, either both indices are odd, or both are even.

Proof. In the case of a block containing two upper indices or lower indices that are either both even or both odd, in the diagram of p they enclose an odd number of indices below their cap or above their cup. This is impossible, as those points need to form pairs and cannot cross the cap or cup. Similarly, if an upper index and a lower index are paired, and one is odd whilst the other one is even, the diagram of the partition is split vertically into two non-interacting parts, but those parts both contain an odd number of points, whilst no non-crossing partitions exist with an odd total number of points. □

The following theorem is an improvement to the negative PPT-result in [BCLY20, Cor. 5.4] (as it holds for all $N \geqslant 3$), and is complementary to the EBT property established in [BCLY20, Thm. 5.2].

Theorem 7.3.1

The quantum channels $\Phi_k^{k+m,(m)}$ are not PPT, and therefore also not EBT.

Proof. The Choi matrix $J(\Phi_k^{k+m,(m)})$ can be calculated as follows, with $\lambda := \frac{[k+1]_q}{[k+m+1]_q} > 0$:

$$J(\Phi_k^{k+m,(m)}) = \sum_{ij} |e_i\rangle\langle e_j| \otimes \Phi_k^{k+m,(m)}(|e_i\rangle\langle e_j|) = \lambda \sum_{ij} |e_i\rangle\langle e_j| \otimes [p_{k+m}\,(|e_i\rangle\langle e_j| \otimes \iota_m)\,p_{k+m}]$$
$$= \lambda\,(\iota_{\mathcal{H}_k} \otimes p_{k+m})\,\Big(\sum_{ij} |e_i\rangle\langle e_j| \otimes |e_i\rangle\langle e_j| \otimes \iota_m\Big)\,(\iota_{\mathcal{H}_k} \otimes p_{k+m})\,. \tag{7.21}$$

Here, $\{|e_i\rangle\}_{i=1}^{[k+1]_q}$ is an orthonormal basis for $\mathcal{H}_k$, *not* for $\mathcal{H}_1^{\otimes k}$ (for those orthonormal bases, we reserve the notation $\{|i\rangle\}_{i:[k]\to[N]}$), as the domain of the quantum channels is restricted to $\mathcal{B}(\mathcal{H}_k)$ (as they are not trace-preserving on the orthogonal complement of $\mathcal{H}_k$ in $H_1^{\otimes k}$). If we apply the partial transpose to the first system, denoted by $\theta_k \otimes \iota_{k+m}$, we get:

$$(\theta_k \otimes \iota_{k+m})\,J(\Phi_k^{k+m,(m)}) = \lambda\,(\iota_{\mathcal{H}_k} \otimes p_{k+m})\,\Big(\sum_{ij} |e_j\rangle\langle e_i| \otimes |e_i\rangle\langle e_j| \otimes \iota_m\Big)\,(\iota_{\mathcal{H}_k} \otimes p_{k+m})\,. \tag{7.22}$$

There is a strong analogy with the fact that the transpose map itself is not 2-positive, although we believe the proof is not yet trivialized by this, as we need to carefully deal with the projections around this expression.

As the diagram of p_{k+m} can be expressed as $|^{\otimes k+m} + f_{k+m}$ for $f_{k+m} \in \mathcal{I}_{k+m}$, where $\mathcal{I}_{k+m}$ is the two-sided ideal generated by the generators $\{u_1, \ldots, u_{k+m-1}\} \subseteq \mathrm{TL}_{k+m}(d)$ (see Lemma 5.1.2), we can pick vectors that are annihilated by *all* diagrams in $\mathrm{NC}_2(k+m)$ that are not the unit: these vectors are then guaranteed to lie in the image of p_{k+m}. As all those diagrams contain at least one cap, by Lemma 7.3.1, we know that these caps always connect an odd and an even index. Thus, if we write $|x\rangle = |i_1, \ldots, i_{k+m}\rangle \in \mathcal{H}_1^{\otimes k+m}$, as long as all pairs $\{i_j, i_k\}$ with an odd j and even k have $i_j \neq i_k$, the vector $|x\rangle$ is annihilated by all diagrams in $\mathrm{NC}_2(k+m)$ except the identity, and we find $p_{k+m}\,|x\rangle = |x\rangle$.

Let us pick $|x_1\rangle = |i_1, \ldots, i_k\rangle$ as a repeating pattern of 0121 until we run out of indices, and $|x_2\rangle = |j_1, \ldots, j_k\rangle$ as a repeating pattern of 2101 until we run out of indices. Note that for any $k \in \mathbb{N}$, $|x_1\rangle \perp |x_2\rangle$. Now put $|y\rangle = |l_1, \ldots, l_m\rangle$ as a repeating pattern of 01 until we run out of indices if k is even, and a repeating pattern of 10 until we run out of indices if k is odd.

The reader is invited to check that $|x_1\rangle\,|y\rangle$ always contains a pattern with all odd indices equal to 1, and all even indices equal to 0 or 2. Similarly for $|x_2\rangle\,|y\rangle$. The clue lies in the fact that $|x_1\rangle$ and $|x_2\rangle$ both lie in the image of p_k, and $|x_1\rangle\,|y\rangle$ and $|x_2\rangle\,|y\rangle$ both lie in the image of p_{k+m}. But then, we see that

$$(\theta_k \otimes \iota_{k+m})\,J(\Phi_k^{k+m,(m)})\,|x_1\rangle\,|x_2\rangle\,|y\rangle = \lambda\,|x_2\rangle\,|x_1\rangle\,|y\rangle \tag{7.23}$$
$$(\theta_k \otimes \iota_{k+m})\,J(\Phi_k^{k+m,(m)})\,|x_2\rangle\,|x_1\rangle\,|y\rangle = \lambda\,|x_1\rangle\,|x_2\rangle\,|y\rangle\,. \tag{7.24}$$

Thus, we conclude that the vector

$$|z\rangle = (|x_1\rangle\,|x_2\rangle - |x_2\rangle\,|x_1\rangle)\,|y\rangle \tag{7.25}$$

is an eigenvector of the partially transposed Choi matrix with eigenvalue $-\lambda < 0$. □

We did not find a counterpart to Theorem 7.3.1 in the case where we take the partial trace over $\mathcal{H}_{k+m}$ (i.e., considering the channels $\Phi_k^{(k+m),m}$ instead of $\Phi_k^{k+m,(m)}$) for general k. However, if we specify to the $k = 1$ case, these complementary channels are interestingly related in the following way:

Theorem 7.3.2

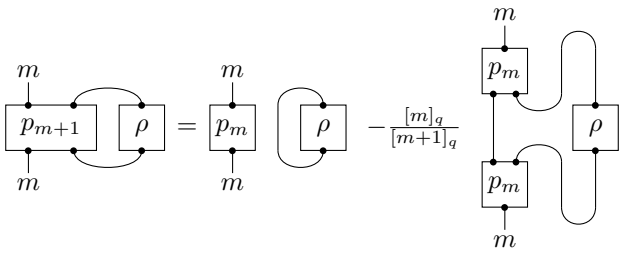
For $k = 1$, $m \in \mathbb{N}$ and $l = m + 1$, the quantum channel $\Phi_1^{(m+1),m}$ can be described as

$$\begin{aligned}\Phi_1^{(m+1),m}(\rho) &= \tfrac{[2]_q}{[m+2]_q}(\iota_m \otimes T_\cap)(p_{m+1} \otimes \rho)(\iota_m \otimes T_\cup) \\ &= \tfrac{[2]_q}{[m+2]_q}\mathrm{Tr}(\rho)p_m - \tfrac{[m]_q}{[m+2]_q}\tfrac{[2]_q}{[m+1]_q}p_m(\iota_{m-1} \otimes \rho)p_m \\ &= T_{x_m}\left(\tfrac{[2]_q}{[m+2]_q}\mathrm{Tr}(\rho)p_m - \tfrac{[m]_q}{[m+2]_q}\Phi_1^{m,(m-1)}(\rho)\right)T_{x_m}. \end{aligned} \tag{7.26}$$

Proof. The first equality follows from the Wenzl recursion formula:

$$p_{m+1} = p_m \otimes \iota - \tfrac{[m]_q}{[m+1]_q}(p_m \otimes \iota)u_{m-1}(p_m \otimes \iota). \tag{7.27}$$

Applied to our problem:



It is not obvious from the description of $\Phi_1^{(m+1),m}$ that this quantum channel is trace-preserving (obviously, this should hold if our diagram manipulations and the original normalization of $\Phi_1^{(m+1),m}$ are correct, so the reader can interpret the next steps as a sanity check). However, we have:

$$\mathrm{Tr}(\Phi_1^{(m+1),m}(\rho)) = \tfrac{[2]_q}{[m+2]_q}[m+1]_q\mathrm{Tr}(\rho) - \tfrac{[2]_q[m]_q}{[m+2]_q[m+1]_q}\tfrac{[m+1]_q}{[2]_q}\mathrm{Tr}(\rho), \tag{7.28}$$

where the second term can be computed by first computing the partial trace of p_m over $\mathcal{H}_{m-1}$ by Lemma 7.2.4. The claim follows from the identity $[2]_q[m+1]_q - [m]_q = [m+2]_q$ [1].

To verify the second equality, note that the term $p_m(\iota_{m-1} \otimes \rho)p_m$ *looks like* $\Phi_1^{m,(m-1)}$, but the density matrix ρ is in the last tensor leg instead of the first. Luckily, by Lemma 7.2.3, we see that $(p_m)_\leftrightarrow = p_m$, equivalently $x_m p_m x_m = p_m$, where x_m is the "reflection partition" introduced in Remark 7.2.1. On the level of partition maps, we have $T_{x_m} : |i_1 \cdots i_m\rangle \mapsto |i_m \cdots i_1\rangle$. The reader is invited to check that $T_{x_m}(\iota_{m-1} \otimes \rho)T_{x_m} = \rho \otimes \iota_{m-1}$, which works because ρ acts on only one strand in the diagram, so trivially $\rho_\leftrightarrow = \rho$. But then, we have

$$p_m(\iota_{m-1} \otimes \rho)p_m = T_{x_m}p_m(\rho \otimes \iota_{m-1})p_m T_{x_m}. \tag{7.30}$$

As T_{x_m} commutes with p_m, we can write

$$\Phi_1^{(m+1),m}(\rho) = T_{x_m}\left(\tfrac{[2]_q}{[m+2]_q}\mathrm{Tr}(\rho)p_m - \tfrac{[m]_q}{[m+2]_q}\Phi_1^{m,(m-1)}(\rho)\right)T_{x_m}. \tag{7.31}$$

□

Using this description for $\Phi_1^{(m+1),m}$ allows us to state the following theorem, which is complementary to the negative PPT result established in [BCLY20, Cor. 5.4], as we discuss the case $k = l - m$ that the authors excluded:

[1] This can be verified using the identity

$$[m+n]_q = q^{-n}[m]_q + q^m[n]_q, \tag{7.29}$$

which itself follows from the well-known identity for q-integers $(m+n)_q = (m)_q + q^m(n)_q$ by using $[n]_q = q^{1-n}(n)_{q^2}$, see the introduction of Section 6.3.

Theorem 7.3.3

For $k = 1$, the quantum channels $\Phi_1^{(m+1),m}$ are PPT.

Proof. From the proof of Theorem 7.3.2, we know that $\Phi_1^{(m+1),m}(\rho) = U\mathcal{N}_m(\rho)U^*$, where U is the unitary T_{x_m} (independent of ρ), and $\mathcal{N}_m$ is the quantum channel

$$\mathcal{N}_m : \rho \mapsto \tfrac{[2]_q}{[m+2]_q}\mathrm{Tr}(\rho)p_m - \tfrac{[m]_q}{[m+2]_q}\Phi_1^{m,(m-1)}(\rho). \tag{7.32}$$

Writing out the definition of the Choi matrix, we see that $J(\Phi_1^{(m+1),m})$ is related by the unitary transformation $\iota_1 \otimes U$ to $J(\mathcal{N}_m)$, which does not change the eigenvalues, so we consider $J(\mathcal{N}_m)$ instead. This Choi matrix equals

$$J(\mathcal{N}_m) = \tfrac{[2]_q}{[m+2]_q}\iota_1 \otimes \iota_{\mathcal{H}_m} - \tfrac{[m]_q}{[m+2]_q}J(\Phi_1^{m,(m-1)})(\rho). \tag{7.33}$$

But then, we see that the partial transpose over the first system, denoted by $\theta_1 \otimes \iota_m$, yields

$$(\theta_1 \otimes \iota_m)J(\mathcal{N}_m) = \tfrac{[2]_q}{[m+2]_q}\iota_1 \otimes \iota_{\mathcal{H}_m} - \tfrac{[m]_q}{[m+2]_q}(\theta_1 \otimes \iota_m)J(\Phi_1^{m,(m-1)})(\rho). \tag{7.34}$$

If we investigate the description of the partially transposed Choi matrix of $\Phi_1^{m,(m-1)}$ in the proof of Theorem 7.3.1, we see that it consists of a composition of contractions, hence its largest possible eigenvalue is $\frac{[2]_q}{[m+1]_q}$. But then, we see that $\mathcal{N}_m$ is PPT if

$$\tfrac{[2]_q}{[m+2]_q} \overset{?}{\geqslant} \tfrac{[m]_q}{[m+2]_q}\tfrac{[2]_q}{[m+1]_q}. \tag{7.35}$$

But this inequality is equivalent to $1 \geqslant \frac{[m]_q}{[m+1]_q}$, which always holds. Thus, $\mathcal{N}_m$ is PPT, and as $\Phi_1^{(m+1),m}$ is a rotated version of $\mathcal{N}_m$, the theorem follows. □

Remark 7.3.1. Note that the upper bound on the largest possible eigenvalue of $J(\Phi_1^{m,(m-1)})$ is actually achieved, as can be deduced from Theorem 7.3.1, using $|z\rangle = (|x_1\rangle|x_2\rangle + |x_2\rangle|x_1\rangle)|y\rangle$. This immediately tells us that

$$\min_\lambda (\theta_1 \otimes \iota_m)J(\Phi_1^{(m+1),m}) = \tfrac{[2]_q}{[m+2]_q} - \tfrac{[m]_q}{[m+2]_q}\tfrac{[2]_q}{[m+1]_q}. \tag{7.36}$$

Remark 7.3.2. Note that Theorem 7.3.2 tells us that

$$\Phi_1^{(m+1),m}(\rho) = \tfrac{[2]_q}{[m+2]_q}p_m\left(\mathrm{Tr}(\rho)\iota_m - \tfrac{[m]_q}{[m+1]_q}\iota_{m-1}\otimes\rho\right)p_m. \tag{7.37}$$

We can write this action as a concatenation of simpler quantum channels: we note that we can first send $\rho \mapsto \mathrm{Tr}(\rho)\iota_1 - \frac{[m]_q}{[m+1]_q}\rho$. With the proper normalization:

$$\mathcal{N} : \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_1) \quad : \quad \rho \mapsto \tfrac{N}{N - [m]_q/[m+1]_q}\left(\mathrm{Tr}(\rho)\tfrac{\iota_1}{N} - \tfrac{[m]_q}{N[m+1]_q}\rho\right), \tag{7.38}$$

and let $\mathcal{M}$ be the quantum channel

$$\mathcal{M} : \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_m) \quad : \quad \rho \mapsto \tfrac{[2]_q}{[m+1]_q}p_m\left(\iota_{m-1}\otimes\rho\right)p_m. \tag{7.39}$$

Then, $\mathcal{M}\circ\mathcal{N} = \Phi_1^{(m+1),m}$ (the multiplication of the normalization constants in $\mathcal{N}$ and $\mathcal{M}$ works out to $\frac{[2]_q}{[2]_q[m+1]_q - [m]_q}$, and $[2]_q[m+1]_q - [m]_q = [m+2]_q$). The map $\mathcal{N}$ *looks like* a depolarization channel Δ_λ with parameters $\lambda = -\frac{N}{N - [m]_q/[m+1]_q}\frac{[m]_q}{N[m+1]_q}$. However, working out this parameter yields

$$\lambda = -\tfrac{[m]_q}{N[m+1]_q - [m]_q} = -\tfrac{[m]_q}{[2]_q[m+1]_q - [m]_q} = -\tfrac{[m]_q}{[m+2]_q}. \tag{7.40}$$

One can check that $-\frac{[m]_q}{[m+2]_q} \leqslant \frac{-1}{N^2-1}$ where equality only holds if $m = 1$. [2] We know that Δ_λ is not completely positive for $\lambda < -\frac{1}{N^2-1}$ [LH16] whence $\mathcal{N}$ is *not* a channel, so we cannot conclude that $\Phi_1^{(m+1),m}$ is entanglement-breaking for $m > 1$.

However, for $m = 1$ we see that $\lambda = -[1]_q/[3]_q = -\frac{1}{N^2-1}$, which leads to the following theorem, which is a positive answer to the question posed in [BCLY20, Remark 5.3] whether the channels $\Phi_{l-m}^{(l),m}$ are entanglement-breaking, but only in the very smallest non-trivial (i.e. $k, l, m \neq 0$) case:

Theorem 7.3.4

For $k = m = 1$, the quantum channel $\Phi_1^{(2),1}$ is entanglement-breaking.

Proof. From the previous remark, we know that $\Phi_1^{(2),1} = \mathcal{M} \circ \mathcal{N}$, where $\mathcal{N} = \Delta_\lambda$ is a depolarizing channel on $\mathcal{B}(\mathcal{H}_1)$ with $\lambda = -\frac{1}{N^2-1}$, which is on the boundary of (but inside) the domain $-\frac{1}{d^2-1} \leqslant \lambda \leqslant \frac{1}{d+1}$ (with d the dimension of the Hilbert space, in our case $d = \dim \mathcal{H}_1 = N$), whence $\mathcal{N}$ is entanglement breaking, see for example [LH16]. Although it is true in general that the composition of a quantum channel with an entanglement-breaking channel is always itself entanglement-breaking [3], in this case the proof simplifies because $\mathcal{M}$ is the identity channel in the case $m = 1$, so $\Phi_1^{(2),1}$ is a proper depolarizing channel Δ_λ with $\lambda = \frac{-1}{N^2-1}$. □

Having established results regarding the PPT and EBT property of the lowest-weight channels, we specify further to the $k = 1$ case (with $l = m \pm 1$), where some considerations lead to the surprising fact that they have constant (and explicitly computable) output entropy.

7.4. CONSTANT OUTPUT ENTROPY FOR O_N^+ CHANNELS WITH $k = 1$

The smallest non-trivial choice of k for the quantum channels $\Phi_k^{k+m,(m)}$ and $\Phi_k^{(k+m),m}$ is $k = 1$, where the fundamental representation $v_1 = u$ of O_N^+ is embedded in $v_{m+1} \otimes v_m$. Remarkably, in this case, we have:

Theorem 7.4.1

In the case of $k = 1$, $m \in \mathbb{N}$, and $l = m + 1$, the quantum channels $\Phi_1^{m+1,(m)}$ and $\Phi_1^{(m+1),m}$ yield an output entropy that is constant on all pure states, because for $\Phi = \Phi_1^{(m+1),m}$ or $\Phi = \Phi_1^{m+1,(m)}$, all pairs of outputs $\Phi(|\psi\rangle\langle\psi|)$ and $\Phi(|\phi\rangle\langle\phi|)$ are related by a unitary (that generally depends on the choice of $|\psi\rangle$ and $|\phi\rangle$).

Before we dive into the proof of this theorem, the following elementary lemma shows that it suffices to prove the previous theorem just for $\Phi_1^{m+1,(m)}$.

Lemma 7.4.1: Same non-zero spectrum for complementary channel output

Given any quantum channel $\Phi : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$. Then, the matrix $\Phi(|\psi\rangle\langle\psi|)$ has the same non-zero spectrum as the matrix $\Phi^c(|\psi\rangle\langle\psi|)$ for any pure state $|\psi\rangle \in \mathcal{H}_A$.

Proof. Describe Φ in its Stinespring representation as $\Phi(\rho) = \mathrm{Tr}_E(V\rho V^*)$ with Stinespring isometry $V : \mathcal{H}_A \to \mathcal{H}_B \otimes \mathcal{H}_E$, where $\mathcal{H}_A$, $\mathcal{H}_B$ and $\mathcal{H}_E$ are finite-dimensional Hilbert spaces.

[2] Use $N^2 - 1 = [3]_q$, so we need to prove $\frac{[m]_q}{[m+2]_q} \geqslant \frac{1}{[3]_q}$, equivalently $[3]_q[m]_q \geqslant [m+2]_q$. Write $[3]_q = [1+2]_q = q^{-2}[1]_q + q[2]_q$ and $[m+2]_q = q^{-2}[m]_q + q^m[2]_q$, so we can equivalently prove $q^{-2}[m]_q + q[2]_q[m]_q \geqslant q^{-2}[m]_q + q^m[2]_q$, equivalently $[m]_q \geqslant q^{m-1}$. This is obviously true for all $m \geqslant 1$ as $[m]_q = \dim \mathcal{H}_{m-1} \geqslant 1$ whilst $q^{m-1} \leqslant 1$ as $q \leqslant 1$. We see that equality only holds for $m = 1$.

[3] This follows from the definition: any channel applied *after* an entanglement-breaking channel $\mathcal{N}$ cannot "restore" the entanglement: by Definition 3.5.2, we know that we can write $(\iota_C \otimes \mathcal{N})(\rho) = \sum_i p_i \sigma_i^C \otimes \rho_i^B$ for any auxiliary Hilbert space $\mathcal{H}_C$, where $\{p_i\}$ is a discrete probability distribution and $\sigma_i^C \in \mathcal{S}(\mathcal{H}_C)$ and $\rho_i^B \in \mathcal{S}(\mathcal{H}_1)$. But then, $(\iota_C \otimes \mathcal{M} \circ \mathcal{N})(\rho) = \sum_i p_i \sigma_i^C \otimes \mathcal{M}(\rho_i^B)$, which shows that $(\iota_C \otimes \mathcal{M} \circ \mathcal{N})(\rho)$ is separable.

Then, note that for any pure state $\rho = |\psi\rangle\langle\psi|$ with $|\psi\rangle \in \mathcal{H}_A$, we have $V\rho V^* = |\phi\rangle\langle\phi|$ with $|\phi\rangle := V|\psi\rangle \in \mathcal{H}_B \otimes \mathcal{H}_E$. Make a Schmidt decomposition of $|\phi\rangle$ by finding an orthonormal system $(|x_i\rangle)_{i=1}^d \subset \mathcal{B}(\mathcal{H}_B)$ and an orthonormal system $(|y_i\rangle)_{i=1}^d \subset \mathcal{B}(\mathcal{H}_E)$ and scalars $\alpha_1 \geqslant \alpha_2 \geqslant \cdots \geqslant \alpha_d \geqslant 0$, where $d = \min\{\dim \mathcal{H}_B, \dim \mathcal{H}_E\}$, such that $|\phi\rangle = \sum_{i=1}^d \sqrt{\alpha_i}\,|x_i\rangle \otimes |y_i\rangle$. Then, we see that the partial trace over E yields $\sum_i \alpha_i |x_i\rangle\langle x_i|$, whilst a partial trace over B yields $\sum_i \alpha_i |y_i\rangle\langle y_i|$. This shows that both non-zero spectra of $\Phi(|\psi\rangle\langle\psi|)$ and $\Phi^c(|\psi\rangle\langle\psi|)$ consist precisely of those $\alpha_i \neq 0$, and the lemma follows. □

Remark 7.4.1. Note that a direct corollary from this lemma is the fact that for any quantum channel Φ, we have

$$H_{\min}(\Phi) = H_{\min}(\Phi^c). \tag{7.41}$$

Before continuing to the proof of Theorem 7.4.1, let us consider a toy example, where we put $k = m = 1$ and thus $l = 2$. In this case, it is not only possible to prove that $\Phi_1^{2,(1)}$ has constant output entropy on pure states, but it is relatively straight-forward to explicitly calculate these eigenvalues. We have the following:

Theorem 7.4.2

In the case of $k = 1$, $m = 1$ and $l = 2$, for any pure input $\rho = |\psi\rangle\langle\psi|$, the output $\Phi_1^{2,(1)}(\rho)$ has the eigenvalues:

- $\lambda = 0$ with multiplicity $N^2 - N - 1$.
- $\lambda = \frac{[2]_q}{[3]_q}\frac{N-1}{N} = \frac{N-1}{N^2-1}$ with unit multiplicity;
- $\lambda = \frac{[2]_q}{[3]_q} = \frac{N}{N^2-1}$ with multiplicity $N - 1$;

Proof. Let $\rho := |\psi\rangle\langle\psi|$ be a pure input state. Let us consider the output ρ_{out}, and for the sake of brevity we will leave out the global normalization constant $\frac{[2]_q}{[3]_q}$ until the very end. We then have

$$\rho_{\text{out}} = p_2(|\psi\rangle\langle\psi| \otimes \iota_1)p_2. \tag{7.42}$$

Note that in general, for a matrix pAp where p is an orthogonal projection and A is any matrix, the following holds: an eigenvector $|x\rangle$ of pA with eigenvalue $\lambda \neq 0$ is also an eigenvector of pAp with eigenvalue λ, as

$$p|x\rangle = \tfrac{1}{\lambda}ppA|x\rangle = \tfrac{1}{\lambda}pA|x\rangle = |x\rangle\,, \tag{7.43}$$

thus $|x\rangle$ lies in the image of p, and $pAp|x\rangle = pA|x\rangle = \lambda|x\rangle$. Conversely, an eigenvector $|y\rangle$ of pAp with eigenvalue $\mu \neq 0$ is also an eigenvector of pA with eigenvalue μ, as a similar calculation shows that $|y\rangle$ is in the image of p, whence $pA|y\rangle = pAp|y\rangle = \lambda|y\rangle$. Hence, we can restrict ourselves to finding the non-zero eigenvalues of $p_2(|\psi\rangle\langle\psi| \otimes \iota_1)$ instead.

In this particular case, we know $p_2 = \iota_2 - \frac{1}{N}T_\cap T_\cup$. We prefer to write $T_\cap = \sum_i |ii\rangle$ as $|\phi^+\rangle$. We claim it suffices to calculate the eigenvalues for just the input state $\rho = |1\rangle\langle 1|$, where $\{|i\rangle\}_{i=1}^N$ is the canonical orthonormal basis of $\mathcal{H}_1 = \mathbb{C}^N$. In that case, the eigenvalues can be exactly calculated, which we will do in Lemma 7.4.3. We are left with the claim that just considering $\rho = |1\rangle\langle 1|$ is sufficient. Firstly, note that the eigenvalues of ρ_{out} are unaffected by a unitary transformation of the form $V\rho_{\text{out}}V^*$. We then pick $V := U \otimes \overline{U}$ where $U := \sum_i |i\rangle\langle\psi_i|$, where $\{|\psi_i\rangle\}_{i=1}^N$ is any orthonormal basis of $\mathbb{C}^N$ with $|\psi\rangle = |\psi_1\rangle$. Here, $\overline{U} := \sum_i |i\rangle\langle\overline{\psi_i}|$ where $|\overline{\psi_i}\rangle$ is the unit vector one gets from conjugating all coefficients of the unit vector $|\psi_i\rangle$. Clearly, $V(|\psi\rangle\langle\psi| \otimes \iota_1)V^* = |1\rangle\langle 1| \otimes \iota_1$. By Lemma 7.4.2, $V|\phi^+\rangle = |\phi^+\rangle$, and we are done. Using the global normalization constant $\frac{[2]_q}{[3]_q}$ and the result of Lemma 7.4.3, the lemma now follows. □

Lemma 7.4.2

The generalized unnormalized Bell state $|\phi^+\rangle := \sum_i |ii\rangle$, where $\{|i\rangle\}_{i=1}^N$ is the canonical orthonormal basis of $\mathcal{H}_1 = \mathbb{C}^N$, is a fixed point of a basis transformation from a basis of the form $\{|\psi_i\rangle\,|\overline{\psi_j}\rangle\}_{i,j=1}^N$ to the canonical tensor product basis $\{|ij\rangle\}_{i,j=1}^N$, where $\{|\psi_i\rangle\}_{i=1}^N$ is

an orthonormal basis of $\mathbb{C}^N$, and $|\overline{\psi_i}\rangle$ denotes the unit vector one gets from conjugating all coefficients from the unit vector $|\psi_i\rangle$.

Proof. This follows directly from a small calculation: let $U := \sum_i |i\rangle\langle\psi_i|$ be the basis transformation, and put $\overline{U} := \sum_i |i\rangle\langle\overline{\psi_i}|$. We see $(U \otimes \overline{U})\,|\phi^+\rangle = \sum_{ijk} |ij\rangle\langle\psi_i\overline{\psi_j}|\,|kk\rangle = \sum_{ij} |ij\rangle \sum_k \langle\psi_i|k\rangle\,\langle\overline{\psi_j}|k\rangle$. Now use the fact that $\langle\overline{\psi_j}|k\rangle = \langle k|\psi_j\rangle$ to see $\sum_k \langle\psi_i|k\rangle\,\langle\overline{\psi_j}|k\rangle = \sum_k \langle\psi_i|k\rangle\,\langle k|\psi_j\rangle = \langle\psi_i|\psi_j\rangle = \delta_{ij}$. Hence, $(U \otimes \overline{U})\,|\phi^+\rangle = |\phi^+\rangle$. □

Lemma 7.4.3

The eigenvalues λ and eigenvectors $|x\rangle$ of

$$|1\rangle\langle 1| \otimes \iota_1 - \tfrac{1}{N}\,|\phi^+\rangle\langle\phi^+|\,(|1\rangle\langle 1| \otimes \iota_1) \tag{7.44}$$

are:

- $\lambda = 0$ with multiplicity $N^2 - N$, with orthogonal eigenvectors $\{|ij\rangle\}_{i\neq 1,j}^N$.
- $\lambda = 1$ with multiplicity $N - 1$, with orthogonal eigenvectors $\{|1j\rangle\}_{j=2}^N$.
- $\lambda = \frac{N-1}{N}$ with unit multiplicity, with eigenvector $|11\rangle + \frac{-1}{N-1}\sum_{i=2}^N |ii\rangle$.

Proof. Note that the last term $|\phi^+\rangle\langle\phi^+|\,(|1\rangle\langle 1|\otimes\iota_1)$ can be evaluated to $|\phi^+\rangle\langle 11|$. Hence, any basis vectors of the form $|ij\rangle$ with $i \neq 1$ is killed by both terms, yielding orthogonal eigenvectors for the eigenvalue 0. Similarly, any basis vectors of the form $|1j\rangle$ with $j \neq 1$ has eigenvalue 1 for the first term, and is killed by the last term, yielding $\lambda = 1$ precisely $N - 1$ times. Because $|1\rangle\langle 1| \otimes \iota_1 - \frac{1}{N}\,|\phi^+\rangle\langle 11|$ is not a normal matrix, it might not be diagonalizable, or if it is, it cannot be unitarily diagonalizable, so we look for an eigenvector that is not orthogonal to the previous ones. Heuristically, as the last term produces a vector in $\mathbb{C}\,|\phi^+\rangle$, it is prudent to start the search with a vector $|x\rangle = \sum_i \alpha_i\,|ii\rangle$. If we apply the first term, we are only left with $\alpha_1\,|11\rangle$, and the second term leaves $-\frac{1}{N}\alpha_1\,|\phi^+\rangle$. Hence, if λ is the corresponding eigenvalue, we must have $\alpha_1 - \frac{1}{N}\alpha_1 = \lambda\alpha_1$, i.e. $\lambda = \frac{N-1}{N}$. Then, the other constraints are $-\frac{1}{N}\alpha_1 = \lambda\alpha_j$ for $j > 1$, hence $\alpha_j = \frac{-1}{N-1}\alpha_1$. We have found N^2 linearly independent eigenvectors, so we can be certain we have found *all* eigenvalues. □

This toy example has given us some insight into the general case. Diagrammatically speaking, Lemma 7.4.2 tells us that any caps do not care if we rotate one strand with a unitary U, as long as we rotate the other with the unitary $\overline{U}$. This is actually a consequence of Lemma 7.2.2, where we showed that we can pull a partition through a cap or cup if we reflect the partition both horizontally and vertically. In our case, the unitary acts only on one "strand" (i.e., one copy of $\mathcal{H}_1$), so it is automatically invariant under the *horizontal flip*. Hence, we know that pulling U through a cap or cup yields the operator U^t (the transpose of U). Hence, if we want it to meet another operator, say V, on the other side of the cap or cup and annihilate with it, i.e. $VU^t = U^tV = \iota$, we must choose $V = \overline{U}$. However, in the general case, more than 2 strands are in play, and if we expand p_{m+1} as a linear combination of diagrams, we do not know which strand ρ is connected to in $p_{m+1}(\rho \otimes \iota_m)p_{m+1}$, and even worse, it is connected to different input and output strands for different diagrams in the expansion of p_{m+1}. Luckily, this is more tractable than it might appear on first sight, through Lemma 7.3.1 that tells us that ρ can't just be connected to *any* input and output: any pair in a NC_2-partition containing two upper indices or two lower indices contains one odd and one even index, and in the case of a pair of one upper index and one lower index, both are even, or both are odd. This now allows us to formulate a proof:

Proof of Theorem 7.4.1. We consider the quantum channel $\Phi_1^{m+1,(m)}$ first, the proof for $\Phi_1^{(m+1),m}$ then follows from Theorem 7.3.2. We see that the output density matrix is described (up to normalization) by $\rho_{\text{out}} = p_{m+1}(|\psi\rangle\langle\psi| \otimes \iota_m)p_{m+1}$. Consider the unitary $V := U \otimes \overline{U} \otimes U \otimes \cdots \otimes \overline{U}$ with $m + 1$ tensor legs (in the case where $m + 1$ is odd, the last tensor leg contains U instead

of $\overline{U}$), where we leave the unitary U undetermined for now. This unitary V lives in $\mathcal{B}(\mathcal{H}_1^{\otimes m+1})$. As $V^*\rho_{\text{out}}V$ has the same eigenvalues as ρ_{out}, we consider the density matrix

$$V^*\rho_{\text{out}}V = V^*p_{m+1}VV^*(|\psi\rangle\langle\psi| \otimes \iota_m)VV^*p_{m+1}V. \tag{7.45}$$

We claim that $V^*pV = p$ for *any* $p \in \text{NC}_2(m+1)$, which implies this equation also holds for the Jones-Wenzl projection. From Lemma 7.3.1 we know that if any block in p contains two upper or two lower indices, then one of those is even and one of them is odd. Hence, one is affected by U^* and the other by $\overline{U}^*$ if they form a cup, or U and $\overline{U}$ if they form a cap. However, by Lemma 7.4.2, we know that $(U \otimes \overline{U})T_\cap = T_\cap$, similarly $T_\cup(U^* \otimes \overline{U}^*) = T_\cup$; hence, pulling U (or U^*) through its cap (or cup) results in an annihilation against $\overline{U}$ (or $\overline{U}^*$) on the other side of the cap (or cup).

Similarly, if any block in p contains a lower and upper index, both are odd or both are even. Hence, the U^* (or $\overline{U}^*$) from the lower index meets and annihilates against the U (or $\overline{U}$) from the upper index. Thus, $V^*pV = p$.

But then, we can pick $U := \sum_i |i\rangle\langle\psi_i|$, where $\{|i\rangle\}_{i=1}^N$ is the canonical orthonormal basis of $\mathbb{C}^N$, and $\{|\psi_i\rangle\}_{i=1}^N$ is some other orthonormal basis with $|\psi\rangle = |\psi_1\rangle$. It is then evident that $V(|\psi\rangle\langle\psi| \otimes \iota_m)V^* = |1\rangle\langle 1| \otimes \iota_m$, thus

$$V\rho_{\text{out}}V^* = p_{m+1}(|1\rangle\langle 1| \otimes \iota_m)p_{m+1}, \tag{7.46}$$

and we see that ρ_{out} has the same eigenvalues for all pure input states $|\psi\rangle$. □

Note that analytically determining the *eigenvectors* of ρ_{out} for any value of $m \in \mathbb{N}$, i.e. a generalization of Theorem 7.4.2, is in all likelihood quite difficult. Namely, as already discussed in Theorem 7.4.2, it suffices to compute the eigenvectors with non-zero eigenvalue of $p_{m+1}(|1\rangle\langle 1| \otimes \iota_m)$. In turn, the eigenvalues of a product of projections are related to the eigenvalues of the *difference* of those two projections (see e.g. [AHT85]), however, this would require intimate knowledge of the eigenvectors of p_{m+1}. In turn, this is equivalent to precisely understanding the image of p_{m+1}, whilst finding direct formulas for the Jones-Wenzl projection (equivalently, direct formulas for its image) has proven to be difficult [BC16b].

We will, however, employ a handy trick that allows us to analytically express the eigenvalues of $\Phi_1^{(m+1),m}$ for *all* values of m in a recurrence relation. Let us first commence with an example that shows how involved a direct calculation with eigenvectors and eigenvalues already gets for $m = 2$:

Lemma 7.4.4

In the case of $k = 1$, $m = 2$ and $l = 3$, for any pure input $\rho = |\psi\rangle\langle\psi|$, the output $\Phi_1^{3,(2)}(\rho)$ has the eigenvalues:

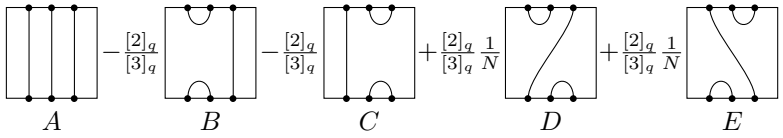
- 0 with multiplicity $N^3 - N^2 - 2N + 1$;
- $\frac{[2]_q}{[4]_q}\frac{N}{N+1} = \frac{1}{N^2-2}\frac{N}{N+1}$ with unit multiplicity;
- $\frac{[2]_q}{[4]_q}\frac{N^2-N-1}{N^2-1} = \frac{1}{N^2-2}\frac{N^2-N-1}{N^2-1}$ with multiplicity $N-1$;
- $\frac{[2]_q}{[4]_q} = \frac{1}{N^2-2}$ with multiplicity $(N-1)^2 + N - 2 = N^2 - N - 1$.

Proof. By the previous result, we know that it suffices to calculate this for $|\psi\rangle = |0\rangle$. We know from Lemma 7.2.5 that the action of $\Phi_1^{3,(2)}(\rho)$ is $p_3(\rho \otimes \iota_2)p_3$ up to normalization. Therefore, we will heuristically consider three *Ansatz* eigenvectors based on the symmetry of the Jones-Wenzl projections, and show that this collection of eigenvectors is exhaustive. Let us commence with the first *Ansatz*.

Ansatz 1. Assume that the eigenvector $|x\rangle$ corresponding to $\lambda \neq 0$ has the form $|x\rangle = \alpha|0\phi^+\rangle + \beta|\phi^+0\rangle + \gamma|000\rangle$ – this Ansatz comes from the observation that $|000\rangle$ is sent to a linear combination of $|000\rangle$, $|\phi^+0\rangle$ and $|0\phi^+\rangle$ by p_3, which we will make exact in the following

calculation. We claim that applying p_3 to $|x\rangle$ yields $|y\rangle := \gamma p_3 |000\rangle$. Note that $T_{u_1} |\phi^+ 0\rangle = |\phi^+ 0\rangle$ and $T_{u_2} |0\phi^+\rangle = |0\phi^+\rangle$, but $p_3 u_i = 0$ for $i = 1, 2$ by the defining property of the Jones-Wenzl projection, whence $p_3 |0\phi^+\rangle = p_3 |\phi^+ 0\rangle = 0$. Finding out what $\gamma p_3 |000\rangle$ is, requires a more detailed description of p_3.

As we know $p_2 = \iota_2 - \frac{1}{N} |\phi^+\rangle\langle\phi^+|$, from the Wenzl recursion formula we find

$p_3 =$ 

Using the fact $[2]_q = \dim \mathcal{H}_1 = N$ and $[3]_q = \dim \mathcal{H}_2 = N^2 - 1$, we see that $A = \iota_3$, $B = -\frac{N}{N^2-1} |\phi^+\rangle\langle\phi^+| \otimes \iota_1$, $C = -\frac{N}{N^2-1} \iota_1 \otimes |\phi^+\rangle\langle\phi^+|$, $D = \frac{1}{N^2-1} \sum_i |\phi^+ i\rangle\langle i\phi^+|$, and $E = \frac{1}{N^2-1} \sum_i |i\phi^+\rangle\langle\phi^+ i|$.

We calculate $|y\rangle = \gamma p_3 |000\rangle$ as:

$$\begin{aligned} A\gamma |000\rangle &= \gamma |000\rangle \quad , \quad B\gamma |000\rangle = -\tfrac{N}{N^2-1}\gamma |\phi^+ 0\rangle \quad , \quad C\gamma |000\rangle = -\tfrac{N}{N^2-1}\gamma |0\phi^+\rangle \\ D\gamma |000\rangle &= \tfrac{1}{N^2-1}\gamma |\phi^+ 0\rangle \quad , \quad E\gamma |000\rangle = \tfrac{1}{N^2-1}\gamma |0\phi^+\rangle . \end{aligned} \tag{7.47}$$

Noting that $\frac{N-1}{N^2-1} = \frac{1}{N+1}$, we find

$$|y\rangle = \gamma p_3 |000\rangle = \gamma |000\rangle - \gamma \tfrac{1}{N+1} |\phi^+ 0\rangle - \gamma \tfrac{1}{N+1} |0\phi^+\rangle . \tag{7.48}$$

Applying $|0\rangle\langle 0| \otimes \iota_2$ to this result yields

$$|z\rangle := (|0\rangle\langle 0| \otimes \iota_2) |y\rangle = \gamma \left(1 - \tfrac{1}{N+1}\right) |000\rangle + \gamma \tfrac{1}{N+1} |0\phi^+\rangle = \gamma \tfrac{N}{N+1} |000\rangle + \gamma \tfrac{1}{N+1} |0\phi^+\rangle . \tag{7.49}$$

We then need to apply p_3 again to the result $|z\rangle$, realizing that the $|0\phi^+\rangle$ term gets killed again, we can quickly conclude

$$p_3 |z\rangle = \gamma \tfrac{N}{N+1} p_3 |000\rangle = \gamma \tfrac{N}{N+1} \left(|000\rangle - \tfrac{1}{N+1} |\phi^+ 0\rangle - \tfrac{1}{N+1} |0\phi^+\rangle \right) . \tag{7.50}$$

Requiring $p_3 |z\rangle = \lambda |x\rangle = \lambda\alpha |0\phi^+\rangle + \lambda\beta |\phi^+ 0\rangle + \lambda\gamma |000\rangle$ we learn:

$$|000\rangle \quad : \quad \left(\tfrac{N}{N+1} - \tfrac{2N}{(N+1)^2}\right) \gamma = \lambda(\alpha + \beta + \gamma) \tag{7.51}$$

$$|0\phi^+\rangle \quad : \quad -\tfrac{N}{(N+1)^2}\gamma = \lambda\beta \tag{7.52}$$

$$|\phi^+ 0\rangle \quad : \quad -\tfrac{N}{(N+1)^2}\gamma = \lambda\alpha. \tag{7.53}$$

We required $\lambda \neq 0$, therefore we can use the second and third equation to learn that $\alpha = \beta = -\frac{N}{(N+1)^2}\frac{1}{\lambda}\gamma$. Plugging this into the first equation yields

$$\left(\tfrac{N}{N+1} - \tfrac{2N}{(N+1)^2}\right) \gamma = \left(\lambda - \tfrac{2N}{(N+1)^2}\right) \gamma \implies \gamma = 0 \vee \lambda = \tfrac{N}{N+1}. \tag{7.54}$$

The option $\gamma = 0$ leads to $\alpha = \beta = 0$, but this leads to $|x\rangle = 0$, and is thus excluded. We are left with the options where $\lambda = \frac{N}{N+1}$, γ is a free parameter, and $\alpha = \beta = -\frac{1}{N+1}\gamma$.

Ansatz 2. Similarly, we can use the *Ansatz* $|x\rangle = \alpha |i\phi^+\rangle + \beta |\phi^+ i\rangle + \gamma |00i\rangle$, where $1 \leqslant i \leqslant N-1$. Again, seeing the first two terms are in the image of u_1 and u_2, we conclude that they are killed by p_3, so $|y\rangle := p_3 |x\rangle = \gamma p_3 |00i\rangle$. This results in:

$$\begin{aligned} A\gamma |00i\rangle &= \gamma |00i\rangle \quad , \quad B\gamma |00i\rangle = -\tfrac{N}{N^2-1}\gamma |\phi^+ i\rangle \quad , \quad C\gamma |00i\rangle = 0. \\ D\gamma |00i\rangle &= 0 \quad , \quad E\gamma |00i\rangle = \tfrac{1}{N^2-1}\gamma |i\phi^+\rangle . \end{aligned} \tag{7.55}$$

Thus, we find

$$|y\rangle = \gamma p_3 |00i\rangle = \gamma |00i\rangle - \tfrac{N}{N^2-1}\gamma |\phi^+ i\rangle + \tfrac{1}{N^2-1}\gamma |i\phi^+\rangle . \tag{7.56}$$

Applying $|0\rangle\langle 0| \otimes \iota_2$ then yields

$$|z\rangle := (|0\rangle\langle 0| \otimes \iota_2)\,|y\rangle = \gamma(|0\rangle\langle 0| \otimes \iota_2)\left(|00i\rangle - \tfrac{N}{N^2-1}\,|\phi^+ i\rangle + \tfrac{1}{N^2-1}\,|i\phi^+\rangle\right) = \gamma\left(1 - \tfrac{N}{N^2-1}\right)|00i\rangle\,. \quad (7.57)$$

Applying p_3 to $|z\rangle$ yields:

$$p_3\,|z\rangle = \gamma\left(1 - \tfrac{N}{N^2-1}\right)\left(|00i\rangle - \tfrac{N}{N^2-1}\,|\phi^+ i\rangle + \tfrac{1}{N^2-1}\,|i\phi^+\rangle\right). \quad (7.58)$$

Requiring $p_3\,|z\rangle = \lambda\,|x\rangle$ yields:

$$|00i\rangle \quad : \quad \left(1 - \tfrac{N}{N^2-1}\right)^2 \gamma = \lambda(\beta + \gamma) \quad (7.59)$$

$$|i\phi^+\rangle \quad : \quad \left(1 - \tfrac{N}{N^2-1}\right)\tfrac{1}{N^2-1}\gamma = \lambda\alpha \quad (7.60)$$

$$|\phi^+ i\rangle \quad : \quad -\left(1 - \tfrac{N}{N^2-1}\right)\tfrac{N}{N^2-1}\gamma = \lambda\beta. \quad (7.61)$$

As we demand $\lambda \neq 0$, we express β in terms of γ from the last equation, and substitute this expression in the first equation to find:

$$\beta = -\left(1 - \tfrac{N}{N^2-1}\right)\tfrac{N}{N^2-1}\tfrac{1}{\lambda}\gamma \quad , \quad \left(1 - \tfrac{N}{N^2-1}\right)^2 \gamma = \left(\lambda - \left(1 - \tfrac{N}{N^2-1}\right)\tfrac{N}{N^2-1}\right)\gamma. \quad (7.62)$$

Excluding the option $\gamma = 0$ again (as this forces $\alpha = \beta = 0$ in the last two equations, forcing $|x\rangle = 0$), we see that we must have $\gamma = 1 - \frac{N}{N^2-1} = \frac{N^2-N-1}{N^2-1}$.

Ansatz 3. Furthermore, it is easy to see that $\{|0ij\rangle\}_{i\neq 0, j\neq i}$ is filled with orthogonal vectors who lie in the image of p_3 *and* in the image of $|0\rangle\langle 0| \otimes \iota_2$, hence those vectors are eigenvectors with eigenvalue 1. Following the same type of calculations as before, it is possible to prove that for $i \neq 0$, we have

$$p_3\,|0ii\rangle = |0ii\rangle - \tfrac{N}{N^2-1}\,|0\phi^+\rangle + \tfrac{1}{N^2-1}\,|\phi^+ 0\rangle\,. \quad (7.63)$$

Hence, $\{|0ii\rangle - |011\rangle\}_{i>1}$ contains $N-2$ linearly independent eigenvectors in the image of p_3, and they also lie in the image of $|0\rangle\langle 0| \otimes \iota_2$, so they are also eigenvectors with eigenvalue 1.

We claim this analysis fully describes all (non-zero) eigenvalues of $\Phi_1^{3,(2)}(|0\rangle\langle 0|)$. Let us count the eigenvalues:

- We have found $\lambda = \frac{N}{N+1}$ with unit multiplicity, with eigenvector $|x\rangle = \gamma(|000\rangle - \frac{1}{N+1}(|0\phi^+\rangle + |\phi^+ 0\rangle))$.
- we have found $\lambda = \frac{N^2-N-1}{N^2-1}$ with multiplicity $N-1$, with eigenvectors $|x_i\rangle = \gamma(|00i\rangle + \alpha\,|i\phi^+\rangle + \beta\,|\phi^+ i\rangle)$, where $\alpha, \beta \neq 0$ are uniquely determined by γ, and $i = 1, \ldots, N-1$. One can quickly check that these vectors are mutually orthogonal.
- We have found $\lambda = 1$ with multiplicity $(N-1)^2 + N - 2 = N^2 - N - 1$, with eigenvectors $|0ij\rangle$ for $i = 1, \ldots, N-1$ and $j \neq i$, and eigenvectors $|0ii\rangle - |011\rangle$ for $i > 1$. They are all linearly independent.

Summing over these eigenvalues with their respective multiplicities yields

$$\tfrac{N}{N+1} + \tfrac{(N-1)(N^2-N-1)}{(N-1)(N+1)} + (N-1)^2 + N - 2 = N^2 - 2, \quad (7.64)$$

and the normalization factor of the quantum channel $\Phi_1^{3,(2)}$ is $\frac{[2]_q}{[4]_q}$. Using $[2]_q = N$ and $[4]_q = N^3 - 2N$, we see that no more space is left for non-zero eigenvalues. Using this normalization factor to scale the eigenvalues of $p_3(|0\rangle\langle 0| \otimes \iota_2)p_3$ to eigenvalues of $\Phi_1^{3,(2)}$, the lemma follows. □

Considering the amount of work the previous lemma took, and considering the Ansatz-approach to finding eigenvectors, how could we hope to generalize this to $m \in \mathbb{N}$? The answer lies in a more careful consideration of $\Phi_1^{m+1,(m)}$, that we have so far ignored because Lemma 7.4.1 told us that the non-zero spectrum of $\Phi_1^{m+1,(m)}$ coincides with the non-zero spectrum of $\Phi_1^{(m+1),m}$. Apart from this lemma, Theorem 7.3.2 tells us the channels $\Phi_1^{(m+1),m}$ and $\Phi_1^{m+1,(m)}$ are strongly related in another sense:

Theorem 7.4.3

The eigenvalues μ_i (with repetition) of $\Phi_1^{(m+1),m}(\rho)$ correspond bijectively with the eigenvalues λ_i of $\Phi_1^{m,(m-1)}(\rho)$ through

$$\mu_i = \tfrac{[2]_q}{[m+2]_q}\mathrm{Tr}(\rho) - \tfrac{[m]_q}{[m+2]_q}\lambda_i. \tag{7.65}$$

Proof. We remind the reader of the statement of Theorem 7.3.2:

$$\Phi_1^{(m+1),m}(\rho) = T_{x_m}\left(\tfrac{[2]_q}{[m+2]_q}\mathrm{Tr}(\rho)p_m - \tfrac{[m]_q}{[m+2]_q}\Phi_1^{m,(m-1)}(\rho)\right)T_{x_m}. \tag{7.66}$$

From the definition of T_{x_m}, it is clear that $T_{x_m}^2 = \iota_m$ and $T_{x_m}^* = T_{x_m}$, so it is a self-adjoint unitary, thus the conjugation by T_{x_m} in the previous equation does not change the eigenvalues. As the image of $\Phi_1^{(m+1),m}$ lies in $\mathcal{B}(\mathcal{H}_m)$, the first term only shifts the value of any eigenvalue of $\Phi_1^{m,(m-1)}$ by $\frac{[2]_q}{[m+2]_q}\mathrm{Tr}(\rho)$, and we recover the last statement of the lemma. □

Theorem 7.4.4: Recurrence relation for eigenvalues of $\Phi_1^{m+1,(m)}$

Let $(\mu_i^{(m)})_{i=1}^{[m+2]_q}$ be the $[m+2]_q$-tuple with eigenvalues (with repetition) of $\Phi_1^{m+1,(m)}(\rho)$. All eigenvalues $\mu_i^{(m)}$ of $\Phi_1^{m+1,(m)}(\rho)$ for a pure input ρ can be computed through the recurrence relation

$$\mu_i^{(m)} = \tfrac{[2]_q}{[m+2]_q} - \tfrac{[m]_q}{[m+2]_q}\mu_i^{(m-1)}, \tag{7.67}$$

which yields $[m+1]_q$ eigenvalues (with repetition). The other $[m+2]_q - [m+1]_q$ eigenvalues $\mu_i^{(m)}$ are 0.

This recurrence relation starts with

$$\left(\mu_i^{(1)} \;:\; 1 \leqslant i \leqslant N^2-1\right) = \left(\tfrac{[2]_q}{[3]_q} \text{ for } 1 \leqslant i \leqslant N-1 \;;\; \tfrac{[2]_q}{[3]_q}\tfrac{N-1}{N} \;;\; 0 \text{ for } 1 \leqslant j \leqslant N^2-N-1\right). \tag{7.68}$$

Proof. The last statement for $\mu_i^{(1)}$ is Theorem 7.4.2. Using Theorem 7.4.3, we see that the eigenvalues $\lambda_i^{(m)}$ of $\Phi_1^{(m+1),m}$ can be found by

$$\lambda_i^{(m)} = \tfrac{[2]_q}{[m+2]_q} - \tfrac{[m]_q}{[m+2]_q}\mu_i^{(m-1)}. \tag{7.69}$$

By Lemma 7.4.1, the non-zero spectrum of $\Phi_1^{m+1,(m)}$ is precisely the non-zero spectrum of $\Phi_1^{(m+1),m}$, whence

$$\mu_i^{(m)} = \lambda_i^{(m)} = \tfrac{[2]_q}{[m+2]_q} - \tfrac{[m]_q}{[m+2]_q}\mu_i^{(m-1)}. \tag{7.70}$$

Finally, the spectrum of $\Phi_1^{m+1,(m)}$ must contain $[m+2]_q$ eigenvalues and we have computed all its non-zero eigenvalues, which yields the theorem. □

Remark 7.4.2. Note that we could have started with the channel $\Phi_1^{1,(0)}$, i.e. $m = 0$, in which case we see from Lemma 7.2.5 that $\Phi_1^{1,(0)}(\rho) = \frac{[2]_q}{[2]_q}p_1\rho p_1 = \rho$, so we get $(\mu_i^{(0)} \;:\; 1 \leqslant i \leqslant N) = (1\;;\; 0 \text{ for } 1 \leqslant j \leqslant N-1)$. And indeed, if we do the recurrence step once, the 1 gets transformed into $\frac{[2]_q}{[3]_q} - \frac{[1]_q}{[3]_q} = \frac{1}{N+1}$, whilst the 0's with multiplicity $N-1$ get transformed into $\frac{[2]_q}{[3]_q}$, which yields the correct eigenvalues according to Theorem 7.4.2. Then, applying the recurrence step again, we see

$$\begin{aligned}
\tfrac{1}{N+1} &\mapsto \tfrac{[2]_q}{[4]_q} - \tfrac{[2]_q}{[4]_q}\tfrac{1}{N+1} = \tfrac{[2]_q}{[4]_q}\tfrac{N}{N+1} && \text{with multiplicity } 1,\\
\tfrac{[2]_q}{[3]_q} &\mapsto \tfrac{[2]_q}{[4]_q} - \tfrac{[2]_q}{[4]_q}\tfrac{N}{N^2-1} = \tfrac{[2]_q}{[4]_q}\tfrac{N^2-N-1}{N^2-1} && \text{with multiplicity } N-1,\\
0 &\mapsto \tfrac{[2]_q}{[4]_q} && \text{with multiplicity } N^2-N-1.
\end{aligned}$$

This is precisely the result of Lemma 7.4.4.

Let us compare these findings to the lower bound of the minimum output entropy, given by $\log \frac{\theta_q(1,m+1,m)}{[2]_q}$, from Theorem 6.5.1. Furthermore, let us compare this with the function $N \mapsto m \log(N)$, which is a close approximation as $N^m[2]_q/\theta_q(1, m+1, m) = 1+\mathcal{O}(N^{-2})$ according to [BCLY20, Lem. 4.2]. For the choice $m = 4$, we record the results in Figure 7.1. We see that the numerical computations of $H_{\min}(\Phi_1^{(5),4})$ using the eigenvalues as described in Theorem 7.4.4 is in excellent agreement with the lower bound from Theorem 6.5.1 (which is asymptotically sharp for $N \to \infty$), and can already be approximated well by $m \log(N)$ even for small values of N.

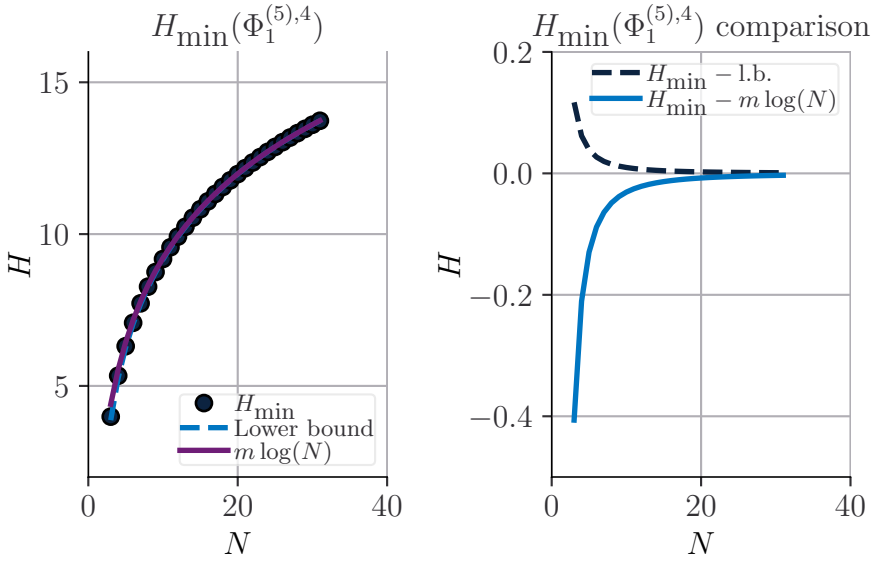


Figure 7.1: The minimum output entropy $H_{\min}$ of $\Phi_1^{(5),4}$ (so, $k = 1$, $l = 5$, $m = 4$), versus the lower bound (abbreviated **l.b.** in the right image) given by $N \mapsto \log \frac{\theta_q(1,5,4)}{[2]_q}$ from Theorem 6.5.1 and versus the function $N \mapsto m \log(N)$.

8

Numerical approximations for $H_{\min}$

When one considers the additivity conjecture for $H_{\min}$ as sketched in Section 3.4, one might wonder how difficult it would be to (numerically) approximate the quantities $H_{\min}(\mathcal{N}_1 \otimes \mathcal{N}_2)$ and $H_{\min}(\mathcal{N}_i)$, where $\mathcal{N}_i$ are quantum channels ($i = 1, 2$), to a sufficient degree to determine whether these channels satisfy the additivity conjecture. In general, calculating the minimum output entropy of a channel is a difficult task: Beigi and Shor showed in [BS07, Thm. 4.2] that the following problem is NP-complete: given a quantum channel $\mathcal{N}$ and a scalar $c > 0$, determine whether $H_{\min}(\mathcal{N}) < c$. They proved that this problem is a reduction from 2-Out-of-4-SAT, which is an NP-complete problem, hence the problem of determining whether $H_{\min}(\mathcal{N}) < c$ is itself an NP-hard problem. Furthermore, it is NP-complete (which are precisely those NP-hard problems that are themselves NP), because checking whether $H(\mathcal{N}(\rho)) < c$ for a given witness state ρ can be achieved in polynomial time.

We will therefore focus on the low-dimensional case, where we can use specific tools that make the computation of $H_{\min}$ tractable. To investigate whether two channels $\mathcal{N}_i$ (for $i = 1, 2$) violate the additivity conjecture, i.e. $H_{\min}(\mathcal{N}_1 \otimes \mathcal{N}_2) < H_{\min}(\mathcal{N}_1) + H_{\min}(\mathcal{N}_2)$, we need to find a good lower bound for the quantities $H_{\min}(\mathcal{N}_i)$, and a good upper bound for the quantity $H_{\min}(\mathcal{N}_1 \otimes \mathcal{N}_2)$. In this chapter, we discuss the theory for both: we use ϵ-covers to find lower bounds for $H_{\min}(\mathcal{N}_i)$, and we use heuristic optimization schemes based on particle swarm optimization to approximate $H_{\min}(\mathcal{N}_1 \otimes \mathcal{N}_2)$ from above. Let us commence with investigating the continuity of the von Neumann entropy, which is essential for the use of ϵ-covers.

8.1. Continuity bounds

It is well-known that, given a finite-dimensional Hilbert space $\mathcal{H}$, the von Neumann entropy $H : \mathcal{S}(\mathcal{H}) \to \mathbb{R}$ given by $H(\rho) = -\mathrm{Tr}(\rho \log \rho)$ is continuous. Hence, one can investigate whether evaluating $H(\Phi(\rho))$, where Φ is a quantum channel, for sufficiently many states ρ, together with the continuity of H, can yield a lower bound for the minimum output entropy $H_{\min}(\Phi)$.

Firstly, note that the von Neumann entropy is concave, i.e. $H(\sum_{i=1}^k p_i \rho_i) \geqslant \sum_{i=1}^k p_i H(\rho_i)$, where $k \in \mathbb{N}$, $\rho_i \in \mathcal{S}(\mathcal{H})$ and $p_i \geqslant 0$ such that $\sum_i p_i = 1$. Given a quantum channel $\Phi : \mathcal{S}(\mathcal{H}_1) \to \mathcal{S}(\mathcal{H}_2)$, we see that $H \circ \Phi : \mathcal{S}(\mathcal{H}_1) \to \mathbb{R}$ is also concave. Hence, as the space of states $\mathcal{S}(\mathcal{H}_1)$ is compact and convex, we know that the minimal output entropy of a quantum channel is always achieved on an extreme point of $\mathcal{S}(\mathcal{H}_1)$, which are the pure states on $\mathcal{H}_1$. We can thus restrict our attention to solving $H_{\min}(\Phi) = \min_{|\psi\rangle} H(\Phi(|\psi\rangle\langle\psi|))$ where the minimum is over all $|\psi\rangle \in \mathcal{H}_1$ with $\||\psi\rangle\| = 1$.

Apart from restricting our attention to the pure input states, we also need a continuity estimate of the output entropy. A famous result is the Fannes continuity inequality, which

reads:

$$|H(\rho) - H(\sigma)| \leqslant 2T \log_2(d) - 2T \log_2(2T) \quad \text{for } 0 \leqslant T \leqslant 1/2e, \tag{8.1}$$

where $T = T(\rho, \sigma) = \|\rho - \sigma\|_1/2$ is the normalized trace distance, where $\|\cdot\|_1$ denotes the trace norm $\|X\|_1 := \mathrm{Tr}\left(\sqrt{X^\dagger X}\right)$, and where $d = \dim \mathcal{H}$. For larger T, a weaker estimate holds [Aud07]. However, Audenaert provided a better bound that holds everywhere:

Theorem 8.1.1: Sharp continuity bound for H

[Aud07] The sharpest possible continuity bound for the von Neumann entropy, using only knowledge of the normalized trace distance T and the dimension d is:

$$|H(\rho) - H(\sigma)| \leqslant T \log_2(d-1) + H_2(T), \tag{8.2}$$

where H_2 is the binary entropy $H_2(p) = -p \log_2(p) - (1-p)\log_2(1-p)$ for $p \in [0,1]$, and $T = T(\rho,\sigma) = \|\rho - \sigma\|_1/2$, so $T \in [0,1]$.

We can combine the previous continuity result for H with the monotonicity property of the trace norm: it is well-known that we have $\|\Phi(\rho) - \Phi(\sigma)\|_1 \leqslant \|\rho - \sigma\|_1$, where $\Phi : \mathcal{S}(\mathcal{H}_1) \to \mathcal{S}(\mathcal{H}_2)$ is a quantum channel and $\rho, \sigma \in \mathcal{S}(\mathcal{H}_1)$, see for example [Wil13]. Let us now combine this with the fact that the minimum output entropy is achieved on a pure state by the concavity of the von Neumann entropy, we can write:

Proposition 8.1.1

Given a quantum channel $\Phi : \mathcal{S}(\mathcal{H}_1) \to \mathcal{S}(\mathcal{H}_2)$, and given two pure states $\rho := |\psi\rangle\langle\psi|$ and $\sigma := |\phi\rangle\langle\phi|$ with the assumption $|\langle\psi|\phi\rangle| \geqslant \frac{1}{2}\sqrt{3}$ (equivalently, $\|\sigma - \rho\|_1 \leqslant 1$). Let $d_2 := \dim \mathcal{H}_2$, then we have:

$$|H(\Phi(\rho)) - H(\Phi(\sigma))| \leqslant \sqrt{1 - |\langle\psi|\phi\rangle|^2}\log_2(d_2 - 1) + H_2(\sqrt{1 - |\langle\phi|\psi\rangle|^2}). \tag{8.3}$$

Proof. We can directly estimate using Audenaert's result:

$$|H(\Phi(\rho)) - H(\Phi(\sigma))| \leqslant \tfrac{1}{2}\|\Phi(\rho) - \Phi(\sigma)\|_1 \log_2(d_2 - 1) + H_2(\|\Phi(\rho) - \Phi(\sigma)\|_1/2). \tag{8.4}$$

Then, note that H_2 is a monotonically increasing function on the interval $[0, 1/2]$, so if we additionally assume $\|\Phi(\rho) - \Phi(\sigma)\|_1 \leqslant 1$, then using the monotonocity property of the trace norm:

$$\frac{1}{2}\|\Phi(\rho) - \Phi(\sigma)\|_1 \log_2(d_2 - 1) + H_2(\|\Phi(\rho) - \Phi(\sigma)\|_1/2) \leqslant \frac{1}{2}\|\rho - \sigma\|_1 \log_2(d_2 - 1) + H_2(\|\rho - \sigma\|_1/2). \tag{8.5}$$

Note that by the same monotonicity property of the trace norm, the previous inequality is in particular valid for $\|\rho - \sigma\|_1 \leqslant 1$. But this is precisely the assumption $|\langle\psi|\phi\rangle| \geqslant \frac{\sqrt{3}}{2}$, using the fact that for pure states the trace distance evaluates to: $\||\psi\rangle\langle\psi| - |\phi\rangle\langle\phi|\|_1 = 2\sqrt{1 - |\langle\phi|\psi\rangle|^2}$. □

Although the previous proposition gives a clear recipe for the continuity of the output entropy for a channel Φ, the trace norm $\|\cdot\|_1$ is a difficult norm to work with. Luckily, we can switch to the familiar Euclidean norm with a penalty of a factor 2:

Proposition 8.1.2

For pure states $\rho := |\psi\rangle\langle\psi|$ and $\sigma := |\phi\rangle\langle\phi|$ for $|\psi\rangle, |\phi\rangle \in \mathcal{H}$, we have:

$$\|\rho - \sigma\|_1 \leqslant 2\|\psi - \phi\|_2. \tag{8.6}$$

Proof. Note that $\|\phi-\psi\|_2^2 = \langle\psi-\phi|\psi-\phi\rangle = 2-2\Re\{\langle\psi|\phi\rangle\}$. We claim that $2-2\Re\{\langle\psi|\phi\rangle\} \geqslant 1-|\langle\phi|\psi\rangle|^2$. Then, using the expression $\|\rho-\sigma\|_1^2 = 4-4|\langle\phi|\psi\rangle|^2$, the proposition follows. To verify the claim, write $\langle\psi|\phi\rangle = a+bi$, then the claim translates to $2-2a \geqslant 1-(a^2+b^2) \iff a(a-2)+b^2 \geqslant -1$, but the right-hand side is always true as trivially $b^2 \geqslant 0$ and $a(a-2) \geqslant -1$ for all $a \in \mathbb{R}$. Hence, the claim follows. □

Remark 8.1.1. Note that the bound in Proposition 8.1.2 can be very loose for certain inputs, for example $|\psi\rangle = -|\phi\rangle$ yields $\|\rho-\sigma\|_1 = 0$, whilst $\|\psi-\phi\|_2 = 2$. However, the bound is the sharpest possible: say $|\psi\rangle = \varepsilon|\phi\rangle + \sqrt{1-\varepsilon^2}|\phi^\perp\rangle$ for some $|\phi^\perp\rangle$ that is perpendicular to $|\phi\rangle$ and some $\varepsilon \in [-1,1]$. One can calculate that

$$\|\psi-\phi\|_2 = \sqrt{2-2\Re\{\langle\phi|\psi\rangle\}} = \sqrt{2-2\varepsilon} \quad ; \quad \|\sigma-\rho\|_1 = \sqrt{4-4|\langle\phi|\psi\rangle|^2} = \sqrt{4-4\varepsilon^2}. \tag{8.7}$$

This yields

$$\frac{\|\sigma-\rho\|_1}{2\|\phi-\psi\|_2} = \frac{2\sqrt{1-\varepsilon^2}}{2\sqrt{2}\sqrt{1-\varepsilon}} = \frac{1}{\sqrt{2}}\sqrt{1+\varepsilon}. \tag{8.8}$$

This can be brought arbitrarily close to 1 by letting $\varepsilon \nearrow 1$. Note that the bound can never be achieved with equality for $|\psi\rangle \neq |\phi\rangle$. Namely, this would demand that $4\,(2-2\Re\{\langle\phi|\psi\rangle\}) = 4-4|\langle\psi|\phi\rangle|^2$, and setting $\langle\phi|\psi\rangle := a+bi$, this reduces to $8-8a = 4-4(a^2+b^2)$, which refactors to $(a-1)^2+b^2 = 0$, i.e. $a=1$ and $b=0$. However, this implies $|\psi\rangle = |\phi\rangle$, which is the trivial case. Hence, no choice of the states $|\psi\rangle \neq |\phi\rangle$ results in $\|\rho-\sigma\|_1 = 2\|\psi-\phi\|_2$.

The previous Proposition 8.1.2 allows us to recast Proposition 8.1.1 to:

Theorem 8.1.2

Given a quantum channel $\Phi : \mathcal{S}(\mathcal{H}_1) \to \mathcal{S}(\mathcal{H}_2)$, and given two pure states $\rho := |\psi\rangle\langle\psi|$ and $\sigma := |\phi\rangle\langle\phi|$ with the assumption $\|\psi-\phi\|_2 \leqslant \frac{1}{2}$. Let $d_2 := \dim \mathcal{H}_2$, then we have:

$$|H(\Phi(\rho)) - H(\Phi(\sigma))| \leqslant \|\psi-\phi\|_2 \log_2(d_2-1) + H_2(\|\psi-\phi\|_2), \tag{8.9}$$

Proof. By $T(\rho,\sigma) = \|\rho-\sigma\|_1/2 \leqslant \|\psi-\phi\|_2$ as seen in Proposition 8.1.2, the assumption $\|\psi-\phi\|_2 \leqslant \frac{1}{2}$ implies $\|\rho-\sigma\|_1 \leqslant 1$, which means that Proposition 8.1.1 holds. Again using the fact that H_2 monotonically increases on $[0,1/2]$, the proposition follows. □

We are now at a stage where we can start to investigate how the previous continuity results can be leveraged to numerically determine a lower bound for the minimum output entropy of a quantum channel Φ: we should sample enough unit vectors $\|\phi\|$ from $\mathcal{H}_1$, calculate the output entropy for these vectors, and estimate the maximal deviation of the output entropy between those "nodes" by using Theorem 8.1.2. This naturally leads us to the investigation of good choices of these "nodes" on the unit sphere of $\mathcal{H}_1$, which we will introduce as ε-covers in the following section.

8.2. ε-COVERS FOR THE UNIT SPHERE $\mathbb{S}^{n-1}$ IN $\mathbb{R}^n$

Ideally, one would like to sample the unit sphere in such a way that the sample points are as equally distributed as possible. Let us establish a few definitions from computational geometry to help us formulate this precisely.

8.2.1. DEFINITIONS AND PROPERTIES OF ε-COVERS

We will follow [HMS16] and [Cla06]. Let us assume we have a metric space (X,d). Given a subset $L \subseteq X$ and a point $x \in X$, the minimal distance of x to L is denoted by $d(x,L)$, i.e. $d(x,L) := \inf_{y\in L} d(x,y)$. Note that, if L is compact, this minimal distance is actually achieved at some point $y \in L$. Note that if L is finite, it is automatically compact.

Definition 8.2.1

Given a subset $L \subseteq X$ of a metric space (X, d), the *separation of* L is defined as the minimal distance between any two distinct points in L, i.e. $\mathrm{Sep}(L) := \min_{\xi \neq \eta \in L} d(\xi, \eta)$. We say that L is ε-*separated* if $\mathrm{Sep}(L) \geqslant \varepsilon$. In this case L is sometimes also called an ε-*packing* .

Definition 8.2.2

Given a subset $L \subseteq X$ of a compact metric space (X, d), the *covering radius of* L is defined as the maximal distance between a point in X and the set L, i.e. $\mathrm{Cov}(L) := \max_{\xi \in X} \min_{\eta \in L} d(\xi, \eta) = \max_{\xi \in X} d(\xi, L)$. We say that L is an ε-*cover* of X if $\mathrm{Cov}(L) \leqslant \varepsilon$.

Definition 8.2.3

A subset $L \subseteq X$ of a compact metric space (X, d) is called an $(\varepsilon_p, \varepsilon_c)$-*Delone set* if L is an ε_p-*packing* and an ε_c-*covering*. If $\varepsilon_p = \varepsilon_c = \varepsilon$, one often calls $(\varepsilon, \varepsilon)$-*Delone* sets also ε-*nets*. In literature, Delone sets are sometimes called *Delaunay* sets.

To summarize, ideally we would like to discretize the unit sphere $\mathbb{S}^{n-1}$ in $\mathbb{C}^n \simeq \mathcal{H}_1$, say by picking a finite set of points $A \subseteq \mathbb{S}^{n-1}$, such that A is an ε-covering of $\mathbb{S}^{n-1}$ for a given small $\varepsilon > 0$. Ideally, we would like A to be an $(\varepsilon, \varepsilon)$-Delone set, i.e. an ε-net, as this would also mean we are not over-sampling parts of the sphere (because ε-nets have to be ε-separated).

8.2.2. DIFFICULTY OF FINDING GOOD ε-COVERS

The problem of finding good distributions of points on the sphere has attracted a lot of scientific attention in computational geometry, combinatorial theory and data sciences, but also beyond those: it is related to biology, for example, through the *Tammes problem*, which is a packing problem that seeks to maximize the minimal distance between N points on a sphere, which is related to the distribution of pores on pollen grains. It is also related to physics, for example through the generalized *Thomson problem*, which seeks to minimize the Riesz s-energy $\sum_{1 \leqslant i < j \leqslant N} \|x_i - x_j\|^{-s}$ for some $s > 0$ and a finite set of variables $\{x_i\}_{i=1}^N \subseteq \mathbb{S}^{n-1}$. For $s = 1$, the Riesz s-energy is the Coulomb law, and thus Thomson's problem seeks to find the optimal placement of N electrons on a sphere such that their electric potential is minimized [SK97]. These problems are highly non-trivial: only in 2015 the Tammes problem was solved for $N = 14$ [MT15]. Our problem does not precisely deal with the packing problem, nor with the energy minimization problem, however, but with a variant of the *covering problem*, where we are given a precision ε, and the task is to find the minimal number of points N and their locations $\{x_i\}_{i=1}^N \subseteq \mathbb{S}^{n-1}$ such that their covering radius $\max_{y \in \mathbb{S}^{n-1}} \min_{1 \leqslant i \leqslant N} \|y - x_i\|_2$ is smaller than or equal to ε, which has also attracted research [Rog63, Don97, BW03, VG04, DPP06].

8.2.3. UPPER AND LOWER BOUNDS FOR THE SIZE OF ε-COVERS

Let us commence with some lower bounds on how large ε-coverings must be, and an upper bound on how large they have to be by considering one particular ε-covering. The following proposition is a small refinement of [HLSW04, Lemma II.4] using the same argument:

Proposition 8.2.1: Upper bound for the size of an ε-covering of $\mathbb{S}^{n-1}$

Given $n \in \mathbb{N}$ and $\varepsilon > 0$, an ε-cover A of $\mathbb{S}^{n-1}$ exists with size

$$|A| \leqslant U(n, \varepsilon) := \left(\tfrac{2}{\varepsilon} + 1\right)^n - \left(\tfrac{2}{\varepsilon} - 1\right)^n . \tag{8.10}$$

Proof. We claim that we can pick a maximal ε-separated subset A_ε of $\mathbb{S}^{n-1}$, i.e. no subset of $\mathbb{S}^{n-1}$ exists that properly contains A_ε and is itself also ε-separated. We apply Zorn's lemma:

consider the collection $\mathcal{P}$ of all subsets $A \subset \mathbb{S}^{n-1}$ that are ε-separated. Partially order $\mathcal{P}$ by set-inclusion. Consider a chain $\mathcal{T} \subseteq \mathcal{P}$, i.e. a totally ordered collection $\{A_i \ : \ i \in \mathcal{I}\}$ of ε-separated subsets of $\mathbb{S}^{n-1}$. We claim that this chain has an upper bound $A := \bigcup_{i\in\mathcal{I}} A_i \in \mathcal{P}$. Pick any $x, y \in A$. Then find $i, j \in \mathcal{I}$ such that $x \in A_i$ and $y \in A_j$. Let $k := \max(i, j)$, then, as $\mathcal{T}$ is totally ordered, $A_i, A_j \subseteq A_k$, so $x, y \in A_k$. But A_k is ε-separated, so $\|x - y\| \geqslant \varepsilon$. This shows that $A \in \mathcal{P}$. Now, by Zorn's lemma, $\mathcal{P}$ has at least one maximal element, say A_ε.

It is clear that A_ε is not only ε-separated, but also an ε-net: if not, one can find an $x_0 \in \mathbb{S}^{n-1}$ such that $\|x_0 - x\| \geqslant \varepsilon$ for all $x \in A_\varepsilon$, but then $A_\varepsilon \cup \{x_0\}$ is ε-separated and properly includes A_ε, which is a contradiction to the maximality of A_ε.

By the ε-separation of A_ε, we know that n-dimensional open balls $B(x, \varepsilon/2)$ of radius $\varepsilon/2$ and with centers $x \in A_\varepsilon$ are disjoint. They are always contained in the set $B(0, 1+\varepsilon/2) \setminus B(0, 1-\varepsilon/2)$. Hence, we have:

$$|A_\varepsilon| V(\varepsilon/2) \leqslant V(1 + \varepsilon/2) - V(1 - \varepsilon/2), \tag{8.11}$$

where $V(a)$ is the volume of an n-ball with radius $a > 0$. We have $V(a) = a^n V_n$, where V_n is the volume of the unit n-ball, which is well known: $V_n = \pi^{n/2}/\Gamma(n/2 + 1)$. Hence, we find:

$$|A_\varepsilon| \leqslant \frac{((1 + \varepsilon/2)^n - (1 - \varepsilon/2)^n) V_n}{(\varepsilon/2)^n V_n} = \left(\frac{2}{\varepsilon} + 1\right)^n - \left(\frac{2}{\varepsilon} - 1\right)^n =: U(n, \varepsilon). \tag{8.12}$$

□

Remark 8.2.1. Note that the size of the bound for $|A_\varepsilon|$ increases rapidly for fixed n and $\varepsilon \downarrow 0$: the leading term of $U(n, \varepsilon)$ is $2n(2/\varepsilon)^{n-1}$. For example, for $n = 5$ and $\varepsilon = 0.01$, the bound for $|A_{0.01}|$ evaluates to approximately $1.6 \cdot 10^{10}$. For $n = 15$ and $\varepsilon = 0.1$, the upper bound is approximately $5.3 \cdot 10^{19}$.

Remark 8.2.2. Note that other upper bounds exist in literature: Rogers established in [Rog63] that for $R > 1$ and $n \geqslant 9$, an n-sphere of radius R has a cover consisting of unit spheres, whose size is smaller than $cn \log(n) R^n$ if $R > n$ and $cn^{5/2} R^n$ when $R < n$, where c is some constant independent of n and R. See also [DPP06]. As [Don97] points out, the number of ε-spheres one needs to cover a ρ-sphere is equal to the number of 1-spheres one needs to cover a ρ/ε-sphere, so we let $\rho := 1$ and thus $R = 1/\varepsilon$, and we see that in our language, an ε-cover B_ε must exist such that

$$|B_\varepsilon| \leqslant \begin{cases} cn \log(n) \varepsilon^{-n} & \varepsilon < 1/n, \\ cn^{5/2} \varepsilon^{-n} & 1/n < \varepsilon < 1. \end{cases} \tag{8.13}$$

Other upper bounds in certain circumstances are also known, see for example [BW03, VG04].

Apart from an upper bound, we can also find lower bounds for the size of ϵ-covers:

Proposition 8.2.2: Lower bound for the size of an ε-covering of $\mathbb{S}^{n-1}$

Given $n \in \mathbb{N}$ and $\varepsilon > 0$, any ε-covering A of $\mathbb{S}^{n-1}$ must have a size of at least:

$$|A| \geqslant L(n, \varepsilon) := \sup_{\delta \in (0,1)} L(n, \varepsilon, \delta) \quad \text{where} \quad L(n, \varepsilon, \delta) := \frac{(1 + \delta)^n - (1 - \delta)^n}{(\varepsilon + \delta)^n}. \tag{8.14}$$

Proof. Given an ε-covering A of $\mathbb{S}^{n-1}$, we wish to use a volumetric argument again. We claim that A is also an $(\varepsilon+\delta)$-covering of $\Theta(\delta) := \overline{B}(0, 1+\delta) \setminus B(0, 1-\delta)$ for any $\delta \in (0, 1)$, where $\overline{B}(0, 1)$ is the closed unit n-ball (hence, $\Theta(\delta)$ is a spherical shell with inner radius $1 - \delta$ and outer radius $1 + \delta$). To see this, take any point $z \in \Theta(\delta)$, and put $y := z/\|z\|$ such that $y \in \mathbb{S}^{n-1}$. By the definition of $\Theta(\delta)$, $\|z - y\| \leqslant \delta$. Pick $x \in A$ such that $\|x - y\| \leqslant \varepsilon$. By the triangle inequality, it follows that $\|x - z\| \leqslant \varepsilon + \delta$, and hence A is an $(\varepsilon + \delta)$-covering of $\Theta(\delta)$.

But then, as $\Theta(\delta) \subseteq \bigcup_{x\in A} \overline{B}(x; \varepsilon + \delta)$, we also have

$$\operatorname{vol}(\Theta(\delta)) \leqslant \operatorname{vol}\left(\bigcup_{x\in A} \overline{B}(x; \varepsilon + \delta)\right) \leqslant \sum_{x\in A} \operatorname{vol}\left(\overline{B}(x; \varepsilon + \delta)\right) = |A|(\varepsilon + \delta)^n V_n. \tag{8.15}$$

The volume of $\Theta(\delta)$ can be evaluated as $((1+\delta)^n - (1-\delta)^n)V_n$, which yields:

$$|A| \geqslant \frac{(1+\delta)^n - (1-\delta)^n}{(\varepsilon+\delta)^n} =: L(n,\varepsilon,\delta) \tag{8.16}$$

As $\delta \in (0,1)$ was arbitrary, the proposition follows. □

Proposition 8.2.3

The lower bound $L_{n,\varepsilon}$ can be estimated with

$$L(n,\varepsilon) \geqslant L_1(n,\varepsilon) \quad \text{where} \quad L_1(n,\varepsilon) = L(n,\varepsilon,\varepsilon/(n-1)) = \frac{\left(1+\frac{\varepsilon}{n-1}\right)^n - \left(1-\frac{\varepsilon}{n-1}\right)^n}{\left(\frac{n}{n-1}\varepsilon\right)^n}, \tag{8.17}$$

and furthermore, the latter quantity can again be estimated by

$$L_1(n,\varepsilon) \geqslant L_2(n,\varepsilon) \geqslant L_3(n,\varepsilon) \quad \text{where} \quad L_2(n,\varepsilon) := 2(1-1/n)^{n-1}\varepsilon^{-n+1} \quad \text{and} \quad L_3(n,\varepsilon) := 2e^{-1}\varepsilon^{-n+1}. \tag{8.18}$$

Proof. The results in this proposition can be found by approximating the unique maximum of $\delta \mapsto L(n,\varepsilon,\delta)$. For a calculation, please see Section A.1. □

Remark 8.2.3. Other lower bounds are also known in literature: for example, Wyner established in [Wyn67] that "the minimum number of caps of half-angle θ required to cover the unit Euclidean n-sphere" is equal to $\exp(-n\log\sin\theta + o(n))$.

Remark 8.2.4. Let us investigate how fast this lower bound grows with n and ε:

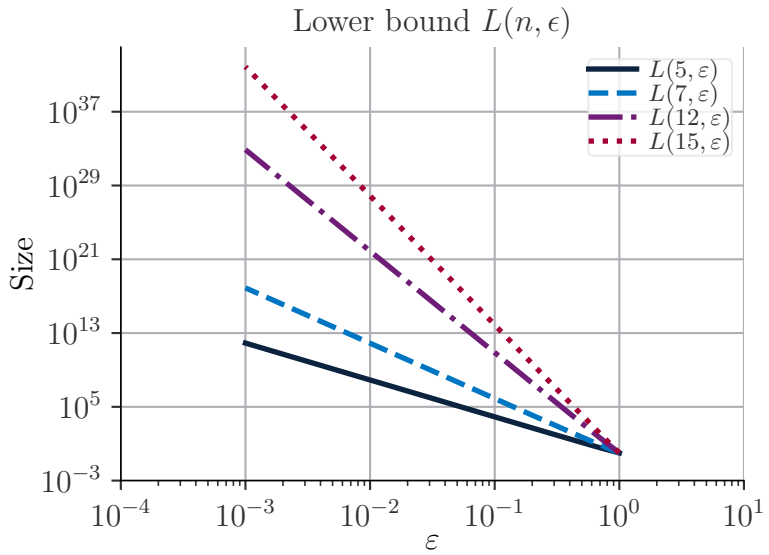


Figure 8.1: The lower bound $L(n,\varepsilon)$ for different values of n and ε.

Now that we have established some notion of the sizes of the ε-coverings involved, we can propose two naive ways of building ε-nets for $\mathbb{S}^{n-1}$, based on either discretizing the Cartesian coordinates, or discretizing the n-spherical coordinates.

8.2.4. ε-COVERING $A_{n,\varepsilon}^{(1)}$: DISCRETIZATION OF CARTESIAN COORDINATES

By discretizing each coordinate x_i for $1 \leqslant i \leqslant n$ separately over the interval $[-1,1]$, we can find an ε-covering of the unit cube $[-1,1]^n$. Using $\mathbb{S}^{n-1} \subseteq [-1,1]^n$, this directly yields an ε-covering $A_{n,\varepsilon}^{(1)}$ for $\mathbb{S}^{n-1}$, although we must stress that the vectors in $A_{n,\varepsilon}^{(1)}$ are generally *not* unit vectors! We discretize the coordinate x_i by setting $x_i^{(j)} := -1 + 2j/(t-1)$ for $1 \leqslant j \leqslant t$ for some natural number t. We choose a parameter $\delta > 0$ that depends on n and ε (we will later precisely specify this dependency), and then we choose t such that:

$$\left|x_i^{(j)} - x_i^{(j+1)}\right| = \frac{2}{t-1} \leqslant 2\delta \implies t \geqslant 1 + \frac{1}{\delta} \text{ so pick } t = 1 + \text{ceil}\left(\frac{1}{\delta}\right). \tag{8.19}$$

Then, we will define $A_{n,\varepsilon}^{(1)}$ as the collection of all vectors $\phi^{(i_1,\ldots,i_n)}$ with components $\phi_m^{(i_1,\ldots,i_n)} = x_m^{(i_m)}$, for any tuple $(i_1,\ldots,i_n)$ with $1 \leqslant i_j \leqslant t$ for all $1 \leqslant j \leqslant n$.

Theorem 8.2.1: $A_{n,\varepsilon}^{(1)}$ is an ε-covering

Given $\varepsilon > 0$ and $n \in \mathbb{N}$, pick $\delta \leqslant \varepsilon/\sqrt{n}$. Construct the set $A_{n,\varepsilon}^{(1)}$ containing all $\phi^{(i_1,\ldots,i_n)}$ for which its m'th coordinate satisfies $\phi_m = x_m^{(i_m)}$. Then, we have for any unit vector $\psi \in \mathbb{S}^{n-1}$:

$$\min_{\phi \in A} \left\|\psi - \phi^{(i_1,\ldots,i_n)}\right\|_2 \leqslant \varepsilon. \tag{8.20}$$

Proof. Given a unit vector ψ, denote with ψ_m the m-th coordinate. For each m, pick an $x_m^{(i_m)}$ such that $|\psi_m - x_m^{(i_m)}| \leqslant \delta$, which is possible as the x_i are distributed uniformly over $[-1,1]$ with spacing 2δ, and $\psi_m \in [-1,1]$, so it is at worst precisely halfway between two x_i. Compose the vector $\phi^{(i_1,\ldots,i_n)} \in A$ consisting of the coefficients $x_m^{(i_m)}$, then clearly $\|\psi - \phi^{(i_1,\ldots,i_n)}\|_2^2 = \sum_{m=1}^n \left|\psi_m - x_m^{(i_m)}\right|^2 \leqslant \sum_{m=1}^n \delta^2 = n\delta^2$. Hence, if we pick $\delta\sqrt{n} \leqslant \varepsilon$, the theorem follows. □

Proposition 8.2.4: The size of $A_{n,\varepsilon}^{(1)}$

Given $n \in \mathbb{N}$ and $\varepsilon > 0$, the ε-net $A_{n,\varepsilon}^{(1)}$ has size $|A_{n,\varepsilon}^{(1)}| = \left(1 + \text{ceil}\left(\frac{\sqrt{n}}{\varepsilon}\right)\right)^n$.

Proof. This follows directly from $t = 1 + \text{ceil}(1/\delta)$ and $\delta = \varepsilon/\sqrt{n}$. □

Remark 8.2.5. For $n = 5$ and $\varepsilon = 0.01$, we have $\delta \approx 0.00447$ and $|A_{n,\varepsilon}^{(1)}| \approx 5.8 \cdot 10^{11}$. We compare this to the lower bound $L_{5,0.01} \approx 8.2 \cdot 10^7$ and upper bound (for the maximal ε-nets, whilst $A_{n,\varepsilon}^{(1)}$ is δ-separated and not ε-separated!) $U_{5,0.01} \approx 1.6 \cdot 10^{10}$, and see that smallers ε-coverings than $A_{n,\varepsilon}^{(1)}$ must exist. For $n = 15$ and $\varepsilon = 0.1$, we have $\delta \approx 0.0258$ and $|A_{n,\varepsilon}^{(1)}| \approx 1.07 \cdot 10^{24}$. We again compare this to the lower bound $L_{15,0.1} \approx 7.6 \cdot 10^{13}$ and upper bound $U_{15,0.1} \approx 5.3 \cdot 10^{19}$.

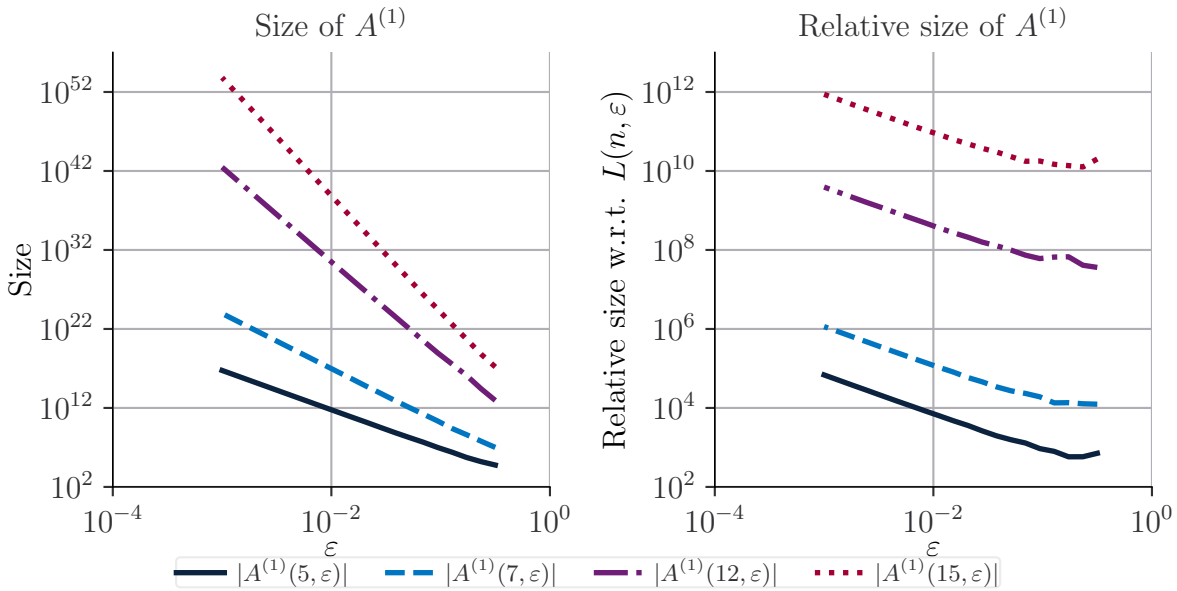


Figure 8.2: **Left**: The size $|A^{(1)}|$ of ε-cover $A^{(1)}$ for different values of n and ε. **Right**: The relative size $|A^{(1)}|/L(n,\varepsilon)$ for different values of n and ε.

8.2.5. ε-COVERING $A^{(2)}_{n,\varepsilon}$: DISCRETIZATION OF n-SPHERICAL COORDINATES

Two large drawbacks occur in the previous construction of the ε-covering: the size of the covering $A^{(1)}_{n,\varepsilon}$ is suboptimal (i.e., smaller ε-coverings must exist), and the vectors in $A^{(1)}_{n,\varepsilon}$ are not actually unit vectors. We can mitigate the second challenge by discretizing *angles*, instead of coordinates. We use generalized n-spherical coordinates, which look like:

$$\begin{aligned} x_1 &= \cos(\alpha_1) \\ x_2 &= \sin(\alpha_1)\cos(\alpha_2) \\ x_3 &= \sin(\alpha_1)\sin(\alpha_2)\cos(\alpha_3) \\ &\vdots \\ x_{n-1} &= \sin(\alpha_1)\cdots\sin(\alpha_{n-2})\cos(\alpha_{n-1}) \\ x_n &= \sin(\alpha_1)\cdots\sin(\alpha_{n-2})\sin(\alpha_{n-1}). \\ 0 \leqslant \alpha_1, \ldots, \alpha_{n-2} \leqslant \pi \quad &, \quad 0 \leqslant \alpha_{n-1} < 2\pi. \end{aligned} \tag{8.21}$$

We discretize $\alpha^{(j)}_i = \pi j/(t_1 - 1)$ for some $t_1 \in \mathbb{N}$ that we will specify later, and $0 \leqslant j \leqslant t_1 - 1$ and $1 \leqslant i \leqslant n-2$. Furthermore, we discretize $\alpha^{(j)}_{n-1} = 2\pi(j-1)/t_2$ for some $t_2 \in \mathbb{N}$ that we will specify later, and $1 \leqslant j \leqslant t_2$, such that $\alpha^{(1)}_{n-1} = 0$ and $\alpha^{(t_2)}_{n-1} = 2\pi(t_2 - 1)/t_2$, as α_{n-1} should be sampled on the half-open interval $[0, 2\pi)$. Let us introduce a parameter $\delta > 0$ that depends on ε and n (we will later precisely specify this dependency), and then pick t_1 such that for any $1 \leqslant i \leqslant n-2$:

$$|\alpha^{(j)}_i - \alpha^{(j+1)}_i| = \pi/(t_1 - 1) < 2\delta \implies t_1 > 1 + \frac{\pi}{2\delta} \text{ so pick } t_1 = 1 + \text{ceil}\left(\frac{\pi}{2\delta}\right). \tag{8.22}$$

Furthermore, let us pick t_2 similarly for α_{n-1}, but now for a range 2π, which yields:

$$|\alpha^{(j)}_{n-1} - \alpha^{(j+1)}_{n-1}| = \frac{2\pi}{t_2} \leqslant 2\delta \implies t_2 \geqslant \frac{2\pi}{2\delta} \text{ so pick } t_2 = \text{ceil}\left(\frac{\pi}{\delta}\right). \tag{8.23}$$

Then, let $A^{(2)}_{n,\varepsilon}$ be the collection of all unit vectors $\phi^{(i_1,\ldots,i_{n-1})}$ whose n-spherical angles $(\alpha_1, \ldots, \alpha_{n-1})$ as described in Equation (8.21) are given by $\alpha_j = \alpha^{(i_j)}_i$ for $1 \leqslant j \leqslant n-1$ as described by Equation (8.21), and $\alpha_{n-1} = \alpha^{(i_{n-1})}_{n-1}$ as described by Equation (8.23). Before we prove that $A^{(2)}_{n,\varepsilon}$ is an ε-cover, we need to pick δ appropriately. We need to be able to bound the distance between two

unit vectors ϕ and ψ with small differences in each n-spherical angle. Firstly, note that if two angles α and β satisfy $|\alpha-\beta| \leqslant \delta$, then by the Mean Value Theorem we have $|\sin(\alpha)-\sin(\beta)| \leqslant \delta$ and $|\cos(\alpha)-\cos(\beta)| \leqslant \delta$. We identify multiple methods to bound the distance between two unit vectors ϕ and ψ with n-spherical angles $(\alpha_1, \ldots, \alpha_{n-1})$ and $(\beta_1, \ldots, \beta_{n-1})$, respectively, where $|\alpha_i - \beta_i|$ is small:

1. Bounding the terms $|\phi_m - \psi_m|$ for $1 \leqslant m \leqslant n$ by using a telescoping sum. As the components are of the form $\sin(\alpha_1)\cdots\sin(\alpha_{m-1})\cos(\alpha_m)$, this can be achieved by using the identity

$$b_1 \cdots b_m - a_1 \cdots a_m = \sum_{k=1}^{m} a_1 \cdots a_{m-k}(b_{m-k+1} - a_{m-k+1})b_{m-k+2} \cdots b_m, \tag{8.24}$$

   in combination with the triangle inequality. Further using that $|\sin(\alpha) - \sin(\beta)| \leqslant \delta$ and $|\cos(\alpha) - \cos(\beta)| \leqslant \delta$ if $|\alpha - \beta| \leqslant \delta$, this yields an estimate of the form $|\phi_m - \psi_m|^2 \leqslant m^2\delta^2$. Summing over these terms, and taking the square root, results in the rule $\delta\sqrt{n(n+1)(2n+1)/6} \leqslant \varepsilon$. The resulting bound is quite loose due to the overuse of the triangle inequality.
2. Bounding the terms $|\phi_m - \psi_m|$ for $1 \leqslant m \leqslant n$ by using a multivariate first-order Taylor approximation of functions f_m given by $f_m(\alpha_1, \ldots, \alpha_m) = \sin(\alpha_1)\cdots\sin(\alpha_{m-1})\cos(\alpha_m)$, where we identify $\phi_m = f_m(\alpha_1, \ldots, \alpha_m)$ and $\psi_m = f_m(\beta_1, \ldots, \beta_m)$. The Lagrange rest term can be estimated as $\delta^2 m^2/2$, and the size of the first order derivatives can be recast into an optimization problem with a unique solution. This yields the estimate

$$|\psi_m - \phi_m| \leqslant \delta\,(1 - 1/m)^{m/2} \cdot \left(\sqrt{m-1} + 1/\sqrt{m-1}\right) + \delta^2 m^2/2, \tag{8.25}$$

   which yields the rule $\delta g(n) \leqslant \varepsilon$, where g scales slightly worse than $\mathcal{O}(n)$. See Lemma A.1.1. The resulting bound is imperfect, because the size of the first order derivatives in the Taylor expansion is bounded individually for each component $|\phi_m - \psi_m|$, instead of a "global" bound for the squared sum of all components.
3. Estimating $\|\phi - \psi\|^2$ by first rotating ψ and ϕ such that either ψ or ϕ is rotated to the vector $(1, 0, \ldots, 0)$, and then combining the components of the other vector using the sum and difference formulae for cosines and sines. This yields the rule $\delta\sqrt{n-1} \leqslant \varepsilon$.

The third method leads to the best bound, and will therefore be fully explained in the following subsection.

ESTIMATING $\|\phi - \psi\|$ IN n-SPHERICAL COORDINATES USING ROTATIONS

Given a unit vector ψ, we pick the best approximation $\phi^{(i_1,\ldots,i_{n-1})}$ that is in the ε-net $A^{(2)}_{n,\varepsilon}$. Then, we rotate ψ back to the "north pole" $(1, 0, \ldots, 0)$, and look at the angles corresponding to the rotated version of $\phi^{(i_1,\ldots,i_{n-1})}$. Let us introduce notation.

Definition 8.2.4: 2D rotation

In two dimensions, as rotation matrix $R(\alpha)$ rotates a vector counterclockwise over an angle α. The matrix $R(\alpha)$ is given by $R(\alpha) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$.

Definition 8.2.5: Embedding of 2D rotation in n dimensions

In n dimensions, we can "embed" a two-dimensional rotation as a rotation of 2 canonical axes around the previous axis as:

$$R_k(\alpha) := \mathbf{1}_{k-1} \oplus R(\alpha) \oplus \mathbf{1}_{n-k-1}, \tag{8.26}$$

where $\mathbf{1}_m$ denotes the $m \times m$ identity matrix. Note that $R_k(\alpha)$ is a square $n \times n$ matrix.

Remark 8.2.6. Note that the matrix $R_k(\alpha)$ are indeed rotations in n dimensions:

$$R_k(\alpha)^\top = \mathbf{1}_{k-1} \oplus R(\alpha)^\top \oplus \mathbf{1}_{n-k-1} = \mathbf{1}_{k-1} \oplus R(\alpha)^{-1} \oplus \mathbf{1}_{n-k-1} = R_k(\alpha)^{-1}, \tag{8.27}$$

and $\det R_k(\alpha) = \det R(\alpha) = 1$, hence $R_k(\alpha) \in \mathrm{SO}(n)$.

Lemma 8.2.1: Approximation lemma

Given a unit vector ϕ with n-spherical angles $(\alpha_1, \dots, \alpha_{n-1})$, and a unit vector ψ with n-spherical angles $(\beta_1, \dots, \beta_{n-1})$. Assume that $|\beta_i - \alpha_i| \leqslant \delta$ for all $1 \leqslant i \leqslant n-1$, then we have:

$$\|\psi - \phi\|_2^2 \leqslant (n-1)\delta^2. \tag{8.28}$$

Proof. Rotate ϕ by $R_{n-1}(-\alpha_{n-1})$, and call the result $\phi^{(1)}$. This only affects the coordinates ϕ_n and ϕ_{n-1}, which become:

$$\begin{aligned}\begin{pmatrix}\phi_{n-1}^{(1)}\\ \phi_n^{(1)}\end{pmatrix} &= R(-\alpha_{n-1})\begin{pmatrix}\phi_{n-1}\\ \phi_n\end{pmatrix} = \begin{pmatrix}\cos\alpha_{n-1} & \sin\alpha_{n-1}\\ -\sin\alpha_{n-1} & \cos\alpha_{n-1}\end{pmatrix} s_{n-2}(\boldsymbol{\alpha})\begin{pmatrix}\cos\alpha_{n-1}\\ \sin\alpha_{n-1}\end{pmatrix}\\ &= s_{n-2}(\boldsymbol{\alpha})\begin{pmatrix}\cos^2\alpha_{n-1} + \sin^2\alpha_{n-1}\\ -\cos\alpha_{n-1}\sin\alpha_{n-1} + \sin\alpha_{n-1}\cos\alpha_{n-1}\end{pmatrix} = s_{n-2}(\boldsymbol{\alpha})\begin{pmatrix}1\\ 0\end{pmatrix},\end{aligned} \tag{8.29}$$

where we use the shorthand notation $s_k(\boldsymbol{\alpha}) := \sin(\alpha_1)\cdots\sin(\alpha_k)$ for $k \in \mathbb{N}$. Hence, precisely by construction, we rotated ϕ_n to 0. Then, rotate $\phi^{(1)}$ by $R_{n-2}(-\alpha_{n-2})$, and call the result $\phi^{(2)}$. Clearly, this only affects $\phi_{n-2}^{(1)}$ and $\phi_{n-1}^{(1)}$. We can again write down:

$$\begin{aligned}\begin{pmatrix}\phi_{n-2}^{(2)}\\ \phi_{n-1}^{(2)}\end{pmatrix} &= R(-\alpha_{n-2})\begin{pmatrix}\phi_{n-2}^{(1)}\\ \phi_{n-1}^{(1)}\end{pmatrix} = \begin{pmatrix}\cos\alpha_{n-2} & \sin\alpha_{n-2}\\ -\sin\alpha_{n-2} & \cos\alpha_{n-2}\end{pmatrix} s_{n-3}(\boldsymbol{\alpha})\begin{pmatrix}\cos\alpha_{n-2}\\ \sin\alpha_{n-2}\end{pmatrix}\\ &= s_{n-3}(\boldsymbol{\alpha})\begin{pmatrix}\cos^2\alpha_{n-2} + \sin^2\alpha_{n-2}\\ -\cos\alpha_{n-2}\sin\alpha_{n-2} + \sin\alpha_{n-2}\cos\alpha_{n-2}\end{pmatrix} = s_{n-3}(\boldsymbol{\alpha})\begin{pmatrix}1\\ 0\end{pmatrix}.\end{aligned} \tag{8.30}$$

So, by the same argument, we're rotating $\phi_{n-1}^{(1)}$ to 0. Continue rotating each $\phi^{(k)}$ until we reach $\phi^{(n-2)}$, which we rotate by $R_1(-\alpha_1)$ to get $\phi^{(n-1)} = (1, 0, \dots, 0)$. We now need to investigate what happens to ψ with the same rotations. Let us consider the first rotation $R_{n-1}(-\alpha_{n-1})$:

$$\begin{pmatrix}\psi_{n-1}^{(1)}\\ \psi_n^{(1)}\end{pmatrix} = \begin{pmatrix}\cos\alpha_{n-1} & \sin\alpha_{n-1}\\ -\sin\alpha_{n-1} & \cos\alpha_{n-1}\end{pmatrix}\begin{pmatrix}\psi_{n-1}\\ \psi_n\end{pmatrix} = \begin{pmatrix}\cos\alpha_{n-1} & \sin\alpha_{n-1}\\ -\sin\alpha_{n-1} & \cos\alpha_{n-1}\end{pmatrix} s_{n-2}(\boldsymbol{\beta})\begin{pmatrix}\cos\beta_{n-1}\\ \sin\beta_{n-1}\end{pmatrix} \tag{8.31}$$

$$= s_{n-2}(\boldsymbol{\beta})\begin{pmatrix}\cos\alpha_{n-1}\cos\beta_{n-1} + \sin\alpha_{n-1}\sin\beta_{n-1}\\ -\sin\alpha_{n-1}\cos\beta_{n-1} + \cos\alpha_{n-1}\sin\beta_{n-1}\end{pmatrix} = s_{n-2}(\boldsymbol{\beta})\begin{pmatrix}\cos(\beta_{n-1} - \alpha_{n-1})\\ \sin(\beta_{n-1} - \alpha_{n-1})\end{pmatrix}. \tag{8.32}$$

The term ψ_n is not changed by further rotations, but the coordinate ψ_{n-1} is changed by the next rotation. We calculate:

$$\begin{pmatrix}\psi_{n-2}^{(2)}\\ \psi_{n-1}^{(2)}\end{pmatrix} = \begin{pmatrix}\cos\alpha_{n-2} & \sin\alpha_{n-2}\\ -\sin\alpha_{n-2} & \cos\alpha_{n-2}\end{pmatrix} s_{n-3}(\boldsymbol{\beta})\begin{pmatrix}\cos\beta_{n-2}\\ \sin\beta_{n-2}\cos(\beta_{n-1} - \alpha_{n-1})\end{pmatrix} \tag{8.33}$$

$$= s_{n-3}(\boldsymbol{\beta})\begin{pmatrix}\cos\alpha_{n-2}\cos\beta_{n-2} + \sin\alpha_{n-2}\sin\beta_{n-2}\cos(\beta_{n-1} - \alpha_{n-1})\\ -\sin\alpha_{n-2}\cos\beta_{n-2} + \cos\alpha_{n-2}\sin\beta_{n-2}\cos(\beta_{n-1} - \alpha_{n-1})\end{pmatrix}. \tag{8.34}$$

We can still approximate the angles in the last expression. We abbreviate $c_k^\alpha := \cos(\alpha_k)$, $c_k^\beta := \cos(\beta_k)$ and similarly $s_k^\alpha := \sin(\alpha_k)$ and $s_k^\beta := \sin(\beta_k)$, and $c_k^{\beta-\alpha} := \cos(\beta_k - \alpha_k)$ and we see:

$$\begin{pmatrix}\psi_{n-2}^{(2)}\\ \psi_{n-1}^{(2)}\end{pmatrix} = s_{n-3}(\boldsymbol{\beta})\begin{pmatrix}c_{n-2}^\alpha c_{n-2}^\beta + s_{n-2}^\alpha s_{n-2}^\beta + s_{n-2}^\alpha s_{n-2}^\beta\left(c_{n-1}^{\beta-\alpha} - 1\right)\\ -s_{n-2}^\alpha c_{n-2}^\beta + c_{n-2}^\alpha s_{n-2}^\beta + c_{n-2}^\alpha s_{n-2}^\beta\left(c_{n-1}^{\beta-\alpha} - 1\right)\end{pmatrix} \tag{8.35}$$

$$= s_{n-3}(\boldsymbol{\beta})\begin{pmatrix}c_{n-2}^{\beta-\alpha} + s_{n-2}^\alpha s_{n-2}^\beta\left(c_{n-1}^{\beta-\alpha} - 1\right)\\ s_{n-2}^{\beta-\alpha} + c_{n-2}^\alpha s_{n-2}^\beta\left(c_{n-1}^{\beta-\alpha} - 1\right)\end{pmatrix}. \tag{8.36}$$

In general, we see that for a parameter q_k we have

$$\begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} \cos\beta \\ \sin(\beta)\cdot q_k \end{pmatrix} = \begin{pmatrix} \cos\alpha\cos\beta + \sin\alpha\sin\beta + \sin\alpha\sin(\beta)\,(q_k-1) \\ -\sin\alpha\cos\beta + \cos\alpha\sin\beta + \cos\alpha\sin(\beta)\,(q_k-1) \end{pmatrix} \tag{8.37}$$

$$= \begin{pmatrix} \cos(\beta-\alpha) + \sin\alpha\sin(\beta)\,(q_k-1) \\ \sin(\beta-\alpha) + \cos\alpha\sin(\beta)\,(q_k-1) \end{pmatrix}. \tag{8.38}$$

Hence, each rotation, the parameter q_k is transformed as $q_k \mapsto \cos(\beta-\alpha)+\sin(\alpha)\sin(\beta)(q_k-1) := q_{k+1}$, as only the topmost element of the 2-dimensional vector is relevant in the next rotation. Starting from $t_0 = 1$ before the first rotation, we see

$$q_0 = 1 \tag{8.39}$$

$$q_1 = c_{n-1}^{\beta-\alpha} \tag{8.40}$$

$$q_2 = c_{n-2}^{\beta-\alpha} + s_{n-2}^{\alpha} s_{n-2}^{\beta}\left(c_{n-1}^{\beta-\alpha} - 1\right) \tag{8.41}$$

$$q_3 = c_{n-3}^{\beta-\alpha} + s_{n-3}^{\alpha} s_{n-3}^{\beta}\left(c_{n-2}^{\beta-\alpha} - 1 + s_{n-2}^{\alpha} s_{n-2}^{\beta}\left(c_{n-1}^{\beta-\alpha} - 1\right)\right). \tag{8.42}$$

We continue this pattern until we have rotated ψ the same as ϕ, and we see that the first coordinate $\psi_1^{(n-1)}$ becomes:

$$\psi_1^{(n-1)} = q_{n-1} = c_1^{\beta-\alpha} + s_1^{\alpha} s_1^{\beta}\left(c_2^{\beta-\alpha} - 1\right) + s_1^{\alpha} s_1^{\beta} s_2^{\alpha} s_2^{\beta}\left(c_3^{\beta-\alpha} - 1\right) + \cdots + s_1^{\alpha} s_1^{\beta} \cdots s_{n-2}^{\alpha} s_{n-2}^{\beta}\left(c_{n-1}^{\beta-\alpha} - 1\right). \tag{8.43}$$

Hence, using the fact that we have

$$\|\phi-\psi\|_2^2 = \langle \psi-\phi|\psi-\phi\rangle = 2 - 2\Re\left\{\langle\phi|\psi\rangle\right\} = 2 - 2\Re\left\{\langle\phi^{(n-1)}|\psi^{(n-1)}\rangle\right\} = 2 - 2\psi_1^{(n-1)}, \tag{8.44}$$

we must estimate the expression found for $\psi_1^{(n-1)}$. We use $\cos(x) \geqslant 1 - x^2/2$ for any $x \in \mathbb{R}$, where the approximation becomes sharper as $|x| \to 0$. By assumption, $|\beta_i - \alpha_i| \leqslant \delta$, hence the first term can be estimated as $c_1^{\beta-\alpha} \geqslant 1 - \delta^2/2$. Similarly, the other terms can be estimated, using the fact that $\sin(\gamma_i) \geqslant 0$ for $1 \leqslant i \leqslant n-2$ and $\gamma = \alpha, \beta$, as the first $n-2$ angles are between 0 and π. As $c_i^{\beta-\alpha} - 1 \leqslant 0$, we have

$$s_1^{\alpha} s_1^{\beta} \cdots s_k^{\alpha} s_k^{\beta}\left(c_{k+1}^{\beta-\alpha} - 1\right) \geqslant c_{k+1}^{\beta-\alpha} - 1 \geqslant -\delta^2/2. \tag{8.45}$$

Thus, in total we get

$$\psi_1^{(n-1)} \geqslant 1 - \delta^2/2 - \underbrace{\delta^2/2 - \delta^2/2 - \cdots - \delta^2/2}_{n-2 \text{ times}} = 1 - (n-1)\delta^2/2, \tag{8.46}$$

hence $\|\phi-\psi\|_2^2 \leqslant (n-1)\delta^2$, completing the proof.

□

Given this approximation lemma we can now prove:

Theorem 8.2.2: $A_{n,\varepsilon}^{(2)}$ is an ε-covering

Given an $\varepsilon > 0$, and $n \in \mathbb{N}$. Pick $\delta \leqslant \varepsilon/\sqrt{n-1}$. Construct the set $A_{n,\varepsilon}^{(2)}$ containing all $\phi^{(i_1,\ldots,i_{n-1})}$ whose n-spherical angles $(\alpha_1,\ldots,\alpha_{n-1})$ as described in Equation (8.21) are given by $\alpha_j = \alpha_i^{(i_j)}$ for $1 \leqslant j \leqslant n-1$ as described by Equation (8.21), and $\alpha_{n-1} = \alpha_{n-1}^{(i_{n-1})}$ as described by Equation (8.23). Then, given any unit vector $\psi \in \mathbb{S}^{n-1}$:

$$\min_{\phi\in A}\left\|\psi - \phi^{(i_1,\ldots,i_n)}\right\|_2 \leqslant \varepsilon. \tag{8.47}$$

Proof. Given a unit vector $\psi \in \mathbb{S}^{n-1}$ with n-spherical angles β_1 up to β_{n-1}, pick $\phi^{(i_1,\dots,i_{n-1})} \in A^{(2)}_{n,\varepsilon}$ with n-spherical angles $\alpha_j^{(i_j)}$ such that $|\alpha_j^{(i_j)} - \beta_j| \leqslant \delta$, which is possible by the uniform discretization of each spherical angle. Then, we satisfy the assumptions of Lemma 8.2.1, hence

$$\|\phi - \psi\|_2^2 \leqslant (n-1)\delta^2 \leqslant \varepsilon^2. \tag{8.48}$$

□

SIZE OF ε-COVERING $A^{(2)}_{n,\varepsilon}$

Perhaps surprisingly, the size of $A^{(2)}_{n,\varepsilon}$ does not simply evaluate to $t_1^{n-2}t_2$, but is actually slightly smaller. This is due to the fact that multiple choices of the discretized angles $\alpha_i^{(j)}$ yield the same unit vector $\phi^{(i_1,\dots,i_{n-1})}$: if any of the angles α_i is either 0 or π, then all coordinates x_{i+1} up to x_n are automatically 0 as they include a factor $\sin(\alpha_i)$. This leads to the following proposition:

Proposition 8.2.5: The size of $A^{(2)}_{n,\varepsilon}$

Given the ε-covering $A^{(2)}_{n,\varepsilon}$ for some $n \in \mathbb{N}$ and $\varepsilon > 0$, we have

$$\left|A^{(2)}_{n,\varepsilon}\right| = (t_1-2)^{n-2}t_2 + 2\sum_{k=1}^{n-2}(t_1-2)^{n-2-k}, \tag{8.49}$$

where $t_1 := 1 + \text{ceil}(\frac{\pi}{2\delta})$ and $t_2 := \text{ceil}(\frac{\pi}{\delta})$, and $\delta := \varepsilon/\sqrt{n-1}$.

Proof. We remind the reader that $\alpha_i^{(j)} := \pi j/(t_1-1)$ for $1 \leqslant i \leqslant n-2$ and $0 \leqslant j \leqslant t_1 - 1$, whilst $\alpha_{n-1}^{(j)} := 2\pi(j-1)/t_2$ for $1 \leqslant j \leqslant t_2$. If we restrict ourselves to the case where $\alpha_i \neq 0, \pi$ for $1 \leqslant i \leqslant n-2$, i.e. we exclude the endpoints of the discretization, we find that $A^{(2)}_{n,\varepsilon}$ contains $(t_1-2)^{n-2}t_2$ unique vectors.

Then, if we assume all $\alpha_i \neq 0, \pi$ for $1 \leqslant i \leqslant n-3$ and $\alpha_{n-2} = 0, \pi$, the value of α_{n-1} is inconsequential, so we have $2 \cdot (t_1-1)^{n-3}$ unique vectors with $\alpha_{n-2} = 0, \pi$. Similarly, if we assume all $\alpha_i \neq 0, \pi$ for $1 \leqslant i \leqslant n-4$ and $\alpha_{n-3} = 0, \pi$, the values of α_{n-2} and α_{n-1} are inconsequential, yielding $2 \cdot (t_1-1)^{n-4}$ unique vectors. Continuing this process until we pick $\alpha_1 = 0, \pi$ (corresponding to the n-dimensional North and South poles $(\pm 1, 0, 0, \dots, 0)$), the expression for $|A^{(2)}_{n,\varepsilon}|$ in the proposition follows. □

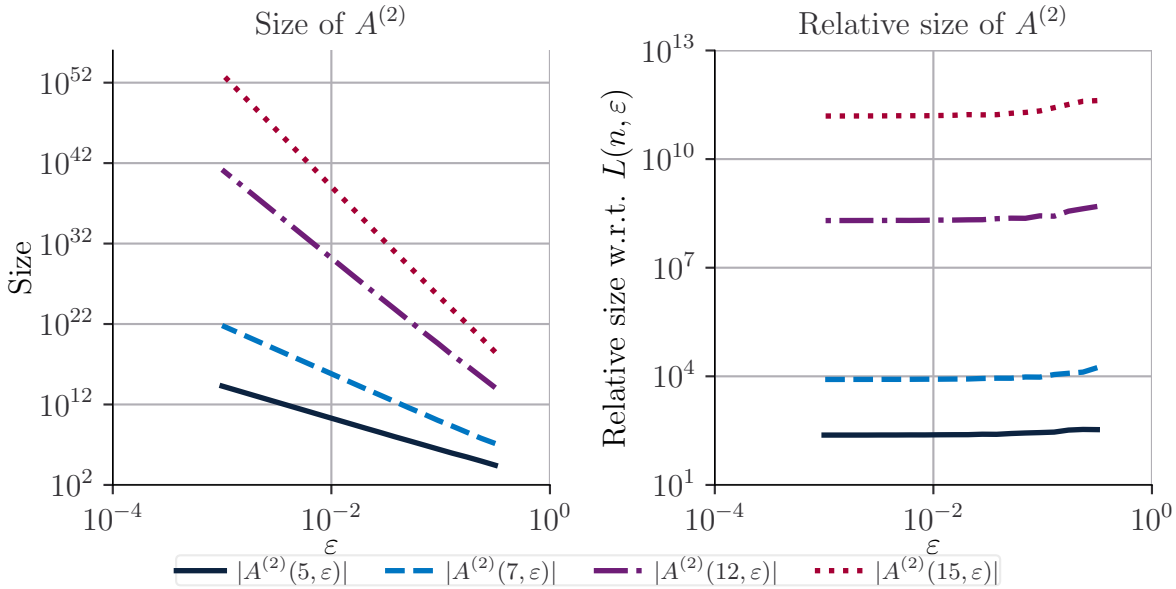


Figure 8.3: **Left**: The size $|A^{(2)}|$ of ε-cover $A^{(2)}$ for different values of n and ε. **Right**: The relative size $|A^{(2)}|/L(n,\varepsilon)$ for different values of n and ε.

8.2.6. ε-COVERING $A^{(3)}_{n,\varepsilon}$: IMPROVED DISCRETIZATION OF n-SPHERICAL COORDINATES

Although discretizing the n-spherical coordinates instead of the Cartesian coordinates gives us an actual ε-cover of the sphere $\mathbb{S}^{n-1}$ (instead of an ε-cover of the n-cube $[-1,1]^n$), it is most certainly not a uniform distribution of points in the sense that it is not ϵ-separated – we can easily visualize this in 3 dimensions, where this discretization procedure yields for example the following:

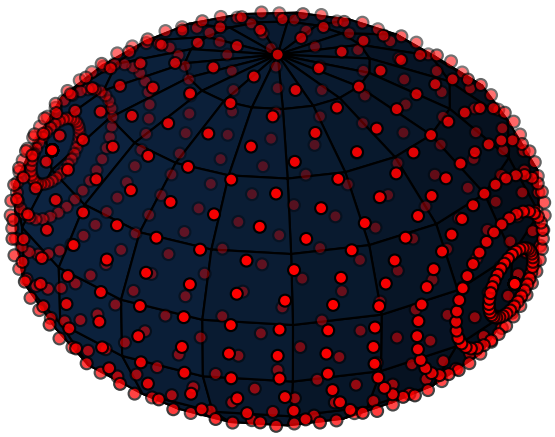


Figure 8.4: Visualization of ε-cover $A^{(2)}_{3,0.125}$.

This non-uniform distribution arises due to oversampling of points close to the North and South pole – in other words, given any latitude given by α_1, we always distribute t_2 points over the longitude, but this is a smaller circle around the North pole than around the equator. We therefore somehow wish to take into account that, when one is closer to one of the poles, one can sample the next angle more coarsely. After applying the method described below, this yields:

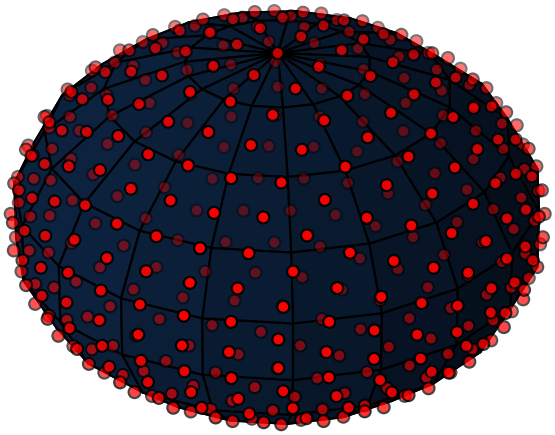


Figure 8.5: Visualization of ε-cover $A^{(3)}_{3,0.125}$.

MORE UNIFORM SAMPLING USING THE EFFECTIVE RADIUS

We remind the reader of the generalized n-spherical coordinates:

$$x_1 = \cos(\alpha_1) \tag{8.50}$$
$$x_2 = \sin(\alpha_1)\cos(\alpha_2) \tag{8.51}$$
$$x_3 = \sin(\alpha_1)\sin(\alpha_2)\cos(\alpha_3) \tag{8.52}$$
$$\vdots \tag{8.53}$$
$$x_{n-1} = \sin(\alpha_1)\cdots\sin(\alpha_{n-2})\cos(\alpha_{n-1}) \tag{8.54}$$
$$x_n = \sin(\alpha_1)\cdots\sin(\alpha_{n-2})\sin(\alpha_{n-1}). \tag{8.55}$$
$$0 \leqslant \alpha_1, \ldots, \alpha_{n-2} \leqslant \pi \quad , \quad 0 \leqslant \alpha_{n-1} < 2\pi. \tag{8.56}$$

Now, presume the angles $\alpha_1, \ldots, \alpha_{n-2}$ are given, and we need to discretize α_{n-1}, which only affects x_{n-1} and x_n. If we consider the 2-dimensional plane where we draw x_{n-1} on the x-axis and x_n on the y-axis, we see that they are located on the circle with radius $r = \sin(\alpha_1)\cdots\sin(\alpha_{n-2})$, see Figure 8.6. If two points on this circle are separated by an angle Δ, one can calculate that their Euclidean distance is $2r|\sin(\Delta/2)|$. Hence, if the previous angles $\alpha_1, \ldots, \alpha_{n-2}$ change, giving a new radius r' instead of r, if we wish to keep the same distance between the sample points by using a separation angle Δ', we require $2r|\sin(\Delta/2)| = 2r'|\sin(\Delta'/2)|$. In the small-angle approximation $\sin(\gamma) \approx \gamma$, we therefore require $r\Delta = r'\Delta'$.

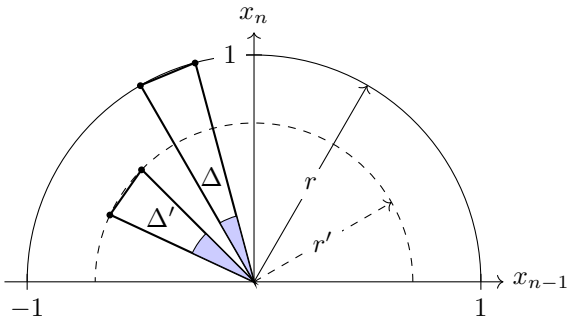


Figure 8.6: Visualization of the distribution of the coordinates x_{n-1} and x_n, given the radii $r := \sin(\alpha_1)\cdots\sin(\alpha_{n-2})$ and $r' := \sin(\alpha'_1)\cdots\sin(\alpha'_{n-2})$, and separation angles Δ and Δ'.

This intuition remains true for the other angles as well: for example, consider that we have to distribute α_2, whilst α_1, α_3 and α_4 are fixed. Let $\phi(\gamma_1, \ldots, \gamma_n)$ refer to the unit vector

whose spherical coordinates are γ_i, then we see:

$$\begin{aligned}\|\phi(\alpha_1,\alpha_2,\alpha_3,\alpha_4)-\phi(\alpha_1,\beta_2,\alpha_3,\alpha_4)\|^2 \\ &= \sin^2(\alpha_1)\left(\cos(\alpha_2)-\cos(\beta_2)\right)^2+\sin^2(\alpha_1)\left(\sin(\alpha_2)-\sin(\beta_2)\right)^2\|\phi(\alpha_3,\alpha_4)\|^2 \\ &= \sin^2(\alpha_1)\left(2-2\cos(\alpha_2-\beta_2)\right)=4\sin^2(\alpha_1)\sin^2((\alpha_2-\beta_2)/2),\end{aligned} \tag{8.57}$$

hence in the small angle approximation for $\alpha_2-\beta_2=\Delta_2\ll 1$ we have the rule that if α_1 changes to α_1', we require $\sin(\alpha_1)\Delta_2=\sin(\alpha_1')\Delta_2'$.

In general, we see that we can sample the angle α_k with discretization steps of Δ_k such that $r(\alpha_1,\ldots,\alpha_{k-1})\Delta_k$ is invariant under the choice of $\alpha_1,\ldots,\alpha_{k-1}$, where r is the effective radius $r(\alpha_1,\ldots,\alpha_{k-1})=\prod_{i=1}^{k-1}\sin(\alpha_i)$. In other words, we must choose

$$\Delta_k(\alpha_1,\ldots,\alpha_{k-1})=\frac{C_k}{\prod_{i=1}^{k-1}\sin(\alpha_i)}, \tag{8.58}$$

where C_k is a constant independent of any of the angles α_i.

Discretization scheme

Taking into account the more uniform sampling from the previous section, we therefore discretize the spherical angles as follows. Firstly, discretize α_1 by $\alpha_1^{(i_1)}=\pi(i_1-1)/(t_1-1)$ for $1\leqslant i_1\leqslant t_1$. Again, we introduce a parameter $\delta>0$ depending on ε and n (which we will precisely specify later), and then pick t_1 such that:

$$|\alpha_1^{(i_1)}-\alpha_1^{(i_1+1)}|=\pi/(t_1-1)\leqslant 2\delta \implies t_1\geqslant 1+\frac{\pi}{2\delta} \text{ so we pick } t_1=1+\text{ceil}\left(\frac{\pi}{2\delta}\right). \tag{8.59}$$

Then, *given* a discretized angle $\alpha_1^{(i_1)}$, we further discretize α_2 by $\alpha_2^{(i_2)}=\pi(i_2-1)/(t_2-1)$, and here the total number of points t_2 is a function of i_1. Namely, the discretization step δ_2 is now dependent on i_1, as we put

$$\delta_2=\min\left\{\frac{\delta}{\sin\alpha_1^{(i_1)}},\frac{\pi}{2}\right\}. \tag{8.60}$$

Here, capping the size of δ_2 by $\pi/2$ ensures that we always have at least two angles in the interval $[0,\pi]$ (separated at most by $2\delta_2$), namely on 0 and π. This is also necessary to bound the error term in the corresponding approximation Lemma 8.2.2. The number of points t_2 can be calculated again as:

$$\left|\alpha_2^{(i_2)}-\alpha_2^{(i_2+1)}\right|=\pi/(t_2-1)\leqslant 2\delta_2 \implies t_2\geqslant 1+\frac{\pi}{2\delta_2} \text{ so we pick } t_2=1+\text{ceil}\left(\frac{\pi}{2\delta_2}\right). \tag{8.61}$$

We continue in this fashion – in general, when discretizing α_k, we assume that the discretized angles $\alpha_1^{(i_1)}$ up to $\alpha_{k-1}^{(i_{k-1})}$ are provided. The total number of α_k's, given by t_k, is therefore a multivariate function depending on $(i_1,\ldots,i_{k-1})$. The discretization step δ_k is given by

$$\delta_k=\min\left\{\frac{\delta}{\prod_{j=1}^{k-1}\sin\alpha_j^{(i_j)}},\frac{\pi}{2}\right\}, \tag{8.62}$$

and we pick $t_k=1+\text{ceil}\left(\frac{\pi}{2\delta_k}\right)$.

For the last angle, we put $\alpha_{n-1}^{(i_{n-1})}=2\pi(i_{n-1}-1)/t_{n-1}$, as we need to sample the half-open interval $[0,2\pi)$. This means that we get the rule:

$$\left|\alpha_{n-1}^{(i_{n-1})}-\alpha_{n-1}^{(i_{n-1}+1)}\right|=2\pi/t_{n-1}\leqslant 2\delta_{n-1} \implies t_{n-1}\geqslant\frac{2\pi}{2\delta_{n-1}} \text{ so we pick } t_{n-1}=\text{ceil}\left(\frac{\pi}{\delta_{n-1}}\right). \tag{8.63}$$

The point set $A_{n,\varepsilon}^{(3)}$ then consists of all those admissible unit vectors $\phi^{(i_1,\ldots,i_{n-1})}$ whose spherical angles α_j correspond to $\alpha_j=\alpha_j^{(i_j)}$, and please note that the indices i_j are not independent – one must know the value of indices $i_1,\ldots,i_{j-1}$ before one can determine whether α_j is indeed one of the discretized $\alpha_j^{(i_j)}$, because this discretization is now dependent on the choices of the previous angles.

ESTIMATING $\|\phi - \psi\|$ IN THE UNIFORM DISCRETIZATION SCHEME

We now wish to find a result similar to the approximation Lemma 8.2.1, but its results are less fortunate, yielding an estimate with a quadratic function of n instead of the linear function in Lemma 8.2.1. Therefore, we have chosen to use the ε-cover $A^{(2)}_{n,\varepsilon}$ in the numerical calculations, but we still record our main findings here – the lengthy proofs are included in the appendix.

Lemma 8.2.2: Approximation lemma in the uniform discretization scheme

Given a unit vector ϕ with n-spherical angles $(\alpha_1, \ldots, \alpha_{n-1})$ and a unit vector ψ with n-spherical angles $(\beta_1, \ldots, \beta_{n-1})$. Assume that for all $1 \leqslant k \leqslant n-1$, we have:

$$|\beta_k - \alpha_k| \leqslant \delta_k \quad , \quad \delta_k := \min\left\{ \frac{\delta}{\prod_{j=1}^{k-1} \sin(\beta_j)}, \frac{\pi}{2} \right\}. \tag{8.64}$$

Then, we have:

$$\|\psi - \phi\|_2^2 \leqslant \left((n-1) + \frac{\pi}{2} \frac{(n-3)(n-2)}{2} \right) \delta^2. \tag{8.65}$$

Proof. Please see Section A.2. □

Given this approximation lemma, we can prove:

Theorem 8.2.3: $A^{(3)}_{n,\varepsilon}$ is an ε-covering

Given $\varepsilon > 0$ and $n \in \mathbb{N}$, pick $\delta > 0$ such that

$$\left((n-1) + \frac{\pi}{2} \frac{(n-3)(n-2)}{2} \right) \delta^2 \leqslant \varepsilon^2. \tag{8.66}$$

Construct the set $A^{(3)}_{n,\varepsilon}$ containing all $\phi^{(i_1,\ldots,i_{n-1})}$ with admissible n-spherical angles $\alpha_j^{(i_j)}$ as specified in the discretization scheme. Then, we have for any unit vector $\psi \in \mathbb{S}^{n-1}$:

$$\min_{\phi \in A} \left\| \psi - \phi^{(i_1,\ldots,i_n)} \right\|_2 \leqslant \varepsilon. \tag{8.67}$$

Proof. Given a unit vector $\psi \in \mathbb{S}^{n-1}$ with n-spherical angles β_1 up to β_{n-1}, pick $\phi^{(i_1,\ldots,i_{n-1})} \in A^{(3)}_{n,\varepsilon}$ with n-spherical angles $\alpha_j^{(i_j)}$ such that $|\alpha_j^{(i_j)} - \beta_j| \leqslant \delta_j$, which is possible by the discretization of each spherical angle. Then, we satisfy the assumptions of Lemma 8.2.2, hence the claim follows. □

Proposition 8.2.6: The size of $A^{(3)}_{n,\varepsilon}$

Given the ε-covering $A^{(3)}_{n,\varepsilon}$ for some $n \in \mathbb{Z}_{>0}$ and $\varepsilon > 0$, we have

$$\left| A^{(3)}_{n,\varepsilon} \right| \approx \left\langle C_{\text{mat}}^{n-2} \boldsymbol{a}, \boldsymbol{\delta}_{\text{vec}} \right\rangle, \tag{8.68}$$

where $\boldsymbol{a} := (2, 2, 0, 0, \ldots, 0)$, $\boldsymbol{\delta}_{\text{vec}} := (1, \pi/(2\delta), \pi^2/(2\delta)^2, \ldots, \pi^{n-1}/(2\delta)^{n-1})$, and the matrix C_{mat} is given by

$$C_{\text{mat}} := \begin{pmatrix} 2 & & & & & \\ 1 & I(1) & & & & \\ 0 & I(1) & I(2) & & & \\ 0 & 0 & I(2) & I(3) & & \\ & & & & \ddots & \\ & & & & I(n-2) & I(n-1) \end{pmatrix} \quad , \quad I(\beta) := \frac{1}{\pi} \int_0^\pi \sin^\beta(y) \mathrm{d}y. \tag{8.69}$$

Proof. Please see Section A.2. □

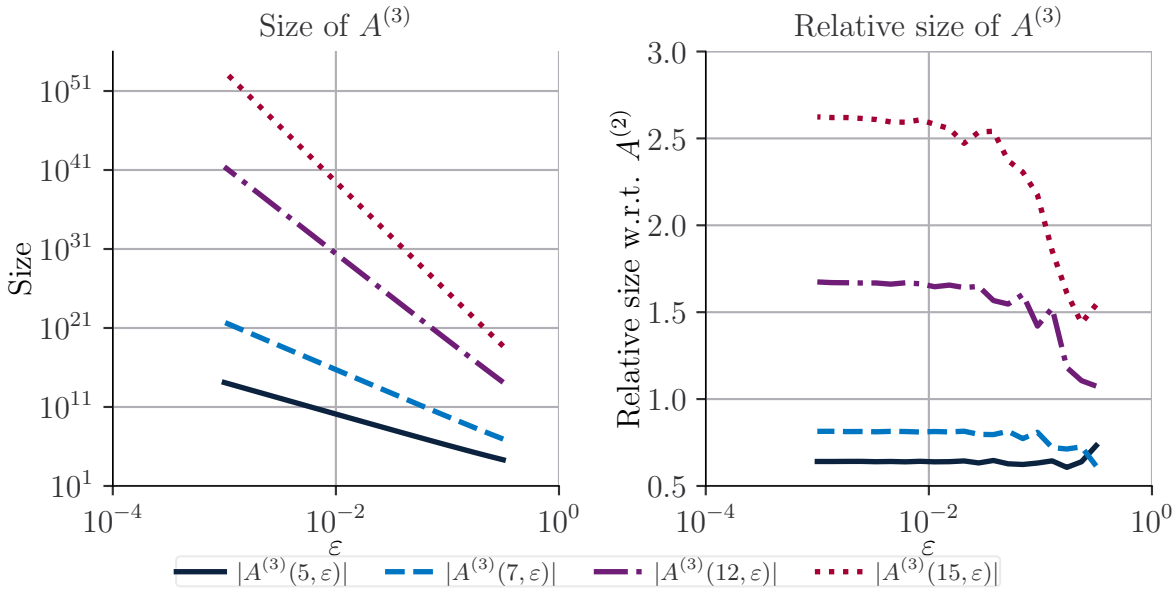


Figure 8.7: **Left**: The size $|A^{(3)}|$ (dots) of ε-cover $A^{(3)}$ for different values of n and ε plotted together with $|A^{(2)}|$ (lines) – for a given n, the same colours are used for $|A^{(2)}|$ and $|A^{(3)}|$. **Right**: The relative size $|A^{(3)}|/|A^{(2)}|$ for different values of n and ε.

8.3. Particle Swarm Optimization

We now turn our attention to approximating $H_{\min}(\mathcal{N}_1 \otimes \mathcal{N}_2)$ from above (here, $\mathcal{N}_i$ for $i = 1, 2$ are two quantum channels). We used the concavity of the von Neumann entropy to deduce that the minimum output entropy is achieved on some pure state, but finding this optimum over the complex n-sphere is in general very hard – although we know that the objective function is continuous, it may contain many local minimums, which makes finding a global optimum difficult. Furthermore, as we will see in the next section, the von Neumann entropy is not differentiable everywhere, further complicating the implementation of gradient-based methods. We therefore investigate a popular *derivative-free optimization* (DFO) method, called *Particle Swarm Optimization* (PSO).

8.3.1. Introduction and background

The particle swarm optimization method (from now on, *PSO*) was created by Eberhart and Kennedy in [EK95, KE95], and is inspired by the sociological behaviour of large collections of individual organisms, such as flocks of birds or schools of fish. As Eberhart and Shi put it, "*[t]he original intent was to graphically simulate the graceful but unpredictable choreography of a bird flock*" [ES01]. The particles are initialized randomly in the search space, and they are given a random velocity. The particles each have a discrete memory, which keeps track of the best solution they have achieved themselves so far. Furthermore, they can communicate to each other to determine the global best solution any member of the flock has so far achieved. Their social interactions are then modelled by an inertial term, a *nostalgic* term that pushes the particle back towards their own best solution position, and a *social* term that pushes the particle towards the global best position. Each timestep, these contributions to the particle's velocity are determined stochastically. Since the conception of PSO in 1995, researchers have applied it successfully in wide variety of fields [ES01, LWJ+05]. Let us commence with the fundamentals of PSO.

Assume we have an objective function $f : \mathbb{R}^d \to \mathbb{R}$ that we wish to minimize over a search domain $\Omega = [a_1, b_1] \times \cdots \times [a_d, b_d] \subseteq \mathbb{R}^d$ for some dimension $d \in \mathbb{N}$, with $-\infty \leqslant a_i < b_i \leqslant +\infty$. At the start of the algorithm, we initialize all d-dimensional positions $\boldsymbol{x}^{[i]}(0)$ uniformly randomly

over the search space Ω. Here, the superscript identifies each particle, with $1 \leqslant i \leqslant n_p$, where n_p is the number of particles, and the quantity in parentheses is the time step. We also initialize all d-dimensional velocities $\boldsymbol{v}^{[i]}(0)$ randomly with a magnitude between 0 and $v_{\max}$. The best known position of the i'th particle at initialization, denoted by $\boldsymbol{p}^{[i]}(0)$, is equal to its first position, $\boldsymbol{x}^{[i]}(0)$. The best global position $\boldsymbol{g}(0)$ is equal to $\boldsymbol{p}^{[i_0]}(0)$ where i_0 is chosen such that $f(\boldsymbol{p}^{[i_0]}(0)) \leqslant f(\boldsymbol{p}^{[j]}(0))$ for all $1 \leqslant j \leqslant n_p$. Each time step, the position and velocity are updated according to the rules

$$\boldsymbol{x}^{[i]}(k+1) = \boldsymbol{x}^{[i]}(k) + \boldsymbol{v}^{[i]}(k) \tag{8.70}$$

$$\boldsymbol{v}^{[i]}(k+1) = w(k)\boldsymbol{v}^{[i]}(k) + \boldsymbol{\phi}^{[i]}(k) \circ (\boldsymbol{p}^{[i]}(k) - \boldsymbol{x}^{[i]}(k)) + \boldsymbol{\psi}^{[i]}(k) \circ (\boldsymbol{g}(k) - \boldsymbol{x}^{[i]}(k)), \tag{8.71}$$

where $w : \mathbb{N}_0 \to [0,1]$ is the inertial weight function, $\boldsymbol{\phi}^{[i]}(k)$ is the nostalgia, uniformly randomly chosen in $[0, \phi_{\max}]^d$, $\boldsymbol{\psi}^{[i]}(k)$ is the social component, uniformly randomly chosen in $[0, \psi_{\max}]^d$, and $\circ$ denotes an element-wise multiplication of two vectors. Furthermore, the norm of $\boldsymbol{v}^{[i]}(k+1)$ is reduced to $v_{\max}$ if it exceeds it, and the components of $\boldsymbol{x}^{[i]}(k+1)$ may not exceed the boundaries of the search domain Ω. After all positions are updated, the $\boldsymbol{p}^{[i]}(k+1)$ are equal to $\boldsymbol{p}^{[i]}(k)$ whenever $f(\boldsymbol{x}^{[i]}(k+1)) \geqslant f(\boldsymbol{p}^{[i]}(k))$, else they are set to $\boldsymbol{x}^{[i]}(k+1)$. Similarly, $\boldsymbol{g}(k+1)$ is equal to $\boldsymbol{p}^{[i_0]}(k+1)$ where i_0 is chosen such that $f(\boldsymbol{p}^{[i_0]}(k+1)) \leqslant f(\boldsymbol{p}^{[j]}(k+1))$ for all $1 \leqslant j \leqslant n_p$. The algorithm terminates when a stopping criterion is reached, such as a maximal number of iterations, or a threshold for the value of $f(\boldsymbol{g}(k))$ is reached.

The hyperparameters n_p, $v_{\max}$, $\phi_{\max}$ and $\psi_{\max}$ need to heuristically optimized, and have a great influence on the performance [LWJ+05].

8.3.2. PSO on the sphere

The original algorithm proposed by Eberhart and Kennedy assumes the search domain Ω is of the shape $[a_1, b_1] \times \cdots \times [a_d, b_d] \subseteq \mathbb{R}^d$, whilst our domain of interest is the complex unit sphere in $\mathbb{C}^n$ for some $n \in \mathbb{N}$. As states are determined up to global phase, we may assume that the first component of a unit vector $|\psi\rangle \in \mathbb{C}^n$ is real. We then identify the set of pure states on $\mathbb{C}^n$ with $\mathbb{S}^{2n-2} \subseteq \mathbb{R}^{2n-1}$, where we map a real vector $\tilde{\psi} \in \mathbb{S}^{2n-2} \subseteq \mathbb{R}^{2n-1}$ to a complex vector $|\psi\rangle := \mathrm{cplx}(\tilde{\psi}) \in \mathbb{C}^n$, with components ψ_k, by:

$$\psi_0 = \tilde{\psi}_0 \quad ; \quad \psi_k = \tilde{\psi}_{2k-1} + i\tilde{\psi}_{2k} \quad \text{for} \quad 1 \leqslant k \leqslant n-1. \tag{8.72}$$

The map $\mathrm{cplx} : \mathbb{R}^{2n-1} \to \{|\psi\rangle \in \mathbb{C}^n \ : \ \psi_0 \in \mathbb{R}\}$ is an $\mathbb{R}$-linear isometric bijection.

Therefore, we now need to investigate how the PSO algorithm works on the real sphere $\mathbb{S}^{2n-2} \subseteq \mathbb{R}^{2n-1}$. We can cast this problem in the framework of optimization on (Riemannian or matrix) *manifolds*, see [AMS08, Bou20]. For this specific purpose, the `ManOpt` toolbox was developed, mainly in the Matlab language, but a port to other languages such as Python and Julia is available [BMAS14]. The velocity vectors $\boldsymbol{v}^{[i]}(k)$ are seen as tangent vectors in the space tangent to the point $\boldsymbol{x}^{[i]}(k+1)$ on the manifold. Considering the update rules for the original PSO algorithm, whenever a velocity vector and a position need to be added, we instead use a *retraction* on the manifold $\mathcal{M}$, and whenever we consider the displacement vector between two points (such as $\boldsymbol{p}^{[i]}(k) - \boldsymbol{x}^{[i]}(k)$), we use the logarithmic manifold function instead (the local diffeomorphic inverse of the exponential map). We then find the update rules:

$$\boldsymbol{v}^{[i]}(k+1) = w(k)\mathcal{T}_{\boldsymbol{x}^{[i]}(k) \leftarrow \boldsymbol{x}^{[i]}(k-1)}(\boldsymbol{v}^{[i]}(k)) + \boldsymbol{\phi}^{[i]}(k) \circ \mathrm{Log}_{\boldsymbol{x}^{[i]}(k)}(\boldsymbol{p}^{[i]}(k)) + \boldsymbol{\psi}^{[i]}(k) \circ \mathrm{Log}_{\boldsymbol{x}^{[i]}(k)}(\boldsymbol{g}(k)), \tag{8.73}$$

$$\boldsymbol{x}^{[i]}(k+1) = \mathcal{R}_{\boldsymbol{x}^{[i]}(k)}(\boldsymbol{v}^{[i]}(k+1)). \tag{8.74}$$

Here, $\mathcal{T}_{y \leftarrow x}(u)$ denotes the *vector transport* $\mathcal{T}_{y \leftarrow x} : T_x\mathcal{M} \to T_y\mathcal{M}$ of a vector $u \in T_x\mathcal{M}$ at $x \in \mathcal{M}$ to a vector in $T_y\mathcal{M}$ (see [Bou20, Def 10.68]), and $\mathcal{R}_x(u)$ denotes the *retraction* $R : T\mathcal{M} \to \mathcal{M}$, a smooth map such that for all curves $c(t) = \mathcal{R}_x(tv)$, we have $c(0) = x$ and $c'(0) = v$ [Bou20, Def. 3.47]. Furthermore, $\mathrm{Log}_x(y) = \arg\min\{\|v\|_x \ : \ v \in \mathcal{O}_x \text{ and } \mathrm{Exp}_x(y) = v\}$, defined on a domain $\mathcal{O}_x$ defined as $\{v \in T_x\mathcal{M} \ : \ (x,v) \in \mathcal{O}\}$, where $\mathcal{O} = \{(x,v) \in T\mathcal{M} \ : \ \gamma_v \text{ is defined on an interval containing } [0,1]\}$, where $\gamma_v : I \to \mathcal{M}$ is a *maximal geodesic*

with $\gamma_v(0) = x$ and $\gamma_v'(0) = v$. This is the set of update rules that is currently used in the `ManOpt` toolbox [BMAS14].

We specialize to the case where $\mathcal{M} = \mathbb{S}^{2n-2} \subseteq \mathbb{R}^{2n-1}$. In this case, we have:

$$\text{Vector transport } \mathcal{T}_{y\leftarrow x}(u) = \mathcal{P}_y(u). \tag{8.75}$$

$$\text{Tangent space projection } \mathcal{P}_y(u) = u - \langle y, u\rangle y. \tag{8.76}$$

$$\text{Retraction } \mathcal{R}_x(u) = (x+u)/\|x+u\|_{\mathbb{R}^{2n-1}}. \tag{8.77}$$

$$\text{Logarithm } \mathrm{Log}_x(y) = \tfrac{d(x,y)}{\|\mathcal{P}_x(y-x)\|}\mathcal{P}_x(y-x). \tag{8.78}$$

$$\text{Great circle distance } d(x,y) = \arccos\langle x, y\rangle. \tag{8.79}$$

Our own `qittoolbox` on GitHub contains a PSO implementation that is largely based on a combination of the original Matlab implementation of `ManOpt` and the `PyManOpt` port to Python, but fully vectorized using `numpy` to obtain a performance increase, and with an important change to the update rules. The vector transport $\mathcal{T}_{y\leftarrow x}(u) = P_y(u)$ is generally strictly contractive, which results in the following unwanted behaviour: when the inertial weight $w(k)$ is set to 1 for all k, and the *nostalgic* and *social* interactions $\phi_i(k)$ and $\psi_i(k)$ to 0 for all i and k, one would expect the particles to fly on a constant trajectory (which indeed happens in the unconstrained original version of PSO). However, due to the contractive nature of the vector transport on the sphere, the particles slow down until they come to a complete stop. We can mitigate this behaviour by rescaling the transport:

$$\tilde{\mathcal{T}}_{y\leftarrow x}(u) = \tfrac{1}{\|\mathcal{T}_{y\leftarrow x}(u)\|}\mathcal{T}_{y\leftarrow x}(u) = \tfrac{1}{\|P_y(u)\|}P_y(u). \tag{8.80}$$

8.3.3. PSO with gradient descent

Although the PSO algorithm is a derivative-free optimization method, it can be enriched with gradient-based optimization – these methods often converge faster as they can identify good search directions, but they perform poorly on flat regions where the gradient all but vanishes, and they get easily stuck in local minima. Thus, a hybrid algorithm combining the stochastic derivative-free PSO exploration technique with a gradient-descent optimization could be valuable [Noe12]. Noel proposed in [Noe12] to apply gradient descent to the current best position $g(k)$ after each iteration of the PSO algorithm, and updating this global best position in case the gradient descent algorithm finds a better position. This hybridization yielded good results on the test functions that were studied in [Noe12]. In our case, we need to investigate two steps before we can utilize this hybridization: we need to determine the (Euclidean) gradient of our cost function at a given point, and we need to relate the Euclidean gradient to the Riemannian gradient on the sphere $\mathbb{S}^{2n-2} \subseteq \mathbb{R}^{2n-1}$.

Determining the Euclidean gradient of $H(\Phi(\rho))$

We identify the following concatenation of maps:

$$\begin{aligned}\psi \in \mathbb{S}^{2n-2} \subseteq \mathbb{R}^{2n-1} &\xrightarrow{\text{cplx}} |\psi\rangle \in \mathbb{C}^n \xrightarrow{\text{mat}} \rho = |\psi\rangle\langle\psi| \in M_n(\mathbb{C}) \\ &\xrightarrow{\Phi} \Phi(\rho) \in M_m(\mathbb{C}) \xrightarrow{H} H(\Phi(\rho)) = -\mathrm{Tr}\left(\Phi(\rho)\log\Phi(\rho)\right) \in \mathbb{R}.\end{aligned} \tag{8.81}$$

Let us investigate the directional derivative of this composition in the direction of ϕ: we consider $\psi + t\phi \in \mathbb{R}^{2n-1}$. Then, as the complexification is linear, we get $|\psi\rangle + t\,|\phi\rangle$, which is mapped to the matrix $\rho(t) = |\psi\rangle\langle\psi| + t\,|\psi\rangle\langle\phi| + t\,|\phi\rangle\langle\psi| + t^2\,|\phi\rangle\langle\phi|$, and using the linearity of Φ, we see

$$\Phi(\rho(t)) = \Phi(|\psi\rangle\langle\psi|) + t\Phi(|\psi\rangle\langle\phi|) + t\Phi(|\phi\rangle\langle\psi|) + t^2\Phi(|\phi\rangle\langle\phi|). \tag{8.82}$$

We then use the derivative rule for a matrix-valued function $t \mapsto A(t)$:

$$\tfrac{\mathrm{d}}{\mathrm{d}t}H(A(t)) = -\mathrm{Tr}\left(A'(t)\left(I + \log A(t)\right)\right). \tag{8.83}$$

We therefore find the directional derivative if we set $\phi = e_i$, where $\{e_i\}_{i=1}^{2n-1}$ is the canonical basis of $\mathbb{R}^{2n-1}$:

$$D_{e_i}H(\Phi(\rho)) = \tfrac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0}H(\rho(t)) = -\mathrm{Tr}\left(\Phi\left(|e_i\rangle\langle\psi| + |\psi\rangle\langle e_i|\right)\left(I + \log\Phi(|\psi\rangle\langle\psi|)\right)\right), \tag{8.84}$$

which is the i-th component of the gradient $(\nabla_{\mathbb{R}^{2n-1}}H)(\psi)$.

RELATING THE EUCLIDEAN GRADIENT TO THE RIEMANNIAN GRADIENT

We must, however, optimize over the compact space $\mathbb{S}^{2n-2} \subseteq \mathbb{R}^{2n-1}$, so we must convert the Euclidean gradient to the Riemannian gradient on $\mathbb{S}^{2n-2}$, which means we use the tangent space projection $\mathcal{P}_y(u)$ as described in Equation (8.76) to calculate:

$$(\nabla_{\mathbb{S}^{2n-2}} H)(\psi) = \mathcal{P}_\psi((\nabla_{\mathbb{S}^{2n-2}} H)(\psi)) = (\nabla_{\mathbb{R}^{2n-1}} H)(\psi) - \langle \psi \ , \ (\nabla_{\mathbb{R}^{2n-1}} H)(\psi) \rangle_{\mathbb{R}^{2n-1}} \cdot \psi. \tag{8.85}$$

This gradient may not always exist, as $\log \Phi(|\psi\rangle\langle\psi|)$ is ill-defined whenever $\Phi(|\psi\rangle\langle\psi|)$ is a singular matrix. This corresponds to the scalar case where we see that the function $x \mapsto -x \log x$ has a derivative $-1 - \log x$ that tends to $+\infty$ for $x \downarrow 0$. However, this divergence is logarithmic, and therefore "extremely slow": for $x = 10^{-16}$, the derivative evaluates to 15. With machine precision in mind, we therefore choose to approximate the von Neumann entropy by the function $\tilde{H} : \sigma \mapsto -\mathrm{Tr}\left((\sigma + \varepsilon I)\log(\sigma + \varepsilon I)\right)$ for some small ε, for which we have the derivative rule

$$\tfrac{\mathrm{d}}{\mathrm{d}t} \tilde{H}(A(t)) = -\mathrm{Tr}\left(A'(t)(I + \log(A(t) + \varepsilon I)\right). \tag{8.86}$$

8.4. DPS HIERARCHY AND DPS* CRITERION

Having discussed the theory of ε-covers to bound the minimum output entropy from below, and having discussed the implementation of particle swarm optimization techniques on the unit sphere to find good upper bounds of the minimum output entropy, we turn our attention to the specific quantum channels that we wish to investigate. Theorem 7.3.3 tells us that the O_N^+-quantum channels $\Phi_1^{(m+1),m}$ (for $m \in \mathbb{N}$) are positive partial transpose, and may therefore be entanglement-breaking. In fact, in the smallest case of $m = 1$ we know from Theorem 7.3.4 that $\Phi_1^{(2),1}$ is indeed entanglement-breaking, and is therefore by Shor's theorem (see Theorem 3.5.1) an MOE-additive channel. To investigate whether $\Phi_1^{(m+1),m}$ for $m \geqslant 2$ are entanglement-breaking, we can utilize the so-called *DPS-hierarchy* and *DPS*-criterion* to determine whether their normalized Choi matrices are separable quantum states (see Definition 3.5.2). Let us denote by $\mathcal{S}_{\mathrm{sep}}(\mathcal{H}_\mathrm{A} \otimes \mathcal{H}_\mathrm{B})$ the set of all separable states on $\mathcal{H}_\mathrm{A} \otimes \mathcal{H}_\mathrm{B}$.

The idea of the DPS hierarchy (named after its inventors Doherty, Parrilo and Spedalieri) was first presented in [DPS02], and two years later the same authors wrote a seminal paper in which they proved the following criterion: [DPS04, Thm. 1]

Theorem 8.4.1: DPS Hierarchy: k-symmetric extensions

Given any bipartite mixed state $\rho_{\mathrm{AB}} \in \mathcal{S}(\mathcal{H}_\mathrm{A} \otimes \mathcal{H}_\mathrm{B})$. Then ρ_{AB} has a (PPT) symmetric extension to k copies of the subsystem B for all $k \in \mathbb{N}$ if and only if ρ_{AB} is separable. Here, ρ_{AB^k} is a k-symmetric extension of ρ_{AB} if it satisfies:

1. $\rho_{\mathrm{AB}^k} \in \mathcal{S}(\mathcal{H}_\mathrm{A} \otimes \mathcal{H}_\mathrm{B}^{\otimes k})$,
2. $\mathrm{Tr}_{\mathcal{H}_\mathrm{B}^{\otimes k-1}}(\rho_{\mathrm{AB}^k}) = \rho_{\mathrm{AB}}$, where the partial trace is performed over the last $k-1$ copies of $\mathcal{H}_\mathrm{B}$,
3. $\rho_{\mathrm{AB}^k} \in \mathcal{S}(\mathcal{H}_\mathrm{A} \otimes \vee^k \mathcal{H}_\mathrm{B})$ where $\vee^k \mathcal{H}_\mathrm{B}$ is the *(Bose) symmetric subspace* of $\mathcal{H}_\mathrm{B}^{\otimes k}$. Equivalently, ρ_{AB^k} is invariant under permuting any of the k tensor legs associated with $\mathcal{H}_\mathrm{B}$, equivalently $\rho_{\mathrm{AB}^k}(\iota_\mathrm{A} \otimes P^k_{\mathrm{sym}}) = \rho_{\mathrm{AB}^k}$, where P^k_{sym} is the *symmetric projection*

$$P^k_{\mathrm{sym}} := \tfrac{1}{k!} \sum_{\pi \in \mathcal{S}_k} P_\pi^{[\dim \mathcal{H}_\mathrm{B}]} \quad \text{where} \quad P_\pi^{[d]} = \sum_{i:[k]\to[d]} |i_{\pi^{-1}(1)}, \ldots, i_{\pi^{-1}(k)}\rangle\langle i_1, \ldots, i_k| \,, \tag{8.87}$$

   and where $\mathcal{S}_k$ is the k-th symmetric group.

We call ρ_{AB^k} a *PPT* k-symmetric extension of ρ_{AB} if additionally

4. ρ_{AB^k} is PPT with respect to the bipartition $AB^{\mathrm{ceil}(N/2)}|B^{\mathrm{floor}(N/2)}$ [NOP09]

Remark 8.4.1. Note that the additional PPT-constraint on ρ_{AB^k} is not necessary for the completeness result (i.e., we already have the equivalence between being separable and having a

k-symmetric extension for all $k \in \mathbb{N}$ without the PPT criterion), but this constraint makes the hierarchy of tests stronger.

Hence, for increasing values of $k \in \mathbb{N}$, one can ask whether $\rho_{\mathrm{AB}} = J(\Phi_1^{(m+1),m})$ has a (PPT) k-symmetric extension, and if not, we conclude that $J(\Phi_1^{(m+1),m})$ is not separable (and hence, $\Phi_1^{(m+1),m}$ is not entanglement-breaking). This feasibility problem can be cast into a positive semi-definite program (PSD), for which efficient algorithms and solvers exist [DPS02]; however, the time complexity scales as $(\dim \mathcal{H}_{\mathrm{B}})^{\mathcal{O}(k)}$ [HNW17] and the space complexity scales as $\mathcal{O}(k^{\dim \mathcal{H}_{\mathrm{B}}-1})$ [DPS04], which means the DPS hierarchy becomes intractable for large k or large $\dim \mathcal{H}_{\mathrm{B}}$.

The DPS hierarchy approaches the set of all separable states from the outside: if we denote by $\mathcal{S}_{k\text{-sym}}(\mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}})$ the set of all k-symmetrically extensible states on $\mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}}$, then we see $\mathcal{S}_{(k+1)\text{-sym}}(\mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}}) \subseteq \mathcal{S}_{k\text{-sym}}(\mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}})$ and $\lim_{k\to\infty} \mathcal{S}_{k\text{-sym}}(\mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}}) = \mathcal{S}_{\mathrm{sep}}(\mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}})$. In contrast, the so-called *DPS* criterion* proposed in [NOP09] approaches the set of all separable states from within by using the following sets, where we let $d := \dim \mathcal{H}_{\mathrm{B}}$ and $\sigma_A = \mathrm{Tr}_{\mathcal{H}_{\mathrm{B}}^{\otimes k}}(\sigma_{\mathrm{AB}^k})$:

$$\tilde{\mathcal{S}}_{k\text{-sym}}(\mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}}) = \{\tfrac{k}{k+d}\sigma_{\mathrm{AB}^k} + \tfrac{1}{k+d}\sigma_{\mathrm{A}} \otimes \iota_{\mathrm{B}} \ : \ \sigma_{\mathrm{AB}^k} \in \mathcal{S}_{k\text{-sym}}(\mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}})\}. \tag{8.88}$$

In this case, we no longer have a strict *hierarchy* because we do not necessarily have $\tilde{\mathcal{S}}_{k\text{-sym}}(\mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}}) \subseteq \tilde{\mathcal{S}}_{(k+1)\text{-sym}}(\mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}})$, but we do have $\tilde{\mathcal{S}}_{k\text{-sym}}(\mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}}) \subseteq \mathcal{S}_{\mathrm{sep}}(\mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}})$ for all k, and a similar completeness criterion: $\lim_{k\to\infty} \tilde{\mathcal{S}}_{k\text{-sym}}(\mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}}) = \mathcal{S}_{\mathrm{sep}}(\mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}})$ [NOP09]. In other words:

Definition 8.4.1: Inner k-symmetric extensions

Given a bipartite mixed state $\rho_{\mathrm{AB}} \in \mathcal{S}(\mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}})$. We say that $\rho_{\mathrm{AB}^k} \in \mathcal{S}(\mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}}^{\otimes k})$ is an *inner k-symmetric extension* of ρ_{AB} if there exists a k-symmetric extension σ_{AB^k} of ρ_{AB}, such that

$$\rho_{\mathrm{AB}^k} = \tfrac{k}{k+d}\sigma_{\mathrm{AB}^k} + \tfrac{1}{k+d}\sigma_{\mathrm{A}} \otimes \iota_{\mathrm{B}}. \tag{8.89}$$

8.5. RESULTS

Using the theory we have developed in the previous sections, we can now numerically investigate the MOE and the MOE-additivity of the O_N^+ and S_N^+-quantum channels. We mainly focus on the lowest-weight channels (where $k = |l - m|$), such that the input dimension of the quantum channels remains relatively small. Let us commence with numerical investigations of the lowest-weight O_N^+ channels.

8.5.1. LOWEST WEIGHT O_N^+-CHANNELS

A deeper analysis of the lowest-weight O_N^+-channels revealed that their output entropy is *constant* on all pure states, and can be quickly computed because the eigenvalues of $\Phi_1^{(m+1),m}(\rho)$ (which are independent of the choice of the pure state ρ by Theorem 7.4.1) can be computed by the recurrence relation established in Theorem 7.4.4. Therefore, we can precisely compute $H_{\min}(\Phi_1^{(m+1),m})$. Nonetheless, before one would utilize the particle swarm optimization (PSO) strategy to find good upper bounds for $H_{\min}(\Phi_1^{(m+1),m} \otimes \Phi_1^{(m'+1),m'})$ (where $m, m' \in \mathbb{N}$), one ought to investigate whether these channels are entanglement-breaking: by Theorem 7.3.3, we know that the channels $\Phi_1^{(m+1),m}$ for $m \in \mathbb{N}$ are PPT, whence they might be entanglement-breaking. We utilize the `QETLAB` toolbox [JCR16] written for Matlab as well as our own implementation in `qittoolbox` on GitHub written in Python, where we implement the DPS hierarchy and DPS* criterion established in Section 8.4. The `QETLAB` toolbox also deploys a list of much simpler separability tests before using the DPS hierarchy and DPS* criterion based on several criterions found in literature.

As a preliminary test, we investigate whether the QETLAB toolbox and the qittoolbox on GitHub implementation report that the Choi matrix of the O_N^+-channel $\Phi_1^{(2),1}$ is separable – we know from Theorem 7.3.4 that this channel is entanglement-breaking, so its Choi matrix should be separable. Indeed, for $3 \leqslant N \leqslant 7$, QETLAB reports that $J(\Phi_1^{(2),1})$ is separable, using its closeness to the maximally mixed state as shown in [GB02]. As expected, $J(\Phi_1^{(2),1})$ has a 4-symmetric PPT extension as reported by our implementation of the DPS hiearchy. However, it does *not* have an i-PPT-inner symmetric extension for $2 \leqslant i \leqslant 4$, so the DPS* criterion is not capable of detecting its separability for small extensions.

For $m = 2$, we can compute the Choi matrix of $\Phi_1^{(3),2}$ by using the results from [BC20]:

$$J(\Phi_1^{(m+1),m}) = \tfrac{[2]_q}{[m+2]_q} \alpha_{m+1}^{m,1} (\alpha_{m+1}^{m,1})^*. \tag{8.90}$$

For $N = 3$ and $m = 2$, we have checked the k-symmetric extensibility of $\Phi_1^{(3),2}$ using qittoolbox: it has a $k = 3$ symmetric extension with the additional PPT-criterion, and it does *not* have a $k = 2$ or $k = 3$ inner symmetric extension with the PPT criterion. This means that the test is inconclusive: it is still possible for the Choi matrix to be separable or entangled.

It was not possible to utilize the DPS hierarchy or the DPS* criterion for larger N and m due to computer memory limitations.

8.5.2. Lowest weight S_N^+-channels

Turning our attention to S_N^+-channels, we investigate whether they are entanglement-breaking for small input and output dimensions. If this is the case, then by Theorem 3.5.1, they cannot have a strictly subbadditive MOE. One could repeat the analysis from [BC20, Section 5] to find analytic expressions for the Choi matrices $J(\hat{\Phi}_k^{(l),m})$ and $J(\hat{\Phi}_k^{l,(m)})$, which would allow us to find an analogous statement to [BCLY20, Thm. 5.2]). This theorem states that in the O_N^+ case, for $k \neq l - m$, $\Phi_k^{(l),m}$ is not EBT, and for $k \neq m - l$, $\Phi_k^{l,(m)}$ is not EBT, because their Choi matrices are multiples of of an orthogonal projection in $\mathcal{B}(\mathcal{H}_k \otimes \mathcal{H}_m)$ or $\mathcal{B}(\mathcal{H}_k \otimes \mathcal{H}_l)$, respectively, whose range is the subrepresentation that is equivalent to $\mathcal{H}_l$ or $\mathcal{H}_m$, respectively. By Proposition 6.5.1, these subrepresentations are highly entangled for $l \neq k + m$ and $m \neq k + l$, respectively. From [BCLY20, Lemma 5.1], it then follows that the Choi matrix is not separable.

However, in our case, the MOE estimate of the S_N^+-channels $\hat{\Phi}_k^{(l),m}$ and $\hat{\Phi}_k^{l,(m)}$ is also trivial whenever the parameter r defined by $r := l+m-k$ is equal to 1, see Theorem 6.5.2. This means that we cannot establish whether any of the channels $\hat{\Phi}_{l-m}^{(l),m}$, $\hat{\Phi}_{m-l}^{l,(m)}$, $\hat{\Phi}_{l-m+1}^{(l),m}$ or $\hat{\Phi}_{m-l+1}^{l,(m)}$ are entanglement-breaking. However, for small input and output dimension, we can numerically check whether the Choi matrices are positive partial transpose: if the channels are not PPT, they are also not entanglement-breaking.

For $N = 4$, we checked for all $l, m \in \{1, 2, 3\}$ and all $\max(|l - m|, 1) \leqslant k \leqslant \min(l + m, 3)$ and found in all cases that $\hat{\Phi}_k^{(l),m}$ and $\hat{\Phi}_k^{l,(m)}$ were both not PPT. For $N = 5$, we checked for all $l, m \in \{1, 2\}$ and $\max(|l - m|, 1) \leqslant k \leqslant \min(l + m, 3)$ and again found that $\hat{\Phi}_k^{(l),m}$ and $\hat{\Phi}_k^{l,(m)}$ were in all cases not PPT. Notably, this includes the channels associated to the tuples $(k, l, m) = (1, 1, 1), (2, 2, 1), (2, 1, 2)$ for both $N = 4, 5$ and additionally $(k, l, m) = (3, 3, 1), (3, 1, 3)$ for $N = 4$. These channels all have the associated parameter $r = l + m - k = 1$, and we could therefore not analytically determine whether they are entanglement-breaking.

Based on this analysis, the channels $\hat{\Phi}_1^{(1),1}$ and $\hat{\Phi}_1^{(2),2}$ are not EBT for $N = 4, 5$, and could therefore have strictly subadditive MOE. Therefore, we turn our attention to casting an ε-cover over these channels to find a lower bound on their MOE.

8.5.3. ε-cover

We chose to investigate the S_N^+-channels $\hat{\Phi}_k^{(l),m}$ for three choices of N, k, l, m and ε. Considering the exponential growth in size of the ε-covers with the dimension of the input states, we investigated the O_N^+ or S_N^+ quantum channels that have the smallest (but non-trivial) input

dimension, which corresponds to $k = 1$, and $N = 3$ in the case of O_N^+ and $N = 4$ in the case of S_N^+. However, a deeper analysis of the lowest-weight O_N^+-channels revealed that their output entropy is *constant* on all pure states (see Theorem 7.4.1), which disqualified them from consideration for ε-covers. We therefore turn our attention to the S_N^+-channels we developed in this thesis. In light of the somewhat surprising findings of Theorem 6.5.2 where the lower bound of the MOE is smaller when r is odd (with r defined by $k = l + m - r$), we choose $k = l = m = 1$ (yielding $r = 1$) with $N = 4, 5$, and $k = 1, l = m = 2$ (yielding $r = 3$) with $N = 4$.

We remind the reader of Theorem 8.1.2: given a quantum channel $\Phi : \mathcal{S}(\mathcal{H}_1) \to \mathcal{S}(\mathcal{H}_2)$, and given two pure states $|\psi\rangle$ and $|\phi\rangle$ with the assumption $\|\psi - \phi\|_2 \leqslant \frac{1}{2}$. Let $d_2 := \dim \mathcal{H}_2$, then we have:

$$|H(\Phi(|\psi\rangle\langle\psi|)) - H(\Phi(|\phi\rangle\langle\phi|))| \leqslant \|\psi - \phi\|_2 \log_2(d_2 - 1) + H_2(\|\psi - \phi\|_2). \tag{8.91}$$

We were granted access to the Dutch National Supercomputer *Snellius* [1] to compute the output entropy on all points of the ε-covers. The following results were obtained [2]:

Table 8.1: Results of the ε-covers over the S_N^+-quantum channels $\hat{\Phi}_k^{(l),m}$ for varying values of ε, N, k, l and m. For these values, the *real dimension* $n = 2[2k+1]_{\hat{q}} - 1$ was calculated for which $\mathbb{S}^{n-1} \subseteq \mathbb{R}^n$ is mapped to $\{|\psi\rangle \in \mathbb{C}^{[2k+1]_{\hat{q}}} \simeq \hat{\mathcal{H}}_k \ : \ \psi_0 \in \mathbb{R}\}$. Furthermore, the total size of the ε-cover is calculated, and the two error terms $\varepsilon \log_2(d_2 - 1)$ and $H_2(\varepsilon)$ are displayed. Lastly, an estimate for the minimum output entropy $\hat{H}_{\min}$ is given, together with the maximal error found by summing the two error terms.

N	(k,l,m)	ε	δ	n	Size	$\varepsilon \log_2(d_2 - 1)$	$H_2(\varepsilon)$	$\hat{H}_{\min}$	Max error
4	$(1,1,1)$	0.01	0.005	5	$1.9 \cdot 10^{10}$	0.01	≈ 0.08079	1.0000	0.0908
4	$(1,2,2)$	0.01	0.005	5	$1.9 \cdot 10^{10}$	0.02	≈ 0.08079	1.7219	0.1008
5	$(1,1,1)$	0.035	≈ 0.0143	7	$3.4 \cdot 10^{12}$	≈ 0.05547	≈ 0.21888	1.0000	0.2744

To put the value of $\hat{H}_{\min}$ into perspective, we note that for $(k,l,m) = (1,1,1)$, the entropy of the maximally mixed output state is $\log_2(N-1) \approx 1.585$ for $N = 4$, and $\log_2(N-1) = 2$ for $N = 5$. For $(k,l,m) = (1,2,2)$, the entropy of the maximally mixed output state is $\log_2(N^2 - 3N + 1) \approx 2.3219$ for $N = 4$. [3]

From these findings, it becomes apparent that the usage of ε-covers is truly bounded to the low-dimensional regime: even in the modest case of covering the unit sphere in 7 dimensions, we had to increase the error ε to 0.035 to reduce the total size (for $\varepsilon = 0.01$, the size would increase to approximately $6.4 \cdot 10^{15}$). The main contribution to the error is clearly $H_2(\varepsilon)$, which grows rapidly for increasing ε: for example, an error of $H_2(\varepsilon) \leqslant 0.01$ requires an ε slightly below 0.00086.

8.5.4. Particle Swarm Optimization

In order to benchmark the application of PSO to the problem of finding the MOE of a quantum channel, let us start by considering the highest-weight channels $\hat{\Phi}_{l+m}^{(l),m}$ and $\hat{\Phi}_{l+m}^{l,(m)}$, where we know by Theorem 6.5.4 that the minimum output entropy is 0. We find:

[1] https://www.surf.nl/en/dutch-national-supercomputer-snellius

[2] *This work was carried out on the Dutch national e-infrastructure with the support of SURF Cooperative.*

[3] Here, we used the identity $[2]_{\hat{q}}[m+1]_{\hat{q}} - [m]_{\hat{q}} = [m+2]_{\hat{q}}$ for $m \in \mathbb{N}$ to find for $m = 3$ the identity $[5]_{\hat{q}} = [2]_{\hat{q}}[4]_{\hat{q}} - [3]_{\hat{q}} = \sqrt{N}\left(N\sqrt{N} - 2\sqrt{N}\right) - (N-1) = N^2 - 3N + 1$

Table 8.2: Results of 100 independent runs of the PSO and PSO-GD algorithms (corresponding to False or True in the **GD** column, respectively) that seek to minimize the output entropy of the S_N^+-quantum channels $\hat{\Phi}_{l+m}^{(l),m}$ and $\hat{\Phi}_{l+m}^{(l),m}$, corresponding to $\mathcal{H}_{\text{out}} = \mathcal{H}_m$ and $\mathcal{H}_{\text{out}} = \mathcal{H}_l$, respectively, for varying values of N, k, l and m. The minimal value of all runs is recorded in the **Best** $\hat{H}_{\min}$ column. The average of all $\hat{H}_{\min}$ over the 100 runs, together with the sample deviation, is recorded in the last column.

N	(k,l,m)	$\mathcal{H}_{\text{out}}$	GD	Best $\hat{H}_{\min}$	avg$(\hat{H}_{\min}) \pm \sigma$
4	$(2,1,1)$	$\mathcal{H}_l$	T	$5.10465 \cdot 10^{-11}$	$1.83636 \cdot 10^{-2} \pm 4.80103 \cdot 10^{-2}$
4	$(2,1,1)$	$\mathcal{H}_l$	F	$3.59177 \cdot 10^{-6}$	$5.05532 \cdot 10^{-5} \pm 5.71389 \cdot 10^{-5}$
4	$(2,1,1)$	$\mathcal{H}_m$	T	$1.02484 \cdot 10^{-10}$	$1.98834 \cdot 10^{-2} \pm 4.08666 \cdot 10^{-2}$
4	$(2,1,1)$	$\mathcal{H}_m$	F	$5.71584 \cdot 10^{-6}$	$8.33604 \cdot 10^{-5} \pm 4.17503 \cdot 10^{-4}$
4	$(3,1,2)$	$\mathcal{H}_l$	T	$1.09430 \cdot 10^{-10}$	$2.39899 \cdot 10^{-2} \pm 6.20306 \cdot 10^{-2}$
4	$(3,1,2)$	$\mathcal{H}_l$	F	$6.40487 \cdot 10^{-5}$	$6.39059 \cdot 10^{-3} \pm 2.80240 \cdot 10^{-2}$
4	$(3,1,2)$	$\mathcal{H}_m$	T	$2.23798 \cdot 10^{-8}$	$8.85733 \cdot 10^{-3} \pm 2.64066 \cdot 10^{-2}$
4	$(3,1,2)$	$\mathcal{H}_m$	F	$7.09647 \cdot 10^{-5}$	$4.21163 \cdot 10^{-3} \pm 6.72610 \cdot 10^{-3}$
5	$(2,1,1)$	$\mathcal{H}_l$	T	$1.50074 \cdot 10^{-9}$	$1.08410 \cdot 10^{-2} \pm 1.83736 \cdot 10^{-2}$
5	$(2,1,1)$	$\mathcal{H}_l$	F	$8.43941 \cdot 10^{-4}$	$9.22410 \cdot 10^{-3} \pm 1.11571 \cdot 10^{-2}$
5	$(2,1,1)$	$\mathcal{H}_m$	T	$9.45717 \cdot 10^{-10}$	$6.44861 \cdot 10^{-3} \pm 1.03997 \cdot 10^{-2}$
5	$(2,1,1)$	$\mathcal{H}_m$	F	$8.54227 \cdot 10^{-4}$	$8.20559 \cdot 10^{-3} \pm 8.84512 \cdot 10^{-3}$
5	$(3,1,2)$	$\mathcal{H}_l$	T	$5.77000 \cdot 10^{-10}$	$4.53873 \cdot 10^{-3} \pm 1.57220 \cdot 10^{-2}$
5	$(3,1,2)$	$\mathcal{H}_m$	T	$6.97742 \cdot 10^{-11}$	$5.13496 \cdot 10^{-3} \pm 1.28646 \cdot 10^{-2}$
5	$(3,1,2)$	$\mathcal{H}_m$	F	$6.38292 \cdot 10^{-2}$	$1.70077 \cdot 10^{-1} \pm 7.85152 \cdot 10^{-2}$

We see that the gradient-descent enriched PSO method outperforms the original PSO method in all cases. It also manages to find good estimates for the MOE (which is 0 in this case) in higher input dimensions (corresponding to higher values of N and k).

With this tool in hand, we can consider a specific case in more depth: we choose the $(k,l,m) = (1,1,1)$ case, that we also analyzed using the ε-covers: this yielded $H_{\min}(\Phi_1^{(1),1}) \geqslant 1.0000 - 0.0908$ and $H_{\min}(\Phi_1^{(1),1}) \geqslant 1.0000 - 0.2744$ for $N = 4$ and $N = 5$, respectively, where the first number denotes the minimal output entropy found on the nodes of the ε-cover, and the second number denotes the maximal deviation possible, see Section 8.5.3. This specific case is of interest because the lower MOE bound found in Theorem 6.5.2 is trivial as it is smaller than 0 (and approaches 0 in the limit $N \to \infty$). We find the following:

Table 8.3: Results of 100 independent runs of the PSO and PSO-GD algorithms (corresponding to the column value **GD** is False or True, respectively) that seek to minimize the output entropy of the S_N^+-quantum channel $\hat{\Phi}_1^{(1),1}$, for varying values of N. The minimal value of all runs is recorded in the **Best $\hat{H}_{\min}$ column**. The average of all $\hat{H}_{\min}$ over the 100 runs, together with the sample deviation, is recorded in the last column.

N	GD	Best $\hat{H}_{\min}$	avg$(\hat{H}_{\min}) \pm \sigma$
4	T	$1.00000 \cdot 10^{0}$	$1.00000 \cdot 10^{0} \pm 9.88631 \cdot 10^{-9}$
4	F	$1.00000 \cdot 10^{0}$	$1.00000 \cdot 10^{0} \pm 4.23084 \cdot 10^{-9}$
5	T	$1.00000 \cdot 10^{0}$	$1.01351 \cdot 10^{0} \pm 2.48793 \cdot 10^{-2}$
5	F	$1.00000 \cdot 10^{0}$	$1.00079 \cdot 10^{0} \pm 6.68804 \cdot 10^{-3}$
6	T	$1.00000 \cdot 10^{0}$	$1.02561 \cdot 10^{0} \pm 3.40027 \cdot 10^{-2}$
6	F	$1.00001 \cdot 10^{0}$	$1.00414 \cdot 10^{0} \pm 1.44261 \cdot 10^{-2}$
7	T	$1.00000 \cdot 10^{0}$	$1.03519 \cdot 10^{0} \pm 1.50104 \cdot 10^{-2}$
7	F	$1.00007 \cdot 10^{0}$	$1.03519 \cdot 10^{0} \pm 1.54949 \cdot 10^{-2}$
8	T	$9.60953 \cdot 10^{-1}$	$9.73042 \cdot 10^{-1} \pm 1.71738 \cdot 10^{-2}$
8	F	$9.60967 \cdot 10^{-1}$	$9.69092 \cdot 10^{-1} \pm 1.27063 \cdot 10^{-2}$
9	T	$8.94484 \cdot 10^{-1}$	$9.05511 \cdot 10^{-1} \pm 2.37725 \cdot 10^{-2}$
9	F	$8.94537 \cdot 10^{-1}$	$9.01560 \cdot 10^{-1} \pm 1.27098 \cdot 10^{-2}$
10	T	$8.36592 \cdot 10^{-1}$	$8.43555 \cdot 10^{-1} \pm 2.42092 \cdot 10^{-2}$
10	F	$8.36699 \cdot 10^{-1}$	$8.42953 \cdot 10^{-1} \pm 1.01063 \cdot 10^{-2}$
11	T	$7.85988 \cdot 10^{-1}$	$8.03928 \cdot 10^{-1} \pm 3.30744 \cdot 10^{-2}$
11	F	$7.86146 \cdot 10^{-1}$	$8.01947 \cdot 10^{-1} \pm 2.59771 \cdot 10^{-2}$
12	T	$7.41490 \cdot 10^{-1}$	$7.71767 \cdot 10^{-1} \pm 5.99180 \cdot 10^{-2}$
12	F	$7.41638 \cdot 10^{-1}$	$7.64464 \cdot 10^{-1} \pm 4.00876 \cdot 10^{-2}$
13	T	$7.02103 \cdot 10^{-1}$	$7.35975 \cdot 10^{-1} \pm 6.08363 \cdot 10^{-2}$
13	F	$7.02819 \cdot 10^{-1}$	$7.33599 \cdot 10^{-1} \pm 4.79379 \cdot 10^{-2}$
14	T	$6.67010 \cdot 10^{-1}$	$6.85970 \cdot 10^{-1} \pm 5.16975 \cdot 10^{-2}$
14	F	$6.68326 \cdot 10^{-1}$	$6.91203 \cdot 10^{-1} \pm 4.42134 \cdot 10^{-2}$
15	T	$6.35549 \cdot 10^{-1}$	$6.70369 \cdot 10^{-1} \pm 7.62529 \cdot 10^{-2}$
15	F	$6.36862 \cdot 10^{-1}$	$6.76487 \cdot 10^{-1} \pm 7.72862 \cdot 10^{-2}$

Interestingly, the best upper bound of the MOE of $\hat{\Phi}_1^{(1),1}$ decreases as as N grows – it is possible that the MOE of $\hat{\Phi}_1^{(1),1}$ decreases to 0 in the limit $N \to \infty$, which would in turn make the trivial lower bound found in Theorem 6.5.2 still asymptotically sharp for $N \to \infty$.

We keep our attention on the $k = l = m = 1$ case, and investigate the best upper bound for $H_{\min}(\hat{\Phi}_1^{(1),1} \otimes \hat{\Phi}_1^{(1),1})$ that we can find. In the case of $N = 4$, The PSO-GD algorithm yielded an upper bound of 2.00000000004, which therefore does not lead to a violation of the MOE-additivity conjecture, but is quite close to twice the value found by the ε-cover for a single channel.

If we turn our attention to the choice $k = 1$, $l = m = 2$ and $N = 4$, the best ε-cover bound we found was $H_{\min} \geqslant 1.7219 - 0.1008$. The PSO-GD algorithm yielded an upper bound of $H_{\min}(\hat{\Phi}_1^{(2),2} \otimes \hat{\Phi}_1^{(2),2})$ of 3.4438562180616468, so although it gets close to twice the lower bound on $H_{\min}$, it does not lead to a violation of the MOE-additivity conjecture.

9

Discussion and conclusion

In this thesis, we have investigated the covariant quantum channels that arise from the representation theory of compact quantum groups. Specifically, we reviewed the work of Brannan and Collins in [BC16b] who introduced the quantum channels associated to the free orthogonal quantum group O_N^+, and we investigated the connection between the intertwiner spaces of O_N^+ and the *Temperley-Lieb* theory, which is at the heart of this construction.

Afterwards, we went beyond the results from Brannan and Collins and investigated the possibility of generalizing this approach to other compact quantum groups. We introduced a similar construction for the quantum permutation group S_N^+ by investigating its intertwiner spaces, and combined this with previous work by Banica and Brannan to give a concrete construction of the quantum channels associated to S_N^+. We also translated the lower bounds on the *minimum output entropy* (from now on, *MOE*) found by Brannan and Collins for the O_N^+-quantum channels to similar lower bounds for the MOE of the S_N^+-quantum channels. Notably, we found the interesting result that our lower bound for the MOE of the S_N^+-channels $\hat{\Phi}_k^{(l),m}$, characterized by a tuple of $(k,l,m) \in \mathbb{N}_0$ such that $k = l + m - r$ for some $0 \leqslant r \leqslant 2\min\{l,m\}$, depends on whether the parameter r is odd or even, which makes the MOE bound trivial for the choice $r = 1$.

We then analyzed the lowest-weight quantum channels in more depth, and we simplified the Stinespring representations of these quantum channels. For the O_N^+-quantum channels, we improved the negative partial transpose results from [BCLY20] to include all fundamental dimensions $N \geqslant 3$. Furthermore, we found the surprising result that it is possible to analytically calculate the MOE of the lowest-weight quantum channels associated to O_N^+ in the case where we embed the fundamental representation in the tensor product of two larger representations, allowing us to directly analyze the sharpness of the lower bound for the MOE found by Brannan and Collins in lower dimensions: the MOE-bound is already adequate in low dimensions and becomes sharper as the dimension grows.

Afterwards, we turned our attention to numerical approximations of the MOE of a general quantum channel. We discussed how one could utilize the theory of ε-covers to cover the unit sphere of pure states together with the continuity of the von Neumann entropy to find lower bounds for the MOE. We showed a proof-of-concept by applying these ε-covers to three of the smallest S_N^+-quantum channels, but also noted that this method quickly becomes intractable for growing dimensions as it scales exponentially in the input dimension.

To bound the MOE from above, we utilized a derivative-free optimization method called *Particle Swarm Optimization* (PSO). We modified the original algorithm, whose search space is an orthotope (hyperrectangle) in $\mathbb{R}^n$ for some $n \in \mathbb{N}$, to work on the unit sphere $\mathbb{S}^{n-1} \subseteq \mathbb{R}^n$ for some $n \in \mathbb{N}$. Furthermore, we analyzed the gradient of the von Neumann entropy and used this additional information to enrich the PSO algorithm with a gradient-descent (GD) component, applied to the best global position after each iteration of PSO. We benchmarked the performance of PSO by approximating the MOE of the highest-weight S_N^+-channels, as we

know that their MOE is 0 for any dimension $N \geqslant 4$. We saw that PSO was adequate at finding good local minima for these channels. Afterwards, we applied PSO to the S_N^+-channel $\hat{\Phi}_1^{(1),1}$ with $k = l = m = 1$, as it has $r = 1$ (with r defined by $r := l + m - k$), and for $r = 1$ the analytical MOE lower bound for S_N^+ is trivial. We found a monotonically decreasing upper bound of the MOE for growing N, which might indicate that the MOE decreases to 0 for $N \to \infty$.

We utilized the PSO algorithm to find upper bounds on the MOE of the tensor product of two S_N^+-channels, but did not find any violation of the MOE-additivity conjecture. We also briefly discussed the utilization of the DPS hierarchy and DPS* criterion to numerically determine whether a quantum channel is entanglement-breaking, by investigating whether its Choi state is separable, but this did not lead to definitive results.

Based on the work in this thesis, we can formulate proposals for future research. Firstly, one could consider investigating the dependency of the MOE lower bound for S_N^+ on whether r is odd or even: in the case where r is odd, in Proposition 6.5.2, we throw away the additional information that the term $T_q^*(e_i \otimes e_j)$ not only forces $j = \check{i}$ (which is also true in the O_N^+ case), but also outputs the quantum state $|i_1\rangle \in \mathcal{H}$. Perhaps it is possible to utilize this additional information to sharpen the bound, possibly removing the additional factor $N^{1/4}$ present in the bound of Proposition 6.5.2, or it might be possible to prove that this bound is still asymptotically sharp for $N \to \infty$, akin to the work done for O_N^+ in [BC16b, BCLY20]. However, this requires a better understanding of the range space of the S_N^+-Jones-Wenzl projections $\hat{p}_k$ to formulate quantum input states for which this lower MOE bound is achieved (asymptotically).

Secondly, one could try to generalize the methods from [BC16b] to other partition quantum groups, whose intertwiner spaces have the same type of diagrammatic interpretations as O_N^+ and S_N^+. A viable candidate is the hyperoctahedral quantum group H_N^+, but its irreducible representations are no longer indexed by $\mathbb{N}_0$, but rather by the free monoid over $\mathbb{Z}_2$, and its fusion rules are more involved than those of O_N^+ or S_N^+ [BV09]. Its intertwiner spaces correspond to non-crossing partitions whose blocks are all of even size. Future research will first have to determine whether the Temperley-Lieb theory, which is instrumental in finding concrete models for the irreducible representations of O_N^+ and S_N^+ through the Jones-Wenzl projections, is still applicable in this situation.

Thirdly, the numerical methods that we employed in this thesis can be refined. We proposed an adaptive ε-cover based on the discretization of the spherical angles, where the stepsize of each spherical angle is dependent on the choice of all previous spherical angles. Although this ε-cover should scale better with the dimension n than the uniform discretization of all spherical angles, we did not succeed in finding a sharp bound on the distance between a unit vector and the ε-cover. If a better bound is found, it could be used to significantly decrease the size of the ε-cover. Furthermore, we converted the norm $\|\cdot\|_1$ that naturally appears in the continuity bounds of the von Neumann entropy to the $\|\cdot\|_2$-norm with a multiplicative penalty of 2 such that we could discretize the Euclidean unit sphere, but it might be possible to formulate an ε-cover using the $\|\cdot\|_1$-norm instead.

Furthermore, although the use of ε-covers is restricted to the case where the input dimension of a quantum channel is small, no such restrictions apply to the output dimension. However, calculating the von Neumann entropy on each node of the ε-cover becomes more expensive with growing output dimension as it requires a diagonalization of the state obtained by applying the quantum channel to the state corresponding to the node of the ε-cover. It is possible to calculate $\mathrm{Tr}(p(\sigma))$ instead of $H(\sigma)$, where $p : \mathbb{R} \to \mathbb{R}$ is a polynomial that approximates $x \mapsto -x \log x$ on $[0, 1]$. We effectively substitute the diagonalization that is necessary for calculating $H(\sigma)$ by a collection of matrix multiplications that are necessary to calculate $\mathrm{Tr}(p(\sigma))$, which is generally a cheaper operation. For large matrices, matrix multiplication also becomes expensive, but $\mathrm{Tr}(p(\sigma))$ can be approximated by a Monte-Carlo scheme using a certain number of inner products $\langle\omega|p(\sigma)|\omega\rangle$ with random vectors $|\omega\rangle$, for details please see [WBS14]. It is then possible to compute the quantity $\mathrm{Tr}(p(\sigma))$ for all $\sigma = \mathcal{N}(|\psi\rangle\langle\psi|)$ where $|\psi\rangle$

is a node in the ε-cover and $\mathcal{N}$ is a quantum channel, and only compute $H(\sigma)$ if the value of $\mathrm{Tr}(p(\sigma))$ is so sufficiently small that a new global minimum of H could be achieved on σ.

The numerical methods utilized in this thesis have been bundled into a quantum information theory toolbox. As ε-covers are also utilized for other optimization problems, such as bounding operator norms [BH15] or solving optimization problems over separable states [SW12], our explicit construction of ε-covers over the unit sphere could be beneficial to other researchers: this method can be utilized to find lower and upper bounds for any continuous quantity over the unit sphere. Similarly, our implementation of the *Particle Swarm Optimization* on the unit sphere, enriched with the gradient-descent algorithm, could be utilized to solve other minimization problems. The toolbox is written in Python and mainly built on the packages `numpy` and `scipy`, and can be publicly accessed at `qittoolbox` on GitHub . There, we also provide a small tutorial that shows how one can utilize the different tools.

Appendices

A

Appendix: Additional proofs

In this appendix, we will provide a selection of additional lemmata and proofs for the treatment of ε-covers. In particular, we will show how we derived bounds on the Euclidean distance between two unit vectors on the $(n-1)$-sphere whose spherical angles are close to each other.

A.1. Additional proofs for the ε-cover $A^{(2)}_{n,\varepsilon}$

Let us state Taylor's theorem for multivariate functions [Edw73, Thm. 7.1]:

Theorem A.1.1: First-order multivariate Taylor's theorem

Given a twice continuously differentiable function $f : \Omega \to \mathbb{R}$ on an open domain $\Omega \subseteq \mathbb{R}^n$ for some $n \in \mathbb{N}$ that contains the line segment L from $\boldsymbol{\alpha}$ to $\boldsymbol{\beta}$, then we can approximate the value of f at $\boldsymbol{\beta}$ by its value at $\boldsymbol{\alpha}$ as follows:

$$f(\boldsymbol{\beta}) = f(\boldsymbol{\alpha}) + \sum_{k=1}^{n} \left.\frac{\partial f}{\partial x_k}\right|_{\boldsymbol{x}=\boldsymbol{\alpha}} (\beta_k - \alpha_k) + R_1(\boldsymbol{\alpha}, \boldsymbol{\beta}), \tag{A.1}$$

where the Lagrange rest term R_1 is defined as

$$R_1(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \sum_{|\gamma|=2} \frac{1}{\gamma!} \cdot (\partial^\gamma f)\,\big|_{\boldsymbol{x}=\boldsymbol{c}} \cdot (\boldsymbol{\beta} - \boldsymbol{\alpha})^\gamma, \tag{A.2}$$

where $\boldsymbol{c} \in L$ is a vector whose components c_i are between β_i and α_i. Here, γ is a multi-index of degree $|\gamma|$: this is an n-tuple $\gamma = (\gamma_1, \ldots, \gamma_n) \in \mathbb{N}_0^n$, with $|\gamma| := \sum_i \gamma_i$ and $\gamma! := \prod_i (\gamma_i!)$. Furthermore, $\boldsymbol{x}^\gamma := \prod_i x_i^{\gamma_i}$, and lastly, the notation $\partial^\gamma f$ refers to:

$$\partial^\gamma f = \partial_1^{\gamma_1} \cdots \partial_n^{\gamma_n} f = \frac{\partial^{|\gamma|} f}{\partial x_1^{\gamma_1} \cdots \partial x_n^{\gamma_n}}. \tag{A.3}$$

We can now state the result that was alluded to in Section 8.2.5, but was not included there as a better result was found using rotations:

Lemma A.1.1: Approximation lemma for $A^{(2)}_{n,\varepsilon}$

Given angles $(\beta_1, \ldots, \beta_k)$ and $(\alpha_1, \ldots, \alpha_k)$ for some $k \in \mathbb{N}$, and assume $|\beta_i - \alpha_i| \leqslant \delta$ for all

$1 \leqslant i \leqslant k$. Then, we have:

$$\begin{aligned}&\left|\sin(\alpha_1)\cdots\sin(\alpha_{k-1})\cos(\alpha_k) - \sin(\beta_1)\cdots\sin(\beta_{k-1})\cos(\beta_k)\right| \\ &\quad\leqslant \delta\cdot\left(1-\frac{1}{k}\right)^{k/2}\cdot\left(\sqrt{k-1}+1/\sqrt{k-1}\right) + \delta^2k^2/2 \leqslant \delta\cdot\left(\sqrt{\frac{k-1}{e}}+\frac{1}{\sqrt{e(k-1)}}\right)+\delta^2k^2/2.\end{aligned} \tag{A.4}$$

Proof. As all $|\beta_i - \alpha_i| \leqslant \delta$, let $\beta_i = \alpha_i + \epsilon_i$ for $|\epsilon_i| \leqslant \delta$. We will use $\boldsymbol{\alpha}$ to denote the vector $(\alpha_1, \ldots, \alpha_k)$, and $\boldsymbol{\epsilon}$ to denote the vector $(\epsilon_1, \ldots, \epsilon_k)$. Then, consider the following function:

$$f(\boldsymbol{x}) = f(x_1, \ldots, x_k) = \sin(x_1)\sin(x_2)\cdots\sin(x_{k-1})\cos(x_k). \tag{A.5}$$

As the function f is a smooth function of all its variables, we can use the multivariate Taylor theorem to approximate f up to first order.

When we look at $R_1(\boldsymbol{\alpha}, \boldsymbol{\beta})$ in our case, we note that every second derivative of f is still a product of sines and cosines (possibly with a minus sign in front). Hence, surely, $|\partial^\gamma f| \leqslant 1$ on the entirety of $\mathbb{R}^k$. Hence, we can estimate:

$$\begin{aligned}|R_1(\boldsymbol{\alpha}, \boldsymbol{\alpha}+\boldsymbol{\epsilon})| &= \left|\sum_{i=1}^k \frac{1}{2}\frac{\partial^2 f}{\partial x_i^2}\bigg|_{\boldsymbol{x}=\boldsymbol{c}}\epsilon_i^2 + \sum_{i<j}\frac{\partial^2 f}{\partial\alpha_i\partial\alpha_j}\bigg|_{\boldsymbol{x}=\boldsymbol{c}}\epsilon_i\epsilon_j\right| \\ &\leqslant \sum_{i=1}^k\frac{\epsilon_i^2}{2} + \sum_{i<j}|\epsilon_i\epsilon_j| \leqslant \sum_{i=1}^k\frac{\delta^2}{2} + \sum_{i<j}\delta^2 = \frac{k}{2}\delta^2 + \frac{k(k-1)\delta^2}{2} = \frac{\delta^2k^2}{2}.\end{aligned} \tag{A.6}$$

We now need to find a good estimate of the factor $\sum_{i=1}^t \frac{\partial f}{\partial x_i}\big|_{\boldsymbol{x}=\boldsymbol{\alpha}}\epsilon_i$. In order to make the proof accessible, let us for now focus on the case $k = 3$, such that we have:

$$f(\boldsymbol{\alpha}) = \sin(\alpha_1)\sin(\alpha_2)\cos(\alpha_3), \tag{A.7}$$

and hence, where we abbreviate $s(x) := \sin(x)$ and $c(x) := \cos(x)$:

$$\sum_{i=1}^k\frac{\partial f}{\partial x_i}\bigg|_{\boldsymbol{x}=\boldsymbol{\alpha}}\epsilon_i = \epsilon_1c(\alpha_1)s(\alpha_2)c(\alpha_3) + \epsilon_2s(\alpha_1)c(\alpha_2)c(\alpha_3) - \epsilon_3s(\alpha_1)s(\alpha_2)s(\alpha_3). \tag{A.8}$$

Thus, we can cast our estimation problem into the optimization problem $\mathcal{P}_3$:

$$\mathcal{P}_3 : \begin{cases}\text{maximize} & |\epsilon_1c(\alpha_1)s(\alpha_2)c(\alpha_3) + \epsilon_2s(\alpha_1)c(\alpha_2)c(\alpha_3) - \epsilon_3s(\alpha_1)s(\alpha_2)s(\alpha_3)| \\ \text{such that} & -\delta \leqslant \epsilon_i \leqslant \delta \quad \forall 1 \leqslant i \leqslant 3 \\ & 0 \leqslant \alpha_1 \leqslant \pi \\ & 0 \leqslant \alpha_2 \leqslant \pi \\ & 0 \leqslant \alpha_3 \leqslant 2\pi.\end{cases} \tag{A.9}$$

This problem and its generalization $\mathcal{P}_k$ to any k can actually be analytically solved, and we will do so in Lemma A.1.2. For now, assuming Lemma A.1.2 is valid, this problem has an optimal value that scales as:

$$C_k = \delta\cdot\left(\frac{k-1}{k}\right)^{k/2}\cdot\frac{k}{\sqrt{k-1}} = \delta\cdot\left(1-\frac{1}{k}\right)^{k/2}\cdot\left(\sqrt{k-1}+1/\sqrt{k-1}\right). \tag{A.10}$$

Now, the first term can be estimated as follows: it is increasing as a function of k [1]. Hence, we can estimate the term by its limit value, which is:

$$\lim_{t\to\infty}(1-1/t)^{t/2} = \lim_{x\to\infty}(1+(-1/2)/x)^x = e^{-1/2} = \sqrt{1/e} \approx 0.607. \tag{A.12}$$

[1] Namely, consider $g(n) = (1-1/n)^n$. Then, by the AM-GM inequality for $p_1 = 1$ and $p_2 = p_3 = \cdots = p_{n+1} = 1 - 1/n$, note that all $p_i \geqslant 0$ and $p_1 \neq p_2$, so

$$\left(1-\frac{1}{n}\right)^{\frac{n}{n+1}} = \sqrt[n+1]{p_1\cdots p_{n+1}} < \frac{p_1+\cdots+p_{n+1}}{n+1} = \frac{1+n(1-1/n)}{n+1} = \frac{n}{n+1} = 1-\frac{1}{n+1}. \tag{A.11}$$

Hence, by taking the $(n+1)$-th power, we see $g(n) < g(n+1)$. The first term is $g(t)^{1/2}$, and the square root is monotonously increasing, from which our claim follows.

Hence, we have

$$\left|\sum_{i=1}^{k} \frac{\partial f}{\partial x_i}\bigg|_{\boldsymbol{x}=\boldsymbol{\alpha}} \epsilon_i\right| = \delta C_k \leqslant \delta \cdot \left(\sqrt{\frac{k-1}{e}} + \frac{1}{\sqrt{e(k-1)}}\right). \tag{A.13}$$

Hence, we have found:

$$|f(\boldsymbol{\beta}) - f(\boldsymbol{\alpha})| \leqslant \left|\sum_{i=1}^{k} \frac{\partial f}{\partial x_i}\bigg|_{\boldsymbol{x}=\boldsymbol{\alpha}} \epsilon_i\right| + |R_1(\boldsymbol{\alpha},\boldsymbol{\beta})| \leqslant \delta \cdot \left(\sqrt{\frac{k-1}{e}} + \frac{1}{\sqrt{e(k-1)}}\right) + \delta^2 k^2/2. \tag{A.14}$$

Filling in the definition of f yields the statement of the lemma. □

What is left, is to prove the claim that the optimization problem in the previous lemma can be analytically solved:

Lemma A.1.2: Optimization lemma for Lemma A.1.1

For $k \in \mathbb{N}$ with $k \geqslant 2$, define the optimization problem $\mathcal{P}_k$ as:

$$\mathcal{P}_k : \begin{cases} \text{maximize} & \left|\left(\epsilon_k \prod_{i=1}^{k} \sin(\alpha_i)\right) + \sum_{i=1}^{k-1} \epsilon_i \cos(\alpha_k)\cos(\alpha_i) \prod_{j=1, j\neq i}^{k-1} \sin(\alpha_j)\right| \\ \text{such that} & -\delta \leqslant \epsilon_i \leqslant \delta \quad \forall 1 \leqslant i \leqslant k \\ & 0 \leqslant \alpha_i \leqslant \pi \quad \forall 1 \leqslant i \leqslant k-1 \\ & 0 \leqslant \alpha_k \leqslant 2\pi. \end{cases} \tag{A.15}$$

The problem $\mathcal{P}_k$ has an optimal solution, and its optimal value C_k is precisely

$$C_k = \delta \cdot \left(\frac{k-1}{k}\right)^{k/2} \cdot \frac{k}{\sqrt{k-1}}. \tag{A.16}$$

Proof. First note that we will always choose $|\epsilon_i| = \delta$, and we can choose $\operatorname{sgn}(\epsilon_i)$ depending on the factor with which it is multiplied, such that all k factors in the objective function are positive. Hence, we can equivalently solve the problem

$$\text{maximize} \quad \left|\prod_{i=1}^{k} \sin(\alpha_i)\right| + \sum_{i=1}^{k-1} \left|\cos(\alpha_k)\cos(\alpha_i) \prod_{j=1, j\neq i}^{k-1} \sin(\alpha_j)\right| \tag{A.17}$$

under the same constraints for α_i. Now let us put $x_i := \sin(\alpha_i)$ for $1 \leqslant i \leqslant k-1$ and $x_k = \cos(\alpha_k)$. Hence, we now have the constraints $0 \leqslant x_i \leqslant 1$ for $1 \leqslant i \leqslant k-1$, and the constraint $-1 \leqslant x_k \leqslant 1$. However, note that the objective function only depends on the absolute value of x_k, so we may as well assume $0 \leqslant x_k \leqslant 1$. Then, note that $|\cos(\alpha_i)| = \sqrt{1-x_i^2}$, and $|\sin(\alpha_k)| = \sqrt{1-x_k^2}$, so our objective is to maximize the function g, given by

$$\begin{aligned} g(\boldsymbol{x}) = x_1 \cdots x_{k-1}\sqrt{1-x_k^2} + x_1 \cdots x_{k-2}\sqrt{1-x_{k-1}^2}x_k + \dots \\ \dots + x_1 \cdots x_{k-3}\sqrt{1-x_{k-2}^2}x_{k-1}x_k + \cdots + \sqrt{1-x_1^2}x_2 \cdots x_k. \end{aligned} \tag{A.18}$$

We now see that the very symmetric function g is also in fact continuous on the compact domain $[0,1]^k$, so we know that the maximum is achieved. Let us define the functions

$$f_i(\boldsymbol{x}) = x_1 \cdots x_{i-1}\sqrt{1-x_i^2}x_{i+1} \cdots x_k \quad ; \quad f_i^{(j)}(\boldsymbol{x}) = f_i(\boldsymbol{x})/x_j, \tag{A.19}$$

where we assume $j \neq i$ and extend $f_i^{(j)}$ so that it is continuous at $x_j = 0$. So, $f_i^{(j)}$ is just the function f_i without the factor x_j. Clearly, $g(\boldsymbol{x}) = \sum_i f_i(\boldsymbol{x})$. Then note that g is smooth in each

variable on $[0,1)^m$, and we can compute:

$$\frac{\partial g}{\partial x_i} = \left(\sum_{j\neq i} f_j^{(i)}(\boldsymbol{x})\right) + \frac{-x_i}{\sqrt{1-x_i^2}}\prod_{j\neq i} x_j, \tag{A.20}$$

We investigate the critical points of g in $(0,1)^k$ (note the exclusion of 0 in this open interval): we see

$$\frac{\partial g}{\partial x_i} = 0 \iff \frac{x_i}{\sqrt{1-x_i^2}} = \frac{\sum_{j\neq i} f_j^{(i)}(\boldsymbol{x})}{\prod_{j\neq i} x_j} = \sum_{j\neq i}\frac{\sqrt{1-x_j^2}}{x_j}. \tag{A.21}$$

Let us define $y_i := x_i/\sqrt{1-x_i^2} \in (0,\infty)$, then we see that the above is equivalent to:

$$y_i = \sum_{j\neq i} 1/y_j \quad , \quad \forall 1 \leqslant i \leqslant k. \tag{A.22}$$

This set of equations has the solution $y_i = \sqrt{k-1}$, which is in fact unique for $k \geqslant 3$. Assume for now $k \geqslant 3$, we will come back to $k = 2$ at the end. Namely, given any such solution $(y_1,\ldots,y_k)$, consider the difference $y_1 - y_2$, and utilize Equation (A.22) to find: [2]

$$y_1 - y_2 = 1/y_2 + 1/y_3 + \cdots + 1/y_k - (1/y_1 + 1/y_3 + \cdots + 1/y_k) = 1/y_2 - 1/y_1. \tag{A.23}$$

Then, multiply both sides by y_1y_2 to find $y_1y_2(y_1 - y_2) = y_1 - y_2$, i.e. $y_1 = y_2$ or $y_1 = 1/y_2$. Considering the constraint $y_1 = 1/y_2 + 1/y_3 + \cdots + y_k$, the latter is impossible as all y_i are strictly positive, hence $y_1 = y_2$. This analysis holds in general for any pair (y_i, y_j). Then, $y_i = (k-1)/y_i$ yields $y_i = \sqrt{k-1}$ for all i.

This critical point has $x_i^* = \sqrt{k-1}/\sqrt{k}$ and thus $\sqrt{1-x_i^{*2}} = \sqrt{1/k}$, and has the value:

$$g(\boldsymbol{x}^*) = k\cdot\left(\frac{\sqrt{k-1}}{\sqrt{k}}\right)^{k-1}\cdot\frac{1}{\sqrt{k}} = (k-1)^{(k-1)/2}k^{1/2-(k-1)/2} = (k-1)^{(k-1)/2}k^{1-k/2}. \tag{A.24}$$

Let us briefly investigate the case of $k = 2$: we start out with the constraint $y_1 = 1/y_2$ (and $y_2 = 1/y_1$, which is equivalent), but in this case starting from $x_1/\sqrt{1-x_1^2} = \sqrt{1-x_2^2}/x_2$, squaring both sides and re-arranging terms reveals $x_1^2 + x_2^2 = 1$, so at these points we have $g(\boldsymbol{x}) = x_1\sqrt{1-x_2^2} + x_2\sqrt{1-x_1^2} = x_1^2 + x_2^2 = 1$, so $g(\boldsymbol{x}^*) = 1$.

All other candidates for optimal values of g must occur at the boundary of the domain $[0,1]^k$. Note that if any $x_i = 0$, all but one term in g vanishes, and the remaining term is $\prod_{j\neq i} x_j$, so the obvious maximal value is 1 in this case. Note that if any $x_i = 1$, the problem simplifies as well: by the symmetry of g, we may as well assume that $x_k = 1$. In this case, we see that we now must determine the maximum of

$$x_1\cdots x_{k-2}\sqrt{1-x_{k-1}^2} + \cdots + \sqrt{1-x_1^2}x_2\cdots x_{k-1}, \tag{A.25}$$

which is simply the function g for $k-1$ variables. The base case of $k = 1$ trivially yields a maximum of 1, and note that g^* is an increasing function of k for $k \geqslant 2$ (and has the same value 1 for both $k = 1$ and $k = 2$), hence this does not yield a global maximum. Therefore, $\boldsymbol{x}^*$ is the critical point at which g achieves its maximum, and the lemma follows. □

We continue with the proof of Proposition 8.2.3, in which we we approximate $L(n,\varepsilon) := \sup_{\delta\in(0,1)} L(n,\varepsilon,\delta)$ with $L(n,\varepsilon,\delta) = [(1+\delta)^n - (1-\delta)^n]\,(\varepsilon+\delta)^{-n}$:

Proof of Proposition 8.2.3. We can approximate $L(n,\varepsilon)$ as follows. Fixing n and ε, we consider

$$f : [0,1] \to [0,\infty) \quad : \quad f(\delta) := L(n,\varepsilon,\delta) = \frac{(1+\delta)^n - (1-\delta)^n}{(\varepsilon+\delta)^n}. \tag{A.26}$$

[2] The author thanks fellow student Stephan Loor for exploiting the symmetry of the optimization problem by looking at terms of the form $y_1 - y_2$ to prove uniqueness.

We see that f is a continuously differentiable function of δ on $[0,1]$, it is also non-negative on this interval. We will show that its derivative at $\delta = 0$ is strictly positive, and we will show that $f'(\delta) = 0$ has precisely one solution for $\delta \in [0,1]$. We will approximate this stationary point which leads to an approximation of the maximal value of f on $[0,1]$.

Let us introduce $T_n(\delta) := (1+\delta)^n - (1-\delta)^n$ and $N_n(\delta) := (1+\delta)^n + (1-\delta)^n$, hence $f(\delta) = T_n(\delta)/(\varepsilon+\delta)^n$. We can directly see:

$$T_n'(\delta) = n(1+\delta)^{n-1} + n(1-\delta)^{n-1} = nN_{n-1}(\delta) \tag{A.27}$$

$$N_n'(\delta) = n(1+\delta)^{n-1} - n(1-\delta)^{n-1} = nT_{n-1}(\delta). \tag{A.28}$$

Furthermore, $T_n(0) = 0$ whilst $N_n(0) = 2$, and $T_n(1) = 2^n$ and $N_n(1) = 2^n$. Let us now calculate:

$$f'(\delta) = \frac{(\varepsilon+\delta)^n \cdot n \cdot N_{n-1}(\delta) - T_n(\delta) \cdot n \cdot (\varepsilon+\delta)^n}{(\varepsilon+\delta)^{2n}}. \tag{A.29}$$

Hence, we find $f'(0) = 2n\varepsilon^n/\varepsilon^{2n} = 2n\varepsilon^{-n} > 0$ for all $\varepsilon > 0$ and $n \in \mathbb{Z}_{>0}$. To find the stationary point(s), we solve:

$$f'(\delta) = 0 \iff (\varepsilon+\delta)N_{n-1}(\delta) - T_n(\delta) = 0. \tag{A.30}$$

Although this might not be analytically solvable for δ, we can approximate a solution by rewriting:

$$f'(\delta) = 0 \iff \varepsilon + \delta = \frac{T_n(\delta)}{N_{n-1}(\delta)} := g(\delta). \tag{A.31}$$

The function g is again a continuously differentiable function of δ on $[0,1]$. Furthermore, $g(0) = 0$ and $g(1) = 2$. We show that $g'(\delta) \geqslant 0$ for all $\delta \in [0,1]$, hence g is a monotonously increasing function and must intersect the line $\delta \mapsto \varepsilon + \delta$ precisely once on $[0,1]$, provided that $0 < \varepsilon < 1$. This point is therefore the unique stationary point of f. Let us show the non-negativity of g' :

$$g'(\delta) = \frac{nN_{n-1}^2(\delta) - nT_n(\delta)T_{n-2}(\delta)}{N_{n-1}^2(\delta)}. \tag{A.32}$$

Hence, showing $g'(\delta) \geqslant 0$ is equivalent to $\dfrac{N_{n-1}^2(\delta)}{T_n(\delta)T_{n-2}(\delta)} \geqslant 1$. We simply expand the functions to find:

$$\begin{aligned}\frac{N_{n-1}^2(\delta)}{T_n(\delta)T_{n-2}(\delta)} &= \frac{(1+\delta)^{2n-2} + 2(1+\delta)^{n-1}(1-\delta)^{n-1} + (1-\delta)^{2n-2}}{(1+\delta)^{2n-2} - (1-\delta)^n(1+\delta)^{n-2} - (1+\delta)^n(1-\delta)^{n-2} + (1-\delta)^{2n-2}} \\ &= \frac{N_{2n-2}(\delta) + 2(1+\delta)^{n-1}(1-\delta)^{n-1}}{N_{2n-2}(\delta) - ((1+\delta)^n(1-\delta)^{n-2} + (1+\delta)^n(1-\delta)^{n-2})} \geqslant 1.\end{aligned} \tag{A.33}$$

We now approximate g by a linear function around $\delta = 0$. As $g(0) = 0$, and $g'(0) = 4n/4 = n$, we see that $g(\delta) \approx n\delta$ for δ around 0. Hence, if we solve $\varepsilon + \delta = g(\delta)$, an approximate solution is $\delta = \varepsilon/(n-1)$. This provides us with the estimate $L(n, \varepsilon, \varepsilon/(n-1))$ for $L(n,\varepsilon)$. The second inequality in Proposition 8.2.3 follows from considering the binomial expansion of $T_n(x)$:

$$T_n(x) = (1+x)^n - (1-x)^n = 2nx + 2\binom{n}{3}x^3 + 2\binom{n}{5}x^5 + \cdots + 2\binom{n}{n'}x^{n'}, \tag{A.34}$$

where $n' = n$ if n is odd and $n' = n-1$ if n is even. Hence, $T_n(x) \geqslant 2nx$ for $x \geqslant 0$, with the inequality increasingly sharp as $x \downarrow 0$. The denominator of $L_1(n,\varepsilon)$ contains a term that can be estimated as well:

$$\left(\frac{n}{n-1}\right)^{-n+1} = \left(1 - \frac{1}{n}\right)^{n-1} \searrow e^{-1} \text{ as } n \to \infty. \tag{A.35}$$

Hence:

$$\begin{aligned}L_1(n,\varepsilon) = L(n,\varepsilon,\varepsilon/(n-1)) = \frac{T_n\left(\frac{\varepsilon}{n-1}\right)}{\left(\frac{n}{n-1}\varepsilon\right)^n} &\geqslant \left(\frac{n}{n-1}\right)^{-n}\varepsilon^{-n}\frac{2n\varepsilon}{n-1} = 2\left(\frac{n}{n-1}\right)^{-n+1}\varepsilon^{-n+1} \\ &= L_2(n,\varepsilon) \geqslant 2e^{-1}\varepsilon^{-n+1} = L_3(n,\varepsilon).\end{aligned} \tag{A.36}$$

□

A.2. ADDITIONAL PROOFS FOR THE ε-COVER $A^{(3)}_{n,\varepsilon}$

We now consider the adaptive ε-cover $A^{(3)}_{n,\varepsilon}$, whose step sizes δ_i depend on the choices of all previous discretized angles α_j for $1 \leqslant j \leqslant i-1$. We need to adapt Lemma 8.2.1, which used rotations to give a bound on the Euclidean distance between two unit vectors when the difference between the spherical angles of these two unit vectors is bounded by a maximal error δ. We find:

Proof of Lemma 8.2.2. We follow the analysis from Lemma 8.2.1, which is independent of the choice of the parameter δ until we find in Equation (8.43):

$$\|\phi - \psi\|_2^2 = 2 - 2\psi_1^{(n-1)}, \tag{A.37}$$

and

$$\psi_1^{(n-1)} = c_1^{\beta-\alpha} + s_1^\alpha s_1^\beta \left(c_2^{\beta-\alpha} - 1\right) + s_1^\alpha s_1^\beta s_2^\alpha s_2^\beta \left(c_3^{\beta-\alpha} - 1\right) + \cdots + s_1^\alpha s_1^\beta \cdots s_{n-2}^\alpha s_{n-2}^\beta \left(c_{n-1}^{\beta-\alpha} - 1\right). \tag{A.38}$$

The analysis now breaks down for the second angle, because we can still approximate $c_2^{\beta-\alpha} - 1 \geqslant -(\beta_2 - \alpha_2)^2/2 \geqslant -\delta_2^2/2$, but for δ_2^2 we only have the estimate

$$\delta_2^2 = \min\{\tfrac{\delta^2}{\sin(\beta_1)^2}, \tfrac{\pi^2}{4}\}. \tag{A.39}$$

Thus, we need a multiplicative term $\sin(\beta_1)^2$ to estimate this term in $\psi_1^{(n-1)}$ quadratically in δ, but we have the multiplicative term $\sin(\alpha_1)\sin(\beta_1)$. However, $\sin(\alpha_1)$ should be close to $\sin(\beta_1)$, so we need to approximate this further. Write $\alpha_i = \beta_i + \varepsilon_i$ with $|\varepsilon_i| \leqslant \delta_i$, consider $\sin(\alpha_i) = \sin(\beta_i)\cos(\varepsilon_i) + \cos(\beta_i)\sin(\varepsilon_i)$ and expand:

$$s_1^\alpha s_1^\beta \left(c_2^{\beta-\alpha} - 1\right) = c_1^\varepsilon \left(s_1^\beta\right)^2 \left(c_2^{\beta-\alpha} - 1\right) + c_1^\beta s_1^\varepsilon s_1^\beta \left(c_2^{\beta-\alpha} - 1\right). \tag{A.40}$$

Note that $|\varepsilon_i| \leqslant \delta_i \leqslant \pi/2$, so we always have $\cos(\varepsilon_i) \geqslant 0$. Then note that the first term is now in the right shape:

$$c_1^\varepsilon \left(s_1^\beta\right)^2 \left(c_2^{\beta-\alpha} - 1\right) \geqslant -\cos(\varepsilon_1)\sin^2(\beta_1)\delta_2^2/2 \geqslant -\delta^2/2. \tag{A.41}$$

The second term requires some tweaking: we only have one $\sin(\beta_1)$, but we now additionally have $\sin(\varepsilon_1)$. Using $|\sin(\gamma)| \leqslant |\gamma|$ for any $\gamma \in \mathbb{R}$:

$$\cos(\beta_1)\sin(\varepsilon_1)\sin(\beta_1)\left(1 - c_2^{\beta-\alpha}\right) \leqslant |\cos(\beta_1)\sin(\varepsilon_1)|\sin(\beta_1)|\delta_2^2/2. \tag{A.42}$$

We now estimate $\delta_2 \leqslant \delta/\sin(\beta_1)$ and also $\delta_2 \leqslant \pi/2$ to arrive at $\sin(\beta_1)\delta_2^2 \leqslant \delta\frac{\pi}{2}$. Then, we get:

$$|\cos(\beta_1)\sin(\varepsilon_1)|\sin(\beta_1)\delta_2^2/2. \leqslant |\varepsilon_1|\frac{\delta}{2}\frac{\pi}{2} \leqslant \frac{\pi}{2}\delta^2/2. \tag{A.43}$$

This gives us insight into the general case. Consider the term $A_k := s_1^\alpha s_1^\beta \cdots s_k^\alpha s_k^\beta \left(1 - c_{k+1}^{\beta-\alpha}\right)$. We first expand s_1^α to find

$$A_k = \underbrace{c_1^\varepsilon (s_1^\beta)^2 s_2^\alpha s_2^\beta \cdots s_k^\alpha s_k^\beta \left(1 - c_{k+1}^{\beta-\alpha}\right)}_{:=A_k^{(1)}} + c_1^\beta s_1^\varepsilon s_1^\beta s_2^\alpha s_2^\beta \cdots s_k^\alpha s_k^\beta \left(1 - c_{k+1}^{\beta-\alpha}\right). \tag{A.44}$$

We can immediately estimate the last term: it contains $\prod_{i=1}^k \sin(\beta_i)$, so estimate $1 - c_{k+1}^{\beta-\alpha} \leqslant \delta_{k+1}^2/2$, then estimate $\delta_{k+1} \leqslant \pi/2$ and use $\prod_{i=1}^k \sin(\beta_i)\delta_{k+1} \leqslant \delta$ to get an upper bound of $\delta\pi/(2\cdot 2)$. As we also have $|\sin(\varepsilon_1)| \leqslant |\varepsilon_1| \leqslant \delta$, we get an upper bound of $\delta^2\pi/(2\cdot 2)$.

The first term $A_k^{(1)}$ is not yet useful, so we expand $\sin(\alpha_2)$ to get:

$$A_k^{(1)} = \underbrace{c_1^\varepsilon c_2^\varepsilon (s_1^\beta)^2 (s_2^\beta)^2 s_3^\alpha s_3^\beta \cdots s_k^\alpha s_k^\beta \left(1 - c_{k+1}^{\beta-\alpha}\right)}_{:=A_k^{(2)}} + c_1^\varepsilon c_2^\beta (s_1^\beta)^2 s_2^\varepsilon s_2^\beta s_3^\alpha s_3^\beta \cdots s_k^\alpha s_k^\beta \left(1 - c_{k+1}^{\beta-\alpha}\right). \tag{A.45}$$

The last term now contains $\sin(\varepsilon_2)$, and we know $|\sin(\varepsilon_2)| \leqslant |\varepsilon_2| \leqslant \delta_2$. As we have two $\sin(\beta_1)$ terms, we can sacrifice one to estimate $\sin(\beta_1)\delta_2 \leqslant \delta$, and we then still have $\prod_{i=1}^{k} \sin(\beta_i)$ left. Estimating the cosine again yields $\delta_{k+1}^2/2$, we estimate $\delta_{k+1} \leqslant \pi/2$, we use the product of sines to get $\prod_{i=1}^{k} \sin(\beta_i)\delta_{k+1} = \delta$, and we thus find an upper bound of $\delta^2\pi/(2\cdot 2)$.

This continues: we expand the next $\sin(\alpha_i)$, we get a term with $\sin(\varepsilon_i)$ and $(1-c_{k+1}^{\beta-\alpha})$ which we can estimate quadratic in δ as we have have terms $\sin(\beta_j)^2$ for $j < i$ in there, which still leaves a term $\prod_{j=1}^{k} \sin(\beta_j)$ after using $\prod_{j=1}^{i-1} \sin(\beta_j)|\varepsilon_i| \leqslant \delta$, so we can estimate $\prod_{j=1}^{k} \sin(\beta_j)(1 - c_{k+1}^{\beta-\alpha}) \leqslant \pi\delta/(2\cdot 2)$.

The last term (which is the only term without any $\sin(\varepsilon_i)$ in it) also works out:

$$c_1^{\varepsilon} \cdots c_k^{\varepsilon} (s_1^{\beta})^2 (s_2^{\beta})^2 \cdots (s_k^{\beta})^2 (1 - c_{k+1}^{\beta-\alpha}) \leqslant \delta^2/2, \tag{A.46}$$

as we precisely have twice the product $\prod_{i=1}^{k} \sin(\beta_i)$ to get $\prod_{i=1}^{k} \sin(\beta_i)\delta_{k+1} = \delta$. In all, we get

$$A_k \leqslant (k-1)\frac{\pi}{2}\frac{\delta^2}{2} + \frac{\delta^2}{2}. \tag{A.47}$$

As k in A_k runs from 1 to $n-2$, we have

$$\sum_{k=1}^{n-2} A_k \leqslant \sum_{k=1}^{n-2} \left((k-1)\frac{\pi}{2}\frac{\delta^2}{2} + \frac{\delta^2}{2} \right) = \frac{\pi}{2}\frac{\delta^2}{2}\frac{(n-3)(n-2)}{2} + (n-2)\frac{\delta^2}{2}. \tag{A.48}$$

Hence,

$$\psi_1^{(n-1)} = c_1^{\beta-\alpha} - \sum_{k=1}^{n-2} A_k \geqslant 1 - \frac{\delta^2}{2} - (n-2)\frac{\delta^2}{2} - \frac{(n-3)(n-2)}{2}\frac{\pi}{2}\frac{\delta^2}{2}. \tag{A.49}$$

Finally,

$$\|\phi - \psi\|_2^2 = 2 - 2\psi_1^{(n-1)} \leqslant \left((n-1) + \frac{\pi}{2}\frac{(n-3)(n-2)}{2} \right) \delta^2. \tag{A.50}$$

□

We now consider the size of the ε-covers $A_{n,\varepsilon}^{(3)}$. Calculating their size is much more involved than calculating the size of $A_{n,\varepsilon}^{(2)}$, because the discretization of α_i depends on the choice of all previous angles α_j for $1 \leqslant j \leqslant i-1$. We find:

Proof of Proposition 8.2.6. Our strategy is to approximate the sum

$$|A_{n,\varepsilon}^{(3)}| = \sum_{(i_1,\ldots,i_n) \text{ admissible}} 1 = \sum_{i_1=1}^{t_1} \sum_{i_2=1}^{t_2} \cdots \sum_{i_{n-2}=1}^{t_{n-2}} t_{n-1}, \tag{A.51}$$

where each nested sum over i_k is dependent on all the choices of i_j for $j < k$, hence we must start with the very last sum.

Approximation of the sum over i_{n-2}. Assume we are given the discretized angles $\alpha_1^{(i_1)}$ up to $\alpha_{n-2}^{(i_{n-2})}$. We then know that $t_{n-1} = \text{ceil}(\pi/\delta_{n-1})$, where $\delta_{n-1} = \delta/\prod_{j=1}^{n-2} \sin(\alpha_j^{(i_j)})$, or $\delta_{n-1} = \pi/2$ if that is smaller. We use $\text{ceil}(x) \leqslant x+1$ for any $x \in \mathbb{R}$ to replace t_{n-1} with a new upper bound $\hat{t}_{n-1}$, defined by:

$$t_{n-1} \leqslant \hat{t}_{n-1} := 2 + \frac{\pi}{\delta} s_{n-2}(\alpha), \tag{A.52}$$

where we again use the notation $s_k(\alpha) = \prod_{i=1}^{k} \sin(\alpha_i)$. Then, let us sum over all possible values of $\alpha_{n-2}^{(i_{n-2})}$, assuming $(i_1, \ldots, i_{n-3})$ is fixed for the moment. This means that the sum over i_{n-2} in the expression of $|A_{n,\varepsilon}^{(3)}|$ can be upper-bounded by:

$$\sum_{i_{n-2}=1}^{t_{n-2}} t_{n-1} \leqslant \sum_{i_{n-2}=1}^{t_{n-2}} 2 + \frac{\pi}{\delta} s_{n-3}(\alpha) \sin(\alpha_{n-2}^{(i_{n-2})}). \tag{A.53}$$

We move all terms that are not dependent on the summand out of the sum and we find:

$$\sum_{i_{n-2}=1}^{t_{n-2}} 2 + \frac{\pi}{\delta} s_{n-3}(\alpha) \sin(\alpha_{n-2}^{(i_{n-2})}) = 2t_{n-2} + \frac{\pi}{\delta} s_{n-3}(\alpha) \sum_{i_{n-2}=1}^{t_{n-2}} \sin\left(\frac{\pi(i_{n-2}-1)}{t_{n-2}-1}\right). \tag{A.54}$$

With foresight, we will investigate this sum in slightly more generality, and note for $\beta \in \mathbb{N}$:

$$\sum_{i=1}^{t} \sin^{\beta}\left(\frac{\pi(i-1)}{t-1}\right) \approx \int_1^t \sin^{\beta}\left(\frac{\pi(x-1)}{t-1}\right) \mathrm{d}x = \frac{t-1}{\pi} \int_0^{\pi} \sin^{\beta}(y) \mathrm{d}y := (t-1) I(\beta), \tag{A.55}$$

where the integrals $I(\beta)$ are well-known and can be expressed in term of the beta function B given by $B(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}\ \mathrm{d}t$, which satisfies $B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$:

$$I(\beta) = \frac{1}{\pi} \int_0^{\pi} \sin^{\beta}(y) \mathrm{d}y = \frac{1}{\pi} B\left(\frac{1}{2}, \frac{\beta+1}{2}\right) = \frac{\Gamma(1/2)\Gamma((\beta+1)/2)}{\pi \Gamma(1+\beta/2)}. \tag{A.56}$$

Let us return to the expression found in Equation (A.54), which we now approximate by:

$$2t_{n-2} + \frac{\pi}{\delta} s_{n-3}(\alpha) \sum_{i_{n-2}=1}^{t_{n-2}} \sin\left(\frac{\pi(i_{n-2}-1)}{t_{n-2}-1}\right) \approx 2t_{n-2} + \frac{\pi}{\delta} s_{n-3}(\alpha)(t_{n-2}-1) I(1). \tag{A.57}$$

Approximation of the sum over i_{n-3} Having approximated the sum over i_{n-2}, we need to sum over all possible values of i_{n-3}. First, note that as $t_{n-2} = 1 + \mathrm{ceil}(\pi/(2\delta_{n-2}))$, and $\delta_{n-2} = \delta / s_{n-3}(\alpha)$ or $\delta_{n-2} = \pi/2$, whichever is smaller, we utilize the upper bound $\mathrm{ceil}(x) \leqslant 1 + x$ for any $x \in \mathbb{R}$ again to find the new upper bound $\hat{t}_{n-2}$, defined by:

$$t_{n-2} \leqslant \hat{t}_{n-2} := 2 + \frac{\pi}{2\delta} s_{n-3}(\alpha). \tag{A.58}$$

Hence, we substitute $\hat{t}_{n-2}$ for t_{n-2} in the summand, which is described in Equation (A.57). We find:

$$2t_{n-2} + \frac{\pi}{\delta} s_{n-3}(\alpha)(t_{n-2}-1) I(1) \leqslant 2 \cdot 2 + \left(\frac{2\pi}{2\delta} + \frac{\pi}{\delta} I(1)\right) s_{n-3}(\alpha) + \frac{\pi^2}{2\delta^2} I(1) s_{n-3}^2(\alpha). \tag{A.59}$$

Now expand $s_{n-3}(\alpha) = s_{n-4}(\alpha) \sin(\alpha_{n-3}^{(i_{n-3})})$, move all terms that are not dependent on i_{n-3} out of the summation, and use the integral approximation as given in Equation (A.55) to find:

$$\begin{aligned} &\sum_{i_{n-3}=1}^{t_{n-3}} 2 \cdot 2 + \left(\frac{2\pi}{2\delta} + \frac{\pi}{\delta} I(1)\right) s_{n-4}(\alpha) \sin(\alpha_{n-3}^{(i_{n-3})}) + \frac{\pi^2}{2\delta^2} I(1) s_{n-4}^2(\alpha) \sin^2(\alpha_{n-3}^{(i_{n-3})}) \\ &\approx 2 \cdot 2t_{n-3} + \left(\frac{2\pi}{2\delta} I(1) + \frac{\pi}{\delta} I(1)^2\right) s_{n-4}(\alpha)(t_{n-3}-1) + \frac{\pi^2}{2\delta^2} I(1) I(2) s_{n-4}(\alpha)^2 (t_{n-3}-1). \end{aligned} \tag{A.60}$$

Approximation of the sum over general i_k. Now, we find the pattern: we find an upper bound $\hat{t}_k$ for t_k, given by the equation $t_k \leqslant \hat{t}_k := 2 + \frac{\pi}{2\delta} s_{k-1}(\alpha)$. Any constant c going into a sum comes out as $ct_k \leqslant 2c + \frac{c\pi}{2\delta} s_{k-1}(\alpha)$, and any term proportional to $s_k^{\beta}(\alpha)$ in a sum over i_k comes out as:

$$s_k^{\beta}(\alpha) \mapsto s_{k-1}^{\beta}(\alpha) I(\beta)(t_k - 1) \leqslant s_{k-1}^{\beta}(\alpha) I(\beta) + \frac{\pi}{2\delta} s_{k-1}^{\beta+1}(\alpha). \tag{A.61}$$

We see that we must keep track of the terms $s_k^{\beta}(\alpha)$ involved, which are the only terms dependent on the summing variables i_j for $1 \leqslant j \leqslant n-2$. We utilize polynomials for the bookkeeping, and those polynomials will be defined recursively. Firstly, let us look back at the upper bound $\hat{t}_{n-1}$ for t_{n-1}, defined by $\hat{t}_{n-1} = 2 + \frac{\pi}{\delta} s_{n-2}(\alpha)$. We can define a polynomial p_{n-2} given by $p_{n-2}(x) := 2 + 2x$, such that $\hat{t}_{n-1} = p_{n-2}(\frac{\pi}{2\delta} s_{n-2}(\alpha))$. Then, for the next summand in the sum

over i_{n-3}, we define $p_{n-3}(x) = 2 \cdot 2 + 2 \cdot x + 2I(1)x + 2I(1)x^2$. We estimated the summand in the sum over i_{n-3} in Equation (A.59), which we can now describe as:

$$2 \cdot 2 + \left(\frac{2\pi}{2\delta} + \frac{\pi}{\delta} I(1) \right) s_{n-3}(\alpha) + \frac{\pi^2}{2\delta^2} I(1) s_{n-3}^2(\alpha) = p_{n-3}\left(\frac{\pi}{2\delta} s_{n-3}(\alpha) \right). \tag{A.62}$$

In this language, we can therefore write:

$$p_{n-3}\left(\tfrac{\pi}{2\delta} s_{n-3}(\alpha) \right) \approx \sum_{i_{n-2}=1}^{t_{n-2}} p_{n-2}\left(\tfrac{\pi}{2\delta} s_{n-2}(\alpha) \right). \tag{A.63}$$

We now wish to generalize this: find a polynomial p_{k-1} such that $p_{k-1}\left(\frac{\pi}{2\delta} s_{k-1}(\alpha)\right)$ approximates $\sum_{i_k=1}^{t_k} p_k\left(\frac{\pi}{2\delta} s_k(\alpha)\right)$. We will give a recursion relation for p_{k-1} in terms of p_k.

We noted at the start of this paragraph that any constant c going into a sum comes out as $ct_k \leqslant 2c + \frac{c\pi}{2\delta} s_{k-1}(\alpha)$, so we see this as a mapping $c \mapsto 2c + cx$. Similarly, according to Equation (A.61), any term proportional to $s_k^\beta(\alpha)$ came out of the sum over i_k as $s_{k-1}^\beta(\alpha) I(\beta)(t_k - 1)$, which we estimated by $s_{k-1}^\beta(\alpha) I(\beta) + \frac{\pi}{2\delta} s_{k-1}^{\beta+1}(\alpha)$. Hence, we see this as a mapping $x^\beta \mapsto I(\beta)x^\beta + I(\beta)x^{\beta+1}$. Note that these mappings applied to $p_{n-2}(x) = 2 + 2x$ indeed yield p_{n-3}.

Then, the dynamics of finding the polynomial p_{k-1} from p_k can be described in matrix-vector notation as follows. Let $p_k := a_0^{(k)} + a_1^{(k)} x + \cdots + a_{n-1}^{(k)} x^{n-1}$, then:

$$\begin{pmatrix} a_0^{(k-1)} \\ a_1^{(k-1)} \\ a_2^{(k-1)} \\ \vdots \\ a_{n-1}^{(k-1)} \end{pmatrix} = \underbrace{\begin{pmatrix} 2 & & & & & \\ 1 & I(1) & & & & \\ 0 & I(1) & I(2) & & & \\ 0 & 0 & I(2) & I(3) & & \\ & & & & \ddots & \\ & & & & I(n-2) & I(n-1) \end{pmatrix}}_{:= C_{\text{mat}}} \begin{pmatrix} a_0^{(k)} \\ a_1^{(k)} \\ a_2^{(k)} \\ \vdots \\ a_{n-1}^{(k)} \end{pmatrix}. \tag{A.64}$$

After evaluating all other sums, we finally end up with $\sum_{i_1=1}^{t_1} p_{n-(n-1)}(\frac{\pi}{2\delta}\sin(\alpha_1))$. This sum can be evaluated just like the previous ones, yielding a polynomial p_0, which we should then evaluate in $\pi/(2\delta)$, so we find

$$|A_{n,\varepsilon}^{(3)}| \approx p_0(\pi/(2\delta)) = \left\langle \begin{pmatrix} a_0^{(0)} \\ \vdots \\ a_{n-1}^{(0)} \end{pmatrix}, \boldsymbol{\delta}_{\text{vec}} \right\rangle = \left\langle C_{\text{mat}}^{n-2} \boldsymbol{a}^{(n-2)}, \boldsymbol{\delta}_{\text{vec}} \right\rangle, \tag{A.65}$$

where $\boldsymbol{a}^{(n-2)} := (2, 2, 0, 0, \ldots, 0)$ is the starting coefficient vector of polynomial $p_{n-2}(x) = 2 + 2x$, and $\boldsymbol{\delta}_{\text{vec}} := (1, \pi/(2\delta), \pi^2/(2\delta)^2, \ldots, \pi^{n-1}/(2\delta)^{n-1})$. □

B

Appendix: Additional Quantum Channel Properties

Definition B.0.1: Hadamard form

We say that Φ can be written in *Hadamard form* if either of the following equivalent conditions is met:

1. The quantum channel Φ^c is entanglement-breaking;
2. We can write
$$\Phi(\rho) = \sum_{jk} |e_j\rangle\langle e_k| \ \langle x_j|x_k\rangle \ \langle w_j|\rho|w_k\rangle = X * W_\rho, \tag{B.1}$$
where $\{|e_i\rangle\}$ is an orthonormal basis of $\mathcal{H}_\mathrm{B}$, $\{|x_i\rangle\} \subset \mathcal{H}_\mathrm{E}$, and $\{|w_i\rangle\} \subset \mathcal{H}_\mathrm{A}$. We put this data into a matrix X with coefficients $[X]_{jk} = \langle x_j|x_k\rangle$ and a matrix W_ρ with coefficients $[W_\rho]_{jk} = \langle w_j|\rho|w_k\rangle$, and we can then describe $\Phi(\rho) = X * W_\rho$, where $*$ is the Hadamard (or Schur) element-wise product.

We say that Φ is a *Hadamard (diagonal) channel* if the vectors $\{|w_k\rangle\} \subset \mathcal{H}_\mathrm{A}$ form an orthonormal basis of $\mathcal{H}_\mathrm{A}$, such that W_ρ is a matrix representation of ρ with respect to this basis. Equivalently, Φ^c is a classical-quantum (CQ) channel (which is an extreme point in the convex set of EBT maps).

Remark B.0.1. The second equivalent condition that allows Φ to be written in *Hadamard form* is implied by the first condition as follows: as Φ^c is entanglement breaking, we can write $\Phi^c(\rho) = \sum_{k=1}^d |x_k\rangle\langle x_k| \langle w_k|\rho|w_k\rangle$ for some $|x_k\rangle \in \mathcal{H}_\mathrm{E}$ and $|w_k\rangle \in \mathcal{H}_\mathrm{A}$. Hence, its Kraus operators F_k are given by $F_k = |x_k\rangle\langle w_k|$. From this, it follows that $\Phi(\rho) = \sum_{jk} |e_j\rangle\langle e_k| \mathrm{Tr}(F_j \rho F_k^*)$, where $\{|e_k\rangle\}_{k=1}^d$ is an orthonormal basis of $\mathcal{H}_\mathrm{B} \simeq \mathbb{C}^d$. Writing out the trace yields $\langle x_j|x_k\rangle \langle w_j|\rho|w_k\rangle$ and we recover the second equivalent condition.

Definition B.0.2: Unital or bistochastic

We say that Φ is *bistochastic* if it sends the maximally mixed input state $\rho = \frac{1}{\dim \mathcal{H}_\mathrm{A}} \iota_{\mathcal{H}_\mathrm{A}}$ to the maximally mixed output state $\frac{1}{\dim \mathcal{H}_\mathrm{B}} \iota_{\mathcal{H}_\mathrm{B}}$. In the case $\dim \mathcal{H}_\mathrm{A} = \dim \mathcal{H}_\mathrm{B}$, this is equivalent to $\Phi(\iota) = \iota$, and in this case Φ is also called *unital*.

Theorem B.0.1: Unital qubit channels are strongly MOE-additive

Given any *unital qubit channel* $\Phi : \mathcal{B}(\mathbb{C}^2) \to \mathcal{B}(\mathbb{C}^2)$. Then, for *any arbitrary* other channel

Ψ (note that this channel does *not* have to be unital, or a qubit channel!), we have

$$H_{\min}(\Phi \otimes \Psi) = H_{\min}(\Phi) + H_{\min}(\Psi), \tag{B.2}$$

Proof. Theorem 1 in [Kin02a]. □

Theorem B.0.2: Depolarizing channel is strongly MOE-additive and χ-additive

The depolarizing channel is a quantum channel $\Delta_\lambda : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ for some finite-dimensional Hilbert space $\mathcal{H}$ and a real parameter $-\frac{1}{d^2-1} \leqslant \lambda \leqslant 1$ with $d := \dim \mathcal{H}$, given by $\Delta_\lambda(\rho) = \lambda\rho + \frac{1-\lambda}{d}\iota_{\mathcal{H}}$. For *any other channel* Ψ, we have

$$H_{\min}(\Delta_\lambda \otimes \Psi) = H_{\min}(\Delta_\lambda) + H_{\min}(\Psi). \tag{B.3}$$

and

$$\chi(\Delta_\lambda \otimes \Psi) = \chi(\Delta_\lambda) + \chi(\Psi). \tag{B.4}$$

Proof. Theorems 2 and 3 in [Kin03], the strong MOE-additivity for $0 < \lambda < 1$ was already established earlier in [AH01]. Also note that Δ_λ is entanglement-breaking if and only if it is PPT (in general, we only know that EBT implies PPT, see Lemma 3.5.1 below), if and only if $\frac{-1}{d^2-1} \leqslant \lambda \leqslant \frac{1}{d+1}$, which was already proven by the brothers Horodecki [HH99], and from Theorem 3.5.1 we know that EBT channels are strongly MOE-additive. Also see [LH16]. □

Theorem B.0.3: Channels in Hadamard form are strongly MOE-additive

Given any channel Φ that can be written in Hadamard form. Then, for *any arbitrary* other channel Ψ (note that this channel does *not* have to be Hadamard!), we have

$$H_{\min}(\Phi \otimes \Psi) = H_{\min}(\Phi) + H_{\min}(\Psi). \tag{B.5}$$

Proof. The Hadamard diagonal case was proven by King [Kin02b, Kin04], the general case was proven by King, Matsumoto, Nathanson and Ruskai in [KMNR05], as this theorem directly follows from their Theorem 5 (see Theorem 3.5.2 in this thesis) using the fact that the complementary channel of a channel that has a Hadamard form is an entanglement-breaking channel, and EBT channels are strongly MOE-additive by Theorem 3.5.1. □

Bibliography

[AH01] G. Amosov and A. Holevo. On the Multiplicativity Conjecture for Quantum Channels. *arXiv e-prints*, pages math–ph/0103015, March 2001, math-ph/0103015.

[AHT85] W. N. Anderson, E. J. Harner, and G. E. Trapp. Eigenvalues of the Difference and Product of Projections. *Linear and Multilinear Algebra*, 17(3-4):295–299, aug 1985. doi:10.1080/03081088508817661.

[AMS08] P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, dec 2008. doi:10.1515/9781400830244.

[ASW10] G. Aubrun, S. Szarek, and E. Werner. Hastings's Additivity Counterexample via Dvoretzky's Theorem. *Communications in Mathematical Physics*, 305(1):85–97, nov 2010. doi:10.1007/s00220-010-1172-y.

[Aud07] K. M. R. Audenaert. A Sharp Continuity Estimate for the Von Neumann Entropy. *Journal of Physics A: Mathematical and Theoretical*, 40(28):8127–8136, jun 2007. doi:10.1088/1751-8113/40/28/s18.

[Ban98] T. Banica. The Representation Theory of Free Orthogonal Quantum Groups. *C. R. Acad. Sci. Paris Ser. I Math. 322 (1996), 241-244*, June 1998, math/9806063.

[Ban99] T. Banica. Symmetries of a Generic Coaction. *Mathematische Annalen*, 314(4):763–780, aug 1999. doi:10.1007/s002080050315.

[Ban02] T. Banica. Quantum Groups and Fuss–Catalan Algebras. *Communications in Mathematical Physics*, 226(1):221–232, mar 2002. doi:10.1007/s002200200613.

[BBC06] T. Banica, J. Bichon, and B. Collins. Quantum Permutation Groups: A Survey. *Banach Center Publ. 78 (2007), 13-34*, December 2006, math/0612724.

[BC07a] T. Banica and B. Collins. Integration over Compact Quantum Groups. *Publications of the Research Institute for Mathematical Sciences*, 43(2):277–302, 2007. doi:10.2977/prims/1201011782.

[BC07b] T. Banica and B. Collins. Integration over Quantum Permutation Groups. *Journal of Functional Analysis*, 242(2):641–657, jan 2007. doi:10.1016/j.jfa.2006.09.005.

[BC16a] M. Brannan and B. Collins. Dual Bases in Temperley-Lieb Algebras, Quantum Groups, and a Question of Jones. volume 9, pages 715–748. European Mathematical Society - EMS - Publishing House GmbH, 9 2016, arXiv:1608.03885v2[math.QA]. doi:10.2307/24715771.

[BC16b] M. Brannan and B. Collins. Highly Entangled, Non-Random Subspaces of Tensor Products from Quantum Groups. *Comm. Math. Phys. 358 (2018), no. 3, 1007-1025*, 358(3):1007–1025, December 2016, 1612.09598. doi:10.1007/s00220-017-3023-6.

[BC20] M. Brannan and B. Collins. Entanglement and the Temperley-Lieb category. *Topological Phases of Matter and Quantum Computation*, 747:27–50, 2020. doi:10.1090/conm/747/15037.

[BCLY20] M. Brannan, B. Collins, H. H. Lee, and S.-G. Youn. Temperley-Lieb Quantum Channels. *Communications in Mathematical Physics*, 376(2):795–839, may 2020. doi:10.1007/s00220-020-03731-2.

[BCN16] S. T. Belinschi, B. Collins, and I. Nechita. Almost One Bit Violation for the Additivity of the Minimum Output Entropy. *Communications in Mathematical Physics*, 341(3):885–909, jan 2016. doi:10.1007/s00220-015-2561-z.

[BH10] F. G. S. L. Brandão and M. Horodecki. On Hastings' Counterexamples to the Minimum Output Entropy Additivity Conjecture. *Open Systems & Information Dynamics*, 17(01):31–52, mar 2010. doi:10.1142/s1230161210000047.

[BH15] F. G. S. L. Brandao and A. W. Harrow. Estimating Operator Norms using Covering Nets. *arXiv e-prints*, page arXiv:1509.05065, September 2015, 1509.05065.

[Bla06] B. Blackadar. *Operator Algebras*. Springer-Verlag GmbH, March 2006. URL https://www.ebook.de/de/product/11430427/bruce_blackadar_operator_algebras.html.

[BLS19] G. Burrull, N. Libedinsky, and P. Sentinelli. p-Jones-Wenzl idempotents. *Advances in Mathematics*, 352:246–264, aug 2019. doi:10.1016/j.aim.2019.06.005.

[BMAS14] N. Boumal, B. Mishra, P.-A. Absil, and R. Sepulchre. Manopt, a Matlab Toolbox for Optimization on Manifolds. *Journal of Machine Learning Research*, 15(42):1455–1459, 2014. URL https://www.manopt.org.

[Bou20] N. Boumal. An introduction to Optimization on Smooth Manifolds. *Available online*, 3, May 2020. URL http://www.nicolasboumal.net/book.

[Bra12] M. Brannan. Reduced Operator Algebras of Trace-Preserving Quantum Automorphism Groups. *arXiv e-prints*, page arXiv:1202.5020, February 2012, 1202.5020.

[BS07] S. Beigi and P. W. Shor. On the Complexity of Computing Zero-Error and Holevo Capacity of Quantum Channels. September 2007, 0709.2090.

[BS09] T. Banica and R. Speicher. Liberation of Orthogonal Lie Groups. *Advances in Mathematics*, 222(4):1461–1501, nov 2009. doi:10.1016/j.aim.2009.06.009.

[BSA18] J. Belletête and Y. Saint-Aubin. Fusion and Monodromy in the Temperley-Lieb Category. *SciPost Physics*, 5(4), oct 2018. doi:10.21468/scipostphys.5.4.041.

[BV09] T. Banica and R. Vergnioux. Fusion Rules for Quantum Reflection Groups. *Journal of Noncommutative Geometry*, pages 327–359, 2009. doi:10.4171/jncg/39.

[BW03] K. Böröczky and G. Wintsche. Covering the Sphere by Equal Spherical Balls. In *Algorithms and Combinatorics*, pages 235–251. Springer Berlin Heidelberg, 2003. doi:10.1007/978-3-642-55566-4_10.

[CFN12] B. Collins, M. Fukuda, and I. Nechita. Towards a State Minimizing the Output Entropy of a Tensor Product of Random Quantum Channels. *Journal of Mathematical Physics*, 53(3):032203, mar 2012. doi:10.1063/1.3695328.

[Cla06] K. L. Clarkson. Building Triangulations Using ϵ-nets. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing - STOC '06*. ACM Press, 2006. doi:10.1145/1132516.1132564.

[CS11] S. Curran and R. Speicher. Quantum Invariant Families of Matrices in Free Probability. *Journal of Functional Analysis*, 261(4):897–933, aug 2011. doi:10.1016/j.jfa.2011.04.004.

[CY19] B. Collins and S.-G. Youn. Additivity Violation of the Regularized Minimum Output Entropy. July 2019, 1907.07856.

[Del09] I. Dell'Ambrogio. Categories of C^*-algebras. *An introduction to C^*-algebras*, 2009. URL https://math.univ-lille1.fr/~dellambr/exercise_C_algebras.pdf.

[Don97] D. L. Donoho. Wald Lecture I: Counting Bits with Kolmogorov and Shannon. 1997.

[DPP06] I. Dumer, M. S. Pinsker, and V. V. Prelov. On the Thinnest Coverings of Spheres and Ellipsoids with Balls in Hamming and Euclidean Spaces. In *Lecture Notes in Computer Science*, pages 898–925. Springer Berlin Heidelberg, 2006. doi:10.1007/11889342_57.

[DPS02] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri. Distinguishing Separable and Entangled States. *Physical Review Letters*, 88(18):187904, apr 2002. doi:10.1103/physrevlett.88.187904.

[DPS04] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri. A Complete Family of Separability Criteria. *Physical Review A*, 69(2):022308, feb 2004. doi:10.1103/physreva.69.022308.

[Dri88] V. G. Drinfel'd. Quantum Groups. *Journal of Soviet Mathematics*, 41(2):898–915, apr 1988. doi:10.1007/bf01247086.

[Edw73] C. H. Edwards. *Advanced Calculus of Several Variables*. Elsevier Science & Techn., May 1973. doi:10.1016/c2013-0-10610-1.

[EK95] R. Eberhart and J. Kennedy. A New Optimizer Using Particle Swarm Theory. In *MHS'95. Proceedings of the Sixth International Symposium on Micro Machine and Human Science*. IEEE, 1995. doi:10.1109/mhs.1995.494215.

[ES75] M. Enock and J.-M. Schwartz. Une Dualité Dans Les Algèbres de von Neumann. *Mémoires de la Société Mathématique de France*, 1:5–144, 1975. doi:10.24033/msmf.179.

[ES01] Eberhart and Y. Shi. Particle Swarm Optimization: Developments, Applications and Resources. In *Proceedings of the 2001 Congress on Evolutionary Computation (IEEE Cat. No.01TH8546)*. IEEE, 2001. doi:10.1109/cec.2001.934374.

[FKM10] M. Fukuda, C. King, and D. K. Moser. Comments on Hastings' Additivity Counterexamples. *Communications in Mathematical Physics*, 296(1):111–143, feb 2010. doi:10.1007/s00220-010-0996-9.

[FN19] M. Fukuda and I. Nechita. Enumerating Meandric Systems with Large Number of Loops. *Annales de l'Institut Henri Poincaré D*, 6(4):607–640, sep 2019. doi:10.4171/aihpd/80.

[FTMS03] M. Fujiwara, M. Takeoka, J. Mizuno, and M. Sasaki. Exceeding the Classical Capacity Limit in a Quantum Optical Channel. *Physical Review Letters*, 90(16):167906, apr 2003. doi:10.1103/physrevlett.90.167906.

[FW16] A. Freslon and M. Weber. On the Representation Theory of Partition (Easy) Quantum Groups. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, 2016(720), jan 2016. doi:10.1515/crelle-2014-0049.

[GB02] L. Gurvits and H. Barnum. Largest Separable Balls Around the Maximally Mixed Bipartite Quantum State. *Physical Review A*, 66(6):062311, dec 2002. doi:10.1103/physreva.66.062311.

[GW19] D. Gromada and M. Weber. Intertwiner Spaces of Quantum Group Subrepresentations. *Communications in Mathematical Physics*, 376(1):81–115, may 2019. doi:10.1007/s00220-019-03463-y.

[Has09] M. B. Hastings. Superadditivity of Communication Capacity using Entangled Inputs. *Nature Physics*, 5(4):255–257, mar 2009. doi:10.1038/nphys1224.

[HH99] M. Horodecki and P. Horodecki. Reduction Criterion of Separability and Limits for a Class of Distillation Protocols. *Physical Review A*, 59(6):4206–4216, jun 1999. doi:10.1103/physreva.59.4206.

[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, jul 2004. doi:10.1007/s00220-004-1087-6.

[HMS16] D. P. Hardin, T. J. Michaels, and E. B. Saff. A Comparison of Popular Point Configurations on $\mathbb{S}^2$. *arXiv e-prints*, page arXiv:1607.04590, July 2016, 1607.04590.

[HNW17] A. W. Harrow, A. Natarajan, and X. Wu. An Improved Semidefinite Programming Hierarchy for Testing Entanglement. *Communications in Mathematical Physics*, 352(3):881–904, mar 2017. doi:10.1007/s00220-017-2859-0.

[Hol05] A. S. Holevo. Additivity Conjecture And Covariant Channels. *International Journal of Quantum Information*, 03(01):41–47, mar 2005. doi:10.1142/s0219749905000530.

[Hol07] A. S. Holevo. Complementary Channels and the Additivity Problem. *Theory of Probability & Its Applications*, 51(1):92–100, jan 2007. doi:10.1137/s0040585x97982244.

[HSR03] M. Horodecki, P. W. Shor, and M. B. Ruskai. Entanglement Breaking Channels. *Reviews in Mathematical Physics*, 15(06):629–641, aug 2003. doi:10.1142/s0129055x03001709.

[Jan10] B. Janssens. Transformation & Uncertainty. Some Thoughts On Quantum Probability Theory, Quantum Statistics, And Natural Bundles. November 2010, http://arxiv.org/abs/1011.3035v1.

[JCR16] N. Johnston, A. Cosentino, and V. Russo. Qetlab: Qetlab V0.9, 2016. doi:10.5281/ZENODO.44637.

[Jon83] V. F. R. Jones. Index for Subfactors. *Inventiones Mathematicae*, 72(1):1–25, feb 1983. doi:10.1007/bf01389127.

[Kac63] G. I. Kac. Ring Groups and the Duality Principle. *Transactions of the Moscow Mathematical Society*, 12:259–301, 1963. URL http://www.mathnet.ru/php/archive.phtml?wshow=paper&jrnid=mmo&paperid=142&option_lang=eng.

[Kac65] G. I. Kac. Ring Groups and the Duality Principle. II. *Trudy Moskovskogo Matematicheskogo Obshchestva*, 13:84–113, 1965. URL http://www.mathnet.ru/php/archive.phtml?wshow=paper&jrnid=mmo&paperid=151&option_lang=eng.

[Kas95] C. Kassel. *Quantum Groups*. Springer New York, 1995. doi:10.1007/978-1-4612-0783-2.

[Kau01] L. H. Kauffman. *Knots and Physics*. WORLD SCIENTIFIC, jul 2001. doi:10.1142/4256.

[KE95] J. Kennedy and R. Eberhart. Particle Swarm Optimization. In *Proceedings of ICNN'95 - International Conference on Neural Networks*. IEEE, 1995. doi:10.1109/icnn.1995.488968.

[Kin02a] C. King. Additivity for Unital Qubit Channels. *Journal of Mathematical Physics*, 43(10):4641, 2002. doi:10.1063/1.1500791.

[Kin02b] C. King. Maximization of Capacity and L_p Norms for Some Product Channels. *Journal of Mathematical Physics*, 43(3):1247–1260, mar 2002. doi:10.1063/1.1433943.

[Kin03] C. King. The Capacity of the Quantum Depolarizing Channel. *IEEE Transactions on Information Theory*, 49(1):221–229, jan 2003. doi:10.1109/tit.2002.806153.

[Kin04] C. King. An Application of a Matrix Inequality in Quantum Information Theory. *arXiv e-prints*, pages quant–ph/0412046, December 2004, quant-ph/0412046.

[KL94] L. H. Kauffman and S. Lins. *Temperley-Lieb Recoupling Theory and Invariants of 3-Manifolds (AM-134)*. Princeton University Press, dec 1994. doi:10.1515/9781400882533.

[KMNR05] C. King, K. Matsumoto, M. Nathanson, and M. B. Ruskai. Properties of Conjugate Channels with Applications to Additivity and Multiplicativity. *arXiv e-prints*, pages quant–ph/0509126, September 2005, quant-ph/0509126.

[KR01] C. King and M. Ruskai. Minimal Entropy of States Emerging From Noisy Quantum Channels. *IEEE Transactions on Information Theory*, 47(1):192–209, 2001. doi:10.1109/18.904522.

[KSW08] D. Kretschmann, D. Schlingemann, and R. F. Werner. The Information-Disturbance Tradeoff and the Continuity of Stinespring' s Representation. *IEEE Transactions on Information Theory*, 54(4):1708–1717, apr 2008. doi:10.1109/tit.2008.917696.

[LH16] L. Lami and M. Huber. Bipartite Depolarizing Maps. *Journal of Mathematical Physics*, 57(9):092201, sep 2016. doi:10.1063/1.4962339.

[Lus10] G. Lusztig. *Introduction to Quantum Groups*. Birkhäuser Boston, 2010. doi:10.1007/978-0-8176-4717-9.

[LVWB21] X. Li, C. Voigt, M. Weber, and C. Budde. C^*-algebras and Dynamics, Lecture Notes. 2021.

[LWJ+05] B. Liu, L. Wang, Y.-H. Jin, F. Tang, and D.-X. Huang. Improved Particle Swarm Optimization Combined With Chaos. *Chaos, Solitons & Fractals*, 25(5):1261–1271, sep 2005. doi:10.1016/j.chaos.2004.11.095.

[MD98] A. Maes and A. V. Daele. Notes on Compact Quantum Groups. March 1998, math/9803122.

[Mor17] S. Morrison. A Formula for the Jones-Wenzl Projections. In *Proceedings of the 2014 Maui and 2015 Qinhuangdao conferences in honour of Vaughan F. R. Jones' 60th birthday*, pages 367–378. Canberra: Australian National University, Centre for Mathematics and its Applications, March 2017, 1503.00384.

[MT15] O. R. Musin and A. S. Tarasov. The Tammes Problem for $N = 14$. *Experimental Mathematics*, 24(4):460–468, jul 2015. doi:10.1080/10586458.2015.1022842.

[Mur14] G. J. Murphy. *C*-Algebras and Operator Theory*. Elsevier, June 2014. doi:10.1016/c2009-0-22289-6.

[Nat02] M. B. Nathanson. Additive Number Theory and the Ring of Quantum Integers. *arXiv Mathematics e-prints*, page math/0204006, March 2002, math/0204006.

[NC09] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2009. doi:10.1017/CBO9780511976667.

[Neg71] J. W. Negrepontis. Duality in Analysis from the Point of View of Triples. *Journal of Algebra*, 19(2):228–253, oct 1971. doi:10.1016/0021-8693(71)90105-0.

[Noe12] M. M. Noel. A New Gradient Based Particle Swarm Optimization Algorithm for Accurate Computation of Global Minimum. *Applied Soft Computing*, 12(1):353–359, jan 2012. doi:10.1016/j.asoc.2011.08.037.

[NOP09] M. Navascués, M. Owari, and M. B. Plenio. Complete Criterion for Separability Detection. *Physical Review Letters*, 103(16):160404, oct 2009. doi:10.1103/physrevlett.103.160404.

[NS06] A. Nica and R. Speicher. *Lectures on the Combinatorics of Free Probability*. Cambridge University Press, 2006. doi:10.1017/cbo9780511735127.

[Nuw13] M. A. Nuwairan. Potential Examples for Non-Additivity of the Minimal Output Entropy. *arXiv e-prints*, page arXiv:1312.2200, December 2013, 1312.2200.

[Nuw14] M. A. Nuwairan. The Extreme Points of SU_2-Irreducibly Covariant Channels. *International Journal of Mathematics*, 25(06):1450048, jun 2014. doi:10.1142/s0129167x14500487.

[ON00] S. Osawa and H. Nagaoka. Numerical Experiments on The Capacity of Quantum Channel with Entangled Input States. *IEICE Trans. Fundamentals, vol.E84-A, no.10, pp.2583-2590, Oct. 2001*, July 2000, quant-ph/0007115.

[Rog63] C. A. Rogers. Covering a Sphere with Spheres. *Mathematika*, 10(2):157–164, dec 1963. doi:10.1112/s0025579300004083.

[SF99] R. P. Stanley and S. Fomin. *Enumerative Combinatorics*. Cambridge University Press, jan 1999. doi:10.1017/cbo9780511609589.

[Sho02] P. W. Shor. Additivity of the Classical Capacity of Entanglement-Breaking Quantum Channels. *Journal of Mathematical Physics*, 43(9):4334–4340, sep 2002. doi:10.1063/1.1498000.

[Sho04] P. W. Shor. Equivalence of Additivity Questions in Quantum Information Theory. *Communications in Mathematical Physics*, 246(3):453–472, apr 2004. doi:10.1007/s00220-003-0981-7.

[SK97] E. B. Saff and A. B. J. Kuijlaars. Distributing Many Points on a Sphere. *The Mathematical Intelligencer*, 19(1):5–11, dec 1997. doi:10.1007/bf03024331.

[SQ15] B. L. Stum and A. Quirós. On Quantum Integers and Rationals, 2015. doi:10.1090/conm/649/13022.

[Sti55] W. F. Stinespring. Positive Functions on C $*$ -Algebras. *Proceedings of the American Mathematical Society*, 6(2):211, apr 1955. doi:10.2307/2032342.

[SW12] Y. Shi and X. Wu. Epsilon-Net Method for Optimizations over Separable States. In *Automata, Languages, and Programming*, pages 798–809. Springer Berlin Heidelberg, 2012. doi:10.1007/978-3-642-31594-7_67.

[Tak72] M. Takesaki. Duality and von Neumann algebras. In *Lectures on Operator Algebras*, pages 665–779. Springer Berlin Heidelberg, 1972. doi:10.1007/bfb0058558.

[Tim08] T. Timmermann. *An Invitation to Quantum Groups and Duality*. European Mathematical Society Publishing House, feb 2008. doi:10.4171/043.

[vDW96] A. van Daele and S. Wang. Universal Quantum Groups. *International Journal of Mathematics*, 07(02):255–263, apr 1996. doi:10.1142/s0129167x96000153.

[VG04] J.-L. Verger-Gaugry. Covering a Ball with Smaller Equal Balls in $\mathbb{R}^n$. *Discrete & Computational Geometry*, 33(1):143–155, jun 2004. doi:10.1007/s00454-004-2916-2.

[Wan98] S. Wang. Quantum Symmetry Groups of Finite Spaces. *Communications in Mathematical Physics*, 195(1):195–211, jul 1998. doi:10.1007/s002200050385.

[Wat18] J. Watrous. *The Theory of Quantum Information*. Cambridge University Press, April 2018. URL https://www.ebook.de/de/product/33640295/john_watrous_theory_of_quantum_information.html.

[WBS14] T. P. Wihler, B. Bessire, and A. Stefanov. Computing the Entropy of a Large Matrix. *Journal of Physics A: Mathematical and Theoretical*, 47(24):245201, jun 2014. doi:10.1088/1751-8113/47/24/245201.

[Web13] M. Weber. On the Classification of Easy Quantum Groups. *Advances in Mathematics*, 245:500–533, oct 2013. doi:10.1016/j.aim.2013.06.019.

[Web17a] M. Weber. Introduction to Compact (Matrix)Qquantum Groups and Banica-Speicher (Easy) Quantum Groups. *Proceedings - Mathematical Sciences*, 127(5):881–933, nov 2017. doi:10.1007/s12044-017-0362-3.

[Web17b] M. Weber. Partition C^*-algebras. October 2017, 1710.06199.

[Wil13] M. M. Wilde. *Quantum Information Theory*. Cambridge University Press, Cambridge, UK, 2013. doi:10.1017/CBO9781139525343.

[Wor80] S. L. Woronowicz. Pseudospaces, Pseudogroups and Pontriagin Duality. In *Mathematical Problems in Theoretical Physics*, pages 407–412. Springer Berlin Heidelberg, 1980. doi:10.1007/3-540-09964-6_354.

[Wor87] S. L. Woronowicz. Compact Matrix Pseudogroups. *Communications in Mathematical Physics*, 111(4):613–665, dec 1987. doi:10.1007/bf01219077.

[Wor88] S. L. Woronowicz. Tannaka-Krein Duality for Compact Matrix Pseudogroups. Twisted SU_N Groups. *Inventiones Mathematicae*, 93(1):35–76, feb 1988. doi:10.1007/bf01393687.

[Wor91] S. L. Woronowicz. A remark on Compact Matrix Quantum Groups. *Letters in Mathematical Physics*, 21(1):35–39, jan 1991. doi:10.1007/bf00414633.

[Wor92] S. L. Woronowicz. Compact Quantum Groups. 1992.

[Wyn67] A. D. Wyner. Random Packings and Coverings of the Unit n-Sphere. *Bell System Technical Journal*, 46(9):2111–2118, nov 1967. doi:10.1002/j.1538-7305.1967.tb04246.x.

[YMP+20] S. Yu, Y. Meng, R. B. Patel, Y.-T. Wang, Z.-J. Ke, W. Liu, Z.-P. Li, Y.-Z. Yang, W.-H. Zhang, J.-S. Tang, C.-F. Li, and G.-C. Guo. Experimental Observation of Coherent-Information Superadditivity in a Dephrasure Channel. *Physical Review Letters*, 125(6):060502, aug 2020. doi:10.1103/physrevlett.125.060502.