Studies in Classical Simulation vs Quantum Advantage for Algorithms and Communication

Vojtěch Havlíček

Keble College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Hillary 2020

This thesis is dedicated to my family.

Acknowledgements

First and foremost, I want to thank Jon Barrett for agreeing to supervise my doctorate after my previous supervisor moved to industry. Words can hardly describe how grateful I am for the chance to work under his academic guidance. Jon gave me a surprising amount of freedom as well as trust by supporting my decision to follow a project that I developed on my own right from the beginning of my doctorate. While the work seemed promising and eventually lead to publications, it ended up far from what I intended it to be. Jon supported me at that time though and helped me to develop a different project that overlapped better with our common research interests, for which I am enormously thankful. This project resulted in a work on communication complexity that was seriously interesting, even though the results I achieved were also somewhat far from my expectations. I still hope that I will be able to close some of the loose ends in the project in the near future.

I want to thank my previous supervisor, Matthias Troyer, for accepting me in his group at ETH Zurich where I spent an enjoyable year. My time at ETH was relatively short, but aside from learning to write papers, I had a chance to meet many remarkable people there. Of them, Alexey Soluyanov, was a great inspiration to me during that time. We kept close contact even after I left ETH until his death a year ago. I regret that we did not finish the research project we started but sincerely hope that I will be able to return to it at some point after the DPhil.

Many thanks belongs to the entire Quantum Computation theory team at IBM Research at Yorktown Heights for giving me the opportunity to spend two enjoyable summers interning for them. I would especially like to thank Kristan Temme for his support and Jay Gambetta for suggesting me to apply for an internship after I gave a talk at the T. J. Watson Lab. It is still hard to believe that the research work initiated during my internship eventually made the cover of Nature. I am also enormously

grateful for the IBM PhD Fellowship that significantly helped in the later stages of my work and paid the bills when I was finishing this thesis.

None of this would happen if James D. Whitfield did not invite me to visit Dartmouth and suggested to visit IBM Research. Many thanks belongs to him for all the nice time I spent at the Wilder Lab.

I want to thank Sergii Strelchuk for his collaboration and for a lot of stimulating discussions, all the Cambridge visits and help with the articles on Permutational Quantum Computing.

A small, but a very important nudge of help came from Scott Aaronson who decided to loop in Greg Kuperberg, Adam Bouland, Saeed Mehraban and Stephen Jordan to our email exchange about the Ball Permutation model of computation, that eventually resulted in the Permutational Quantum Computing project.

I want to thank Richard Jozsa and Samson Abramsky for kindly agreeing to examine my thesis.

I would also like to, in no particular order, acknowledge the following people for their support and/or inspiration during my doctorate: Terry Rudolph, Bob Coecke, Damian Steiger, Subhayan Roy Moulik, Konstantinos Meichanetzidis, Robin Lorenz, Leo Wossning, Juani Bermejo-Vega, Gorjan Alagic, Greg Kuperberg, Steve Simon, Juan-Miguel Arrazzola, Chris Weedbrook, Peter D. Johnson, Pierre Luc Dalarie-Demers, Ryan Babbush, Jarrod McClean, Matt Pusey, Matty Hoban, Sivert Aasnæss, Will Kirby, Nathan Wiebe, Alex Moylett, Raul Garcia-Patron, Peter Turner, Ashley Montanaro, Aram Harrow, and Giuseppe Carleo.

I would also like to acknowledge the support from the Clarendon and Keble de Breyne scholarships and the IBM PhD Fellowship.

Contents

Abstract

This thesis studies computational advantages that could be achieved by using quantum resources in two different contexts.

The first part studies Permutational Quantum Computing, a quantum computational model developed to study computation by spin recoupling. It was a conjecture, stated by Jordan in [Quantum Inf. Comput., 10, 470–497 (2010)], that the model solves problems that have no efficient classical algorithms by approximating its transition amplitudes to polynomial precision. Here we give a classical algorithm that disproves this conjecture and also prove that the computational complexity class PQP, defined to capture the computational power of the model in this regime, is entirely classical. The simulation algorithms are derived by utilizing the computational tractability framework of van den Nest [Quantum Inf. Comput., 9–10, 784–812 (2011)]. We look for other sources of quantum advantage in the computational model and rule out the possibility that it encodes answers to computationally hard problems in patterns of small number of significantly large output probabilities. This analysis extends the results of Schwarz and van den Nest [arXiv:1310.6749].

The second part of the thesis studies advantages achievable in the context of communication complexity. We take inspiration in the proof of the Pusey-Barrett-Rudolph theorem and devise a simple communication task based on the concept of quantum state antidistinguishability. We show that there is an exponential separation between one-way exact classical and quantum communication complexities for the task, but also that the separation for this particular task disappears if we allow for interaction or require only a bounded error solution. We conclude by stating and numerically justifying a conjecture about quantum state antidistinguishability that could have some implications for quantum foundations.

Introduction

Thesis Overview

One of the central goals of Quantum Information Theory is to understand how quantum devices could outperform their classical counterparts in terms of computational power or communication costs. This thesis addresses this question in two different contexts and achieves two distinct results.

Part I

Topological Quantum Computing is a fascinating model of quantum computation that computes by splitting, braiding, and fusing quasiparticles called non-abelian anyons. I studied this computational model at the beginning of my PhD with the aim to adapt the quantum supremacy argument of Bremner, Montanaro and Shepherd [1] in this context. Unfortunately, I was not fast enough to prove the results before Mann and Bremner posted Ref. [2] that did essentially the same thing and I had to look for alternative research directions.

This brought me to study Permutational Quantum Computing – a toy model of Topological Quantum Computing studied in Ref. [3] that forgets the topology of the quasiparticle trajectories but nevertheless performs nontrivial computation by particle permutations. The model was originally developed with the aim to study the computational power of spin networks in Refs. [4, 5] and while it was thought to solve problems without efficient classical solution, the supporting evidence for this was weak.

I first attempted to strengthen this evidence by proving a quantum computational supremacy argument for the model which ultimately failed as the model was too discrete for the usual proof techniques. I therefore turned to literature on classical simulation of quantum computations and, following a recommendation of Kristan Temme, stumbled upon the work of van den Nest in Ref. [6] that was a perfect fit for

analysis of the hardness conjecture. Additionally, Sergii Strelchuk got interested in the project at that time and being able to discuss the work with him regularly made the work substantially easier. We eventually found a classical algorithm that ruled out the conjecture about the model's non-classicality and extended it to show that the computational complexity class PQP, characterizing a large class of efficiently solvable problems in the model, can be simulated on a conventional computer. The results are presented in Chapter 4.

It however seemed that the class PQP captured the computational power of Permutational Quantum Computing inadequately. While its definition was partly justified by BQP-completeness of an analogous algorithm in Topological Quantum Computing (the Aharonov-Jones-Landau algorithm [7]), showing that PQP is classical did not preclude the possibility that the model could achieve computational advantage by other means. It remained open if there is a classical algorithm that samples the output distributions of the model or finds all large transition amplitudes in polynomial time and we study this possibility in Chapter 5. This work is largely an extension of the results of Schwarz and van den Nest from Ref. [8].

Part II

This part studies a problem at the intersection of communication complexity and quantum foundations. It had a two-fold motivation. First, it was an attempt to find a problem to collaborate on with my supervisor. Second, the aim was to prove a separation between the power of quantum and classical resources from very simple principles in a way that would minimally depend on conjectures.

This effort resulted in a very simple and self-contained proof of an exponential separation in exact one-way communication complexity that we derived from the foundational concept of quantum state antidistinguishability. Using connections between symmetric, informationally complete sets, mutually unbiased bases and complex spherical codes, we first obtain nontrivial complexity separations for the communication task in small dimensions. We then give a spherical packing argument that leads to the asymptotically exponential separation, which may be of independent interest. We subsequently fully characterize the communication complexity of the communication task and prove that the separations disappear in interactive protocols or

if we require a bounded error solution. The key advantage of our construction is that it is, compared to the previous proofs, rather simple and essentially self-contained.

We conclude by stating a conjecture about a sufficient condition for quantum state antidistinguishability. If the conjecture holds, it implies $\Omega(d \log d)$ vs $O(\log d)$ separation of quantum and classical exact communication complexities, which is so far unproven as far as we are aware. We support the conjecture with numerical evidence and briefly remark on some of its potential consequences for the Pusey-Barrett-Rudolph theorem from quantum foundations.

Excluded Work

I worked on other projects throughout my DPhil that lead to publications.

The work on proof-of-principle classification using parametrized quantum circuits (Ref. [9]) was initiated during my internship at IBM Research. The assumption is that despite IBM allowed me to do so, I cannot include it in this thesis.

Another line of work on fermion-to-qubit mappings (Refs. [10, 11]) was initiated while I was transferring from ETH Zurich, but most of it was completed while I was a student at Oxford. I felt that the work does not play well with the other topics and excluded it for that reason.

Summary of Contributions

This thesis is based on the following work:

1. Havlíček, V., Barrett, J.,
   Simple Communication Complexity Separation from Quantum State Antidistinguishability
   Phys. Rev. Research, 2, 1, 013326, 2020, ArXiv: arxiv.org/abs/1911.01927

2. Havlíček, V., Strelchuk, S., Temme, K.,
   Classical Algorithm for $SU(2)$ Schur Sampling
   Phys. Rev. A 99, 062336, 2019, ArXiv: arxiv.org/abs/1809.05171

3. Havlíček, V., Strelchuk, S.,
   Quantum Schur Sampling Circuits can be Strongly Simulated,
   Phys. Rev. Lett., 121, 6, 060505, 2018, ArXiv: arxiv.org/abs/1801.04795

Other work completed during my DPhil that is not included in the thesis:

1. Havlíček, V., Córcoles, A. D., Temme, K., Harrow, A. W., Kandala, A., Chow, J. M., Gambetta, J. M.,
   Supervised learning with quantum enhanced feature spaces,
   Nature, 567, 141-278, (Cover article), ArXiv: arxiv.org/abs/1804.11326

2. Havlíček, V. and Troyer, M. and Whitfield, J. D.,
   Operator Locality in Quantum Simulation of Fermionic Models,
   Phys. Rev. A 95, 032332, 2017, ArXiv: arxiv.org/abs/1701.07072,

3. McClean, J., et.al.,
   OpenFermion: The Electronic Structure Package for Quantum Computers,
   (Submitted to Quantum Science and Technology), ArXiv: arxiv.org/abs/1710.07629

On Permutational Quantum Computing

Introduction

The effort for building quantum computers is propelled by the belief that they efficiently solve computational tasks beyond the capabilities of classical computers. Not every computational task could however enjoy such quantum advantage and it is an important theoretical challenge to identify those that possibly do.

Suppose that someone proposes a quantum algorithm to solve some computational problem – how can one argue that such solution is genuinely advantageous compared to the solutions achievable by classical means? It would be ideal to *prove* that the quantum algorithm solves the problem in time that scales polynomially with the input size, while there was no classical polynomial time solution. No such problem is however known and is unlikely to be found since ruling out the possibility of classical polynomial-time solution would also imply that $\mathsf{PSPACE} \neq \mathsf{P}$, resolving one of the central problems of computational complexity (see Fig. 1).

It is for that reason that essentially all proposals for significant quantum/classical separations in computational complexity are formulated as conditional separations that hold under plausible conjectures of varying strength.[1] An illustration of this is Shor's factoring algorithm that gives a strong evidence for quantum/classical time complexity separation. The algorithm provides significant computational advantage for factoring integers when compared to the best known classical algorithm, the general number field sieve [13, 14]. While it remains unproven if the sieve is indeed the best classical algorithm for factoring, no classical polynomial algorithm for factoring has been yet found despite decades of intensive search. In fact, the belief that none will be found is put to test with every use of a credit card – believing that

[1] A notable exception are the recent works on shallow depth classical vs. constant depth quantum circuit complexity separations that are unconditional separations of complexity classes contained in P [12].

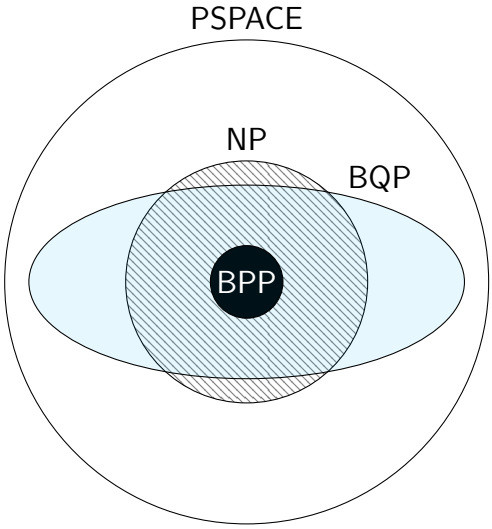


Figure 1: Proving that there is no classical polynomial-time algorithm for a problem that is efficiently solvable in BQP would imply that P ⊆ BPP ⊂ PSPACE, thus resolving one of the major problems in computational complexity. Note that the most interesting subset of BQP problems are those for which it is also possible to efficiently verify correctness of the solution in polynomial time by classical means. Proving an exponential lower bound for classical time complexity of these problems would imply P ⊂ NP. See the above diagram for the conjectured relationship of the computational complexity classes.

one's bank account is safe against classical attacks can be then taken as an argument in favor of quantum computers outperforming the classical ones (albeit in an error-corrected regime and with several thousands or even millions of qubits; see for example Ref. [15]).

From a purely theoretical perspective however, it could be that a classical polynomial-time factoring algorithm will be eventually discovered since the decision version of the problem is likely to be NP-intermediate and if it was in P, it would not have too dramatic consequences for computational complexity theory (see Fig. 2) [13, 16]. It is therefore desirable to instead use conjectures whose violation is implausible from a complexity-theoretic viewpoint. An illustration of this is the approach developed by DiVincenzo, Terhal, and Aaronson, who proved that probability distributions arising from a certain family of constant-depth quantum circuits cannot be exactly sampled from by classical means, unless the polynomial hierarchy collapses [17, 18, 19]. The polynomial hierarchy is an infinite tower of computational complexity classes that are believed to be increasingly more powerful. The 'collapse' means that the classes have the same computational power from a certain point, which contradicts this belief.

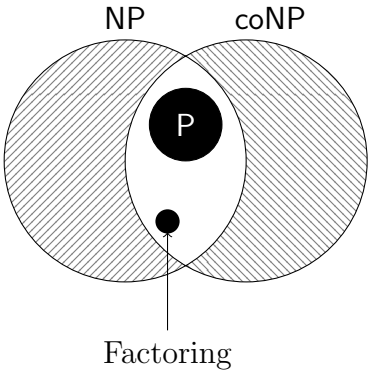


Figure 2: The decision version of the factoring problem is thought to be NP-intermediate. This is because it belongs both to coNP and NP and if it was NP-complete, then coNP = NP which is unlikely. No polynomial time algorithm for factoring has been found despite a lot of effort, so it is usually believed that factoring lies outside of P. Inclusion of factoring in P would however not have too dramatic complexity-theoretical consequences, making the conjecture about general number field sieve optimality rather weak.

As the collapse is often argued to be only marginally less surprising than P = NP, the argument is considered a stronger evidence for hardness of the sampling tasks compared to the hardness conjecture backing Shor's algorithm [16, 20].[2]

The central object of study of this part of the thesis is a quantum algorithm from a class of algorithms for approximating matrix elements of group representations, discovered by Jordan in Ref. [24], which was believed to provide superpolynomial quantum speedup. The evidence backing this conjecture was however much weaker than in the case of Shor's factoring algorithm, as the classical algorithms for these problems are relatively unexplored.[3] An interesting feature of the algorithm, shown in Ref. [3] and outlined in this thesis, is that it can be naturally implemented by approximating the transition amplitudes in Permutational Quantum Computing – a restricted model of quantum computation that computes by spin recoupling.[4] This

[2]Significant progress has been made in the last decade in the context of the 'quantum computational supremacy' research program built around this argument. This effort was ignited by Arkhipov and Aaronson's proposal of Boson Sampling and the IQP-circuits program developed by Bremner, Shepherd, Jozsa and Montanaro [21, 16, 1, 22, 19]. It eventually culminated in the recent claim of verification of quantum computational supremacy put forward by Google last year – a key milestone for quantum computing, as it is the first experimentally verified quantum computational task that seems to transcend the capabilities of classical computers [23].

[3]To make this notion precise, we borrow the following definition from Jordan's Quantum Algorithm Zoo [25]: A problem has a superpolynomial quantum speedup if there exists a constant $\alpha > 0$ such that the runtime $C(n)$ of the best known classical algorithm and the runtime $Q(n)$ of the quantum algorithm satisfy $C = 2^{\Omega(Q^\alpha)}$.

[4]Restricted quantum computational models are models of quantum computation that are be-

computational model, which can be also seen as a toy model of Topological Quantum Computing, was introduced by Marzuoli and Rasetti in Refs. [4, 5] and further formalized by Jordan in Ref. [3]. The conjectured superpolynomial speedup of the quantum algorithm then supported the claim about non-classicality of the model.

The most ambitious goal, which I ultimately failed to achieve, was to find a complexity-theoretical backing for the hardness conjecture. Somewhat surprisingly, the main result derived here ended up on the completely opposite side of the spectrum – it shows that the algorithm can be classically simulated in polynomial time. To that end, we give an efficient classical algorithm that approximates the transition amplitudes of Permutational Quantum Computing to polynomial precision, which suffices to give a classical algorithm for a problem that was believed to be classically hard. The algorithm is then found by using methods from the probabilistic simulation framework developed by van den Nest in Ref. [6].

We extend this result and show that any problem in the computational class PQP, proposed in Ref. [3] to capture the set of problems solvable in the model, is also efficiently simulatable by a classical probabilistic algorithm. We subsequently look for other possible sources of computational advantage in the model and show that no superpolynomial speedup can be achieved by encoding solutions to computationally hard problems by patterns of significantly large output probabilities. This is proved by extending the work of Schwarz and van den Nest [8].

This part of the thesis is structured as follows:

- Chapter 1 – Permutational Quantum Computing can be seen as a toy model of Topological Quantum Computing. This brief chapter outlines the key ideas of this model and sketches the Aharonov-Jones-Landau algorithm with the aim to make this connection apparent later.

- Chapter 2 – Introduces spin eigenfunctions, spin recoupling and the Yamanouchi-Kotani representation needed to introduce the quantum algorithm.

lieved to be computationally weaker than universal quantum computers, but could still solve some problems that show superpolynomial speedups. Examples of such restricted quantum computational models are Permutational Quantum Computing [5, 3], the One Clean Qubit model [26, 27, 28], Ball Permuting model [29] or IQP circuits [21] and Boson Sampling [19].

- Chapter 3 – Introduces Permutational Quantum Computing and the computational complexity class PQP.
- Chapter 4 – Describes a classical algorithm that disproves the conjecture about hardness of the model. This is the main result of this part of the thesis. Based on work with S. Strelchuk [30].
- Chapter 5 – Studies additional consequences of the simulation algorithm. Based on work with S. Strelchuk and K. P. Temme [31].

Chapter 1

Prologue: A Sketch of Topological Quantum Computing

1.1 Introduction

Permutational Quantum Computing can be seen as a toy model of Topological Quantum Computing. This chapter briefly outlines this computational model so that this connection becomes apparent later.

1.2 Anyons

Two fundamental kinds of particles, bosons and fermions, exist in $3+1$ dimensional spacetime. They have either exchange-symmetric or antisymmetric quantum states, which is modeled by one-dimensional representations of the symmetric group. There are only two such representations; the trivial one that corresponds to bosonic particles and the alternating representation that corresponds to fermions.[1]

In Ref. [33], Leinaas and Myrrheim realized that it is possible for 2-dimensional quantum systems to support particle-like excitations with exchange symmetries that are intermediate to fermions and bosons. Roughly stated, this possibility occurs because the trajectory of a particle wound around another one in 2 spatial dimensions cannot be contracted to the starting point without touching the other particle, which leads to a well-defined notion of 'braiding' [33, 32]. Particles of this kind are called anyons as they allow for "any" intermediate exchange phase between fermions and

[1] Particles transforming under higher-dimensional representation of S_n are called parafermions and parabosons. They can be decomposed into fermions and bosons with an additional quantum number attached to each particle [32].

bosons [34, 35, 36, 37]. For example, the wavefunction may instead transform as:

$$\Psi(\vec{r}_1, \vec{r}_2) \to e^{i\theta}\Psi(\vec{r}_2, \vec{r}_1),$$

for some nontrivial phase θ upon a right-handed twist of the two particles [38]. The symmetry group can be then presented by its generators σ_i – right-handed nearest-neighbor twists – that are subject to:

$$\sigma_i\sigma_j = \sigma_j\sigma_i \text{ for } |i-j| \geq 2, \qquad \sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1}; \tag{1.1}$$

This group presentation is very similar to the presentation of symmetric group, except that it does not impose the $\sigma_i^2 = \mathbf{1}$ 'monodromy' condition. It is the Artin braid group B_n.

Anyons are not just a mathematical curiosity but appear in superconducting thin films. The presence of anyons forces the film to have an unusual Hall voltage response to external magnetic fields when a current is running through it – this is called the Fractional Quantum Hall effect. The first explanation of Fractional Quantum Hall in terms of anyonic excitations is due to Laughlin who was awarded the 1998 Nobel prize for this discovery with Störmer and Tsui, who verified the effect experimentally [39, 40].

Certain physical systems may support anyons with symmetries of higher-dimensional representations of the braid group [41, 42, 43, 44, 45, 46]. Quantum states of such systems carry 'topological degeneracy' and the B_n representation matrices implement unitary operations in this space. For example, a d-degenerate three-nonabelion wavefunctions $\Psi_\alpha(\vec{r}_1, \vec{r}_2, \vec{r}_3)$ with degeneracy $\alpha \in [d]$ transform as:

$$\Psi_\alpha(\vec{r}_1, \vec{r}_2, \vec{r}_3) \to \rho(\sigma_1)_{\alpha\beta}\Psi_\beta(\vec{r}_2, \vec{r}_1, \vec{r}_3), \qquad \Psi_\alpha(\vec{r}_1, \vec{r}_2, \vec{r}_3) \to \rho(\sigma_2)_{\alpha\beta}\Psi_\beta(\vec{r}_1, \vec{r}_3, \vec{r}_2),$$

where $\rho(\sigma_1)$ and $\rho(\sigma_2)$ are $d \times d$ unitaries that represent effects of the counterclockwise $1 \leftrightarrow 2$ and $2 \leftrightarrow 3$ exchanges respectively and act on the degeneracy space. The actions generally do not commute: $[\rho(\sigma_1), \rho(\sigma_2)] \neq 0$, in which case the symmetries of such systems realize non-abelian anyons. The proposal of Topological Quantum Computing rests on the fact that it is possible to implement arbitrary unitaries acting on the topological degeneracy space by braiding and measuring multiparticle states of such particles [43, 47, 48, 49, 32].

1.3 Fusion and Splitting

Any anyonic system has at least two particles types, one of which one corresponds to the vacuum. Composing two anyons into a single particle is referred to as fusion, . Two anyons of types ϕ_a and ϕ_b fuse as:

$$\phi_a \times \phi_b = \sum_{k=0}^{c} N_{ab}^{k}\, \phi_k,$$

where the N_{ab}^{c} integer coefficients are called fusion channels. Non-abelian anyonic theories have at least one pair of anyons fusing into multiple fusion channels. The simplest possible example of this, the Ising anyon theory relevant to physics of $\nu = 5/2$ Fraction Quantum Hall effect, consists of three anyon types, σ, ψ and $\mathbf{1}$ (the vacuum), which combine as:

$$\begin{aligned}
&\sigma \times \sigma \to \psi + \mathbf{1}, && \psi \times \psi \to \mathbf{1}, && \sigma \times \psi \to \sigma,\\
&\mathbf{1} \times \mathbf{1} \to \mathbf{1}, && \mathbf{1} \times \sigma \to \sigma, && \mathbf{1} \times \psi \to \psi.
\end{aligned}$$

A dual operation to fusion is called splitting.

1.4 Angular Momentum Analogy

The above fusion rules are analogous to angular momentum recoupling rules in quantum physics. Two systems of definite total angular momenta $0 \leq j_1, j_2 \leq 1$ decompose as:

$$\begin{aligned}
&\frac{1}{2} \otimes \frac{1}{2} \to 1 \oplus 0, && 1 \otimes 1 \to 0 \oplus 1 \oplus 2, && \frac{1}{2} \otimes 1 \to \frac{1}{2} \oplus \frac{3}{2},\\
&0 \otimes 0 \to 0, && 0 \otimes \frac{1}{2} \to \frac{1}{2}, && 0 \otimes 1 \to 1.
\end{aligned}$$

A general rule is that two such 'particles' j_1 and j_2 fuse into a 'particle' with angular momentum $j = |j_1 - j_2|, |j_1 - j_2| + 1, \ldots j_1 + j_2$. As outlined in Chapter 2, every total angular momentum eigenstate j corresponds to an irreducible representation of $SU(2)$ labelled by a Young diagram with two rows and overhang $2j$. This analogy can be seen as the key motivating principle behind Permutational Quantum Computing.

The Ising anyon theory fusion rules can be described in terms of 'k-deformations' of $SU(2)$, denoted as $SU(2)_k$. For an integer k, there exist $k+1$ 'particles' labeled

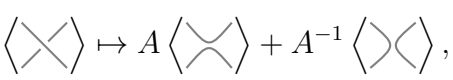
by $j = 0, 1/2, 1, 3/2 \ldots k/2$. The fusion rules for 'particles' j_1 and j_2 become: $j = |j_1 - j_2|, |j_1 - j_2| + 1, \ldots \min\{j_1 + j_2,\, k - (j_1 + j_2)\}$. For $k = 2$, this gives the fusion rules for Ising theory by setting $\mathbf{1} = 0$, $\sigma = 1/2$, $\psi = 1$ [50].

1.5 Links and Topological Invariants

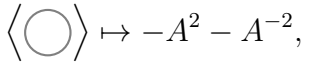
Vacuum splitting, braiding and a subsequent fusion of anyons to vacuum draws a link in $(2 + 1)$ dimensional spacetime. A knot is a circle embedded in three dimensions – examples of some non-trivial knots are shown in Fig. 1.1. A link is then an embedding of multiple knots. Topologically distinct links cannot be continuously deformed into one another without cutting, which is referred to as regular isotopy [51].

Figure 1.1: Knots - an unknot, a right handed trefoil and a cinquefoil. By Alexander's theorem, any knot (and link) can be obtained as a closure of a braid word [51].

A topological invariant is a number assigned to a link such that distinct numbers correspond to topologically distinct links. An important example of such topological invariant is the Kauffman bracket [51, 42]. It is can be defined by recursive application of the "skein" rule, replacing a left-handed crossing with a linear combination of diagrams with the crossing removed:



where A is a complex number. Formally, this is a homomorphism from links to the Temperley-Lieb algebra $\mathsf{TL}_n(d)$ and the rule that replaces each unknot with a number $d = -A^2 - A^{-2}$:



then assigns each link L a polynomial in A and A^{-1} that is invariant under regular isotopy. Kauffman bracket is equivalent to another topological invariant, called the Jones' polynomial, up to a normalization factor [51].

The obvious algorithm for computing the Jones polynomial by recursive application of the skein rule already hints on the difficulty of computing it exactly, since the

number of diagrams to evaluate grows exponentially with the number of crossings. It was shown by Kuperberg that a broad class of 'value-distinguishing' approximations of the invariant is a #P-hard problem [52].

1.6 Aharonov-Jones-Landau Algorithm

Freedman, Kitaev, Larsen and Wang proposed Topological Quantum Computing in Ref. [48] and proved that the model is universal for quantum computation by showing that a polynomially precise approximation to the Jones polynomial for the Fibonacci anyons is a BQP-complete problem.[2] Aharonov, Jones and Landau later extended this argument to certain other anyon theories [53, 54]. A key contribution of their work is an explicit formulation of the algorithm of Ref. [48].

The Aharonov-Jones-Landau algorithm solves the following problem: form a link by connecting ends of the neighboring strands of a braid b on $2n$ strands (the plat closure) and approximate the value of the Jones polynomial of it to polynomial precision. The algorithm is based on a unitary representation of the braid group called the path model representation due to Jones [55], which is interpreted as a quantum circuit. The value of the Jones polynomial is encoded into matrix elements of this representation: for a braid b on n pairs of anyons prepared in a particle eigenstate $|\alpha\rangle$ and measured in an eigenstate $|\beta\rangle$, the Jones polynomial corresponds, up to normalization, to the transition amplitude $\langle\alpha|Q(b)|\beta\rangle$ where $Q(b)$ is the path-model representation of the braid b (see Fig. 1.2). The algorithm approximates both real and imaginary parts of this transition amplitude to polynomial precision by the Hadamard test.

The central object of study in this part of the thesis is an analog of this algorithm which uses angular momentum eigenstates in place of anyons and thus works with the symmetric group in place of the braid group [5, 3]. The input and output states are replaced by spin eigenfunctions introduced in Chapter 2, the representation matrix $Q(b)$ is replaced with a permutation matrix and the transition amplitudes correspond to matrix elements of the Yamanouchi-Kotani representations (described later). It was a conjecture, largely supported by BQP-completeness of the Aharonov-Jones-Landau algorithm, that there is no classical polynomial time algorithm for approximating these transition amplitudes [3]. We discuss this conjecture later in Chapter 3.

[2] This approximation is not a value distinguishing approximation.

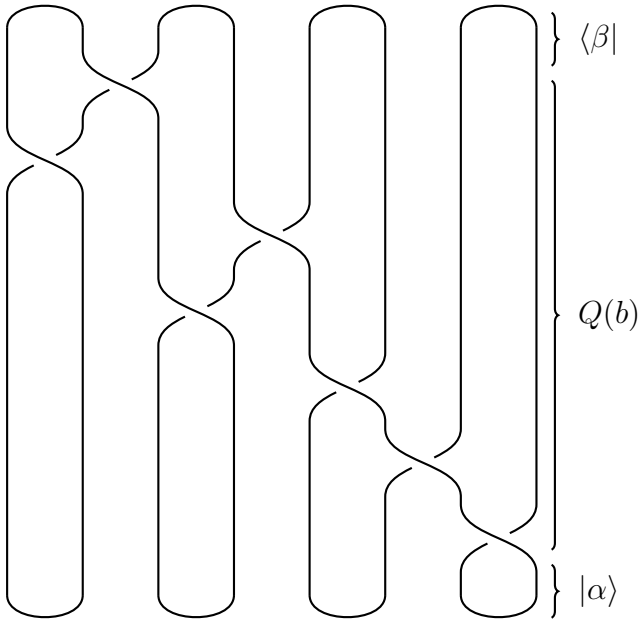


Figure 1.2: Graphical representation of the overlap approximated by the Aharonov-Jones-Landau algorithm. Suppose that we work with Fibonacci anyons, so that there are only two particle types, τ and (vacuum) $\mathbf{1}$. Then $|\alpha\rangle = |(\tau\tau)(\tau\tau)(\tau\tau)(\tau\tau)\rangle$ corresponds to the initial state, while $|\beta\rangle = |(\tau\tau)(\tau\tau)(\tau\tau)(\tau\tau)\rangle$ is the final state, measured with probability $|\langle\beta|Q(b)|\alpha\rangle|^2$, where $Q(b)$ is the unitary representation of the applied braid. Using either the Hadamard test or interferometry, it is in principle possible to approximate both the real and imaginary part of the transition amplitude $\langle\beta|Q(b)|\alpha\rangle$ to polynomial precision. Ref. [7] shows that this problem is BQP-complete.

Chapter 2

Spin Eigenfunctions

2.1 Introduction

Permutational Quantum Computing works with n distinguishable spin-1/2 particles (qubits) labelled by elements of $[n] := \{1, 2 \ldots n\}$ and uses bases defined by a set of commuting spin measurements. The corresponding basis states are called 'spin eigenfunctions' in Ref. [56].

Spin eigenfunctions are connected to the computational basis by the $SU(2)$ Quantum Schur Transformation [57, 58, 59, 60] and underpin many applications in quantum information processing, such as decoherence-free subspaces [61], communication without a shared reference frame [62, 63] or quantum color coding [64]. Angular momentum recoupling also plays an important role in spin networks and the Ponzano-Regge model of quantum gravity [3, 65, 66].

We introduce the spin eigenfunctions and describe some of their properties needed to define the computational model.

2.2 Spin Operators

Let:

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \qquad Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \tag{2.1}$$

be the Pauli matrices. They are Hermitian, square to identity and obey the following equalities:

$$ZX = -XZ = iY, \qquad YZ = -ZY = iX, \qquad XY = -YX = iZ. \tag{2.2}$$

With a convention that $\hbar = 1$, a spin of k-th qubit is a triple of operators:

$$\vec{S}_k = \frac{1}{2}(X_k, Y_k, Z_k),$$

where X_k denotes a Pauli X operator on k-th qubit given by:

$$X_k := \overbrace{\underbrace{\mathbf{1} \otimes \ldots \otimes \mathbf{1}}_{k-1} \otimes X \otimes \mathbf{1} \ldots \otimes \mathbf{1}}^{n}, \tag{2.3}$$

where $\mathbf{1}$ is the 2×2 identity matrix. Operators Z_k and Y_k are defined similarly.

Definition 1 (Total Spin Operator[1])**.** *The total spin operator on $A \subseteq [n]$ is:*

$$\vec{S}_A := \sum_{k \in A} \vec{S}_k, \qquad\qquad S_A^2 := \vec{S}_A \cdot \vec{S}_A, \tag{2.4}$$

where $\cdot$ is a three-dimensional dot product of the operator vectors. The operator is Hermitian and can be written in terms of Pauli operators as:

$$S_A^2 = \frac{1}{4} \sum_{k \in A} \sum_{l \in A} (X_k X_l + Y_k Y_l + Z_k Z_l). \tag{2.5}$$

We write $S_{[n]}^2 = S^2$ for brevity.

Example 1.

$$S_{[2]}^2 = \frac{1}{2}[(\mathbf{1} + X_1 X_2) + (\mathbf{1} + Y_1 Y_2) + (\mathbf{1} + Z_1 Z_2)] = \frac{3}{2}\mathbf{1} + \frac{1}{2}(X_1 X_2 + Y_1 Y_2 + Z_1 Z_2).$$

Definition 2 (Azimuthal Spin Operator)**.** *The azimuthal spin operator on $A \subseteq [n]$ is given by:*

$$Z_A = \frac{1}{2} \sum_{k \in A} Z_k. \tag{2.6}$$

It is Hermitian and we write $Z_{[n]} = Z$. Operators X_A and Y_A are defined similarly.

Note that Eq. 2.5 can be written as $S_A^2 = Z_A^2 + X_A^2 + Y_A^2$ in terms of these operators. We now prove a series of facts about the spin operators.

[1]We will not make a careful distinction between angular momentum and spin in this thesis as the S^2 operators have been defined without reference to specific physical systems. The total Spin operator is referred to as the Casimir element in the language of Lie groups.

Claim 1. *The total spin operator acts trivially on qubits:*

$$S^2_{\{\bullet\}} = \frac{3}{4}\mathbf{1}, \tag{2.7}$$

where $\{\bullet\}$ *denotes a singleton set.*

Proof. From Eq. 2.5:

$$S^2_{\{\bullet\}} = \frac{1}{4}\left(X^2_\bullet + Y^2_\bullet + Z^2_\bullet\right) = \frac{3}{4}\mathbf{1}. \tag{2.8}$$

□

Eigenvalues of S^2 are denoted by $J(J+1)$, are all non-negative and correspond to total spin J eigenstates. From this definition, the 'total spin' of a qubit is $J_{\{\bullet\}} = 1/2$. We often identify the computational basis states $|0\rangle, |1\rangle$ with azimuthal angular momentum eigenstates of qubits:

$$\begin{aligned} \left|J_{\{\bullet\}} = \frac{1}{2},\, M = -\frac{1}{2}\right\rangle &:= \left|-\frac{1}{2}\right\rangle \equiv |0\rangle\,, \\ \left|J_{\{\bullet\}} = \frac{1}{2},\, M = \ \ \frac{1}{2}\right\rangle &:= \left|\ \ \frac{1}{2}\right\rangle \equiv |1\rangle\,. \end{aligned} \tag{2.9}$$

Claim 2. Z_A *and* S^2_B *commute if* $A \supseteq B$.

Proof. From Def. 2, we can write:

$$Z_A = \frac{1}{2}\sum_{k\in B} Z_k + \frac{1}{2}\sum_{k\in A/B} Z_k = Z_B + Z_{A/B}. \tag{2.10}$$

It follows that:

$$[Z_A, S^2_B] = [Z_B, S^2_B], \tag{2.11}$$

because S^2_B and $Z_{A/B}$ act on disjoint subset of qubits and therefore commute. Using Def. 1, Def. 2 and linearity, the commutator in Eq. 2.11 can be expanded as:

$$\begin{aligned} 8[Z_B, S^2_B] &= \sum_{k,l,m\in B} [Z_k, (X_l X_m + Y_l Y_m + Z_l Z_m)] \\ &= \sum_{k,l,m\in B} [Z_k, X_l]X_m + X_l[Z_k, X_m] + [Z_k, Y_l]Y_m + Y_l[Z_k, Y_m]. \end{aligned} \tag{2.12}$$

Note that it follows from Eq. 2.2 that:

$$[Z_k, X_l] = \delta_{kl}(Z_k X_k - X_k Z_k) = 2i\delta_{kl} Y_k, \tag{2.13}$$

where:

$$\delta_{kl} = \begin{cases} 1 & \text{if } k = l, \\ 0 & \text{otherwise,} \end{cases} \tag{2.14}$$

is the Krönecker delta and similar relations hold for the other Pauli operators. Using these relations in Eq 2.12, we obtain:

$$8[Z_B, S_B^2] = 2i \sum_{k,l \in B} Y_l X_k + X_l Y_k - Y_l X_k - X_l Y_k = 0.$$

□

Operators S_A^2 and S_B^2 commute if A and B are disjoint or one is subset of another:

Claim 3 (Commutation of Total Spin Operators)**.** *S_A^2 and S_B^2 commute if $A \supseteq B$ or if $A \cap B = \emptyset$.*

Proof. For $A \supseteq B$, write:

$$\sum_{k \in A} \vec{S}_k = \overbrace{\sum_{k \in B} \vec{S}_k}^{\vec{\beta}} + \overbrace{\sum_{k \in A/B} \vec{S}_k}^{\vec{\alpha}}.$$

Note that $S_A^2 = (\vec{\alpha} + \vec{\beta}) \cdot (\vec{\alpha} + \vec{\beta})$, $\quad S_B^2 = \vec{\beta} \cdot \vec{\beta}$ and that from Eq. 2.3, the components of $\vec{\alpha}$ and $\vec{\beta}$ act on disjoint subsets of qubits and therefore commute. It follows that:

$$\vec{\alpha} \cdot \vec{\beta} = \vec{\beta} \cdot \vec{\alpha}, \qquad\qquad [\vec{\beta} \cdot \vec{\beta},\, \vec{\alpha} \cdot \vec{\alpha}] = \vec{0}.$$

This leads to:

$$[S_A^2, S_B^2] = 2[\vec{\alpha} \cdot \vec{\beta},\, S_B^2]. \tag{2.15}$$

Observe that similarly to Eq. 2.5:

$$2\,\vec{\alpha} \cdot \vec{\beta} = \frac{1}{2} \sum_{l \in A/B} \sum_{k \in B} (X_l X_k + Y_l Y_k + Z_l Z_k). \tag{2.16}$$

Using this in the commutator of Eq 2.15, we have that:

$$\left[2\,\vec{\alpha} \cdot \vec{\beta},\, S_B^2\right] = \sum_{l \in A/B} X_l \left[X_B, S_B^2\right] + Y_l \left[Y_B, S_B^2\right] + Z_l \left[Z_B, S_B^2\right] = 0, \tag{2.17}$$

where we used Claim 2. The operators commute trivially for $A \cap B = \emptyset$ because they act on a disjoint sets of qubits. □

Operators Z and S^2 commute by Claim 2 and can be simultaneously diagonalized by a set of $\{|J, M\rangle\}$ eigenstates:

$$S^2 |J,\, M\rangle = J(J+1)\, |J,\, M\rangle \qquad\qquad Z\, |J,\, M\rangle = M\, |J,\, M\rangle\,, \tag{2.18}$$

where J is the total spin of all qubits and M is their azimuthal spin. It follows from Hermiticity of the operators that these states are orthogonal to one another:

$$\langle J', M' | J, M\rangle = \delta_{J',J}\, \delta_{M',M}. \tag{2.19}$$

However, the $|J, M\rangle$ eigenstates of S^2 and Z do not form a basis for Hilbert spaces of 3 or more qubits. To see this, we need an additional fact about the S_A^2 and Z_A eigenvalues known as the angular momentum addition theorem:

Theorem 1 (Addition of Angular Momentum in Quantum Mechanics)**.** *The eigenvalues $J_{A\cup B}, J_A$ and J_B of operators $S^2_{A\cup B}$, S^2_A and S^2_B are related by the angular momentum addition rules [67, 56]:*

$$|J_A - J_B| \leq J_{A\cup B} \leq J_A + J_B, \qquad\qquad J_{A\cup B} + J_A + J_B \in \mathbb{Z}. \tag{2.20}$$

The eigenvalues M of Z_A for any $A \subseteq [n]$ take values in $\{-J_A, -J_A + 1, \ldots, J_A\}$.

The proof of Theorem 1 is standard (see for example Ref. [68]) and I omit it here. As a corollary, spin eigenvalues $J_{[k+1]}$ and $J_{[k]}$ that correspond to operators $S^2_{[k+1]}$ and $S^2_{[k]}$ are subject to:

$$J_{[k+1]} \in \left\{ \left| J_{[k]} + \frac{1}{2} \right|, \left| J_{[k]} - \frac{1}{2} \right| \right\}. \tag{2.21}$$

Another consequence of Claim 1 is that the values of total spin J_A, that correspond to the eigenvalue of the S^2_A operator for $A \subseteq [n]$, are upper-bounded by $n/2$. There are $2J + 1$ distinct values of M for every fixed J and the number of distinct Z, S^2 eigenstates $|J, M\rangle$ is therefore given by:

$$\sum_{J=1}^{n/2} J\,(2J+1) = \frac{(n+2)(n+1)n}{12} + \frac{(n+2)n}{8} = O(n^3). \tag{2.22}$$

If operators act on n-qubits and hence a 2^n-dimensional Hilbert space, this falls an exponential number of states short of a basis.

2.3 Genealogical Spin Basis

The S^2, Z basis can be completed by listing eigenvalues of an additional set of commuting total spin operators on the set of qubits. One way to do this is by choosing the additional subsets of qubit as follows:

$$\{1,2\} \subseteq \{1,2,3\} \subseteq \ldots \subseteq \{1,2,3 \ldots n-1\} \subseteq [n], \tag{2.23}$$

along with the corresponding spin operators $S^2_{\{12\}}$, $S^2_{\{123\}}$, $\ldots S^2$, Z. These operators are all Hermitian and commute by Claims 2 and 3. They share orthogonal eigenstates that can be used to define a basis – this basis choice will be called the *genealogical basis* [56].

Definition 3 (Genealogical basis)**.** *The joint eigenstates of spin operators:*

$$S^2_{[2]}, S^2_{[3]}, \ldots S^2, Z, \tag{2.24}$$

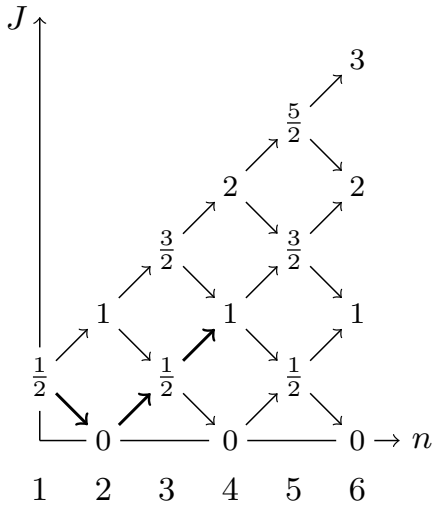
form a basis called the genealogical basis [56].

We now prove that the basis is complete for n-qubit Hilbert space and establish additional facts about genealogical spin eigenfunctions for that purpose. Specifically, we introduce representation of genealogical spin eigenfunction in terms of Bratteli paths, Yamanouchi symbols and standard Young tableaux on two rows. We then use standard results from enumerative combinatorics to prove the completeness. This treatment loosely follows Pauncz [56] who attributes the proofs to Yamanouchi [69] and Kotani [70].

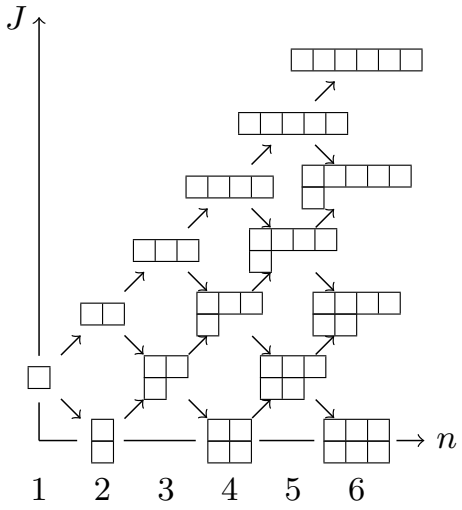
Definition 4 (Bratteli diagrams (informal))**.** *The angular momentum Bratteli diagram is a graph that depicts the relation between the possible states of angular momenta on a given number of qubits. For every number n of qubits, the graph contains a vertex for all possible values of n-particle spin, labelled by the spin value. Vertices corresponding to $n+1$ qubits are connected to the vertices on n qubits if their labels differ by $\pm 1/2$. See Fig. 2.1.*

Definition 5 (Bratteli paths)**.** *A $|J_{[2]}, J_{[3]} \ldots J_{[k]}\rangle$ eigenstate of the $S^2_{[2]}, S^2_{[3]} \ldots S^2_{[k]}$ total spin operators corresponds bijectively to a path of length k in the Bratteli diagram, given by $1/2 \to J_{[2]} \to \ldots \to J_{[k]}$. See Fig. 2.1.*

Figure 2.1



(a) Bratteli diagram. The highlighted path $\frac{1}{2} \to 0 \to \frac{1}{2} \to 1$ corresponds to total spin eigenstate $\big|J = 1, J_{[3]} = \frac{1}{2}, J_{[2]} = 0\big\rangle$. The path correponds to a Yamanouchi symbol 011.



(b) Bratteli paths are bijective with standard Young tableaux on two rows. A Young diagram with $n/2 + J$ boxes in the top and $n/2 - J$ boxes in the bottom row labels a set of paths that correspond to the Young tableaux with a shape of this diagram.

Bratteli paths can be described as sequences of $\nearrow$, $\searrow$ steps:

Claim 4 (Bratteli Paths and Yamanouchi Symbols)**.** *Any length-k Bratteli path corresponds uniquely to a $(k-1)$-bit string such that any prefix of length $m \leq k$ contains at most $\lceil m/2 \rceil$ zeroes. These bitstrings are referred to as Yamanouchi symbols in [56].*

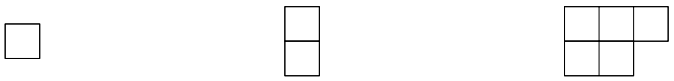
Proof. By Def. 5, the length-k Bratteli path corresponds to an eigenstate $|J_{[2]}, J_{[3]} \dots J_{[k]}\rangle$. It follows from Eq. 2.21 that $J_{[k]} - J_{[k-1]} = \pm 1/2$ for all $2 \leq k \leq n$. The corresponding Yamanouchi symbol is then given by a bitstring $a_2 a_3 a_4 \dots a_n$ for $a_k := J_{[k]} - J_{[k-1]} + 1/2$. □

Yamanouchi symbols correspond bijectively to standard Young tableaux with two rows.

Definition 6 (Standard Young Diagrams and Tableaux [3, 57, 71])**.** *A standard Young diagram is a partition of n boxes into rows, such that no row is longer than the one above it. A standard Young tableau on n boxes is a Young diagram with the*

boxes filled with integers in $[n]$, such that the numbers in any column are increasing downwards and the numbers in each row are increasing to the right and no number is repeated.

Example 2. *Example of standard Young diagrams:*



Examples of standard Young tableaux:

1	2
3	

1	2	4	5
3			

1	2
3	4

Claim 5 (Bijection between Genealogical Spin Eigenfunctions and Standard Young Tableaux)**.** *There is a bijection between the set of genealogical spin eigenfunctions on n qubits and standard Young tableaux with two rows and n boxes.*

Proof. From Def. 5 and Claim 4, it suffices to establish a bijection between Yamanouchi symbols and Young tableaux. Let $a = a_2 a_3 \dots a_n$ be a Yamanouchi symbol and let u, v be empty lists. Add an element 1 to u. If $a_i = 0$, append i to v, if $a_i = 1$, append i to u. It trivially follows that both u and v contain number that increase from left to right. Suppose that at some point, say step j, both u and v have the same length, which means that length-j prefix of a has the same number of 0 and 1. By definition of Yamanouchi symbol, this means that $a_j = 1$, otherwise there is a contradiction with the prefix condition of Claim 4. It follows that $|u| \geq |v|$.

Let u and v define the upper and lower rows of the Young tableau respectively. Numbers in every row increase from left to right, and the second row is at most as long as the first one. It not hard to see from the previous that the numbers in every column increase from top to bottom. The Young tableau is hence in the standard form. Let $[u, v]$ denote standard Young tableau with rows u and v. Set a to be an empty bitstring. For each $1 \leq k \leq n$, append 0 to a if $k \in v$ and append 1 to a otherwise. The prefix condition of Claim 4 follows from the definition of standard Young tableau. Two Yamanouchi symbols are equal if and only if they correspond to the same Young tableau. □

Claim 6. *The $S^2_{[2]}, S^2_{[3]}, \dots S^2_{[k-1]}, S^2_{[k]}$ eigenstate $|J_{[2]}, J_{[3]} \dots J_{[k]}\rangle$ corresponds to a Young tableau with two rows and shape $[\mu, \nu]$, such that $\mu + \nu = k$ and $\mu - \nu = 2J_{[k]}$.*

Proof. Observe that $\mu - 1$ is the number of $\nearrow$ steps in the Bratteli path. Similarly, ν is the number of $\searrow$ steps in the Bratteli path. The path starts at $J_1 = 1/2$. It follows that: $2J_{[k]} - 1 = \mu - 1 - \nu$. □

Claims 5 and Claim 6 are now combined to prove the basis completeness:

Claim 7. *The genealogical basis is complete for 2^n-dimensional Hilbert space [56].*

Proof. Claim 6 shows that any spin eigenstate $|J_{12}, J_{123} \dots J_{[n-1]}, J\rangle$ corresponds to a standard Young Tableau of shape $[\mu, \nu]$ such that $\mu + \nu = n$ and $\mu - \nu = 2J$. The number of such tableaux can be found using the Frame-Robinson-Thrall hook-length formula [72, 56, 71] from which the number of standard Young Tableaux of a shape λ on n boxes is given by:

$$f^\lambda = \frac{n!}{\prod_{(ij)} h_{ij}}, \tag{2.25}$$

where h_{ij} is called the hook-length and the product $\prod_{(ij)}$ runs over all boxes in λ. The hook-length h_{ij} is defined as 1 + the number of boxes to the right of the (i, j)-th box + the number of the boxes below the (i, j)-th box in the standard Young diagram λ (See Fig. 2.2).

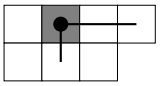


Figure 2.2: Hook-length $h_{12} = 4$ associated to the $(1, 2)$ box (in gray) in a standard Young diagram of shape $\lambda = [4, 3]$. It is the number of boxes crossed by the thick black line, called the 'hook' of $(1, 2)$.

A hook diagram is a Young diagram with h_{ij} written in each box (i, j). A two-row standard Young diagram $\lambda = [\mu, \nu]$ with μ boxes in the first row and ν boxes in the second row has a hook diagram given by:

$\mu+1$	$\cdots$	$\mu-\nu+2$	$\mu-\nu$	$\mu-\nu-1$	$\cdots$	2	1
ν	$\cdots$	1					

(2.26)

Combining this with Eq. 2.25 leads to:

$$f^{[\mu,\nu]} = \frac{(\mu+\nu)!\,(\mu-\nu+1)}{\nu!(\mu+1)!} = \binom{n}{n/2-J} - \binom{n}{n/2-J-1} := f(J), \tag{2.27}$$

where we assume that $\binom{n}{k} = 0$ for all nonpositive k. By Claim 1, there are $(2J+1)$ genealogical spin eigenfunctions for each two-row Young diagram with overhang $2J$. The number of them is therefore given by:

$$\sum_J (2J+1) f(J), \tag{2.28}$$

where the sum of J runs in integer steps from 0 to $n/2$ for n even and from $1/2$ to $n/2$ for n odd. Choose n to be even to simplify the proof statement (the n odd case can be proved similarly). Eq. 2.28 can be then written as:

$$\sum_{J=0}^{n/2} (2J+1) f(J) =$$

$$\begin{array}{rccccc|c}
f(0) + & f(1) + & f(2) + & \cdots & f(n/2) & & \binom{n}{n/2} \\
 & f(1) + & f(2) + & \cdots & f(n/2) & & \binom{n}{n/2-1} \\
 & f(1) + & f(2) + & \cdots & f(n/2) & & \binom{n}{n/2-1} \\
 & & f(2) + & \cdots & f(n/2) & & \binom{n}{n/2-2} \\
 & & f(2) + & \cdots & f(n/2) & & \binom{n}{n/2-2} \\
 & & & & \vdots & & \vdots \\
 & & & & f(n/2) & & \binom{n}{0}
\end{array}$$

where the summation over all rows is implied and each row sums up to the quantity to the right of |. It follows that:

$$\sum_{J=0}^{n/2} (2J+1) f(J) = \binom{n}{n/2} + 2 \sum_{k=0}^{n/2-1} \binom{n}{k} = 2^n, \tag{2.29}$$

which shows that the genealogical basis is complete. □

Claim 2.3 also implies that there exist sets of states with fixed J and M with size that scales as $O(2^n)$ because the summation in Eq. 2.28 runs only over $\lceil (n+1)/2 \rceil$ terms. For example, degeneracy of the $J = 0$ basis states is given by the number of standard Young tableaux with shape λ corresponding to n (even) boxes with no overhang:

$$f^\lambda = \frac{2}{2+n} \binom{n}{n/2} \approx O\left(2^n\right). \tag{2.30}$$

This observation will be important later when we study a class of computations in a subspace with fixed J, M quantum numbers.

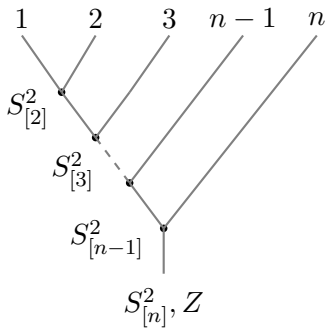


Figure 2.3: Genealogical spin basis on n qubits.

2.4 Spin Recoupling

Spin eigenfunctions can be represented by spin recoupling diagrams used in Refs. [3, 5]. An example of a diagram for the genealogical basis is in Fig. 2.3. The diagrams are to be read from top to bottom. The numbers at the leaf nodes are qubit indices and every internal vertex • carries a total spin operator S^2_A acting on the qubits in the leaves A of the subtree rooted at it. The Z and S^2 operators are associated to the root of the tree. Spin eigenfunctions can be described similarly – as an example, all genealogical spin eigenfunctions for $n = 3$ are shown in Fig 2.4.

2.4.1 $SU(2)$ Clebsch-Gordan Transform

Coupling of two spin eigenstates in an internal vertex of a recoupling diagram is described by the $SU(2)$ Clebsch-Gordan transform:

Definition 7 (Clebsch-Gordan transform [73])**.** *Joint eigenstates* $|J_A, M_A\rangle$ *and* $|J_B, M_B\rangle$ *of spin operators* S^2_A, Z_A *and* S^2_B, Z_B *on* $A, B \subseteq [n]$ *and* $A \cap B = \emptyset$ *can be related to eigenstates of spin operators* $S^2_{A\cup B} = (\vec{S}_A + \vec{S}_B)^2$, $Z_{A\cup B} = Z_A + Z_B$ *by:*

$$|J_{A\cup B}, M_{A\cup B}, J_A, J_B\rangle = \sum_{M_A, M_B} C^{J_{A\cup B}, M_{A\cup B}}_{J_A, M_A; J_B, M_B} |J_A, M_A\rangle |J_B, M_B\rangle , \tag{2.31}$$

where:

$$C^{J_{A\cup B}, M_{A\cup B}}_{J_A, M_A; J_B, M_B} := \langle J_A, M_A; J_B, M_B | J_{A\cup B}, M_{A\cup B}, J_A, J_B\rangle , \tag{2.32}$$

is the SU(2) Clebsch-Gordan coefficient. Since $[S^2_A, S^2_{A\cup B}] = [S^2_B, S^2_{A\cup B}] = 0$, *the* $S^2_{A\cup B}$

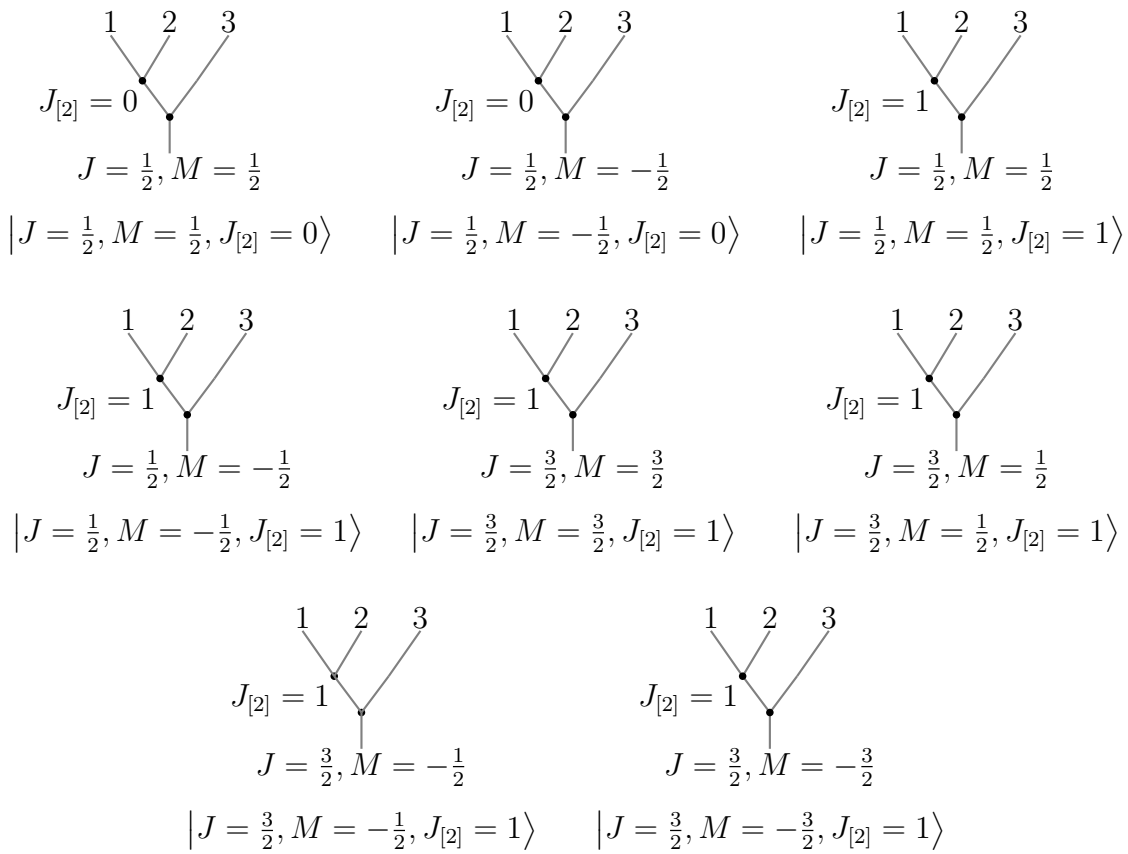


Figure 2.4: All three-qubit genealogical basis states in diagrammatic representation.

eigenstate is also an eigenstate of S_A^2 and S_B^2 and thus retains the J_A, J_B labels.[2] *Aside from the obvious orthogonality relations, the coefficients also satisfy following identity:*

$$C^{J_{A\cup B},\, M_{A\cup B}}_{J_A,M_A;J_B,M_B} = \delta^{M_{A\cup B}}_{M_A+M_B} C^{J_{A\cup B},\, M_{A\cup B}}_{J_A,M_A;J_B,M_B}, \tag{2.34}$$

that can be understood as a conservation of the azimuthal angular momentum.

Clebsch-Gordan coefficients are elements of an orthogonal transform with matrix elements given by Eq. 2.32. This transform can be thought of as the internal vertex

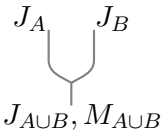
[2]The same also works for states that carry additional degeneracy labels α, β determined by operators that commute with $Z_{A\cup B}$, $S^2_{A\cup B}$ as well as one another:

$$|J_{A\cup B}, M_{A\cup B}, J_A, J_B, \alpha, \beta\rangle = \sum_{M_A,M_B} C^{J_{A\cup B},\, M_{A\cup B}}_{J_A,M_A;J_B,M_B} |J_A, M_A, \alpha\rangle\, |J_B, M_B, \beta\rangle\,. \tag{2.33}$$

Here we use this property (which is just basis fixing) with a sequence of total spin operators that each act on subsets of qubits in A or B that are either contained in one another or disjoint (so that the operators can be arranged onto vertices defined by a coupling tree).

We would work with $SU(d)$ Clebsch Gordan coefficients if we coupled a tensor product of d-level systems and used generalized Paulis in place of qubits and Pauli matrices. Such coefficients are harder to deal with than $SU(2)$, as there is for example no known closed-form solution for $d \geq 3$.

of the spin recoupling diagrams:



Claim 8. *Clebsch-Gordan coefficients can be computed to $\pm\epsilon$ error in $poly(n, \log(1/\epsilon))$ time.*

Proof. Clebsch-Gordan coefficients have the following closed-form solution:

$$C^{J_{A\cup B},M_{A\cup B}}_{J_A,M_A;J_B,M_B} = (-1)^{M+J_A-J_B}\sqrt{2J_{A\cup B}+1}\begin{pmatrix} J_A & J_B & J_{A\cup B} \\ M_A & M_B & M_{A\cup B}\end{pmatrix}, \tag{2.35}$$

where the Wigner $3j$ symbol can be computed using the Racah formula [74]:

$$\begin{pmatrix} a & b & c \\ d & e & f\end{pmatrix} = \tag{2.36}$$

$$\sum_t (-1)^{t+a-b-f}\frac{\sqrt{\Delta(abc)(a+d)!(b+e)!(c+f)!(a-d)!(b-e)!(c-f)!}}{t!(c-b+t+d)!(c-a+t-e)!(a+b-c-t)!(a-d-t)!(b-t-e)!}, \tag{2.37}$$

where $\Delta(abc)$ is the triangle coefficient given by:

$$\Delta(abc) = \frac{(a+b-c)!(a-b+c)!(-a+b+c)!}{(a+b+c+1)!},$$

and the summation $\sum_t$ runs over all integer t for which the factorials in the expression are well-defined. It is known that there are $\nu+1$ such terms, where [75]:

$$\nu = \min\{a\pm d, b\pm e, c\pm f, a+b-c, b+c-a, c+a-b\}\ .$$

In the context considered here $a, b\ldots f$ are total spins of up to n qubits and it always holds that $-n/2 \leq a, b\ldots f \leq n/2$. Hence, the summation in Eq. 2.36 runs over $O(n)$ terms. Each of these terms requires evaluation of factorials of quantities upper bounded by n that can be computed exactly efficiently. Aside from that, the computation only requires division and a square root, each of which can be approximated to $\pm\epsilon$ error in $poly(n, \log(1/\epsilon))$ time. Evaluation of the $SU(2)$ Clebsch-Gordan coefficient up to $\pm\epsilon$ error hence takes $poly(n, \log(1/\epsilon))$ time. □

2.4.2 F-moves

This diagrammatic representation can be conveniently used to describe the unitary transformation between distinct spin bases, called the F-move. It allows for a basis change between the right and left recoupling at a three qubit level (see Fig. 2.5). Any basis defined that can be obtained by applying a sequence of F-moves to the

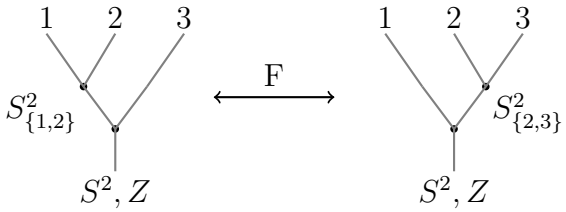
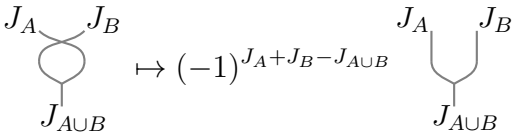


Figure 2.5: Two distinct bases of spin eigenfunctions on three qubits are related by a unitary transformation called the F-move. The left diagram corresponds to coupling in the sequentially coupled basis, while the right one corresponds to the basis defined by the operators $S^2, Z, S^2_{\{2,3\}}$. The F-move is a unitary transformation from the elements $|J, M, J_{12}\rangle$ of the left basis to $|J, M, J_{23}\rangle$ elements of the right basis.

genealogical basis, along with the Z operator measurement is complete. Since the F-move leaves the number of internal vertices in a recoupling diagram invariant, we also conclude that any n-qubit spin eigenfunction contains exactly $n-1$ internal vertices.

We remark that the set of operations on spin eigenfunctions can be complemented by accounting for permutations that act on the spin eigenstates as follows:



The above rules applied to spin eigenfunctions provide a graphical calculus for the Wigner-Racah algebra that fully describes spin recoupling [56, 76, 77].

2.4.3 Quantum Schur Transform

Genealogical spin eigenfunctions can be related to the computational basis by a unitary transformation called the $SU(2)$ Quantum Schur Transform [58]. This transformation acts on a system of n qubits in computational basis $|y_1\rangle \ldots |y_n\rangle$, $y_i = \pm 1/2$ as follows: it starts by Clebsch-Gordan coupling $|y_1\rangle$ with $|y_2\rangle$ into $|J_{[2]}, M_{[2]}\rangle$ as:

$$|J_{[2]}, M_{[2]}\rangle = \sum_{y_1, y_2} C^{J_{[2]}, M_{[2]}}_{y_1, y_2} |y_1, y_2\rangle \,. \tag{2.38}$$

It then couples $|y_3\rangle$ with $|J_{[2]}, M_{[2]}\rangle$ into $|J_{[3]}, J_{[2]}, M_{[2]}\rangle$ by cascading Clebsch-Gordan transformations. As discussed previously, this state still carries the label $J_{[2]}$, since the total angular momentum operators $S^2_{[3]}$ and $S^2_{[2]}$ commute. The label $M_{[2]}$ is in contrast discarded because $Z_{[2]}$ no longer commutes with $S^2_{[3]}$. This process is continued until all the qubits are coupled into a genealogical basis state $|J, M, \ldots J_{[3]}, J_{[2]}\rangle$.

It is easy to see from the above that the matrix elements of this transformation are products of Clebsch-Gordan coefficients. For example for $n = 3$:

$$\begin{aligned} |J, M, J_{[2]}\rangle &= \sum_{y_1,\, y_2} \sum_{M_{[2]},\, y_3} C^{J,M}_{J_{[2]},M_{[2]};\frac{1}{2},y_3} C^{J_{[2]},M_{[2]}}_{\frac{1}{2},y_1;\frac{1}{2},y_2} |y_1, y_2, y_3\rangle \\ &= \sum_{y_1,y_2,y_3} [U_{\text{Sch}}]^{J,M,J_{[2]}}_{y_1,y_2,y_3} |y_1, y_2, y_3\rangle \,, \end{aligned}$$

from where:

$$[U_{\text{Sch}}]^{J,M,J_{[2]}}_{y_1,y_2,y_3} = C^{J,M}_{J_{[2]},y_1+y_2;\frac{1}{2},y_3} C^{J_{[2]},y_1+y_2}_{\frac{1}{2},y_1;\frac{1}{2},y_2}, \tag{2.39}$$

where U_{Sch} is the Quantum Schur Transform unitary. Two examples of spin eigenfunctions in the computational basis are:

$$\begin{aligned} \left| J = \frac{1}{2}, M = -\frac{1}{2}, J_{[2]} = 1 \right\rangle &= \sqrt{\frac{2}{3}} |001\rangle - \frac{|010\rangle + |100\rangle}{\sqrt{6}}, \\ \left| J = \frac{1}{2}, M = -\frac{1}{2}, J_{[2]} = 0 \right\rangle &= \frac{|010\rangle - |100\rangle}{\sqrt{2}}. \end{aligned} \tag{2.40}$$

The transform can be implemented efficiently using the algorithms of Bacon, Harrow and Chuang, Kirby and Strauch and Krovi [78, 58, 59, 60, 79]. The transform is based on the Schur-Weyl duality that relates the irreducible representations of the symmetric group and the special unitary group $SU(d)$. It thus relates the actions performed by tensor products of d-dimensional unitaries to those performed by permutation of the qudits. A permutation $\pi \in S_n$ induces a unitary U_π by acting on n qubits in the computational basis as follows:

$$U_\pi |x_1\rangle \otimes |x_2\rangle \cdots \otimes |x_{n-1}\rangle \otimes |x_n\rangle = |x_{\pi(1)}\rangle \otimes |x_{\pi(2)}\rangle \otimes \ldots \otimes |x_{\pi(n)}\rangle \,. \tag{2.41}$$

The group $SU(2)$ has a natural representation on n qubits as follows:

$$Q(U) |x_1\rangle \otimes |x_2\rangle \cdots \otimes |x_{n-1}\rangle \otimes |x_n\rangle = U |x_1\rangle \otimes U |x_2\rangle \otimes \ldots \otimes U |x_n\rangle \,. \tag{2.42}$$

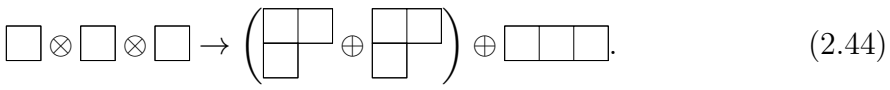
The two actions clearly commute. Schur-Weyl duality [71, 58] then states that:

$$\left(\mathbb{C}^2\right)^{\otimes n} \cong \bigoplus_{\lambda \vdash n} Q_\lambda \otimes P_\lambda,$$

where the direct sum runs over all partitions $\lambda \vdash n$, Q_λ labels irreducible representations of $SU(2)$ and P_λ labels S_n irreps. For example, the three-qubit Hilbert space $(\mathbb{C}^2)^{\otimes 3}$ decomposes, in terms of $SU(2)$ irreps, into:

$$\frac{1}{2} \otimes \frac{1}{2} \otimes \frac{1}{2} \to \left(\frac{1}{2} \oplus \frac{1}{2}\right) \oplus \frac{3}{2}, \tag{2.43}$$

where the brackets also correspond to a two dimensional, mixed-symmetry representation of the symmetric group. In terms of Young diagrams:[3]



(2.44)

The decomposition can be hence also written as $1/2 \otimes 1/2 \otimes 1/2 \to (Q_{3/2} \otimes P_{\square\square\square}) \oplus (Q_{1/2} \otimes P_{\square\!\square})$, where $P_{\square\square\square}$ is the trivial representation of the symmetric group and $P_{\square\!\square}$ is the two-dimensional mixed-symmetry representation of S_3 labeled by $\square\!\square$ [58]. This works similarly with a larger number of qubits.

2.5 Yamanouchi-Kotani Representations

Spin eigenfunctions generate irreducible representations of the symmetric group, called the Yamanouchi-Kotani representations. These representations were discovered by Yamanouchi [69] and Kotani [70]. The presentation here follows Pauncz [56].

The unitary U_π from Eq. 2.41 commutes with both S^2 and Z operators as they are both invariant under permutation of qubit coordinates. This way, the J and M eigenvalues remain invariant under the action of U_π and the corresponding Young diagram λ also remains unchanged. Let λ be a Young diagram with two rows and overhang $2J$. Let $f^\lambda := f$ and $\{\lambda_i\}_{i \in [f]}$ be the set of standard Young tableaux of shape λ. It follows that:

$$U_\pi \left|J, M, \lambda_i\right\rangle = \sum_{j=1}^{f} \left|J, M, \lambda_j\right\rangle W_{ij}(\pi), \tag{2.45}$$

[3]The fact that the dimension of the symmetric group irrep is the multiplicity of Young diagrams in this decomposition is called the Young's rule [71]. The dimensions can be also computed with the Frame-Robinson-Thrall formula of Eq. 2.25.

where $W_{ij}(\pi) \in \mathbb{C}$ are scalars subject to $\sum_j |W_{ij}(\pi)|^2 = 1$. They can be found by:

$$\langle J', M', \lambda_j | \, U_\pi \, | J, M, \lambda_i \rangle = W_{ij}(\pi). \tag{2.46}$$

As both J and M remain invariant under the permutation action, we can write that:

$$\langle \lambda_j \, | \, U_\pi \, | \, \lambda_i \rangle \, \delta_{JJ'} \, \delta_{MM'} = W_{ij}(\pi). \tag{2.47}$$

J-conservation means that the coefficients are nonzero only if λ_i and λ_j have the same shape. Applying another permutation $\sigma \in S_n$ to the eigenfunction gives:

$$U_\sigma \, U_\pi \, | J, M, \lambda_i \rangle = \sum_{k=1}^{f} \sum_{j=1}^{f} | J, M, \lambda_k \rangle \, W_{ij}(\pi) W_{jk}(\sigma). \tag{2.48}$$

We also have that $U_\sigma \, U_\pi = U_{\sigma \circ \pi}$,

$$U_{\sigma \circ \pi} \, | J, M, \lambda_i \rangle = \sum_{j=1}^{f} | J, M, \lambda_k \rangle \, W_{ij}(\sigma \circ \pi). \tag{2.49}$$

Comparing Eqs. 2.48 and 2.49, it follows that:

$$W(\sigma \circ \pi)_{ik} = \sum_{j=1}^{f} W(\sigma)_{ij} W(\pi)_{jk}, \tag{2.50}$$

which shows that the matrices W form a representation of the symmetric group. It will follow later from Theorem 2 that these representations are in fact irreducible. Yamanouchi-Kotani representation matrices of the S_n generators are derived in Appendix A.

Chapter 3

Permutational Quantum Computing

3.1 Introduction

Permutational Quantum Computing is a computational model that can be seen as a toy model of Topological Quantum Computing. It computes not by splitting, braiding and fusing anyons but rather by splitting, permuting and coupling spin eigenstates.

The computational model was introduced by Marzuoli and Rassetti in Refs. [4, 5] that studied the computational power of spin networks. This line of research was inspired by works in quantum gravity that originated from Penrose's proposal in Ref. [66] to use total spin eigenstates as basic building block for modeling granular spacetime. The model was subsequently formalized by Jordan in Ref. [3], where he defined the computational complexity class PQP in an attempt to capture the class of problems solvable by polynomial approximation of the transition amplitudes in the model. He proved that $\mathsf{PQP} \subseteq \mathsf{BQP}$ and conjectured that some of the PQP problems have no efficient classical algorithm.

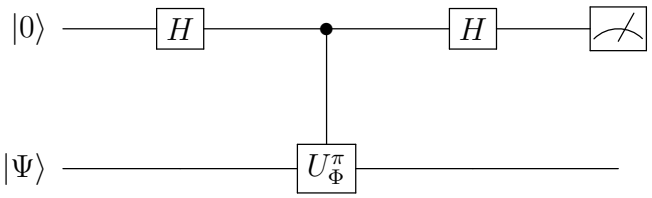
This chapter introduces the computational model and formalizes the definition of the computational complexity class PQP. We discuss the problem of approximating the S_n irreps that belongs to this class and conclude by discussion of Jordan's hardness conjecture.

3.2 Permutational Quantum Computing

Definition 8 (Permutational Quantum Computing [3]). *Given a permutation $\pi \in S_n$ and a classical description of two spin eigenfunctions $|\Psi\rangle$, $|\Phi\rangle$ given as lists of eigenvalues of spin operators on n qubits, the Permutational Quantum Computing model comprises of all problems that can be efficiently solved by sampling a probability distribution $p_\pi(\Psi|\Phi) = |\langle\Psi|U_\pi|\Phi\rangle|^2$ where U_π is an n-qubit unitary induced by permuting the qubits:*

$$U_\pi |x_1\rangle \otimes |x_2\rangle \cdots \otimes |x_{n-1}\rangle \otimes |x_n\rangle = |x_{\pi(1)}\rangle \otimes |x_{\pi(2)}\rangle \otimes \ldots \otimes |x_{\pi(n)}\rangle . \tag{3.1}$$

By repeating the above process, it is possible to estimate the probabilities $p_\pi(\Psi|\Phi) = |\langle\Psi|U_\pi|\Phi\rangle|^2$ to $\pm\epsilon$ error with $O(1/\epsilon^2)$ retries with high probability of success. Jordan also defined mathematically more convenient version of the model capable of estimating both the real and imaginary part of the transition amplitudes $\langle\Psi|U_\pi|\Phi\rangle$. In this model, it is assumed that one can perform an interferometric experiment that determines the real and imaginary parts of $\langle\Psi|U_\pi|\Phi\rangle$ to $\pm\epsilon$ error in $poly(n, 1/\epsilon)$ time with high probability of success. While Jordan's definition does not elaborate on how to do this in practice, such computation could be in principle done by translating the computation to the circuit model using the Quantum Schur Transform and performing the Hadamard test (see Fig 3.1).



(3.2)

Figure 3.1: The Hadamard test shown above has the output probability of measuring the outcome 0 on the first qubit given by $(1+\mathrm{Re}\,\langle\Psi|U_\Phi^\pi|\Psi\rangle)/2$. By taking $O(1/\epsilon^2)$ samples, this can be used to approximate the real part of the transition amplitude to $\pm\epsilon$ error. If U_Φ is a unitary such that $U_\Phi|\Psi\rangle = |\Phi\rangle$. Such unitary can be for example found by applying the angular momentum recoupling rules from Sec 2.4. Setting $U_\Phi^\pi = U_\pi U_\Phi$ in the Hadamard test allows for estimation of $\mathrm{Re}\,\langle\Psi|U_\pi|\Phi\rangle$ to $\pm\epsilon$ error in $O(1/\epsilon^2)$ time. A variant of the test can be also used to approximate the imaginary part of the transition amplitude [53].

The conventional quantum circuit model of quantum computation measures the length of computation in terms of the number of elementary computational gates.

Since any permutational gate U_π decomposes into $O(n^2)$ transpositions and state preparation/measurement are both efficient, any distribution generated by the model in Def. 8 can be sampled from in polynomial time. Jordan thus decided to ignore the computational length and defined the complexity class PQP roughly as '*the set of problems solvable by approximating the transition amplitudes* $\langle\Psi|U_\pi|\Phi\rangle$ *to polynomial precision with polynomially many qubits*' [3].

Ref. [3] also discussed the kind of a classical computer that should control the quantum experiment – that is the machine that transmits the instructions to the experimental apparatus, receives measurement outcomes from the quantum experiment and processes them in order to solve the computational problem. Jordan argues that choosing a P machine would make PQP trivially contain P and that it is more meaningful to use a logspace machine (L) instead for that reason. While he does not elaborate on this choice further, the key idea seems to be an attempt to capture the power of computation by estimating constant number of the transition amplitudes, without the possibility of performing additional (computationally complex) post-processing on them. With this, Jordan updates the previous informal definition of PQP to be '*the set of problems solvable by a logspace machine with access to an oracle that provides amplitudes* $\langle\Psi|U_\pi|\Phi\rangle$ *for polynomially many spins to polynomial precision*' [3]. This definition however assumes access to an oracle that solves a counting problem, which can be defined in many ways. Since Ref. [3] does not formalize the computational class beyond the definition given above, we attempt to deal with this deficiency now.[1]

The computational complexity class L is a class of problems that can be solved by a deterministic Turing machine with a logarithmic amount of work space. The formal definition is as follows:

Definition 9 (Space Bounded Computations and L [80])**.** $s : \mathbb{N} \to \mathbb{N}$ *and* $L \subseteq \{0,1\}^*$*.* $L \in \mathsf{SPACE}(s(n))$ *if there exists a constant* c *and a Turing machine* M *that decides* L*, such that at most* $c \cdot s(n)$ *cells of* M*'s worktapes (excluding the input tape) are ever visited by* M*'s head during its computation on any input. The class* L *of logspace computations is defined as* $\mathsf{L} = \mathsf{SPACE}(\log(n))$*.*

[1] We remark that the simulation results proved later in this work also hold if the controller is a P or a BPP machine as well as for any scenario captured by Jordan's informal definition.

We formalize the class PQP as a logspace oracle class L^O, where O is, roughly speaking, an oracle that provides $poly(n)$ precise additive approximations to the transition amplitudes of Permutational Quantum Computing. Defining logspace oracle machines is subject to several subtleties, so we describe this in some detail now. As the first step, we state the conventional definition of oracle Turing machines:

Definition 10 (Oracle Turing Machines (Arora, Barak [80]))**.** *An oracle Turing machine M is a Turing machine that has a special read-write tape, called oracle tape, along with three special states: q_{query}, q_{yes}, q_{no}. To execute M, one specifies an additional language $O \subseteq \{0,1\}^*$ decided by the oracle. When M enters the state q_{query}, the machine moves to q_{yes} if $q \in O$ and q_{no} if $q \notin O$; where q is the content of the oracle tape.*

Definition 10 becomes problematic for logspace oracle machines whenever the query q written on the oracle tape is of larger than logarithmic size. If we count oracle cells among the cells that were 'visited' by the Turing machine's head, then such oracle machine can only make queries that are of logarithmic size. This is problematic if one needs to pass a large part of the input to the oracle. On the other hand, if we do not count the oracle cells among the 'visited' cells in the definition of space-bounded computations, then the Turing machine can cheat by using the oracle tape instead of its worktape, free of the space bound penalty. This was addressed by Ruzzo, Simon and Tompa in Ref. [81], who updated the definition of the oracle machines for space bounded computation as follows:

Definition 11 (Oracle Turing Machines for Space Bounded Computations (Ruzzo, Simon, Tompa [81]))**.** *The Oracle Machine M is a Turing machine that has a special write-only tape, the oracle tape, whose cells are not subject to the space penalty. The machine has also three special states: q_{query}, q_{yes}, q_{no}. To execute M, one specifies an additional language $O \subseteq \{0,1\}^*$ decided by the oracle for M. When M enters the state q_{query}, the machine state changes to q_{yes} if $q \in O$ and q_{no} if $q \notin O$; where q is the content of the oracle tape. After executing an oracle query, the oracle tape is erased. The oracle tape also operates deterministically from the time some symbol is written on it, until the oracle query is made.*

PQP can be then defined as:

Definition 12 (PQP). *Let O be a Ruzzo, Simon, Tompa oracle that takes as an input:*

- *a classical description of a pair of n-qubit spin eigenfunctions $|\Phi\rangle, |\Psi\rangle$ in terms of eigenvalues in the spin operator basis.*
- *a permutation π.*

The oracle (which mimics the output of the Hadamard test) returns 'yes' with probability $\frac{1}{2}(1+\langle\Phi|U_\pi|\Psi\rangle)$, otherwise it returns 'no' (note that the overlap $\langle\Phi|U_\pi|\Psi\rangle$ is real valued). $\mathsf{PQP} = \mathsf{L}^O$*, where* L *is a logspace machine and the oracle Turing machine is defined as in Def. 11.*

The logspace restriction on the Turing machine worktape guarantees that the machine cannot store or manipulate better than $poly(n)$-precise additive approximation to the transition amplitudes. To obtain a polynomially precise approximation to the transition amplitude with high probability of success, the algorithm queries the oracle $p(n)$-times for some fixed polynomial $p(n)$ and keeps track of the number of 'yes' and 'no' answers in its worktape (this takes $O(log(n))$ space). The transition amplitude is then approximated by the average $\bar{\mu}$ of those two numbers. By the Chernoff-Hoeffding inequality for the Bernoulli process, we have that:

$$\Pr\left(|\bar{\mu} - \mu| > \epsilon\right) \leq 2\exp(-2p(n)\epsilon^2). \tag{3.3}$$

Setting $\epsilon = 1/q(n)$ for some polynomial $q(n)$, we obtain that $\bar{\mu}$ is a $\pm 1/q(n)$ approximation to $\mu = \frac{1}{2}(1 + \langle\Phi|U_\pi|\Psi\rangle)$ with probability at least $1 - 2\exp(-2p(n)/q(n)^2)$. Choosing $p(n) > q(n)^2$, we can upper-bound $p(n)/q(n)^2$ by another polynomial, $r(n)$, to obtain that $\bar{\mu}$ is a good approximation with probability of at least $1 - 2\exp(-2r(n)) = 1 - \delta$. Note that owing to the Ruzzo, Simon and Tompa oracle definition, the Turing machine can now pass an arbitrarily large part of its input as a part of the oracle query, writing to the oracle tape bit by bit.

See Fig. 3.2 for the relation between P, PQP and BQP that was conjectured in [3].

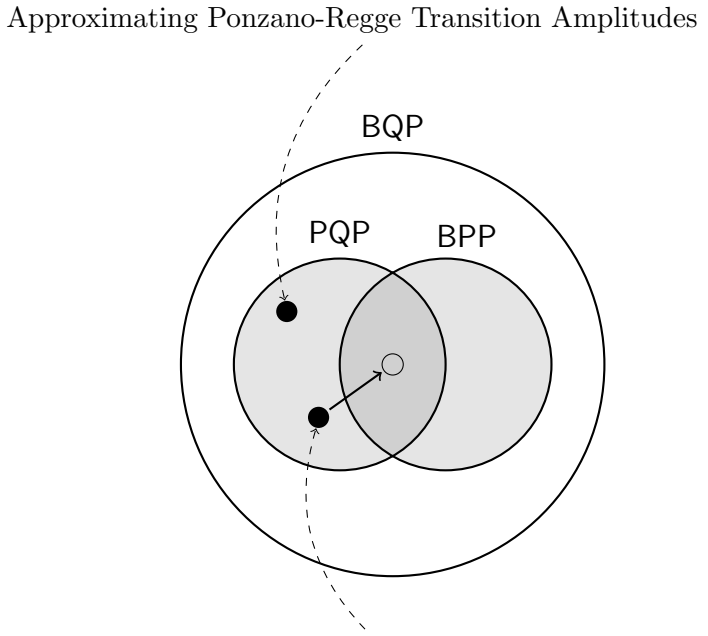


Figure 3.2: The computational complexity class PQP and its conjectured relationship to BPP and BQP. Jordan proved that PQP $\subseteq$ BQP and conjectured that P and PQP were incomparable. Following the discussion about the definition of PQP, we extend this to a conjecture that BPP and PQP are incomparable. To support the conjecture, Jordan studied two problems with no known efficient classical algorithms that are solvable within PQP: Approximation of a class of Ponzano-Regge transition amplitudes and Approximating Matrix Elements of S_n Irreducible representations in the Young orthogonal form. The contribution of our work [30] is a classical algorithm that shows that the irrep matrix element approximation algorithm is in fact in BPP (the thick arrow in the above diagram). We extend this result to show that PQP $\subseteq$ BPP by extending this algorithm to the problem of approximating the transition amplitudes between any pair of spin eigenfunctions. This classical probabilistic algorithm simulates the oracle O in polynomial time.

3.3 Approximating Irreps of the Symmetric Group

Jordan presented two problems as arguments that the class PQP could have supra-classical computational power in Ref. [3]. One of them, introduced in Ref. [5], is concerned with the approximation of a set of transition amplitudes in the Ponzano-Regge theory. Ponzano-Regge is a 3-dimensional Topological Quantum Field Theory that gives rise to non-trivial topological invariants that play role in the study of quantum gravity. We do not discuss this problem here as its statement is quite complex, but mention that it is contained in the class PQP. The second problem discussed in Ref. [3, 24] is however the central to this part of the thesis:

Problem 1 (Approximating Matrix Elements of S_n Irreducible Representations in Young's orthogonal form [3])**.**

Problem: *Approximate a given matrix element of Young's orthogonal form matrices to $\pm 1/poly(n)$ precision.*

Input*: A permutation $\pi \in S_n$, a Young diagram λ on two rows, a pair of standard Young tableaux λ_1, λ_2 of shape λ and a parameter $\epsilon = 1/poly(n)$.*

Output: *With high success probability, a matrix element $W(\pi)_{\lambda_1\lambda_2}$ of Young's orthogonal form representation matrix $W(\pi)$ to $\pm\epsilon$ precision.*

We now describe a PQP algorithm for Problem 1, which follows from the correspondence of Permutational Quantum Computing transition amplitudes and matrix elements of Young's orthogonal form matrices. In order to explicitly construct the S_n representation matrices, it is necessary to fix a basis. Young's orthogonal form is a canonical choice of a basis, subgroup-adapted to a chain $S_2 \subset S_3 \dots S_{n-1} \subset S_n$ [3]. The subgroup adaptation property means that if the representation of the group S_n becomes restricted to its subgroup S_{n-1}, the representation matrices are going to be matrix direct sums of irreducible representation of S_{n-1}. It is defined by the following prescription for computing its generators:

Definition 13 (Young's orthogonal form [56])**.** *For a fixed Young diagram λ, number the standard Young tableaux $\{\lambda_i\}_i$ of shape λ from 1 to f^λ, where f^λ is given by the Frame-Robinson-Thrall formula of Eq. 2.25. An orthogonal representation of S_n – the Young's orthogonal form – can be obtained by representing the nearest-neighbor transpositions σ_k as $f^\lambda \times f^\lambda$ orthogonal matrices $W(\sigma_k)$ with the following matrix elements:*

$$W(\sigma_k)_{ii} = -1/d^i_{k,k+1} := \rho, \tag{3.4}$$

$$W(\sigma_k)_{ij} = \begin{cases} 0 & \text{if the standard tableau } \lambda_i \text{ cannot be obtained from the standard tableau } \lambda_j \text{ by swapping the } k\text{-th and } k+1\text{-st elements.} \\ \sqrt{1-\rho^2} & \text{otherwise.} \end{cases} \tag{3.5}$$

Here $d^i_{k,k+1}$ denotes the axial distance between the numbers k and $k+1$ in a standard Young tableau λ_i, which is the number of steps necessary to get from k to $k+1$ when steps left and down are counted as positive and steps to the right and up as negative.

Since any permutation can be composed by at most $O(n^2)$ nearest-neighbor transpositions it is possible to represent any permutation from S_n by multiplying at most $O(n^2)$ representation matrices. Observe that while the matrix elements of transposition matrices in Young's orthogonal form can be easily found, finding matrix elements for general permutations may be difficult, especially if f^λ is very large. The following theorem is then key to the quantum algorithm for approximating the matrix elements of the Young's orthogonal form with PQP.

Theorem 2 (Pauncz [56])**.** *Yamanouchi-Kotani representation and Young's orthogonal representations are equivalent.*

The proof is deferred to Appendix A.2. The theorem implies that the transition amplitudes between a pair of genealogical spin eigenfunctions with an applied permutation are matrix elements of Young's orthogonal form matrices. The algorithm for Problem 1 then simply queries the oracle O for these transition amplitudes (See Algorithm 1).

Algorithm 1: Approximating S_n irreps with PQP

Input:

A permutation π and two lists of genealogical spin operator eigenvalues:

$$(J, M, J_{[n-1]}, J_{[n-2]} \dots J_{\{1,2\}}),$$
$$(J, M, J'_{[n-1]}, J'_{[n-2]} \dots J'_{\{1,2\}}).$$

Algorithm:

Use the PQP oracle O to estimate the overlap $\langle \Psi | U_\pi | \Phi \rangle$ for:

$$|J, M, J_{[n-1]}, J_{[n-2]} \dots J_{\{1,2\}}\rangle := |\Psi\rangle\,,$$
$$|J, M, J'_{[n-1]}, J'_{[n-2]} \dots J'_{\{1,2\}}\rangle := |\Phi\rangle\,.$$

By definition of PQP, it is possible to estimate $\langle \Psi | U_\pi | \Phi \rangle$ to $\pm 1/poly(n)$ precision in $poly(n)$ time with high success probability (See Fig. 3.3).

By the correspondence outlined in Chapter 2, the total spin eigenvalues define standard Young tableaux with two rows, so this solves Problem 1.

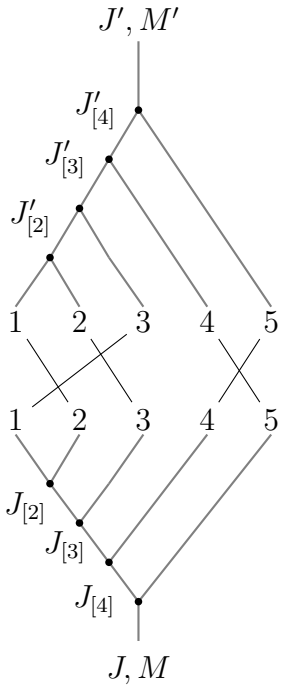


Figure 3.3: Permutational Quantum Computing circuit used to estimate the Young's orthogonal form matrix elements. The applied permutation is $\pi = (1, 2, 3)(4, 5)$ and the transition amplitudes $\langle \Phi | U_\pi | \Psi \rangle$ between genealogical spin eigenfunctions $|\Phi\rangle, |\Psi\rangle$ correspond to the Young's orthogonal form matrix $W(\pi)$. The quantum algorithm that underlies Jordan's hardness conjecture approximates any of these transition amplitudes to polynomial precision.

3.4 Jordan's Hardness Conjecture

There are several reasons to believe that Problem 1 might be hard to solve classically. In Ref. [24], Jordan points out that there are optimized algorithms for finding the Young orthogonal form matrices [82, 83, 84, 85, 86, 87, 88]. All of these algorithms have exponential worst-case running time but also calculate the entire representation matrices. Aside from these, there seems to be no literature on approximating the representation matrix elements and, to the best of Jordan's and my knowledge, no other algorithm than the one we later derive here. The reason why classical algorithms for the task are missing seems tied to the question if a solution to Problem 1 could be used for anything practical. While its solution may be used to find polynomially precise approximations to all polynomially large elements in rows of the representation matrices, it remains unclear how to use such approximations for any computation with the symmetric group representations. Similarly to Jordan, we do not have an answer

to this question; the possibility seems intriguing though.

Furthermore, Jordan also argues that he expects Problem 1 to be trivial on most inputs. Since an average matrix element of a $d \times d$ orthogonal matrix U has magnitude of a square root of its inverse dimension [24]: $\sqrt{\frac{1}{d^2}\sum_{i,j\in d}|U_{ij}|^2} = \sqrt{\frac{1}{d^2}\mathrm{Tr}(U^\dagger U)} = 1/\sqrt{d}$, the typical matrix elements of an exponentially-sized U are exponentially small and will not be resolved by polynomially-precise approximation. At the same time, the interesting instances of Problem 1 are precisely those with representation matrices that have exponentially large dimensions, as the polynomially large representation matrices can be computed efficiently exactly. Jordan concludes that an algorithm that simply guesses '0' at all times achieves similar results to the quantum algorithm. It may be possible that the hardness of classically simulating the model comes not from approximating the individual transition amplitudes, but rather from finding the inputs for which the approximation becomes non-trivial. We study this in Chapter 5.

The fact that the PQP-complete algorithm is average-case trivial however does not mean that it becomes trivial in the worst-case. Another problem that shows similar behavior is in fact the Jones' polynomial approximation by Aharonov-Jones-Landau algorithm outlined in Sec. 1.6 [7]. It also relies on estimating matrix elements of large unitaries, now of braid group representations. A typical matrix element is exponentially small by the same argument as here, but approximating the transition amplitudes of the model to polynomial precision on worst-case inputs remains a BQP-complete problem. This analogy lead Jordan to conjecture that:

Conjecture 1. *There is no classical polynomial time algorithm for Problem 1 [3].*

Chapter 4

Classical Algorithm for Approximation of the Transition Amplitudes

This chapter is based on the article [30] written in collaboration with S. Strelchuk. I also benefited from discussions with S. P. Jordan and K. P. Temme. The analysis of the computational model in terms of cascaded $SU(2)$ Clebsch-Gordan coefficients and their simulation was initiated by myself, as was the application of van den Nests's result to compute the transition amplitudes. The cascading algorithm for sampling the output in the computational basis is my idea, but I benefited from discussions with S. Strelchuk when developing it. Van den Nest's results were pointed out to me as a possible approach to the analysis of the model by K. Temme. S. Strelchuk helped me to work out the details of the $SU(2)$ Clebsch-Gordan coefficient computations and to solidify the results to a publishable form. S. Jordan told me about previous attempts at the proof of the hardness conjecture and contributed with several ideas about where to take it next.

4.1 Introduction

Contrary to Conjecture 1, we now derive a classical algorithm that solves Problem 1 in polynomial time. We subsequently extend this algorithm to approximating transition amplitudes between arbitrary spin eigenfunctions and show that the class PQP can be efficiently simulated classically.

The algorithm is derived by using methods for probabilistic simulation of quantum computation developed by van den Nest in Ref. [6]. We briefly outline this framework

and show that spin eigenfunctions satisfy the computational tractability property. This allows us to use a simple probabilistic algorithm based on the Chernoff-Hoeffding bound for approximating their overlaps to the required precision.

4.2 Computational Tractability

When quantum computation is to be simulated with a classical algorithm, the aim could be to efficiently compute the probability of any outcome, as well as any marginal of the output, to ϵ additive error in $poly(n, \log(1/\epsilon))$ time. This is referred to as strong simulation.[1] Strong simulation is generally #P-hard, as it asks for much more than what would be possible to solve with a quantum computer – it is not expected to resolve the output probabilities to better than $\pm\epsilon$ additive error in $poly(n, 1/\epsilon)$ time in this context. It therefore seems more appropriate to impose that the classical simulator only *samples* from the output probability distribution of the quantum circuit, which is referred to as weak simulation [6].

Weak simulation is *essentially* implied by strong. This is shown, for a somewhat stricter definition of strong simulation than we use here, by Jozsa and van den Nest in Ref. [90] by using the result of Terhal and DiVincenzo from Ref. [18]. The technique specifically uses telescoping marginal expansion to produce samples from the output distribution and is similar to one of the proofs given here in Sec. 4.3. Let the output probability distribution of a quantum circuit be given by $p : \{0,1\}^n \to [0,1]$. Using the chain rule, this can be expanded into a product of marginals as:

$$p(x_1, x_2, \ldots, x_n) = p(x_1)p(x_2|x_1)p(x_3|x_2, x_1)\ldots p(x_n|x_{n-1}\ldots, x_2, x_1).$$

Assuming first that the conditional probabilities can be sampled from *exactly*, one can sample the output by flipping a coin with bias $p(x_1)$, fixing x_1 to the outcome, flipping a biased coin with bias $p(x_2|x_1) = p(x_2, x_1)/p(x_1)$, fixing x_2 to the outcome, and so on. The resulting bitstring $x_1, x_2 \ldots x_n$ from this random process is then distributed according to $p(x_1, x_2 \ldots x_n)$. If the conditionals can be instead only approximated

[1] Other definitions are commonly used in the literature. For example, Ref. [89] defines strong simulation as computation of output probabilities on all marginals up to multiplicative errors while Ref. [16] as computation of thereof to m significant digits in $poly(n, m)$ time (which gives $\pm\epsilon$ additive approximation in $poly(n, \log(1/\epsilon))$ time and matches the definition used here).

to $\pm\epsilon$ error, then the same technique samples from $p(x_1, x_2, \ldots x_n) \pm O(n\epsilon)$.[2] The sampled distribution is $O(\epsilon\, n\, 2^n)$ away from the original one in the total variational distance and since the samples will be usually produced in $O(\log \epsilon)$ time, it is possible to bind ϵ to n in a way that the overall error decreases with n in a $poly(n)$-time computation – this will be the case in the proof in Sec. 4.3. I am however not aware of any result that would show that strong implies weak, assuming *only* that the marginals can be computed to only $\pm\epsilon$ additive precision in $poly(n, \log(1/\epsilon))$ time.

In Ref. [6], van den Nest developed a set of classical simulation algorithms built around sampling methods that can be applied to weak simulation of quantum computations. In that context, he identified a class of quantum states called *computationally tractable* (CT) states:

Definition 14 (Computational Tractability [6])**.** *An n-qubit state $|\psi\rangle$ is computationally tractable (CT) if:*

1. *the overlaps $\langle x|\psi\rangle$ can be approximated efficiently and with large success probability to $\pm\epsilon$ error in time $poly(n, \log(1/\epsilon))$ for any computational basis state $|x\rangle$.*

2. *there exists a classical randomized algorithm that draws samples from $p(x) = |\langle x|\psi\rangle|^2$ for $x \in \{0,1\}^n$ in polynomial time.*

It will suffice here if the classical algorithm in condition 2. only samples from a distribution $\tilde{p}$ for which $|\tilde{p}(x) - |\langle x|\psi\rangle|^2| \leq \epsilon$ for all $x \in \{0,1\}^n$ in $poly(n, \log(1/\epsilon))$ time. We justify this technical detail in Sec. 4.5.

Examples of CT states listed in Ref. [6] include product states, states obtained by a polynomially-sized Toffoli circuits applied to a product state, matrix product

[2] A justification of this is as follows: let $\kappa_i \in [-\epsilon, \epsilon]$, $\forall i \in [n]$ and consider the telescoping expansion that uses approximations to conditionals: $\tilde{p}(x_1) = p(x_1) + \kappa_1$, $\tilde{p}(x_2|x_1) = p(x_2|x_1) + \kappa_2$, ... etc.. By the chain rule,

$$\tilde{p}(x_1, x_2, x_3 \ldots x_n) = \tilde{p}(x_1)\tilde{p}(x_2|x_1)\ldots\tilde{p}(x_n|x_{n-1}\ldots x_2, x_1) = (p(x_1) + \kappa_1)(p(x_2|x_1) + \kappa_2)\ldots \quad (4.1)$$

To the leading order in ϵ:

$$|p(x_1, x_2, x_3 \ldots x_n) - \tilde{p}(x_1, x_2, x_3 \ldots x_n)| \leq n\epsilon \frac{p(x_1, x_2 \ldots)}{\min[p(x_1), p(x_2|x_1)\ldots]} + O(\epsilon^2) \leq n\epsilon + O(\epsilon^2), \quad (4.2)$$

where the last inequality again follows from the chain rule.

states with a polynomial bond dimension, stabilizer circuits or Quantum Fourier Transform applied to a computational basis state. We extend this list by showing that genealogical spin eigenfunctions are computationally tractable.

4.3 Genealogical Spin Eigenfunctions are CT

We first prove that the overlap condition of Def. 14 holds for genealogical spin eigenfunctions.

Claim 9. *The overlap $\langle x|J, M, \ldots\rangle$ for a genealogical spin eigenfunction $|J, M, \ldots, J_{[2]}\rangle$ and a computational basis state $|x\rangle$, $x \in \{0,1\}^n$, can be computed to $\pm\epsilon$ precision in $poly(n, \log(1/\epsilon))$ time.*

Proof. The Clebsch-Gordan coefficients satisfy:

$$C^{J,M}_{J_A,M_A;\,J_B,M_B} = \delta_{M,\,(M_A+M_B)}\, C^{J,M}_{J_A,M_A;\,J_B,M_B}. \tag{4.3}$$

A genealogical spin eigenstate $|J, M, \ldots J_{[2]}\rangle$ can be written as:

$$\begin{aligned}|J, M, \ldots J_{[2]}\rangle &= \sum_{M_{[2]},\ldots M_{[n-1]}} \sum_{M_1,\ldots M_n} C^{J,M}_{J_{[n-1]},\,M_{[n-1]};M_n} \ldots C^{J_{[3]},M_{[3]}}_{J_{[2]},M_{[2]};M_3} C^{J_{[2]},M_{[2]}}_{M_1;M_2} |M_1, \ldots M_n\rangle \\ &= \sum_{M_1,\ldots M_n} C^{J,M}_{J_{[n-1]},\sum_{k=1}^{n-1} M_k;\,M_n} \ldots C^{J_{[3]},M_1+M_2+M_3}_{J_{[2]},M_1+M_2;M_3} C^{J_{[2]},M_1+M_2}_{M_1;M_2} |M_1 \ldots M_n\rangle\,,\end{aligned}$$

where Eq. 4.3 was used to cancel the summation over $M_{[2]}, \ldots M_{[n-1]}$ and we omitted qubit total spins from the arguments of the Clebsch-Gordan coefficients. The overlap $\langle x|J, M, J_{[n-1]}, \ldots, J_{[2]}\rangle$ is then given by:

$$\left\langle x \,\middle|\, J, M, J_{[n-1]}, \ldots, J_{[2]}\right\rangle = C^{J,M}_{J_{[n-1]},\,\sum_{k=1}^{n-1} x_k;\,\frac{1}{2},\,x_n} \ldots C^{J_{[3]},x_1+x_2+x_3}_{J_{[2]},\,x_1+x_2;\,\frac{1}{2},\,x_3} C^{J_{[2]},x_1+x_2}_{\frac{1}{2},\,x_1;\,\frac{1}{2},\,x_2}.$$

Each of the Clebsch-Gordan coefficients can be computed to $\pm\epsilon$ error in $poly(n, \log(1/\epsilon))$ time by results of Sec. 2.4. The product of $n-1$ of them can be computed to this precision in $poly(n, \log(1/\epsilon))$ time. □

We now prove the second computational tractability condition:

Claim 10. *Given a genealogical spin eigenfunction $|\Phi\rangle = |J, M, J_{[n-1]}, \ldots, J_{[2]}\rangle$, there exists a classical randomized algorithm that draws samples from a distribution $\tilde{p}(x)$ such that $|\tilde{p}(x) - |\langle x|\Phi\rangle|^2| \leq \epsilon$ for all $x \in \{0,1\}^n$. The algorithm runs in $poly(n, \log(1/\epsilon))$ time.*

Proof. Importantly note that M and all values of the total spins $J, J_{[n-1]}, \ldots, J_{[2]}$ are fixed and known for all distributions computed in this proof. We first show that it is possible to efficiently compute the 'suffix' marginals of the output probability distribution $p(x) := |\langle x|J, M, J_{[n-1]}, \ldots J_{[2]}\rangle|^2$ defined for $k \leq n$ by:

$$p(x_k \ldots x_n) := \sum_{x_1 \ldots x_{k-1}} p(x_1 \ldots x_n).$$

and $p(x_k \ldots x_n) = 1$ otherwise. This is achieved by successive application of:

$$\begin{aligned} p(x_k \ldots x_n) &= \sum_{x_1 \ldots x_{k-1}} p(x_1 \ldots x_n) \\ &= \sum_{x_1 \ldots x_{k-1}} \left| C^{J,M}_{J_{[n-1]}, \sum_{i=1}^{n-1} x_i;\, \frac{1}{2}, x_n} C^{J_{[n-1]}, \sum_{i=1}^{n-1} x_i}_{J_{[n-2]}, \sum_{i=1}^{n-2} x_i;\, \frac{1}{2}, x_{n-1}} \ldots \right|^2 \\ &= \left| C^{J,M}_{J_{[n-1]}, (M-x_n);\, \frac{1}{2},\, x_n} \right|^2 p(x_k \ldots x_{n-1}), \end{aligned}$$

where we again used the correspondence of Eq. 2.9 and Eq. 4.3 was used to proceed to the last line. Repeating this $n-k$ times and noting that by normalization:

$$\sum_{x_1 \ldots x_{k-1}} p(x_1 \ldots x_{k-1}) = 1,$$

gives the marginals as factors of at most $n-1$ Clebsch-Gordan coefficients.

Note that the very reason we compute the 'suffix' marginals – that is starting from the root of the corresponding recoupling diagram – is that all the samples are subject to the constraint of $\sum_{i=1}^{n} x_i = M$ that fixes all the inputs of the CG coefficients. For example:

$$\begin{aligned} p(x_n) &= \left| C^{J,M}_{J_{[n-1]}, (M-x_n);\, \frac{1}{2},\, x_n} \right|^2, \\ p(x_n,\, x_{n-1}) &= \left| C^{J,M}_{J_{[n-1]}, (M-x_n);\, \frac{1}{2},\, x_n} \right|^2 \left| C^{J_{[n-1]}, (M-x_n)}_{J_{[n-2]}, (M-x_n-x_{n-1});\, \frac{1}{2},\, x_{n-1}} \right|^2. \end{aligned} \tag{4.4}$$

We can similarly compute conditional probabilities such as:

$$p(x_{n-1}|x_n) = \frac{p(x_{n-1}, x_n)}{p(x_n)} = \left| C^{J_{[n-1]}, (M-x_n)}_{J_{[n-2]}, (M-x_n-x_{n-1});\, \frac{1}{2},\, x_{n-1}} \right|^2. \tag{4.5}$$

Note that the fact that both $p(x_n)$ and $p(x_{n-1}, x_n)$ are both computable to the right precision would not immediately imply that the conditionals could be computed to the same precision. We can see here however that the conditional is simply another

Clebsch-Gordan coefficient that can be approximated to machine precision. This property is a key to the classical simulation technique developed here.

Using the chain rule for probability distributions, it is then possible to write:

$$p(x_1 \dots x_n) = p(x_n)\, p(x_{n-1}|\, x_n) \dots p(x_1|x_n, x_{n-1}, \dots, x_2).$$

This way, the distribution can be sampled from by successively drawing individual bits x_i as follows. First, compute $b_n = p\,(x_n = 1/2)$ and flip a coin with a bias b_n. Set $x_n = \pm 1/2$ depending on the outcome. Fix x_n and compute $b_{n-1} = p\,(x_{n-1} = 1/2\,|x_n)$. Set $x_{n-1} = \pm 1/2$ depending on the outcome. Proceed in the same way up to x_1. The resulting sample $x = x_1 x_2 \dots x_n$ (treated as bitstring) is drawn from $p(x) = p(x_1 x_2 \dots x_n)$ by the chain rule.

All Clebsch-Gordan coefficients (including the products of up to $n-1$ of them) were computed with precision of $\pm\epsilon$ in $poly(n, \log(1/\epsilon))$ time. This means that the algorithm samples from $\tilde{p}(x)$, such that $|\tilde{p}(x) - |\,\langle x|\Psi\rangle\,|^2| \leq \epsilon$ for all $x \in \{0,1\}^n$. $\square$

4.4 Classical Algorithm

Here we give an efficient classical algorithm for Problem 1 that refutes Conjecture 1. The algorithm follows from the above results combined with the following:

Theorem 3 (Van den Nest's algorithm [6])**.** *Given two computationally tractable states $|\Phi\rangle$ and $|\Psi\rangle$, the overlap $\langle\Psi|\Phi\rangle$ can be approximated to $\pm\epsilon$ error in $poly(n, 1/\epsilon)$ time.*

Proof. Define the two distributions induced by sampling in the computational basis:

$$p(x) := |\,\langle x|\Psi\rangle\,|^2, \qquad\qquad q(x) := |\,\langle x|\Phi\rangle\,|^2\,.$$

Consider an indicator function:

$$\Upsilon(x) = \begin{cases} 1 \text{ if } p(x) > q(x), \\ 0 \text{ if } p(x) \leq q(x), \end{cases}$$

and define:

$$F(x) := \frac{\langle\Psi|x\rangle\,\langle x|\Phi\rangle}{p(x)}\Upsilon(x), \qquad\qquad G(x) := \frac{\langle\Psi|x\rangle\,\langle x|\Phi\rangle}{q(x)}(1-\Upsilon(x)).$$

Both of these functions have efficient $\pm\epsilon'$ approximation computable in $poly(n, \log(1/\epsilon'))$ time, since $|\Psi\rangle$ and $|\Phi\rangle$ are both computationally tractable. Moreover, notice that $|F|, |G| \leq 1$. We can write:

$$\langle\Psi|\Phi\rangle = \sum_x p(x)F(x) + \sum_x q(x)G(x)\,.$$

A polynomially-precise additive approximation to the state overlaps is then obtained by applying the Chernoff-Hoeffding bound (see Ref. [6], Appendix A). This implies that for $T = \frac{4}{\epsilon^2}\log\left(\frac{2}{\delta}\right)$ samples $\{F_i\}$ and a real random variable F with expectation $\langle F\rangle$ and $|F| \leq 1$, we have that:

$$Pr\left(\left|\frac{1}{T}\sum_k^T F_k - \langle F\rangle\right| \leq \epsilon\right) \geq 1 - \delta\,. \tag{4.6}$$

This can be used to estimate both the imaginary and real parts of the overlap $\langle\Psi|\Phi\rangle$ to $\pm\epsilon$ error by taking $O\left(\frac{1}{\epsilon^2}\log(\frac{1}{\delta})\right)$ samples of $F \sim p$ and $G \sim q$. □

If $|\Psi\rangle$ and $|\Phi\rangle$ are two genealogical spin eigenfunctions, it is a consequence of the above that the transition amplitude $\langle\Psi|U_\pi|\Phi\rangle$ can be approximated efficiently classically to polynomial precision by fixing $\epsilon = 1/poly(n)$ in the above algorithm. This follows from the fact that if $|\Phi\rangle$ is computationally tractable, then $U_\pi\,|\Phi\rangle$ is, which can be seen by noting that the overlap $\langle x|U_\pi|\Phi\rangle = \langle\pi^{-1}(x)|\Phi\rangle$ can be efficiently computed and the distribution $q_\pi(x) = |\,\langle x|U_\pi|\Phi\rangle\,|^2$ can be efficiently sampled from by sampling $q(x) = |\,\langle x|\Phi\rangle\,|^2$ and applying π^{-1} to the sample.

This gives a classical algorithm for Problem 1 and refutes Conjecture 1.

4.5 Remark about Precision

The proposed algorithm does not sample from exact distributions, but instead works with distributions that have output probability of every outcome exponentially close to exact one. Here we argue that this error is irrelevant for the approximation given by Eq. 4.6.

Suppose that the classical algorithm computes the approximation:

$$\widetilde{\langle\Psi|\Phi\rangle} = \sum_x \tilde{p}(x)\tilde{F}(x) + \sum_x \tilde{q}(x)\tilde{G}(x),$$

where we can efficiently sample $\tilde{q}$ and $\tilde{p}$ in $poly(n, \log(1/\epsilon))$ time, such that $|\tilde{p}(x) - p(x)| \leq \epsilon$ and $|\tilde{q}(x) - q(x)| \leq \epsilon$ for all x and that $\tilde{F}(x)$ and $\tilde{G}(x)$ are also both exponentially close to $F(x)$ and $G(x)$. For simplicity, we set that $\left|\tilde{F}(x) - F(x)\right| \leq \epsilon$. By $|F(x)| \leq 1$, this implies that $|\tilde{F}(x)| \leq 1 + \epsilon$. Consider:

$$\left|\widetilde{\langle\Psi|\Phi\rangle} - \langle\Psi|\Phi\rangle\right| \leq \left|\sum_x \tilde{p}(x)\tilde{F}(x) - p(x)F(x)\right| + \left|\sum_x \tilde{q}(x)\tilde{G}(x) - q(x)G(x)\right|. \quad (4.7)$$

Focusing on the first term, it follows by application of triangle inequality that:

$$\begin{aligned}
\left|\sum_x \tilde{p}(x)\tilde{F}(x) - p(x)F(x)\right| &= \left|\sum_x \tilde{p}(x)\tilde{F}(x) - p(x)\tilde{F}(x) + p(x)\tilde{F}(x) + p(x)F(x)\right| \\
&\leq \left|\sum_x \tilde{p}(x)\tilde{F}(x) - p(x)\tilde{F}(x)\right| + \left|\sum_x p(x)\tilde{F}(x) + p(x)F(x)\right| \\
&\leq \left|\sum_x \tilde{p}(x) - p(x)\right| |\tilde{F}(x)| + \max_x \left|\tilde{F}(x) - F(x)\right| \\
&\leq \sum_x |\tilde{p}(x) - p(x)|(1 + \epsilon) + \epsilon \lessapprox 2^n\epsilon + O(\epsilon^2).
\end{aligned}$$

Both sampling and computing the $\tilde{F}(x)$ approximation can be done in $poly(n, \log(1/\epsilon))$ time and the same holds for $\tilde{G}(x)$. Setting (say) $\epsilon = 4^{-n}$ then leads to:

$$\left|\widetilde{\langle\Psi|\Phi\rangle} - \langle\Psi|\Phi\rangle\right| \lessapprox 2^{-n}, \quad (4.8)$$

which shows that the expectation values are exponentially close to one another. The best approximation to the overlap in Eq. 4.6 by the Chernoff-Hoeffding bound is however only polynomially precise. This means that the error due to sampling of an approximating distribution becomes insignificant.

4.6 Spin Eigenfunctions are CT

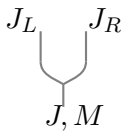
We now extend the previous result to arbitrary spin eigenfunctions. This gives an efficient classical probabilistic algorithm for all PQP problems.

Claim 11. *The overlap $\langle x|\Phi\rangle$ of an n-qubit spin eigenfunction $|\Phi\rangle$ and a computational basis state $|x\rangle$ for $x \in \{0,1\}^n$ can be computed to $\pm\epsilon$ precision in $poly(n, \log(1/\epsilon))$ time.*

Proof. A spin eigenfunction $|\Phi\rangle$ has a coupling diagram with $n-1$ internal vertices, each of which corresponds to a Clebsch-Gordan transform. The state can be therefore expressed as a linear combination of computational basis states, each weighted by a product of $n-1$ Clebsch-Gordan coefficients – all intermediate sums of the coefficients vanish by azimuthal angular momentum conservation. The overlap $\langle x|\Phi\rangle$ is then given by a product of $n-1$ Clebsch-Gordan coefficients. Each of the coefficients can be computed to $\pm\epsilon$ error in $poly(n, \log(1/\epsilon))$ time and their product can be computed with the same precision. □

Claim 12. *Given a spin eigenfunction $|\Phi\rangle$, there exists a classical algorithm that draws samples from $\tilde{p}(x)$ such that $|\tilde{p}(x) - |\langle x|\Phi\rangle|^2| \leq \epsilon$ for all $x \in \{0,1\}^n$ in time $poly(n, \log 1/\epsilon)$.*

Proof. Starting at the root of the coupling diagram of $|\Phi\rangle$:

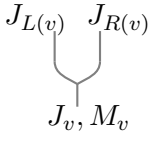


draw a sample:

$$M_L \sim p(M_L) = \left| C^{J,M}_{J_L,M_L;\, J_R,(M-M_L)} \right|^2 . \tag{4.9}$$

Fix $M_R = M - M_L$ and proceed recursively to the internal vertices rooted at J_L, M_L and J_R, M_R; all the way to the leaves. Output the set of leaf values $x_1 \ldots x_n$ converted to a bitstring.

Let V be the set of internal vertices of the coupling diagram. Denote by J_v, M_v the 'incoming' values of the total spin and azimuthal angular momenta to the vertex, and $J_{R(v)}, M_{R(v)}$ and $J_{L(v)}, M_{L(v)}$ the 'outcoming' values on the left/right branch respectively:



With this definition, the output distribution of the algortihm is given by:

$$p(x_1 x_2, \ldots x_n) = \prod_{v \in V} \left| C^{J_v,M_v}_{J_{L(v)},M_{L(v)};\, J_{R(v)},M_{R(v)}} \right|^2 = |\langle x|\Psi\rangle|^2 , \tag{4.10}$$

where the values of M_v for all internal vertices are fixed by the bitstring x of azimuthal momenta associated to the leaves. There are at most $n+1$ distinct values of M_v at every vertex $v \in V$, because all J_v values are all upper bounded by $n/2$. This shows that the samples can be drawn efficiently. □

Example 3. *We give an example of the above algorithm for a spin eigenfunction* $|\Psi\rangle = |J, M, J_R, J_L\rangle$ *with a recoupling diagram that is a binary tree on* 4 *qubits (see Fig. 4.1).*

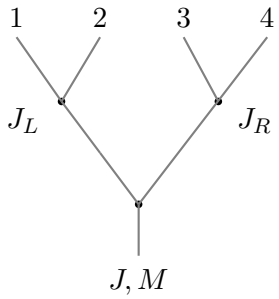


Figure 4.1: Binary tree spin eigenfunction on n qubits.

The algorithm starts at the root by sampling:

$$M_L \sim \left| C^{J,M}_{J_L,M_L;\, J_R,(M-M_L)} \right|^2 = p(M_L). \tag{4.11}$$

Since the value of M *is fixed, sampling* M_L *also fixes* $M_R = M - M_L$ *by Eq. 4.3. Proceeding to the left and right subtree, draw:*

$$\begin{aligned} x_1 &\sim \left| C^{J_L,M_L}_{\frac{1}{2},x_1;\, \frac{1}{2},(M_L-x_1)} \right|^2 = p(x_1|\, M_L). \\ x_3 &\sim \left| C^{J_R,M_R}_{\frac{1}{2},x_3;\, \frac{1}{2},(M_R-x_3)} \right|^2 = p(x_3|\, M_L). \end{aligned} \tag{4.12}$$

Note that x_1, x_3 *fix the value of* $x_2 = M_L - x_1$ *and* $x_4 = M_R - x_3$ *respectively and that the distributions* $p(x_1|\, M_L)$ *and* $p(x_3|\, M_L)$ *are conditionally independent:*

$$p(x_1|\, M_L)p(x_3|\, M_L) = p(x_1, x_3|\, M_L). \tag{4.13}$$

The joint distribution on the samples is given by:

$$\begin{aligned} p(x_1, x_3, M_L) &= p(x_1, x_3|M_L)p(M_L) \\ &= \left| C^{J,M}_{J_L,M_L;\, J_R,(M-M_L)} \right|^2 \left| C^{J_L,M_L}_{\frac{1}{2},x_1;\, \frac{1}{2},(M_L-x_1)} \right|^2 \left| C^{J_R,M_R}_{\frac{1}{2},x_3;\, \frac{1}{2},(M_R-x_3)} \right|^2 \end{aligned} \tag{4.14}$$

Substituting $x_2 = M_L - x_1$ *and* $x_4 = M - M_L - x_3$*, we obtain that:*

$$p(x_1, x_2, x_3, x_4) = \left| C^{J,M}_{J_L,x_1+x_2;\, J_R,x_3+x_4} \right|^2 \left| C^{J_L,x_1+x_2}_{\frac{1}{2},x_1;\, \frac{1}{2},x_2} \right|^2 \left| C^{J_R,x_3+x_4}_{\frac{1}{2},x_3;\, \frac{1}{2},x_4} \right|^2 = |\, \langle x|\Psi\rangle\, |^2. \tag{4.15}$$

The classical algorithm defined in Theorem 3 then extends without change to general spin eigenfunctions. The algorithm shows that the oracle O that was used to define the computational class PQP in Def. 12 can be simulated in polynomial time with a classical probabilistic algorithm. It follows from the definition of PQP that $\mathsf{PQP} = \mathsf{L}^O \subseteq \mathsf{BPP}$, which means that the class PQP is classical.

Claim 13. $\mathsf{PQP} \subseteq \mathsf{BPP}$.

Proof. First notice that $\mathsf{PQP} = \mathsf{L}^O \subseteq \mathsf{P}^O \subseteq \mathsf{BPP}^O$. We aim to prove that $\mathsf{BPP}^O = \mathsf{BPP}$. To that end, let L be a language decided by BPP^O, so that L is decided by a probabilistic Turing machine M that makes up to polynomially many (say n^k) calls to the oracle O and gives the correct answer with probability at least $2/3$ on any input. To see that $L \in \mathsf{BPP}$, we use the classical algorithm derived above to simulate the oracle O. For every simulated oracle call, running on a probabilistic Turing machine in $poly(n, 1/\epsilon, \log 1/\delta)$ time for some $\epsilon = 1/poly(n)$, gives wrong approximation with probability $1 - \delta$. The overall chance of all n^k simulated calls succeeding is then at least $(1 - n^k\delta)$ and the algorithm then succeeds with probability at least $2/3\,(1 - n^k\delta)$. For the sake of the argument, we can fix $\delta = n^{-k}/10$, in which case the algorithm runs in $poly(n)$ time and succeeds with probability $3/5$. This can be amplified to $2/3$ by the standard error reduction techniques (See Ref. [80], Thm 7.10). As a result, we have that $\mathsf{BPP}^O \subseteq \mathsf{BPP}$. $\mathsf{BPP} \subseteq \mathsf{BPP}^O$ clearly holds. Hence $\mathsf{PQP} \subseteq \mathsf{BPP}^O = \mathsf{BPP}$. □

As a consequence, Permutational Quantum Computing cannot provide significant quantum advantage by empirical estimation of the transition amplitudes.

4.7 Quantum Schur Sampling Circuits

The algorithm of Sec 4.4 straightforwardly extends to a wider class of circuits sandwiched between a pair of spin eigenfunctions. Suppose that U is a unitary such that for any spin eigenfunction $|\Psi\rangle$, the state $U|\Psi\rangle$ is CT. The algorithm of Sec 4.4 then remains practically unchanged for computing the overlaps of $\langle\Phi|U|\Psi\rangle$ – the only difference being that both real and complex part of the amplitude have to be estimated. Van den Nest made this observation in Ref. [6], where he calls the set of unitaries U 'efficiently computable unitary basis-preserving'. The class can be further extended to 'efficiently computable sparse' unitaries, defined in the same paper. We referred

to various classes of these computations as the Quantum Schur Sampling circuits in Refs. [30, 31]. An example that will be of some significance in the discussion of the following chapter is the circuit class where U is an efficiently computable reversible classical circuit.

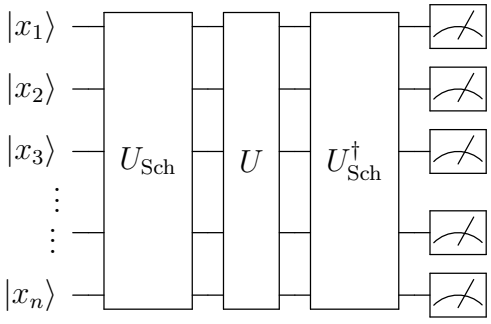


Figure 4.2: Quantum Schur Sampling circuit. The U_{Sch} gates are $SU(2)$ Quantum Schur Transforms, mapping from the computational basis to the spin eigenfunction basis. The unitary gate U preserves computational tractability. If U is a permutation, and the standard basis is used for QST, the above circuit corresponds to Permutational Quantum Computing with genealogical basis.

Chapter 5

Classical Algorithm for Sifting Probabilities

This chapter is based on a work [31] done in collaboration with S. Strelchuk and K. P. Temme. The Schwarz and van den Nest paper [8] that this section extends was pointed to me by K. Temme. The work on adaptation of the Kushilevitz-Mansour algorithm in this context was initiated by me, but S. Strelchuk helped me to work out most of the technical results such as the generalized swap test or the marginal projection lemma. Both K. Temme and him helped me to solidify the results to a publishable form.

5.1 Introduction

The previous chapter described an efficient classical algorithm for approximating of the transition amplitudes of Permutational Quantum Computing to polynomial precision. It refuted Conjecture 1 and lead to a proof that any problem in PQP can be simulated with a classical probabilistic algorithm. An intriguing question remained: Is it classically hard to find the transition amplitudes for which the approximation in Problem 1 becomes nontrivial? More specifically, is there an efficient classical algorithm for the following problem (see Fig. 5.1)?

Problem 2 (Large Elements in the Rows of Young's Orthogonal Form)**.**
Problem: *Find all matrix elements larger than* $1/poly(n)$ *in a row of Young's orthogonal matrix for some fixed polynomial* $poly(n)$*.*
Input*:* $\pi \in S_n$*, a Young diagram* λ *on two rows and an integer* $k \in [f^\lambda]$*, where* f^λ *is given by Eq. 2.25.*

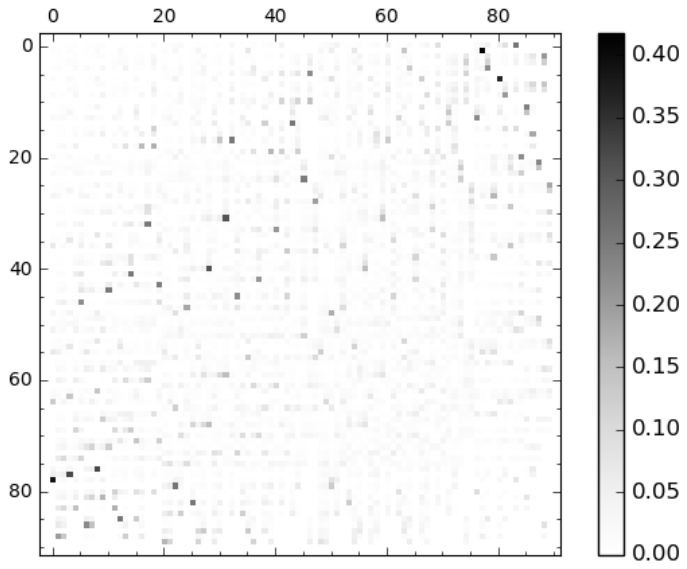


Figure 5.1: All output probabilities for a randomly chosen permutation in Permutational Quantum Computing with genealogical basis on 10 qubits and $J = 1$. Each row corresponds to an output distribution for fixed input state and is normalized. Typical matrix elements scale as $\sim 1/2^n$ and are indistinguishable from zeros by the inverse-polynomially precise algorithm. The central question that we pose here is: If there are $1/poly(n)$ large elements in any row of the output distribution pattern, is it hard to find them with a classical algorithm?

Output: *A set L of column indices of all $1/poly(n)$ large elements in the k-th row of the Young's orthogonal matrix that represents π in an irrep labeled by λ.*

Following the analysis outlined in the previous chapter, it is not hard to see that Problem 2 has an efficient quantum algorithm that uses the model in Def. 8: Fix a spin eigenfunction that corresponds to the Young's diagram λ and the k-th row of the corresponding Young's orthogonal matrix, apply the permutation π and take about $\sim poly(n)^2$ samples in the genealogical basis. The output distribution probabilities are resolved to about $\sim 1/poly(n)$ precision this way, so it is possible to output all significantly large ones (if there are any) with high probability.

In contrast, there can be exponentially many possible outputs and it seems difficult to find the sufficiently large ones by classical means. Could the quantum algorithm for Problem 2 lead to a superpolynomial quantum advantage? The main result of this chapter answers this question negatively. We describe a classical algorithm that finds all $1/poly(n)$ large output probabilities in Permutational Quantum Computing with genealogical bases in polynomial time. The work again builds on the simulation results

of van den Nest [6] and its extension by Schwarz and van den Nest [8], where the authors studied a similar problem in the context of the Quantum Fourier Transform. Their results are derived by applying a variant of the Kushilevitz-Mansour algorithm from computational learning theory [91] and the main technical contribution here is an adaptation of these techniques to Permutational Quantum Computing.

Another question that remained open from the last chapter was that of approximately sampling the output distributions in the model. In Ref. [8], the authors focused on a weak simulation of circuit classes with a QFT-Toffoli-QFT^{-1} structure, followed by a computational basis measurement on subset of qubits after the final QFT.[1] Circuits of this form are used in Shor's algorithm and the key result of Ref. [8] was to show that – under additional promise on sparsity of the output distribution – the circuits can be approximately sampled from classically. Here we show that the same holds for Permutational Quantum Computing.

5.2 Sifting Output Probabilities

We address the central question of this part by giving an algorithm that finds large probabilities in the output of Permutational Quantum Computing as defined in Def. 8. Our approach once again revolves around the concept of computational tractability from Def. 14 and extends the following result of Schwarz and van den Nest:

Theorem 4 (Schwarz and van den Nest, Theorem 5 [8])**.** *Let $C = U_2U_1$ be a unitary n-qubit circuit composed of two blocks and let $|\Psi\rangle$ be the input state. Suppose that the state $U_1|\Psi\rangle$ is CT and that U_2 is either:*

1. *QFT modulo 2^n or its inverse or,*

2. *$U_2 = u_1 \otimes u_2 \ldots \otimes u_n$, where u_i are arbitrary qubit unitaries.*

Then there exists a classical randomized algorithm with runtime $poly(n, 1/\theta, \log 1/\pi)$ that with probability at least $1 - \pi$ outputs a list $L = \{x_1, x_2 \ldots, x_l\}$ of n-bit strings such that $l \leq 2/\theta$ and each x_i is an n-bit string for which:

- *for all $x \in L$ it holds that: $|\langle x|C|\Psi\rangle|^2 \geq \theta/2$.*

[1] QFT stands for the Quantum Fourier Transform, Toffoli stands for a circuit composed by $poly(n)$ Toffoli gates.

- *every n-bit string x, such that $|\langle x|C|\Psi\rangle|^2 \geq \theta$ belongs to L.*

Specifically, by fixing $\theta = 1/poly(n)$, the algorithm lists all $1/poly(n)$ output probabilities in polynomial time. The key result here is an extension of the above theorem to the case where U_2 is the Quantum Schur Transform.

We denote the Bratelli paths on n qubits by boldface in this chapter – such a path is a list of total spin eigenvalues $\boldsymbol{J} = (J_{[2]}, J_{[3]}, \ldots J_{[n]})$ in the genealogical basis, that satisfy the angular momentum addition rules of Eq. 2.21.[2] Any genealogical spin eigenfunction can be then written as $|\boldsymbol{J}, M\rangle$. In the rest of this chapter, we for simplicity assume a fixed spin eigenfunction $|\Phi\rangle$ with a permutation U_π applied to it on the input.[3] This way, the quantum number M becomes conserved throughout the computation and we drop it where appropriate. We will also work with genealogical basis for simplicity. The output distribution is then given by:

$$p(\boldsymbol{J}) = |\langle \Phi|\boldsymbol{J}\rangle|^2. \tag{5.1}$$

The central result that we prove here is the following extension of Theorem 4:

Claim 14. *There is a classical randomized algorithm with runtime $poly(n, 1/\theta, \log 1/\pi)$ that with probability at least $1-\pi$ outputs a list $L = \{\boldsymbol{J}_1, \boldsymbol{J}_2 \ldots \boldsymbol{J}_l\}$ of n-qubit paths, such that $l \leq 2/\theta$ and:*

$$\forall\, \boldsymbol{J} \in L:\ p(\boldsymbol{J}) \geq \frac{\theta}{2},$$
$$\forall\, \boldsymbol{J}:\ p(\boldsymbol{J}) > \theta \implies \boldsymbol{J} \in L.$$

By fixing $\theta = 1/poly(n)$, the algorithm finds all polynomially large probabilities in any row of Fig 5.1 in polynomial time, which answers the central question of this part negatively.

What follows is a proof of the above claim. We first show how to approximate a set of output probability marginals that will allow us to use the techniques of Ref. [8].

For a path $\boldsymbol{j}$ on $k \leq n$ qubits, define the *output marginal* $p(\boldsymbol{j})$:

$$p(\boldsymbol{j}) := \sum_{\boldsymbol{J} \supset \boldsymbol{j}} p(\boldsymbol{J}) = \langle \Phi| \sum_{\boldsymbol{J} \supset \boldsymbol{j}; M} |\boldsymbol{J}, M\rangle \langle \boldsymbol{J}, M|\, \Phi\rangle := \langle \Phi|\, \Pi(\boldsymbol{j})\, |\Phi\rangle,$$

[2]The paths on n-qubits are bijective to Young tableaux with n boxes and we could formulate the algorithm this way as well instead.

[3]The input spin eigenfunction does not have to be genealogical. The results translate with minor modification to the case where $|\Phi\rangle$ is only a computationally tractable state.

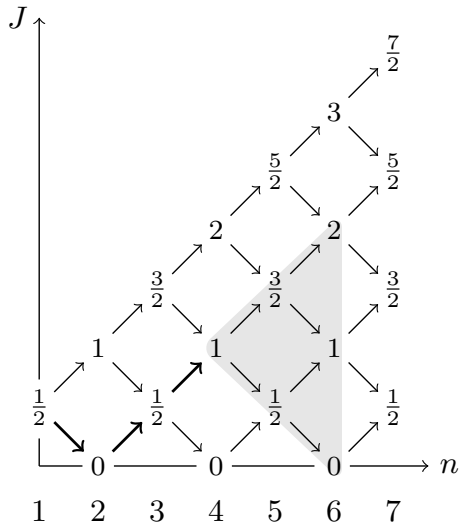


Figure 5.2: The summation $\sum_{\boldsymbol{J}\supseteq\boldsymbol{j}}$ runs over all paths $\boldsymbol{J}$ on n qubits that share a common path $\boldsymbol{j}$ for the first k qubits. As an example, in the diagram above $\boldsymbol{j} = \frac{1}{2} \to 0 \to \frac{1}{2} \to 1$ and $k = 4$, $n = 6$, where the summation runs over all Bratteli paths of length n within the shaded region.

where the summation $\sum_{\boldsymbol{J}\supseteq\boldsymbol{j}}$ sums all paths $\boldsymbol{J}$ on n qubits that share a common path $\boldsymbol{j}$ for the first k qubits (see Fig. 5.2). The summation $\sum_M$ runs from $-J$ to J in integer steps and we use $\sum_{\boldsymbol{J}\supseteq\boldsymbol{j};M}$ as a shorthand for $\sum_{\boldsymbol{J}\supseteq\boldsymbol{j}}\sum_M$.

Claim 15. *We have that:*

$$\Pi\left(\boldsymbol{j}\right) := \sum_{\boldsymbol{J}\supseteq\boldsymbol{j};M} \left|\boldsymbol{J}, M\right\rangle\left\langle \boldsymbol{J}, M\right| = \sum_{m} \left|\boldsymbol{j}, m\right\rangle\left\langle \boldsymbol{j}, m\right| \otimes \mathbf{1}_{n-k} := \sum_{m} \left|\boldsymbol{j}, m\right\rangle\left\langle \boldsymbol{j}, m\right| .$$

where $\boldsymbol{J}$ is a path on n qubits, $\boldsymbol{j}$ is a path on $k \leq n$ qubits, $\mathbf{1}_{n-k}$ is the identity on the last $n-k$ qubits. The sum $\sum_m$ runs over $m \in \{-j, -j+1, \ldots j\}$ where j is the endpoint of the path $\boldsymbol{j}$.

The proof is technical and deferred to Sec. 5.6.1. We now show that the output marginals have an efficient polynomial approximation:

Claim 16. *For a Bratteli path $\boldsymbol{j}$ on $k \leq n$ qubits, $p(\boldsymbol{j})$ can be approximated to ϵ error with probability $1-\delta$ in poly $(n, 1/\epsilon, \log(1/\delta))$ time.*

Proof. (Note: builds on the idea of using generalized SWAP test used in the proof of Corollary 2 of Ref. [6]) We first show that the marginal $p(\boldsymbol{j})$ on $k \leq n$ qubits can

be written as a transition amplitude of a larger, $(n+k)$-qubit circuit. We do this be noticing that:

$$\langle\Phi|\boldsymbol{j},m\rangle\langle\boldsymbol{j},m|\Phi\rangle = \left(\langle\boldsymbol{j},m|\langle\Phi|\right)U_{\mathsf{SWAPS}}\left(|\Phi\rangle|\boldsymbol{j},m\rangle\right),$$

where U_{SWAPS} is a permutation gate on $k+n$ qubits. Write symbolically $|\boldsymbol{j},m\rangle = |\Psi\rangle = |\Psi_1\Psi_2\ldots\Psi_k\rangle$ and $|\Phi\rangle = |\Phi_1\Phi_2\ldots\Phi_n\rangle$, where Φ_i labels the i-th qubit in the multiqubit state Φ. We have that:

$$\langle\Phi|\boldsymbol{j},m\rangle\langle\boldsymbol{j},m|\Phi\rangle = \langle\Phi|\Psi\rangle\langle\Psi|\Phi\rangle = \langle\Phi_1\ldots\Phi_n|\Psi_1\ldots\Psi_k\rangle\langle\Psi_1\ldots\Psi_k|\Phi_1\ldots\Phi_n\rangle.$$

Let U_{SWAPS} swap the $(n+i)$-th and $(n-k+i)$-th qubits for all $1\leq i\leq k$:

$$U_{\mathsf{SWAPS}}|\Psi_1\ldots\Psi_k\rangle|\Phi_1\ldots\Phi_n\rangle = |\Phi_1\ldots\Phi_k\rangle|\Psi_1\ldots\Psi_k,\Phi_{k+1}\ldots\Phi_n\rangle.$$

This gives:

$$\begin{aligned}&\langle\Phi|\langle\Psi|U_{\mathsf{SWAPS}}|\Psi\rangle|\Phi\rangle = \langle\Phi_1\ldots\Phi_n|\langle\Psi_1\ldots\Psi_k|\Phi_1\ldots\Phi_k\rangle|\Psi_1\ldots\Psi_k,\Phi_{k+1}\ldots\Phi_n\rangle\\ &= \langle\Phi_1\ldots\Phi_n|\Psi_1\ldots\Psi_k\rangle\langle\Psi_1\ldots\Psi_k|\Phi_1\ldots\Phi_n\rangle = \langle\Phi|\Psi\rangle\langle\Psi|\Phi\rangle.\end{aligned}$$

Since both $|\boldsymbol{j},m\rangle$ and $|\Phi\rangle$ states are computationally tractable and U_{SWAP} is a permutation on up to $2n$ objects, it follows that:

$$\left(\langle\boldsymbol{j},m|\langle\Phi|\right)U_{\mathsf{SWAPS}}\left(|\Phi\rangle|\boldsymbol{j},m\rangle\right).$$

With probability at least $1-\delta$, this can be approximated to $\pm\epsilon$ error in $poly(n,1/\epsilon,\log 1/\delta)$ time by Theorem 3. Therefore:

$$p(\boldsymbol{j}) = \langle\Phi|\Pi(\boldsymbol{j})|\Phi\rangle = \sum_m \langle\Phi|\boldsymbol{j},m\rangle\langle\boldsymbol{j},m|\Phi\rangle = \sum_m \left(\langle\boldsymbol{j},m|\langle\Phi|\right)U_{\mathsf{SWAPS}}\left(|\Phi\rangle|\boldsymbol{j},m\rangle\right).$$

Since $\sum_m$ sums at most $2j+1\leq n+1$ terms, it follows that $p(\boldsymbol{j})$ can be also approximated to $\pm\epsilon$ error in $poly(n,1/\epsilon,\log(1/\delta))$ time. □

We now describe a classical algorithm that finds large elements in the output distribution. It is an adaptation of the Kushilevitz-Mansour algorithm from Ref. [91].

Theorem 5 (Kushilevitz-Mansour [8, 91])**.** *Let $p(\boldsymbol{J})$ be a probability distribution on paths of length n. There is a classical algorithm that outputs a set L of length-n paths, such that for $\theta > 0$:*

$$\forall\, \boldsymbol{J} \in L:\ p(\boldsymbol{J}) \geq \frac{\theta}{2},$$
$$\forall\, \boldsymbol{J}:\ p(\boldsymbol{J}) > \theta \implies \boldsymbol{J} \in L.$$

Both hold with probability at least $1-\pi$. The algorithm runs in poly $(n, 1/\theta, \log(1/\pi))$.

Proof. (See Algorithm 2). The algorithm runs in n steps. In every step, a list L_k

Algorithm 2: Kushilevitz-Mansour Algorithm

1. Set $L_2 = \emptyset$. Choose:
$$\delta < \frac{\theta}{2n},$$
and compute $\tilde{p}(\boldsymbol{j}_2)$ for both paths on two qubits by Claim 16 , such that:
$$|p(\boldsymbol{j}_2) - \tilde{p}(\boldsymbol{j}_2)| \leq \frac{\theta}{4}.$$
Add $\boldsymbol{j}$ to L_2 if $\tilde{p}(\boldsymbol{j}_2) \geq \frac{3}{4}\theta$.
2. Continue for $k = 3, \ldots, n$. Assume L_{k-1} has been found. For any path $\boldsymbol{j}_{k-1} \in L_{k-1}$, take all possible steps in the branching diagram.
This gives k-qubits paths $\boldsymbol{j}_k$ that end at $j_k = |j_{k-1} \pm \frac{1}{2}|$.
For every such path, compute the approximation $\tilde{p}(\boldsymbol{j}_k)$ such that:
$$|\tilde{p}(\boldsymbol{j}_k) - p(\boldsymbol{j}_k)| \leq \frac{\theta}{4}.$$
If $\tilde{p}(\boldsymbol{j}_k) \geq \frac{3}{4}\theta$, add the path $\boldsymbol{j}_k$ to L_k.
3. In every step of the computation, check if $|L_k| > \frac{2}{\theta}$.
If true, halt and output $\emptyset$.
(The algorithm never halts if all approximation steps succeed.)
4. Output $L = L_n$.

of marginal path approximations is constructed using Claim 16. Each such approximation is computed to error $\theta/4$ and confidence $1 - \delta$ in time $poly(n, 1/\theta, \log 1/\delta)$.

The probability that every one of these approximations succeeds is $(1-\delta)^{|L_k|}$. Since $|L_k| \leq 2/\theta$, this is at least:

$$(1-\delta)^{|L_k|} \geq (1-\delta)^{2n/\theta} \geq 1 - \frac{2\delta n}{\theta} := 1 - \pi.$$

The algorithm therefore terminates in $poly\,(n,\, 1/\theta,\, \log 1/\pi)$ time as it halts whenever the number of elements in any list exceeds $2/\theta$. Since $p(\boldsymbol{j}_k) \geq \theta/2$ for each $\boldsymbol{j}_k \in L_k$, the final list L contains at most $2/\theta$ elements by normalization. So if all approximation steps succeed, the algorithm does not halt before it outputs L. *(We remark that this analysis follows closely the one in Ref. [8], which in turn follows the analysis of Ref. [91])* □

Algorithm 2 has an interesting consequence: as it runs in polynomial time whenever $\theta = 1/poly(n)$, it finds all paths with polynomially large $p(\boldsymbol{J})$ in polynomial time. This gives a negative answer to the central question in this part. It does not seem difficult to extend this result to general spin eigenfunctions, although we do not do this here as the algorithm is already quite complex.

5.3 Approximate Sampling

Following the work of Schwarz and van den Nest in Ref. [8], we now use the above algorithm to approximately sample the quantum Schur circuits under additional sparsity constraint on their output distribution:

Definition 15 (ϵ-approximate t-sparsity [8])**.** *A probability distribution $p(\boldsymbol{J})$ is t-sparse if it has at most t nonzero elements $p(\boldsymbol{J})$. A probability distribution $\tilde{p}(\boldsymbol{J})$ is ϵ-approximately t-sparse if there exists a t-sparse distribution $p(\boldsymbol{J})$ such that:*

$$\sum_{\boldsymbol{J}} |p(\boldsymbol{J}) - \tilde{p}(\boldsymbol{J})| \leq \epsilon, \tag{5.2}$$

where the sum runs over all n-qubit paths.

We adapt a technical lemma of Kushilevitz and Mansour in Ref. [91] to spin eigenfunctions:

Claim 17. *Let $p(\boldsymbol{J})$ be an ϵ-approximate t-sparse distribution. Let S be the set of all $\boldsymbol{J}$ for which $p(\boldsymbol{J})$ is greater than ϵ/t. Then:*

$$\sum_{\boldsymbol{J}\notin S} p(\boldsymbol{J}) \leq 2\epsilon.$$

The proof is deferred to Sec 5.6.2. From now on, assume that $p(\boldsymbol{J})$ is ϵ-approximately t-sparse. The aim is now to construct an approximating distribution that can be efficiently sampled from, which for most part follows from Ref. [8]. Let L be the set of n-qubit paths generated by the Kushilevitz-Mansour algorithm with threshold $\theta = \epsilon/t$. Choose:

$$\epsilon' = \min\left(\frac{\epsilon}{|L|}, \frac{\epsilon}{4t}\right), \tag{5.3}$$

and compute ϵ' approximations $\tilde{p}(\boldsymbol{J})$ for all $\boldsymbol{J} \in L$ by Theorem 3. Let N be the number of all n-qubit paths $N = \sum f^{[\mu,\nu]}$, where $f^{[\mu,\nu]}$ is the number of standard tableaux with μ and ν boxes in the first and second row and the sum runs over all $\mu \geq \nu$ such that $\mu+\nu = n$. This sum has less than n terms and can be computed fast with the Frame-Robinson-Thrall formula of Eq. 2.25. Define a normalization factor $\alpha = 1/(N - |L|)$ such that $\sum_{\boldsymbol{J}\notin L} \alpha = 1$. Use the ϵ'-approximations $\tilde{p}(\boldsymbol{J})$ to define:

$$\tilde{q}(\boldsymbol{J}) = \begin{cases} \tilde{p}(\boldsymbol{J}) \text{ for } \boldsymbol{J} \in L, \\ \tilde{p}_\circ := \alpha(1 - \sum_{\boldsymbol{J}\in L} \tilde{p}(\boldsymbol{J})) \text{ otherwise,} \end{cases}, \tag{5.4}$$

so that $\tilde{q}$ becomes uniform on all $\boldsymbol{J}$ outside L and the constant $\tilde{p}_\circ$ is chosen so that it is normalized. This way, we can prove that:

Claim 18.

$$\|\tilde{q} - p\|_1 \leq 6\epsilon.$$

The proof is deferred to Sec 5.6.3. We now show how to sample $\tilde{q}$.

Claim 19. *Assume that $p(\boldsymbol{J})$ is ϵ-approximate t-sparse. It can be sampled classically in poly$(n, 1/\epsilon, t)$ time to 6ϵ error in the total variational distance.*

Proof. Use the Kushilevitz-Mansour algorithm in Theorem 5 with threshold $\theta = \epsilon/t$ to find the set L. Compute $b = \sum_{\boldsymbol{J}\in L} \tilde{p}(\boldsymbol{J})$ and flip a coin with a bias b.

- With probability b, output a sample from $\tilde{q}(\boldsymbol{J})$ for $\boldsymbol{J} \in L$.

- With probability $1 - b$, output $\boldsymbol{J} \notin L$ uniformly randomly.

It is easy to see that this random process samples the probability distribution $\tilde{q}$ defined above – the bias b is the chance of sampling a path $\boldsymbol{J} \in L$ and the the distribution is uniform on all paths outside of L.

To sample $\boldsymbol{J} \notin L$ uniformly randomly, sample $[\mu, \nu]$ according to $p([\mu, \nu]) \sim f^{[\mu,\nu]}/N$. With the shape of the Young diagram fixed, generate uniformly randomly a standard Young tableau using the algorithm of Greene, Nijenhuis and Wilf [92] (See Sec. 5.6.4 for its description). The probability of choosing a particular tableau is given by $1/N$. By the correspondence of Chapter 2, the Young tableau gives a path $\boldsymbol{J}$. Resample if $\boldsymbol{J} \in L$. (*Resampling is exponentially unlikely as we expect that* $|L| \sim poly(n)$ *and* $N \sim exp(n)$) □

5.4 Remark about Output Sparsity

Could the above sampling technique work for Permutational Quantum Computing? To get at least some intuition about this, we devise two proxy criteria to gain some intuition about the form of the output distribution. The output distributions are all of size d, where $d < 2^n$. As the first proxy criterion, we construct an at most n-sparse distribution by taking the first $\log(d) < n$ largest elements, setting all other elements to zero and renormalizing. We then compute the total variational distance between the renormalized sparse distribution and the original one. If the total variational distance becomes smaller than n^{-1}, the new distribution is n^{-1}-approximately n-sparse – we say that the distribution is consistent with the first proxy. The second proxy we consider is to construct an at most n^2-sparse distribution in the same way. If the total variational distance becomes smaller than n^{-1}, the distribution is n^{-1}-approximately n^2-sparse. We say that the distribution is consistent with the second proxy. The proxies are applied to a selected set of irreps on up to 9 qubits and a set of 50 uniformly randomly chosen permutations. For every irrep, we plot the flattest distribution among all of them – which is the distribution that is the closest to uniform – and verify the consistency with the sparsity hypothesis. See Fig. 5.3.

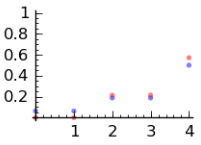
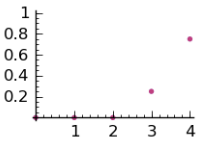
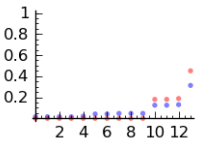
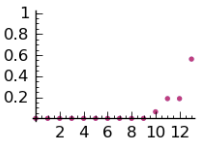
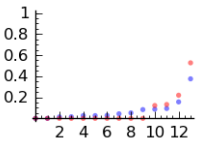
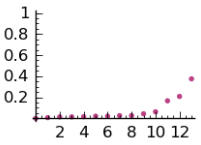
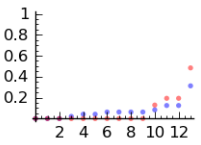
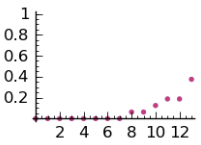
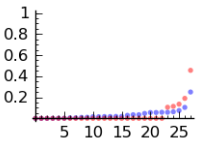
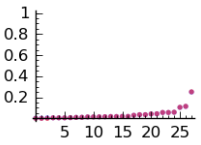
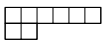
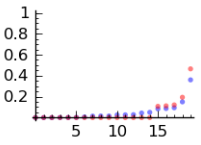
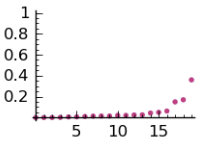
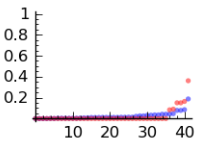
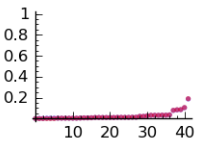
Irrep.	Proxy 1			Proxy 2		
	$\epsilon = 0.13$		Yes	$\epsilon = 0$		Trivial
	$\epsilon = 0.31$		No	$\epsilon = 0$		Yes
	$\epsilon = 0.29$		No	$\epsilon = 0$		Yes
	$\epsilon = 0.35$		No	$\epsilon = 0$		Yes
	$\epsilon = 0.45$		No	$\epsilon = 0.01$		Yes
	$\epsilon = 0.22$		No	$\epsilon = 0$		Yes
	$\epsilon = 0.48$		No	$\epsilon = 0.01$		Yes

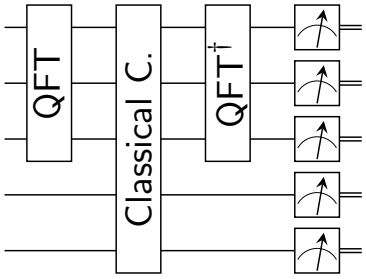
Figure 5.3: The two proxies for output sparsity of Permutational Quantum Computing. For every irrep, we generated 50 uniformly random permutations, computed all output probability distributions selected the one that was closest to being sparse (plotted in blue). We then prepared the approximately sparse distribution depending on the proxy (plotted in red) and computed the total variational distance ϵ between the two distributions. For each pair of distributions, we indicate if they are consistent with the proxy or not by (Yes/No) – we have that almost no distribution is consistent with the first proxy condition, while all (except for the first, where the condition becomes trivial) pairs satisfy the second proxy condition.

The second proxy seems to indicate that the typical output distribution in Permutational Quantum Computing could be sufficiently sparse and may allow applying the above sampling algorithm in some form. Running the same analysis for larger number of qubits becomes prohibitively slow as the Young orthogonal matrices have to be

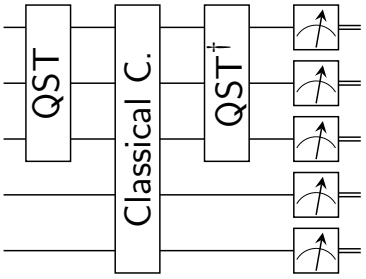
computed in full. The validity of the second proxy condition for most of the studied distributions should be therefore only taken as a mere intuition that approximately sparse distributions could be common in Permutational Quantum Computing.

5.5 Additional Remarks

It may seem that Permutational Quantum Computing (or its extension to Quantum Schur Sampling) is unlikely to offer a super-polynomial quantum computational advantage. There is however interesting structural similarity between quantum circuits used in Shor's algorithm and quantum Schur sampling – and one could wonder if there could be an analog of Shor's algorithm for the quantum Schur sampling circuits (see Fig. 5.4.)



(a) Quantum circuit used in Shor's factoring algorithm [8]. The classical circuits between the transforms represent a polynomially-long sequence of Toffoli gates.



(b) Quantum Schur Sampling circuit. QST denotes the $SU(2)$ Quantum Schur Transformation.

Figure 5.4

In both cases, the quantum state immediately before the second quantum Fourier or Schur transform is computationally tractable, which means that both circuit classes can be approximately sampled from if their outputs are sufficiently sparse. One may then question what precludes efficient simulation of Shor's algorithm with this technique. Schwarz and van den Nest observed in Ref. [8] that Shor's algorithm typically does not satisfy the sparsity condition as the output distribution of the algorithm is conjectured to have $\Omega(2^n/n)$ significantly large probabilities in its output. While this means that the output probabilities cannot be all resolved even in quantum computation, Shor's algorithm uses clever post-processing that requires finding only $poly(n)$-

many nontrivial probabilities on average, thus possibly leading to a quantum computational advantage. The post-processing step is, loosely, based on reconstructing group generators by sampling $\log|G|$-many group elements for a super-polynomially large $|G|$. Inspired by this analogy, it becomes an intriguing possibility if there exists an analogous post-processing technique that could lead to quantum advantage for quantum Schur sampling circuits. We remark that aside from the work on Permutational Quantum Computing, the question of using the Schur basis in quantum algorithms seems relatively unexplored, perhaps with the exception of work in Ref. [78] that studied limitations of the quantum 'Fourier-Schur' sampling.

Problem 2 concerned with finding large matrix elements for a fixed row of the Young's orthogonal form matrix. It remained open if it is possible to find some of the $1/poly(n)$ large matrix elements without fixing the input. A quantum algorithm for this task that goes somewhat beyond the Permutational Quantum Computing model could be as follows – prepare: $|\omega\rangle = (1/2^n)(U_\pi \otimes \mathbf{1}) \sum_x |x\rangle_A |x\rangle_B$ and measure the n-qubit registers A and B in the genealogical basis. The probability of getting an outcome $\boldsymbol{J}, \boldsymbol{K}$ is then given by $p(\boldsymbol{J}, \boldsymbol{K}) \propto |\langle \boldsymbol{J}|U_\pi|\boldsymbol{K}\rangle|^2$, unfortunately with an exponentially large normalization factor. The adaptation of the Kushilevitz-Mansour algorithm derived here cannot be applied in this case for this reason[4] and Claim 14 does not extend to this context. It would be interesting to find out if the above quantum procedure has a high enough probability of resolving large matrix elements, as the corresponding simulation algorithm may need to use different techniques than those discussed here.

Kushilevitz-Mansour algorithm was originally conceived in the context of boolean function analysis and applied almost exclusively to the Fourier transform. Our extension of Schwarz and van den Nest's work is an application of this technique to quantum Schur transform. It would be interesting to explore if the Schur basis expansion has any classical analog that could be of use for the analysis of boolean functions.

[4]This is because Theorem 3 does not work with subnormalized states, which means that the sequence of marginals in the telescoping product cannot be efficiently approximated.

5.6 Deferred Material

5.6.1 Simplified Form of the Projector

Claim 15. *The projector:*

$$\Pi(\boldsymbol{j}) := \sum_{\boldsymbol{J}\supseteq\boldsymbol{j};M} |\boldsymbol{J},M\rangle\langle\boldsymbol{J},M|,$$

can be simplified to:

$$\Pi(\boldsymbol{j}) = \sum_{m} |\boldsymbol{j},m\rangle\langle\boldsymbol{j},m|,$$

where the sum $\sum_m$ *runs over* $m \in \{-j, -j+1, \dots j\}$.

Proof. To do so, we repeatedly use the Clebsch-Gordan orthogonality:

$$\sum_{JM} C^{JM}_{jm,j'm_2} C^{JM}_{jm',j'm_2'} = \delta_{m_2,m_2'}\delta_{mm'},$$

This becomes straightforward by noting that:

$$\langle J,M|jm,j'm_2\rangle = \langle jm,j'm_2|J,M\rangle = C^{JM}_{jm,j'm_2}. \tag{5.5}$$

One of the j-values is always $1/2$ in the sequential basis. We can drop it from the subscript of the coefficient and write:

$$\sum_{JM} C^{JM}_{jm,m_2} C^{JM}_{jm',m_2'} = \delta_{m_2,m_2'}\delta_{mm'}.$$

We have for the projector $\Pi(\boldsymbol{j})$ that:

$$\begin{aligned}\Pi(\boldsymbol{j}) &= \sum_{\boldsymbol{J}\supseteq\boldsymbol{j};M} |\boldsymbol{J},M\rangle\langle\boldsymbol{J},M| \\ &= \sum C^{J,M}_{J_{n-1}M_{n-1};m_n} |m_n\rangle |\boldsymbol{J}_{n-1},M_{n-1}\rangle \left\langle \boldsymbol{J}_{n-1},M'_{n-1}\right| \langle m_n'| C^{J,M}_{J_{n-1},M'_{n-1};m_n'},\end{aligned}$$

where in the second line $\sum = \sum_{\boldsymbol{J}_{n-1}\supseteq\boldsymbol{j}} \sum_{J,M} \sum_{m_n,m_n'} \sum_{M_{n-1},M'_{n-1}}$. The $\sum_{\boldsymbol{J}_{n-1}\supseteq\boldsymbol{j}}$ runs over all $n-1$ qubit paths $\boldsymbol{J}_{n-1}$ that contain $\boldsymbol{j}$. The sum $\sum_J$ runs over all allowed J. Clebsch-Gordan coefficients are only non-zero for $J = \left|J_{n-1} \pm \frac{1}{2}\right|$. Using

the Clebsch-Gordan orthogonality, this simplifies to:

$$\begin{aligned}\Pi(\boldsymbol{j}) &= \sum_{\boldsymbol{J}_{n-1}\supseteq\boldsymbol{j}} \sum_{m_n,m'_n} \sum_{M_{n-1},M'_{n-1}} \delta_{m_n,m'_n}\delta_{M_{n-1},M'_{n-1}} |m_n\rangle |\boldsymbol{J}_{n-1},M_{n-1}\rangle \left\langle \boldsymbol{J}_{n-1},M'_{n-1}\right| \langle m'_n| \\ &= \sum_{\boldsymbol{J}_{n-1}\supseteq\boldsymbol{j}} \sum_{M_{n-1}} \underbrace{\sum_{m_n} |m_n\rangle \langle m_n|}_{\mathbf{1}_{2\times 2}} \otimes |\boldsymbol{J}_{n-1},M_{n-1}\rangle \langle \boldsymbol{J}_{n-1},M_{n-1}| \\ &= \left(\sum_{M_{n-1}} \sum_{\boldsymbol{J}_{n-1}\supseteq\boldsymbol{j}} |\boldsymbol{J}_{n-1},M_{n-1}\rangle \langle \boldsymbol{J}_{n-1},M_{n-1}|\right) \otimes \mathbf{1}.\end{aligned}$$

This has the same form as the initial expression, but the projector is now defined by summing over paths with $n-1$ steps. It is possible to continue recursively and write:

$$\Pi(\boldsymbol{j}) = \sum_{M_{n-i}} \sum_{\boldsymbol{J}_{n-i}\supseteq\boldsymbol{j}} |\boldsymbol{J}_{n-i},M_{n-i}\rangle \langle \boldsymbol{J}_{n-i},M_{n-i}|,$$

for any integer $0 \leq i \leq n-k$. For $i = n-k$, one obtains that:

$$\Pi(\boldsymbol{j}) = \sum_{M_k} \sum_{\boldsymbol{J}_k\supseteq\boldsymbol{j}} |\boldsymbol{J}_k,M_k\rangle \langle \boldsymbol{J}_k,M_k|.$$

Since $\boldsymbol{j} \in \mathcal{A}_k$, there is only one path contributing to $\sum_{\boldsymbol{J}_k\supseteq\boldsymbol{j}}$, the path $\boldsymbol{j}$ itself. It follows that:

$$\Pi(\boldsymbol{j}) = \sum_{M_k} |\boldsymbol{j},M_k\rangle \langle \boldsymbol{j},M_k|.$$

We can write:

$$\Pi(\boldsymbol{j}) = \sum_{m} |\boldsymbol{j},m\rangle \langle \boldsymbol{j},m|.$$

where the final summation runs over $m \in \{-j,-j+1,\ldots j\}$.

□

5.6.2 Support Lemma

Claim 17. *Let $p(\boldsymbol{J})$ be an ϵ-approximate t-sparse distribution. Let S be the set of all $(\boldsymbol{J})$ for which $p(\boldsymbol{J})$ is greater than ϵ/t. Then:*

$$\sum_{\boldsymbol{J}\notin S} p(\boldsymbol{J}) \leq 2\epsilon.$$

Proof. Let $p^t(\boldsymbol{J})$ be a t-sparse distribution that is ϵ-close to $p(\boldsymbol{J})$. Define T to be the set of all $(\boldsymbol{J})$ for which p^t is nonzero, i.e. the support of p^t. Trivially, $S \cap T \subseteq S$, which implies that:

$$\sum_{\boldsymbol{J} \notin S} p(\boldsymbol{J}) \leq \sum_{\boldsymbol{J} \notin S \cap T} p(\boldsymbol{J}).$$

Define the indicator $I_A : A \to \{0, 1\}$ on the set A as follows:

$$I_A(a) := \begin{cases} 1, \text{ if } a \in A, \\ 0, \text{ otherwise,} \end{cases}$$

so that:

$$\sum_{\boldsymbol{J} \notin S \cap T} p(\boldsymbol{J}) = \sum_{\boldsymbol{J}} p(\boldsymbol{J}) \left(I_{S \cap T}(\boldsymbol{J}) - 1\right) = \sum_{\boldsymbol{J}} |p(\boldsymbol{J}) I_{S \cap T}(\boldsymbol{J}) - p(\boldsymbol{J})| := \|pI_{S \cap T} - p\|_1$$

By the triangle inequality:

$$\|pI_{S \cap T} - p\|_1 \leq \|p - pI_T\|_1 + \|pI_{S \cap T} - pI_T\|_1.$$

Since p is ϵ-approximate t-sparse, it follows that:

$$\|p - pI_T\|_1 \leq \|p - p^t\|_1 \leq \epsilon.$$

We also have that:

$$\|pI_{S \cap T} - pI_T\|_1 = \sum_{\boldsymbol{J} \in T} p(\boldsymbol{J}) I_{S \cap T}(\boldsymbol{J}) = \sum_{\boldsymbol{J} \in T/S} p(\boldsymbol{J}) \leq \frac{\epsilon}{t} t = \epsilon. \tag{5.6}$$

because all elements in T/S are $\leq \epsilon/t$ and there is at most t of them. This gives:

$$\sum_{\boldsymbol{J} \notin S} p(\boldsymbol{J}) \leq 2\epsilon.$$

□

5.6.3 Approximate Distribution

Claim 18. *For the approximate distribution $\tilde{q}$ of Eq. 5.4, it holds that:*

$$\|\tilde{q} - p\|_1 \leq 6\epsilon.$$

Proof.

$$\|\tilde{q} - p\|_1 = \sum_{\boldsymbol{J} \in L} |\tilde{q}(\boldsymbol{J}) - p(\boldsymbol{J})| + \sum_{\boldsymbol{J} \notin L} |\tilde{q}(\boldsymbol{J}) - p(\boldsymbol{J})| \leq \epsilon + \sum_{\boldsymbol{j} \notin L} |\tilde{q}(\boldsymbol{J}) - p(\boldsymbol{J})|,$$

since from Eqs. 5.4 and 5.3:

$$\sum_{\boldsymbol{J} \in L} |\tilde{p}(\boldsymbol{J}) - p(\boldsymbol{J})| \leq \sum_{\boldsymbol{J} \in L} \epsilon' \leq |L|\epsilon' \leq \epsilon,$$

where ϵ' is the approximation from 5.3 to which was $\tilde{p}(\boldsymbol{J})$ was computed for all $\boldsymbol{J} \in L$. Define also:

$$p_\circ = \alpha \left(1 - \sum_{\boldsymbol{J} \in L} p(\boldsymbol{J})\right),$$

and notice that:

$$\sum_{\boldsymbol{J} \notin L} |p_\circ - \tilde{p}_\circ| \leq \left| \sum_{\boldsymbol{J} \in L} p(\boldsymbol{J}) - \sum_{\boldsymbol{J} \in L} \tilde{p}(\boldsymbol{J}) \right| \leq \sum_{\boldsymbol{J} \in L} |p(\boldsymbol{J}) - \tilde{p}(\boldsymbol{J})| \leq \epsilon.$$

By triangle inequality:

$$\begin{aligned} \sum_{\boldsymbol{J} \notin L} |\tilde{q}(\boldsymbol{J}) - p(\boldsymbol{J})| &= \sum_{\boldsymbol{J} \notin L} |\tilde{p}_\circ - p(\boldsymbol{J})| \leq \sum_{\boldsymbol{J} \notin L} |p_\circ - \tilde{p}_\circ| + \sum_{\boldsymbol{J} \notin L} |p_\circ - p(\boldsymbol{J})| \\ &\leq \epsilon + \sum_{\boldsymbol{J} \notin L} |p_\circ - p(\boldsymbol{J})|. \end{aligned}$$

We now use the set S from from Claim 17. Since $S \subseteq L$ by the defining property of L, it follows that:

$$\sum_{\boldsymbol{J} \notin L} p(\boldsymbol{J}) \leq \sum_{\boldsymbol{J} \notin S} p(\boldsymbol{J}) \leq 2\epsilon.$$

Notice that:

$$\sum_{\boldsymbol{J} \notin L} p_\circ = \sum_{\boldsymbol{J} \notin L} \left(\alpha \sum_{\boldsymbol{J}' \notin L} p(\boldsymbol{J}') \right) = \sum_{\boldsymbol{J}' \notin L} p(\boldsymbol{J}').$$

This gives:

$$\sum_{\boldsymbol{J} \notin L} |p_\circ - p(\boldsymbol{J})| \leq \sum_{\boldsymbol{J} \notin L} p_\circ + \sum_{\boldsymbol{J} \notin L} p(\boldsymbol{J}) = 2 \sum_{\boldsymbol{J} \notin L} p(\boldsymbol{J}) \leq 4\epsilon,$$

which leads to:

$$\sum_{\boldsymbol{J} \notin L} |\tilde{q}(\boldsymbol{J}) - p(\boldsymbol{J})| \leq 5\epsilon,$$

and therefore:

$$\|\tilde{q} - p\|_1 \leq 6\epsilon.$$

□

5.6.4 Greene, Nijenhuijs, Wilff algorithm

Here we describe the algorithm of [92] and its correctness proof. It is also a proof of the Frame-Robinson-Thrall formula that was used in Eq. 2.25 [72]. See Algorithm 4.

Algorithm 3: Greene, Nijenhuis, Wilff [92]

Input: Standard Young diagram $\lambda \vdash n$.

Algorithm:

1. Set $\lambda' \leftarrow \lambda$ and $k \leftarrow n$.
2. Choose a uniformly random box (a, b) of λ'.
3. Then:
   - If (a, b) is at the end of a row a of λ' and there is no box below it, remove the box (a, b) from λ'. Insert k to the (a, b) box of λ and set $k \leftarrow k - 1$. If $k = 0$, terminate the algorithm and output λ; otherwise go to step 2.
   - If (a, b) is not at the end of a row a of λ, choose an element (a', b') uniformly randomly from the hook of (a, b) (excluding (a, b)) and set $(a, b) \leftarrow (a', b')$. Repeat step 3.

Output: A uniformly random Young tableau of shape λ.

A crude bound on the algorithm running time is $O(n^3)$. A proof of correctness of the algorithm is given in Appendix A.3.

Conclusion

Isolating computational advantages that could be achieved by quantum algorithms is a central motivation of quantum computation theory. Here we studied a conjectured computation speedup for the problem of approximating the matrix elements of S_n irreducible representations in Young's orthonormal basis (Problem 1). The quantum algorithm for this task is naturally implemented in a restricted model of quantum computation – Permutational Quantum Computing – which can be seen as an attempt to capture the power of computation with spin networks, and some of our results translate to that context as well.

Our main result disproved a conjecture about superpolynomial speedup achieved by the quantum algorithm for this problem. To that end, we derived a classical simulation algorithm for approximation of transition amplitudes in Permutational Quantum Computing – a problem that was thought to be difficult to solve by classical means. The classical algorithm was developed from the following sequence of observations:

- Spin Eigenfunctions are related to the computational basis by the $SU(2)$ Quantum Schur Transform, which can be expressed as a cascade of $SU(2)$ Clebsch-Gordan transformations. The main motivation behind this insight was an attempt to translate Permutational Quantum Computing to quantum circuit model, which is something that was not done in Ref. [3].

- That $SU(2)$ Clebsch-Gordan coefficients can be computed to machine precision in polynomial time using the Racah formula (Eq. 2.36).

- The overlaps of a spin eigenfunction with any computational basis state can be computed efficiently to machine precision.

- The distribution $p(x) = |\langle x|\psi\rangle|^2$, given by the overlap between computational basis and spin eigenfunctions, can be efficiently approximately sampled from by classical means using a telescoping product expansion.

- The above can be used, in combination with van den Nest's results on computational tractability, to approximate transition amplitudes in the model to polynomial precision.

We subsequently extended these results to show that the computational complexity class PQP, defined with the aim to capture all computational tasks solvable by polynomial approximation of the transition amplitudes of Permutational Quantum Computing, is in fact classical. Motivated by this result, we also investigated other sources of quantum advantage that may be obtained in the model. Specifically, we studied if the computational model could offer superpolynomial speedups by finding all nontrivial instances of Problem 1 in its output for a fixed input state, but ruled this possibility out.

Lastly, while there is a classical algorithm that approximates the transition amplitudes of the model to polynomial precision in polynomial time, there seems to be no straightforward way of extending our results to efficient sampling of its output distributions. The closest we got to this was the adaptation of the simulation algorithm of Schwarz and van den Nest [8], that can sample the output under the additional sparsity promise. We gave some indication that the typical instances of the computational model satisfy this sparsity promise by simulating the model numerically, but did not manage to extend this analysis beyond 9 qubits, which makes this analysis rather weak. It would be thus interesting to attempt to prove the sparsity condition. If it was indeed the case that the output distributions are typically close to sparse (concentrated on small fraction of outcomes), it could rule out the possibility of a quantum supremacy argument along the lines of Ref. [1].

Communication Complexity and Antidistinguishability

Introduction

Communication complexity is a basic computational complexity model that addresses the following problem [93, 94, 95]:

"If two or more people want to compute something about the information that they possess, how long does the conversation need to be?" — [95]

This part of the thesis studies the difference in the conversation length between quantum and classical protocols in a two-party communication problem that we derive from the concept of quantum state antidistinguishability. This setup allows for a particularly simple proof of a large separation between quantum and classical one-way exact communication complexity that works in all dimensions $d \geq 3$.

Quantum state antidistinguishability is a property of a finite set of pure quantum states that guarantees existence of a (POVM) measurement that never outputs an outcome 'k' on the k-th quantum state chosen from the set. Here we use a lemma of Caves, Fuchs and Schack from Ref. [96] to show that any triple of states, such that no pair of distinct states has modulus-squared overlap greater than $1/4$, is antidistinguishable. A set S of d-dimensional pure quantum states with upper-bounded modulus-squared overlap δ^2 is called a spherical code. It follows that any triple of states drawn from a spherical code with $\delta \leq 1/2$ is antidistinguishable.

We use this to study a communication task in which Alice obtains a state drawn from the set S and Bob gets a measurement that antidistinguishes some three states contained in S. His aim is to produce outcomes of this measurement on Alice's state. The trivial quantum protocol in which Alice simply sends her quantum state to Bob uses $\lceil \log(d) \rceil$ qubits, where d is the dimension of the state that we tie to the input size $|S|$. Using an argument that lies at the heart of the PBR theorem – a no-go theorem that rules out a class of ψ-epistemic ontological models of quantum theory

[96, 97] – we derive a lower bound of $\lceil \log |S| \rceil - 1$ on the number of bits communicated in any zero-error classical one-way randomized communication protocol. This leads to complexity separations whenever $2d < |S|$. We are thus interested in finding the largest spherical code in a fixed dimension d.

Using connections between spherical codes, symmetric, informationally complete sets of states (SIC) and mutually unbiased bases (MUBs), we first derive nontrivial communication complexity separations in $d = 3$ and $d = 4$ and describe a method for deriving separations in higher dimensions using numerical computation. We then show a $0.451(d-1)-1$ lower bound that holds for general d. This gives an exponential separation between one-way quantum and zero-error randomized classical communication complexities. We ask if the separation also holds for two-way communication complexity and prove that any significant communication complexity separation for this task disappears if interaction is allowed by giving an explicit deterministic communication protocol. A significant asymptotic separation also vanishes if we allow for bounded error communication, which we prove by a reduction to equality testing. It remains an open problem if it is possible to modify our construction so that the separation becomes robust to noise – we at least show that the separation holds for multiplicative-error sampling.

We conclude by studying quantum state antidistinguishability – the enabling concept behind our results. We give a conjecture about antidistinguishability of $d > 3$ states that, if proven, gives a lower bound of $\Omega(d \log(d))$ for the zero-error randomized communication of the problem. As it is known that the best upper bound for bounded error communication goes as $O(d \log(1/\epsilon))$ it would imply a separation between exact and bounded error communication of simulating a transmission of a qudit, which was previously unknown as far as we are aware. We support the conjecture numerically.

The work was done under the supervision of Jon Barrett, who, together with Matt Pusey, suggested the idea of porting a classical simulation lower bound result of Karanjai, Bartlett and Wallman [98] when I reviewed it at a group seminar. I also benefited from discussions with John-Mark Allen who previously worked on a similar problem. This part is structured as follows:

- Chapter 6 – Introduces communication complexity.
- Chapter 7 – Reviews quantum state discrimination and antidistinguishability.
- Chapter 8 – Presents the main result and concludes by discussing the antidistinguishability conjecture.

Chapter 6

Communication Complexity

6.1 Introduction

This chapter reviews the basic concepts of communication complexity needed to present the main result. It is structured as follows: We first define the two-party communication complexity model, deterministic communication complexity and outline a basic lower-bound technique. We define randomized protocols and discuss the communication complexity of relations. We introduce quantum communication complexity and review several key results in the field. Lastly, we briefly comment on the connections between communication complexity and the ontological models framework of quantum theory.

6.2 Deterministic Communication Complexity

Communication complexity is a model of distributed computing introduced by Yao [94]. In this model, a system aims to compute a function, partial function or a relation, that depends on information distributed among its parts [93, 94, 95, 99]. These parts, usually referred to as 'players', need to communicate with one another to perform such computation. The only computational resource accounted for is the amount of this communication [100].

Possibly the simplest model of communication complexity is the following: Two players, Alice and Bob, aim to compute a function $f : X \times Y \to Z$ for finite sets X, Y, Z. Alice is given an input $x \in X$ and Bob receives $y \in Y$. They do not see each other's input and communicate some information about their inputs to one another. While doing so, they follow a fixed protocol agreed upon beforehand. This gives:

Definition 16 (Yao's model [94, 99])**.** *Let X, Y, Z be finite sets and let $f : X \times Y \to Z$ be a function. Consider a game played by Alice and Bob: For $(x, y) \in X \times Y$, give x to Alice and y to Bob. They aim to compute $f(x, y)$ according to the following rules: In every step of the game, one of the players sends a single bit to the other. The players can fix a strategy for the game beforehand, that describes in every step:*

1. *A way to determine who speaks based on the sequence of bits transmitted up to that step.*

2. *A way for the communicating player to send one bit of information based only on their input and the all messages already transmitted.*

A protocol is a strategy such that one of the players knows the value of $f(x, y)$ after the last step.

Communication protocols can be thought of as binary trees with every internal node v labelled by a boolean function. We call these functions $a_v : X \to \{0, 1\}$ for steps during which Alice speaks or $b_v : Y \to \{0, 1\}$ for steps when Bob speaks. Every leaf of the tree is labelled by an element of $z \in Z$. A strategy is an assignment of exactly one of these functions to every internal vertex of the tree. The players execute this strategy by traversing this tree from the root to a leaf, so that at every internal node v labeled by a_v, Alice transitions to the left child if $a_v(x) = 0$ and to the right child if $a_v(x) = 1$. Similarly, for nodes labeled by b_v, Bob transitions to the left child if $b_v(y) = 0$ and to the right child if $b_v(y) = 1$. The output of a strategy on the input (x, y) is the value of the leaf reached starting from the root and traversing the tree. The *communication cost* is the height of this tree [93].

Example 4 (One-way protocol for equality)**.** *Consider a communication protocol for the equality function f, defined as:*

$$f : X \times Y \to \{0, 1\}, \qquad f(x, y) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{if } x \neq y \end{cases} \tag{6.1}$$

on $X = Y = \{00, 01, 10, 11\}$ and described in Fig 6.1. Upon getting the inputs (x, y), Alice and Bob follow the protocol drawn in the diagram. Only Alice speaks in this protocol: she sends her input to Bob in two steps after which Bob announces the result. The communication cost is 2 bits. If both parties are required to know the

result, Bob sends his answer back to Alice which changes the protocol cost by one bit. Observe that every bit that Alice sends partitions the set of inputs $X \times Y$. The first

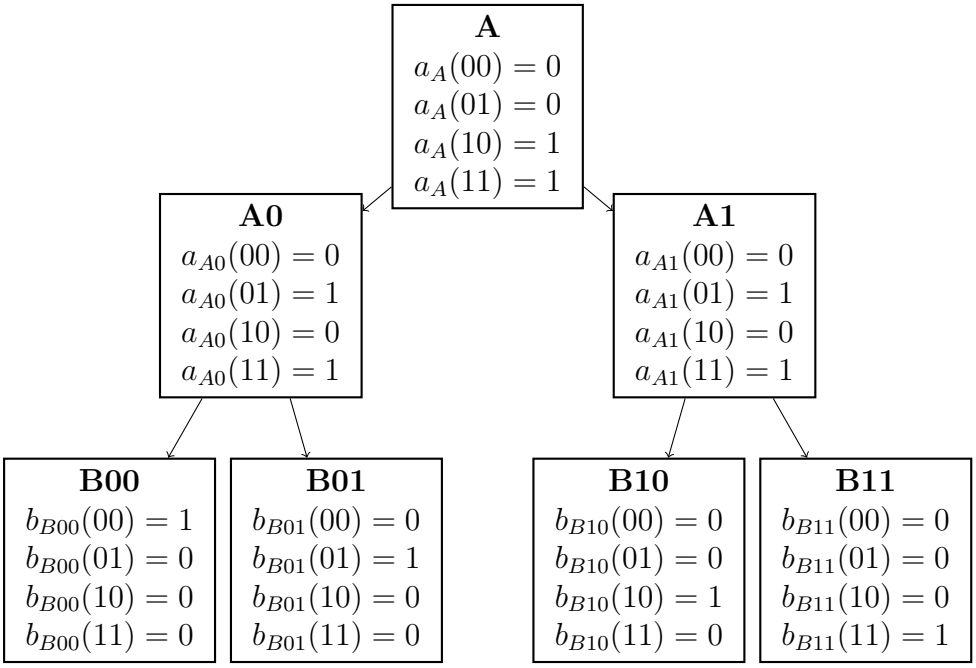


Figure 6.1: A communication protocol for equality. Alice speaks first and sends two bits to Bob. Upon receiving the two bits, Bob checks if they are equal to his input. If they are, he outputs 1, otherwise he outputs 0. Every node in the tree is labeled by the party that takes the turn and a communication transcript up to that point: for example $B01$ means that it is Bob's turn to speak and both parties have so far communicated a bitstring 01. For every leaf node, Bob announces the value that his last vertex function evaluated to.

bit partitions it into subsets $A0 = \{00, 01\} \times Y$ and $A1 = \{10, 11\} \times Y$, the second bit into: $A00 = \{00\} \times Y$, $A01 = \{01\} \times Y$, $A10 = \{10\} \times Y, A11 = \{11\} \times Y$.

Deterministic communication complexity quantifies the minimum amount of communication needed to solve the communication task with a deterministic protocol. It is defined as follows:

Definition 17 (Deterministic Communication Complexity [93, 99])**.** *The deterministic computation complexity of a function $f : X \times Y \to Z$ is the minimum cost over all (deterministic) protocols for f.*

Every function $f : X \times Y \to Z$ has trivial deterministic communication complexity of $O(\log |X|)$ since the first player can send their input to the other player in a single round. The second player then evaluates $f(x, y)$ – this was the case in Example 5.

It is a central question of communication complexity to find out if a communication problem has more efficient protocol than this, or to rule this possibility out.

An appealing feature of communication complexity is that it is possible to prove communication complexity lower bounds without any additional assumptions or complexity conjectures – this in turn allows to prove unconditional complexity separations between quantum and classical communication. Here we only discuss a rudimentary lower bound technique called the *tiling bound* based on counting the number of monochromatic *combinatorial rectangles*:

Definition 18 (Combinatorial Rectangles [93])**.** *Let R_v be the set of inputs (x, y) for which the traversal through a protocol P passes through an internal vertex $v \in P$. Notice that $R_v \cap R_{v'} \neq \emptyset$ if and only if v is a descendant of v' or v' is a descendant of v. This means that the set $\{R_l\}_{l \in L}$, where L is the set of leaves of P, partitions the inputs $X \times Y$. The subsets R_v do not have an arbitrary structure and are called combinatorial rectangles: we have that $R_v = A \times B$ for some $A \subseteq X$ and $B \subseteq Y$.*

This property can be seen by induction. At the root of the protocol P, $R = X \times Y$, which is a combinatorial rectangle. For an internal node w, assume that $R_w = A_w \times B_w$ and let v be a child of w. Without loss of generality, assume that Alice speaks so that $a_w : X \to \{0, 1\}$ and that v is the left child of w. Then $R_v = R_w \cap \{(x, y) |\, a_w(x) = 0\} = (A_w \cap \{x |\, a_w(x) = 0\}) \times B_w$, which is a rectangle.

A subset $R \subseteq X \times Y$ is called *f-monochromatic* if the value of f is fixed on R; that is:

$$\forall\, (x, y),\, (x', y') \in R : f(x, y) = f(x', y'). \tag{6.2}$$

Any protocol P for f partitions $X \times Y$ into f-monochromatic rectangles because it fixes values of f on its leaves – recall that the leaves partition the set of inputs $X \times Y$, as shown in Def. 18. If any partition of $X \times Y$ into f-monochromatic rectangles requires at least T such rectangles, then so does the partition induced by the protocol P. This means that the deterministic communication complexity is lower-bounded by $\log(T)$ – this is the *tiling bound* [93]. The idea is to show that there are many monochromatic rectangles, usually by proving that the size of each f-monochromatic rectangle is small or counting them directly:

Example 5 (Tiling Lower Bound for Exact Communication of Equality [93])**.** *Consider the equality function over $X \times Y$, $X = Y = \{0,1\}^n$. Let $S_{EQ} = \{(k,k) | \, k \in X\}$ be the set of the 2^n diagonal entries of the input. Suppose that two distinct elements of S_{EQ}, (i,i) and (j,j) were in the same 1-monochromatic rectangle – then (i,j) and (j,i) would have to be, which is not the case. Any deterministic protocol must partition the inputs into at least 2^n monochromatic rectangles and the deterministic communication complexity becomes lower-bounded by n. Since Alice can trivially send her input to Bob using n bits, the one-way strategy from Ex. 5 is optimal and the bound is tight. We show later that a randomized strategy for this task can do significantly better and achieves $O(1)$ complexity in the bounded error case.*

6.3 Randomized Communication

A randomized protocol is a communication protocol where each party has access to a random string in addition to the inputs. A protocol $P : X \times Y \to \{0,1\}$ computes $f : X \times Y \to \{0,1\}$ with zero error if:

$$\forall (x,y) \in X \times Y : \Pr\left[P(x,y) = f(x,y)\right] = 1, \tag{6.3}$$

The protocol computes f with ϵ-error if for every (x,y):

$$\forall (x,y) \in X \times Y : \Pr\left[P(x,y) = f(x,y)\right] \geq 1 - \epsilon. \tag{6.4}$$

The averages are taken over the random strings in both of these cases. A protocol is called *public coins* if the same random string can be accessed by all parties, otherwise it is called a private coins [95]. Unless explicitly mentioned, all protocols are public coins in this thesis. A randomized protocol can be also seen as a distribution over deterministic protocols, indexed by a value of shared randomness drawn from a distribution π:

$$P(x,y) = \sum_s P(x,y|\, s)\pi(s). \tag{6.5}$$

Alice and Bob generate a string $s \sim \pi$ and follow a deterministic protocol $P(x,y|\, s)$ for the task. Randomized communication cost is defined as follows:

Definition 19 (Randomized Communication Cost [93]). *The worst case running time of a randomized protocol P on an input (x, y) is the maximum number of bits communicated for any choice of s. The worst case cost of P is the worst case running time for any choice of $(x, y) \in X \times Y$. The average case running time of a randomized protocol P is the expected number of communicated bits, averaged over $s \sim \pi$. The average case cost is the average case running time on the worst input $(x, y) \in X \times Y$.*

The above definition leads to distinct complexity measures of which we use the zero-error and bounded error randomized communication complexities.

Definition 20 (Randomized Communication Complexity). *The zero-error randomized communication complexity of f is the average communication cost of the best protocol that computes f. It will be denoted by $R_0(f)$. The ϵ-error randomized communication complexity of f is the worst case communication cost of the best protocol that computes f with error at most $1/2 > \epsilon > 0$. It is denoted by $R_\epsilon(f)$.*

Note that the zero-error randomized complexity counts the average communication cost instead of the worst case one – it can be shown that the worst case zero-error randomized complexity is the same as deterministic complexity, but also that there can be gaps between average and deterministic communication costs. Randomization can significantly reduce cost of communication protocols, as we now illustrate with the randomized protocol for equality:

Example 6 (Public Coin Bounded-Error Protocol for Equality [101]). *The equality function has a $O(1)$ bounded error protocol. Work with input sets $X = Y = \{0, 1\}^n$. Alice gets $x \in X$, Bob gets $y \in Y$. Suppose that Alice and Bob also hold a shared random string $s \in \{0, 1\}^n$. Alice computes $a = \sum_i x_i s_i$ (mod(2)) and sends it to Bob (one bit). Bob computes $b = \sum_i y_i s_i$ (mod(2)) and checks if $a = b$. If $x = y$, then $a = b$ with certainty. If $a \neq b$, then $a = b$ w.p. $1/2$. Alice and Bob can perform $O(\log(1/\epsilon))$ protocol runs to win the game with at most ϵ error.*

It is usually more difficult to lower bound randomized communication complexity compared to the deterministic case. A very useful tool that addresses this problem is Yao's principle [102] that establishes an equivalence between randomized complexity and complexity of deterministic protocols with some distribution on inputs:

Definition 21 (Distributional Communication Complexity [93]). *Let $\mu : X \times Y \to [0,1]$ be a distribution on inputs. The μ-distributional communication complexity of f with error ϵ, denoted by $D^{\mu}_{\epsilon}(f)$, is the cost of the best deterministic protocol that gives a correct answer for f on at least $1-\epsilon$ fraction of inputs $X \times Y$ weighted by μ.*

Theorem 6 (Yao's principle [102]).

$$R_{\epsilon}(f) = \max_{\mu} D^{\mu}_{\epsilon}(f). \tag{6.6}$$

To prove a lower bound on $R_{\epsilon}(f)$, it therefore suffices to find a hard distribution on inputs for the deterministic protocols for f. Since the principle is used for R_0 in our work, we show that $\forall \mu : R_{\epsilon}(f) \geq D^{\mu}_{\epsilon}(f)$ for all $\epsilon \geq 0$ here – the proof is given in Appendix B.1.

6.4 Relations

The main result of this part concerns the communication complexity of relations:

Definition 22 (Relations). *A relation R is a subset $R \subseteq X \times Y \times Z$, where X, Y should be understood as the set of inputs passed to Alice and Bob respectively and Z as the set of possible outcomes. The corresponding communication problem is as follows: Alice gets $x \in X$, Bob gets $y \in Y$ and aim to find $z \in Z$ such that $(x, y, z) \in R$.*

It may be that for a certain input pair (x, y), there is no value z, such that $(x, y, z) \in R$. Such input is called illegal and it is assumed that it is never given to Alice and Bob. Relations that are at most single valued – that is, for all legal inputs, there exists a unique $(x, y, z) \in R$ – are called partial functions. Relations that have no illegal inputs are called total.

Definition 23 (Deterministic Communication Complexity of Relations [93]). *A protocol P computes the relation $R \subseteq X \times Y \times Z$ if for every legal input (x, y), it reaches a leaf marked by z, such that $(x, y, z) \in R$. The deterministic communication complexity for relations is the number of bits sent on the worst (legal) input pair (x, y) by the best protocol that computes R.*

The tiling lower bound for relations works similarly as before, up to a modification to the definition of a monochromatic rectangle:

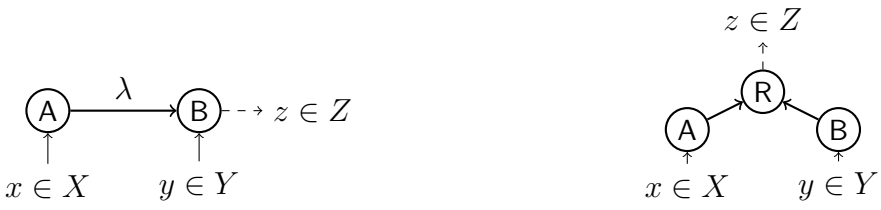
Definition 24 (Monochromatic Rectangles for Relations [93])**.** *$A \times B \subseteq X \times Y$ is a monochromatic rectangle with respect to a relation R if there exists a value z, such that $\forall (x, y) \in A \times B : (x, y, z) \in R$ or (x, y) is illegal.*

As before, any T-bit protocol P that computes a relation R partitions the inputs $X \times Y$ into at most 2^T monochromatic rectangles. The tiling lower bound applies up to this modification and so does Yao's principle [93].

6.5 Simultaneous Message Passing and One-Way Communication

We briefly mention that other communication layouts were studied in two-party communication complexity. In the *simultaneous message passing* model, there is a third party – the referee. Alice and Bob are given inputs, are not allowed to talk to one another, but can each send only one message (comprising of a set of bits) to the referee.

In the *one-way communication model*, Alice is only allowed to send a single message to Bob. The one-way communication complexity of total relations is especially important to us, as it is closely related to simulation of prepare-measure experiments in quantum theory. Both layouts are described in Fig. 6.2.



(a) One-Way Communication **(b)** Simultaneous Message Passing

Figure 6.2: Other commonly studied two-party communication layouts include Simultaneous Message Passing and One-Way Communication.

6.6 Quantum Communication Complexity

Randomization can significantly reduce communication complexity and it is natural to ask if quantum resources provide further improvement. Quantum communication was first studied by Yao in Ref. [103] who defined the model of quantum communication complexity. In a two-party quantum communication scenario, the parties hold qubits, some of which are initialized to their inputs. In a communication round – where a round is a sequence of steps during which only one of the players talks – a player performs a unitary operation on their qubits and sends some of them to the other player. At the end of the protocol, one of the players measures his or her qubits and reports the outcome [104]. Quantum communication complexity is the number of qubits exchanged in the optimal bounded-error protocol. There are other models of quantum computational complexity – for example, Cleve and Buhrmann [105] introduced a model in which Alice and Bob share an unlimited number of entangled states, but can only communicate classically and the only resource counted is the amount of communication.

Quantum strategies can be indeed more powerful than classical ones. The first significant separations between quantum and classical communication complexity were found by Buhrman, Cleve and Wigderson [106] who initiated the study of distributed versions of quantum query algorithms. They proved that a communication version of Deutsch-Jozsa algorithm gives a promise-equality problem that gives exponential separation between quantum and classical exact communication complexities:

Example 7 (Exponential Communication Gap for Promise-Equality [106])**.** *Without loss of generality, assume that d is a power of two. Let $x, y \in \{0,1\}^d$ be Alice's and Bob's inputs respectively. The problem is to find if $x = y$, under the promise that either $x = y$ or that $h(x, y) = d/2$, where h denotes the Hamming distance. A quantum protocol is as follows:*

1. *Alice sends $\frac{1}{\sqrt{d}} \sum_i (-1)^{x_i} |i\rangle$.*
2. *Bob applies $|i\rangle \mapsto (-1)^{y_i} |i\rangle$, and measures each qubit in the $|\pm\rangle$ basis.*
3. *His output is 1 if the measurement gave '+' outcome on every measured qubit. He outputs 0 otherwise.*

Denote $n = \log_2 d$. The overlap of Bob's pre-measurement state with $|+\rangle^{\otimes n}$ is:

$$\frac{1}{\sqrt{d}} \langle +|^{\otimes n} \sum_{i}^{N} (-1)^{x_i + y_i} |i\rangle = \frac{1}{d} \sum_{i}^{d} (-1)^{x_i + y_i}. \tag{6.7}$$

If $x = y$, then $(-1)^{x_i+y_i} = 1$ for all i, in which case Bob receives the all-'+' outcome with certainty. If $h(x, y) = d/2$, the overlap becomes zero. There is a $0.01n$ classical lower bound for the problem (with little effort, this can be strengthened to $0.02n$), leading to an exponential separation between quantum and classical communication. The proof is however complicated and hinges on a combinatorial result of Frankl and Rödl [107]. Thanks to the small scaling constant, the result also only gives separation for large inputs and is only meaningful in the asymptotic regime. Our results, introduced later, address both of these issues in a different communication problem. Notice also, that the asymptotic separation disappears if we compare to a bounded error protocols – Example 6 described a $O(1)$ protocol for equality that also applies for the promise problem studied here.

The field of quantum communication complexity is relatively new and underwent significant developments recently. Some of them are summarized below:

- (1998) *Promise Equality* and *Set Disjointness*: Buhrman, Cleve and Wigderson gave an $\Omega(d)$ vs $O(\log d)$ separation between classical and quantum exact communication complexities for the promise task of deciding whether two d-bit inputs are either equal or have Hamming distance $d/2$ [106]. We have outlined this in Ex. 7. The work also shows a quadratic gap between randomized and quantum communication complexity of set disjointness by considering a distributed variant of Grover's algorithm. This gap has been for a very long time the largest known separation for total functions, surpassed only recently by [108].

- (1999) *Vector in a Subspace*: Raz [109] gave a partial function with efficient quantum two-way protocol but no efficient classical two-way protocol. This was the first separation between randomized and quantum communication complexities that held up to bounded error. The studied two-party promise problem is derived from the vector in a subspace problem, originally proposed by Kremer in [100]. In such task, Player I gets as an input a real vector $x \in \mathbb{R}^d$ and

two orthogonal subspaces $M_0, M_1 \subset \mathbb{R}^d$. Player II gets an orthogonal matrix $T : \mathbb{R}^d \to \mathbb{R}^d$. Their aim to answer 0 if $T(x) \in M_0$ and 1 if $T(x) \in M_1$ and any answer in any other case. For a discretized version of the problem, where the parties get inputs of finite length, while retaining the complexity lower bounds, the task exhibits a separation $O(\log n)$ vs $\Omega(n^{1/4}/\log n)$ between two-way classical communication complexity and quantum communication for inputs of length n [109].

- (2001) *Quantum Fingerprinting*: A total function was found that had an efficient quantum simultaneous message passing protocol without shared randomness but no efficient classical simultaneous message passing protocol without shared randomness. The task is called 'quantum fingerprinting' and we borrow some proof techniques from this paper for our purposes [110]. Specifically, equality in simultaneous message passing model can be solved with $O(\log n)$ qubits rather than $\Theta(\sqrt{n})$ bits in a bounded error classical simultaneous message passing protocol with private randomness. With shared randomness however, only $O(1)$ bits suffice classically.

- (2004) *Hidden Matching Problem*: Bar-Yossef, Jayram and Kerenidis [104] proved the first exponential separation between randomized (bounded error) one-way communication complexity and quantum exact communication complexity with public coins for a relation called the Hidden Matching Problem.

  In the task, Alice receives a string $x \in \{0,1\}^n$, Bob gets a perfect matching M on $\{1, \ldots n\}$. The aim is to output a triple $(i, j, x_i \oplus x_j)$ for some $(i, j) \in M$.

  There is a $O(\log(n))$ quantum protocol for this task. Alice sends $\frac{1}{\sqrt{n}} \sum_i^n (-1)^{x_i} |i\rangle$. Bob performs a projective measurement in the orthonormal basis:

$$\left\{ \frac{1}{\sqrt{2}} \left(|j\rangle \pm |k\rangle \right) \mid (j,k) \in M \right\}. \tag{6.8}$$

  The probability of an outcome $\frac{1}{\sqrt{2}} \left(|j\rangle + |k\rangle \right)$ is given by:

$$\left| \langle \psi | \left(\frac{1}{\sqrt{2}} |j\rangle + |k\rangle \right) \right|^2 = \frac{1}{2n} \left((-1)^{x_j} + (-1)^{x_k} \right)^2, \tag{6.9}$$

  which is $2/n$ if $x_j \oplus x_k = 0$ and 0 otherwise. If this is the outcome of the measurement, Bob outputs $(j, k, 0)$. Similarly, the probability of $\frac{1}{\sqrt{2}} \left(|j\rangle - |k\rangle \right)$

outcome is 0 if $x_j \oplus x_k = 0$ and $2/n$ otherwise, in which case Bob outputs $(j, k, 1)$.

Bar-Yossef, Jayram and Kerenidis also show a $\Omega(\sqrt{n})$ lower bound for a randomized one-way communication complexity for the problem, which gives an exponential asymptotic separation. Due to the error tolerance, this asymptotic separation is strictly stronger than the one we manage to prove here. The proof is however much more involved than ours.

- (2007) *One-way Separation for Partial Functions*: Gavinsky, Kempe, Kerenidis, Raz and de Wolf [111] strengthened the previous result to a separation in one-way communication complexity of partial Boolean function.

- (2010) *One-way Vector in a Subspace*: Klartag and Regev proved an exponential separation between a one-way quantum protocol for the vector in a subspace problem and a two-way classical communication protocols [112].

- (2016) *Distributed Fourier Sampling*: Montanaro [13] proved that any classical two-way communication protocol with shared randomness that approximately simulates the result of applying an arbitrary measurement to a quantum state, up to a constant accuracy, must transmit at least $\Omega(2^n)$ bits. The lower bound is optimal, and based on a communication complexity of Distributed Fourier Sampling. The separation is however derived for sampling problems, where the task is to output a sample from a distribution D_x drawn from a family $\{D_x\}_{x \in X}$ on an input $x \in X$.

- (2016) *Power-5/2 Sepatation for a Total Function*: Anshu et al. [108] broke the power-2 barrier for separation in communication complexity of total functions by finding a power-5/2 separation between quantum and randomized complexities for a total function.

- (2019) *Quantum SMP vs Classical Two-Way*: Gavinsky found a relation with efficient quantum simultaneous message passing protocol, but no efficient classical two-way protocol in [113].

6.7 Quantum Communication and Ontological Models

Quantum theory is extremely accurate in predicting experimental outcomes, making it the central building block for physical theories developed in the last century. That despite, it remains disturbingly close to a set of rules of thumb. Even the very notion of a quantum state remains a topic of an active debate – is it a representation of agent's knowledge (epistemic; episteme means 'to know' in Greek) or a 'state of reality' (ontic; 'ontos' loosely translates from Greek as 'that which is') [114, 115]? Attempts to address this question propelled the development of *operational theories* and *ontological models*, that, for our purposes, are conveniently intertwined with communication complexity [116, 117]. Since this connection largely motivated our work, we briefly outline it here.

An operational theory is specified by primitive laboratory operations – preparations, transformations and measurements – together with a prescription for finding probabilities $p(k|\, M, P)$ for each outcome k, given a measurement M and a preparation P [116, 117]. Both classical probability theory and quantum theory fit in this framework: classical preparations correspond to probability distributions, transformations to functions applied to random variables distributed according to these distributions and measurements correspond to sampling. Quantum preparations are quantum states (including mixtures), transformations are completely positive maps and measurements are POVMs. We will do without transformations in what follows.

An ontological model of an operational theory associates to every physical system a set Λ called the ontic state space. Every ontic state $\lambda \in \Lambda$ defines all physical properties of the system. A preparation P of a system samples the system's ontic state λ from a probability distribution $\mu_P(\lambda) : \Lambda \to [0, 1]$ that is associated to it. This probability distribution is called an 'epistemic state', because it aims to reflect the knowledge (or lack of thereof) about the ontic state of the system. A measurement M on the system has its outcome k sampled randomly from a distribution characterized by the ontic state λ. The probability distribution on the outcomes k, given a measurement M and λ is called the response function: $\xi_M(k|\lambda)$. An ontological model must reproduce its empirical predictions and if we require it to do so exactly,

it follows that:

$$p(k|\, M, P) = \int_{\Lambda} d\lambda\, \xi_M(k|\lambda)\mu_P(\lambda). \tag{6.10}$$

Here we only do with discrete ontic spaces Λ and also avoid discussions of Spekkens' contextuality for brevity [117].

The consistency condition can be related to a one-way communication complexity in the following way: Let Alice's input be P, a classical description of a preparation and Bob's input M, a classical description of a measurement. Assume first that the parties can only generate private randomness. In this case, Alice prepares a message λ with probability $\mu_P(\lambda)$ and Bob outputs an answer k with probability $\xi_M(k|\lambda)$. The probability of the communication protocol giving an output k for inputs M, P is the same as Eq. 6.10. Every private coins one way communication protocol can be therefore also interpreted as an ontological model. If we assume that the parties can access shared randomness drawn from some probability distribution π, the comparison becomes less straightforward. While any ontological model also gives a public coin randomized protocol,[1] it does not mean that every public coin protocol gives an ontological model. It is however the case for bounded error protocols, justified by Newman's theorem that shows that difference between public and private coin communication complexity randomized ϵ-error protocols is at most $O(\log\log(|X||Y|) + \log(1/\epsilon))$ [118]. Any public coin protocol can be turned into private coins one with this overhead that corresponds to an 'ontological model' in which the consistency condition of Eq. 6.10 holds only up to ϵ-error.

Several results in foundations touch upon the connections of communication complexity and ontological models. For example, Hardy's ontological excess baggage theorem [119] asserts that if one wants to exactly communicate an identity of an arbitrary quantum state with classical resources, an infinite amount of bits must be sent in the absence of shared random data. Exactly reproducing the results of communication by quantum channels may hence seem unlikely. Contrasting to that however, it was proved by Toner and Bacon in Ref. [120] that just two bits of communication suffice to reproduce any projective measurement on a qubit state if the parties share randomness. More recently, one of the key results in favor of 'onticity' of quantum states,

[1] this follows because any private coin is straightforwardly a public coin protocol – we just ask the respective parties to ignore parts of the shared randomness.

the PBR theorem [97], was used to study separations in communication complexity in [121] where large separations between one-way communication and information complexities for instances of the so-called 'exclusion games' have been proven. A followup work [122] proved a quantum communication complexity lower bound for an exclusion game that implies a sub-constant versus logarithmic gap between the quantum information and communication. A sequence of works by Montina investigated connections between ontological models and communication complexity [123, 124, 125, 126, 127] and we use some of his results in what follows.

Chapter 7

Antidistinguishablity

7.1 Introduction

This chapter reviews some results in quantum detection theory needed to introduce our result. We also reproduce a lemma of Caves, Fuchs and Schack that will be central to our proof. We then briefly review some other results on quantum state antidistinguishability.

7.2 Quantum State Discrimination

It is an important feature of quantum theory, implicit to its postulates, that non-orthogonal states are not perfectly distinguishable [128]. Discriminating quantum states can be formulated as a one-way quantum communication task: Alice prepares a quantum state out of some set of states S and sends it to Bob, who applies a measurement. Both parties know what states are contained in S, so that Bob's task is to correctly identify the state that Alice chose. How much can Bob learn about the transmitted state by a quantum measurement?

While there are various figures of merit, the most relevant to us will be the minimum error single-shot distinguishability. A typical single-shot state distinguishing game looks as follows: Alice prepares a d-dimensional quantum state ρ_k, drawn from some set of states $S = \{\rho_1, \rho_2 \dots \rho_N\}$ with probability q_k such that $\sum_k^N q_k = 1$. Upon the receipt of his state, Bob applies a POVM $M = \{M_k\}_{k=1}^N$.

$$M_k \succeq 0, \qquad \sum_k^N M_k = \mathbf{1}. \tag{7.1}$$

For a state ρ_k, Bob receives an outcome i with probability:

$$p(i|k) = \mathrm{Tr}\left(M_i \rho_k\right). \tag{7.2}$$

He aims to minimize the mismatch error:

$$\epsilon := \sum_{k \neq i} q_k \, p(i|\, k) = 1 - \sum_k q_k \, p(k|\, k). \tag{7.3}$$

Rewritten in terms of the measurement operators, this can be cast to the following optimization problem:

$$\begin{aligned} \text{maximize: } & \sum_k \mathrm{Tr}(M_k q_k \rho_k), \\ \text{s.t.: } & \sum_k M_k = \mathbf{1}, \\ & M_k \succeq 0, \end{aligned} \tag{7.4}$$

which is a semidefinite program that can be solved numerically. Analytic optimal solutions are so far known only in special cases. For example the optimal measurement for a two-state minimum error discrimination was (independently) found by Holevo and Helstrom [129], an approximately optimal measurements ('pretty good measurements') where derived by Knill and Barnum. Recently, an analytic solution to three qubit state discrimination was found in Ref. [130]. See Ref. [128] for a 2015 review.

7.3 Quantum State Antidistinguishability

Yuen, Kennedy and Lax studied multiple state hypothesis in Ref. [131]. For an ensemble of states $S = \{q_i, \rho_i\}$, where q_i's denote the prior state preparation probabilities and ρ_i's are the prepared states, they defined the Bayes' costing function as:

$$\bar{C} = \sum_{i,j} C_{ij} \, \mathrm{Tr}\left(q_i M_j \rho_i\right), \tag{7.5}$$

for some costing matrix C_{ij}. The minimum error discrimination problem becomes:

$$\text{minimize } \bar{C}, \tag{7.6}$$

$$\text{s.t.: } M_i \succeq 0, \ \forall i, \tag{7.7}$$

$$\sum_i M_i = \mathbf{1}. \tag{7.8}$$

For example, using $C = (E - \mathbf{1})$ in Eq. 7.5, where E is an all unit $d \times d$ matrix corresponds to the minimum error state discrimination cost described in the previous section. One can however choose different costings – specifically, choosing $C = \mathbf{1}$ corresponds to 'single state exclusion', which will play role in what follows.

We also consider modification of the state discrimination game studied in the previous section, in which Alice sends Bob a pure quantum state $|\rho_i\rangle \in S$, but Bob's task is now to rule out with certainty (but without aborting) what state it is not. Following Ref. [132], we call the task *conclusive exclusion* and call the sets of quantum states that allow for perfect conclusive exclusion *antidistinguishable*. More formally, a set S of n pure quantum states $|\rho_1\rangle, |\rho_2\rangle, \ldots, |\rho_n\rangle$ will be called antidistinguishable if there exists an n-outcome measurement $M = \{M_z \mid z \in [n]\}$ that never outputs the outcome z on the quantum state $|\rho_z\rangle$, i.e.,

$$M_z |\rho_z\rangle = 0, \ \forall\, z \in [n]. \tag{7.9}$$

Antidistinguishability of quantum states lies at the heart of the PBR theorem [97], which rules out ψ-epistemic ontological models of quantum theory, under the assumption of *preparation independence*. The concept has been therefore well studied in foundations literature [97, 96, 133, 115, 134]. Caves, Fuchs and Schack derived necessary and sufficient conditions for three states to be antidistinguishable by a projective measurement in Ref. [96]. They specifically show that:

Lemma 1 (Antidistinguishability [96])**.** *Three distinct pure quantum states $|\rho_i\rangle, |\rho_j\rangle, |\rho_k\rangle$ are antidistinguishable with a projective measurement if and only if there exist angles θ_k, $0 < \theta_k < \pi/2$ for $k = 1, 2, 3$, such that:*

$$\begin{aligned} a &:= |\langle\rho_1|\rho_2\rangle|^2 = (\sin\theta_1 \cos\theta_2)^2, \\ b &:= |\langle\rho_2|\rho_3\rangle|^2 = (\sin\theta_2 \cos\theta_3)^2, \\ c &:= |\langle\rho_1|\rho_3\rangle|^2 = (\sin\theta_3 \cos\theta_1)^2. \end{aligned} \tag{7.10}$$

Proof. If the states are antidistinguishable, there exist a basis $|1\rangle, |2\rangle, |3\rangle$, such that:

$$\begin{aligned} |\rho_1\rangle &= e^{i\xi_1}(\cos\theta_1 |2\rangle + e^{\phi_1} \sin\theta_1 |3\rangle), \\ |\rho_2\rangle &= e^{i\xi_2}(\cos\theta_2 |3\rangle + e^{\phi_2} \sin\theta_2 |1\rangle), \\ |\rho_3\rangle &= e^{i\xi_3}(\cos\theta_3 |1\rangle + e^{\phi_3} \sin\theta_3 |2\rangle). \end{aligned} \tag{7.11}$$

Taking the inner products of these states gives Eq. 7.10.

Conversely, if the conditions of Eq. 7.10 are satisfied, it is possible to find the angles ξ_k and ϕ_k for $k \in \{1, 2, 3\}$, such that the inner products satisfy Eq. 7.11, since the pairwise inner products define the states up to a unitary transformation. This means that there exists an orthonormal basis $|1\rangle, |2\rangle, |3\rangle$ such that the states can be written as Eq. 7.11 and therefore a projective measurement that antidistinguishes them. □

Ref. [96] followed on to give a simpler antidistinguishability condition in terms of state overlaps. We reproduce their proof in detail for completeness:

Lemma 2 (Three-state antidistinguishability [96])**.** *Any three pure states are antidisitnguishable by a projective measurement if and only if:*

$$a = |\langle \rho_i | \rho_j \rangle|^2, \qquad b = |\langle \rho_i | \rho_k \rangle|^2, \qquad c = |\langle \rho_k | \rho_j \rangle|^2, \tag{7.12}$$

we have that:

$$a + b + c < 1, \qquad \textit{and} \qquad (1 - a - b - c)^2 \geq 4abc. \tag{7.13}$$

Proof. Let $x_k = \sin^2 \theta_k$ for $k \in \{1, 2, 3\}$ as in Eq. 7.10. Assume again that the states are distinct, so that $0 < x_1, x_2, x_3 \leq 1$. Eq 7.12 reads:

$$x_1(1 - x_2) = a, \qquad x_2(1 - x_3) = b, \qquad x_3(1 - x_1) = c. \tag{7.14}$$

Solving this system for x_2 in terms of the variables a, b, c, leads to:

$$x_2 = \frac{1 - a + b - c \pm \sqrt{(1 - a + b - c)^2 - 4b(1 - a)(1 - c)}}{2(1 - c)}, \tag{7.15}$$

$$= \frac{1 - a + b - c \pm \sqrt{(a + b + c - 1)^2 - 4abc}}{2(1 - c)}, \tag{7.16}$$

and analogous equations for x_1 and x_3 follow from cyclic permutations of a, b, c. Real solutions x_1, x_2, x_3 exist for:

$$(a + b + c - 1)^2 \geq 4abc, \tag{7.17}$$

The sign of these solutions is determined by $1 - a + b - c$ by Eq. 7.15, because:

$$|1 - a + b - c| \geq \sqrt{(1 - a + b - c)^2 - 4b(1 - a)(1 - c)}. \tag{7.18}$$

This means that the condition $0 < x_2$ is equivalent to:

$$0 < 1 - a + b - c. \tag{7.19}$$

The condition that there exists $x_2 \leq 1$ is then:

$$1 - a + b - c - \sqrt{(a + b + c - 1)^2 - 4abc} \leq 2(1 - c). \tag{7.20}$$

Noting that $2(1 - c) = (1 - a + b - c) + (1 + a - b - c)$ leads to:

$$-(1 - b + a - c) \leq \sqrt{(a + b + c - 1)^2 - 4abc}, \tag{7.21}$$

which is already implied by the condition $1 - b + a - c > 0$. The full set of conditions is then given by:

$$1 - a + b - c > 0, \qquad 1 - b + c - a > 0, \qquad 1 - c + a - b > 0, \tag{7.22}$$

$$(a + b + c - 1)^2 \geq 4abc. \tag{7.23}$$

Eqs. 7.22 can be rewritten as:

$$|a - b| < 1 - c, \qquad a + b < 1 + c. \tag{7.24}$$

Eq. 7.23 is a conic section that can be transformed to a standard form by noting that: $(a + b - 1)^2 - (a - b)^2 = 4ab - 2a - 2b + 1$ as follows:

$$\begin{aligned}
(a + b + c - 1)^2 &\geq 4abc, \\
(a + b - 1)^2 + 2c(a + b - 1) + c^2 &\geq 4abc, \\
(a + b - 1)^2 + c(2a + 2b - 1 - 4ab) - c + c^2 &\geq 0, \\
(1 - c)(a + b - 1)^2 + c(a - b)^2 - c(1 - c) &\geq 0, \\
\frac{(a + b - 1)^2}{c} + \frac{(a - b)^2}{(1 - c)} &\geq 1.
\end{aligned}$$

For a fixed $c \in (0, 1)$, substituting $x = a + b$ and $y = a - b$, the boundary of the inequality gives an ellipse centered at $x = 1$, $y = 0$ with x-semiaxis of length $\sqrt{c}$ and y-semiaxis of length $\sqrt{1 - c}$. In the a–b coordinates, this means that $a = b = 1/2$, with

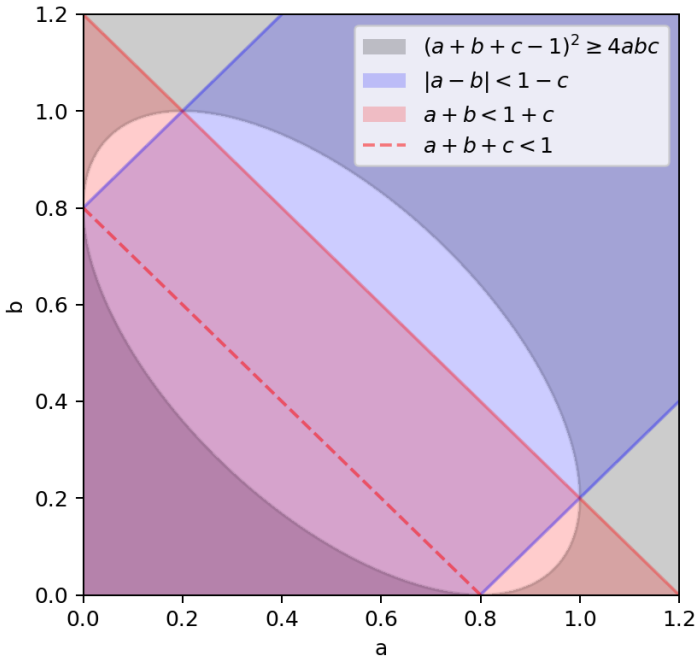


Figure 7.1: Conditions Eq. 7.22 and Eq. 7.23 imply that the necessary and sufficient antidistinguishability inequalities for three states reduce to: $a+b+c < 1$ and $(a+b+c-1)^2 \geq 4abc$. The simplification can be deduced from the above diagram, where the inequalities are plotted for a fixed value of c. A filled region implies that a given inequality is satisfied. The antidistinguishing region is then the region where all the three colors overlap. By simplifying the constraints, the region can be defined as all solutions of $(a+b+c-1)^2 \geq 4abc$ for $0 < a, b$ in the halfspace defined by the dashed line, given by is $a + b + c < 1$. Boundaries of the plot were chosen between as $a, b \in [0, 1.2]$ for clarity, note however that $a, b \leq 1$.

semiaxis rotated by 45° with respect to the coordinate axes with semiaxes of length $\sqrt{c/2}$ and $\sqrt{(1-c)/2}$. The ellipse has a single intersection point with both a and b axes at $(a, b) = (1-c,\ 0)$ and at $(a, b) = (0,\ 1-c)$. It follows that Eq. 7.22 and Eq. 7.23 are both satisfied if and only if $a+b+c < 1$ and $(a+b+c-1)^2 \geq 4abc$. This conclusion is not straightforward and a diagrammatic justification is provided in Fig. 7.1. It follows that Eq. 7.13 is necessary and sufficient condition for antidistinguishability of three pure states. □

We remark that the above result was also shown to be a sufficient and necessary condition for antidistinguishabiliy by POVMs by Molina in Ref. [135]. The lemma is the key ingredient to the proof of Pusey, Barrett and Rudolph in Ref. [97] and attracted some attention since the publication of this proof. Bandyopadhyay, Jain,

Oppenheim and Perry [132] studied 'quantum state exclusion', formulated antidistinguishability as a semidefinite program (essentially by choosing the right costings in Eq. 7.5) and – translating directly from the SDP framework – gave necessary and sufficient conditions for optimality.

Heinosaari and Kerppo studied algebraic conditions for a set of pure states to be antidistinguishable in Ref. [133]. They show that a finite set $S = \{P_i\}_i$ of pure states (treated as rank-1 projectors in the equation below) is antidistinguishable if there exist nonnegative reals $\{t_i\}_{i \in S}$, $t_i \geq 0$, $\forall i$ such that:

$$\sum_j t_j P_j = R, \tag{7.25}$$

where R is a projection with kernel $\cap_j \ker P_j$. We will conjecture another sufficient condition and provide numerical evidence of thereof in Chapter 8.

The authors later studied the task of 'communicating partial ignorance' in Ref. [136] that came out while we were finishing our project and has overlap with our work. The key object of study in their work is, similarly to us, a communication task between Alice and Bob, in which Alice is trying to transmit to Bob information about what alternatives he should not chose. They provide full characterization of the communication task for qubits and derive certain solutions for qudits, discuss connections to symmetric informationally complete states (SICs) and also show that a real Hilbert space qubit cannot communicate as much 'ignorance' as the ordinary qubit. Compared to us however, they do not derive asymptotic complexity separations the packing argument or discuss the two-way communication complexity of the task for which we derive an explicit protocol in Chapter 8.

Chapter 8

Communication Complexity Separation from Antidistinguishability

The following chapter is based on an article with Jon Barrett [137].

8.1 Introduction

How difficult is it to communicate classically the identity of a quantum state in an entanglement-unassisted scenario? Specifying a pure state of a qubit requires two real numbers, so communicating its identity exactly seemingly requires an infinite amount of classical information [138]. This suspicion is confirmed by the results of Ref. [119] that shows that if the two communicating parties do not share random data and a small error is tolerated, an unbounded amount of classical communication is needed to exactly simulate results of quantum experiments. Assuming unbounded amount of shared randomness however, Bacon and Toner proved in Ref. [120] that just two bits suffice to exactly reproduce results of any projective qubit measurement.

The main result of this chapter is a simple lower bound for the classical communication cost of simulating the transmission of a d-dimensional quantum state. We find it by describing a communication problem based on a relation, and proving an exponential separation between the classical and quantum communication complexities. Our proof uses quantum state antidistinguishability, a concept that has been studied in the foundations literature [97, 96, 133, 115, 134], and lies behind PBR theorem [97] that rules out ψ-epistemic ontological models of quantum theory.

It is already well known that there can be exponential separation between classical and quantum communication complexities – some of the key results have been reviewed in Sec. 6.6. For example, the promise-equality problem studied in Ref. [106], as reviewed in Ex. 7, presents an $\Omega(d)$ vs $O(\log d)$ separation between zero-error classical and quantum communication complexities for the task of deciding whether two d-bit inputs are either equal or have Hamming distance $d/2$. We mentioned earlier that the proof involves a highly non-trivial combinatorial result of Ref. [107], and results in a classical lower bound of the form cd, for a constant $c = 0.01$. Slightly strengthened, the same proof yields a lower bound of cd for any $c < 0.02$, which means that a non-trivial separation between the number of classical bits that must be communicated and the number of qubits can be established only for $d \geq 512$.

Subsequent works [109, 139, 104, 140, 141, 111, 111, 112, 142] established exponential separations between zero-error quantum and bounded-error classical protocols in various settings. The separation of Ref. [112] in particular is strong in the sense that the separation holds between zero-error one-way quantum protocols and bounded-error classical protocols that allow two-way communication. Montanaro [143], by considering a task based on distributed Fourier sampling, derives an $\Omega(d)$ vs $O(\log d)$ separation which is robust against constant additive error, and also holds when two-way classical communication is allowed.

In Ref. [127], Montina studied a scenario in which Alice's input can be any pure state of a d dimensional quantum system, and Bob's input can be a classical description of any two-outcome measurement consisting of a rank-1 projector and the orthogonal projector. For the case of exact, one-way classical communication, Montina derives a classical lower bound of cd, for $c \approx 0.293$, where the proof uses a result concerning volumes of subsets of a hypersphere due to Raigorodskii [144]. It is also shown that a classical lower bound of $d - 1$ follows from (a complex generalization of) a mathematical conjecture known as the *double cap conjecture*.

Finally, other works have had slightly different aims from that of establishing quantum-classical separations in the standard communication complexity setting, but nonetheless contain techniques related to those that we use. These include Ref. [110], which presents quantum fingerprinting protocols that show an exponential separation between bounded-error quantum and classical simultaneous message passing with pri-

vate randomness. Refs. [121, 122] then use the notion of antidistinguishability to give separations between one-way communication and information complexity in exclusion games – this shows that the parties may need to communicate very large number of bits to exactly solve the exclusion task, while an adversary can learn almost nothing from the communication transcript. They also include Ref. [98], where lower bounds on the size of a classical memory needed to simulate quantum processes are derived and applied to the stabilizer subtheory of quantum theory. Ref. [136], that appeared while we were finishing this work, defines and studies tasks of communicating 'partial ignorance', including communication tasks using antidistinguishable quantum states that are similar to those used here.[1]

One of the main motivations for our work is to present a proof of exponential separation between classical and quantum communication complexity that is very simple and essentially self-contained. Aside from this, the advantages of the approach include (i) a lower bound for zero-error one way classical communication of cd, with $c \approx 0.415$, which is the strongest that we have seen, and (ii) a separation between quantum and classical one-way communication complexity for any $d \geq 3$. On the negative side, we show that the classical communication lower bound disappears if two-way classical communication is allowed. Although the result is robust against a limited amount of additive noise, the lower bound also disappears if bounded error classical protocols are allowed.

8.2 Antidistinguishability and Spherical codes

Recall the CFS lemma proved in Sec 7.3:

Lemma 1 (Three-state antidistinguishability [96])**.** *Any three pure states are antidisitnguishable if:*

$$a = |\langle \rho_i | \rho_j \rangle|^2, \qquad b = |\langle \rho_i | \rho_k \rangle|^2, \qquad c = |\langle \rho_k | \rho_j \rangle|^2, \tag{8.1}$$

we have that:

$$a + b + c < 1, \qquad \textit{and} \qquad (1 - a - b - c)^2 \geq 4abc. \tag{8.2}$$

[1] We provide additional details about this result later.

In what follows, we use the conditions of Eq. 8.2, but weaken them to a simpler sufficient condition in the following corollary:

Corollary 1. *Let* $S = \{|\rho_i\rangle, |\rho_j\rangle, |\rho_k\rangle\}$ *be a triple of states and let:*

$$a = |\langle\rho_i|\rho_j\rangle|^2, \qquad b = |\langle\rho_i|\rho_k\rangle|^2, \qquad c = |\langle\rho_k|\rho_j\rangle|^2. \tag{8.3}$$

The triple S *is antidistinguishable if:*

$$a, b, c \leq \frac{1}{4}. \tag{8.4}$$

Proof. Since $a, b, c \leq 1/4$, we have that $a + b + c \leq 3/4$. It immediately follows that $a + b + c < 1$ in Eq. 7.22. For $0 \leq a, b, c \leq 1/4$, the condition in Eq. 7.23 reduces to:

$$1 - a - b - c \geq 2\sqrt{abc}. \tag{8.5}$$

The left hand side is monotonically decreasing, attaining a minimum of $1/4$ for $a = b = c = 1/4$. The right hand side is monotonically increasing, attaining a maximum of $1/4$ for $a = b = c = 1/4$. It follows that $0 \leq a, b, c \leq 1/4$ suffices for three state antidistinguishability. □

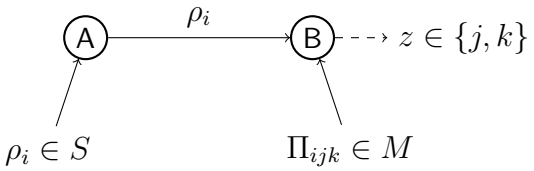
Now consider a finite set S of pure states $|\rho_1\rangle, |\rho_2\rangle, \ldots, |\rho_{|S|}\rangle$, for which:

$$|\langle\rho_i|\rho_j\rangle| \leq \delta; \ \forall\, i \neq j,$$

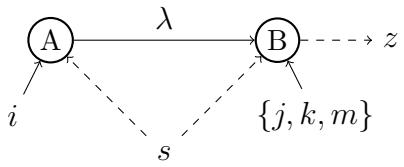
where $\delta \in [0, 1)$. Such sets are called *complex spherical codes*, and have been much studied, with applications in frame theory, classical signal processing, error correction, and quantum information [145, 146, 147, 110, 148, 149, 150, 151]. Our results will be obtained from the following observation:

Claim 20. *Any triple of states drawn from a complex spherical code* S *with* $\delta \leq 1/2$ *is antidistinguishable.*

We have from Corollary 1 that any triple of states drawn from a finite set S of pure quantum states, such that all mod-squared overlaps are upper-bounded by $1/4$ is antidistinguishable. Any triple of states drawn from a complex spherical code S with $\delta \leq 1/2$ is hence antidistinguishable.



(a) One-way quantum communication protocol for the task. After receiving $i \in [n]$, Alice prepares a state $|\rho_i\rangle \in S$. Given $\{i,j,k\}$ as an input, Bob prepares a projective measurement $\Pi \in M$ defined by the three states $\{\rho_i, \rho_j, \rho_k\}$. Alice sends ρ_i to Bob, after which he performs the measurement that assigns nonzero probability to outcomes j or k.



(b) A classical protocol. Alice gets an integer $i \in [|S|]$, and Bob gets three integers $j,k,m \in [|S|]$. Alice sends a message λ, after which Bob outputs $z \in \{j,k,m\}$. The classical strategy can use a shared random string s, drawn according to a probability distribution $P(s)$.

Figure 8.1: Quantum (a) and classical (b) protocols for the antidistinguishability relation R given in Eq. 8.6.

8.3 The Communication Task

The separation between classical and quantum communication complexities that we establish is for solution of a relational task, defined as follows: For a d dimensional Hilbert space, let S be a complex spherical code $S = \{|\rho_1\rangle, |\rho_2\rangle, \ldots |\rho_{|S|}\rangle\}$, with $\delta = 1/2$. Alice's input is an integer $i \in [|S|]$. Bob's input is a set of 3 integers $j,k,m \in [|S|]$. Bob must output one of the integers j,k,m, under the constraint that his output must not be equal to Alice's input. Setting $X = Z = [|S|]$, and $Y = \{\, T \,|\, T \subseteq S, |T| = 3 \,\}$, the relation is thus given by:

$$\begin{aligned} R &\subseteq X \times Y \times Z : \\ (i, \{j,k,m\}, z) &\in R \text{ iff } z \in \{j,k,m\} \text{ and } z \neq i. \end{aligned} \tag{8.6}$$

The quantum solution is straightforward (see Fig. 8.1a).

1. Given input i, Alice prepares a quantum system in the state $|\rho_i\rangle \in S$ and sends it to Bob.

2. Given input $\{j,k,m\}$, Bob performs an antidistinguishing measurement for the three states $|\rho_j\rangle, |\rho_k\rangle, |\rho_m\rangle$ on the system he receives from Alice. Label the three outcomes Π'_z, for $z \in \{j,k,m\}$, such that $\Pi'_z |\rho_z\rangle = 0$. If Bob obtains outcome Π'_z, then he outputs z. (This means that Bob never outputs an outcome 'i'.)

The communication complexity is the number of qubits transmitted, which is $\lceil \log d \rceil$.

A classical protocol for the task is illustrated in Fig. 8.1b. In the next section, we prove a lower bound on the number of bits that needs to be transmitted on average in a randomized protocol to solve it with certainty.

8.4 One-way Communication Complexity

We now prove a lower bound on the one-way communication complexity of the task in terms of the size of the complex spherical code S.

Claim 21. *The zero-error one-way classical communication complexity of the task is at least* $\lceil \log |S| - 1 \rceil$.

Proof. In order to establish the claim, consider first deterministic protocols, in which the message that Alice sends to Bob is a function of her input, and Bob's output is a function of his input and the message. Suppose that there are three distinct values of Alice's input, $|\rho_j\rangle, |\rho_k\rangle, |\rho_m\rangle$, such that the same message λ is sent for each of them. If Bob receives λ, and his input is the triple $\{j, k, m\}$, then there is no output he can give that will not sometimes generate an error. Therefore Alice needs at least a distinct message per two states of S. This gives:

$$|\Lambda| \geq \frac{|S|}{2}, \tag{8.7}$$

where Λ is the set of possible values of Alice's message. It follows that on the worst case input, Alice needs to send at least $\lceil \log |S| - 1 \rceil$ bits. In the presence of shared randomness s, each value of s defines a deterministic protocol which in the zero-error case must respect the relation. If communication complexity is evaluated as the communication cost on the worst case values of s (which corresponds to the deterministic communication complexity as defined in Def 17), this concludes the proof. With communication complexity given by the communication cost averaged over s, and evaluated on the worst case input (which corresponds to zero-error randomized complexity R_0 defined in Def. 20), we use Yao's principle (Eq. 6). For any probability distribution Q over input pairs, the randomized zero-error communication complexity is lower bounded by the communication cost averaged both over values of s and over inputs drawn according to the distribution Q. This is in turn lower bounded

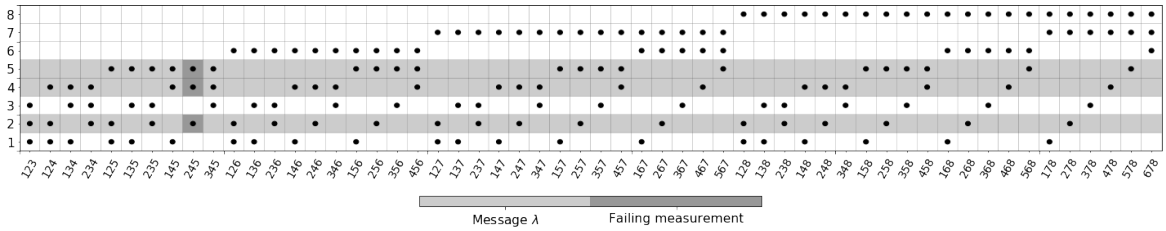


Figure 8.2: Any classical deterministic one-way protocol that uses a message λ for three or more Alice's inputs fails to reproduce the output of a quantum protocol on at least one input pair. In the above, Alice's input is one of the states $2, 4$ or 5 and Bob's input is 245. Alice chooses to send a message corresponding to the subset $2, 4, 5$. Bob cannot output any answer without risking an error.

by the communication cost averaged over inputs drawn according to the distribution Q, of the deterministic protocol that achieves the lowest value for this cost – this is the distributional complexity $D_0^Q(R)$ defined in Def 21. Choosing Q as the uniform distribution over all input pairs, the argument above establishes that the lower bound of $\lceil \log |S| - 1 \rceil$ bits still holds. □

We briefly remark that the lower bound derived in Claim 21 also straightforwardly applies to a (seemingly) less constrained problem of multiplicative error sampling from distributions generated by the quantum experiment. Let S be the set of states defined in Claim 20 of the main text and let $\Pi' = \{\Pi'_z |\, z \in [n]\}$ be a measurement antidistinguishing a triple of states $\{|\rho_j\rangle, |\rho_k\rangle, |\rho_m\rangle\} \subseteq S$, chosen such that $\Pi'_z |\rho_z\rangle = 0$ for all $z \in \{j, k, m\}$. Let $p(z|\, \Pi', |\rho_i\rangle)$ denote the probability of measuring an outcome 'z' by applying Π' to $|\rho_i\rangle$. A ϵ-multiplicative sampling protocol samples from a distribution $\tilde{p}(z|\, \Pi', |\rho_i\rangle)$, such that:

$$|\tilde{p}(z|\, \Pi', |\rho_i\rangle) - p(z|\, \Pi', |\rho_i\rangle)| \leq \epsilon\, p(z|\, \Pi', |\rho_i\rangle), \tag{8.8}$$

for some $\epsilon \geq 0$ and all inputs. Notice that:

$$p(z|\, \Pi', |\rho_i\rangle) = 0 \implies \tilde{p}(z|\, \Pi', |\rho_i\rangle) = 0, \tag{8.9}$$

holds for arbitrary ϵ. Even as $\epsilon \to \infty$, the mutliplicative error simulation protocol cannot output an outcome that is assigned zero probability and hence also solves the relation R defined in Eq. 8.6 of the main text. This means that classical protocols for the sampling task are subject to the same classical lower bounds as the zero-error protocol for R.

The lower bound of Claim 21 is determined by $|S|$ – the size of the spherical code S. In contrast, for a fixed dimension d regardless of the size of $|S|$, the quantum protocol uses only $\lceil \log d \rceil$ qubits. This gives a communication advantage whenever $|S| > 2d$. For the best separation, the problem then becomes: how large can a complex spherical code in d dimensions be, with $\delta = 1/2$?

8.5 Complexity Separations in Small Dimensions

In $d = 3$, the largest such set S is given by an equiangular complex spherical code, otherwise known as a symmetric, informationally-complete set (SIC) [152, 153, 154, 155]. A SIC in dimension 3 consists of 9 unit vectors such that:

$$|\langle \rho_i | \rho_j \rangle|^2 = \frac{1}{4}, \ \forall\, i \neq j. \tag{8.10}$$

That a larger set cannot be found follows from the Welch bound [145], which states that:

$$\max_{i \neq j} |\langle \rho_i | \rho_j \rangle|^2 \geq \frac{|S| - d}{d(|S| - 1)}, \tag{8.11}$$

for any set of d-dimensional pure states $S = \{|\rho_1\rangle, |\rho_2\rangle, \ldots |\rho_{|S|}\rangle\}$. This shows that Alice needs to send at least a 5-valued message vs a 3-dimensional quantum system, or in terms of whole numbers of bits and qubits, at least 3 bits vs. 2 qubits.

In $d \geq 4$, mutually unbiased bases (MUBs) yield larger sets than SICs. It is known that for d power prime, there exist $d + 1$ distinct MUBs [156], which satisfy $|\langle \rho_i | \rho_j \rangle|^2 \leq 1/d$. Taking S as the union of the vectors in the MUBs gives $|S| = d(d+1)$, hence a $\lceil \log(d^2 + d) - 1 \rceil$ lower bound on classical communication. In $d = 4$, Alice needs to send at least a 10-valued message vs. a 4-dimensional quantum system, or in terms of whole numbers of bits and qubits, at least 4 bits vs. 2 qubits. A bound due to Levenshtein [146, 147, 150] implies that in $d = 4$, the 20-element set of vectors given by MUBs is the largest that can be achieved with $\delta = 1/2$. In $d \geq 5$, larger sets than those given by MUBs have been found numerically in Refs. [150, 151].

8.6 Asymptotic Separations

We now present two arguments that give asymptotically exponential separations between the classical zero-error randomized communication complexity and quantum

complexity. The first one is based on packing bound – specifically the argument about packing hyperspherical caps attributed independently to Shannon, Wyner and Chabauty [157, 158, 159, 160]. It gives a good lower bound, but relies on volume integration of complex hyperspherical caps which may be seen as difficult. The second method, which gives a somewhat weaker lower bound, is based on a probabilistic method argument for Hadamard codes that we borrowed from the quantum fingerprinting paper (Ref. [110]) and Alon and Spencer's book on the probabilistic method [161].

8.6.1 Lower Bound by Packing

We first establish the bound by adapting a spherical packing lower-bound of Chabauty, Shannon and Wyner [157, 158, 159, 160]. The original bound is derived by associating to every unit vector $\vec{e} \in \mathbb{R}^d$ a spherical δ-cap, defined as a set of vectors $\vec{g} \in \mathbb{R}^d$ such that $\vec{e} \cdot \vec{g} > \delta$ for some constant $\delta \in (0, 1)$. Suppose that we found the optimal packing configuration of unit vectors such that no two vectors have overlap greater than δ. Then the union of their associated spherical δ-caps fully covers the unit hypersphere – we could add another vector to the set otherwise. Any pair of vectors has overlap at least δ in this configuration, and the number of vectors in the optimal packing configuration can be thus lower-bounded by the ratio unit sphere surface to the surface of a spherical δ-cap [160]. We apply a modification of this technique to quantum states.

Recall first that pure quantum states in d-dimensions can be treated as points in complex projective space $\mathbb{C}P^{d-1}$ [162]. Following Refs. [163, 162], one can use the Hurwitz parametrization of $\mathbb{C}P^{d-1}$ to describe these points as:

$$\begin{aligned} |\psi\rangle = (&\cos\theta_{d-1}, \\ &\sin\theta_{d-1}\cos\theta_{d-2}\, e^{i\phi_{d-1}}, \\ &\sin\theta_{d-1}\sin\theta_{d-2}\cos\theta_{d-3}\, e^{i\phi_{d-2}}, \\ &\ldots, \\ &\sin\theta_{d-1}\ldots\sin\theta_2\cos\theta_1\, e^{i\phi_2}, \\ &\sin\theta_{d-1}\ldots\sin\theta_2\sin\theta_1\, e^{i\phi_1}), \end{aligned} \tag{8.12}$$

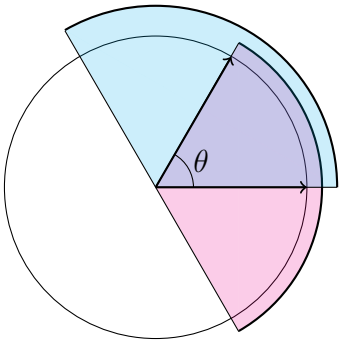


Figure 8.3: The closest possible packing of two spherical $\cos\theta$ caps associated to two unit vectors (bold arrows) in $d = 2$. The packing bound uses the fact that in an optimal configurations, the union of caps fully cover the unit sphere – in this case, it is the arcs covering the circle in full. A reciprocal of the optimal cap volume then gives the lower bound on the number of caps needed to cover the sphere and thus the number of vectors in the largest spherical code in that dimension. The argument extends trivially to double-cap packings.

where $\theta_k \in [0, \pi/2]$ and $\phi_k \in [0, 2\pi)$ for $k \in [d]$. The volume element is given by:

$$d\psi = \prod_{k=1}^{d-1} \cos\theta_k (\sin\theta_k)^{2k-1} d\theta_k d\phi_k. \tag{8.13}$$

Consider the ratio:

$$\frac{\int d\psi}{\int I[|\phi\rangle] d\phi}, \tag{8.14}$$

where the integration runs uniformly over all pure d-dimensional quantum states and:

$$I(|\phi\rangle) = \begin{cases} 1 & \text{if } |\langle\phi|0\rangle| \geq \delta, \\ 0 & \text{otherwise }, \end{cases} \tag{8.15}$$

for some (arbitrary) fixed reference state $|0\rangle$. In analogy to the Chabauty, Shannon and Wyner argument, the size of the largest set of d-dimensional pure quantum states:

$$S = \{|\psi\rangle \in \mathbb{C}P^{d-1} \mid |\psi\rangle, |\phi\rangle \in S : |\psi\rangle \neq |\phi\rangle \implies |\langle\psi|\phi\rangle| \geq \delta\}, \tag{8.16}$$

becomes lower bounded by:

$$|S| \geq \frac{\int d\psi}{\int I[|\phi\rangle] d\phi}. \tag{8.17}$$

Claim 22. *For any d, there exists a set S with $\delta = 1/2$, such that $|S| \geq (4/3)^{d-1}$.*

Proof. Choose $|0\rangle = (1, 0 \ldots)$ as the reference state, which leads to $|\langle 0|\psi\rangle| = |\cos\theta_{d-1}| \leq \delta$ in the Hurwitz parametrization. For the permitted range of angles, this means that we integrate over $\theta_{d-1} \in [0, \pi/3]$ when evaluating $\int I[|\psi\rangle]$. Using the volume element of $\mathbb{C}P^{d-1}$ from Eq. 8.13 gives:

$$\begin{aligned}\int I[|\psi\rangle]d\psi &= \prod_{k=1}^{d-1}\int_0^{2\pi} d\phi_k \left(\prod_{k=1}^{d-2}\int_0^{\pi/2}\cos\theta_k(\sin\theta_k)^{2k-1}d\theta_k\right)\int_0^{\pi/3}\cos\theta_{d-1}(\sin\theta_{d-1})^{2d-3}d\theta_{d-1} \\ &= \frac{\pi^{d-1}}{(d-1)!}\left(\frac{3}{4}\right)^{d-1}.\end{aligned} \tag{8.18}$$

Integrating leads to:

$$\int d\psi = \frac{\pi^{d-1}}{(d-1)!}, \tag{8.19}$$

from which it follows that $|S| \geq (4/3)^{d-1}$. □

We remark that the double-cap volume integral appears in Ref. [127] as a part of a different method for establishing another communication complexity separation.

8.6.2 By Hadamard States

An alternative proof that there exist sets S with $|S|$ exponentially large, which results in a worse bound, but some may find even simpler, comes from Refs. [110, 161] and is reproduced here. It shows that complex spherical codes exist that are exponentially large in the dimension. Take two random d-dimensional pure states:

$$|v\rangle = \frac{1}{\sqrt{d}}\sum_{i=1}^{d} v_i |i\rangle, \quad |w\rangle = \frac{1}{\sqrt{d}}\sum_{i=1}^{d} w_i |i\rangle, \tag{8.20}$$

where w_i, v_i are Rademacher random variables with $\Pr(v_i = \pm 1) = \Pr(w_i = \pm 1) = 1/2$. Their inner product is:

$$\langle v|w\rangle = \frac{1}{d}\sum_{i=1}^{d} v_i w_i = \frac{1}{d}\sum_{i=1}^{d} X_i, \tag{8.21}$$

where X_i is again a Rademacher random variable with $\Pr(X_i = \pm 1) = 1/2$. The probability that $|\langle v|w\rangle| > 1/2$ is upper bounded by the Chernoff-Hoeffding inequality:

$$\Pr\left(|\langle v|w\rangle| > \frac{1}{2}\right) = \Pr\left(\left|\sum_i^d X_i\right| > \frac{d}{2}\right) \leq 2e^{-\frac{d}{8}}. \tag{8.22}$$

The probability that a set S of such random vectors contains a pair $|v_i\rangle, |w_j\rangle \in S,\ i \neq j$ with overlap greater than $1/2$ is given by:

$$\Pr\left(|\langle v_i|w_j\rangle| > \frac{1}{2}\right) \leq 2\binom{|S|}{2} e^{-\frac{d}{8}} < |S|^2 e^{-\frac{d}{8}}. \tag{8.23}$$

As soon as this probability falls below 1, there exists a set S of such random states, so that $|\langle v_i|w_j\rangle| \leq 1/2$ for any pair. From Eq. (8.23), $|S| = e^{\frac{d}{16}}$. Using such a set S for the task defined in the main text, the classical one-way communication complexity is at least $\lceil 0.09\, d - 1 \rceil$ bits.

8.7 Two-way Communication

We study two-way communication complexity of the task and give a deterministic communication protocol with $O(\log\log|S|)$ communication complexity. This shows that the separation does not hold if we allow the two parties to interact.

Claim 23. *Assuming $|S| = 2^q$, for integer q, the two-way classical communication complexity of R is at most $\lceil \log\log|S| \rceil + 1$ bits.*

Proof. The assumption that $|S| = 2^q$ is not essential, but allows a short statement of the proof. The protocol has two rounds and starts with a message from Bob to Alice. Let Bob's input be $\{j, k, m\}$, and assume without loss of generality that $j < k < m$. Bob determines the largest integer $r \leq q = \log|S|$, such that for some integer $s \geq 0$ either:

$$\begin{gathered} s\,2^r < j, k \leq (s+1)\,2^r \text{ and} \\ (s+1)\,2^r < m \leq (s+2)\,2^r, \end{gathered} \tag{8.24}$$

or

$$\begin{gathered} s\,2^r < j \leq (s+1)\,2^r \text{ and} \\ (s+1)2^r < k, m \leq (s+2)2^r. \end{gathered} \tag{8.25}$$

Bob sends r to Alice using $\lceil \log\log|S| \rceil$ bits. Note that r determines a subset $Y_r \subseteq Y$ of Bob's possible inputs. For an input $i \in [|S|]$, Alice computes the parity p of $\lceil \frac{i}{2^r} \rceil$, and sends it to Bob. Note that the set X of Alice's inputs is partitioned by p and r into subsets $X_{r,p} \subseteq X$. By Eqs. 8.24 and 8.25, at least one of j, k, m is not in $X_{r,p}$. Bob chooses such a value for his output. Communicating $\lceil \log\log|S| \rceil + 1$ bits hence suffices in the two-way scenario. Fig. 8.4 illustrates the protocol for $|S| = 8$. □

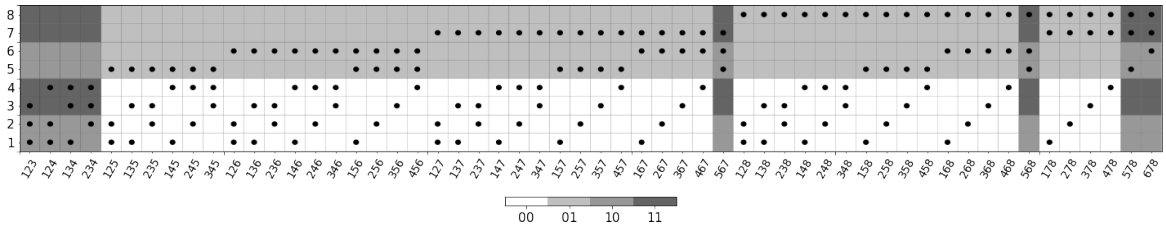


Figure 8.4: An illustration of the two-way classical protocol with $|S| = 8$. Rows correspond to different values of Alice's input, columns to different values of Bob's input. In single run of the protocol, Bob sends an integer r to Alice, and Alice sends an integer p to Bob. Each value of the conversation (r, p) is compatible with a subset of the joint inputs, where the subset is of the form $X_{r,p} \times Y_r$ for $X_{r,p} \subset X$ and $Y_r \subset Y$. In the $|S| = 8$ example, four distinct conversations suffice to partition the inputs in a way that each of Bob's inputs is split across two distinct rectangles - enough to allow him to solve the relation.

We present additional intuition for the above proof here. The general idea is that Alice and Bob can give a correct answer whenever they communicate enough information in order to separate Bob's inputs into two disjoint sets, in a way that Alice can tell Bob which set contains the 'dangerous' input. The key observation is that this can be done very efficiently if these subsets are combinatorial rectangles $X_A \times Y_B$ for $X_A \subseteq X$ and $Y_B \subseteq Y$, such that $|X_A| = |S|/2$.

Fix $|S| = 2^q$ for some integer $q > 1$ for simplicity. Start with an example in which we assume that Bob's input is such that $|S|/2 < m \leq |S|$ and $0 < j, k \leq |S|/2$. Bob chooses $r = \log[|S|/2] = q - 1$ and sends it to Alice using at most $\log q \leq \log\log|S|$ bits. This way, Alice knows about a 'horizontal cut' at $|S|/2$ that separates Bob's inputs, as in the 00, 01 rectangles of Fig. 8.4. Alice now needs to tell Bob what part of the cut is her input i contained in, for which we chose the cut in a way that allows this to be communicated very efficiently. In this case, she needs to tell Bob if $0 < i \leq |S|/2$ or if $|S|/2 < i \leq |S|$ by sending a single bit. By symmetry, the above also deals with inputs where $0 < j \leq |S|/2$, $|S|/2 < k, m \leq |S|$.

As an example of a more complicated case, suppose now that either:

$$0 < j, k, m \leq |S|/2, \qquad 0 < j \leq |S|/4, \qquad |S|/4 < k, m \leq |S|/2, \tag{8.26}$$

or that:

$$|S|/2 < j, k, m \leq |S|, \qquad |S|/2 < j \leq 3|S|/4, \qquad 3|S|/4 < k, m \leq |S|. \tag{8.27}$$

Notice that both of the sets can be separated by a pair of 'horizontal cuts' at $|S|/4$ and $3|S|/4$. The trick that allows for the additional logarithmic compression is to

communicate this pair of cuts irrespectively of which of the above cases holds. Bob sends $r = \log[|S|/4] = q - 2$ using at most $\log\log|S|$ bits. Knowing that $r = |S|/4$, Alice can tell Bob if her input lies in the range $X_1 = \{1, \ldots |S|/4\} \cup \{|S|/2 + 1, \ldots 3|S|/4\}$ or in the complement of X_1 in X – this takes a single bit. By symmetry, the above protocol also deals with the case when either $0 < j, k \leq |S|/4$ and $|S|/4 < m \leq |S|/2$ or $|S|/2 < j, k \leq 3|S|/4$ and $3|S|/4 < m \leq |S|$.

The protocol runs similarly on the remaining inputs and it does not take much from there to formalize it as we did in Claim 23. It would be interesting to see if the protocol could be formulated in terms of the Fourier transform over the inputs, since the information that Bob communicates to Alice is essentially the frequency of horizontal cuts between two combinatorial rectangles that Bob must choose in order to place one of his elements into a distinct rectangle from the remaining two.

8.8 Bounded Error

Here we show that the separation disappears in one-way classical communication if the classical players need only solve the relation with bounded error. In such case, we demand from the ϵ-error protocol P for the task R that for all input pairs $(x, y) \in X \times Y$:

$$\Pr\left[P(x, y) \in R(x, y)\right] \geq 1 - \epsilon, \tag{8.28}$$

where the average is taken over the shared random data. As we assumed that $\epsilon > 0$, the randomized complexity of R is the minimum worst case cost (see Def. 19) of a randomized protocol that solves the relation.

A simple way to see that the lower bound is not noise resilient is to consider reduction to the bounded-error protocol for equality testing, described in Ex 6. Suppose that Alice's input is i and Bob's input is $\{j, k, l\}$. The parties repeat the following in several runs: On every run of the protocol, they generate a uniformly random string $s \in \{0, 1\}^{|S|}$. Alice computes the parity p_i of $\sum_{r=1}^{|S|} i_r s_r$ where i_r and s_r denote the r-th bit of her input i and the shared randomness s respectively. She sends p_i (one bit) to Bob. Bob computes the parities p_j, p_k, p_m of $\sum_{r=1}^{|S|} j_r s_r$, $\sum_{r=1}^{|S|} k_r s_r$, and $\sum_{r=1}^{|S|} m_r s_r$. Without loss of generality, suppose that $j = i$. Then $p_i = p_j$ with certainty, while $p_k, p_m = p_i$ with probability $1/2$ as we choose s to be uniform. Repeating

this $\lceil \log(1/\epsilon) \rceil$ times, the probability of Bob wrongly concluding that either $i = k$ or $i = m$ is at most ϵ. He therefore outputs one of k or m, making a mistake with less than ϵ probability. On the other hand, if i is not equal to any j, k, m, Bob cannot do a mistake by outputting any of these.

The low error sensitivity is a serious drawback of our results as it is tied to noise resilience of the lower bound – not even the quantum communication itself will be noiseless. I unfortunately did not manage to come up with any modification of the protocol that would allow for noise resilience in the asymptotic setting. On the other hand, the finite dimensional results retain *some* noise resilience and may be verifiable in a sufficiently precise experiment. It is also possible that the multiplicative sampling version of the exclusion task can be more noise robust – although, despite many attempts, I did not manage to prove this either.

8.9 Antidistinguishability Conjecture

In Ref. [164], Montina and Wolf studied exact classical simulation of a scenario in which Alice chooses an arbitrary quantum pure state and sends it to Bob, who performs an arbitrary two-outcome projective measurement. A lower bound of $\Omega(d \log d)$ bits for this communication task is shown to follow from two mathematical conjectures. Using essentially the same proof that we presented here, the following conjecture concerning antidistinguishability would also give a separation of $\log d$ qubits vs $\Omega(d \log d)$ bits for exact simulation. This would be asymptotically stronger than all existing results, and would be particularly interesting given the $O(\log(1/\epsilon)d)$ upper bound for bounded-error classical simulation. The argument we give is very different from those outlined in Ref. [125], and the conjecture is also very different from the two conjectures of Ref. [164].

Conjecture 2. *Let $|\rho_1\rangle, \ldots, |\rho_d\rangle$ be d pure states. If $|\langle \rho_i | \rho_j \rangle| \leq (d-2)/(d-1)$ for all $i \neq j$, then the states are antidistinguishable.*

We remark that the $\log d$ vs $\Omega(d \log d)$ separation is established by combining Conjecture 2 with essentially the same proof of zero-error one-way separation for the communication task that was described in the previous sections. The key difference is that the spherical code has $\delta = (d-2)/(d-1)$ and that Bob obtains instead a

measurement that is antidistinguishing on d states. The packing lower bound on $|S|$ then gives:

$$|S| \geq \frac{V^d}{V_\delta^d} = \frac{\int_0^1 x^{2d-3} dx}{\int_0^{\sqrt{1-(d-2)^2/(d-1)^2}} x^{2d-3} dx} = \left[\frac{2d-3}{(d-1)^2} \right]^{1-d}, \tag{8.29}$$

and the number of distinct messages used by Alice becomes lower bounded by:

$$|\Lambda| \geq \frac{|S|}{d-1} \geq \frac{1}{d-1} \left(\frac{2d-3}{(d-1)^2} \right)^{1-d}. \tag{8.30}$$

By collecting the orders of d, it follows that $\log |\Lambda| = \Omega(d \log d)$, which implies the classical exact communication complexity lower bound of $\Omega(d \log d)$.

We now discuss the motivations behind our conjecture. It is obvious that it holds with $d = 2$, since any two perfectly distinguishable states are also antidistinguishable. It also follows from the Caves-Fuchs-Schack lemma [96], reviewed in Sec. 7.3, that the conjecture holds for $d = 3$.

In order to gain some intuition for why the conjecture might be true in all dimensions, first consider a generic set of states $|\rho_1\rangle, \ldots, |\rho_d\rangle$ in d dimensions. Heinosaari and Kerppo [133] show that if any triple of the states is antidistinguishable, then it follows that the set of d states is antidistinguishable. On the other hand, it does not follow that if the set of d states is antidistinguishable, then any triple must be antidistinguishable. Hence, the condition that d states is antidistinguishable is logically weaker than the condition that any triple of them is antidistinguishable. If the states satisfy $|\langle \rho_i | \rho_j \rangle| \leq 1/2$ for all $i \neq j$, then the stronger condition holds [96]. It is hence natural to suppose that a similar statement, with $1/2$ on the right hand side replaced by a larger number, suffices for the weaker condition.

Second, consider the set of states:

$$\begin{aligned}
|\rho_1\rangle &= \frac{1}{\sqrt{d-1}} \left(|e_2\rangle + |e_3\rangle + \cdots + |e_d\rangle \right) \\
|\rho_2\rangle &= \frac{1}{\sqrt{d-1}} \left(|e_1\rangle + |e_3\rangle + \cdots + |e_d\rangle \right) \\
&\cdots \\
|\rho_d\rangle &= \frac{1}{\sqrt{d-1}} \left(|e_1\rangle + |e_2\rangle + \cdots + |e_{d-1}\rangle \right),
\end{aligned}$$

where $\{|e_i\rangle\}$ is an orthonormal basis. This set is antidistinguishable by construction and satisfies $|\langle \rho_i | \rho_j \rangle| = (d-2)/(d-1)$ for all $i \neq j$, which motivates the particular choice of function on the right hand side of the conjectured sufficient condition.

Finally, Fig. 8.5 displays numerical evidence for Conjecture 2. Note that for a given set of states in d dimensions, determining whether they are antidistinguishable or not corresponds to solving a semi-definite program (SDP) [131, 132]. For each $d = 2, 3, 4, 5, 6$, we generated two sets of vectors as follows:

1. 150000 sets of d vectors in d dimensions, where each vector is chosen independently, and uniformly according to the Haar measure and

2. a set of vectors:

$$\begin{aligned} |\rho_1\rangle &= \cos\theta\,|e_1\rangle + \sin\theta\,|e_2\rangle\,, \\ |\rho_2\rangle &= \cos\theta\,|e_1\rangle + \sin\theta\,|e_3\rangle\,, \\ &\cdots \\ |\rho_{d-1}\rangle &= \cos\theta\,|e_1\rangle + \sin\theta\,|e_d\rangle\,, \\ |\rho_d\rangle &= |e_1\rangle \end{aligned}$$

   for a set of 10000 angles $\theta \in [0, \frac{\pi}{2}]$. The reasoning behind this choice of states is that $\theta = 0$ corresponds to non-antidistinguishable states, while $\theta = \frac{\pi}{2}$ gives a basis that is trivially antidistinguishable.

The second set was used because non-antidistinguishable ensembles are very rare in $d \gtrsim 5$ — in fact, none of the 150000 uniformly random ensembles in $d = 6$ tested as non-antidistinguishable.

For every set of vectors, the SDP is solved in order to determine whether the set is antidistinguishable. For those sets that are not antidistinguishable, the quantity $\alpha = \max_{i \neq j} |\langle \rho_i | \rho_j \rangle|$ is recorded. The shaded region of the graph shows, for each dimension, the minimum value of α obtained over the union of the above two sets, and the dashed line shows the value of $(d-2)/(d-1)$, with lines rather than points used for clarity. A counterexample to the conjecture would appear as the dashed line crossing the non-shaded region. The evidence for the conjecture consists in the fact that the shaded region cleaves fairly closely to the dashed line, yet no counterexamples were found.

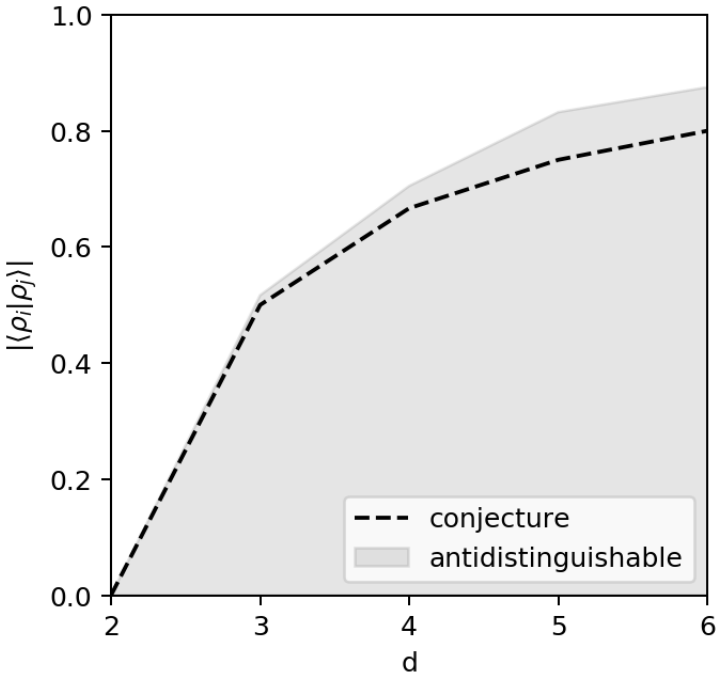


Figure 8.5: Numerical evidence for Conjecture 2. Line graph used for clarity.

Conclusion

Quantum communication complexity shows unconditional complexity separations between classical and quantum strategies. Proofs of these separations, even in the very simplest cases, however tend to rely on complicated theorems in combinatorics (Ref. [106]) or geometry (Ref. [109]) and are rarely self-contained.

Here we turn to quantum foundations and use the principle of quantum state antidistinguishability to give a very simple proof of separation between classical and quantum exact one-way communication complexities. Using connections to symmetric, informationally complete sets (SICs) and mutually unbiased bases (MUBs), we prove complexity separations in any dimension $d \geq 3$. Using a lower bound for the achievable size of a suitable complex spherical code, we showed an exponential separation of $\log d$ qubits vs. $0.415(d-1)-1$ bits using a spherical packing argument.

The results were stated in terms of quantum and classical players solving a relation defined on the finite sets of possible inputs for Alice and Bob. As expanding the sets of inputs can only make things more difficult for classical players, the lower bounds also apply to the one-way classical communication cost of simulating an experiment in which Alice prepares an arbitrary pure state of a d-dimensional quantum system, and sends it to Bob who performs an arbitrary projective measurement. (Naturally, the same can be said for the results of any of Refs. [106, 165, 109, 139, 104, 140, 141, 166, 112, 142, 143, 127].) In this more general scenario, if error ϵ is tolerated in the measurement outcome probabilities in the classical simulation, then it is easy to see that transmission of $O(\log(1/\epsilon)d)$ classical bits is sufficient. As discussed by Montanaro in Ref. [143], this means that the results of Ref. [143] are asymptotically optimal for bounded error simulation. If exact simulation is required, on the other hand, then Ref. [119] shows that without shared randomness, bounded classical communication becomes insufficient. If exact simulation is required, and shared random data

permitted, then surprisingly little is known apart from the exponential lower bounds. Ref. [120] shows that for $d = 2$, transmission of two bits is sufficient (see also Ref. [138] which considers positive operator-valued measurements). But we are unaware of any previous demonstration, even for $d = 3$, that classical simulation of arbitrary preparations and projective measurements is possible with bounded communication cost, let alone a demonstration of a specific protocol with bounded communication, or a finite upper bound [126, 137].

Exponential lower bounds on communication complexity are of interest for the foundations of quantum theory as well as for communication complexity per se. Paraphrasing Montanaro in Ref. [143]:

"On a fundamental, conceptual level, the question asks: are quantum states 'really' like an exponentially-long string of numbers, or do they have a more efficient representation?"

Restated in the language of quantum foundations, any lower bound on the size of the classical message in a simulation of the transmission of a quantum state becomes a lower bound on the size of the set of ontic states that the system must have available to it in an ontological model of quantum theory [167, 97, 137]. We thus believe that our results, owing particularly to their simplicity, could find further applications in quantum foundations. A specific direction in this respect is given by the works of Montina in Refs. [124, 126] and Montina and Wolf in Refs. [125, 164]. Their results imply that a proof of the antidistinguishability conjecture stated in our work would strengthen one of the seminal results in quantum foundations, the PBR theorem [97]. We have previously mentioned, somewhat in passing, that ontological models attempt to provide a realist description of quantum theory by associating to every physical system a space of ontic states Λ and to every quantum state preparation P a probability distribution $\mu : \Lambda \to [0, 1]$. Every measurement M is then associated a response function $\xi_M(k|\lambda)$, such that the probability of it giving the outcome k given a preparation P becomes:

$$p(k|\, M, P) = \int d\lambda\, \mu_P(\lambda)\xi_M(k|\lambda). \tag{8.31}$$

This way, the quantum state thus looks similar to a probability distribution that represents like a state of knowledge about the quantum system. The ontic distributions

in a theory could even overlap, in which case the ontic states in the overlaps provide incomplete information about the quantum state preparations. Ontological models with this feature are called ψ-epistemic and their complement are the ψ-ontic models.

The PBR theorem shows that ψ-epistemic theories are incompatible with predictions of quantum theory under the assumption of *preparation independence*. Preparation independence corresponds to the assumption that quantum states that are prepared independently have independently distributed ontic states – for example the state $|\psi\rangle\,|\phi\rangle$ would be described by an ontic state $\lambda = (\lambda_\psi, \lambda_\phi)$ drawn from the distribution $\mu_{\psi,\phi}(\lambda) = \mu_\psi(\lambda_\psi) \times \mu_\phi(\lambda_\phi)$. Soon after the publication of the theorem, Lewis, Jennings, Barrett and Rudolph gave a counterargument to it that did not assume the preparation independence in Ref. [168]. However, as described by Montina in Ref. [124], their model has several undesirable properties. For example, some ontic distributions in the model still contain full information about the corresponding quantum state and the model also reduces to an ψ-ontic model in the limit of infinite number of qubits.

In Ref. [126], Montina shows a close relationship between ψ-epistemic theories and communication complexity which goes beyond what was outlined earlier in Sec. 6.7. He specifically proved that a completely ψ-epistemic theory exists if and only if the quantum communication can be simulated exactly by a classical protocol with finite amount of classical communication. Such protocol was outlined for qubits in Ref. [120] (the corresponding ψ-epistemic model seems close to the Kochen-Specker model of a qubit), but generalizations to higher dimension are not known. Montina then reasons that, even if the extension of ψ-epistemic models could be possible beyond qubits, it could happen that the ψ-epistemic theory would not differ much from a ψ-ontic one. For example, similarly to the construction of Ref. [168], the overlaps of the ontological distributions could vanish rapidly as the number of qubits is increased, which would cause the ψ-epistemic model to reduce to the ψ-ontic one in the limit of infinite number of qubits. Refs. [125, 164] prove that this is indeed the case whenever the minimal communication cost of the corresponding communication protocol grows more than exponentially with the number of qubits. The proof is given under the assumption called *probability equipartition property*, which is a logically weaker condition than preparation independence. As mentioned earlier in Sec. 8.9, Ref. [125] then gives two

mathematical conjectures that would imply, similarly to our conjecture, a $\Omega(d \log d)$ communication lower bound. Their conjectures are, interestingly, very different from the antidistinguishability conjecture that we proposed here.

Montina's reasoning thus implies that a proof of our conjecture would weaken the assumption of preparation independence in the proof of the PBR theorem to the probability equipartition property and thus strengthen the result. Additionally, Montina claims in Ref. [124] that a proof of such lower bound implies that a ψ-epistemic theory does not provide a descriptional advantage over ψ-ontic theories, even without the 'equipartition property' assumption. Such proof could be a significant contribution to the ontological models approach to quantum theory.

Bibliography

[1] M. J. Bremner, A. Montanaro, and D. J. Shepherd. "Average-Case Complexity Versus Approximate Simulation of Commuting Quantum Computations". In: *Phys. Rev. Lett.* 117.8 (2016), p. 080501.

[2] R. L. Mann and M. J. Bremner. "On the Complexity of Random Quantum Computations and the Jones Polynomial". In: *arXiv e-prints* (2017). arXiv: 1711.00686.

[3] S. P. Jordan. "Permutational Quantum Computing". In: *Quantum Info. Comput.* 10.5 (2010), pp. 470–497.

[4] A. Marzuoli and M. Rasetti. "Spin Network Quantum Simulator". In: *Physics Letters A* 306.2 (2002), pp. 79–87.

[5] A. Marzuoli and M. Rasetti. "Computing Spin Networks". In: *Annals of Physics* 318 (2005), pp. 345–407.

[6] M. Van Den Nest. "Simulating Quantum Computers with Probabilistic Methods". In: *Quantum Info. Comput.* 11.9–10 (2011), pp. 784–812.

[7] D. Aharonov and I. Arad. "The BQP-hardness of Approximating the Jones Polynomial". In: *New Journal of Physics* 13.3 (2011), p. 035019.

[8] M. Schwarz and M. Van den Nest. "Simulating Quantum Circuits with Sparse Output Distributions". In: *arXiv e-prints* (2013). arXiv: 1310.6749.

[9] V. Havlíček et al. "Supervised Learning with Quantum-Enhanced Feature Spaces". In: *Nature* 567.7747 (2019), pp. 209–212.

[10] V. Havlíček, M. Troyer, and J. D. Whitfield. "Operator Locality in the Quantum Simulation of Fermionic Models". In: *Phys. Rev. A* 95.3 (2017), p. 032332.

[11] J. R. McClean et al. "OpenFermion: The Electronic Structure Package for Quantum Computers". In: *arXiv e-prints* (2017). arXiv: 1710.07629.

[12] S. Bravyi, D. Gosset, and R. König. "Quantum Advantage with Shallow Circuits". In: *Science* 362.6412 (2018), pp. 308–311.

[13] A. Montanaro. "Quantum Algorithms: An Overview". In: *npj Quantum Mechanics* 2 (2016), p. 15023.

[14] D. Aggarwal and U. Maurer. "Breaking RSA Generically Is Equivalent to Factoring". In: *Advances in Cryptology - EUROCRYPT 2009*. Ed. by Antoine Joux. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 36–53.

[15] C. Gidney and M. Ekerå. "How to Factor 2048 bit RSA Integers in 8 Hours using 20 Million Noisy Qubits". In: *arXiv e-prints* (2019). arXiv: 1905.09749.

[16] M. J. Bremner, R. Jozsa, and D. J. Shepherd. "Classical Simulation of Commuting Quantum Computations Implies Collapse of the Polynomial Hierarchy". In: *Proceedings of the Royal Society of London Series A* 467 (2011), pp. 459–472.

[17] S. Aaronson. "Limits on Efficient Computation in the Physical World". PhD thesis. PhD Thesis, 2004.

[18] B. M. Terhal and D. P. DiVincenzo. "Adaptive Quantum Computation, Constant Depth Quantum Circuits and Arthur-Merlin Games". In: *Quant. Inf. Comp.* 4.2 (2004), pp. 134–145.

[19] S. Aaronson and A. Arkhipov. "The Computational Complexity of Linear Optics". In: *Proceedings of the Forty-Third Annual ACM Symposium on Theory of Computing*. STOC '11. San Jose, California, USA: Association for Computing Machinery, 2011, pp. 333–342.

[20] S. Aaronson. "BQP and the Polynomial Hierarchy". In: *Proceedings of the Forty-Second ACM Symposium on Theory of Computing*. STOC '10. Cambridge, Massachusetts, USA: Association for Computing Machinery, 2010, pp. 141–150.

[21] D. Shepherd and M. J. Bremner. "Temporally Unstructured Quantum Computation". In: *Proceedings of the Royal Society of London Series A* 465.2105 (2009), pp. 1413–1439.

[22] A. W. Harrow and A. Montanaro. "Quantum Computational Supremacy". In: *Nature* 549.7671 (2017), pp. 203–209.

[23] F. Arute et al. "Quantum Supremacy Using a Programmable Superconducting Processor". In: *Nature* 574.7779 (2019), pp. 505–510.

[24] S. P. Jordan. "Fast Quantum Algorithms for Approximating some Irreducible Representations of Groups". In: *arXiv e-prints* (2008). arXiv: 0811.0562.

[25] S. P. Jordan. *Quantum Algorithm Zoo*. https://quantumalgorithmzoo.org/.

[26] E. Knill and R. Laflamme. "Power of One Bit of Quantum Information". In: *Phys. Rev. Lett.* 81 (1998), pp. 5672–5675.

[27] D. Poulin et al. "Testing Integrability With a Single Bit of Quantum Information". In: *Phys. Rev. A* 68.2 (2003), p. 022302.

[28] D. Poulin et al. "Exponential Speedup with a Single Bit of Quantum Information: Measuring the Average Fidelity Decay". In: *Phys. Rev. Lett.* 92.17 (2004), p. 177906.

[29] S. Aaronson et al. "The Computational Complexity of Ball Permutations". In: *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*. STOC 2017. Montreal, Canada: Association for Computing Machinery, 2017, pp. 317–327.

[30] V. Havlíček and S. Strelchuk. "Quantum Schur Sampling Circuits can be Strongly Simulated". In: *Phys. Rev. Lett.* 121.6 (2018), p. 060505.

[31] V. Havlíček, S. Strelchuk, and K. Temme. "Classical Algorithm for Quantum SU(2) Schur Sampling". In: *Phys. Rev. A* 99.6 (2019), p. 062336.

[32] C. Nayak et al. "Non-Abelian Anyons and Topological Quantum Computation". In: *Rev. Mod. Phys.* 80 (2008), pp. 1083–1159.

[33] J. M. Leinaas and J. Myrheim. "On The Theory of Identical Particles". In: *Il Nuovo Cimento B (1971-1996)* 37.1 (1977), pp. 1–23.

[34] F. Wilczek. "Quantum Mechanics of Fractional-Spin Particles". In: *Phys. Rev. Lett.* 49.14 (1982), pp. 957–959.

[35] F. Wilczek. "Anyons for Anyone". In: *Physics World* 4.1 (1991), p. 40.

[36] F. Wilczek. "New Kinds of Quantum Statistics". In: *The Spin: Poincaré Seminar 2007*. Ed. by B. Duplantier, J.-M. Raimond, and V. Rivasseau. 2009, p. 61.

[37] F. Wilczek. "New Kinds of Quantum Statistics". In: *The Spin: Poincaré Seminar 2007*. Ed. by B. Duplantier, J.-M. Raimond, and V. Rivasseau. 2009, p. 61. arXiv: 0812.5097.

[38] A. Roy and D. P. DiVincenzo. "Topological Quantum Computing". In: *arXiv e-prints* (2017). arXiv: 1701.05052.

[39] R. B. Laughlin. "Anomalous Quantum Hall Effect: An Incompressible Quantum Fluid with Fractionally Charged Excitations". In: *Phys. Rev. Lett.* 50.18 (1983), pp. 1395–1398.

[40] D. C. Tsui, H. L. Stormer, and A. C. Gossard. "Two-Dimensional Magnetotransport in the Extreme Quantum Limit". In: *Phys. Rev. Lett.* 48.22 (1982), pp. 1559–1562.

[41] J. Fröhlich. "Statistics of Fields, the Yang-Baxter Equation, and the Theory of Knots and Links". In: *Adv. Ser. Math. Phys.* 15 (1987), 575–604. 30 p.

[42] S. Simon. *Lecture Notes on Topological Quantum Physics*. Unpublished, 2016.

[43] M. H. Freedman, A. Kitaev, and Z. Wang. "Simulation of Topological Field Theories by Quantum Computers". In: *Communications in Mathematical Physics* 227 (2002), pp. 587–603.

[44] A. Stern. "Non-Abelian States of Matter". In: *Nature* 464.7286 (2010), pp. 187–193.

[45] G. Moore and N. Read. "Nonabelions in the Fractional Quantum Hall Effect". In: *Nuclear Physics B* 360.2 (1991), pp. 362–396.

[46] A. Y. Kitaev. "Quantum Computing: Unpaired Majorana Fermions in Quantum Wires". In: *Physics Uspekhi* 44 (2001), p. 131.

[47] M. Freedman, M. Larsen, and Z. Wang. "A Modular Functor which is Universal for Quantum Computation". In: *Communications in Mathematical Physics* 227.3 (2002), pp. 605–622.

[48] M. Freedman et al. "Topological Quantum Computation". In: *Bull. Amer. Math. Soc.* 40 (2003), pp. 31–38.

[49] Z. Wang. *Topological Quantum Computation*. Conference Board of the Mathematical Sciences. CBMS regional conference series in mathematics. Conference Board of the Mathematical Sciences, 2010.

[50] S. Trebst et al. "A Short Introduction to Fibonacci Anyon Models". en. In: vol. 176. Progress of Theoretical Physics Supplement. Kyoto: Publication Office, Progress of Theoretical Physics, 2008, pp. 384–407.

[51] L. Kauffman. *Knots and Physics, Third Edition*. K & E series on knots and everything. World Scientific, 2001.

[52] G. Kuperberg. "How Hard Is It to Approximate the Jones Polynomial?" In: *Theory of Computing* 11.6 (2015), pp. 183–219.

[53] D. Aharonov, V. Jones, and Z. Landau. "A Polynomial Quantum Algorithm for Approximating the Jones Polynomial". In: *Algorithmica* 55.3 (2009), pp. 395–421.

[54] D. Aharonov et al. "Polynomial Quantum Algorithms for Additive approximations of the Potts model and other Points of the Tutte Plane". In: *arXiv e-prints* (2007). arXiv: quant-ph/0702008.

[55] V. F. R. Jones. "A Polynomial Invariant for Knots via von Neumann Algebras". In: *Bull. Amer. Math. Soc. (N.S.)* 12.1 (1985), pp. 103–111.

[56] R. Pauncz. *Alternant Molecular Orbital Method*. Studies in Physics and Chemistry, no. 4. Saunders, 1967.

[57] A. W. Harrow. "Applications of Coherent Classical Communication and the Schur Transform to Quantum Information Theory". PhD thesis. 2005. arXiv: quant-ph/0512255.

[58] D. Bacon, I. L. Chuang, and A. W. Harrow. "Efficient Quantum Circuits for Schur and Clebsch-Gordan Transforms". In: *Phys. Rev. Lett.* 97.17 (2006), p. 170502.

[59] D. Bacon, I. L. Chuang, and A. W. Harrow. "The Quantum Schur Transform: I. Efficient Qudit Circuits". In: *arXiv e-prints* (2006). arXiv: quant-ph/0601001.

[60] W. M. Kirby and F. W. Strauch. "A Practical Quantum Algorithm for the Schur Transform". In: *Quantum Info. Comput.* 18.9–10 (2018), pp. 721–742.

[61] J. Kempe et al. "Theory of Decoherence-Free Fault-Tolerant Universal Quantum Computation". In: *Phys. Rev. A* 63.4 (2001), p. 042307.

[62] S. D. Bartlett, T. Rudolph, and R. W. Spekkens. "Classical and Quantum Communication without a Shared Reference Frame". In: *Phys. Rev. Lett.* 91.2 (2003), p. 027901.

[63] S. D. Bartlett, T. Rudolph, and R. W. Spekkens. "Reference Frames, Superselection Rules, and Quantum Information". In: *Rev. Mod. Phys.* 79 (2007), pp. 555–609.

[64] J. v. Korff and J. Kempe. “Quantum Advantage in Transmitting a Permutation”. In: *Phys. Rev. Lett.* 93.26 (2004), p. 260502.

[65] J. .W Barrett and Ileana N. - G. “The Ponzano–Regge model”. In: *Classical and Quantum Gravity* 26.15 (2009), p. 155014.

[66] R. Penrose. “Angular Momentum: an Approach to Combinatorial Space-Time”. In: *Quantum Theory and Beyond*. Ed. by T. Bastin. Cambridge: Cambridge University Press, 1971.

[67] P. Woit. *Quantum Theory, Groups and Representations*. Springer, 2017.

[68] A.R. Edmonds. *Angular Momentum in Quantum Mechanics*. Investigations in physics. Princeton University Press, 1960.

[69] T. Yamanouchi. In: *Proc. Phys. Math. Soc. Jpn.* 18.623 (1936).

[70] M. Kotani. *Tables of Molecular Integrals*. Maruzen Co., 1963.

[71] W. Fulton and J. Harris. *Representation Theory: A First Course*. Graduate Texts in Mathematics. Springer New York, 1991.

[72] J. S. Frame, G. de B. Robinson, and R. M. Thrall. “The Hook Graphs of the Symmetric Group”. In: *Canadian Journal of Mathematics* 6 (1954), pp. 316–324.

[73] H. Georgi. *Lie Algebras In Particle Physics: from Isospin To Unified Theories*. Frontiers in Physics. Avalon Publishing, 1999.

[74] G. Racah. “Theory of Complex Spectra. II”. In: *Phys. Rev.* 62.9-10 (1942), p. 438.

[75] E. W. Weisstein. *Wigner 3j-Symbol*. MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/Wigner3j-Symbol.html.

[76] L. C. Biedenharn and J. D. Louck. *The Racah-Wigner Algebra in Quantum Theory*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1984.

[77] V. Aquilanti et al. “Combinatorics of Angular Momentum Recoupling Theory: Spin Networks, Their Asymptotics and Applications”. In: *Theoretical Chemistry Accounts* 123.3 (2009), p. 237.

[78] A. M. Childs, A. W. Harrow, and P. Wocjan. “Weak Fourier-Schur Sampling, the Hidden Subgroup Problem, and the Quantum Collision Problem”. In: *arXiv e-prints* (2006). arXiv: 0609.110.

[79] H. Krovi. “An Efficient High Dimensional Quantum Schur Transform”. In: *Quantum* 3 (2019), p. 122.

[80] S. Arora and B. Barak. *Computational Complexity: A Modern Approach*. 1st. New York, NY, USA: Cambridge University Press, 2009.

[81] W. L. Ruzzo, J. Simon, and M. Tompa. “Space-Bounded Hierarchies and Probabilistic Computations”. In: *Proceedings of the Fourteenth Annual ACM Symposium on Theory of Computing*. STOC ’82. San Francisco, California, USA: Association for Computing Machinery, 1982, pp. 215–223.

[82] M. Hamermesh. *Group Theory and Its Application to Physical Problems*. Addison Wesley Series in Physics. Dover Publications, 1989.

[83] H. Boerner et al. *Representations of Groups: With Special Consideration for the Needs of Modern Physics*. Series in physics. North-Holland Publishing Company, 1963.

[84] W. Wu and Q. Zhang. "An Efficient Algorithm for Evaluating the Standard Young-Yamanouchi Orthogonal Representation with Two-Column Young tableaux for Symmetric Groups". In: *Journal of Physics A: Mathematical and General* 25.13 (1992), pp. 3737–3747.

[85] W. Wu and Q. Zhang. "The Orthogonal and the Natural Representation for Symmetric Groups". In: *International Journal of Quantum Chemistry* 50.1 (1994), pp. 55–67.

[86] Ö. Eğecioğlu. "Algorithms for the Character Theory of the Symmetric Group". In: *EUROCAL '85*. Ed. by Bob F. Caviness. Berlin, Heidelberg: Springer Berlin Heidelberg, 1985, pp. 206–224.

[87] S. Rettrup. "A Recursive Formula for Young's Orthogonal Representation". In: *Chemical Physics Letters* 47.1 (1977), pp. 59–60.

[88] R. Pauncz. *The Symmetric Group in Quantum Chemistry*. Mathematical Chemistry. Taylor & Francis, 1995.

[89] S. Bravyi et al. "Simulation of Quantum Circuits by Low-Rank Stabilizer Decompositions". In: *Quantum* 3 (2019), p. 181.

[90] Richard Jozsa and Marten Van Den Nest. "Classical Simulation Complexity of Extended Clifford Circuits". In: *Quantum Info. Comput.* 14.7 (2014), pp. 633–648.

[91] E. Kushilevitz and Y. Mansour. "Learning Decision Trees using the Fourier Spectrum". In: *SIAM Journal on Computing* 22.6 (1993), pp. 1331–1348.

[92] C. Greene, A. Nijenhuis, and H. S. Wilf. "A Probabilistic Proof of a Formula for the Number of Young Tableaux of a Given Shape". In: *Advances in Mathematics* 31.1 (1979), pp. 104–109.

[93] E. Kushilevitz and N. Nisan. *Communication Complexity*. New York, NY, USA: Cambridge University Press, 1997.

[94] A. C. C. Yao. "Some Complexity Questions Related to Distributive Computing (Preliminary Report)". In: *Proceedings of the Eleventh Annual ACM Symposium on Theory of Computing*. STOC '79. Atlanta, Georgia, USA: ACM, 1979, pp. 209–213.

[95] A. Rao and A. Yehudayoff. *Communication Complexity: and Applications*. Cambridge University Press, 2019.

[96] C. M. Caves, C. A. Fuchs, and R. Schack. "Conditions for Compatibility of Quantum-State Assignments". In: *Phys. Rev. A* 66.6 (2002), p. 062111.

[97] M. F. Pusey, J. Barrett, and T. Rudolph. "On the Reality of the Quantum State". In: *Nature Physics* 8 (2012), pp. 476–479.

[98] A. Karanjai, J. J. Wallman, and S. D. Bartlett. "Contextuality Bounds the Efficiency of Classical Simulation of Quantum Processes". In: *arXiv e-prints* (2018). arXiv: 1802.07744.

[99] R. Raz. "Fourier Analysis for Probabilistic Communication Complexity". In: *computational complexity* 5.3 (1995), pp. 205–221.

[100] I. Kremer. "Quantum Communication". MA thesis. The Hebrew University of Jerusalem, 1995.

[101] H. Buhrman et al. "Nonlocality and Communication Complexity". In: *Rev. Mod. Phys.* 82.1 (2010), pp. 665–698.

[102] A. C. C. Yao. "Probabilistic Computations: Toward a Unified Measure of Complexity". In: *Proceedings of the 18th Annual Symposium on Foundations of Computer Science*. SFCS '77. Washington, DC, USA: IEEE Computer Society, 1977, pp. 222–227.

[103] A. C. C. Yao. "Quantum Circuit Complexity". In: *Proceedings of 1993 IEEE 34th Annual Foundations of Computer Science*. 1993, pp. 352–361.

[104] Z. Bar-Yossef, T. S. Jayram, and I. Kerenidis. "Exponential Separation of Quantum and Classical One-Way Communication Complexity". In: *SIAM J. Comput.* 38.1 (2008), pp. 366–384.

[105] R. Cleve and H. Buhrman. "Substituting Quantum Entanglement for Communication". In: *Phys. Rev. A* 56.2 (1997), pp. 1201–1204.

[106] H. Buhrman, R. Cleve, and A. Wigderson. "Quantum vs. Classical Communication and Computation". In: *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing*. STOC '98. Dallas, Texas, USA: Association for Computing Machinery, 1998, pp. 63–68.

[107] P. Frankl and V. Rodl. "Forbidden Intersections". In: *Transactions of the American Mathematical Society* 300.1 (1987), pp. 259–286.

[108] A. Anshu et al. "Separations in communication complexity using cheat sheets and information complexity". In: *CoRR* abs/1605.01142 (2016). arXiv: 1605.01142.

[109] R. Raz. "Exponential Separation of Quantum and Classical Communication Complexity". In: *Proceedings of the Thirty-first Annual ACM Symposium on Theory of Computing*. STOC '99. Atlanta, Georgia, USA: ACM, 1999, pp. 358–367.

[110] H. Buhrman et al. "Quantum Fingerprinting". In: *Phys. Rev. Lett.* 87.16 (2001), p. 167902.

[111] D. Gavinsky et al. “Exponential Separations for One-Way Quantum Communication Complexity, with Applications to Cryptography”. In: *Proceedings of the Thirty-Ninth Annual ACM Symposium on Theory of Computing*. STOC ’07. San Diego, California, USA: Association for Computing Machinery, 2007, pp. 516–525.

[112] O. Regev and B. Klartag. “Quantum One-way Communication Can Be Exponentially Stronger Than Classical Communication”. In: *Proceedings of the Forty-third Annual ACM Symposium on Theory of Computing*. STOC ’11. San Jose, California, USA: ACM, 2011, pp. 31–40.

[113] D. Gavinsky. “Bare Quantum Simultaneity Versus Classical Interactivity in Communication Complexity”. In: *arXiv e-prints* (2019). arXiv: 1911.01381.

[114] A. Montina. “Exponential Communication Gap between Weak and Strong Classical Simulations of Quantum Communication”. In: *Phys. Rev. A* 87.4 (2013), p. 042331.

[115] M. S. Leifer. “ψ-Epistemic Models are Exponentially Bad at Explaining the Distinguishability of Quantum States”. In: *Phys. Rev. Lett.* 112.16 (2014), p. 160404.

[116] D. Schmid and R. W. Spekkens. “Contextual Advantage for State Discrimination”. In: *Phys. Rev. X* 8.1 (2018), p. 011015.

[117] R. W. Spekkens. “Contextuality for Preparations, Transformations, and Unsharp Measurements”. In: *Phys. Rev. A* 71.5 (2005), p. 052108.

[118] I. Newman. “Private vs. Common Random Bits in Communication Complexity”. In: *Inf. Process. Lett.* 39.2 (1991), pp. 67–71.

[119] L. Hardy. “Quantum Ontological Excess Baggage”. In: *Studies in the History and Philosophy of Modern Physics* 35.2 (2004), pp. 267–276.

[120] B. F. Toner and D. Bacon. “Communication Cost of Simulating Bell Correlations”. In: *Phys. Rev. Lett.* 91.18 (2003), p. 187904.

[121] C. Perry, R. Jain, and J. Oppenheim. “Communication Tasks with Infinite Quantum-Classical Separation”. In: *Phys. Rev. Lett.* 115.3 (2015), p. 030504.

[122] Z. Liu et al. “Doubly Infinite Separation of Quantum Information and Communication”. In: *Phys. Rev. A* 93.1 (2016), p. 012347.

[123] A. Hansen, A. Montina, and S. Wolf. “Simple Algorithm for Computing the Communication Complexity of Quantum Communication Processes”. In: *Phys. Rev. A* 93.4 (2016), p. 042315.

[124] A. Montina. “Communication Complexity and the Reality of the Wave Function”. In: *Modern Physics Letters A* 30.01 (2015), p. 1530001.

[125] A. Montina and S. Wolf. “Necessary and Sufficient Optimality Conditions for Classical Simulations of Quantum Communication Processes”. In: *Phys. Rev. A* 90.1 (2014), p. 012309.

[126] A. Montina. “Epistemic View of Quantum States and Communication Complexity of Quantum Channels”. In: *Phys. Rev. Lett.* 109 (2012), p. 110501.

[127] A. Montina. “Approximate Simulation of Entanglement with a Linear Cost of Communication”. In: *Phys. Rev. A* 84 (2011), p. 042307.

[128] J. Bae and L-C. Kwek. “Quantum State Discrimination and its Applications”. In: *Journal of Physics A Mathematical General* 48.8 (2015), p. 083001.

[129] C. W. Helstrom. “Quantum Detection and Estimation Theory”. In: *Journal of Statistical Physics* 1.2 (1969), pp. 231–252.

[130] D. Ha and Y. Kwon. “Complete Analysis for Three-Qubit Mixed-State Discrimination”. In: *Phys. Rev. A* 87.6 (2013), p. 062302.

[131] H. Yuen, R. Kennedy, and M. Lax. “Optimum Testing of Multiple Hypotheses in Quantum Detection Theory”. In: *IEEE Transactions on Information Theory* 21.2 (1975), pp. 125–134.

[132] S. Bandyopadhyay et al. “Conclusive Exclusion of Quantum States”. In: *Phys. Rev. A* 89.2 (2014), p. 022336.

[133] T. Heinosaari and O. Kerppo. “Antidistinguishability of Pure Quantum States”. In: *Journal of Physics A: Mathematical and Theoretical* 51.36 (2018), p. 365303.

[134] J. Barrett et al. “No ψ-Epistemic Model Can Fully Explain the Indistinguishability of Quantum States”. In: *Phys. Rev. Lett.* 112 (2014), p. 250403.

[135] A. Molina. “POVMs are equivalent to projections for perfect state exclusion of three pure states in three dimensions”. In: *Quantum* 3 (2019), p. 117.

[136] T. Heinosaari and O. Kerppo. “Communication of Partial Ignorance with Qubits”. In: *Journal of Physics A: Mathematical and Theoretical* 52.39 (2019), p. 395301.

[137] V. Havlíček and J. Barrett. “Simple Communication Complexity Separation from Quantum State Antidistinguishability”. In: *Phys. Rev. Research* 2.1 (2020), p. 013326.

[138] N. J. Cerf, N. Gisin, and S. Massar. “Classical Teleportation of a Quantum Bit”. In: *Phys. Rev. Lett.* 84 (2000), pp. 2521–2524.

[139] A. Ambainis et al. “The Quantum Communication Complexity of Sampling”. In: *SIAM J. Comput.* 32 (2003), pp. 1570–1585.

[140] D. Gavinsky, J. Kempe, and R. de Wolf. “Exponential Separation of Quantum and Classical One-Way Communication Complexity for a Boolean Function”. In: *arXiv e-prints* (2006). arXiv: quant-ph/0607174.

[141] I. Kerenidis and R. Raz. “The One-Way Communication Complexity of the Boolean Hidden Matching Problem”. In: *arXiv e-prints* (2006). arXiv: quant-ph/0607173.

[142] A. Montanaro. “A New Exponential Separation Between Quantum and Classical One-way Communication Complexity”. In: *Quantum Info. Comput.* 11.7-8 (2011), pp. 574–591.

[143] Ashley Montanaro. “Quantum states cannot be transmitted efficiently classically”. In: *Quantum* 3 (2019), p. 154.

[144] A. M. Raigorodskii. “On a bound in Borsuk’s problem”. In: *Russian Mathematical Surveys* 54.2 (1999), pp. 453–454.

[145] L. Welch. “Lower Bounds on the Maximum Cross Correlation of Signals”. In: *IEEE Trans. Inf. Theory* 20.3 (1974), pp. 397–399.

[146] V. I. Levenshtein and G. A. Kabatiansky. “On Bounds for Packings on a Sphere and in Space”. In: *Probl. Peredachi Inf.* 14.1 (1978), pp. 14–25.

[147] V. I. Levenshtein. “Bounds for Packings of Metric Spaces and Some of Their Applications”. In: *Probl. Kibern.* 40 (1983), pp. 43–110.

[148] J. M. Renes. “Frames, Designs, and Spherical Codes in Quantum Information Theory”. PhD thesis. The University of New Mexico, 2004.

[149] A. Roy and S. Suda. “Complex Spherical Designs and Codes”. In: *Journal of Combinatorial Designs* 22.3 (2014), pp. 105–148.

[150] H. Zorlein and M. Bossert. “Coherence Optimization and Best Complex Antipodal Spherical Codes”. In: *IEEE Transactions on Signal Processing* 63 (2015), pp. 6606–6615.

[151] M. Heredia Conde and O. Loffeld. “Fast Approximate Construction of Best Complex Antipodal Spherical Codes”. In: *CoRR* abs/1705.03280 (2017). arXiv: 1705.03280.

[152] J. M. Renes et al. “Symmetric Informationally Complete Quantum Measurements”. In: *Journal of Mathematical Physics* 45 (2004), pp. 2171–2180.

[153] A. J. Scott and M. Grassl. “Symmetric informationally complete positive-operator-valued measures: A new computer study”. In: *Journal of Mathematical Physics* 51.4 (2010), pp. 042203–042203.

[154] A. J. Scott. “SICs: Extending the List of Solutions”. In: *arXiv e-prints* (2017). arXiv: 1703.03993.

[155] B. Stacey. “SIC-POVMs and Compatibility among Quantum States”. In: *Mathematics* 4.2 (2016), p. 36.

[156] W. K. Wootters and B. D. Fields. “Optimal State-Determination by Mutually Unbiased Measurements”. In: *Annals of Physics* 191 (1989), pp. 363–381.

[157] C. Chabauty. “Résultats sur l’Empilement de Calottes Égales sur une Périsphère de $\mathbb{R}^n$ et Correction à un Travail Antérieur”. In: *C. R. Acad. Sci.* 236 (1953), pp. 1462–1464.

[158] C. E. Shannon. “Probability of Error for Optimal Codes in a Gaussian Channel”. In: *The Bell System Technical Journal* 38.3 (1959), pp. 611–656.

[159] A. D. Wyner. “Capabilities of Bounded Discrepancy Decoding”. In: *The Bell System Technical Journal* 44.6 (1965), pp. 1061–1122.

[160] M. Jenssen, F. Joos, and W. Perkins. “On Kissing Numbers and Spherical Codes in High Dimensions”. In: *Advances in Mathematics* 335 (2018), pp. 307–321.

[161] A. Noga and J. H. Spencer. *The Probabilistic Method*. 4th. Wiley Publishing, 2016.

[162] K. Zyczkowski and H.-J. Sommers. “Induced measures in the space of mixed quantum states”. In: *Journal of Physics A: Mathematical and General* 34.35 (2001), pp. 7111–7125.

[163] I. Bengtsson and K. Zyczkowski. *Geometry of Quantum States: An Introduction to Quantum Entanglement*. Cambridge University Press, 2006.

[164] A. Montina and S. Wolf. “Lower Bounds on the Communication Complexity of Two-Party (Quantum) Processes”. In: *2014 IEEE International Symposium on Information Theory*. 2014, pp. 1484–1488.

[165] G. Brassard, R. Cleve, and A. Tapp. “Cost of Exactly Simulating Quantum Entanglement with Classical Communication”. In: *Phys. Rev. Lett.* 83 (1999), pp. 1874–1877.

[166] Dmitry Gavinsky. “Classical Interaction Cannot Replace a Quantum Message”. In: *Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing*. STOC ’08. Victoria, British Columbia, Canada: Association for Computing Machinery, 2008, pp. 95–102.

[167] N. Harrigan and R. W. Spekkens. “Einstein, Incompleteness, and the Epistemic View of Quantum States”. In: *Foundations of Physics* 40.2 (2010), pp. 125–157.

[168] P. G. Lewis et al. “Distinct Quantum States Can Be Compatible with a Single State of Reality”. In: *Phys. Rev. Lett.* 109.15 (2012), p. 150404.

Appendix

Appendix A

Part I

A.1 Yamanouchi-Kotani Representation

The following theorem derives the form of Yamanouchi-Kotani matrices for transpositions. This is used in Theorem 2 to show that the Yamanouchi-Kotani representation is in fact equivalent to the Young's orthogonal form – a canonical basis for orthogonal irreducible representations of S_n. This is used to derive the quantum algorithm that approximates matrix elements of this representation – a centerpiece to the conjecture about non-classicality of Permutational Quantum Computing.

Theorem 4 (Yamanouchi-Kotani representation matrices for transpositions [69, 56])**.** *The representation matrices for transpositions for $n = 3$ and $J = 1/2$, $M = -1/2$ are given by:*

$$W(\sigma_1) = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad W(\sigma_2) = \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{bmatrix}. \tag{A.1}$$

Representation matrices for transpositions σ_k on $n \geq 3$ qubits and $k \leq n-2$ for a Young diagram $\lambda = [\mu, \nu]$, such that $\mu + \nu = n$ can be defined recursively as:

$$W(\sigma_k) = \left[\begin{array}{c|c} W'(\sigma_k) & 0 \\ \hline 0 & W''(\sigma_k) \end{array}\right] \tag{A.2}$$

where $W'(\sigma_k)$ is the representation matrix for $n-1$ qubits and $\lambda = [\mu - 1, \nu]$, $W''(\sigma_k)$ is the representation matrix for $n-1$ qubits and $\lambda = [\mu, \nu - 1]$. The representation matrix $W(\sigma_{n-1})$ is given by:

$$W(\sigma_{n-1}) = \left[\begin{array}{c|ccc|c} \mathbf{1} & 0 & 0 & 0 & \\ \hline 0 & -a\mathbf{1} & b\mathbf{1} & 0 & f^{[\mu-1,\nu-1]} \\ 0 & b\mathbf{1} & a\mathbf{1} & 0 & f^{[\mu-1,\nu-1]} \\ \hline 0 & 0 & 0 & \mathbf{1} & f^{[\mu,\nu-2]} \end{array}\right] \begin{array}{c} f^{[\mu-2,\nu]} \\ \\ \\ \\ \end{array} \tag{A.3}$$

where $a = 1/(\mu - \nu + 1)$ *and* $b = \sqrt{1 - a^2}$ *and* f^λ *denote sizes of the respective matrix blocks.*

Proof. Base case: Using the eigenfunctions from Eq. 2.40 that form a basis for an eigenspace with $n = 3$ and $J = 1/2$, $M = -1/2$, we have that:

$$W(\sigma_1) = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad W(\sigma_2) = \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{bmatrix}. \tag{A.4}$$

Induction: Assume that all representation matrices for $(n-1)$ and all possible values of $J_{[n-1]}$ have been found. It follows for a Young tableau λ_i of shape $[\mu, \nu]$, such that $\mu + \nu = n$ and $\mu - \nu = 2J$ that:

$$|J, M, \lambda_i\rangle = C^{J,M}_{J-\frac{1}{2},M\pm\frac{1}{2};\frac{1}{2},\mp\frac{1}{2}} \left| J - \frac{1}{2}, M \pm \frac{1}{2}, \lambda_i' \right\rangle \left| \mp \frac{1}{2} \right\rangle \tag{A.5}$$

$$+ C^{J,M}_{J+\frac{1}{2},M\pm\frac{1}{2};\frac{1}{2},\mp\frac{1}{2}} \left| J + \frac{1}{2}, M \pm \frac{1}{2}, \lambda_i'' \right\rangle \left| \mp \frac{1}{2} \right\rangle, \tag{A.6}$$

where λ_i' is a standard Young tableau of shape $[\mu - 1, \nu]$ and λ_i'' is a Young tableau of shape $[\mu, \nu - 1]$. Any qubit transposition U_{σ_k} for $k \leq n - 2$ acting on this state does not alter the shapes of λ_i' or λ_i''. Combined with orthogonality of genealogical spin functions, this means that the $f \times f$ unitary matrix:

$$W(\sigma_k)_{\lambda_i,\lambda_j} = \langle \lambda_i | U_{\sigma_k} | \lambda_j \rangle, \tag{A.7}$$

block-diagonalizes as:

$$W(\sigma_k) = \left[\begin{array}{c|c} W'(\sigma_k) & 0 \\ \hline 0 & W''(\sigma_k) \end{array}\right] \begin{array}{c} f^{[\mu-1,\nu]} \\ f^{[\mu,\nu-1]} \end{array}, \quad \text{for } k \leq n - 2. \tag{A.8}$$

where $W'(\sigma_k)$ and $W''(\sigma_k)$ are representations of transpositions for their Young diagrams of shapes $[\mu - 1, \nu]$ and $[\mu, \nu - 1]$ respectively.

The matrix $W(\sigma_{n-1})$ can be obtained by going another step back in Eqs. A.5 and expanding the Clebsch-Gordan coefficients with the Racah formula of Eq. 2.35. This calculation is due to Yamanouchi [69] and we only sketch it here – it results in [56]:

$$\begin{aligned} |J, M, \lambda_i\rangle = {} & \alpha_1 |J + 1, M - 1, \lambda_i'\rangle |11\rangle + \alpha_2 |J + 1, M, \lambda_i'\rangle (|10\rangle + |01\rangle) \\ & + \alpha_3 |J + 1, M + 1, \lambda_i'\rangle |00\rangle + \beta_1 |J - 1, M - 1, \lambda_i''\rangle |11\rangle \\ & + \beta_2 |J - 1, M, \lambda_i''\rangle (|10\rangle + |01\rangle) + \beta_3 |J - 1, M, \lambda_i''\rangle |00\rangle \\ & + \gamma_1 |J, M - 1, \lambda_i'''\rangle |11\rangle + \gamma_2 |J, M, \lambda_i'''\rangle |01\rangle \\ & + \gamma_3 |J, M, \lambda_i'''\rangle |01\rangle + \gamma_4 |J, M + 1, \lambda_i'''\rangle |00\rangle . \end{aligned}$$

for some coefficients α, β, γ that are given by the appropriate products of Clebsch-Gordan coefficients. The only part of the wavefunction that transforms nontrivially under the exchange of the n-th and $(n-1)$-st particles is the one with $\{\gamma_i\}_i$ coefficients. Substituting for the α, β, γ coefficients, the transposition matrix becomes:

$$W(\sigma_{n-1}) = \left[\begin{array}{c|cc|c} \mathbf{1} & 0 & 0 & 0 \\ \hline 0 & -a\mathbf{1} & b\mathbf{1} & 0 \\ 0 & b\mathbf{1} & a\mathbf{1} & 0 \\ \hline 0 & 0 & 0 & \mathbf{1} \end{array}\right] \begin{array}{l} f^{[\mu-2,\nu]} \\ f^{[\mu-1,\nu-1]} \\ f^{[\mu-1,\nu-1]} \\ f^{[\mu,\nu-2]} \end{array} \tag{A.9}$$

where $a = 1/(\mu - \nu + 1)$ and $b = \sqrt{1-a^2}$. □

A.2 Equivalence of Young's Orthogonal Form and Yamanouchi-Kotani Representation

Theorem 2. *Yamanouchi-Kotani representation and Young's orthogonal representations are equivalent [56].*

Proof. The proof is inductive and uses the bijection between spin eigenfunctions in the sequentially coupled basis and Young tableaux in the standard form from Claim 5. *Base:* Base case at $n = 3$, $J = \frac{1}{2}$. Under the bijection of Claim 5, these correspond to the standard Young tableaux of shape $\ydiagram{2,1}$, which we order as 1:$\begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1}\end{array}$ and 2:$\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array}$. The Yamanouchi-Kotani matrices are (see Eq. A.1):

$$W(\sigma_1) = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad W(\sigma_2) = \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{bmatrix}. \tag{A.10}$$

We now verify that these matrices are equivalent to the Young's orthogonal form ones. The axial distances in these tableaux are given by:

$$d^1_{12} = 1, \qquad d^2_{12} = -1, \qquad d^1_{23} = -2, \qquad d^2_{23} = 2,$$

which fixes the diagonal elements of the representation matrices (see Def. 13). We also have that:

$$\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array} \neq (1,2) \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1}\end{array}, \qquad \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array} = (2,3) \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1}\end{array},$$

which fixes the off-diagonals. The matrices obtained from transition amplitudes and Young orthogonal form are equivalent.

Induction: Assume for induction that the transition amplitudes between genealogical spin eigenfunctions correspond to matrix elements of Young's orthogonal matrices for any transposition and any Young diagram with up to two rows and up to $n-1$ boxes. We show that the correspondence then holds for Young diagrams with two rows and n boxes as well. Consider the set of standard Young tableaux with shape $\lambda = [\mu, \nu]$, such that $\mu + \nu = n$. The set of tableaux can be divided into two groups:

1. The first group of the tableaux has the element n in the second row. From Claim 5, the n-qubit spin eigenfunctions in this set were obtained from the $n-1$ qubit spin eigenfunctions with Young diagrams of shape $[\mu,\, \nu - 1]$ – that is with $J_{[n-1]} = J + 1/2$. See Fig. A.1.

Figure A.1: Young tableaux in the first group were obtained from $n-1$ qubit spin eigenfunctions by adding a box with n to the second row. The overhang changes from $2J+1$ to $2J$, meaning that $J_{[n-1]} = J + 1/2$. The example above shows a Young tableau of shape $[4,3]$ with $J = 1/2$ built from a Young tableau $[4,2]$ with $J_{[n-1]} = 1$.

2. The second group of the tableaux has the element n in the first row. From Claim 5, the n-qubit spin eigenfunctions in the set were obtained from the $n-1$ qubit spin eigenfunctions with Young diagrams of shape $[\mu - 1,\, \nu]$ – that is with $J_{[n-1]} = J - 1/2$. See Fig. A.2.

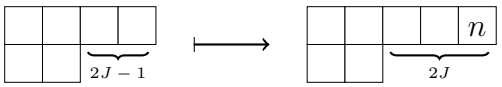


Figure A.2: Young tableaux in the second group were obtained from $n-1$ qubit spin eigenfunctions by adding a box with n to the first row. The overhang changes from $2J-1$ to $2J$, meaning that $J_{[n-1]} = J - 1/2$. The example above shows a Young tableau of shape $[5,2]$ with $J = 3/2$ built from a Young tableau $[4,2]$ with $J_{[n-1]} = 1$.

The proof is first given for transpositions σ_k acting on the first $n-1$ qubits. When such transposition acts on a spin eigenfunctions in the first group, it never results in a spin eigenfunction from the second group, as it only acts on the first $n-1$ qubits,

commutes with $S^2_{[n-1]}$ and therefore preserves $J_{[n-1]}$. The matrices corresponding to these transpositions are therefore block diagonals of the following form:

$$W(\sigma_k) = \left[\begin{array}{c|c} W_1 & 0 \\ \hline 0 & W_2 \end{array}\right], \; k \leq n-2. \tag{A.11}$$

Both of the non-zero blocks W_1 and W_2 correspond to $n-1$ qubit representation matrices, for which the claim holds by the inductive assumption. It remains to establish the identity of the transposition σ_{n-1}. For this, the set of standard Young tableaux on n qubits can be partitioned into four subsets:

1. The first subset f_{00} is the set of all standard tableaux of shape $\lambda = [\mu, \nu]$ with both n and $n-1$ in the second row. These are all obtained by taking a standard Young tableau of shape $[\mu, \nu-2]$ and adding the two boxes to the bottom row. The axial distance of $(n-1)$ and n is -1 and the $(n-1, n)$ interchange applied to the tableau does not yield a standard tableau. It follows from Def. 13 that this gives a $|f_{00}| \times |f_{00}|$ identity block in $W(\sigma_{n-1})$.

2. The second subset is a set f_{11} of all standard tableaux of shape λ with both $n-1$ and n in the first row. These are all obtained by taking a standard Young tableau of shape $[\mu - 2, \nu]$ and adding both boxes to the top row. The axial distance of $(n-1)$ and n is again -1 and the $(n-1, n)$ interchange applied to the tableau also does not yield a standard tableau. It follows from Def. 13 that this gives a $|f_{11}| \times |f_{11}|$ identity block in $W(\sigma_{n-1})$.

3. The set f_{01} of all standard tableaux of shape λ with $n-1$ in the first and n in the second row. The axial distance between the two boxes is $d^{01}_{n,n-1} = \mu - \nu + 1$. By Claim 6, we have that $d^{01}_{n,n-1} = 2J + 1$.

4. The set f_{10} of all standard tableaux of shape λ with n in the first and $n-1$ in the second row. By applying the $(n, n-1)$ transposition to this tableau, one obtains a standard tableau from the group f_{01}. It follows that $d^{10}_{n,n-1} = -d^{01}_{n,n-1} := -\frac{1}{a}$ and that there is a nonzero off-diagonal element $b = \sqrt{1-a^2}$ that links the two tableaux.

This fully determines the generator matrix:

$$W(\sigma_n) = \left[\begin{array}{c|ccc} \mathbf{1} & 0 & 0 & 0 \\ \hline 0 & -a\mathbf{1} & b\mathbf{1} & 0 \\ 0 & b\mathbf{1} & a\mathbf{1} & 0 \\ \hline 0 & 0 & 0 & \mathbf{1} \end{array}\right] \begin{array}{l} f_{00} \\ f_{01} \\ f_{10} \\ f_{11} \end{array}, \tag{A.12}$$

where $a = 1/(\mu - \nu + 1)$ and $b = \sqrt{1 - a^2}$. The theorem follows by comparison to the Yamanouchi-Kotani representation introduced in Sec 2.5. □

A.3 Greene-Nijenhuis-Wilff algorithm

Here we describe the algorithm of [92] and its correctness proof. It is also a proof of the Frame-Robinson-Thrall formula that was used in Eq. 2.25 [72]. See Algorithm 4.

Algorithm 4: Greene, Nijenhuis, Wilff [92]

Input: Standard Young diagram $\lambda \vdash n$.

Algorithm:

1. Set $\lambda' \leftarrow \lambda$ and $k \leftarrow n$.
2. Choose a uniformly random box (a, b) of λ'.
3. Then:
   - If (a, b) is at the end of a row a of λ' and there is no box below it, remove the box (a, b) from λ'. Insert k to the (a, b) box of λ and set $k \leftarrow k - 1$. If $k = 0$, terminate the algorithm and output λ; otherwise go to step 2.
   - If (a, b) is not at the end of a row a of λ, choose an element (a', b') uniformly randomly from the hook of (a, b) (excluding (a, b)) and set $(a, b) \leftarrow (a', b')$. Repeat step 3.

Output: A uniformly random Young tableau of shape λ.

We now prove the algorithm correctness that also proves the Frame-Robinson-Thrall formula as a byproduct. Write $\lambda = [\lambda_1, \lambda_2, \dots \lambda_m]$ for the standard Young diagram λ, where λ_i is the length of its i-th row. Define:

$$f^\lambda = f(\lambda_1, \lambda_2 \dots \lambda_m) = \begin{cases} \frac{n!}{\prod_{ij} h_{ij}} & \text{if } \lambda_1 \geq \lambda_2 \dots \geq \lambda_m, \\ 0 & \text{otherwise.} \end{cases} \tag{A.13}$$

where as before, h_{ij} is the hook-length of the box (i, j). The central observation[1] is that the box with n can only appear at the 'corners' of standard Young tableaux –

[1] Similarly to the proofs related to the Yamanouchi-Kotani and Young's orthogonal form representation equivalence

that is end of rows that have is no box under it. Removing this box leaves a standard Young tableau of a smaller shape. So we expect that:

$$f = f(\lambda_1, \lambda_2 \dots \lambda_m) = \sum_{\alpha=1}^{m} f(\lambda_1, \lambda_2 \dots \lambda_{\alpha-1}, \lambda_\alpha - 1, \dots, \lambda_m) = \sum_{\alpha} f_\alpha. \tag{A.14}$$

By to Eq. A.13, this summation olnly runs over all Young tableaux on $n-1$ boxes given by removing the corner boxes from the original tableau. If this holds, the Frame-Robinson-Thrall formula follows by inducing on n.

Consider now the random process defined in Algorithm 4. The algorithm chooses a sequence of boxes $(a, b) = (a_1, b_1) \to \dots \to (a_j, b_j) = (\alpha, \beta)$, where (α, β) is a corner box and will be called a terminal box. The intermediate boxes are each chosen from the hook of the preceeding box. The probability of choosing a specific terminal box (α, β) is then given by:

$$p(\alpha, \beta) = \frac{1}{n} \left(\frac{1}{h_{a_1 b_1} - 1} \right) \left(\frac{1}{h_{a_2 b_2} - 1} \right) \dots \left(\frac{1}{h_{a_{j-1} b_{j-1}} - 1} \right). \tag{A.15}$$

The correctness proof follows by verifying that:

$$p(\alpha, \beta) = f_\alpha / f. \tag{A.16}$$

To show this, first notice that:

$$\frac{f_\alpha}{f} = \frac{1}{n} \prod_{1 \le i < \alpha} \left(\frac{h_{i\beta}}{h_{i\beta} - 1} \right) \prod_{1 \le j < \beta} \left(\frac{h_{\alpha j}}{h_{\alpha j} - 1} \right) \tag{A.17}$$

This can be shown by noticing that removing a terminal box (α, β) only changes the hooklengths in row α and column β. See Fig. A.3.

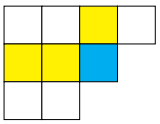


Figure A.3: Intuition for Eq. A.17. Suppose that the blue terminal box (α, β) is removed from the above Young diagram. The hooklengths decrease by 1 for all the yellow boxes, but stay the same for all the white boxes. The only terms that contribute to the product in f_α / f correspond to the yellow boxes – these are all boxes in the row α and all boxes in the column β.

Now suppose that a sequence $(a, b) = (a_1, b_1) \to \dots \to (a_w, b_w) = (\alpha, \beta)$ was sampled by the random process defined in Algorithm 4. Define the row projector $A =$

$\{a_1, a_2 \dots a_w\}$ and the column projector $B = \{b_1, b_2 \dots b_w\}$. Then the probability:

$$p(A, B|\, a, b) = \prod_{i \in A, i \neq \alpha} \frac{1}{h_{i\beta} - 1} \prod_{j \in B, j \neq \beta} \frac{1}{h_{\alpha j} - 1} := \prod . \tag{A.18}$$

The proof of this equation is derived in Lemma 3 of [92] – we give the proof below and give additional intuition for it.

Proof. First, it follows that:

$$p(A, B|\, a, b) = \frac{1}{h_{ab} - 1} \left(p(A - a, B|\, a_2, b_1) + p(A, B - b|\, a_1, b_2) \right). \tag{A.19}$$

Eq. A.19 is explained in Fig. A.4.

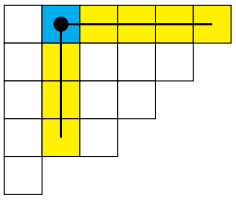


Figure A.4: Intuition for Eq. A.19. The probability of getting a pair (A, B) of projectors given that the blue box (a, b) was chosen initially is the same as getting a pair of projectors $(A - a, B)$ given that the second box $(a_2, b_2) = (a_2, b)$ was chosen from the vertical part of the hook of (a, b) and the probability of choosing projector pairs $(A, B - b)$, given that the second box $(a_2, b_2) = (a, b_2)$ was chosen from the horizontal part of the hook of (a, b). The second box is chosen u.a.r. with probability $1/(h_{ab} - 1)$.

By induction on w, assume that:

$$p(A - a_1, B|\, a_2, b_1) = (h_{a_1\beta} - 1) \cdot \prod , \tag{A.20}$$

$$p(A, B - b_1|\, a_1, b_2) = (h_{\alpha b_1} - 1) \cdot \prod , \tag{A.21}$$

where $\prod$ is the right hand side of Eq. A.18. It follows that:

$$p(A, B|\, a, b) = \frac{1}{h_{ab} - 1} \left((h_{a_1\beta} - 1) + (h_{\alpha b_1} - 1) \right) \cdot \prod . \tag{A.22}$$

But:

$$h_{ab} - 1 = (h_{a\beta} - 1) + (h_{\alpha b} - 1), \tag{A.23}$$

(see Fig. A.5 for a graphical proof of this), which shows that $p(A, B|\, a, b) = \prod$.

□

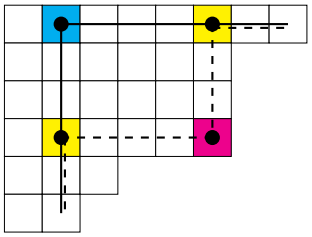


Figure A.5: Intuition for Eq. A.23. Let (a, b) be the initial box (blue) and (α, β) be the terminal box (violet). Then $h_{ab} - 1$ is the number of white or yellow boxes crossed by the solid thick black line. We have that $h_{a\beta} - 1 + h_{\alpha b} - 1$ is the number of white or violet boxes crossed by the dashed thick line, counting the violet box twice. Comparing the length of the dashed 'path' to the length of the solid 'path' in the above diagram, it is easy to see that the two numbers are equal.

The probability $p(\alpha, \beta)$ is given by summing the conditional probabilities over the first box chosen and then summing over all possible projectors $A \subseteq \{1, 2, \ldots, \alpha\}$ and $B \subseteq \{1, 2, \ldots, \beta\}$ and $a = \min A$, $b = \min B$:

$$p(\alpha, \beta) = \sum p(A, B|\, a, b). \tag{A.24}$$

By the above proof, this is the same as expanding the product on the right hand side of Eq. A.17. This implies that:

$$p(\alpha, \beta) = f_\alpha / f, \tag{A.25}$$

which implies correctness of the algorithm by induction. By normalization, it also follows that:

$$\sum_\alpha f_\alpha / f = 1. \tag{A.26}$$

This proves the Frame-Robinson-Thrall formula of Eq. 2.25.

Appendix B

Part II

B.1 Yao's Principle

Theorem 6.

$$R_\epsilon(f) \geq \max_\mu D^\mu_\epsilon(f). \tag{B.1}$$

Proof. A randomized $\epsilon > 0$ error protocol P is correct on all inputs with probability $1 - \epsilon$:

$$\forall (x, y) \in X \times Y : \Pr\left[P(x, y) = f(x, y)\right] \geq 1 - \epsilon. \tag{B.2}$$

As discussed previously, P can be seen as a distribution over all deterministic protocols $\{P(x, y|\, s)\}_s$ for the task, weighted by the shared randomness $s \sim \pi$:

$$P(x, y) = \sum_s P(x, y|\, s)\pi(s), \tag{B.3}$$

Let $[P(x, y|\, s) = f(x, y)]$ be the indicator function of the event $P(x, y|\, s) = f(x, y)$. We can then rewrite Eq. B.2 as:

$$\forall (x, y) \in X \times Y : \sum_s \left[P(x, y|\, s) = f(x, y)\right] \pi(s) \geq 1 - \epsilon. \tag{B.4}$$

But this also means that for any distribution $\mu : X \times Y \to [0, 1]$ on inputs:

$$\sum_{(x,y)} \sum_s \left[P(x, y|\, s) = f(x, y)\right] \mu(x, y)\pi(s) \geq 1 - \epsilon. \tag{B.5}$$

This means that there exists a deterministic protocol (a fixed value of s), such that:

$$\sum_{(x,y)} \left[P(x, y|\, s) = f(x, y)\right] \mu(x, y) \geq 1 - \epsilon. \tag{B.6}$$

Assume for contradiction that the deterministic protocol has strictly better success probability (over the input distribution) than the randomized protocol (over the internal randomness). By reducing the success rate to ϵ, it could achieve better complexity than the randomized strategy. This gives a contradiction.

Similar argument holds for zero-error complexity R_0. Recall that the communication complexity measure is different here - it is the average over communication costs of all deterministic protocols on the worst case input (see Definition 19). Let $\{P(x, y|\, s)\}_s$ be the set of all deterministic communication protocols for the task and $C[P(x, y|\, s')]$ be the cost of the deterministic protocol $P(x, y|\, s') \in \{P(x, y|\, s)\}_s$ that achieves the minimum distributional complexity on inputs drawn from μ. We have that:

$$D_0^{\mu}(f) = \sum_{(x,y)} \mu(x, y) C[P(x, y|\, s')]. \tag{B.7}$$

Since $P(x, y|s')$ achieves the minimum distributional complexity, it follows that:

$$\forall s' : \sum_{(x,y)} \mu(x, y) C[P(x, y|\, s')] \leq \sum_{(x,y)} \mu(x, y) C[P(x, y|\, s)]. \tag{B.8}$$

In particular, we have for any distribution over the shared randomness that:

$$\sum_s \pi(s) \sum_{(x,y)} \mu(x, y) C[P(x, y|\, s')] \leq \sum_{(x,y)} \mu(x, y) \sum_s C[P(x, y|\, s)]\pi(s). \tag{B.9}$$

Note that $\sum_{s'} C[P(x, y|\, s')]\pi(s')$ is the average running time on input (x, y). It follows that (see Definition 19):

$$D_0^{\mu}(f) \leq \sum_{(x,y)} \mu(x, y) \sum_{s'} C[P(x, y|\, s')]\pi(s') \leq \max_{(x,y)} \sum_{s'} C[P(x, y|\, s')] = R_0(f). \tag{B.10}$$

□

The equality in Yao's minimax principle is derived using the Von Neumann's minimax principle from zero-sum games, which can be in turn related to linear programming duality [93]. The proof is standard, lengthy and somewhat complicated and I omit it here.