QUANTUM SIZES: COMPLEXITY, DIMENSION, AND MANY-BOX LOCALITY

CAI YU

(B.Sc.(Hons.), NUS)

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

in the

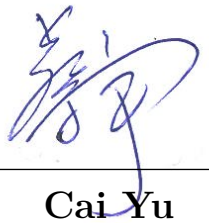
Centre for Quantum Technologies
National University of Singapore

2015

DECLARATION

I hereby declare that this thesis is my original work and has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.



Cai Yu

November 2, 2015

LIST OF PUBLICATIONS

- H. N. Le, Y. Cai, X. Wu, and V. Scarani. "Tree-size complexity of multiqubit states". *Phys. Rev. A* **88** 1 (2013), p. 012321
- H. N. Le, Y. Cai, X. Wu, R. Rabelo, and V. Scarani. "Maximal tree size of few-qubit states". *Phys. Rev. A* **89** 6 (2014), p. 062333.
- Y. Cai, H. N. Le, and V. Scarani. "State complexity and quantum computation". *Ann. Phys.* **527** 9 (2015), pp. 684-700.

Publication not forming part of this thesis:

- X. Wu, Y. Cai, T.H. Yang, Le Huy Nguyen, J.D. Bancal, and V. Scarani. "Robust self testing of the 3-qubit W state". *Phys. Rev. A* **90** (2014), p. 042339.

ACKNOWLEDGEMENTS

I could not have accomplished the my PhD candidature and this thesis without support from various people.

First and foremost, I would like to express my sincere gratitude to my PhD supervisor, Professor Valerio Scarani, for his continuous guidance throughout my PhD study. He is superb at making complex ideas digestible, always optimistic in research and gives valuable advice for both academic and non-academic matters.

I owe much to two post-docs in our group, Le Huy Nguyen and Jean-Daniel Bancal. They have spent much time guiding me through my projects. They taught me much on how to tackle problems and think critically.

My PhD life would not be so enjoyable without support from friends and colleagues. First thanks to past and present members of our group: Le Phuc Thinh, Melvyn Ho, Colin Teo, Rafael Rabelo, Alex Roulet, Wu Xingyao, Haw Jingyan, Goh Koon Tong, Yang Tzyh Haur, Wang Yimin, Lana Sheridan, Jiri Minar, and Daniel Cavalcanti. I also need to thank my collaborators and everyone who have generously shared their scientific expertise with me, Jacquiline Romero, Poh Hou Shun, Nicolas Brunner, Tamás Vértesi, Nelly Ng, Bobby Tan, Jedrzej Kaniewski, Marek Wajs, Andy Chia and many more. I am grateful to CQT for its conducive environment and exceptionally efficient staffs.

Last but not least, I would like to thank my family, especially my wife, for their unconditional love and support.

CONTENTS

Acknowledgements iii

Contents vi

Abstract 1

1 Introduction 2

1.1 Quantum weirdness . . . 2
1.2 Complexity . . . 4
1.3 Nonlocality . . . 5

2 Tree size complexity 7

2.1 Motivation . . . 7
2.1.1 An alternative size for Schrödinger's cat . . . 7
2.1.2 Quantum computation . . . 8
2.1.3 Types and measures of complexity . . . 10
2.2 Definition and basic properties of tree size . . . 11
2.3 Tree size of few-qubit states . . . 13
2.3.1 Two qubits . . . 14
2.3.2 Three qubits . . . 14
2.3.3 Four qubits . . . 19
2.4 Simple states . . . 30
2.5 Complex states . . . 32
2.5.1 Methods to obtain lower bounds on tree size . . . 32
2.5.2 Immanant states . . . 35
2.5.3 Deutsch-Jozsa states . . . 36
2.5.4 Shor's states . . . 39

2.5.5 Subgroup states . . . 41
2.5.6 2D cluster state . . . 43
2.6 Witnessing complex states . . . 46
2.6.1 Subgroup states and their generators . . . 46
2.6.2 Complexity witness based on stabilizer witness . . . 47
2.7 Relation to quantum computation . . . 49
2.7.1 Measurement-based quantum computation (MBQC) . . . 49
2.7.2 Weaker version of the TreeBQP conjecture . . . 51
2.8 Conclusion . . . 53
2.8.1 Technical open problems . . . 54

3 Nonlocality **55**
3.1 Bell scenario . . . 55
3.1.1 No-signalling polytope . . . 56
3.1.2 Local polytope . . . 57
3.1.3 An example: the CHSH scenario . . . 58
3.2 The quantum set . . . 60

4 Dimension witness **64**
4.1 Motivation . . . 64
4.1.1 Dimension of quantum systems . . . 64
4.1.2 Dimension witnesses . . . 64
4.1.3 An overlooked feature . . . 66
4.2 The scenario . . . 67
4.3 CGLMP inequality . . . 67
4.3.1 Maximal violation of CGLMP inequality . . . 68
4.3.2 Depolarization . . . 70
4.3.3 CGLMP4 polytope . . . 71
4.4 Dimension witness with $CGLMP_4$ . . . 72
4.4.1 Upper bound on the maximum qutrit violation . . . 73
4.4.2 Lower bound on the maximum qutrit violation . . . 75
4.5 Discussion . . . 78
4.5.1 The overlooked problem . . . 78
4.5.2 Conclusion . . . 80

5 Many-box locality **81**
5.1 Motivation . . . 81
5.1.1 Quantum physics through physical principles . . . 81
5.2 Notation and definition . . . 83

5.2.1 Example: 2-box distribution . . . 85
5.2.2 Example: N-box distribution . . . 85
5.3 New tool: Fourier transformation . . . 86
5.4 On the symmetric slice . . . 89
5.4.1 Numerical analysis of the many-box local set . . . 91
5.4.2 Analytical characterisation of many-box local set . . . 93
5.4.3 Solution for any number of copies . . . 98
5.4.4 Solution for infinite copies . . . 104
5.5 On another slice . . . 107
5.6 Conclusion . . . 111

6 Conclusion and outlook **112**
6.1 Tree size complexity . . . 112
6.2 Dimension witness . . . 113
6.3 Many-box locality . . . 114
6.4 Quantum–classical boundary . . . 114

APPENDICES **115**

A Big O notation **116**

ABSTRACT

Quantum physics is probably the most successful and fascinating physical theory of the last century. Alongside with its success, quantum mechanics has some features that are less intuitive to our classical mind. This thesis examines some of these features with the concept of "size" measures: tree size complexity, dimension and many-box locality; and studies how these sizes elucidate the quantum–classical boundary.

Tree size (TS) complexity is a complexity measure of quantum states proposed by Aaronson. A (family of) state is complex if its tree size scales superpolynomial in the number of qubits. By studying a mathematical theorem that puts superpolynomial lower bound on tree size, we exhibit explicit examples of complex states, and efficient witnessing of them. Moreover, the relation between tree size complexity and quantum computation is discussed.

Dimension witnesses (DW) are tests that allow one to certify the lower bound of the dimension, the number of perfectly distinguishable states of the physical system. By violating a device independent dimension witness, one can certify the presence of states of high dimension. We discuss a device independent dimension witness for entangled four dimensional systems (ququarts) based on the CGLMP_4 inequality.

We propose the notion of many-box locality (MBL) as a possible physical principle that defines the quantum set of correlations. Novel tools are developed to analyse MBL_N, the sets of correlations that become local when N copies are measure together. The set MBL_∞ matches the quantum set on a slice of the two-party, two-input, two-outcome no-signalling polytope.

CHAPTER 1

INTRODUCTION

1.1 Quantum weirdness

> *Quantum mechanics is very much more than just a theory; it is a completely new way of looking at the world, involving a change in paradigm perhaps more radical than any other in the history of human thought.*
>
> —Tony Leggett

Quantum physics is a fascinating theory. It successfully explains what classical physics failed to, such as the photo-electric effect, the stability of the atom, the ultraviolet catastrophe; it explains virtually every phenomenon (except gravity) from elementary particles to chemistry to semiconductors. Quantum physics is at variance with classical physics both conceptually and mathematically. Quantum states reside in the complex Hilbert space; one can predict only the probability of outcomes of a measurement given by the trace of operators. This abstract formalism of quantum physics, along with its explanation power, makes predictions that are counter-intuitive to our classically trained mind. We do not experience quantum effects, such as wave particle duality, in our everyday life. Extending concepts from quantum mechanics to everyday objects might result in seemingly paradoxical situations.

In 1935, Einstein, Podolsky and Rosen questioned about the completeness of the quantum description of reality [1]. They consider a scenario where two correlated particles, A and B, are sent to two distant locations. Measurement of a variable (e.g. position) performed on the first particle A, will cause the conjugate variable (e.g. momentum) of the second particle B to be indeterminate, otherwise

Heisenberg's uncertainty relation would be violated. This "spooky action" cannot be explained by signals sent from A to B, since an instantaneous signal would violate Einstein's special relativity. They then concluded that quantum mechanics is not complete, and there might be some "hidden variables" shared by the particles that determine the outcome of all possible measurements, whether performed or not. This has remained as a philosophical dispute until 1964, when John Bell came up with an experimental testable criterion for the local realistic assumption in the EPR argument [2]. It is formulated as linear inequalities on the probability distribution of the outcomes, called *Bell inequalities*. Quantum mechanics does allow the violation of Bell inequalities hence it must violate at least one of the assumptions of locality and realism. The probability distribution of measurement statistics that leads to a Bell violation is called *Bell nonlocal*, or simply *nonlocal*.

Numerous experiments have confirmed the violation of Bell's inequalities [3–9][1]. Nonetheless, the particles that lead to a Bell violation are usually photons or ions, under stringent laboratory conditions. Can a Bell inequality be violated with macroscopic objects?

The EPR paper had also inspired one of the founders of quantum mechanics to question the interpretation of quantum mechanics with his cat gedankenexperiment. In 1935, Erwin Schrödinger raised the problem of extending the superposition principle to everyday object [13, 14]. In his thought experiment, a cat was kept in a sealed box with a flask of poison possibly triggered by the radioactive decay of a single atom. At the time of the half life of the atom, the atom is in the equal superposition of decayed and not decayed. The cat, the flask of poison and the atom will be in an entangled state of $\frac{1}{\sqrt{2}}(|\text{decayed}\rangle\,|\text{released}\rangle\,|\text{dead}\rangle + |\text{not decayed}\rangle\,|\text{not released}\rangle\,|\text{alive}\rangle)$. Can a macroscopic object, like a cat, be in the superposition of dead and alive? A lot of experimental effort has been devoted to the realization of some forms of Schrödinger's cat, that is putting something *big* in superposition. Different notions of size were used [15–17], including amplitude of the coherent state, mass of molecule in an interference experiment, magnitude of current in the superconducting circuit. In order to compare across different setups, one of these size measures aims to put bounds on the parameter in collapse models [16]. Though the size is still far from a cat, it seems we can put bigger and bigger objects into superposition. Is there a limit on the "size" of the objects where superposition will persist?

[1] Previous experiments are still open to one or more of the loopholes, including the locality loophole and detection loophole. Loopholes have been closed in different experiment, see for example [10, 11]. Only recently, a loophole free Bell test is reported [12].

Traditionally, the quantum-to-classical transition is explained through decoherence [18] and collapse models [19, 20]. This thesis takes a different approach and focuses on two aspects of quantum weirdness: complexity and nonlocality. Three size measures are introduced: tree size complexity, dimension of quantum system and many-box distributions. Besides motivations in quantum information, we will try to discuss how these sizes are related to the quantum–classical boundary.

1.2 Complexity

Quantum mechanics has been tested by numerous experiments with no disagreement. Nonetheless, straightforward application of the laws of quantum mechanics to everyday objects may lead to counter-intuitive paradoxes. Among others, the thought experiment of Schrödinger's cat [13] is such a paradox where a macroscopic object (a cat) could be put into a superposition of two distinct states (dead and alive). First conceived to describe the microscopic world of atoms and electrons, the predictions of quantum mechanics have been tested with objects of increasing sizes: motional state of atom [21], optomechanical systems [22], superconducting qubits [23], large molecules [24] and so on. These experiments can be interpreted as Schrödinger's cats of increasing size. However, a cat is not only large in physical size, large in mass, consisting of many elementary particles, but also it is a complex object. We may treat complexity as another axis to test the limit of quantum mechanics.

We propose "complexity" as an alternative "size" measure of the Schrödinger's cat: a cat is not only big, but also complex. The proposed cat state usually takes two forms: for spin systems, it takes the form of the Greenberger–Horne–Zeilinger (GHZ) state [15]; for double slit type of interference experiments, it involves the superposition of the single degree of freedom of the centre of mass. If we are able to generate larger and larger quantum states, can we generate more and more complex quantum states? To qualitatively discuss complexity, we need to introduce a complexity measure.

In Chapter 2, we will explore a complexity measure for quantum states, called the tree size, its property and its possible relation with quantum computation, a promising application of quantum information technology.

Information is physical [25]. The physical property of the information carrier determines how we can manipulate the information. The quantum mechanical laws of nature may empower us with computation capability stronger than what is allowed by classical physics. Quantum computation is computation with the aid of quantum states and operations. Some quantum algorithms are significantly faster

than all known classical algorithms. But where exactly the power of quantum computing lies remains unclear. We phrase the question from another angle: what are the states that are *useless* for quantum computation. One possible answer is those that can be efficiently simulated by a classical computer. We called those states *simple*, otherwise *complex*. With the tree size complexity, we discuss in more detail simple and complex states, how we can certify the complexity of a quantum state and evidences that at least in some models of quantum computation, complex quantum states are necessary.

1.3 Nonlocality

Bell's theorem states that violation of Bell inequalities implies that no physical theory of local hidden variables can reproduce all the predictions of quantum mechanics. It is arguably one of the most profound scientific discoveries and has deeply influenced our perception and understanding of physics. It has been the subject of analysis, discussion and development by both physicists and philosophers of science. Readers may refer to Ref. [26] for a review on the foundational perspective of nonlocality.

Other than foundational interest, recently nonlocality is considered from a operational perspective where its relation with quantum information science are investigated [27]. To name a few, nonlocality has led to applications in various device independent (DI) quantum information processing tasks, such as quantum key distribution [28, 29], randomness generation [30, 31], as well as certification of various quantum resources, including device independent entanglement witness [32], device independent dimension witness [33], and device independent tomography, better known as self-testing [34–37].

In this thesis we study two size measures related to nonlocality: dimension of quantum systems and many-box distributions.

In Chapter 3, we will review some basic tools used in the study of nonlocality. The main object under study is the space of probability distributions. In this geometric view, sets of distributions subject to certain conditions are seen as geometric bodies. These include the no-signalling polytope, the local polytope and the quantum set. Bell inequalities are simply the facets of the local polytope. Materials in Chapter 3 are not new but they introduce the framework and vocabulary for the next two chapters.

In Chapter 4, we are going to explore the power of Bell inequality serving as a device independent dimension witness (DIDW). Dimension refers to the number of perfectly distinguishable states in a physical system. In quantum mechanics, this

is the dimension of the Hilbert space that describes the system. A lower bound on the dimension can be certified by the violation of a Bell inequality. This dimension witness certifies lower bound on the dimension of the system while making no assumption about the state of the system or the measurement performed. Other forms of dimension witness rely on some assumptions or do not distinguish classical from quantum dimension. Towards the end of the chapter, we discuss a feature of the dimension witness we considered, that had been overlooked and that makes it superfluous to consider an experimental realization.

Quantum mechanics is not the most nonlocal theory under the no-signalling condition. Whether there is a physical principle that limits the amount of nonlocality that can be achieved by quantum mechanics remains an open question. On the other hand, nonlocality may diminish in the macroscopic limit. To exhibit nonlocality with quantum states, it usually requires careful preparation of the states, precise manipulation and measurements on microscopic systems. Can nonlocality be macroscopic? Can a negative answer to this question lead to a principle that defines the set of correlation achievable by quantum mechanics?

In Chapter 5, we will analyse how nonlocality diminishes or remains under a local data processing. Similar to the previously proposed principle of Macroscopic Locality (ML), we consider the local processing of combining several copies of the same distribution (boxes) and measuring collectively. The set of correlation defined by ML alone is know to be larger than the set of quantum distributions. By removing some assumptions in ML, we examine the sets of correlations whose nonlocality vanishes when N copies of the boxes are measured, the many-box local set MBL_N. With novel mathematical technique, we characterize these MBL_N sets on slices of the no-signalling polytope, numerically and analytically. On a specific slice, we show that the MBL_∞ coincides with the quantum set. This suggest that MBL could be a candidate of a physical principle that defines the quantum set.

In Chapter 6, we summarize the results presented in this thesis and an outlook on possible future directions.

CHAPTER 2

TREE SIZE COMPLEXITY

The content of this Chapter has been partially published in References [38], [39] and [40].

2.1 Motivation

2.1.1 An alternative size for Schrödinger's cat

As mentioned in the Introduction, one of the motivations of studying state complexity is to use it as an alternative "size" for Schrödinger's cat. Schrödinger's *gedankenexperiment* was meant to question the validity of quantum linear superposition at a macroscopic scale. As such, a "cat state" must contain *superposition* of *macroscopically* distinct states. Different approaches have been proposed to capture the notion of macroscopicity. For example, in a double slit interference experiment [24, 41], one can denote the state as $|\Psi\rangle = c_0 |\Psi_0\rangle + c_1 |\Psi_1\rangle$ with

$$
\begin{aligned}
|\Psi_0\rangle &= |x_1\rangle \otimes |x_2\rangle \otimes \cdots \otimes |x_n\rangle \,, \\
|\Psi_1\rangle &= |x_1 + \Delta x\rangle \otimes |x_2 + \Delta x\rangle \otimes \cdots \otimes |x_n + \Delta x\rangle \,,
\end{aligned}
$$

where x_i is the position of the particle if it pass through the left slit, and $x_i + \Delta x$ the right slit. One may call it a macroscopic superposition when the number of particles n is large. However, one soon realized that the number of particles depends on what is considered elementary. Do we count the number of molecules, atoms or nucleons? Moreover, under a change of basis, the number of degrees of freedom in the superposition can be small. If one choose to specify the state by the position of the centre of mass x_{cm} and the position of the other $(n-1)$ particles

relative to the centre of mass x_i', then $|\Psi_0\rangle$ and $|\Psi_1\rangle$ become

$$|\Psi_0\rangle = |x_{cm}\rangle \otimes |x_2'\rangle \otimes \cdots \otimes |x_n'\rangle\,,$$
$$|\Psi_1\rangle = |x_{cm} + \Delta x\rangle \otimes |x_2'\rangle \otimes \cdots \otimes |x_n'\rangle\,,$$

and the state becomes

$$|\Psi\rangle = c_0\,|\Psi_0\rangle + c_1\,|\Psi_1\rangle = \Big(c_0\,|x_{cm}\rangle + c_1\,|x_{cm} + \Delta x\rangle\Big) \otimes |x_2'\rangle \otimes \cdots \otimes |x_n'\rangle\,,$$

where now the superposition is only in a single degree of freedom. How macroscopic is this superposition?

For discussion about macroscopicity of quantum states and systems, we would like to refer the readers to the topical review by Leggett [42], as well as various ways proposed to quantify macroscopicity based on: overlap between states of individual subsystems [43]; quantum metrology [44]; number of measurements to collapse into one branch of the superposition [45]; number of single particle operation to map one state onto the other [46]; phase space structures of quantum states [47]; quantum Fisher information [15] and minimal modification of quantum mechanics [16, 48]. Experiments have demonstrate progressively macroscopic quantum systems and there does not seem to be a breakdown of quantum physics at the macroscopic scale[1].

Macroscopicity is one of the possible axes to test the limitation of quantum mechanics. We propose complexity as an alternative axis due to its possible relation to a promising application of quantum information — quantum computing. Whether a quantum computer is eventually possible depends on if quantum mechanics breaks down in the limit of high complexity. One of the original motivations for Aaronson to propose tree size complexity measure is to refine the discussion of this possible break down [49].

2.1.2 Quantum computation

Some problems become intractable on conventional classical computers, not because they are insoluble, but because the resources required to solve them are immeasurable. The promise of quantum computers is to enable quantum algorithms that renders some of these problems feasible [50].

[1] See Table I in [16] for the macroscopicity μ of various conceivable experiments with μ up to 23.3, while an idealized model of Schrödinger's cat has a μ of ~ 57, where $\mu = \log_{10}(\tau)$ and τ is the coherence time parameter in units of seconds

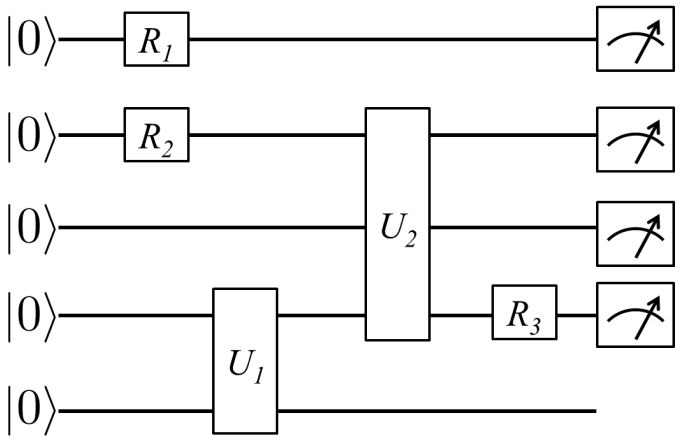


Figure 2.1: An example of a quantum circuit. Wires represent qubits and boxes represent unitary transformations.

A quantum computer can be modelled by a set of wires and gates, as shown in Fig. 2.1, called the quantum circuit model. Wires represent qubits carrying quantum information, while gates represent unitary transformations on qubits. The state is initially prepared in the product state $|0\rangle^{\otimes N}$, and the input to the problem and the algorithm is encoded in the transformation of qubits. Finally, measurement is performed on the qubits and the solution is read out from the classical outcomes of the measurements.

Another model of quantum computation, called the measurement-based quantum computation (MBQC) model, though conceptually different, is equivalent to the quantum circuit model. In MBQC, an initial resource state with multipartite entanglement is prepared, after which single qubit measurement is performed sequentially. Measurement basis of subsequent qubits may depend on the outcome of previous measurements. The input to the problem and the algorithm is encoded in the choice of measurement basis. Finally the output to the problem is encoded in the outcomes of measurements of a subsets of qubits.

There are three main types of quantum algorithms that have shown advantages over known classical algorithms [50]: quantum search algorithm, quantum Fourier transform and quantum simulation. It is not obvious to pinpoint what gives quantum computation its power. One may naïvely think that *quantum superposition*, where all the possible solutions are checked simultaneously, captures the power of quantum computation. However, if this were the case, solution to any problem could be sped up by this manner, while quantum algorithms show advantage over classical ones for only few classes of problems: only problems where the "bad" solutions miraculously destructively interfere with each other, for example

the Deutsch-Jozsa algorithm. The presence of interference seems to be necessary in this case. A quantitative measure of interference and its role in quantum algorithms is investigated in [51, 52].

Entanglement, is at the heart of many quantum information applications, and its role in quantum computing has also been extensively explored. The conclusion heavily depends on the measure of entanglement chosen. For example, quantum computation with state which has polynomial Schmidt rank across any bipartition at each step can be efficiently simulated by classical computer [53]; universal quantum computation is possible with a pure state whose entanglement entropy of every bipartition may tend to zero [54]. More surprisingly, in the measurement-based quantum computation, high geometric measure of entanglement in the resource state precludes it from offering universal quantum computational speed up [55]. For quantum computation with pure states, multipartite entanglement with number of parties increases unboundedly with input size is necessary for exponential speed up [56]. The Gottesman-Knill theorem clearly shows that (even multipartite) entanglement is not sufficient for quantum speed up [57]. For the case of mixed states, in the absence of entanglement, discord is proposed to capture the computational resource in the model of deterministic quantum computation with one qubit (DQC1) [58, 59].

Entanglement, initially devised to capture the non-separability of quantum states may not represent the resource for quantum computation speed up. Instead, we turn to state complexity which might be more directly related to quantum computation.

2.1.3 Types and measures of complexity

Everyone has an intuitive notion of complexity. Complex systems all consist of many parts and they interact in a non-trivial way. We refer to the type of "organised complexity" according to Weaver [60]. A fully ordered object, like a crystal, is simple. On the other end of the spectrum, a completely random object, like gas, is also simple (or "disorganised complexity"). In the middle of spectrum is where the complex objects are, like a human brain. Complexity has precise meanings in different fields of science. In complexity system science, complexity is about how the interaction between the different components of the system give rise to organised but hard-to-predict behaviours. In information science, complexity is a measure of resources. For example, computation complexity describes how much time, memory space, communication or oracle access is required to complete a computation task. Here, we are about to explore state complexity, which is the

minimum amount of information required to either describe or simulate a quantum state.

The complexity of classical strings can be measured by the Kolmogorov complexity [61]: the length of the shortest programme that generates the string. Its generalization to quantum Kolmogorov is not straightforward, and various definitions have been proposed: quantum computer programme to generate classical strings [62] or classical description of quantum strings [63] or quantum input to generate quantum strings [64]. Quantum Kolmogorov complexity is useful in some context but suffers from the same setback as its classical counterpart that they are not computable.

Circuit complexity [65–70] is another possible candidate for the complexity of quantum states. The circuit complexity of a state $|\psi\rangle$ is defined as the smallest quantum circuit that produced $|\psi\rangle$ from a product state. This captures the preparation complexity of the quantum state but has little to do with the description or simulation complexity. A complex state can be generated with a simple circuit [71].

Tree size (TS) introduced by Aaronson [49] captures the shortest bra-ket notation of a quantum state. It has two noticeable advantages. It is in principle *computable*, and non trivial *lower bounds* can be shown for family of states. We focus on this tree size complexity measure and extend the work of Aaronson. The rest of this Chapter is organized as follows: In Sec. 2.2, we will introduce the concept of tree size complexity and some basic properties; in Sec. 2.3, we will identify the most complex two-, three- and four-qubit states; Sec. 2.4 gives examples of simple states while Sec. 2.5 introduces superpolynomial lower bound on tree size, and gives examples of complex states; Sec. 2.6 shows how to efficiently verify complex states using a complexity witness; Sec. 2.7 discusses the possible link between tree size complexity and the usefulness of states for quantum computation; and we conclude and suggest future direction in the last section.

2.2 Definition and basic properties of tree size

The basic units for quantum information are qubits, analogous to bits for classical information. Naïvely, an n qubit quantum state requires 2^n complex amplitudes to fully specify the state:

$$|\Psi\rangle = \sum_{x\in\{0,1\}^n} c_x |x\rangle , \tag{2.1}$$

where x is an n bit string, $|x\rangle$ is a vector in the computational basis of an n qubit state, and c_x are complex coefficients that satisfy the normalization condition

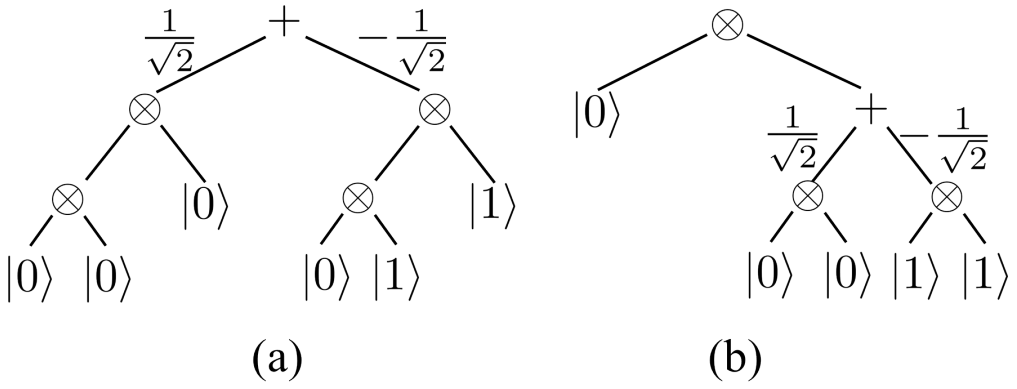


Figure 2.2: An example of tree representations of the biseparable state $|\psi\rangle = (|000\rangle - |011\rangle)/\sqrt{2}$. Representation (a) has a size of 6 and representation (b) has a size of 5, which is the minimum for this state. Hence $\mathrm{TS}(|\psi\rangle) = 5$.

$\sum |c_x|^2 = 1$. This is not the shortest representation. For small number of qubits, there exist *ad hoc* decompositions that are more compact. For two qubit this is the Schmidt decomposition. For three qubits, the Acín decomposition [72] provides a form with four real and one complex parameters. Generalization to more qubits is also possible [73]. Tree size (TS) a complexity measure of quantum states that captures the minimum number of parameters that is required to specify the state, first introduced by Aaronson [49]. Here we rephrase its definition:

Definition 2.1. *Any multiqubit state written in Dirac's bra-ket notation can be represented with a rooted tree consists of* $\otimes$ *and* $+$ *gates; each leaf vertex is labelled with a single qubit state* $\alpha|0\rangle + \beta|1\rangle$ *which need not be normalized; the outgoing edges of each* $+$ *gate are labelled with a complex number representing the amplitude of each component. The size of the tree is defined as the number of leaves.*

Finally, the tree size (TS) *of a quantum state is defined as the size of the* minimum *tree representation of the state.*

We illustrate the definition of tree size with the following example. Let $|\psi\rangle = (|000\rangle - |011\rangle)/\sqrt{2}$. One way to represent $|\psi\rangle$ is $\frac{1}{\sqrt{2}}|0\rangle \otimes |0\rangle \otimes |0\rangle + (-\frac{1}{\sqrt{2}})|0\rangle \otimes |1\rangle \otimes |1\rangle$, with size 6 (see Fig. 2.2(a)); or the shortest possible representation, $|0\rangle \otimes (\frac{1}{\sqrt{2}}|00\rangle + (-\frac{1}{\sqrt{2}})|11\rangle)$, with size 5 (see Fig. 2.2(b)).

This measure of complexity is in principle *computable*, though we lack efficient algorithms. It can always be exhaustively computed because tree size is both lower and upper bounded. A trivial lower bound for an n qubit is state is n. The most trivial upper bound is $n2^n$ if we expand the state in the computational basis. A less trivial upper bound is shown below:

Proposition 2.2. *The tree size of an n qubit state is trivially bounded by*

$$n \leq \mathrm{TS}(|\psi_n\rangle) \leq 3 \cdot 2^{n-1} - 2. \tag{2.2}$$

Proof. The upper bound can be shown by repeated Schmidt decomposition. We can single out the first qubit of the n qubit state,

$$|\psi_n\rangle = |0\rangle\,|\psi^0_{n-1}\rangle + |1\rangle\,|\psi^1_{n-1}\rangle\,, \tag{2.3}$$

so we have $\mathrm{TS}(|\psi_n\rangle) \leq 2 \cdot \mathrm{TS}(|\psi_{n-1}\rangle) + 2$. By the fact that tree size of a single qubit state $\mathrm{TS}(|\psi_1\rangle) = 1$, solving this recursive relation we have $\mathrm{TS}(|\psi_n\rangle) \leq 3 \cdot 2^{n-1} - 2$. □

This upper bound is still exponential in n, the number of qubits. In computer science, a polynomial scaling is often considered as tractable, conversely a super-polynomial scaling is considered intractable. Following this idea and recall that tree size represents the shortest bra-ket description of a state, we call a state with polynomial tree size simple, and a state with superpolynomial tree size complex. Besides being computable in principle, another advantage of tree size complexity measure is that superpolynomial lower bound can be derived for some states, and this will the subject of Sec. 2.5. Before studying tree size for many qubit states, let us examine the tree size for few-qubit states.

2.3 Tree size of few-qubit states

In this section, we will examine the tree size of states consisting of two, three or four qubits. We will identify the most complex state, its minimal tree representation as well as their ϵ-approximate tree size.

Entanglement, especially multipartite entanglement, is essential for a state to be complex. Tools from entanglement classification can be utilized to study the tree size classification of few qubit states. Formally speaking, we have the following proposition [38],

Proposition 2.3. *If $|\psi\rangle = A_1 \otimes \cdots \otimes A_n\,|\phi\rangle$, where all the single-qubit operators A_i are invertible, then* $\mathrm{TS}(|\psi\rangle) = \mathrm{TS}(|\phi\rangle)$.

Any two states that can be transformed to each other by invertible local operators (ILOs) as above are said to be equivalent under stochastic local operation and classical communication (SLOCC). The above proposition implies that all states in a SLOCC equivalent class have the same TS. The converse is not necessarily

true: two states can have the same tree size while belonging to different SLOCC equivalent classes, for example, states biseparable with respect to different partitioning of the parties. In this sense, tree size for few qubits can be seen as a coarse graining of SLOCC classification.

For the number of qubits $n = 2, 3, 4$, we will first find an upper bound on the tree size by an explicit form. By listing down all the rooted trees with size less or equal to that bound, we find the tree size of a representative states from each SLOCC equivalence class. By the virtue of Prop. 2.3, then all the states in the same SLOCC class will have the same tree size.

2.3.1 Two qubits

Any two-qubit state can be written in the Schmidt decomposition as [50]:

$$|\psi\rangle = c\,|0\rangle \otimes |0\rangle + s\,|1\rangle \otimes |1\rangle\,, \tag{2.4}$$

where c and s are nonnegative real numbers satisfying $c^2 + s^2 = 1$, and $\{|0\rangle\,, |1\rangle\}$ form an orthonormal basis. The state is said to be separable if one of the coefficients c or s vanishes and entangled otherwise. The Schmidt decomposition has size at most 4, hence the TS of any two-qubit state is at most 4. There are only two different rooted trees of size at most 4 that can describe a two-qubit state, which are shown in Fig. 2.3.

A general two qubit state can be written as:

$$|\psi\rangle = c_{00}\,|00\rangle + c_{01}\,|01\rangle + c_{10}\,|10\rangle + c_{11}\,|11\rangle\,, \tag{2.5}$$

from which we can extract the coefficient matrix:

$$C = \begin{pmatrix} c_{00} & c_{01} \\ c_{10} & c_{11} \end{pmatrix}. \tag{2.6}$$

It is straightforward to check that the state is entangled if and only if (iff) $\det(C) \neq 0$, i.e. $c_{00}c_{11} \neq c_{01}c_{10}$. Thus $\det(C) \neq 0$ implies the state is entangled and cannot be represented by Fig. 2.3(a), hence $\mathrm{TS}(|\psi\rangle) = 4$; while $\det(C) = 0$ implies the state is a product state so $\mathrm{TS}(|\psi\rangle) = 2$. This concludes the case for two qubits.

2.3.2 Three qubits

For three qubits, a useful decomposition that has a similar role as the Schmidt decomposition does for two qubits is the canonical form derived by Acín *et al.*

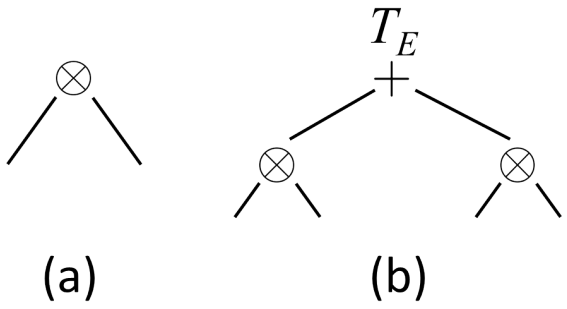


Figure 2.3: Possible rooted trees with size at most 4 for two-qubit states.

[72]: Any three-qubit state can be written as

$$|\psi\rangle = \cos\theta\,|000\rangle + \sin\theta\,|1\rangle\,(\cos\omega\,|0'0''\rangle + \sin\omega\,|1'1''\rangle)\,, \tag{2.7}$$

where the prime and double prime indicate different bases. The TS is upper bounded by the size of this decomposition and thus is at most 8.

It is known that three qubit pure states can be categorized into six different SLOCC equivalent classes: the product class, three biseparable classes due to permutation, the GHZ class and the W class [74]. Examples of states in these classes are, respectively,

$$\begin{aligned} |\mathrm{P}\rangle &= |000\rangle\,, \\ |\mathrm{B}\rangle &= \frac{1}{\sqrt{2}}\,|0\rangle\,(|01\rangle + |10\rangle)\,, \\ |\mathrm{GHZ}\rangle &= \frac{1}{\sqrt{2}}\,|000\rangle + |111\rangle\,, \\ |W\rangle &= \frac{1}{\sqrt{3}}\,(|001\rangle + |010\rangle + |100\rangle)\,. \end{aligned}$$

So, a state $|\psi\rangle$ is said to be in a particular SLOCC class, say the W class, if there exist ILOs A_1, A_2, A_3 such that $|\psi\rangle = A_1 \otimes A_2 \otimes A_3\,|W\rangle$.

Classifying a general three qubit state

$$\begin{aligned} |\psi\rangle =& c_{000}\,|000\rangle + c_{001}\,|001\rangle + c_{010}\,|010\rangle + c_{011}\,|011\rangle \\ &+ c_{100}\,|100\rangle + c_{101}\,|101\rangle + c_{110}\,|110\rangle + c_{111}\,|111\rangle \end{aligned} \tag{2.8}$$

into one of the SLOCC classes is done by examining the coefficient matrices:

$$C_0 = \begin{pmatrix} c_{000} & c_{001} \\ c_{010} & c_{011} \end{pmatrix},\ C_1 = \begin{pmatrix} c_{100} & c_{101} \\ c_{110} & c_{111} \end{pmatrix}. \tag{2.9}$$

Notice that we have chosen the partition $1|23$ in writing the matrices in this way. Since identifying product and biseparable states are rather obvious, we shall focus on telling apart the GHZ and the W class. We now state the condition for a state to be in the GHZ or the W class, first derived by Lamata *et al.* in Ref. [75]:

Proposition 2.4. *Let $|\psi\rangle$ be a three-qubit pure state. Then*

1. *$|\psi\rangle$ is in the GHZ class iff one of the following conditions holds:*
   (a) *There is a partition $i|jk$ for which C_0 and C_1 are linearly independent, $\det(C_0) \neq 0$, and $C_0^{-1}C_1$ has two distinct eigenvalues.*
   (b) *Same as (a) but with C_0 and C_1 exchanged.*
   (c) *For* all *partitions $i|jk$, C_0 and C_1 are linearly independent, and there is a partition such that $\det(C_0) = \det(C_1) = 0$.*
2. *$|\psi\rangle$ is in the W class iff one of the following conditions holds:*
   (a) *There is a partition $i|jk$ for which C_0 and C_1 are linearly independent, $\det(C_0) \neq 0$, and $C_0^{-1}C_1$ has only one eigenvalue.*
   (b) *Same as (a) but with C_0 and C_1 exchanged.*

Notice that the eigenvalue equation of a 2×2 matrix M is $\lambda^2 - \mathrm{tr}(M)\lambda + \det(M) = 0$. Hence M has only one eigenvalue when $(\mathrm{tr}(M))^2 - 4\det(M) = 0$, otherwise M has two distinct eigenvalues. A generic state of three qubits will satisfy (1a) or (1b) of Proposition 2.4, since the others require the coefficients to satisfy an equation, restricting the coefficients to a set of measure zero. As a result, most of the three qubit states belong to the GHZ class.

A related observation is that the W class is *unstable* in a sense that an arbitrarily small perturbation in a certain direction will make a state in the W class into the GHZ class. This is the basis for finding the ϵ-tree size of W states that we will introduce later in Definition 2.6. Consider an infinitesimal change μ in the direction of $|111\rangle$ to $|W\rangle$. For $|W\rangle$, the coefficient matrices are:

$$C_0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \; C_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \tag{2.10}$$

$\det(C_0) \neq 0$ and $M = C_0^{-1}C_1$ has one eigenvalue, satisfying condition (2a) of Prop. 2.4. After the perturbation $|W\rangle \to |W\rangle + \mu\,|111\rangle$, C_1 becomes

$$C_1 = \begin{pmatrix} 1 & 0 \\ 0 & \mu \end{pmatrix}, \tag{2.11}$$

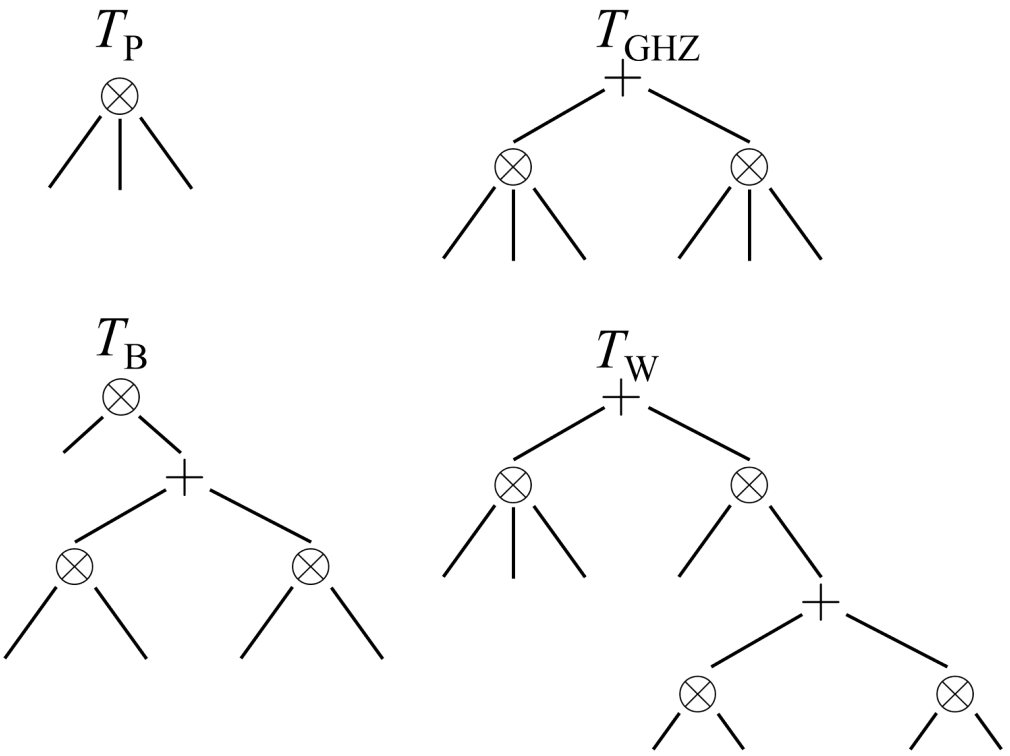


Figure 2.4: Rooted trees of three qubits with size at most 8. There are four of them, denoted as T_{P}, T_{GHZ}, T_{B} and T_W respectively.

and now $M = C_0^{-1}C_1$ has two distinct eigenvalues. By (2a) of Proposition 2.4, $|W\rangle + \mu\,|111\rangle$ is in the GHZ class. In fact any state from the W class, $|\phi_W\rangle = A_1 \otimes A_2 \otimes A_3\,|W\rangle$, the effect of a small perturbation

$$|\phi_W\rangle \to |\phi_W\rangle + \mu A_1 \otimes A_2 \otimes A_3\,|111\rangle = A_1 \otimes A_2 \otimes A_3(|W\rangle + \mu\,|111\rangle), \qquad (2.12)$$

results in a state in the GHZ class.

Next we find the tree size of W state and GHZ state by exhausting trees that have size less or equal to 8. In total there are four types of trees that have at most 8 leaves, listed in Fig. 2.4.

Proposition 2.5. *The most complex states of three qubits are the states in the W class, and they have tree size 8.*

Proof. The proof is by ruling out the smaller trees in Fig. 2.4 as possible representation of the W state. First we notice that T_{P} is a special case of T_{B} with one branch of the $+$ gate having weight 0; similarly, T_{B} can be cast into the form of T_{GHZ} if we distribute the top $\otimes$ gate through the $+$. Let us denote a general state

in the form of T_{GHZ},

$$|T_{\text{GHZ}}\rangle = \bigotimes_{i=1}^{3}(x_i|0\rangle + y_i|1\rangle) + \bigotimes_{i=1}^{3}(x_i'|0\rangle + y_i'|1\rangle), \tag{2.13}$$

where x_i, y_i, x_i' and y_i' are some complex coefficient. It is straightforward to see that there is no value for x_i, y_i, x_i' and y_i' such that $|T_{\text{GHZ}}\rangle = |W\rangle$. Therefore, $|W\rangle$ cannot be represented by $|T_{\text{GHZ}}\rangle$, ruling out all those trees in Fig. 2.4 except T_W. Hence, $\text{TS}(|W\rangle) = \text{TS}(T_W) = 8$. □

The second complex state after $|W\rangle$ is the $|\text{GHZ}\rangle$, with size 6. The smaller trees T_{B} and T_{P} do not contain tripartite enlargement cannot describe $|\text{GHZ}\rangle$. Finally, $\text{TS}(|\text{B}\rangle) = 5$, and $\text{TS}(|\text{P}\rangle) = 3$.

Approximate tree size

In practice, states are determined only up to finite precision, it is necessary to consider a more physical definition of tree size that allow some small uncertainty — the ϵ-tree size (TS_ϵ) [38, 49]:

Definition 2.6. *For $0 \leq \epsilon < 1$, the ϵ-tree size of a state $|\psi\rangle$ is defined as*

$$\text{TS}_\epsilon(|\psi\rangle) = \min_{|\langle\phi|\psi\rangle|^2 \geq 1-\epsilon} \text{TS}(|\phi\rangle). \tag{2.14}$$

Due to the instability of $|W\rangle$ we mentioned before, we can see that $\text{TS}_\epsilon(|W\rangle) = 6$ for small $\epsilon > 0$. We can find the largest ϵ where $\text{TS}_\epsilon(|W\rangle) = 6$ by finding the largest overlap between a biseparable state and the W state. Let us define

$$|T_{\text{B}}\rangle = |u\rangle\,(\alpha|0\rangle|v\rangle + \beta|1\rangle|w\rangle), \tag{2.15}$$

with $|\alpha|^2 + |\beta|^2 = 1$ and $|u\rangle, |v\rangle, |w\rangle$ some normalized single qubit states. Since the W state is invariant under permutation of qubits, we do not need to consider the other partitions. We can see that,

$$\max|\langle W|T_{\text{B}}\rangle|^2 = \frac{2}{3}, \tag{2.16}$$

with the maximum is achieved by for example $|T_{\text{B}}\rangle = |\text{B}\rangle$ as in Eqn. 2.8. Thus $\text{TS}_\epsilon(|W\rangle) = 6$ for $0 < \epsilon < \frac{1}{3}$.

2.3.3 Four qubits

As in the case for three qubits, SLOCC equivalent classes can be used to find the tree size of four-qubit states. First, we introduce the concept of irreducible $A|BCD$ form and show that state of that form is in fact the most complex four-qubit state. Then we show the best tree representation of these states has 16 leaves, in an unusual form. Incidentally, these states belong to a entanglement class that was not described in previous classifications [76]. Finally, we discuss the ϵ tree size of these states.

Irreducible $A|BCD$ form

We begin with the observation that any four qubit state can be written as:

$$|\psi\rangle = |0\rangle |\phi_0\rangle + |1\rangle |\phi_1\rangle , \tag{2.17}$$

with $|\phi_0\rangle$ and $|\phi_1\rangle$ being some arbitrary three qubit state. We called this the $A|BCD$ form. The size of this decomposition is at most 18, since $|\phi_0\rangle$ and $|\phi_1\rangle$ have size at most 8, achieved when they are in the W class. Now we define *irreducible $A|BCD$ form* as follows:

Definition 2.7. *A state $|\psi\rangle$ is said to have the irreducible $A|BCD$ form, or simply to be irreducible, if for all $A \in \{1,2,3,4\}$ and all ILOs A_1, the $A|BCD$ form of $A_1 |\psi\rangle$ always has $|\phi_0\rangle$ and $|\phi_1\rangle$ in the W class.*

Note that in the above definition, we allow all permutations of parties and ILOs on the first party, that is equivalent to $A_1 \otimes A_2 \otimes A_3 \otimes A_4 |\psi\rangle$ having $|\phi_0\rangle$ and $|\phi_1\rangle$ in the W class for all one against three partition, and any ILOs $A_1 \otimes A_2 \otimes A_3 \otimes A_4$.

Conditions to be $A|BCD$ irreducible

Let us compute the conditions on the coefficients for which the state will be irreducible. First we pick a certain one-against-three partition, and apply ILO on the last three qubits to transform $|\phi_1\rangle$ to a $|W\rangle$ state, resulting in:

$$|\psi'\rangle = A_2 \otimes A_3 \otimes A_4 |\psi\rangle = |0\rangle |\phi_W\rangle + |1\rangle |W\rangle , \tag{2.18}$$

where $|\phi_W\rangle$ is a state in the W class with coefficient matrices

$$C_0 = \begin{pmatrix} c_1 & c_2 \\ c_3 & c_4 \end{pmatrix}, \; C_1 = \begin{pmatrix} c_5 & c_6 \\ c_7 & c_8 \end{pmatrix}. \tag{2.19}$$

We can always set $c_2 = 0$ by applying A_1 such that $A_1 |0\rangle = |0\rangle$ and $A_1 |1\rangle = -c_2 |0\rangle + |1\rangle$.

Next we apply a ILO to the first qubit

$$A_1 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \tag{2.20}$$

with

$$\det(A_1) = a_{11}a_{22} - a_{12}a_{21} \neq 0, \tag{2.21}$$

we have

$$A_1 |\psi'\rangle = |0\rangle (a_{11} |\phi_W\rangle + a_{12} |W\rangle) + |1\rangle (a_{21} |\phi_W\rangle + a_{22} |W\rangle). \tag{2.22}$$

For $|\psi\rangle$ to be irreducible we must find $|\phi_W\rangle$ such that both (i) $a_{11} |\phi_W\rangle + a_{12} |W\rangle$ and (ii) $a_{21} |\phi_W\rangle + a_{22} |W\rangle$ that remain in the W class for all a_{ij} satisfying the invertibility condition (2.21). This constraint makes sure a_{11} and a_{12} are not both zero, and neither are a_{21} and a_{22}. If $a_{11} = 0$, then $a_{12} \neq 0$, and (i) is satisfied. If $a_{11} \neq$, let $\lambda = a_{12}/a_{11}$ and (i) becomes finding $|\phi_W\rangle$ such that $|\phi_W\rangle + \lambda |W\rangle$ remains in W class for all λ. Similar arguments for a_{21} and a_{22} result in the same requirement.

Recall that we set $c_2 = 0$, the coefficient matrices of $|\phi_W\rangle + \lambda |W\rangle$ are

$$C_0 = \begin{pmatrix} c_1 & \lambda \\ c_3 + \lambda & c_4 \end{pmatrix}, \; C_1 = \begin{pmatrix} c_5 + \lambda & c_6 \\ c_7 & c_8 \end{pmatrix}. \tag{2.23}$$

Now we can check when do these coefficients satisfy condition 2 stipulated by Prop. 2.4. Condition 2(a), which requires $\det C_0 \neq 0$, cannot be satisfied by all λ, because $\det C_0 = -\lambda^2 - c_3\lambda + c_1c_4$ has at least one zero regardless of the values of c_1, c_3 and c_4. So we turn to 2(b): $\det C_1 = (c_5 + \lambda)c_8 - c_6c_7$, which is not zero for all λ if and only if $c_8 = 0$ and $c_6c_7 \neq 0$; $C_1^{-1}C_0$ having only one eigenvalue amounts to the quadratic equation, $(\mathrm{tr}[C_1^{-1}C_0])^2 = 4 \det[C_1^{-1}C_0]$:

$$a_2\lambda^2 + a_1\lambda + a_0 = 0, \tag{2.24}$$

where a_0, a_1 and a_2 are functions of the c_i's,

$$\begin{aligned} a_0 &= (c_3c_6 - c_4c_5)^2 + 4c_1c_4c_6c_7, \\ a_1 &= 2((c_3c_6 - c_4c_5)(c_6 + c_7 - c_4) - 2c_3c_6c_7), \\ a_2 &= (c_6 + c_7 - c_4)^2 - 4c_6c_7. \end{aligned}$$

For this quadratic equation to be true for all λ one has to set a_0, a_1 and a_2 to zero. Condition 2(b) also requires both C_0 and C_1 to be linearly independent. It is not hard to see that this is true for all λ if and only if $c_4 \neq 0$.

The same analysis must be carried out for the other partitions. The algebra is lengthy but straightforward. As it turns out, no new constraints on the coefficients is needed. The state $|\psi\rangle$ in Eqn. (2.17) has the $A|BCD$ irreducible form is the coefficients of $|\phi_W\rangle$ obey the following:

$$\begin{cases} c_4, c_6, c_7 \neq 0, \\ c_2 = c_8 = 0, \\ (c_6 + c_7 - c_4)^2 = 4c_6c_7, \\ (c_3c_6 - c_4c_5)(c_6 + c_7 - c_4) = 2c_3c_6c_7, \\ (c_3c_6 - c_4c_5)^2 = -4c_1c_4c_6c_7. \end{cases} \tag{2.25}$$

These constraints can be simplified if we consider separately the case $c_1 = 0$ and $c_1 \neq 0$. If $c_1 = 0$ the above constraints become:

$$\begin{cases} c_4, c_6, c_7 \neq 0, \\ c_1 = c_2 = c_3 = c_5 = c_8 = 0, \\ c_4 = (\sqrt{c_6} \pm \sqrt{c_7})^2, \end{cases} \tag{2.26}$$

where $\sqrt{z}$ denotes the principal square root of the complex number z.

If $c_1 \neq 0$, by applying A_1 on the first qubit such that $A_1 |0\rangle = -\frac{1}{c_1} |0\rangle$ and $A_1 |1\rangle = |1\rangle$ we can get a new $|\phi_W\rangle$ with $c_1 = -1$. Substituting this into the set of equations we obtain the following constraints:

$$\begin{cases} c_4, c_6, c_7 \neq 0, \quad c_1 = -1, \quad c_2 = c_8 = 0, \\ c_4 = \left(\frac{c_3}{2}\right)^2, \quad c_6 = \left(\frac{c_5}{2}\right)^2, \quad c_7 = \left(\frac{c_3 - c_5}{2}\right)^2. \end{cases} \tag{2.27}$$

A representative of irreducible states

Now let us consider one example of states in the irreducible $A|BCD$ class. One possible assignment is all $c_i = 0$ except $c_4 = 4, c_6 = c_7 = 1$, which yields the state

$$|\psi\rangle = |0\rangle\,(4\,|011\rangle + |101\rangle + |110\rangle) + |1\rangle\,(|001\rangle + |010\rangle + |100\rangle).$$

We may find an ILO that converts this state to a more symmetric form:

$$|\psi^{(4)}\rangle = \frac{1}{\sqrt{2}}(|0\rangle\,|W_0\rangle + |1\rangle\,|W_1\rangle), \tag{2.28}$$

where

$$|W_0\rangle = \frac{1}{\sqrt{6}}(|110\rangle + |101\rangle - 2\,|011\rangle),$$
$$|W_1\rangle = \frac{1}{\sqrt{6}}(|001\rangle + |010\rangle - 2\,|100\rangle),$$

or explicitly,

$$|\psi^{(4)}\rangle = \frac{1}{\sqrt{12}}\Big(\big(\,|0110\rangle + |0101\rangle + |1001\rangle + |1010\rangle\,\big) - 2\,|0011\rangle - 2\,|1100\rangle\Big).$$

This is a representative of states in the irreducible $A|BCD$ form. Incidentally, this state is the four qubit singlet state and has already been realized in a down-conversion experiment, where the genuine four-partite entanglement was confirmed by measuring an entanglement witness [77, 78]. This state is similar to the four-qubit Dicke state with two excitations, that is

$$|D_{4,2}\rangle = \frac{1}{\sqrt{6}}(|0011\rangle + |0101\rangle + |0110\rangle + |1001\rangle + |1010\rangle + |1100\rangle), \tag{2.29}$$

except for the two factors of -2. Despite this similarity, the state $|D_{4,2}\rangle$ does not have irreducible $A|BCD$ form. One can apply, for example,

$$A_1 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

on the first qubit, yielding $|0\rangle\,|\mathrm{GHZ}'\rangle + |1\rangle\,|\mathrm{GHZ}''\rangle$. Next we aim to show that a four-qubit state has maximal tree size if and only if it has irreducible $A|BCD$ form. Moreover, while a direct decomposition yields a size of 18, the tree size of a four-qubit state in that form is 16.

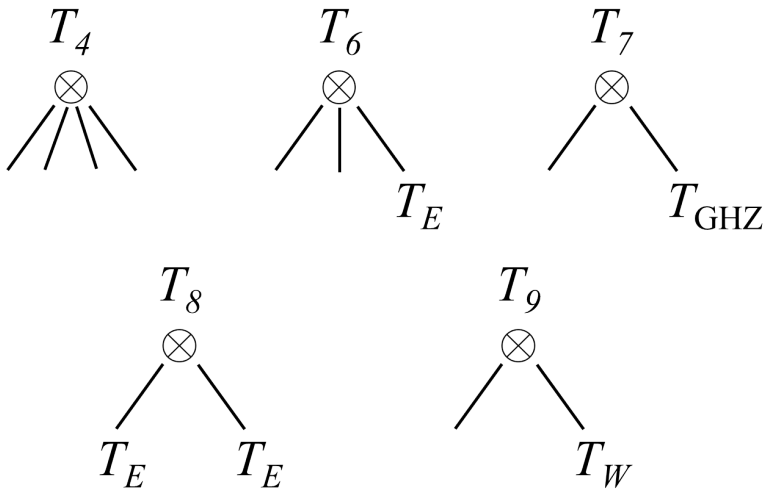


Figure 2.5: Trees for four qubit state rooted with a $\otimes$ gate. The number at the subscript indicates the size of the tree. The two qubit tree T_E and three qubit tree T_{GHZ} and T_W are used as branches to construct these trees.

Maximal tree size

First we list all the four-qubit trees rooted with a $\otimes$ as shown in Fig. 2.5. The number at the subscript indicates the size of that tree. All possible trees are combinations of these trees. We then proceed to prove that the tree size of the states with irreducible $A|BCD$ form is 16; and any state not in this form can be described by a tree with at most 15 leaves.

Proposition 2.8. *A four-qubit state $|\psi\rangle$ with the irreducible $A|BCD$ form has the following decomposition*

$$|\psi\rangle = |\phi_{12}\rangle\,|\varphi_{34}\rangle + |\phi'_{13}\rangle\,|\varphi'_{24}\rangle\,, \tag{2.30}$$

where $|\phi\rangle\,, |\varphi\rangle\,, |\phi'\rangle$ and $|\varphi'\rangle$ are two-qubit entangled states, and the subscripts denote the label of the qubits.

Here $|\phi_{12}\rangle$ is an entangled states of the first and second qubit while $|\phi'_{13}\rangle$ is an entangled states of the first and the third qubit. This tree can be represented by $T_8 + T_8$ since $|\phi\rangle\,, |\varphi\rangle\,, |\phi'\rangle$ and $|\varphi'\rangle$ are members of T_E. Note that the order of qubits in the two branches are not the same, hence this decomposition is not of the form of repeated Schmidt decomposition. This "crossing" of qubits is required to obtain the minimal tree for states with irreducible $A|BCD$ form.

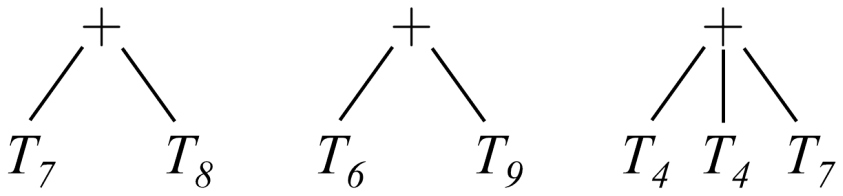


Figure 2.6: Rooted trees of four qubit states with exactly 15 leaves. All the trees with fewer than 15 leaves are special cases of these trees.

Proof. We show this by an explicit construction. For the case where $c_1 = 0$ and the others satisfying (2.25), one can verify that:

$$\begin{aligned}
|\phi_{12}\rangle &= |10\rangle \mp \sqrt{c_6}\sqrt{c_7}\,|01\rangle\,, \\
|\varphi_{34}\rangle &= \mp\frac{\sqrt{c_6}}{\sqrt{c_7}}\,|01\rangle + |10\rangle\,, \\
|\phi'_{13}\rangle &= \sqrt{c_7}(\sqrt{c_7} \pm \sqrt{c_6})\,|01\rangle + |10\rangle\,, \\
|\varphi'_{24}\rangle &= \frac{\sqrt{c_7} \pm \sqrt{c_6}}{\sqrt{c_7}}\,|01\rangle + |10\rangle\,.
\end{aligned}$$

These states are well defined and are entangled because we choose $c_6, c_7 \neq 0$.

For the case $c_1 = -1$ and the others satisfying (2.27), we have:

$$\begin{aligned}
|\phi_{12}\rangle &= \frac{4}{c_5(c_5-c_3)}\,|10\rangle + |0\rangle\left(\frac{2c_3}{c_5(c_5-c_3)}\,|0\rangle + |1\rangle\right), \\
|\varphi_{34}\rangle &= \frac{c_5^2}{4}\,|01\rangle + \left(\frac{c_5}{2}\,|0\rangle + \frac{c_5(c_5-c_3)}{4}\,|1\rangle\right)|0\rangle\,, \\
|\phi'_{13}\rangle &= |0\rangle\left(\frac{c_3}{c_3-c_5}\,|0\rangle + \frac{c_3}{2}\,|1\rangle\right) + \frac{2}{c_3-c_5}\,|10\rangle\,, \\
|\varphi'_{24}\rangle &= \frac{c_3-c_5}{2}\,|10\rangle + |0\rangle\left(|0\rangle + \frac{c_3}{2}\,|1\rangle\right),
\end{aligned}$$

which are again well defined because the constraint (2.27) requires $c_3 - c_5 \neq 0$. □

The decomposition of Eqn. (2.30) turns out to be optimal for all states with irreducible $A|BCD$ form. In other words, these states cannot have decompositions with size smaller than 16.

Proposition 2.9. *If $|\psi\rangle$ is a state with irreducible $A|BCD$ form, its minimal tree is $T_8 + T_8$. Thus, the tree size of these states is 16.*

Proof. First we list down all the trees with 15 leaves or fewer. There are plenty of them, but most are special cases of others. Let us first consider the set of trees shown in Fig. 2.5. One can check that T_4 is a special case of T_6 since the two

qubit product state is a special case of T_E (with one branch equals to 0). We denote this relation as $T_4 \subset T_6$. Similarly, one can check other relations, such as, $T_4 \subset T_6 \subset T_7 \subset T_9$, $T_6 \subset T_8$ and $T_7 \subset T_4 + T_4$. From these relations one has $T_6 + T_6 \subset T_7 + T_7 \subset T_4 + T_4 + T_7$ and so on. After listing all the trees with at most 15 leaves one finds that they are all special cases of trees with exactly 15 leaves. The set of trees with exactly 15 leaves is shown in Fig. 2.6.

We now proceed by showing that if a state is described by a tree with at most 15 leaves, it cannot have irreducible $A|BCD$ form. We show the argument for $T_6 + T_9$, the argument for the other two trees are similar. For better clarity, we can write down $T_6 + T_9$ explicitly as in Fig. 2.7. Let us denote the single qubit state of the leaf at the root of T_9 by $\gamma |u_1\rangle$, where $|u_1\rangle$ is normalized. This qubit may correspond to any branch of T_6. There are two inequivalent cases. First, this qubit is assigned to one of the two leaves at the root of T_6. Let us denote its state as $\alpha |u_0\rangle + \beta |u_1\rangle$ where $|u_0\rangle$ and $|u_1\rangle$ form a orthonormal basis. The four-qubit state described by $T_6 + T_9$ can then be written as:

$$\begin{aligned} |\psi\rangle &= (\alpha |u_0\rangle + \beta |u_1\rangle) |T_B\rangle + \gamma |u_1\rangle |T_W\rangle \\ &= \alpha |u_0\rangle |T_B\rangle + |u_1\rangle (\beta |T_B\rangle + \gamma |T_W\rangle). \end{aligned}$$

Note that $|T_B\rangle$ and $\beta |T_B\rangle + \gamma |T_W\rangle$ are both three-qubit states. In this $A|BCD$ partition, the state is reducible, since $|T_B\rangle$ is clearly not in the W class.

The second case is when the qubit $\gamma |u_1\rangle$ in T_9 corresponds to one of the qubits in $|T_E\rangle$ in T_6. Let us write $|T_E\rangle = \alpha |u_0\rangle |v_0\rangle + \beta |u_1\rangle |v_1\rangle$. Expand $|\psi\rangle$ and group term with $|u_1\rangle$, we have

$$|\psi\rangle = \alpha |u_0\rangle |v_0\rangle |\phi\rangle |\varphi\rangle + |u_1\rangle (\beta |v_1\rangle |\phi\rangle |\varphi\rangle + \gamma |T_W\rangle).$$

Again, this state in this $A|BCD$ partition is clearly reducible since $|v_0\rangle |\phi\rangle |\varphi\rangle$ is in the class T_P.

For $T_7 + T_8$ and $T_4 + T_4 + T_7$, denote the single-qubit state at the root of T_7 as $\gamma |u_1\rangle$ and similar argument shows that these states are reducible. Since all other trees with fewer than 15 leaves are special cases of these three, we can conclude that all tree with at most 15 leaves can not describe a state in the irreducible $A|BCD$ form. Therefore, $T_8 + T_8$ with size 16 is the optimal decomposition for states in the irreducible $A|BCD$ form. □

Now we have found the tree size for states in the irreducible $A|BCD$ form. What is left to show is that these states are the most complex, that is, the reducible states have tree size at most 15.

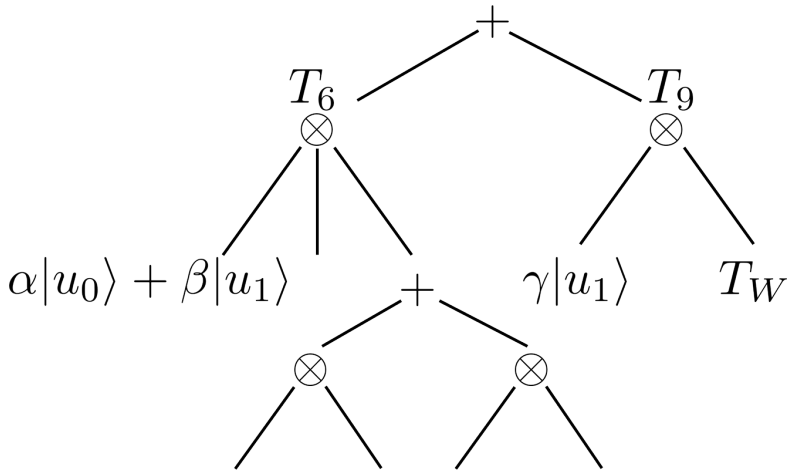


Figure 2.7: Explicit drawing of the $T_6 + T_9$ tree. Note that we have not labelled the leaves with qubits, so $\gamma\,|u_1\rangle$ at the root of T_9 can correspond to any qubit on the left branch T_6.

Proposition 2.10. *If a four-qubit state $|\psi\rangle$ does not have an irreducible $A|BCD$ form, it can be described by a tree with at most 15 leaves.*

Proof. If $|\psi\rangle$ does not have the irreducible $A|BCD$ form, there exist a partition and an ILO A_1 such that after the application of A_1 the state becomes:

$$|\psi\rangle = |0\rangle\,|\phi_0\rangle + |1\rangle\,|\phi_1\rangle\,,$$

where at least one of $|\phi_0\rangle$ or $|\phi_1\rangle$ is not in the W class, say $|\phi_1\rangle$. If $|\phi_1\rangle$ is biseparable, then clearly $|\psi\rangle$ can be described by the tree $T_6 + T_9$ with 15 leaves. If $|\phi_1\rangle$ is in the GHZ class, we can apply ILOs on the last three qubit to transform $|\psi\rangle$ into

$$|\psi\rangle = |0\rangle\,|\phi_w\rangle + |1\rangle\,|\mathrm{GHZ}\rangle\,.$$

This decomposition has size 16, but it is not optimal. Consider the state $|\phi_W\rangle + \lambda\,|\mathrm{GHZ}\rangle$, the condition for this state to remain in the W class is the $C_0C_1^{-1}$ of $C_1^{-1}C_0$ has only one eigenvalue. Both cases yield a fourth-power polynomial equation in λ, which has at most four solutions. It is always possible to find a λ^* such that it does not satisfy the equation. Then $|\phi_W\rangle + \lambda^*\,|\mathrm{GHZ}\rangle$ will not be in the W class and hence can be described by a T_{GHZ}. Applying ILO A_1 on the first qubit of $|\psi\rangle$ such that $A_1\,|0\rangle = |0\rangle$ and $A_1\,|1\rangle = \lambda^*\,|0\rangle + |1\rangle$, the state becomes

$$|\psi\rangle = |0\rangle\,\Big(\,|\phi_W\rangle + \lambda^*\,|\mathrm{GHZ}\rangle\,\Big) + |1\rangle\,|\mathrm{GHZ}\rangle\,,$$

which can be described by $T_7 + T_7$ with 14 leaves. □

Combining Proposition 2.9 and 2.10 we can conclude that the most complex four qubit states are those that have irreducible $A|BCD$ form and its tree size is 16.

Approximate tree size

We are going to discuss the ϵ-approximate tree size for four-qubit state. First recall that for arbitrarily small $\epsilon > 0$, $\mathrm{TS}_\epsilon(|\phi_W\rangle) = 6$. A consequence of this is for a state $|\psi\rangle$ in the irreducible $A|BCD$ class and arbitrarily small $\epsilon > 0$, we can always find a

$$|\phi\rangle = |0\rangle\,|\mathrm{GHZ}'\rangle + |1\rangle\,|GHZ''\rangle\,,$$

such that $|\,\langle\psi|\phi\rangle\,|^2 \geq 1-\epsilon$. Here $|\mathrm{GHZ}'\rangle$ and $|\mathrm{GHZ}''\rangle$ are states in the GHZ class hence the size of $|\phi\rangle$ is at most 14. As an example, for the representative state $|\psi^{(4)}\rangle$ in Eqn. (2.28) with the irreducible $A|BCD$ form, we show that

Proposition 2.11. *The ϵ-tree size of $|\psi^{(4)}\rangle$ is 14 for $0 < \epsilon < \frac{1}{12}$.*

Proof. We need to show that if $|\phi\rangle$ is described by a tree with fewer than 14 leaves, then $|\,\langle\psi^{(4)}|\phi\rangle\,|^2 \leq 1-\frac{1}{12}$. For this purpose we employ the same elimination procedure used previously. First we list all the trees with 13 leaves or fewer. We observe that all of these trees are special cases of the four trees listed in Fig. 2.8. Thus we are left of show what is the maximal over lap between states represented by these trees and $|\psi^{(4)}\rangle$.

Eliminating T_4+T_8, T_6+T_7, and T_4+T_9. Using the argument similar to Proposition 2.9, one can see that a state described by T_4+T_8, T_6+T_7, and T_4+T_9 has the form:

$$|\varphi\rangle = \alpha\,|u_0\rangle\,|T_B\rangle + \beta\,|u_1\rangle\,|\phi\rangle\,, \tag{2.31}$$

where $|\alpha|^2+|\beta|^2=1$, $|u_0\rangle$ and $|u_1\rangle$ form a orthonormal basis. $|T_B\rangle$ is a normalized biseparable state and $|\phi\rangle$ also normalized.

The state $|\varphi\rangle$ has the form of $A|BCD$ where A may correspond to any of the four qubits. Let us consider A correspond to the first qubit, the other case will be dealt with later. One can bound the overlap:

$$\begin{aligned}|\,\langle\psi^{(4)}|\varphi\rangle\,|^2 &= |\alpha\,\langle\psi^{(4)}|u_0,T_B\rangle + \beta\,\langle\psi^{(4)}|u_1,\phi\rangle\,|^2\\ &\leq |\,\langle\psi^{(4)}|u_0,T_B\rangle\,|^2 + |\,\langle\psi^{(4)}|u_1,\phi\rangle\,|^2,\end{aligned}$$

where the last line follows from the Cauchy-Schwarz inequality. Moreover, we have

$$\begin{aligned} |\langle\psi^{(4)}|u_0, T_B\rangle|^2 &= \frac{1}{2}|\langle 0|u_0\rangle\langle W_0|T_B\rangle + \langle 1|u_1\rangle\langle W_1|T_B\rangle|^2 \\ &\leq \frac{1}{2}\Big(|\langle W_0|T_B\rangle|^2 + |\langle W_1|T_B\rangle|^2\Big), \end{aligned}$$

and

$$\begin{aligned} |\langle\psi^{(4)}|u_1, \phi\rangle|^2 &= \frac{1}{2}|\langle 0|u_0\rangle\langle W_0|\phi\rangle + \langle 1|u_1\rangle\langle W_1|\phi\rangle|^2 \\ &\leq \frac{1}{2}\Big(|\langle W_0|\phi\rangle|^2 + |\langle W_1|\phi\rangle|^2\Big). \end{aligned}$$

Since $|W_0\rangle$ and $|W_1\rangle$ are orthogonal, we have $|\langle W_0|\phi\rangle|^2 + |\langle W_1|\phi\rangle|^2 \leq 1$. Together, we have

$$|\langle\psi^{(4)}|\varphi\rangle|^2 \leq \frac{1}{2}\Big(\underbrace{|\langle W_0|T_B\rangle|^2 + |\langle W_1|T_B\rangle|^2}_{f} + 1\Big). \tag{2.32}$$

We still need to bound $f \equiv |\langle W_0|T_B\rangle|^2 + |\langle W_1|T_B\rangle|^2$. Let us write a biseparable state as:

$$|T_B\rangle = \Big(a\,|0\rangle + b\,|1\rangle\Big) \otimes \Big(c_{00}\,|00\rangle + c_{01}\,|01\rangle + c_{10}\,|10\rangle + c_{11}\,|11\rangle\Big),$$

where $|a|^2 + |b|^2 = 1$ and $|c_{00}|^2 + |c_{01}|^2 + |c_{10}|^2 + |c_{11}|^2 = 1$. Substituting this into f, we obtain

$$f = \frac{1}{6}\Big(|b(c_{01} + c_{10}) - 2a\,c_{11}|^2 + |a(c_{01} + c_{10}) - 2b\,c_{00}|^2\Big).$$

Maximizing f with respect to those constraints gives $f_{\max} = \frac{2}{3}$.

Now we turn to the case where A the partition $A|BCD$ does not correspond to the first qubit in $|\varphi\rangle$. Since we are concerned with the overlap $|\langle\psi^{(4)}|\varphi\rangle|$, we can keep $|\varphi\rangle$ unchanged and permute the qubits in $|\psi^{(4)}\rangle$, which amounts to placing the factor -2 at different places. Due to the symmetry present in $|\psi^{(4)}\rangle$, there are only two other inequivalent cases:

$$|\psi_1^{(4)}\rangle = \frac{1}{\sqrt{12}}\Big(|0\rangle\Big(|110\rangle - 2\,|101\rangle + |011\rangle\Big) + |1\rangle\Big(|001\rangle - 2\,|010\rangle + |100\rangle\Big)\Big), \tag{2.33}$$

$$|\psi_2^{(4)}\rangle = \frac{1}{\sqrt{12}}\Big(|0\rangle\Big(-2\,|110\rangle + |101\rangle + |011\rangle\Big) + |1\rangle\Big(-2\,|001\rangle + |010\rangle + |100\rangle\Big)\Big). \tag{2.34}$$

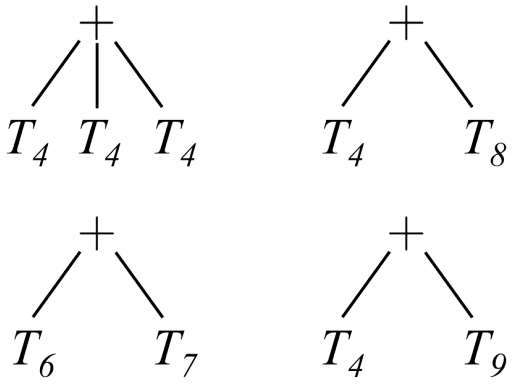


Figure 2.8: Rooted trees of four-qubit states with size 12 and 13.

Following the same argument as for $|\psi^{(4)}\rangle$, we arrive at the same bound as Eqn. 2.32 but with different value of f,

$$f_1 = \frac{1}{6}\Big(|b(-2c_{01}+c_{10})+a\,c_{11}|^2+|a(c_{01}-2c_{10})+b\,c_{00}|^2\Big),$$
$$f_2 = \frac{1}{6}\Big(|b(c_{01}-2c_{10})+a\,c_{11}|^2+|a(-2c_{01}+c_{10})+b\,c_{00}|^2\Big),$$

for $|\psi_1^{(4)}\rangle$ and $|\psi_2^{(4)}\rangle$ respectively. Notice that f_1 and f_2 are identical upon switching $c_{01} \leftrightarrow c_{10}$, so they have the same maximum. Maximizing f_1 with respect to the constraint gives $f_{1\,\max} = \frac{5}{6}$. Since $f_{1\,\max} > f_{\max}$, we have for all permutation of qubits,

$$|\,\langle\psi^{(4)}|\varphi\rangle\,|^2 \leq \frac{1+f_{1\,\max}}{2} = \frac{11}{12}. \tag{2.35}$$

Eliminating $T_4+T_4+T_4$. We still have to rule the last tree $T_4+T_4+T_4$, which does not take the form of Eqn. 2.31. Let us label the leaves at each branch as $x_i\,|0\rangle+x_i'\,|1\rangle$, $y_i\,|0\rangle+y_i'\,|1\rangle$ and $z_i\,|0\rangle+z_i'\,|1\rangle$, with $i=1,2,3,4$ for the three branches respectively. We then write the state described by $T_4+T_4+T_4$ as:

$$|\varphi\rangle = \bigotimes_{i=1}^{4}(x_i\,|0\rangle+x_i'\,|1\rangle)+\bigotimes_{i=1}^{4}(y_i\,|0\rangle+y_i'\,|1\rangle)+\bigotimes_{i=1}^{4}(z_i\,|0\rangle+z_i'\,|1\rangle),$$

with the constraint $|\,\langle\varphi|\varphi\rangle\,| = 1$. Numerical optimization shows that that $|\,\langle\psi^{(4)}|\varphi\rangle\,|^2 \leq \frac{8}{9} < \frac{11}{12}$.

We hence conclude that the maximum overlap $|\,\langle\psi^{(4)}|\varphi\rangle\,|^2 \leq \frac{11}{12}$ for any $|\varphi\rangle$ with at most 13 leaves. In other words, $\mathrm{TS}_\epsilon(|\psi^{(4)}\rangle) = 14$ for $0<\epsilon<\frac{11}{12}$. □

2.4 Simple states

From this section onwards, we will be focusing on the scaling behaviour of tree size: how tree size scales with respect to the number of qubits. We called states (more precisely, a family of states indexed by the number of qubits, n) whose tree size scales polynomially *simple* states, and those superpolynomially *complex*. The motivation for this is that tree size provides the shortest description of a state in the braket notation, and if that description is polynomial in n, it can be considered as an efficient description.

As we have seen in the last section, tree size is closely related to entanglement, especially genuine multipartite entanglement. However, tree size is an evidently different concept from multipartite entanglement. States that are genuinely multipartite entangled can be simple. To show a state to be simple, it is sufficient to show an explicit decomposition with polynomial tree size. In this section, we will explore some examples of simple states.

Product states

The product state $|0\rangle^{\otimes n}$ is the simplest state in terms of tree size. It is usually regarded as the input for the circuit model of quantum computation. Obviously $\mathrm{TS}(|0\rangle^{\otimes n}) = n$, which is the minimal TS for n-qubit states.

GHZ states

The n-qubit GHZ state, $(|0\rangle^{\otimes n} + |1\rangle^{\otimes n})/\sqrt{2}$, which saturates most of the macroscopicity measures [15], has $\mathrm{TS}(|\mathrm{GHZ}_n\rangle) = 2n$, which is linear in the number of qubits. This is a clear evidence that complexity is a different notion from macroscopicity. A maximally macroscopic state can yet be very simple.

Dicke states

The Dicke state $|D_{n,k}\rangle$ represents the equal superposition of n-qubit state with k excitations; formally it is the (unnormalized) uniform superposition of all the n-bit strings with Hamming weight k:

$$|D_{n,k}\rangle = \sum_{\{\alpha\}} \bigotimes |0\rangle_{i\notin\{\alpha\}} \bigotimes |1\rangle_{i\in\{\alpha\}} \tag{2.36}$$

where the summation is over $\{\alpha\}$, all the distinct k element subset of $\{1, \cdots, n\}$. When $k = 1$, $|D_{n,1}\rangle$ is also known as the n-qubit W state.

We show that

Proposition 2.12. *The state $|D_{n,k}\rangle$ has tree size $O(n^2)$.*

Proof. To see this, one can consider the uniform superposition of the following Fourier form (omitting normalization):

$$|\psi_{n,k}\rangle = \sum_{j=0}^{k-1} \left(|0\rangle + \exp(\frac{2\pi ij}{k}) |1\rangle \right)^{\otimes n}. \tag{2.37}$$

A direct expansion yields

$$|\psi_{n,k}\rangle = \sum_{p=0}^{n} \left(|D_{n,p}\rangle \sum_{j=0}^{k-1} \exp(2\pi ij\frac{p}{k}) \right).$$

When $p = mk$, for some integer m, $\exp(2\pi ij\frac{p}{k}) = 1$; when p is not a multiple of k, $\sum_{j=0}^{k-1} \exp(2\pi ij\frac{p}{k}) = 0$. Hence, $|\psi_{n,k}\rangle = \sum_{m=0}^{\lfloor n/k \rfloor} |D_{n,mk}\rangle$. Note the range of m is from 0 to $\lfloor \frac{n}{k} \rfloor$: for $k > \frac{n}{2}$, m can be only 0 or 1, thus $|D_{n,k}\rangle = |\psi_{n,k}\rangle - |0\rangle^{\otimes n}$; for $k = \frac{n}{2}$, m can be 0, 1 and 2; thus $|D_{n,k}\rangle = |\psi_{n,k}\rangle - |0\rangle^{\otimes n} - |1\rangle^{\otimes n}$; for $k < n/2$, by interchanging 0 and 1 we obtain the $k > \frac{n}{2}$ case. So for any k, $\mathrm{TS}(|D_{n,k}\rangle) = O(n^2)$. □

The n-qubit W state, which is $|D_{n,1}\rangle$, though representing the most complex class for the three-qubit case, has polynomial tree size $O(n^2)$.

Matrix product states

Matrix product states (MPS) are defined by

$$|\psi\rangle = \bigotimes_{i=1}^{n} (A_0^{(i)} |0\rangle + A_1^{(i)} |1\rangle),$$

where $A_0^{(i)}$ and $A_1^{(i)}$ are matrices of dimension at most χ. We may take $A^{(1)}$ to be a row vector and $A^{(n)}$ to be a column vector so that the coefficient of $|x_1, x_2, \cdots, x_n\rangle$ is $A_{x_1}^{(1)} A_{x_2}^{(2)} \cdots A_{x_n}^{(n)}$ is a complex number rather than a matrix. Any state can be written in this matrix product form if we do not restrict the rank of the matrix $A^{(i)}$'s. MPS is an efficient way of representing a quantum state if the bond dimensions (rank of $A^{(i)}$) do not increase with the number of qubits n. The ground state of one-dimensional gapped Hamiltonians can be well approximated by MPS with low bond dimensions [79, 80]. We will show that, MPS with a bounded bond dimension χ is simple.

Proposition 2.13. *An MPS with bond dimension χ has tree size $O(n^{\log 2\chi})$.*

Table 2.1: Summary of n-qubit simple states

Product state	n
GHZ_n states	$2n$
Dicke states $D_{n,k}$	$O(n^2)$
MPS with rank χ	$O(n^{\log 2\chi})$

Proof. This can be shown by an inductive argument. First consider partitioning the qubits into two halves, we have

$$|\psi\rangle = \sum_{s=1}^{\chi} |\psi_{n/2}^{s,1}\rangle\, |\psi_{n/2}^{s,2}\rangle\,, \tag{2.38}$$

where now $|\psi_{n/2}^{s}\rangle$ is an MPS of $n/2$ qubits. We can see that $\mathrm{TS}(|\psi_n\rangle) \leq 2\chi \cdot \mathrm{TS}(|\psi_{n/2}\rangle)$. Then by repeating this partitioning, we have $\mathrm{TS}(|\psi_n\rangle) = O((2\chi)^{\log n}) = O(n^{\log 2\chi})$. Thus, if χ is bounded as n increases, then TS is polynomial. □

One example of MPS, the 1D cluster state, is the ground state of the three-body interaction $\sum_i \sigma_i^z \sigma_{i+1}^x \sigma_{i+2}^z$. It has an MPS representation with

$$A_0 = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, \; A_1 = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix},$$

with bond dimension $\chi = 2$, hence its tree size is $O(n^2)$.

A natural generalization of MPS to higher spatial dimension are the Projected Entangled Pair States (PEPS) [81]. The same recursive argument can be applied to PEPS, though the upper bound are in general superpolynomial, $\chi^{O(\sqrt{n})}$ and $\chi^{O(n^{2/3})}$ for 2D and 3D PEPS respectively. And indeed, as we are going to discuss later in Sec. 2.7, PEPS that are universal for measurement-based quantum computation (MBQC) should have superpolynomial tree size. Specifically, in Sec. 2.5.6, we will show that the 2D cluster states has a superpolynomial tree size.

Examples of simple states are summarized in Table 2.1.

2.5 Complex states

2.5.1 Methods to obtain lower bounds on tree size

One of the advantages of tree size as a complexity measure is that there are tools for proving lower bound on tree size, hence certifying complex states. One way is to use a *counting argument* as Aaronson did in Theorem 7 of [49]. The fact

that there are fewer states that has polynomial tree size than there are in the whole Hilbert space (or the state space of interest), some states are bound to have superpolynomial or even exponential tree size.

Another method, which will be discussed more often here, is *a theorem first proved by Raz* in the context of multilinear formula size (MFS) [49, 82]. Although counting arguments could show that states with superpolynomial tree size must exist, Raz's theorem allows us to construct explicit examples. Let us present here this important theorem, first in Raz's original formulation, then in an equivalent way in terms of Schmidt rank.

Raz's theorem

We first need to introduce the concept of multilinear formulae and their size. A multilinear formula, $f(x_1, x_2, \cdots)$, is a formula that is linear in all of its inputs $x_1, x_2, \cdots$. The MFS of a multilinear formula is defined as the number of leaves in its *minimal tree representation* similar to the tree size of a quantum state. Next we introduce the concept of the partial derivative matrix. Consider a multilinear formula $f : \{0,1\}^n \to \mathbb{C}$, let P be a bipartition of the input variables $x_1, \cdots, x_n$ into two sets, $y_1, \cdots, y_{n/2}$ and $z_1, \cdots, z_{n/2}$. We now view $f(x)$ as a function $f_P(y,z) : \{0,1\}^{n/2} \times \{0,1\}^{n/2} \to \mathbb{C}$. Then define the partial derivative matrix, $\mathcal{M}_{f|P}$, to be the $2^{n/2} \times 2^{n/2}$ matrix whose rows and columns are labelled by y and $z \in \{0,1\}^{n/2}$, respectively, and the entry (y,z) of this matrix is defined as $\mathcal{M}_{f|P}(y,z) = f_P(y,z)$. Finally, let $\text{rank}(\mathcal{M}_{f|P})$ be the rank of $\mathcal{M}_{f|P}$ over the complex numbers, and $\mathcal{P}$ be the uniform distribution over all the possible bipartitions P.

Alternatively, one can define $\mathcal{M}'_{f|P}(y,z)$ as the coefficient of pq in $f(Y,Z)$, where $p = Y_1^{y_1} \cdots Y_{n/2}^{y_{n/2}}$ is a monomial in Y, and similarly for q. $\mathcal{M}'_{f|P}$ defined this way has same rank as $\mathcal{M}_{f|P}$.

We now state Raz's theorem [82]:

Theorem 2.14. *If*

$$\Pr_{P\in\mathcal{P}}\left[\text{rank}(\mathcal{M}_{f|P}) > 2^{\frac{n-n^{1/8}}{2}}\right] = n^{-o(\log n)}, \tag{2.39}$$

then $\text{MFS}(f) = n^{\Omega(\log n)}$.

Note the rank of $\mathcal{M}_{f|P}$ is at most $2^{n/2}$. The theorem says that if $\mathcal{M}_{f|P}$ is close to full rank for not too small fraction of all the bipartitions, then $\text{MFS}(f) = n^{\Omega(\log n)}$. A short note on the big O notation and related notations can be found in Appendix. A.

Aaronson pointed out the relation between multilinear formula size and tree size. For any quantum state $|\psi\rangle$, we can define the associated multilinear formula $f_\psi(x) = \langle x|\psi\rangle$. Note that this formula computes the coefficients in the computational basis expansion of $|\psi\rangle$. Given a tree representation of a quantum state, a tree for the associated multilinear formula can be obtained by replacing $|0\rangle_i \to (1 - x_i)$ and $|1\rangle_i \to x_i$. For example, the two qubit state $|\psi\rangle = c\,|00\rangle + s\,|11\rangle$. The corresponding formula obtained from the replacement is $f(x_1, x_2) = c\,(1 - x_1)(1 - x_2) + s\,x_1x_2$. One can easily check that $f(0, 0) = c$, $f(1, 1) = s$ and $f(0, 1) = f(1, 0) = 0$, satisfying $f(x) = \langle x|\psi\rangle$.

Given the minimal tree of quantum state, we can obtain a tree for the associated multilinear formula with the same size. The true MFS of the formula can only be smaller, therefore [49]:

Theorem 2.15. $\mathrm{MFS}(f_\psi) = O(\mathrm{TS}(|\psi\rangle))$. *Therefore, if f_ψ satisfies Raz's theorem, then* $\mathrm{TS}(|\psi\rangle) = n^{\Omega(\log n)}$.

In fact, in the original paper of Aaronson [49], he showed that the reverse of the inequality is also true up to n, $\mathrm{TS}(|\psi\rangle) = O(\mathrm{MFS}(f_\psi)+n)$. But for our application to lower bound tree size, $\mathrm{MFS}(f_\psi) \leq \mathrm{TS}(|\psi\rangle)$ is sufficient.

In previous sections we have seen that tree size complexity is related to multipartite entanglement though no simple connection can be drawn. Here, we present another link between entanglement and tree size complexity, by phrasing Raz's theorem in terms of an entangle measure for pure multipartite states, the Schmidt rank [50]:

Theorem 2.16. *For a pure quantum state of n qubits, $|\psi\rangle$, consider all the uniformly distributed $(\frac{n}{2}, \frac{n}{2})$ bipartitions, if*

$$\Pr\left[SR > 2^{\frac{n-n^{1/8}}{2}}\right] = n^{-o(\log n)}, \tag{2.40}$$

where SR is the Schmidt rank across a particular bipartition, then $\mathrm{TS} = n^{\Omega(\log n)}$.

Proof. The statement follows indeed from Raz's theorem, because partitioning of the input x of the associated multilinear formula f_ψ is the same as partitioning the qubits of the state $|\psi\rangle$. Note that $\mathcal{M}_{f_\psi|P}$ is a matrix each of whose element is a coefficient of the state $|\psi\rangle$ in its computational basis. Thus, the rank of $\mathcal{M}_{f_\psi|P}$ is exactly the Schmidt rank of the state $|\psi\rangle$ for the bipartition P [50]. □

In light of this, complex states are those not only contain genuine multipartite entanglement, but whose entanglement is also high (almost full Schmidt rank) across many bipartitions.

With Raz's theorems, we shall identify some explicit multiqubit states with superpolynomial tree size.

2.5.2 Immanant states

An explicit family of states with superpolynomial tree size can be constructed based on the immanant of a (0,1) matrix [38]. Consider the case when the number of qubits is a square number, $n = m^2$, for each bit string $|x\rangle = |x_1, \ldots, x_n\rangle$ we arrange the bits $x_1, \ldots, x_n$ row by row to an $m \times m$ matrix $M(x)$ such that

$$M(x) = \begin{pmatrix} x_1 & x_2 & \cdots & x_m \\ x_{m+1} & x_{m+2} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n-m} & x_{n-m+1} & \cdots & x_n \end{pmatrix},$$

so $M(x)_{ij} = x_{m(i-1)+j}$. The immanant states are defined in its computational basis expansion as

$$|\mathrm{Imm_n}\rangle = \sum_{x \in \{0,1\}^n} \mathrm{Imm}(M(x)) \, |x\rangle \, . \tag{2.41}$$

Here the immanant $\mathrm{Imm}(M)$ of a matrix M is given by

$$\mathrm{Imm}(M) = \sum_{\sigma \in S_m} c_\sigma \prod_{i=1}^{m} x_{i\sigma_i}, \tag{2.42}$$

where σ is an element of the symmetric group S_m of all the $m!$ permutation of $\{1, 2, \ldots, m\}$, and c_σ is the corresponding complex coefficient. When $c_\sigma = 1$ for all σ the immanant reduces to the permanent, and when $c_\sigma = 1$ for even permutations and -1 for odd permutations it reduces to the determinant. It is proved in Ref. [38] that

Theorem 2.17. *The immanant states as defined above have* $\mathrm{TS} = n^{\Omega(\log n)}$ *if the coefficients* c_σ *are all nonzero.*

Proof. The proof follows from Raz's proof for the superpolynomial MFS for determinant and permanent. For any bipartition P, adopting the alternative definition of $\mathcal{M}_{f|P}(y, z)$ and the definition of immanant (2.42), for every p there is exactly one q such that $pq = \prod_{i=1}^{m} x_{i\sigma_i}$ and the coefficient of pq is c_σ, otherwise $\mathcal{M}_{f|P}(y, z) = 0$. This implies that $\mathcal{M}_{f|P}$ is a permutation of a diagonal matrix filled with c_σ, thus having full rank $2^{n/2}$ if all c_σ are nonzero. By Raz's theorem, $\mathrm{MFS(Imm)} = n^{\Omega(\log n)}$.

Following from Theorem 2.15, tree size being lower bounded by MFS, so $\mathrm{TS}(|\mathrm{Imm}_n\rangle) = n^{\Omega(\log n)}$. □

The permanent and the determinant states,

$$|\mathrm{Perm}_n\rangle = \sum_{x\in\{0,1\}^n} \mathrm{Perm}(M(x))\,|x\rangle\,,$$
$$|\mathrm{Det}_n\rangle = \sum_{x\in\{0,1\}^n} \mathrm{Det}(M(x))\,|x\rangle\,, \tag{2.43}$$

are two examples in this family of complex states. Let us discuss possible upper bounds on the tree size of these states. For the particular case of the permanent, the smallest known formula for computing it is the Ryser's formula [83], which is multilinear and given as the following: Let S be one of the 2^m subsets of $\{1, 2, \ldots, m\}$ and $|S|$ the number of its elements, then the permanent of the matrix M is

$$\mathrm{Perm}(M) = \sum_S (-1)^{m+|S|} \prod_{i=1}^{m} \sum_{j\in S} M_{ij}. \tag{2.44}$$

By substituting this formula to the permanent state and carrying out the summation over x we obtain a decomposition with size $n^{\frac{3}{2}}2^{\sqrt{n}}$. Since this must be an upper bound on TS, we have $\mathrm{TS}(\mathrm{Perm}_n) = 2^{O(\sqrt{n})}$. We conjecture that $\mathrm{TS}(\mathrm{Perm}_n) = 2^{\Omega(\sqrt{n})}$ based on a strong evidence: computing the permanent of $(0, 1)$ matrices is a #P-complete problem [84, 85], and hence it is very likely that the smallest multilinear formula for doing so has exponential size.

The determinant is known to be much easier to compute than the permanent: In fact, there exists a formula that computes the determinant of a $m \times m$ matrix with size $O(m^4)$ [86]. However, this formula is not multilinear; and only multilinear formulae can be used to find an upper bound on the TS of the corresponding state. As shown by Raz [82], any multilinear formula computing the determinant of a $m\times m$ matrix has size $m^{\Omega(\log m)}$. Thus, the determinant state also has superpolynomial TS.

2.5.3 Deutsch-Jozsa states

A brief description of the algorithm

The Deutsch-Jozsa algorithm [87] outperforms its classical counterparts in the deterministic case [50]. It is an algorithm that solves the following hypothetical question: a function $f : \{0, 1\}^n \to \{0, 1\}$ is called *balanced* if exactly half of its input is mapped to 0 and the other half to 1, and *constant* if all the inputs

are mapped to 0 or 1. Given the promise that the function is either balanced or constant, how many queries do we need to find out whether the function is balanced or constant? Classically, in the deterministic and worse case scenario, it requires $2^{n-1}+1$ queries, in which case the function outputs all 0 or all 1 for the first 2^{n-1} queries.

The Deutsch-Jozsa algorithm solves the quantum version of this problem with only one query, which is exponentially faster than the classical algorithm. In the quantum version, a query is replaced by the quantum oracle $|x\rangle |y\rangle \to |x\rangle |y \oplus f(x)\rangle$. In this algorithm, one first prepares the input state as $|0\rangle^n |1\rangle$, then applies the Hadamard transformation to all the registers, resulting in

$$\frac{1}{2^{(n+1)/2}} \sum_{x\in\{0,1\}^n} |x\rangle (|0\rangle - |1\rangle).$$

After applying the oracle, the state becomes

$$\frac{1}{2^{(n+1)/2}} \sum_{x\in\{0,1\}^n} |x\rangle (|f(x)\rangle - |1 \oplus f(x)\rangle).$$

Since $f(x)$ is either 0 or 1, we can simplify this to $|\psi\rangle = |\psi_{DJ}\rangle \otimes |-\rangle$, where

$$|\psi_{DJ}\rangle = \frac{1}{2^{n/2}} \sum_{x\in\{0,1\}^n} (-1)^{f(x)} |x\rangle . \tag{2.45}$$

The last qubit register can be left out at this point. Applying the Hadamard transformation to all the qubits once again, we have

$$\frac{1}{2^n} \sum_{y\in\{0,1\}^n} \left(\sum_{x\in\{0,1\}^n} (-1)^{f(x)+x\cdot y} |y\rangle \right),$$

where $x \cdot y$ represents the sum of bitwise product. Finally, a projection onto $|0\rangle^n$ has probability

$$\left| 2^{-n} \sum_{x\in\{0,1\}^n} (-1)^{f(x)} \right|^2 = \begin{cases} 0, \text{if } f(x) \text{ is constant,} \\ 1, \text{if } f(x) \text{ is balanced.} \end{cases}$$

Thus we can find out if the function f is constant or balanced with only a single application of the oracle. This concludes the algorithm, now we switch the focus to the tree size of the state $|\psi_{DJ}\rangle$.

The complexity of the state

If f is constant, then the state $|\psi_{DJ}\rangle$ is a simple product state $|+\rangle^n$ with tree size n. If f is balanced, we would like to show that an overwhelmingly large fraction of balanced functions correspond to states with superpolynomial tree size. Consider a function f randomly drawn from the uniform distribution of all the $\binom{2^n}{2^{n/2}}$ balanced functions, let P be a random equal bipartition of the input x into y and z, then $\mathcal{M}_{f|P}$ the $2^{n/2} \times 2^{n/2}$ matrix whose entries are $(-1)^{f(y,z)}$. Note that for a balanced f, the matrix $\mathcal{M}_{f|P}$ has exactly half entries equal to $+1$ and half equal to -1. Let $\mathcal{E}_1$ be the event that $\mathcal{M}_{f|P}$ has full rank $2^{n/2}$, we need to compute the probability that $\mathcal{E}_1$ happens, in order to see whether the balanced function f leads to a state with superpolynomial TS (c.f. Theorem 2.14).

Let us call a matrix with exactly half entries equal to 1 and the other half -1 a *balanced* $(1,-1)$ *matrix*. Denote by $\mathcal{M}_R$ a random balanced $(1,-1)$ matrix, $\mathcal{M}_R$ can be chosen by first drawing a random balanced function f, then picking a random bipartition P and assigning $\mathcal{M}_R = \mathcal{M}_{f|P}$. Now let $\mathcal{E}_2$ be the event that $\mathcal{M}_R$ has full rank, we have

$$\Pr(\mathcal{E}_2) = \sum_f \Pr(f) \Pr(\mathcal{E}_1|f). \tag{2.46}$$

Next, we split the set of f into those which give rise to a complex state (i.e. satisfy Raz's theorem) and those which do not. Explicitly, let $C = \{f | \Pr(\mathcal{E}_1|f) \geq q\}$, where q is a constant to be specified later, and $\bar{C}$ be the complement of C, then

$$\Pr(\mathcal{E}_2) = \sum_C \Pr(f) \Pr(\mathcal{E}_1|f) + \sum_{\bar{C}} \Pr(f) \Pr(\mathcal{E}_1|f). \tag{2.47}$$

Since $\Pr(\mathcal{E}_1|f) \leq 1$ for all f and $\Pr(\mathcal{E}_1|f) < q$ for all $f \in \bar{C}$, we have

$$\Pr(\mathcal{E}_2) \leq \sum_C \Pr(f) + q \sum_{\bar{C}} \Pr(f). \tag{2.48}$$

Note that the sums of the probability that f is chosen from C and $\bar{C}$ give the fraction of states in the respective sets, that is, $\sum_C \Pr(f) = \frac{N_C}{N_f} = F_C$ and $\sum_{\bar{C}} \Pr(f) = \frac{N_f - N_C}{N_f} = 1 - F_C$. Substituting this to the above inequality, we arrive at

$$F_C \geq \frac{\Pr(\mathcal{E}_2) - q}{1 - q}. \tag{2.49}$$

Thus, to know how large F_C is we need to know $\Pr(\mathcal{E}_2)$, the probability that $\mathcal{M}_R$ is invertible where $\mathcal{M}_R$ is a random $2^{n/2} \times 2^{n/2}$ balanced $(1,-1)$ matrix. Numerical evidence shows that $\Pr(\mathcal{E}_2)$ approaches 1 quickly as n becomes large

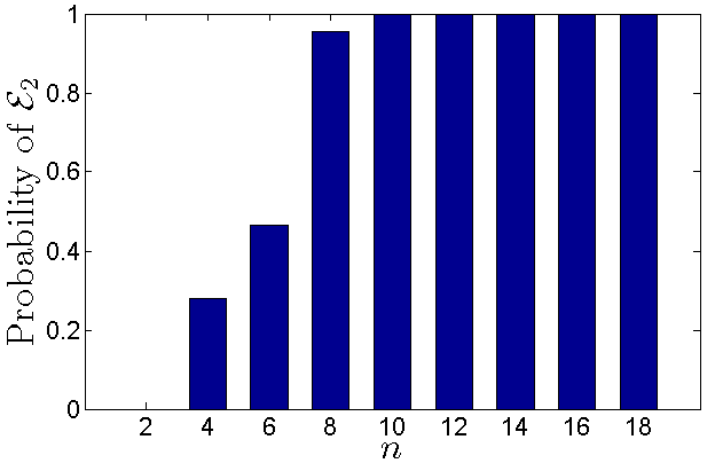


Figure 2.9: The probability of the event $\mathcal{E}_2$ versus the number of input bits n. $\mathcal{E}_2$ is the event that a random $2^{n/2} \times 2^{n/2}$ balanced $(1, -1)$ matrix has full rank. For large n, the probability is estimated by sampling.

(see Figure 2.9). If one believes that $\Pr(\mathcal{E}_2) \approx 1$ for large n, which is strongly suggested by the numerical evidence, then by setting q to a constant not close to 1, say 0.5, we see that $F_C \approx 1$. This means that nearly all balanced functions give rise to states with superpolynomial tree size.

One may argue that the large tree size that arises from the Deutsch-Jozsa algorithm has its root in the oracle's access to completely-random balanced function. The link between large tree size and the usefulness of the algorithm is unclear. Nonetheless, this provides us with an example of complex states that appear in a quantum algorithm. More on the relation between state complexity and quantum computation will be discussed in Sec. 2.7.

2.5.4 Shor's states

Shor's algorithm factors an integer N in time $O((\log N)^3)$ [88], which is exponentially faster than the most efficient known classical algorithm [89]. Do states arising from this algorithm have superpolynomial tree size? Aaronson showed the answer is yes assuming a number-theoretic conjecture [49]. For completeness, we present his proof and the number-theoretic conjecture that he used here.

To factorize N, one picks a pseudo random integer $x < N$, coprime to N, and finds the period of the function $g(r) = x^r \bmod N$ with the quantum subroutine. Let us call the following state of $2n = 2\log(N)$ qubits in the period finding subroutine the Shor's state, defined as

$$|\text{Shor}\rangle = \frac{1}{2^{n/2}} \sum_{r=0}^{2^n-1} |r\rangle \, |x^r \bmod N\rangle \, . \tag{2.50}$$

To prove of the lower bound on the TS of the Shor's state, it is convenient to measure the second register. Since a projective measurement on part of the state does not increase tree size, we can focus on the resulting state of the first register,

$$|a + p\mathbb{Z}\rangle = \frac{1}{\sqrt{I}} \sum_{i=0}^{I} |a + pi\rangle \,, \tag{2.51}$$

where p is the period of $g(r)$ and $I = \lfloor (2^n - 1)/p \rfloor$. Here $a + pi$ is represented in binary with n bits, so $|a + p\mathbb{Z}\rangle$ is a n-qubit state. $\mathrm{TS}(|a + p\mathbb{Z}\rangle)$ provides a lower bound for the tree size of the state of the two registers given in (2.50).

The associated formula for this state is a function of a n-bit string such that $f_{n,p,a}(x) = 1$ if $x \equiv a \bmod p$ and $f_{n,p,a}(x) = 0$ otherwise. Without loss of generality, we can take $a = 0$ and denote $f_{n,p,a}$ as $f_{n,p}$. $\mathrm{MFS}(f_{n,p})$ lower bounds $\mathrm{TS}(|a + p\mathbb{Z}\rangle)$, so we shall focus on this formula.

We now state the number-theoretic conjecture that allows us to show the superpolynomial lower bound for Shor's state:

Conjecture 2.18. *There exists $\gamma, \delta \in (0, 1)$ and a prime $p = \Omega(2^{n^\delta})$ for which the following holds. Let A consists of $l = n^\delta$ element from the set $\{2^0, 2^1, \cdots, 2^{n-1}\}$ chosen uniformly at random. Let S consists of all the 2^l sums of subsets of A, and let $S \bmod p = \{x \bmod p : x \in S\}$. Then*

$$\Pr_A \left[|S \bmod p| \geq (1 + \gamma)\frac{p}{2} \right] = n^{-o(\log n)}. \tag{2.52}$$

It was shown by Aaronson that [49]:

Theorem 2.19. *Conjecture 2.18 implies that $\mathrm{MFS}(f_{n,p}) = n^{\Omega(\log n)}$ and hence $\mathrm{TS}(|\mathrm{Shor}\rangle) = n^{\Omega(\log n)}$.*

Proof. Let $f = f_{n,p}$. Let R be a restriction of f that relabels $2l$ variables to $y_1, \cdots, y_l$ and $z_1, \cdots, z_l$ and sets the other $n - 2l$ variables to 0. This defines a new function $f_R(y, z)$ that equals to 1 when $y + z + c \equiv 0 \bmod p$ and 0 otherwise, for some constant c. Here y and z are defined through a binary encoding $y = 2^{a_1} y_1 + \cdots + 2^{a_l} y_l$ and $z = 2^{b_1} z_1 + \cdots + 2^{b_l} z_l$ where $a_1, \cdots, a_l, b_1, \cdots, b_l$ are appropriate place values. Now suppose $y \bmod p$ and $z \bmod p$ both assume at least $(1 + \gamma)p/2$ distinct values as we range over $x \in \{0, 1\}^n$ (This occurs with probability $n^{-o(\log n)}$ by the conjecture). Then by the pigeon hole principle, for at least γp possible values of $y \bmod p$, there exists a unique value of $z \bmod p$ for which $y + z + c \equiv 0 \bmod p$ hence $f_R(y, z) = 1$. So $\mathrm{rank}(\mathcal{M}_{f|R}) = \gamma p$, where $\mathcal{M}_{f|R}$ is the $2^l \times 2^l$ partial derivative

matrix where $\mathcal{M}_{f|R}(y,z) = f_R(y,z)$. From the conjecture 2.18, so we have

$$\Pr_{R\in\mathcal{R}_l}\left[\operatorname{rank}(\mathcal{M}_{f|R}) \geq \gamma p\right] = n^{-o(\log n)}.$$

Further more, $\gamma p \geq 2^{l-l^{1/8}/2}$ for sufficiently large n since $p = \Omega(2^{n^l})$. Therefore by Raz's theorem $\mathrm{MFS}(f) = n^{\Omega(\log n)}$. □

2.5.5 Subgroup states

Subgroup states used in quantum error correction also exhibit superpolynomial tree size. Let the element of $\mathbb{Z}_2^n$ be labelled by n-bit strings. Given a subgroup $S \subseteq \mathbb{Z}_2^n$, a subgroup state is defined as

$$|S\rangle = \frac{1}{\sqrt{|S|}} \sum_{x\in S} |x\rangle\,. \tag{2.53}$$

One way to construct a subgroup state is by considering the subgroup to be the null space of a $(0,1)$ matrix over the field $\mathbb{Z}_2$. Given a $n/2 \times n$ binary matrix A, a bit string x is in the null space of A if

$$Ax = 0 \bmod 2; \tag{2.54}$$

and the subgroup state is the equal superposition of all such bit strings. Aaronson shows in Ref. [49] that, if A is drawn from the set of all possible $n/2 \times n$ binary matrices, then at least 4% of these matrices give rise to subgroup states with superpolynomial TS.

Let us describe briefly how to prove that a subgroup state has superpolynomial tree size. Consider a random equal bipartition of $x = \{x_1, \ldots, x_n\}$ into $y = \{y_1, \ldots, y_{n/2}\}$ and $z = \{z_1, \ldots, z_{n/2}\}$. Denote by A_y the $n/2 \times n/2$ submatrix of the columns in A that applies to y (see Eq. (2.54)), and similarly A_z the submatrix of the columns that applies to z. Then, the element of the partial derivative matrix $\mathcal{M}_{f|P}(y,z)$ is 1 when $Ax = 0 \mod 2$, which means $A_y y + A_z z = 0 \mod 2$, and 0 otherwise. So long as both A_y and A_x are invertible, for each y there is only one unique value of z that gives $\mathcal{M}_{f|P}(y,z) = 1$. In other words, $M_{f|P}$ is a permutation of the identity matrix, hence it has full rank. Based on this observation, one sees that

Theorem 2.20. *Let A be a $n/2 \times n$ binary matrix and $S = \ker(A)$ over the field $\mathbb{Z}_2$. For random equal bipartitions of x into y and z as described above, if*

A_y and A_z are both invertible with probability $n^{-o(\log n)}$, then $\mathrm{TS}(|S\rangle) = n^{\Omega(\log n)}$. *Moreover,* $\mathrm{TS}_\epsilon(|S\rangle) = n^{\Omega(\log n)}$ *with* $\epsilon \leq 1 - \mu_n$, *where* $\mu_n = 2^{-(n/2)^{1/8}/2}$.

Proof. The first part follows from the fact that both A_y and A_z being invertible implies that $M_{f|P}$ has full rank. If this happens with probability $n^{-o(\log n)}$, then Raz's theorem is satisfied, hence $\mathrm{TS}(|S\rangle) = n^{\Omega(\log n)}$.

For the second part, we use a lemma proved by Aaronson in Ref. [49]: Denote by $|\psi\rangle$ a state close to a complex state $|S\rangle$ that satisfies Theorem 2.20, such that $|\langle\psi|S\rangle|^2 \leq 1 - \epsilon$. Then, for a fraction of $n^{-o(\log n)}$ of all equal bipartitions, the rank of the partial derivative matrix is

$$\mathrm{rank}(\mathcal{M}_{\psi|P}) \geq (1-\epsilon)2^{n/2}. \tag{2.55}$$

In order to satisfy Raz's theorem, we want $\mathrm{rank}(\mathcal{M}_{\psi|P}) \geq 2^{n/2-(n/2)^{1/8}/2}$. A comparison with the above equation gives $\epsilon \leq 1-\mu_n$ where $\mu_n = 2^{-(n/2)^{1/8}/2}$. Therefore, $\mathrm{TS}_\epsilon(|\psi\rangle) = n^{\Omega(\log n)}$ if $\epsilon \leq 1 - 2^{-(n/2)^{1/8}/2}$. □

Since μ_n is exponentially small in $n^{1/8}$, one might think that most states in the Hilbert space satisfy $|\langle\psi|S\rangle|^2 \geq \mu_n$, and hence Theorem 2.20 can be used to show that most states have superpolynomial TS. This is not correct: Indeed, if $|\psi\rangle$ is randomly and uniformly chosen from the Hilbert space according to a Haar measure, the probability that $|\langle\psi|S\rangle|^2 \geq \mu_n$ is smaller than $\exp[-(2^n-1)\mu_n]$, which is exponentially small [55]. However, it is true that most states in the Hilbert space have *exponential* tree size, as showed by a counting argument in Ref. [49].

Explicit construction

Aaronson first showed an explicit construction by Vandermonde matrix that leads to a superpolynomial complex subgroup state [49]. Here we present a different construction of the matrix A, for which strong numerical evidence suggests that the corresponding subgroup state has superpolynomial TS. Consider the matrix $A_J = (\mathbb{1}|Q)$, where $\mathbb{1}$ is the identity matrix and Q a binary Jacobsthal matrix, both of size $n/2 \times n/2$. Jacobsthal matrices are used in the Paley construction of Hadamard matrices [90]. The binary version is defined as follows: For a prime number q, one can define the quadratic character $\chi(a)$ that indicates whether the finite field element $a \in \mathbb{Z}_q$ is a perfect square. We have $\chi(a) = 1$ if $a = b^2$ for some non-zero element $b \in \mathbb{Z}_q$; and $\chi(a) = 0$ otherwise. Then $Q_{i,j}$ is equal to $\chi(i-j)$.

We study the partitioning of A_J into A_y and A_z randomly. Numerical evidence (see Fig. 2.10) shows that when q *is a prime and* $q = 8k+3$ *with* $k \in \mathbb{N}$, then, A_y

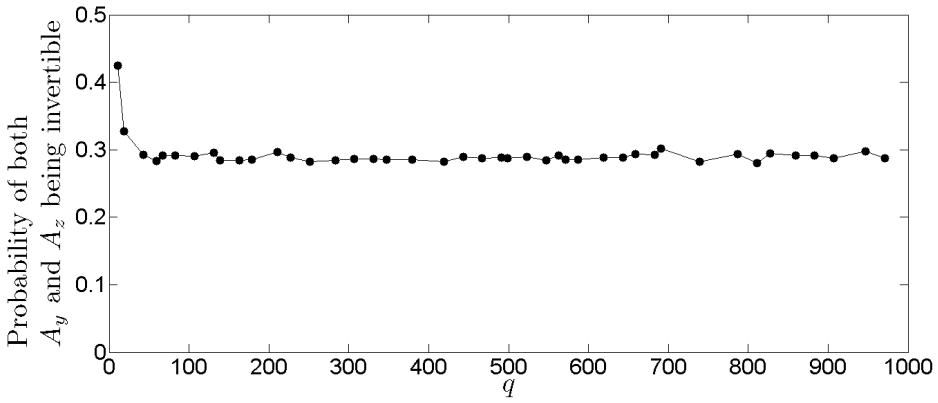


Figure 2.10: The probability of both A_y and A_z being invertible over random equal bipartitions of A_J. A_J is the $q \times 2q$ matrix $(\mathbb{1}|Q)$, where Q is the Jacobsthal matrix of size $q \times q$, where $q = 3 \mod 8$ and is a prime. For large q, the probability estimated from random sampling approaches a constant around 30%.

and A_z are both invertible with a probability approaching to a constant around 30%. From Theorem 2.20 we see that the subgroup state defined by A_J has $\mathrm{TS} = n^{\Omega(\log n)}$ where $n = 2q$.

2.5.6 2D cluster state

It is known that measurement-based quantum computation (MBQC) on the 2D cluster state is as strong as the circuit model of quantum computation [91–93]. In this scheme of computation, after the initial resource state is prepared, one only performs single qubit projective measurements and feedforward the outcomes. The power of the computation seems to lie in the initial resource state. Therefore, an initial state that is universal for quantum computation, such as the 2D cluster state, should be highly complex. It is conjectured in Ref. [49] that the 2D cluster state has superpolynomial TS. By studying the generation of a complex subgroup state via MBQC on the 2D cluster state, we prove that this conjecture is true:

Theorem 2.21. *The 2D cluster state of N qubits has* $\mathrm{TS} = N^{\Omega(\log N)}$.

Proof. Suppose we aim to produce an n-qubit complex subgroup state $|S_C\rangle$ (as described in Sec. 2.5.5) that has tree size $n^{\Omega(\log n)}$. These states are known to be stabilizer states [50]. Aaronson and Gottesmann showed that any n-qubit stabilizer state can be prepared using a stabilizer circuit with $O(n^2/\log n)$ number of gates [71] . A stabilizer circuit is one that consists of only CNOTs, $\pi/2$-phase gates and Hadamard gates. In the MBQC scheme, each of these gates can be implemented by measuring a constant number of qubits: 15 qubits for CNOT, and 5 qubits for the phase gate and the Hadamard gate [93]. In order to obtain a n-qubit

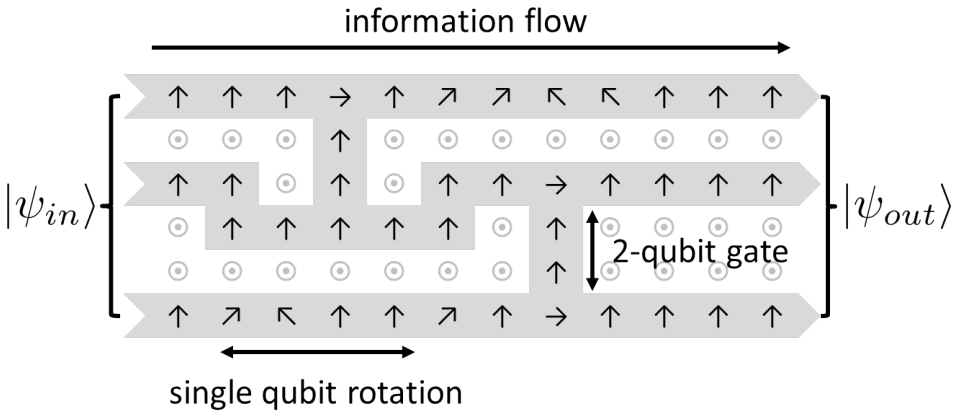


Figure 2.11: A schematic diagram of measurement-based quantum computation. Starting from a 2D cluster state, single qubit measurements are performed. $\odot$ represents a Z measurement, and other arrows represent measurements in the XY plane. The logical input state enters from the left and propagates to the right. Single qubit rotations and controlled gates are realized by a certain sequence of adaptive measurements, from left to right. Implementing a circuit on n qubits with m gates requires a cluster state of size $O(n)$-by-$O(m)$.

complex subgroup state, one needs to prepare a $O(n)$-by-$O(n^2/\log n)$ lattice (see Fig. 2.11), so the number of qubits in the 2D cluster state is $N = O(n^3/\log n)$. Since single qubit projective measurements only decrease tree size (c.f. the proof of Theorem 2.25), we have

$$\mathrm{TS}(\text{2D cluster}) \geq \mathrm{TS}(|S_C\rangle) = n^{\Omega(\log n)} = N^{\Omega(\log N)}. \tag{2.56}$$

So, the N-qubit 2D cluster state has the superpolynomial tree size. □

We show that the ϵ-tree size of the 2D cluster state is also superpolynomial:

Theorem 2.22. *For $\epsilon \leq 1/2$ and N large enough,* $\mathrm{TS}_\epsilon(\text{2D cluster}) = N^{\Omega(\log N)}$.

Proof. Assume that we have prepared a state close to the 2D cluster state, $|\mathrm{2D}_\epsilon\rangle$, such that the fidelity $F(|\mathrm{2D}\rangle, |\mathrm{2D}_\epsilon\rangle) = |\langle \mathrm{2D}|\mathrm{2D}_\epsilon\rangle| \geq \sqrt{1-\epsilon}$. Then we apply the same measurement sequence to the erroneous 2D cluster state as if we would to the ideal 2D cluster for preparing a complex subgroup state $|S_C\rangle$. Consider the state after one of the single-qubit measurement in the orthonormal basis $\left\{|\eta\rangle, |\eta^\perp\rangle\right\}$; the single-qubit projectors are $P_0 = |\eta\rangle\langle\eta|$ and $P_1 = |\eta^\perp\rangle\langle\eta^\perp|$. We now show that one of these outcomes will increase the fidelity between the two cases. If the measurement outcome is not observed, the resulting states on the ideal and

ϵ-deviated 2D cluster states are:

$$\begin{aligned} |2D\rangle \to \rho &= P_0 |2D\rangle \langle 2D| P_0^\dagger + P_1 |2D\rangle \langle 2D| P_1^\dagger \\ &= p_0 |\eta\rangle \langle \eta| \otimes |\psi_0\rangle \langle \psi_0| + p_1 |\eta^\perp\rangle \langle \eta^\perp| \otimes |\psi_1\rangle \langle \psi_1| , \qquad (2.57) \\ |2D_\epsilon\rangle \to \sigma &= P_0 |2D_\epsilon\rangle \langle 2D_\epsilon| P_0^\dagger + P_1 |2D_\epsilon\rangle \langle 2D_\epsilon| P_1^\dagger \\ &= p'_0 |\eta\rangle \langle \eta| \otimes |\psi'_0\rangle \langle \psi'_0| + p'_1 |\eta^\perp\rangle \langle \eta^\perp| \otimes |\psi'_1\rangle \langle \psi'_1| , \qquad (2.58) \end{aligned}$$

where $|\psi_{0,1}\rangle$ and $|\psi'_{0,1}\rangle$ are the states of the remaining qubits in the cluster; and $p_{0,1}$ and $p'_{0,1}$ are the probability of the measurement outcomes. Clearly, the above map is completely positive and trace-preserving (CPTP). Thus, the fidelity of these two states should not decrease due to contractivity of trace-preserving maps [50],

$$F(\rho, \sigma) \geq F(|2\mathrm{D}\rangle , |2\mathrm{D}_\epsilon\rangle) = \sqrt{1-\epsilon}. \qquad (2.59)$$

With a bit of algebra, we can express $F(\rho, \sigma)$ in terms of the fidelity of the post-selected states for the same outcome:

$$\begin{aligned} F(\rho, \sigma) &= \mathrm{tr}\sqrt{\rho^{1/2}\sigma\rho^{1/2}} \\ &= \sqrt{p_0 p'_0}| \langle \psi_0|\psi'_0\rangle | + \sqrt{p_1 p'_1}| \langle \psi_1|\psi'_1\rangle |. \qquad (2.60) \end{aligned}$$

Let us denote $x = \max\{| \langle \psi_0|\psi'_0\rangle |, | \langle \psi_1|\psi'_1\rangle |\}$ to be the larger overlap between the two, then

$$F(\rho, \sigma) \leq x \left(\sqrt{p_0 p'_0} + \sqrt{p_1 p'_1}\right) \leq x, \qquad (2.61)$$

since $\sqrt{p_0 p'_0} + \sqrt{p_1 p'_1} \leq (p_0 + p'_0 + p_1 + p'_1)/2 = 1$. Combining Eq. (2.60) and Eq. (2.61), we have

$$x = \max \{| \langle \psi_0|\psi'_0\rangle |, | \langle \psi_1|\psi'_1\rangle |\} \geq \sqrt{1-\epsilon}. \qquad (2.62)$$

Therefore, for at least one of the outcomes, we have a non-decreasing fidelity on the unmeasured parts of the states. For every measurements we post-select on the outcome that do not decrease the fidelity. Note that the complex subgroup states can be realized by a Clifford circuit, which can be implemented by a series of non-adaptive measurements. This means that, regardless of the outcome, the state obtained from the ideal 2D cluster is a complex subgroup state $|S_C\rangle$ upto local Pauli operators. For the erroneous 2D cluster state, we would obtain a state $|S_\epsilon\rangle$ such that $| \langle S_\epsilon|S_C\rangle | \geq \sqrt{1-\epsilon}$. From Theorem 2.20, we see that when n is large enough, $\mathrm{TS}(|S_\epsilon\rangle) = n^{\Omega(\log n)}$, and hence $\mathrm{TS}(|2D_\epsilon\rangle) = N^{\Omega(\log N)}$, if $\epsilon \leq 1/2$.

Thus, $\mathrm{TS}_\epsilon(|2\mathrm{D}\rangle) = N^{\Omega(\log N)}$ for $\epsilon \leq 1/2$. □

2.6 Witnessing complex states

In this section we address the problem of verifying the large TS of complex states. Suppose one wants to create complex states such as the complex subgroup states or the 2D cluster state in the lab, in reality the produced states are at some distance away from the target states due to experimental imperfection. How do we verify that the produced state is superpolynomially complex? Full state tomography requires exponentially many operations and is hence not practical. Nonetheless, for complex states that are *stabilizer states*, there exists a complexity witness that can be measured with only a *polynomial* number of basic operations. This witness can be used for verifying the superpolynomial TS of *pure states*.

2.6.1 Subgroup states and their generators

The subgroups states described in Sec. 2.5.5 belong to the class of stabilizer states. An n-qubit stabilizer state $|S\rangle$ is uniquely defined by n mutually commutative stabilizing operators in the Pauli group, $g_1, \ldots, g_n$, satisfying the eigenvalue equation:

$$g_i |S\rangle = |S\rangle . \tag{2.63}$$

Recall that a subgroup state can be defined with the null space of a $(0,1)$ matrix A. The generators of the subgroups states can be read off from the matrix A. Let $R = \mathrm{rank}(A)$; then there are R linearly independent rows r_i $(1 \leq i \leq R)$ in A. For the first R generators, one simply replaces 0 by I and 1 by Z for each of the first R linear independent row. For example, if row r_i is $(0,0,1,0)$, we write $g_i = IIZI$, where the position of the operators denotes the qubit on which they operate on. The remaining generators can be found from the $n-R$ linearly independent vectors c_i that span the null space of A. One replaces 0 with I and 1 with X for each vector, and the generator is the ordered product of these operators.

Proposition 2.23. *The operators g_i defined above are the generators of the stabilizer of $|S\rangle$.*

Proof. Recall that $|S\rangle$ is the uniform superposition of $|x\rangle$ where x is a vector in the null space of A. For the first R generators, we have $g_i |x\rangle = (-1)^{r_i \cdot x} |x\rangle = |x\rangle$, for all $x \in \ker(A)$, hence $g_i |S\rangle = |S\rangle$ for $i = 1, \cdots, R$. For the generators obtained from the $n - R$ linearly independent vectors c_i in the null space of A, we have $g_i |x\rangle = |x \oplus c_i\rangle$, where $\oplus$ is the bitwise addition modulo 2. Note that c_i is in the

null space of A, so $\ker(A) + c_i = \ker(A)$, and hence $g_i |S\rangle = |S\rangle$. This shows that the g_is stabilize $|S\rangle$.

For the commutation relation, it is obvious that the first R generators commutes with each other and so do the $n-R$ obtained from the null space. It remains to show that g_i from row r_i commutes with g_j from c_j. $r_i \cdot c_j = 0 \bmod 2$ implies the number of positions where the entries of both r_i and c_j are 1 must be even. The single-qubit operators in $g_i g_j$ at these positions are $ZX = -XZ$; and since there are an even number of these pairs we see that $g_i g_j = g_j g_i$. □

2.6.2 Complexity witness based on stabilizer witness

Now we show how to construct a complexity witness based on the complex subgroup states. Consider a state $|S_C\rangle$ that satisfies Theorem 2.20. For large n, $\mu_n = 2^{-(n/2)^{1/8}/2}$ is small so one can choose $\epsilon = 1/2 \leq 1 - \mu_n$. Thus $\mathrm{TS}_\epsilon(|S_C\rangle) = n^{\Omega(\log n)}$ implies that any n-qubit state $|\psi\rangle$ such that $|\langle\psi|S_c\rangle|^2 \geq 1/2$ must have $\mathrm{TS} = n^{\Omega(\log n)}$. The superpolynomial TS of these states can be verified by measuring the witness

$$W = \frac{1}{2}\mathbb{1} - |S_C\rangle\langle S_C| . \tag{2.64}$$

A negative value of $\langle W\rangle$ implies that the overlap of the produced state and $|S_C\rangle$ is larger than $1/2$, and hence the TS of the produced state is superpolynomial. However, W as such is not measurable in practice, under the natural constraint that only local measurements are feasible. If one decomposes W into a sum of locally measurable operators, the number of such measurements increases exponentially with the number of qubits [77, 94, 95]. Nonetheless, when $|S_C\rangle$ is a stabilizer state, it is possible to construct a *stabilizer witness* W' with the following properties: If $\langle W'\rangle < 0$ then $\langle W\rangle < 0$; and W' can be decomposed into a sum of a linear number of operators in the Pauli group, which in turn can be measured by a *polynomial* number of basic operations [96]. The stabilizer witness is defined as:

$$W' = (n-1)\mathbb{1} - \sum_{i=1}^{n} g_i. \tag{2.65}$$

Proposition 2.24. $\langle W'\rangle < 0$ *implies* $\langle W\rangle < 0$.

Proof. Consider all the eigenvalue equations of the form (2.63) but with possible eigenvalues ± 1. This defines the set of 2^n common eigenstates of the generators g_i's. Since all the generators are Hermitian operators, the common eigenstates are mutually orthogonal and form a complete basis. One can verify that, in this

basis, the operator $W' - 2W$ is a diagonal matrix with non-negative diagonal entries. Thus, $W' - 2W$ is a positive semi-definite operator; so $\langle W' \rangle < 0$ implies $\langle W \rangle < 0$. □

If in an experiment the expectation value of the stabilizer witness W' is found to be negative, then one can certify that the produced state indeed has $\mathrm{TS} = n^{\Omega(\log n)}$.

While the witness W detects all complex states with a fidelity (with respect to $|S_C\rangle$) larger than $1/2$, W' detects a smaller set. It is necessary to know how close to $|S_C\rangle$ a state $|\psi\rangle$ needs to be for $\langle\psi| W' |\psi\rangle$ to be negative. If the required fidelity is exponentially close to 1 then no state would be detected by W' in practice. For this purpose, we first expand $|\psi\rangle$ as

$$|\psi\rangle = c_1 |S_C\rangle + c_2 |S^{\perp}\rangle , \tag{2.66}$$

where $|S^{\perp}\rangle$ is a state orthogonal to $|S_C\rangle$ and $|c_1|^2 + |c_2|^2 = 1$. We have

$$\langle\psi| W' |\psi\rangle = n - 1 - n|c_1|^2 - |c_2|^2 \sum_{i=1}^{n} \langle S^{\perp}| g_i |S^{\perp}\rangle . \tag{2.67}$$

Since $1 + g_i$ is a positive semi-definite matrix, $\langle S^{\perp}| g_i |S^{\perp}\rangle \geq -1$. Therefore,

$$\langle\psi| W' |\psi\rangle \leq n - 1 - n|c_1|^2 + n|c_2|^2 = 2n - 1 - 2n|c_1|^2. \tag{2.68}$$

Thus, $\langle\psi| W' |\psi\rangle < 0$ when the overlap $|\langle\psi|S_C\rangle|^2 = |c_1|^2 > 1 - 1/(2n)$. So, the loss of fidelity must be smaller than $1/(2n)$ for a state to be detected by W'.

One needs to measure all the generators to estimate $\langle W' \rangle$. With the help of an ancilla qubit, each generator can be measured by applying a circuit of size $O(n)$ followed by a measurement on the ancilla qubit [50] (see Fig. 2.12). But each measurement need to be repeated for obtaining the desired accuracy. When the produced state has the fidelity $|\langle\psi|S_C\rangle|^2 = 1 - \alpha/(2n)$ with $\alpha < 1$, we have $\langle W' \rangle < -(1 - \alpha)$. If the random error in each g_i is δg then $\delta W' = n\delta g$. Thus, to be confident that $\langle W' \rangle < 0$ one needs $n\delta g < 1 - \alpha$, or $\delta g < (1 - \alpha)/n$, which is achievable with a polynomial number of repetitions. Therefore, a correct negative expectation value of W' can be obtained with polynomial effort.

There is a similar stabilizer witness for detecting complex states close to the 2D cluster states. Indeed, the 2D cluster state has $\mathrm{TS}_{1/2} = n^{\Omega(\log n)}$ and is also a stabilizer state. Thus, the witness for the 2D cluster state has the same form as W', with the g_is replaced by the generators of the 2D cluster state. These generators are described in Ref. [93].

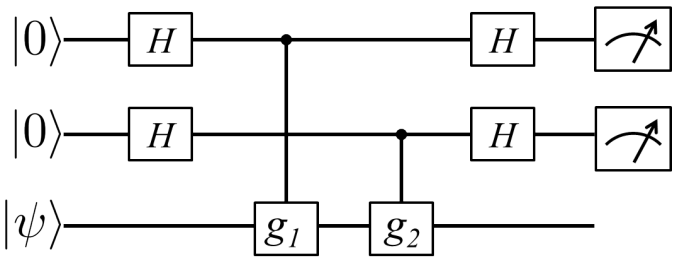


Figure 2.12: A circuit for measuring the generator g_i. The controlled-g_i gate can be decomposed into at most n two-qubit controlled-Pauli gates. A projective measurment of the ancilla qubit in the computational basis gives the outcome of g_i.

2.7 Relation to quantum computation

One of the main motivation of the study of tree size is to investigate the relation between state complexity and quantum computation. To elaborate on this, we can divide all the quantum states into four categories according to their preparation complexity and state complexity (see Fig. 2.13). The set of states with large preparation complexity but small state complexity is presumably empty because preparing simple states should not be too difficult. The states with small state complexity are not useful for quantum computation because they are too simple and hence a classical computer can simulate them efficiently. The states with large preparation complexity are not useful either because quantum computation with these states requires too much resource in space and time. States that are useful for quantum computation should be the ones that have large state complexity yet small preparation complexity. If tree size is a good measure of state complexity, then we might ask: is superpolynomial tree size a necessary condition for the state to provide advantage in some computational task? In this section, we are going to discuss this link in the framework of measurement-based quantum computation and the circuit model of quantum computation.

Note that the complex subgroup states presented in Sec. 2.5.5 belongs to the class of stabilizer states. They have superpolynomial tree size and can be realized by a quantum circuits consisting of $O(n^2/\log n)$ gates [71]. Therefore, these states belong to the bottom left corner of Fig. 2.13. But a computation only consists of only this class of states, does not provide a quantum speed up since stabilizer circuits can be simulated efficiently on a classical computer [71, 97].

2.7.1 Measurement-based quantum computation (MBQC)

There are several theoretical models of quantum computation, including the circuit model and the MBQC model. For the circuit model, the input state can always

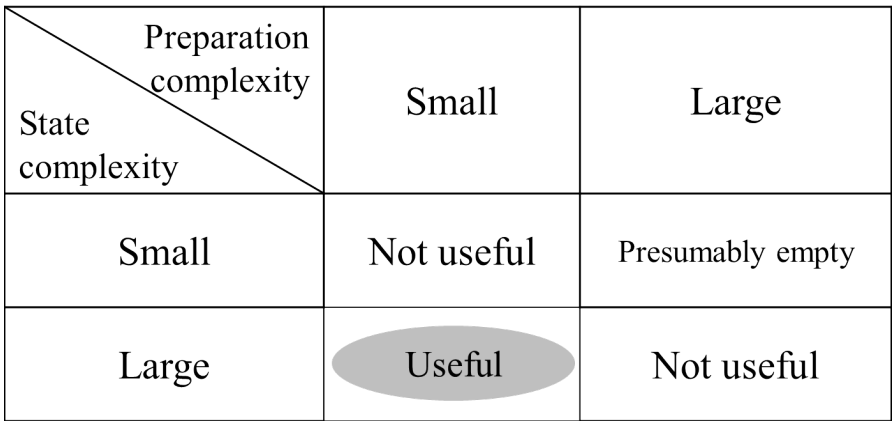


Figure 2.13: Dividing all quantum states into four categories according to their state complexity and preparation complexity. One out of the categories is presumably empty, two are not useful for quantum computation. The states that are useful for quantum computation should have large state complexity and small preparation complexity.

be the simple product state. The quantum power of the computation lies in the gates applied for coherently manipulating single qubits and entangling different qubits [50]. On the contrary, for MBQC, after the initial resource state is prepared, we perform projective measurements on single qubits and feedforward the results for choosing the basis of the next round of measurement [92, 93]. Loosely speaking, all the quantum advantage is contained in the resource state. If this resource state is simple, then MBQC will not offer any real speed up over classical computation. To make this intuition more rigorous, we prove that:

Theorem 2.25. *If the resource state has* TS $=$ poly(n)*, then MBQC can be simulated efficiently with classical computation.*

Proof. Consider the resource state in its minimal tree representation, one sees that at the lowest layer there are a polynomial number of leaves. We will show that it requires polynomial effort to update the tree given a measurement outcome: Assume we measure the ith qubit in the basis $\left\{|\eta\rangle, |\eta^{\perp}\rangle\right\}$ and obtain the the result $|\eta\rangle$, then for every leaf containing qubit i, say $c_\alpha |\alpha\rangle_i + c_\beta |\beta\rangle_i$, we update it to $(c_\alpha \langle\eta|\alpha\rangle + c_\beta \langle\eta|\beta\rangle)|\eta\rangle$. This requires evaluation of the inner products for a polynomial number of leaves. The size of the tree after updating can only get smaller and thus is still polynomial. So, both the tree representation of the state at each step of the computation and the update of the state after a measurement can be carried out with polynomial effort. It follows that MBQC on resource states with polynomial TS can be simulated on a classical computer with polynomial overhead. □

2.7.2 Weaker version of the TreeBQP conjecture

For the circuit model, rather than checking for each algorithm, one would like to have a general proof that small tree size does not provide any computational advantage. In [49], Aaronson raised the question of whether TreeBQP = BPP. This remains an open conjecture, here we prove a weaker version of it.

First let us define what TreeBQP is. Bounded-error quantum polynomial-time (BQP) is the class of decision problems solvable with a quantum Turing machine, with at most 1/3 probability of error. TreeBQP is essentially BQP with the restriction that at each step of the computation, the state is exponentially close to a state with polynomial tree size. In other words, the TS_ϵ of the state is polynomial with $\epsilon = 2^{-\Omega(n)}$ (See Eqn. (2.14)). Since we impose more restrictions, clearly TreeBQP $\subseteq$ BQP. BPP, the classical counterpart of BQP, is the class of decision problem solvable by an NP machine with at most 1/3 probability of error. We can also simulate BPP in TreeBQP: One simply implements reversible classical computation, applies a Hadamard gate on a single qubit and measures in its computational basis to generate random bits if needed. Since each classical bit string can be represented by a quantum product state, TS is n at every steps, so this simulation is in TreeBQP. Thus, we have [49]:

Theorem 2.26. BPP $\subseteq$ TreeBQP $\subseteq$ BQP.

If TreeBQP = BPP, then large tree size is a necessary condition for quantum computers to outperform classical ones. Unfortunately, we can only prove a weaker version of this. For this purpose, we first show a proposition that relates tree size and Schmidt rank.

Note that one can draw a rooted tree in a binary form (each gate has only two children) without changing the number of leaves (its size). Next, for any gate w we denote $S(w)$ as the set of qubits in the state described by the subtree with w as the root. Let $Y|Z$ be a bipartition of the qubits into two sets Y and Z. A $\otimes$ gate is called *separating with respect to* $Y|Z$ when at least one of its children u has the property $S(u) \subseteq Y$ or $S(u) \subseteq Z$. A $\otimes$ gate is called *strictly separating* if its children u_1, u_2 satisfy $S(u_1) \subseteq Y$ and $S(u_2) \subseteq Z$. Then,

Proposition 2.27. *For a bipartition of the qubits into Y and Z, if there exists a polynomial sized tree such that all the $\otimes$ gates are separating with respect to $Y|Z$, then the Schmidt rank of the state with respect to the bipartition $Y|Z$ is polynomial.*

Proof. Identify all the strictly separating $\otimes$ gates in the binary tree. Since the number of leaves N_L is polynomial and the total number of gates in the binary tree is $N_G = N_L - 1$, the number of strictly separating gates, N_S, is also polynomial.

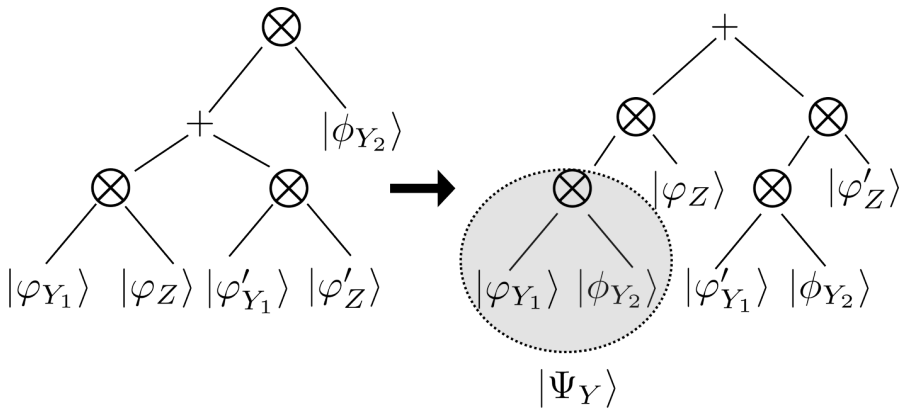


Figure 2.14: A + gate that joins two ⊗ gates with the following property: One of its children are contained in the set of qubit Y and the other contained in Z. The sibling $|\phi_{Y_2}\rangle$ of such a + gate must be strictly contained in either Y or Z, for ⊗ being separating. Now we can exchange the order of + and ⊗ by distributing $|\phi_{Y_2}\rangle$ to $|\varphi_{Y_1}\rangle$ and $|\varphi'_{Y_1}\rangle$. This + gate has the same property as before; and this process can be repeated upward until it reaches the root, transforming the tree into a form similar to the Schmidt decomposition.

It is clearer to look at a representative example in Fig. 2.14. Focus on the + gate that joins two such ⊗ gates, $|\varphi_{Y_1}\rangle\otimes|\varphi_Z\rangle$ and $|\varphi'_{Y_1}\rangle\otimes|\varphi'_Z\rangle$. Since this + gate contain qubits in both Y and Z, and the ⊗ gate at the top is separating, the qubits under the sibling of the + gate must be contained strictly in either Y or Z. Without lost of generality, let them be contained in Y and denote their state as $|\phi_{Y_2}\rangle$. We can exchange the + gate and the ⊗ gate at the top so that the state becomes $(|\varphi_{Y_1}\rangle \otimes |\phi_{Y_2}\rangle) \otimes |\varphi_Z\rangle + (|\varphi'_{Y_1}\rangle \otimes |\phi_{Y_2}\rangle) \otimes |\varphi'_Z\rangle$. Now let us relabel $|\varphi_{Y_1}\rangle \otimes |\phi_{Y_2}\rangle$ as $|\Psi_Y\rangle$ and $|\varphi'_{Y_1}\rangle \otimes |\phi'_{Y_2}\rangle$ as $|\Psi'_Y\rangle$, the state can be written as $|\Psi_{Y_1}\rangle \otimes |\varphi_Z\rangle$ and $|\Psi'_{Y_1}\rangle \otimes |\varphi'_Z\rangle$. The same process can be applied upward until these + gates joins at the root. In the final form of the tree, one sees that the state has a form similar to the Schmidt decomposition:

$$|\psi\rangle = \sum_{i=1}^{N_S} |\Psi_Y\rangle_i \, |\Psi_Z\rangle_i \,, \tag{2.69}$$

where $|\Psi_Y\rangle_i$ contain qubits in Y and $|\Psi_Z\rangle_i$ qubits in Z. N_S, the number of terms in this Schmidt-like decomposition upper bounds the true Schmidt rank, hence the Schmidt rank is polynomial. □

Now suppose that at every step of the quantum computation, Proposition 2.27 is satisfied for all bipartitions, then the Schmidt rank is polynomial for all bipartitions. It follows from a theorem by Vidal [53] that the computation can be efficiently simulated with classical computers.

There are states that do not satisfy the condition of Proposition 2.27, one example is the optimal tree of the most complex four qubit states (see Eqn. (2.30)). There are also states with polynomial TS that do not satisfy Vidal's criteria, hence do not satisfy Proposition 2.27 for some bipartitions. For example, the state $\left(\frac{|00\rangle+|11\rangle}{\sqrt{2}}\right)^{\otimes n/2}$ has polynomial TS, but there is a bipartition for which the Schmidt rank is $2^{(n/2)}$.

2.8 Conclusion

In this chapter, we have introduced the concept of tree size, a complexity measure for pure multiqubit states. The results of tree size of few qubit states shows that tree size is closely related to multipartite entanglement. Examples of most complex two-, three- and four-qubit states were given together with their ϵ-tree size. It is worth mentioning that the most complex four qubit state admits a compact decomposition that is not recursive; changing of the order of qubits in the branches is necessary.

Moving on to the many qubit case, we divide states into simple and complex with polynomial tree size as the separation. By explicit decomposition, we have shown a few examples of simple states, though they contain multipartite entanglement. Matrix product states with bounded bond dimension are also simple.

A mathematical theorem by Raz allows us to prove superpolynomial lower bounds on tree size. Examples of complex states include the immanant state, Deutsch-Jozsa state, Shor's state and subgroup states. We proved that the 2D cluster state, a universal resource state for measurement-based quantum computation has superpolynomial tree size. The superpolynomial tree size of subgroup states can also be verified by a witness involving polynomial number of local measurements.

We discussed the relation between tree size quantum computation. In the measurement-based model, we have shown that superpolynomial tree size of the initial resource states is necessary otherwise the computation can be simulated classically with a polynomial overhead. In the circuit model, we are able to show a weaker version of the TreeBQP conjecture.

In conclusion, tree size, though related, is a different concept from multipartite entanglement and macroscopicity. The relation between tree size and quantum computation can be seen in two perspective. First, states present in Deutsch-Jozsa algorithm and Shor's algorithm are complex; Second, useful resource states in the MBQC model must be complex.

2.8.1 Technical open problems

Besides the TreeBQP=BPP conjecture, there are also technical open problems related to tree size:

Tree size for mixed states

We can formally generalize tree size to mixed states via a min-max definition:

$$\mathrm{TS}(\rho) = \min \Big[\max \mathrm{TS}(|\psi_i\rangle) \Big],$$

where the minimization done over all decomposition $\rho = \sum p_i \, |\psi_i\rangle \, \langle\psi_i|$. Can techniques similar to that of Raz theorem be used to bound the tree size of mixed states? Can this be applied to rule out useless mixed states in quantum algorithms?

Exponential lower bound

Is there a way, other than number counting argument, to show *exponential* lower bound on tree size? Can one construct an explicit state with exponential tree size?

Experimental demonstration

Have states with superpolynomial tree size been produced in experiments? Scaling of tree size can only be defined for *a family of* states, how can one conclude if the states created in the experiment have superpolynomial tree size?

CHAPTER 3

NONLOCALITY

In this chapter, we will review some basic notions of nonlocality, as well as some tools that are useful in the study of nonlocality.

3.1 Bell scenario

A Bell experiment consists of a source that emits particles to a number of spatially separated labs, $\{A, B, \cdots\}$. In each lab the experimentalists may carry out a number of measurements on the particle, $\{X, Y, \cdots\}$, each producing outcomes $\{a, b, \cdots\}$. Finally, the joint distribution of the outcomes given certain input is recorded. The experiment is described by the probability distribution $P(a, b, c, \cdots, |X, Y, Z, \cdots)$, which is the main subject of study in this framework of nonlocality. A Bell scenario is characterized by the number of parties, number of measurements each party could apply and the number of outcomes of each measurement.

Let us consider the simple case where there are two parties, two inputs, d outcomes, the so-called $(2, 2; d, d)$ scenario. There are a few conditions that we impose on the probability distribution:

(1) Positivity:

$$P(a, b|X, Y) \geq 0, \forall X, Y. \tag{3.1}$$

(2) Normalization:

$$\sum_{a,b} P(a, b|X, Y) = 1, \forall X, Y. \tag{3.2}$$

The first two are natural requirements for a probability distribution to be valid. Furthermore, we require

(3) No-signalling:

$$\sum_b P(a,b|X,0) = \sum_b P(a,b|X,1) \equiv P(a|X), \forall X \tag{3.3}$$

for Alice, and similarly for Bob,

$$\sum_a P(a,b|0,Y) = \sum_a P(a,b|1,Y) \equiv P(b|Y), \forall Y. \tag{3.4}$$

The no-signalling condition says that one cannot send a signal with such a probability distribution, i.e. Alice's measurement outcomes do not depend on Bob's inputs and vice versa. If this condition were to be violated, spatially separated Alice and Bob can achieve faster-than-light communication hence violate Einstein's special relativity.

We can study these distributions from a geometric point of view. A particular probability distribution is regarded as a vector in the space of probability distributions. A convex combination of points in the probability space is also a valid probability distribution. Geometrically, $P(v) = vP_1 + (1-v)P_2$ represent a point along the line joining P_1 and P_2, while operationally $P(v)$ is the probability distribution of statistically mixing P_1 and P_2 with weight v and $1-v$ respectively. Within the space of distributions, there are a few set of particular interest to us.

3.1.1 No-signalling polytope

The three conditions above are linear inequalities and equalities in probabilities, hence defining a polytope in the space of probability distributions, called the *no-signalling polytope*. A polytope is a convex set with finitely many extremal points. A convenient way to represent a point in the no-signalling space is by the array notation:

$$P^{\text{full}}(a,b|X,Y) := \left(\begin{array}{c|c} P(a,b|0,0) & P(a,b|0,1) \\ \hline P(a,b|1,0) & P(a,b|1,1) \end{array}\right). \tag{3.5}$$

If we choose a and b in the array to run from 0 to $d-1$, then not all the numbers in the array can freely vary due to the normalization and no-signalling constraint. We can remove this redundancy in the representation by deleting the last row and

column for each input, let a, b range from 0 to $d-2$ and adding the marginals:

$$P^{\mathrm{CG}}(a,b|X,Y) := \left(\begin{array}{c||c|c} 1 & P(b|Y=0) & P(b|Y=1) \\ \hline\hline P(a|X=0) & P(a,b|0,0) & P(a,b|0,1) \\ \hline P(a|X=1) & P(a,b|1,0) & P(a,b|1,1) \end{array}\right), \tag{3.6}$$

where $a, b \in \{0, 1, \cdots, d-2\}$. The probability involving the last outcome can be inferred from the marginal, e.g., $P(a, d-1|X,Y) = P(a|X) - \sum_{b=0}^{d-2} P(a,b|X,Y)$, and $P(d-1, d-1|X,Y)$ can be inferred from normalization. The form of Eqn. (3.5) is called the full form, while Eqn. (3.6) is called the no-signalling form or Collins-Gisin form. The dimension of the no-signalling polytope in this $(2,2;d,d)$ scenario is $4d(d-1)$, which is the number of variables in the no-signalling form.

3.1.2 Local polytope

Within the no-signalling polytope, there resides the set of distributions that admit a local hidden variable model, or simply *local distributions*:

$$P(a,b|X,Y) = \sum_{\lambda} P(a|X,\lambda)P(b|Y,\lambda)P(\lambda). \tag{3.7}$$

Such a local distribution can always be written as a convex sum of local deterministic distributions, of the following form:

$$P(a = f(X), b = g(Y)|X,Y) = 1. \tag{3.8}$$

A local deterministic distribution is one in which that the outcome deterministically depends on the input. In the table form of representation, the 1's in the table should be in a square grid and the other entries are all 0. For example,

$$P_{LD} = \left(\begin{array}{ccc|ccc} & \vdots & & & \vdots & \\ \cdots & 1 & \cdots & \cdots & 1 & \cdots \\ & \vdots & & & \vdots & \\ \hline & \vdots & & & \vdots & \\ \cdots & 1 & \cdots & \cdots & 1 & \cdots \\ & \vdots & & & \vdots & \end{array}\right)$$

For each input X, there are d possible outcomes, same for Y, so there are d^4 local deterministic strategies in total. In general, there are $d_a^{m_a}\, d_b^{m_b}\, d_c^{m_c} \cdots$ local deterministic points in the $(m_a, m_b, m_c, \cdots; d_a, d_b, d_c, \cdots)$ scenario.

The set of local distributions is a polytope, called the *local polytope*, its facets being Bell inequalities and positivity constraints. The local deterministic points provide a natural way to characterize the local set. It translates to the problem of finding all the facets of the convex hull of a set of extremal points. Unfortunately, this problem soon becomes intractable, because the number of extremal points and the space of probability grows exponentially in the number of parties [98]. Checking whether a point belongs to the local set, however, is efficiently solvable. It can be described as the following linear programme:

$$\begin{aligned} \text{maximise} \quad & v, \\ \text{subject to} \quad & P(v) = vP_{test} + (1-v)P_{mix}, \\ & P(v) = \sum_i c_i P_{LD_i}, \\ & c_i \geq 0, \\ & \sum_i c_i = 1, \end{aligned}$$

where P_{mix} is the completely random distribution:

$$P_{mix}(a,b|X,Y) = \frac{1}{d^2}, \quad \forall a,b,X,Y.$$

$P(v)$ is a point along the line connecting P_{mix} and P_{test}, when $v = 1$, $P(v) = P_{test}$. The constraints requires $P(v)$ to be local. If $v_{\max} < 1$, the test distribution is non-local; if $v_{\max} \geq 1$, the test distribution is local, especially when $v_{\max} = 1$, we say that the point is on the boundary of the local set. See Fig. 3.1 for an geometric illustration of this linear programme.

3.1.3 An example: the CHSH scenario

The simplest non-trivial case is the $(2,2;2,2)$ scenario, also known as the CHSH scenario. In this case, the no-signalling polytope is 8-dimensional. To specify a point in this no-signalling polytope, one can write:

$$P = \left(\begin{array}{c||c|c} & P(0|Y=0) & P(0|Y=1) \\ \hline\hline P(0|X=0) & P(0,0|0,0) & P(0,0|0,1) \\ \hline P(0|X=1) & P(0,0|1,0) & P(0,0|1,1) \end{array}\right). \tag{3.9}$$

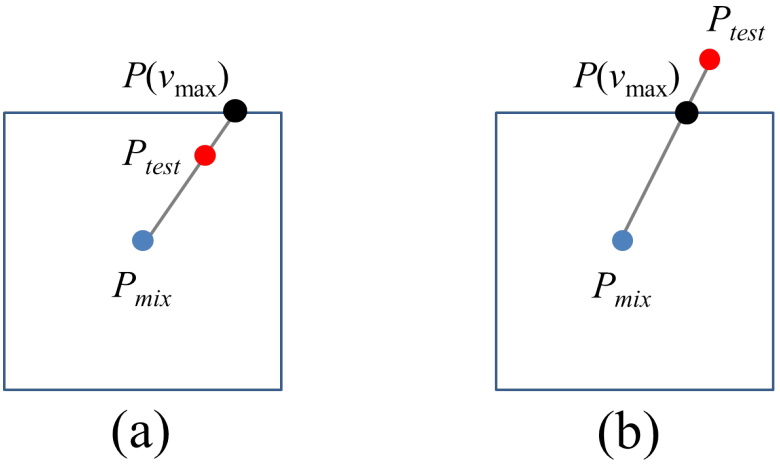


Figure 3.1: A geometric view of the linear programme that tests the locality of the distribution, P_{test}. The local polytope is represented by the square, with the maximally mixed distribution, P_{mix}, at its centre. $P(v)$ is a point on the straight line connecting P_{mix} and P_{test}. The programme aims to maximize v while maintaining $P(v)$ local. (a) If the test distribution is local, $v_{\max} \geq 1$; (b) If the test distribution is nonlocal, $v_{\max} < 1$.

For example, one of the 16 local deterministic points:

$$p_{LD} = \left(\begin{array}{c||c|c} & 1 & 1 \\ \hline\hline 1 & 1 & 1 \\ \hline 1 & 0 & 0 \end{array}\right),$$

or a PR-box [99]:

$$p_{PR} = \left(\begin{array}{c||c|c} & 1/2 & 1/2 \\ \hline\hline 1/2 & 1/2 & 1/2 \\ \hline 1/2 & 1/2 & 0 \end{array}\right).$$

A Bell inequality can also be written in this tabular form by writing the Bell coefficient at the corresponding entries. For example, the CHSH inequality [100]:

$$c_{CHSH} = \left(\begin{array}{c||c|c} & -1 & 0 \\ \hline\hline -1 & 1 & 1 \\ \hline 0 & 1 & -1 \end{array}\right).$$

The bound for local bound for CHSH written in this way is $-1 \leq c_{CHSH} \cdot p_{Local} \leq 0$. Besides positivity, all the facets of the local polytope in this scenario are defined by the CHSH inequalities with permutation of parties or relabelling of inputs and

outcomes: 4 ways of relabelling that defines 2 facets each, so 8 facets in total. For each CHSH inequality, one will find that 8 out of the 16 local deterministic point saturate the bound 0 and the other 8 at -1. There are 8 PR-boxes, one above each facet.

This concludes the example of the CHSH scenario. Next we will turn to another set of distributions that requires more sophisticated treatment: the quantum set.

3.2 The quantum set

Quantum distributions are defined as those arise from measurements on a quantum system. If we consider only a bipartite distribution:

Definition 3.1. *A distribution P is said to be quantum if there exists a quantum state $|\psi\rangle$ and $\{E_A^{a|X}\}$ be a set of measurement operators for Alice and $\{E_B^{b|Y}\}$ be a set of measurement operators for Bob, such that:*

$$P(a,b|X,Y) = \langle\psi| E_A^{a|X} E_B^{b|Y} |\psi\rangle\,, \tag{3.10}$$

and the following holds,

1. $E_A^{a|X\dagger} = E_A^{a|X}$ *and* $E_B^{b|Y\dagger} = E_B^{b|Y}$ *(hermiticity);*
2. $E^{a|X}E^{a'|X} = \delta_{aa'}E^{a|X}$ *and* $E^{b|Y}E^{b'|Y} = \delta_{bb'}E^{b|Y}$ *(orthogonality);*
3. $\sum_a E_A^{a|X} = \mathbb{1}$ *and* $\sum_b E_B^{b|Y} = \mathbb{1}$ *(completeness);*
4. $[E_A^{a|X}, E_B^{b|Y}] = 0$ *(commutativity).*

The state can be taken to be pure and measurements to be projective since we do not restrict the dimension of the state. Condition 3 ensures the normalization condition, also implies that there is a redundancy in the definition of the measurement operators.

Quantum distributions can violate Bell inequality and any local deterministic point also admits a quantum model, so clearly $L \subset Q$. From the commutativity requirement, one can see that quantum distribution are also no-signalling. But quantum mechanics cannot reproduce all the no-signalling distributions [101], so $Q \subset NS$.

The quantum set does not have a simple characterisation as the local or no-signalling set. It is convex but not a polytope. The best known characterisation of the quantum set is via the Navascués–Pironio–Acín (NPA) hierarchy [102]. NPA

introduced an infinite hierarchy of conditions necessarily satisfied by any quantum correlations, each level formulated as the existence of a positive semidefinite matrix.

Let us illustrate the basic idea of the hierarchy. First, consider the following mathematical lemma,

Lemma 3.2. *Let $\mathcal{O} = \{O_1, O_2, \cdots, O_n\}$ be the set of n operators and any state $|\psi\rangle$, then the n-by-n matrix Γ defined by*

$$\Gamma_{ij} = \langle\psi| O_i^\dagger O_j |\psi\rangle , \tag{3.11}$$

is semi-definite positive, $\Gamma \geq 0$.

Proof. For any vector x, we have

$$\begin{aligned} x^\dagger \Gamma x &= \sum_{ij} x_i^* \Gamma_{ij} x_j \\ &= \langle\psi| \left(\sum_i x_i^* O_i^\dagger\right) \left(\sum_j x_j O_j\right) |\psi\rangle \\ &= || \sum_j x_j O_j |\psi\rangle ||^2 \geq 0, \end{aligned}$$

so Γ is a positive matrix. □

Since the lemma is true for any set of operators, in particular, we can take $\mathcal{O}$ to be a set of operators that are linear combinations of measurement operators. As such, the observed probability will enter the matrix Γ as constraints on its entries, for example, $\Gamma_{ij} = \langle\psi| E^{a|X} E^{b|Y} |\psi\rangle = P(a, b|X, Y)$. We can abstractly represent this as:

$$\sum_{ij} (F_k)_{ij} \Gamma_{ij} = g_k(P), \quad k = 1, \cdots, m. \tag{3.12}$$

Depending on how one chooses the set of operators $\mathcal{O}$, different level of restriction is imposed on the set of distributions. Any $n \times n$ positive semidefinite matrix $\Gamma \geq 0$ satisfying the linear constraints (3.12) will be called a certificate associated to $\mathcal{O}$. Here we present two ways of constructing a hierarchy of conditions.

The first way, the NPA hierarchy introduced in [102], is based on the length of sequences in $\mathcal{O}$. Let a *sequence* S be a product of operators from the set $\{\mathbb{1}\} \cup \{E_A^{a|X}\} \cup \{E_B^{b|Y}\}$. Examples of a sequence are $E_A^{a|X}$, $E_A^{a|X} E_A^{a'|X'} E_B^{b|Y}$. Some sequence may produce the null operator, for example $E_A^{a|X} E_A^{a'|X} = 0$ for $a \neq a'$. We ignore the null sequence. The *length* of a sequence, $|S|$ is defined as the minimum

number of projectors that generates the sequence. For example, $|E_A^{a|X} E_B^{b|Y} E_A^{a|X}| = |E_A^{a|X} E_A^{a|X} E_B^{b|Y}| = |E_A^{a|X} E_B^{b|Y}| = 2$. By convention, the length of the identity operator is $|\mathbb{1}| = 0$. Then we define $\mathcal{S}_n$ to the set of sequences of length at most n:

$$\begin{aligned}
\mathcal{S}_0 &= \{\mathbb{1}\}, \\
\mathcal{S}_1 &= \mathcal{S}_0 \cup \{E_A^{a|X}\} \cup \{E_B^{b|Y}\}, \\
\mathcal{S}_2 &= \mathcal{S}_1 \cup \{E_A^{a|X} E_A^{a'|X'}\} \cup \{E_B^{b|Y} E_B^{b'|Y'}\} \cup \{E_A^{a|X} E_B^{b|Y}\}, \\
\mathcal{S}_3 &= \mathcal{S}_2 \cup \cdots \\
&\vdots
\end{aligned}$$

It is clear that $\mathcal{S}_0 \subseteq \mathcal{S}_1 \subseteq \mathcal{S}_2 \cdots$, and any operator O can be written as linear combination of operators in $\mathcal{S}_n$ when n is large enough. The set Q_n is the set of distributions with a positive certificate associated to $\mathcal{S}_n$. For example, a probability distribution is said to be in Q_1 if the following semi-definite programme return a non-negative λ:

$$\begin{aligned}
\max \quad & \lambda \\
\text{subject to} \quad & \Gamma - \lambda \mathbb{1} \geq 0,
\end{aligned}$$

where Γ is symmetric and the entries are

$$\Gamma = \begin{pmatrix}
1 & P(0|X=0) & P(0|X=1) & P(0|Y=0) & P(0|Y=1) \\
 & P(0|X=0) & v_1 & P(00|00) & P(00|01) \\
 & & P(0|X=1) & P(00|10) & P(00|11) \\
 & & & P(0|Y=0) & v_2 \\
 & & & & P(00|11)
\end{pmatrix},$$

where v_1 and v_2 are not observable probabilities, hence enter as SDP variables.

Due to the inclusion relation of $\mathcal{S}$, one has $Q_1 \supseteq Q_2 \supseteq \cdots \supseteq Q_\infty = Q$. The sufficient part of the hierarchy ($Q_\infty = Q$) was shown in [102], and we refer the readers to it for more detail.

The second way is by restricting the length of the sequence on the local level, introduced by Moroder et al. [103], so we refer to it as the Moroder hierarchy. Let us define

$$\begin{aligned}
\mathcal{L}_1 &= \{\mathbb{1}, E_A^{a|X}\} \times \{\mathbb{1}, E_B^{b|Y}\} \\
\mathcal{L}_2 &= \{\mathbb{1}, E_A^{a|X}, E_A^{a|X} E_A^{a'|X'}\} \times \{\mathbb{1}, E_B^{b|Y}, E_B^{b|Y} E_B^{b'|Y'}\} \\
&\vdots
\end{aligned}$$

A hierarchy of certificate is similarly defined as Q_n. The set of local level n, Q'_n, is the set of distribution with a positive certificate associated to $\mathcal{L}_n$. For example, the certificate correspond to local level 1, is a symmetric $\Gamma \geq 0$, where

$$\Gamma = \begin{pmatrix} 1 & p_A(0|0) & p_A(0|1) & p_B(0|0) & p(00|00) & p(00|10) & p_B(0|1) & p(00|01) & p(00|11) \\ & p_A(0|0) & v_1 & p(00|00) & p(00|00) & v_2 & p(00|01) & p(00|01) & v_3 \\ & & p_A(0|1) & p(00|10) & v_2 & p(00|10) & p(00|11) & v_3 & p(00|11) \\ & & & p_B(0|0) & p(00|00) & p(00|10) & v_4 & v_5 & v_6 \\ & & & & p(00|00) & v_2 & v_5 & v_5 & v_7 \\ & & & & & p(00|10) & v_6 & v_8 & v_6 \\ & & & & & & p_B(0|1) & p(00|01) & p(00|11) \\ & & & & & & & p(00|01) & v_3 \\ & & & & & & & & p(00|11) \end{pmatrix}.$$

Note that the constraints on the SDP variable: even though they are not observables, but since they correspond to the same operator, their expectation value must be the same. For example, we have assigned $\langle E_A^{0|0}(E_A^{0|1}E_B^{0|0})\rangle = \langle (E_A^{0|0}E_A^{0|1})E_B^{0|0}\rangle = v_2$. Incidentally, the set Q'_1 is called Q_{1+AB} in the work of NPA. One can see that $\mathcal{S}_n \subset \mathcal{L}_n \subset \mathcal{S}_{2n}$, so in the limit $n \to \infty$, Q'_∞ also converges to the quantum set Q.

Between the above-mentioned two different ways of constructing the hierarchy of certificates that converge to the quantum set in the limit, we found the Moroder hierarchy more suitable for some quantum information applications. The tensor product structure of the set $\mathcal{L}_n$ allows imposing constraints that are not possible in the NPA hierarchy, such as the positivity of the partial transpose of the moment matrix. This has led to applications including dimension witness, Bell violation with positive partial transpose states and device-independent entanglement witness involving an arbitrary number of parties [103].

CHAPTER 4

DIMENSION WITNESS

4.1 Motivation

4.1.1 Dimension of quantum systems

The number of perfectly distinguishable states in the physical system, or *dimension*, is crucial to the information carrying and processing capacity of the system. For classical systems, the dimension is given by the number of states. For example, a switch has two states, on and off; so it can encode a *bit* of information. In our notation, we denote it as 2_c; similarly we can have 3_c (a trit) and 4_c (a quart) for a classical system with three or four different states and so on. This is not so for quantum systems for there are infinitely many states even for a two level quantum system. The dimension of the quantum system is defined as the dimension of the Hilbert space where the states reside. A two level quantum system is called a qubit, or 2_q, analogously we can have 3_q (a qutrit), 4_q (a ququart) and so on.

4.1.2 Dimension witnesses

If someone claims to have control over a classical or quantum d dimensional system, the claim can be verified by some criteria. These criteria are termed dimension witnesses (DWs).

The simplest DW consists in testing whether the system can encode and decode faithfully one $d$it of information. However, due to Holevo's bound, one qu$d$it can encode at most one $d$it, hence this elementary DW cannot distinguish whether the information carrier is quantum or classical.

Recently, more elaborated tests have been proposed that discriminate classical and quantum dimension in a "device-independent" framework. The first device

independent dimension witness (DIDW) was proposed by Brunner et al. [104], based on violation of a Bell inequality. The intuition behind is that the strong correlation required to violate the Bell inequality is only achievable by states that are entangled in many dimensions. Later dimension witnesses in the prepare-and-measure (PM) scenario have been considered [105]. In the PM scenario, a black box called the *state preparator* prepares a classical or quantum state on request. The box has a set of N buttons and when the button $X \in \{1, \cdots, N\}$ is pressed the state ρ_x is sent to the *measurement device*. The measurement device performs a measurement $Y \in \{1, \cdots, y\}$ and yielding an outcome $b \in \{1, \cdots k\}$. The outcome probability distribution $P(b|X, Y)$ is then analysed to lower bound the classical or quantum dimension of the states prepared.

Since any system is fundamentally quantum mechanical, what do we mean by a system with classical dimension d? If the states that lead to a violation of the DW have to possess some quantum properties, we then say that the system has quantum dimension d. In the PM scenario, if the dimension of the system d is known, to violate the dimension witness of quantum dimension d, one needs to prepare a set of states $\{\rho_x\}_{x=0,1,\cdots,N}$, where $\rho_x \in \mathcal{H}^d$ and not all ρ_x commute with each other. If one can prepare only commuting states of dimension d, we say that one has access to classical dimension d. In the Bell scenario, since entanglement is required to violate a Bell inequality, the dimension certified in that case is always genuine quantum dimension.

The DW based on the PM scenario are handy because they can bound the dimension of the system produced by a single source. However, the shortcoming of their approach is that any prepare-and-measure statistics can be simulated with sufficiently high-dimension classical systems. In other words, if we know the dimension d of the system, we can discriminate whether it is a classical $d$it or quantum $d$it by the violation of the dimension witness $W(d_q) > W(d_c)$; but there also exists a D such that $W(D_c) > W(d_q)$. Moreover, for many proposed DW, $D = d + 1$, for example we only need a classical quart to simulate (to violate the dimension witness as much as) a qutrit. It is in principle possible to construct a prepare-and-measure DW which requires D_c exponential in d_q to simulate the violation (see for example the distributed Deutsch-Jozsa problem in [106]).

Dimension witnessing in other scenarios has also been investigated, such as those based on quantum random access code [107], contextuality [108], observable dynamics [109] as well as Hardy's paradox [110].

There has been plenty of effort in experimental demonstrations of dimension witnesses, for instance, Hendrych et al. and Ahrens et al. have witnessed classical and quantum dimension up to 4 with the prepare-and-measure scheme [111, 112];

Dada et al. have certified quantum system up to dimension 12 [113]; Krenn et al. generated a 100×100 dimension quantum system with the spatial mode of light [114], holding the record of the largest dimension so far. However, the certifications of the last two experiments is not device independent.

In this chapter we will present a fully device independent dimension witness in the Bell scenario, based on the CGLMP_4 inequality, with no assumptions on the form the state used in the experiment.

4.1.3 An overlooked feature

Some part of this project was carried out during my Honours project. That includes the characterization of the CGLMP_4 polytope, and bound computed based on three outcome measurements. During my PhD candidature, a more rigorous bound based on negativity were derived, and a simplified measurement scheme for experiment were proposed. An overlooked feature of the dimension witness in this project was discovered in the very late stage of my PhD during the preparation of the paper. This feature makes experimental demonstration of this dimension witness superfluous.

First, we realize that high dimension entanglement *per se* is trivial. A ququart can be encoded with two qubits, for example with the standard binary encoding, $|00\rangle \to |0\rangle, |01\rangle \to |1\rangle, |10\rangle \to |2\rangle, |11\rangle \to |3\rangle$. A maximally entangled ququart pair is then two maximally entangled qubits pairs (omitting normalization):

$$|00\rangle + |11\rangle + |22\rangle + |33\rangle \simeq (|00\rangle + |11\rangle)_{A_1B_1} \otimes (|00\rangle + |11\rangle)_{A_2B_2}. \tag{4.1}$$

In this sense, the generation of ququarts is not different than the generation of two qubits. This is almost trivial, because what we have is a source that repeatedly generates pairs of particles, we can generate two pairs if we can generate one.

Second, as we will see in Sec. 4.5.1, the measurements required to violation this dimension witness can be performed on each qubit separately and sequentially. The measurement basis on the second qubit depends on the outcome of the measurement on the first qubit, but this can be done by classical processing; no coherent manipulation between the two qubits is required. In other words, the violation of the dimension witness, certifies the generation of entangled ququarts (trivially equivalent to two pairs of entangled qubits) and coherent manipulation on *qubit* systems only.

The rest of this chapter is organized as follows: we first describe our scenario in which we apply our dimension witness in Sec. 4.2; in Sec. 4.3 we introduce the CGLMP inequality, which is the inequality that we use as the dimension witness;

in order to find the maximal violation of CGLMP_4 inequality, in Sec. 4.4 we derive an upper bound of this maximal violation based on negativity, and a lower bound based on a specific form of measurement; finally in Sec. 4.5, we discuss more about the overlooked feature of this dimension witness that makes it superfluous for any experimental realization.

4.2 The scenario

Let us consider a Bell scenario. A source sends out particles to Alice and Bob, who then apply two suitable chosen measurements with d outcomes. They collect statistics about the experiment $P(a, b|X, Y)$ and compute the violation of a Bell inequality.

What information can one obtain from the measurement statistics alone? We already know that if $P(a, b|X, Y)$ violates a Bell inequality (modulo the various loopholes), then the source must have prepared some entangled quantum state. Surprisingly, more information can be deduced from the statistics alone. To the extreme, self-testing of a quantum state is possible: one can certify the state produced from the source is equivalent to an ideal state up to local isometry. Self-testing is black box tomography.

Our task here is simpler. We do not need to know the precise form of the state, but we would like to give a lower bound on the dimension of the system. The dimension of the system maybe large, but we are interested in the dimension in which entanglement is present, the *entangled dimension*. We would like to give this lower bound device independently, with no assumption being made on the state of the particle the source prepares nor on the measurements that Alice and Bob performs.

The simplest bipartite Bell inequality the CHSH inequality can be violated up to $2\sqrt{2}$ quantum mechanically. This maximal value can be achieved with qubit in the maximally entangled state. So CHSH inequality cannot be used as dimension witness to certify higher entangled dimensions. To accomplish this task, other Bell inequalities are needed.

4.3 CGLMP inequality

A Bell inequality for this $(2, 2; d, d)$ scenario is the CGLMP inequality first introduced in [115]. Here, we are going to use an equivalent form as in Eqn.(41) of

Ref. [116]:

$$I_d = \langle \mathcal{I}_d, P \rangle - 2 \leq 0, \tag{4.2}$$

with

$$\mathcal{I}_d = \left(\begin{array}{c|c} J_d & J_d^T \\ \hline J_d^T & -J_d^T \end{array} \right), \tag{4.3}$$

where J_d an upper triangular array filled with 1, ^{T} the transposition, and $\langle \cdot, \cdot \rangle$ denotes the sum of term-by-term multiplication. The local bound of I_d is $I_d \leq 0$, achievable by some local deterministic points, for example,

$$P(0,0|X,Y) = 1, \quad \text{for all } X, Y.$$

The maximal no-signalling violation $I_d = \frac{d-1}{d}$ is achieved by the generalized PR-box [117]:

$$P(a,b|X,Y) = \begin{cases} \frac{1}{d}, & a + b = XY \bmod d, \\ 0, & \text{otherwise.} \end{cases} \tag{4.4}$$

The completely mixed distribution, $P(a,b|X,Y) = \frac{1}{d^2}$, which models the white noise, gives a violation of $I_d = -\frac{d-1}{d}$.

4.3.1 Maximal violation of CGLMP inequality

In this section, we will derive the maximal violation of CGLMP inequality for $d = 3$ and $d = 4$. This is done by first guessing the form of the optimal measurement, computing the maximal violation with these measurements. This provides a lower bound on the maximal violation. Then this is compared with the upper bound on the maximal violation given by the NPA hierarchy.

To maximally violation the CGLMP$_d$ inequality with qudits[1], the following measurement settings are conjectured to be optimal [118–120]:

$$A_X = \{|\Psi_X(a)\rangle\}_{a=0}^{d-1}, \qquad |\Psi_X(a)\rangle = \sum_{k=0}^{d-1} \frac{\omega^{ak}}{\sqrt{d}} (e^{ik\phi_X} |k\rangle), \tag{4.5}$$

$$B_Y = \{|\Phi_Y(b)\rangle\}_{b=0}^{d-1}, \qquad |\Phi_Y(b)\rangle = \sum_{k=0}^{d-1} \frac{\omega^{-bk}}{\sqrt{d}} (e^{ik\theta_Y} |k\rangle), \tag{4.6}$$

[1] the same violation can be achieved by higher dimensional by an embedding of this d dimensional state

where $\omega = e^{\frac{2\pi i}{d}}$ is the d-th root of unity, and the choice for the phase ϕ and θ is:

$$\phi_0 = 0, \phi_1 = \frac{\pi}{d}, \text{ and } \theta_0 = -\frac{\pi}{2d}, \theta_1 = \frac{\pi}{2d}. \tag{4.7}$$

This measurement can be seen as first applying a phase $e^{ik\phi_X}$ (for Bob, $e^{ik\theta_Y}$) to the computational basis, followed by a (for Bob, inverse) discrete Fourier transformation.

From these measurements one can construct the Bell operator,

$$\mathcal{B} = \sum_{a,b,X,Y} c_{abXY} \Pi_X(a) \otimes \Pi_Y(b), \tag{4.8}$$

where $\Pi_X(a) = |\Psi_X(a)\rangle \langle \Psi_X(a)|$ and $\Pi_Y(b) = |\Phi_Y(b)\rangle \langle \Phi_Y(b)|$, and c_{abXY} is the Bell coefficient of CGLMP for $P(a, b|X, Y)$ (cf. Eqn. (4.3)). The eigenvector with the maximum eigenvalue is the maximal violation state (MVS) under these measurement settings.

Let us study the MVS and its corresponding violation for $d = 3$ and $d = 4$. For $d = 3$, numerical evidence shows that the MVS has the following form:

$$|\psi\rangle = \frac{1}{\sqrt{2+\gamma^2}}(|00\rangle + \gamma |11\rangle + |22\rangle), \tag{4.9}$$

for a real positive γ. $\gamma = 1$ correspond to the maximally entangled state (MES). The maximal violation is achieved when $\gamma = \frac{\sqrt{11}-\sqrt{3}}{2}$ [119]. One can check in the Table 1 in [102], this matches the upper bound given by the NPA hierarchy.

For $d = 4$, again numerical evidence suggests a certain form of MVS:

$$|\psi\rangle = \frac{1}{\sqrt{2}}(\cos\theta |00\rangle + \sin\theta |11\rangle + \sin\theta |22\rangle + \cos\theta |33\rangle). \tag{4.10}$$

The violation as a function of the parameter θ can be written as:

$$I_4(\theta) = (-\frac{3}{4} + \frac{C}{2}) + (\frac{1}{2\sqrt{2}} + \frac{S}{\sqrt{2}}) \sin 2\theta + \frac{S}{2} \cos 2\theta, \tag{4.11}$$

where we introduce constants C and S, $C = \cos\frac{\pi}{8}$, $S = \sin\frac{\pi}{8}$. The relation $C - S = \sqrt{2}S$ and $C + S = \sqrt{2}C$ are used to simplify the expression. By taking the derivative, one can find that maximal of value of $I(\theta)$ is attained at θ such that $\tan 2\theta = \frac{1+2S}{\sqrt{2}S}$, and the maximal value is:

$$I_4^* = -\frac{3}{4} + \frac{C}{2} + \frac{1}{2\sqrt{2}}\sqrt{6S^2 + 4S + 1} \approx 0.364\,762. \tag{4.12}$$

This matches the upper bound given by the NPA hierarchy up to numerical precision, hence is the true maximal violation of CGLMP_4 inequality with quantum systems.

The maximally entangled states also has a large violation of the inequality, $I_4(\frac{\pi}{4}) = 0.336\,091$.

4.3.2 Depolarization

In this section, we digress to discuss a local classical processing of the data called depolarization. It will be applied to reduce the number of variables in the probability distribution, as we will see in Sec. 4.4, while keeping violation of CGLMP unchanged. This depolarization is described in [117], but we would like reproduce it for the benefit of clarity.

Due to the symmetry present in the inequality, the number of parameters in P that are relevant for the value of I_d can be reduced. Indeed, there exists a classical post processing, also called a depolarization, that maps all the points in the probability distribution space to a lower dimension space, while keeping the CGLMP violation unchanged [117]. Under the action of this map, every probability distribution is projected to a slice of the no-signalling polytope. This procedure can be described as follows:

Step 1. Alice and Bob add a number k, uniformly chosen from $\{0, \cdots, d-1\}$, to their outcomes:

$$a \to a + k, \quad b \to b + k.$$

Step 2. With probability $\frac{1}{4}$ according shared randomness, Alice and Bob perform one of the four possible processes:

Proc 1.	A : Do nothing,	B : Do nothing;
Proc 2.	$A : x \to \bar{x}, a \to -a,$	$B : b \to -b + y;$
Proc 3.	$A : a \to -a - x,$	$B : y \to \bar{y}, b \to -b;$
Proc 4.	$A : x \to \bar{x}, a \to a + x,$	$B : y \to \bar{y}, b \to b + \bar{y};$

where $\bar{x} = 1 - x$ and the operation on the outcome are done modulo d. These implements $P \to P'$ such that it only depends on the difference of the outcome $\Delta = a - b$ as follows:

$$P'(\Delta|0,0) = P'(-\Delta|0,1) = P'(-\Delta|1,0) = P'(\Delta+1|1,1).$$

The number of free variables is thus reduced to $d-1$.

4.3.3 CGLMP4 polytope

CGLMP with $d=2$ is simply the CHSH inequality. The case for $d=3$ is described in the work of Brunner et al [104]., who first propose CGLMP inequality as a dimension witness. Here, we describe the dimension witness based on CGLMP_4. Before that, we shall describe the no-signalling polytope of $d=4$ outcomes.

After depolarization, a point in the depolarized polytope can be represented in the following table form:

$$P=\frac{1}{4}\left(\begin{array}{cccc|cccc} p_0 & p_3 & p_2 & p_1 & p_0 & p_1 & p_2 & p_3 \\ p_1 & p_0 & p_3 & p_2 & p_3 & p_0 & p_1 & p_2 \\ p_2 & p_1 & p_0 & p_3 & p_2 & p_3 & p_0 & p_1 \\ p_3 & p_2 & p_1 & p_0 & p_1 & p_2 & p_3 & p_0 \\ \hline p_0 & p_1 & p_2 & p_3 & p_1 & p_0 & p_3 & p_2 \\ p_3 & p_0 & p_1 & p_2 & p_2 & p_1 & p_0 & p_3 \\ p_2 & p_3 & p_0 & p_1 & p_3 & p_2 & p_1 & p_0 \\ p_1 & p_2 & p_3 & p_0 & p_0 & p_3 & p_2 & p_1 \end{array}\right), \tag{4.13}$$

with normalization $p_0+p_1+p_2+p_3=1$. We may represent points in this depolarized slice as (p_0,p_1,p_2,p_3). The dimension of the depolarized slice of the no-signalling polytope is three, so we can have a convenient geometric view in 3D space that we are familiar with. It can be represented by a tetrahedron, with the four vertices corresponding to $p_i=1$, a generalized PR-box [117], as shown in Fig. 4.1. The surfaces of this tetrahedron represents the positivity constraints.

One can identify the local polytope by applying the depolarization to all the local deterministic points. Resultant local extremal points are:

$$\begin{aligned} &L_1=(\tfrac{3}{4},\tfrac{1}{4},0,0), \quad L_2=(0,\tfrac{3}{4},\tfrac{1}{4},0), \quad L_3=(0,0,\tfrac{3}{4},\tfrac{1}{4}), \quad L_4=(\tfrac{1}{4},0,0,\tfrac{3}{4}),\\ &L_5=(\tfrac{1}{2},0,\tfrac{1}{4},\tfrac{1}{4}), \quad L_6=(\tfrac{1}{4},\tfrac{1}{2},0,\tfrac{1}{4}), \quad L_7=(\tfrac{1}{4},\tfrac{1}{4},\tfrac{1}{2},0), \quad L_8=(0,\tfrac{1}{4},\tfrac{1}{4},\tfrac{1}{2}). \end{aligned}$$

The convex hull of these points is the depolarized local set. The facets of the local polytope are Bell inequalities or trivial facets given by the positivity constraint.

The facet defined by CGLMP_4 inequality is given by the formal inner product, $\langle \mathcal{I}, P\rangle - 2 = 0$, simplifies to

$$p_0 - 2p_1 - p_2 - \frac{1}{4} \leq 0. \tag{4.14}$$

In fact, the other three CGLMP$_4$ inequalities can be found by simultaneously permuting the outcome of Alice and Bob, as the following operations:

Operation 1. $b \to b + 2$;

Operation 2. $a_{X=1} \to a_{X=1} + 2,\ b_{Y=0} \to b_{Y=0} - 1,\ b_{Y=1} \to b_{Y=1} + 1$;

Operation 3. $a_{X=1} \to a_{X=1} + 2,\ b_{Y=0} \to b_{Y=0} + 1,\ b_{Y=1} \to b_{Y=1} - 1$.

The four CGLMP inequalities on this slice can be summarized as

$$p_i - 2p_{i+1} - p_{i+2} - \frac{1}{4} \leq 0, \quad i \in \{0, 1, 2, 3\}. \tag{4.15}$$

One can easily check that the PR-boxes corresponding $p_i = 1$ violates each of these inequalities up to the algebraic maximum of $\frac{3}{4}$.

Compare this with the convex hull of the extremal points of the local polytope, one can find that there is another class of non-trivial facets that is not given by CGLMP$_4$ inequality. These two facets in this class are:

$$\frac{1}{4} \leq p_0 + p_2 \leq \frac{3}{4}. \tag{4.16}$$

Analysis shows that these are simply liftings of the CHSH inequality, with 0 and 2 grouped to be outcome“0” and 1 and 3 to be outcome “1”.

A geometric representation of the CGLMP$_4$ polytope in three dimension is shown in Fig. 4.1.

4.4 Dimension witness with CGLMP$_4$

In this section, we aim to bound the violation of the CGLMP inequality if we restrict ourselves to only qutrits. This cannot be achieve by directly using the NPA hierarchy, since the hierarchy do not restrict the dimension.

We first derive an upper bound on the maximum qutrit violation, $I_4^{(3)}$. The method is similar to that of Moroder et al. [103]. A lower bound on the maximum qutrit violation is also discussed, by restricting ourselves to three outcome measurements.

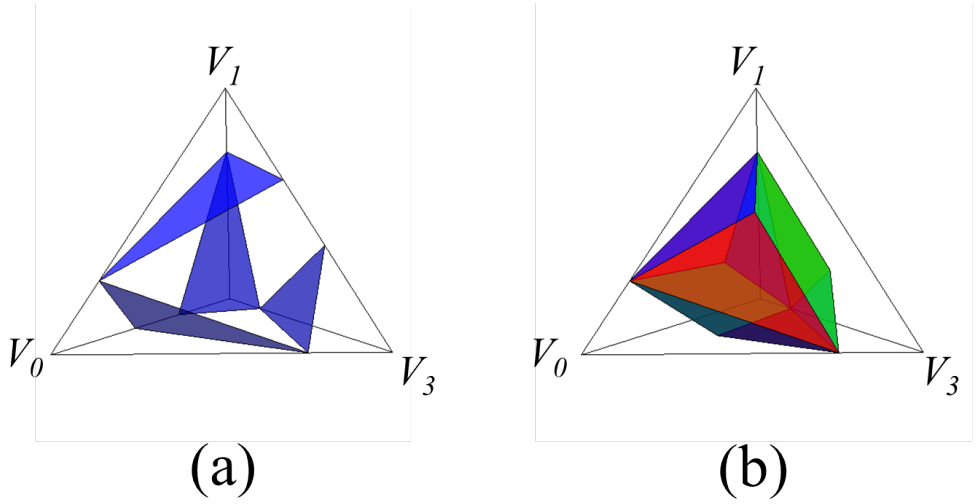


Figure 4.1: The CGLMP$_4$ polytope. The tetrahedron represents the NS polytope with each vertices being a PR-box. Shown in (a) are the CGLMP$_4$ inequalities. Shown in (b) is the local polytope defined by three types of facets: the CGLMP facets (blue), the lifted CHSH facets (green) and positivity facets (red).

4.4.1 Upper bound on the maximum qutrit violation

Following the method in [103], one can derive an upper bound on $I_4^{(3)}$ based on negativity. Negativity is a measure of bipartite entanglement defined as:

$$\mathcal{N}(\rho) = \frac{||\rho^{T_A}|| - 1}{2} \tag{4.17}$$

where ρ^{T_A} denotes the partial transposition of the state ρ. If we fixed the dimension of the system, we have

Proposition 4.1. *The negativity of a d-by-d quantum state is bounded by* $\frac{d-1}{2}$.

Proof. Negativity is convex, so one needs to consider only the case where the state is pure. One can write it in the Schmidt form

$$|\psi\rangle = \sum_{i=0}^{d-1} \lambda_i |ii\rangle , \tag{4.18}$$

with $\sum \lambda_i = 1$. The negativity of this state can be calculated as

$$\mathcal{N}(|\psi\rangle \langle\psi|) = \sum_{i \neq j} \sqrt{\lambda_i \lambda_j}. \tag{4.19}$$

This can be optimized using the method of Lagrange multiplier, with the constraint $\sum_i \lambda_i = 1$. The maximal value $\frac{d-1}{2}$ is achieved when $\lambda_i = \frac{1}{d}$ for all i, which corresponds to the maximally entangled state. □

From this proposition, we can see that a large negativity will give a lower bound to the dimension fo the system.

Next we shall relate negativity with observed statistics, the Bell violation in this case. This is done through the matrix of moments χ at local level ℓ (c.f. Sec. 3.2). For a quantum state ρ and measurements $M^A_{a|X}$ and $M^B_{b|Y}$, this matrix is defined as:

$$\chi[\rho] = \sum_{i,j,k,l} |ij\rangle_{\bar{A}\bar{B}} \langle kl| \, \chi^{kl}_{ij}, \tag{4.20}$$

where $\chi^{kl}_{ij} = \mathrm{tr}[\rho_{AB} A^\dagger_{\bar{k}} A_{\bar{i}} \otimes B^\dagger_{\bar{l}} B_{\bar{j}}]$, and $A_{\bar{i}} = A_{i_1} A_{i_2} \cdots A_{i_\ell}$ is product of ℓ operators chosen from the set of identity and projectors of measurements, $\{\mathbb{1}, M^A_{a|X}\}$, and similarly for $B_{\bar{j}}$s. Here, i, j indicates the rows while k, l indicates the columns of χ. By construction, χ can be seen as a local processing of the state ρ, hence entanglement in χ can only be less, $\mathcal{N}(\chi[\rho]) \leq \mathcal{N}(\rho)$. Lower bound on $\mathcal{N}(\chi[\rho])$ provides a lower bound on $\mathcal{N}(\rho)$.

The lower bound on $\mathcal{N}(\chi)$ can be found in a device-independent manner, by solving the following semi-definite programme:

$$\begin{aligned} \mathcal{N}(\chi[\rho]) \geq \min_{\chi,\sigma_+,\sigma_-} \quad & \mathrm{tr}[\sigma_-] \\ \text{subject to} \quad & \chi_{\bar{A}\bar{B}} = \sigma_+ - \sigma_-, \\ & \sigma_\pm^{T_A} \geq 0, \\ & I[\chi] = I_4. \end{aligned} \tag{4.21}$$

Here, $I[\chi]$ is the violation of the CGLMP inequality value of the correlations issued from the χ matrix and $\sigma_\pm$ are moment matrices of the same form as χ.

For each fixed value of I_4, one can solve for the minimum of $\mathcal{N}(\chi)$, hence lower bounding the dimension of the system that is compatible with the observation of this I_4. This lower bound gets better with increasing local level ℓ, but becomes also computationally intractable. To reduce the number of independent variables, we make use of the previously-mentioned depolarization process.

The effect of depolarization amounts to local relabelling of inputs and outcomes, which can be taken into account in the moment matrix χ by applying some suitable permutations on the row and columns. Regard the indices i, j, k, l as function of local inputs and outcomes, i.e. $i = i(i, a)$, then the matrix after relabelling is:

$$\mathcal{D}(\chi)^{kl}_{ij} = \chi^{f(k)g(l)}_{f(i)g(j)}, \tag{4.22}$$

where f and g are bijective maps from the index space to itself. Since the relabelling is local, moreover the same map is applied to the columns and rows, hence we have:

Proposition 4.2. *Let $\mathcal{D}$ be some permutation of the form Eqn. (4.22), then:*

1. $\chi \geq 0 \implies \mathcal{D}(\chi) \geq 0$,
2. $\chi^{T_A} \geq 0 \implies \mathcal{D}(\chi)^{T_A} \geq 0$.

Proof. For the first part, since simultaneous permutation of rows and columns does not change the eigenvalue of a matrix, χ positive implies $\mathcal{D}(\chi)$ positive.

For the second part, observe that

$$(\mathcal{D}(\chi)^{T_A})^{kl}_{ij} = (\mathcal{D}(\chi))^{il}_{kj} = \chi^{f(i)g(l)}_{f(k)g(j)} = (\chi^{T_A})^{f(i)g(j)}_{f(k)g(l)} = \mathcal{D}(\chi^{T_A})^{kl}_{ij}, \tag{4.23}$$

that is, the permutation $\mathcal{D}(\cdot)$ commutes with the partial transposition $(\cdot)^{T_A}$. Then by the first part, $\chi^{T_A} \geq 0 \implies \mathcal{D}(\chi^{T_A}) \geq 0 \implies \mathcal{D}(\chi)^{T_A} \geq 0$. □

CGLMP violation is invariant under the depolarization considered, that is $I[\chi] = I[\mathcal{D}(\chi)]$. Moreover the other constraints in programme 4.21 remain satisfied under the depolarization. Thus for any solution to the programme, there exist another solution that is invariant under the depolarization. One can then solve the programme with matrices that are depolarized from the start, hence reducing the number of free parameters, making the optimization tractable.

With this simplification at hand, we could compute a lower bound on the negativity given any violation of CGLMP I_4. The result is plotted in Fig. 4.2. We observe that in order to violate I_4 up to $I_N \sim 0.315$, the minimum negativity has to exceed 1, which is the maximum negativity of two qutrit states. In other words, $I_4^{(3)} \leq I_N$: a violation of I_4 more than I_N certifies the dimension of the entangled system is at least four.

4.4.2 Lower bound on the maximum qutrit violation

To see the tightness of this bound, we derive a lower bound on the maximum qutrit violation. This is done by considering a subclass of measurements on three qubit, namely those measurements where only three outcomes occur. We aim to derive a lower bound on the maximum qutrit violation by demonstrating an example of a state and measurements. Numerical evidence suggests that the true maximum qutrit violation is not larger than this.

Without lost of generality, we can assume that it is the fourth outcome that never occurs, $P(3, b|X, Y) = P(a, 3|X, Y) = 0$ for all a, b, X, Y. Due to the null

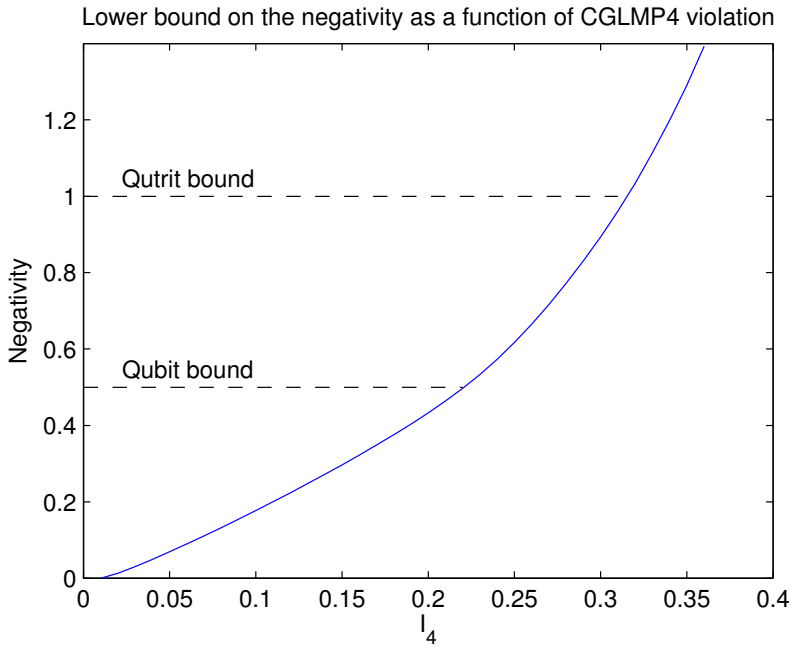


Figure 4.2: Plot of lower bound on negativity versus observed violation of the CGLMP$_4$ inequality. The lower bound is computed via solving the semi-definite programme (4.21). Maximal negativity for qubit and qutrit systems are shown. Violation above these values certifies a lower bound on the dimension of the entangled system in the experiment.

probability, the coefficients in the last rows and columns of $\mathcal{I}_4$ becomes irrelevant. We can rewrite it as:

$$\mathcal{I}_4' = \left(\begin{array}{cc|cc} J_3 & \mathbf{0} & J_3^T & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline J_3^T & \mathbf{0} & -J_3^T & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{array}\right), \tag{4.24}$$

which is identical to $\mathcal{I}_3$. Thus the maximal violation of the CGLMP$_3$ inequality, I_3^*, gives a lower bound on the maximal qutrit violation.

From Sec. 4.3.1, we have $I_3^* = \frac{\sqrt{33}-3}{9} \approx 0.30495$, and this is a lower bound on the maximal qutrit violation, $I_3^* \leq \max I_4^{(3)}$.

To summarize, so far we know that the maximal violation of CGLMP$_4$ inequality is bounded by $I_3^* \leq \max I_4^{(3)} \leq I_N$. Numerical evidence indicates that in fact $\max I_4^{(3)} = I_3^*$.

To optimize $I_4^{(3)}$ over all the possible measurement, we make use of an iterative numerical optimization procedure, called the "see-saw" method, introduced in [121] and further developed in [122]. The "see-saw" method works as follows:

1. Randomly choose measurement settings for Alice and Bob;

2. With the measurement setting of Alice and Bob, write down the Bell operator, let the state be the eigenvector with the largest eigenvalue;

3. With the state found in step 2, fix the measurement of Bob, optimize the measurement of Alice;

4. With the state and measurement of Alice from step 3, optimize the measurement of Bob;

5. Repeat step 2 to 4 until the violation do not increase any more.

By this method, one can try to find the maximal violation of CGLMP$_4$ with qutrits. Each step of this method, the violation monotonically increases. Step 2 corresponds to diagonalizing a 9-by-9 matrix. Step 3 and 4 is done be an SDP routine: optimizing a function linear in terms of the variables (violation), subjected to a positivity constraint (the POVM elements are positive). Both of these procedure guarantee convergence. Due to the possibility of attaining to a local minimum, this numerical method does not guarantee to converge globally. Nonetheless, it has given remarkable results in similar context [122]. Maximizing the CGLMP$_4$ inequality over all qutrit states and all POVM measurements, we did not find any violation exceeding the bound I_3^*.

In summary, we have:

Theorem 4.3. *The maximal violation of CGLMP$_4$ inequality with entangled qutrit,* $\max I_4^{(3)}$ *is bounded by*

$$I_3^* \leq \max I_4^{(3)} \leq I_N. \tag{4.25}$$

Any violation of CGLMP$_4$ inequality greater than I_N certifies the presence of entangled systems of dimension at least four.

Recall that the maximum violation achievable is $I_4^* \approx 0.364\,762$ with MVS and $0.336\,091$ with MES, thus they both violate the dimension witness $I_4 \leq I_N$. Consider a small amount of white noise which correspond to a violation of $-\frac{3}{4}$, MVS and MES can still violate the dimension witness with 95.54% and 98.06% visibility respectively. Hence this dimension witness could in principle be test in an experiment. However, a feature of this dimension witness makes any experimental demonstration of this superfluous.

4.5 Discussion

4.5.1 The overlooked problem

As mentioned in the beginning of this chapter, the dimension witness based on CGLMP_4 has a feature: one can encode a pair entangled ququarts with two pairs of entangled qubits; moreover, the optimal measurement factorizes, i.e. it can be performed on each qubit separately and sequentially. Let us re-examine the state and measurement used to violate the dimension witness.

A ququart state can be seen as composed of two qubits, so a pair of entangled ququart can be viewed as two pairs of entangled qubits. For example, the maximally entangled two ququart state:

$$\frac{1}{2}(|00\rangle + |11\rangle + |22\rangle + |33\rangle)_{AB} = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)_{A_1B_1} \otimes \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)_{A_2B_2}, \tag{4.26}$$

with the encoding $|00\rangle_{A_1A_2} = |0\rangle_A$, $|01\rangle_{A_1A_2} = |1\rangle_A$, $|10\rangle_{A_1A_2} = |2\rangle_A$ and $|11\rangle_{A_1A_2} = |3\rangle_A$. The dimension of the system does not seem to be a scarce resource: operationally, generating n pairs of entangled qubit is the same as generating a pair of entangled qu*d*it with $d = 2^n$.

So far we have only a formal re-writing of the states. When one claims to have a high dimensional quantum system, we would expect coherent control over the dimension as well. Not only we can apply unitaries to each qubit individually, but also unitaries to all the qubits on each party, effecting a unitary in the qudit space. Can the dimension witness based on CGLMP inequalities certify this coherent manipulation device independently? Unfortunately, the answer is no for the specific case when $d = 4$ (in fact, for $d = 2^n$ for $n = 2, 3, \cdots$). Let us go back to the optimal measurement that allow us to violate the dimension witness. From Eqn. (4.5), each vector in the measurement bases can be written as:

$$\begin{aligned} |\Psi_X(a)\rangle &= \frac{1}{2}\sum_{k=0}^{3} \bar{\omega}^k |k\rangle \\ &= \frac{1}{2}(|0\rangle + \bar{\omega}|1\rangle + \bar{\omega}^2|2\rangle + \bar{\omega}^3|3\rangle), \end{aligned} \tag{4.27}$$

where $\bar{\omega} = e^{(\frac{2a\pi}{d} + \phi_X)i}$. If we use the same encoding as how we write the maximally entangled state in Eqn. (4.26), one can see that they all factorizes. Let me spell

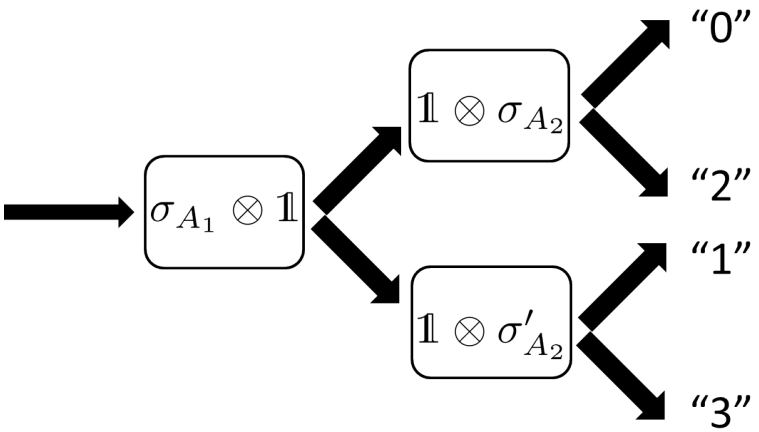


Figure 4.3: The optimal measurement can be carried in a sequential manner, with measurement outcome of the first measurement fed forward to determine the basis of the second measurement. Each measurement acts only on one qubit.

out for example Alice's measurement X:

$$
\begin{aligned}
|\Psi_X(0)\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + e^{(2\phi_x)i}|1\rangle)_{A_1} \otimes \frac{1}{\sqrt{2}}(|0\rangle + e^{(\phi_X)i}|1\rangle)_{A_2}, \\
|\Psi_X(1)\rangle &= \frac{1}{\sqrt{2}}(|0\rangle - e^{(2\phi_x)i}|1\rangle)_{A_1} \otimes \frac{1}{\sqrt{2}}(|0\rangle + e^{(\frac{\pi}{2}+\phi_X)i}|1\rangle)_{A_2}, \\
|\Psi_X(2)\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + e^{(2\phi_x)i}|1\rangle)_{A_1} \otimes \frac{1}{\sqrt{2}}(|0\rangle - e^{(\phi_X)i}|1\rangle)_{A_2}, \\
|\Psi_X(3)\rangle &= \frac{1}{\sqrt{2}}(|0\rangle - e^{(2\phi_x)i}|1\rangle)_{A_1} \otimes \frac{1}{\sqrt{2}}(|0\rangle - e^{(\frac{\pi}{2}+\phi_X)i}|1\rangle)_{A_2}.
\end{aligned}
$$

Moreover, the vectors of the first qubit form a valid measurement basis $\{|0\rangle \pm e^{(2\phi_x)i}|1\rangle\}$, while the vectors of the second form two different measurements, $\{|0\rangle \pm e^{(\phi_X)i}|1\rangle\}$ and $\{|0\rangle \pm e^{(\frac{\pi}{2}+\phi_X)i}|1\rangle\}$ separated according to the outcome of the first measurement. In light of this, the measurement on the ququart system can be done sequentially (see Fig. 4.3): first measure the first qubit with σ_{A_1}, if it outputs $+$, proceed to the second measurement σ_{A_2} to see if it is outcome 0 or 2; if it outputs $-$, proceed to a different second measurement σ'_{A_2} to see if it is outcome 1 or 3.

This is true for all the vectors in different bases, for both Alice and Bob. In other words, the measurement can be sequentially carried out on the two qubits, with only feed-forwarding of classical information. Hence, one can violate the CGLMP_4 dimension witness, with entangled qubits and qubit measurements only.

We expect this separation of the optimal measurement happens to any DW based on CGLMP for d equals powers of 2. Whenever d is power of 2, we can decompose the state into pairs of qubits, and the measurement can similarly be performed on each qubit sequentially, with the outcomes feeding forward to the

choice the settings of the measurement of the next qubits. One question remains to be answered is whether this is true for d not a power of 2, for example the dimension witness based on CGLMP_3. Now we need to bound the violation of CGLMP_3 inequality subject to pairs of qubits and three outcome measurement with the possibility feeding forward of outcomes.

A side note is that, to create the MVS instead of the MES, one might opt to choose a different encoding than the one we used here. Recall that the MVS is of the form $|\psi\rangle_{MVS} = \frac{1}{\sqrt{2}}(\cos\theta\,|00\rangle + \sin\theta\,|11\rangle + \sin\theta\,|22\rangle + \cos\theta\,|33\rangle)$, one can conveniently produce

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)_{A_1B_1} \otimes (\cos\theta\,|00\rangle + \sin\theta\,|11\rangle)_{A_2B_2}, \tag{4.28}$$

with the encoding being $|00\rangle_{A_1A_2} = |0\rangle_A$, $|01\rangle_{A_1A_2} = |1\rangle_A$, $|11\rangle_{A_1A_2} = |2\rangle_A$ and $|10\rangle_{A_1A_2} = |3\rangle_A$, and similarly for Bob. Under this encoding, the measurement do not factorize as before. One could proceed to violate the dimension witness with a maximally entangled state, however for CGLMP inequalities, the violation with MES decreases (the gap between MES and MVS widens, see [119]) as d increases . It is left to check that at which d the MES ceases to violate the dimension witness. Alternatively, one could keep the natural binary encoding but generate different states in order to maximize the violation.

4.5.2 Conclusion

To conclude, we have shown that a violation of the CGLMP_4 inequality of $I_N \approx 0.315$ or above, device independently certifies the generation of a ququart system. However, due to the form of the measurements that achieves this bound, what we can certify from this dimension witness are generation of entangled ququart pairs (equivalent to two-qubit pairs) and coherent manipulation only on qubits.

We speculate that this is a coincidence for CGLMP_d, where d is a power of 2. For the case where d is not a power of 2, the tensor structure of the optimal measurement no long exists; we expect that the optimal measurement would then require coherent manipulations on all the dimension rather than on each qubit separately.

CHAPTER 5

MANY-BOX LOCALITY

5.1 Motivation

5.1.1 Quantum physics through physical principles

Quantum mechanics is most frequently *defined* by its mathematical formulation: states are vectors in the Hilbert space and probability of an outcome occurring is given by the trace of operators. But can we define quantum mechanics from some physical principles, in the same way as special relativity is defined by the principle of constancy of speed of light in all reference frames? In the context of quantum information, the physical principles have often taken the form of a computation or communication task.

In order to define quantum theory, that is describing its limit, one must embed it in a family of more general theories. Two choices have been made. The first consists in working within the framework of *generalized probabilistic theories (GPTs)*, following the suggestion of Hardy [123]. These works, which can be seen as the revival of the axiomatic approach of Ludwig [124, 125], Piron [126] and others some decades ago, have achieved their goals [127–130]. Works in this framework rely on a separation between states and observables. Furthermore, axioms from this approach are formulated in terms of abstract entities, such as states, are not directly testable in experiments.

The second consists in working in the framework of *no-signalling (NS)* theories, following the insights of Popescu and Rohrlich [99]. We will be working in this second framework. Popescu and Rohrlich first questioned whether quantum theory could be defined as the theory that allows Bell violation without allowing the user to send a signal through the channel. This is to avoid open conflict with

special relativity. They noticed that this is not the case: there exist no-signalling boxes that violate Bell inequality stronger than what quantum physics allows (the "PR-box"). What other physical principle on top of no-signalling would recover the set of quantum distributions, Q? Five such principles has been proposed in recent years: Non-trivial Communication Complexity [131], No Advantage for Nonlocal Computation [132], Information Causality (IC) [133], Macroscopic Locality (ML) [134] and Local Orthogonality [135]. All of these criteria can rule out the PR-box but none reaches Q. It has been shown that a set of correlations, dubbed as "almost quantum" correlation [136], satisfied all the principles above except IC; numerical evidence strongly suggests that IC is satisfied by the almost quantum set too. One way to define the almost quantum set is through the SDP hierarchies: local level one in the Moroder hierarchy or Q_{1+AB} in the NPA hierarchy, and it is shown to be strictly larger than the quantum set Q [136].

Our approach: many-box locality

We would like to propose a modification of the macroscopic locality principle, called many-box locality (MBL), that might be stricter than the original ML, hence defining a possibly smaller set of correlations.

Let us first describe the principle of ML. The principle of ML states that nonlocal distributions will become consistent with classical physics in the macroscopic limit. A *microscopic* scenario is the familiar Bell scenario. At each run of the experiment, the experimentalists choose a setting and a *single* click is registered at one of the detectors of each party. One records the outcome a and repeats the experiment to estimate the statistics $P(a, b|X, Y)$. The nonlocality of the distribution can be then studied.

By *macroscopic* scenario, we refer to the case where at each run of the experiment, N pairs of pairs of particles were emitted from the source. Each party will register a total of N clicks distributed among their detectors. Instead of a particular outcome a, the intensity in each detector, I_a, is recorded. After sufficient repetition of the experiment, the probability density $P(I_a, I_b|X, Y)dI_a dI_b$ can be estimated. Furthermore, in the principle of ML, two crucial assumptions are made: (1) the intensities are measured with a fluctuation of the order of $\sqrt{N}$. This is a classical local data processing, a "smoothing" of the probability distribution, that could in principle reduce nonlocality; (2) N is taken to the limit of infinity. Consistency with classical physics in the macroscopic limit implies that when $N \gg 1$, $P(I_a, I_b|X, Y) = \int d\lambda P(\lambda)P(I_a|X, \lambda)P(I_b|Y, \lambda)$ admits a local model. When this is true, it is said that the original microscopic system exhibits macroscopic locality. The two assumptions allow the application of the central limit theorem and lead

to the result: correlations are macroscopically local if and only if they belong to the set Q_1, the first level of the NPA hierarchy.

We propose the principle of *many-box locality (MBL)* by making two modifications to the original ML principle. First, instead of N taken to be infinity, we allow N to be any natural number. We can speak about the N-local set: distributions that become local when N copies are measured. Second, we do not perform the smoothing. The flux I_a is measured with an accuracy of a single click. We say a distribution (box) satisfies MBL_N if the statistics of N such boxes measured together is local. With these two modifications, the set of correlations defined by MBL_N is possibly smaller than that defined by ML.

With this definition of MBL, one can ask the following questions: how do we characterize these N-local sets? Is there a specific N_0 such that quantum distributions are N local for all $N \geq N_0$? If not, are there quantum distributions that do not satisfy the principle of many-box locality for all N? In the limit N approach infinity, does MBL define the set of quantum distributions?

This rest of this chapter is organized as follows. In Sec. 5.2, we will first introduce some notation, the definition of many-box locality and the N-local sets; in Sec. 5.3, we will introduce the novel tool of Fourier transformation of probabilities and inequalities, which is then applied in Sec. 5.4 to study the MBL_N sets on a symmetric slice of the NS polytope; in Sec. 5.5 we turn to another slice of the NS polytope; we would then conclude the findings in the last section.

5.2 Notation and definition

In this section, we would like to introduce the concept of N-box distributions and N-locality.

Bell inequalities can be violated by measuring pairs of entangled particles with suitable measurements. Let us consider a two party (Alice and Bob), two input (A_X and B_Y, with $X, Y \in \{0, 1\}$), two outcome ($a, b \in \{0, 1\}$) Bell experiment. In each *run* of the experiment, we prepare *one* pair of entangled particles, we choose to measure A_0 for Alice, B_1 for Bob and obtain outcomes for example, $a = 0$, $b = 1$. Then in the next run, we prepare another pair of particles, choose some measurement settings, and obtain some outcomes. Multiple runs of the experiment with different inputs is repeated until we have obtained an estimation of the probability $P(a, b|X, Y)$ (see Fig. 5.1(a)). Then we compute the value of

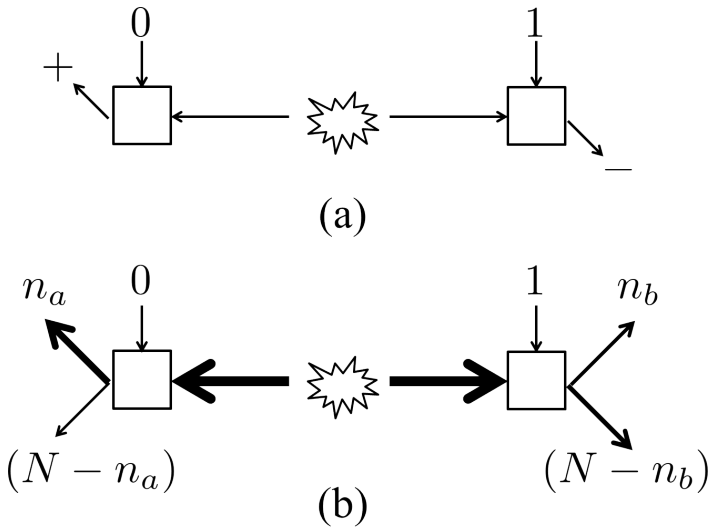


Figure 5.1: Schematic diagram for the many-box scenario. (a) A microscopic scenario: only one pair of particle is being measured in each run of the experiment. (b) An N-box scenario: N pairs of particle are being measured in each run of the experiment. The thickness of the arrows symbolizes the number of particles. The outcomes are then coarse grained to only the number of 0's (and 1's) for each party, where we denote n_a and n_b, respectively.

the Bell expression and check for violation:

$$\sum_{a,b,X,Y} c_{abxy} P(a,b|X,Y) \leq \beta, \tag{5.1}$$

where β is the local bound.

Now consider a different scenario, called an N-box scenario, where in each run of the experiment, we measure N pairs of entangled particles simultaneously, in principle we obtain outcomes, for example, $(a_1 = 0, a_2 = 1, \cdots, a_N = 1; b_1 = 1, b_2 = 1, \cdots, b_N = 0)$. Since the N pairs are emitted simultaneously, we do not keep track of the correspondence of pairs of particles; instead, we only have the coarse grained information of the number of 0's and 1's for each run. We then record the outcomes of this run as (n_a, n_b), for n_a 1's for Alice and n_b 1's for Bob. By repeating many runs of experiments and choosing different inputs, we obtain the *N-box distribution* (or *many-box distribution*) of the original microscopic distribution, $P^{*N}(a,b|X,Y)$, where $a, b \in \{0, 1, \cdots, N\}$ (see Fig. 5.1(b)). Similar scenarios has been consider in the multipair scenario [137], where collective measurements are performed on many pairs of particles. Here, in the device independent framework, we do not consider the states and measurements, only the probability distributions.

If the microscopic distribution $P(a,b|X,Y)$ is nonlocal, will the N-box distribution $P^{*N}(a,b|X,Y)$ remain nonlocal for some N? Clearly, $P^{*N}(a,b|X,Y)$ do not become nonlocal if $P(a,b|X,Y)$ is local, since we apply local post-processing of the data to obtain $P^{*N}(a,b|X,Y)$. To study the nonlocality of many-box distributions, we first need to be able to compute the many-box distribution from a microscopic one.

5.2.1 Example: 2-box distribution

To illustrate the concept of many-box distributions, let us consider the simplest example of two-box distribution.

Denote the microscopic probability distribution in the no-signalling form:

$$P(a,b|x,y) = \begin{array}{c|c||c|c} 1 & & p(0|y=0) & p(0|y=1) \\ \hline\hline p(0|x=0) & & p_{00} & p_{01} \\ \hline p(0|x=1) & & p_{10} & p_{11} \end{array} \tag{5.2}$$

where the 1 in the top left corner denotes the normalization. Then the 2-box distribution $P^{*2}(a,b|X,Y)$, with $a,b \in \{0,1,2\}$, can be computed as follows. To get an outcome $a = 0$, both boxes must output 0, the probability of this is $P^{*2}(a|X) = p(0|X)^2$. To get an outcome $a = 2$, both boxes must output 1, so similarly $P^{*2}(2|X) = p(1|X)^2$. To get an outcome, we may have 0 from the first box and 1 from the second, or 1 from the first and 0 from the second, so $P^{*2}(2|X) = p(0|X)\cdot p(1|X) + p(1|X)\cdot p(0|X) = \binom{2}{1} p(0|X)\cdot p(0|X)$. One can recognize this to be the binomial distribution. For the correlation terms, take $P^{*2}(1,1)$ as an example, we may have four possible combinations, $P^{*2}(1,1) = p(0,0)\cdot p(1,1) + p(1,1)\cdot p(0,0) + p(0,1)\cdot p(1,0) + p(1,0)\cdot p(0,1)$. One can recognize this to be the sum of multinomial distributions. Next, we shall derive the probability distribution of combining N copies of the same box.

5.2.2 Example: N-box distribution

Proposition 5.1. *The N-box distribution of a microscopic distribution $P(a,b|X,Y)$ is given by:*

$$P^{*N}(a,b) = \sum_{k=0}^{N} (a-k, b-k, k, N-a-b+k)!\; p_{00}^{N-a-b+k} \cdot p_{01}^{b-k} \cdot p_{10}^{a-k} \cdot p_{11}^{k}, \tag{5.3}$$

and the marginals:

$$P^{*N}(a) = \binom{N}{a} p_{a=1}^{a} \cdot p_{a=0}^{N-a}, \tag{5.4}$$

$$P^{*N}(b) = \binom{N}{b} p_{b=1}^{b} \cdot p_{b=0}^{N-b}, \tag{5.5}$$

where $p_{ab} = P(a,b)$, $p_{a=a'} = P(a')$, $p_{b=b'} = P(b')$ *for each* X *and* Y, *and* $(n_1, n_2, \cdots, n_k)!$ *are the multinomial coefficients.*

Proof. It follows from the binomial and multinomial distribution. For example, to obtain the outcome (a,b), we may have k boxes that output $(1,1)$, $a-k$ boxes that output $(1,0)$, $b-k$ boxes that output $(0,1)$ and the rest output $(0,0)$, the number of ways to divide a set of N elements to four sets of size $(k, a-k, b-k, N-a-b+k)$ is given by the multinomial coefficient:

$$(k, a-k, b-k, N-a-b+k)! = \frac{N!}{k!(a-k)!(b-k)!(N-a-b+k)!}, \tag{5.6}$$

when $k \leq \min(a,b)$ and $k \geq \max(0, a+b-N)$, and 0 otherwise.

Finally, we need to sum over all the possible value of k from 0 to N. □

Given microscopic distribution, the N-box distribution can be computed. We now define what we call an N-box local distribution:

Definition 5.2. *A distribution* P *is said to be* N*-box local* ($P \in MBL_N$), *if* P^{*N} *is local* ($P^{*N} \in L$).

The locality of many-box distributions can be tested numerically via the linear programme described in Chapter 3. However, to analytically characterize of the many-box local sets, especially to investigate the limiting behaviour when N tends to infinity, another method to compute .

5.3 New tool: Fourier transformation

In this section, we introduce a novel tool to compute and characterize the N-box distributions, namely Fourier transformation on probability distributions. Note that a N-box distribution $P^{*N}(a,b|X,Y)$ has a, b range from 0 to N, we denote the number of outcomes as $d = N+1$.

Now consider combining two boxes, not necessary identical. Focusing on a specific input of Alice, denote the two box as P_1 and P_2. Then the combined box,

$P_1 * P_2$ outputs a with the probability of:

$$P_1 * P_2(a) = \sum_{a'} P_1(a') * P_2(a - a'), \tag{5.7}$$

which is in fact a discrete version of convolution. Adding one box at a time, by an inductive argument, one can see that the N-box distribution P^{*N} is P convoluted with itself N times:

$$P^{*N} = \underbrace{P * P * \cdots * P}_{N}. \tag{5.8}$$

The advantage of treating the N-box distribution as a convolution is the *convolution theorem*: The Fourier transformation of convolution of two functions is the product of the two functions Fourier transformed, formally,

$$\mathcal{F}[f * g] = \mathcal{F}[f] \cdot \mathcal{F}[g]. \tag{5.9}$$

Here, the (discrete) Fourier transformation on probabilities are applied for each input X and Y, and independently for Alice and Bob, as follows:

Definition 5.3. *For a joint probability distribution $P(a, b)$, we define its Fourier transform, $\tilde{P}(k, l)$ as*

$$\tilde{P}(k, l) = \mathcal{F}_d[P(a, b)](k, l) = \sum_{a,b=0}^{d-1} e^{\frac{2\pi i}{d}(ak+bl)} P(a, b), \tag{5.10}$$

where d must be larger than or equal to the number of outcomes of each party in $P(a, b)$.

$\tilde{P}(k, l)$ can be complex and its physical interpretation is not immediately clear. Nonetheless, it is related to the probability via a linear transformation. By applying the inverse transformation to the Bell coefficient, we obtained a valid "Bell inequality" for the Fourier transformed probabilities. Mathematically,

Definition 5.4. *For a Bell inequality c_{abXY}, its inverse Fourier transform, $\tilde{c}_{klXY}$ is defined as*

$$\tilde{c}_{klXY} = \frac{1}{d^2} \sum_{a,b=0}^{d-1} e^{-\frac{2\pi i}{d}(ak+kl)} c_{abXY}, \tag{5.11}$$

where d must be larger than or equal to the number of a and b in c_{abXY}, and $\frac{1}{d^2}$ is for normalization.

With this definition, one can verify that

Proposition 5.5.

$$\tilde{c}\cdot\tilde{P}\equiv\sum_{k,l,X,Y}\tilde{c}_{klXY}\tilde{P}(k,l|X,Y)=\sum_{a,b,X,Y}c_{abXY}P(a,b|X,Y)\equiv c\cdot P, \tag{5.12}$$

and if $c\cdot P\leq\beta$ is a valid Bell inequality, $\tilde{c}\cdot\tilde{P}\leq\beta$ implies that both $P(a,b|X,Y)$ and $\tilde{P}(k,l|X,Y)$ has a local decomposition.

Proof. Let us denote $\omega=e^{\frac{2\pi i}{d}}$ as the dth root of unity. For each X and Y, from the definition of $\tilde{c}$ and $\tilde{P}$:

$$\begin{aligned}\sum_{k,l}\tilde{c}_{kl}\tilde{P}(k,l)=&\sum_{k,l}\left(\frac{1}{N^2}\sum_{a,b=0}^{N-1}\omega^{-(ak+kl)}c_{ab}\right)\left(\sum_{a',b'=0}^{N-1}\omega^{(a'k+b'l)}P(a',b')\right)\\=&\sum_{a,b,a',b'}c_{ab}P(a',b')\\=&\sum_{a,b}c_{ab}P(a,b)\end{aligned}$$

where we use the fact that $\sum_k\omega^{(a'-a)k}=N\delta_{a,a'}$.

Summing over all the X and Y, we have $c\cdot P=\tilde{c}\cdot\tilde{P}$. Clearly $P(a,b|X,Y)$ admits a local decomposition when $\tilde{c}\cdot\tilde{P}\leq\beta$, since $c\cdot P=\tilde{c}\cdot\tilde{P}\leq\beta$. □

Now let us state our version of the convolution theorem:

Theorem 5.6. *Let P_1 and P_2 be two probability distribution with d_1 and d_2 outcomes respectively, let $d\geq d_1+d_2-1$, then*

$$\mathcal{F}_d[P_1*P_2]=\mathcal{F}_d[P_1]\cdot\mathcal{F}_d[P_2]. \tag{5.13}$$

Proof.

$$\begin{aligned}\mathcal{F}_d[P_1*P_2](k,l)=&\sum_{a,b=0}^{d-1}\omega^{(ak+bl)}P_1*P_2(a,b)\\=&\sum_{a,b=0}^{d-1}\omega^{(ak+bl)}\sum_{a',b'=0}^{d-1}P_1(a',b')P_2(a-a',b-b')\\=&\sum_{a,b,a',b'=0}^{d-1}\omega^{(a'k+b'l)}\,P_1(a',b')\,\omega^{(a-a')k+(b-b')l}\,P_2(a-a',b-b')\\=&\sum_{a',b'=0}^{d-1}\omega^{(a'k+b'l)}\,P_1(a',b')\sum_{a'',b''=0}^{d-1}\omega^{(a''k+b''l)}\,P_2(a'',b'')\\=&\mathcal{F}_d[P_1]\cdot\mathcal{F}_d[P_2],\end{aligned}$$

where at the fourth line we made a change of variable $(a-a') \to a'' \in \{0, \cdots, d-1\}$ and similarly for $(b-b')$. □

By inductively applying the convolution theorem, we can express the N-box distribution in terms of the Fourier transformed microscopic distribution:

Corollary 5.7. *The Fourier transformation of the N-box distribution is the Fourier transformation of microscopic distribution raised to the power of N:*

$$\widetilde{P^{*N}} = \mathcal{F}_d[P^{*N}](k,l|X,Y) = (\mathcal{F}_d[P](k,l|X,Y))^N = \tilde{P}^N. \tag{5.14}$$

This provides us with another way to describe the N-box local sets. The N-box distribution lies in the d outcome probability space. Let us denote the Bell inequalities in this space by their coefficients c_i and their corresponding bound β_i. Then

Proposition 5.8. *If there exist an inequality i such that, $\tilde{c}_i \cdot \tilde{P}^N \geq \beta_i$, then $P \notin MBL_N$.*

This allows us to exclude points from being in the MBL_N sets. To certify a point is within MBL_N, the point has to satisfy all the inequalities,

Proposition 5.9.

$$P \in MBL_N \iff \tilde{c}_i \cdot \tilde{P}^N \leq \beta_i \quad \forall i. \tag{5.15}$$

In general, the set of all the Bell inequalities in the d outcome space is not known for large d. The number of Bell inequalities grows at least exponentially in d. In the next section, we will restrict ourselves to a slice of the NS polytope. On that slice, only one class of Bell inequalities is relevant for determining MBL_N based on numerical evidence.

5.4 On the symmetric slice

To study these sets of N-box local distribution, we consider the simplest scenario of two-party, two-input and two-outcome experiment. Moreover, we are going to focus on two two-dimensional slices of the no-signalling polytope. In this section, we first consider the slice that passes through two PR-boxes and the completed mixed distribution (see Fig. 5.2):

$$p(x,y) := x\,(P_{PR1} - P_{mix}) + y\,(P_{PR2} - P_{mix}) + P_{mix}, \tag{5.16}$$

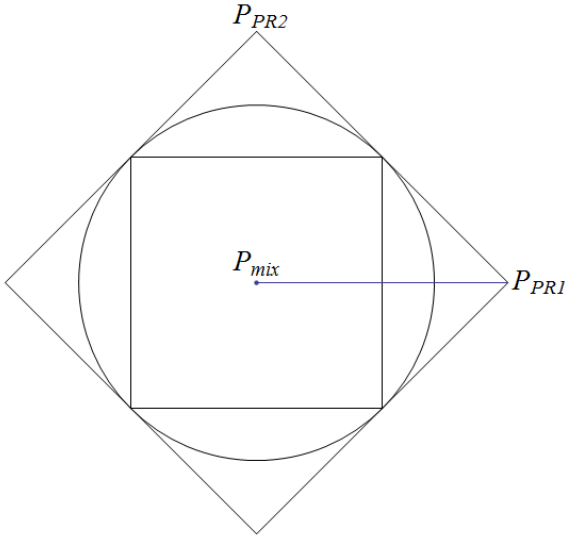


Figure 5.2: A geometric representation of the symmetric slice. Any point (x, y) represents a distribution defined via Eqn. (5.16). The no-signalling polytope is represented by the outer square with PR-boxes as its vertices. The local polytope is represented by the inner square, with its edge being CHSH inequalities. The quantum set is represented by the circle between the NS polytope and the local polytope.

where

$$P_{PR1} = \left(\begin{array}{cc||c|c} 1 & \frac{1}{2} & \frac{1}{2} \\ \hline\hline \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \hline \frac{1}{2} & \frac{1}{2} & 0 \end{array}\right), \quad P_{PR2} = \left(\begin{array}{cc||c|c} 1 & \frac{1}{2} & \frac{1}{2} \\ \hline\hline \frac{1}{2} & 0 & \frac{1}{2} \\ \hline \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{array}\right), \quad P_{mix} = \left(\begin{array}{cc||c|c} 1 & \frac{1}{2} & \frac{1}{2} \\ \hline\hline \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \hline \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \end{array}\right), \tag{5.17}$$

and $x, y \in [0, 1]$, or in the full correlation form:

$$p(x, y) = \frac{1}{4}\left(\begin{array}{cc|cc} 1+x-y & 1-x+y & 1+x+y & 1-x-y \\ 1-x+y & 1+x-y & 1-x-y & 1+x+y \\ \hline 1+x+y & 1-x-y & 1-x+y & 1+x-y \\ 1-x-y & 1+x+y & 1+x-y & 1-x+y \end{array}\right). \tag{5.18}$$

Distributions on this slice have all the symmetries of the PR boxes, with the following generators: exchanging the parties, (A, B); simultaneous permutation of the inputs and outputs $X_1(0, 1)Y(0, 1)$ or $Y_1(0, 1)X(0, 1)$; and output permutations $A_0(0, 1)A_1(0, 1)B_0(0, 1)B_1(0, 1)$. Any distribution in the no-signalling polytope can be depolarized onto this slice by equally mixing the resultant distribution of each permutation. Note that the marginals are always maximally mixed on this slice, hence there are no local deterministic point on this slice, only mixtures of them.

5.4.1 Numerical analysis of the many-box local set

For any point on the slice, we can compute its N-box distribution either via the multinomial distribution (see Eqn. (5.3)) or the Fourier transformation and raising the power (see Eqn. (5.14)). The locality of P^{*N} is determined by the linear programme in Sec. 3. Notice that we are now trying to decompose P^{*N} into a convex combination of local deterministic points in the d outcome space, instead of the 2 outcome one. We can regard the output of the linear programme, $v_{\max}$ as a function of the parametrisation of the points on the slice, and let us define $h_N(x, y) := v_{\max}(x, y) - 1$.

Along the isotropic line

As a first step into the investigation of the MBL_N local sets, let us look at how does the nonlocality degrade when we increase N for point along the isotropic line $(x, 0)$, for $x \in \{0, 1\}$. Points on the isotropic line represent distribution obtain from a PR box mixing with different amount of white noise.

There are a few special points along this line: for $x \leq \frac{1}{2}$, $(x, 0)$ belongs to the local set; $(\frac{1}{\sqrt{2}}, 0)$ is the point that leads to maximal CHSH violation, in other words the most non-local quantum point; and $(1, 0)$ is the PR-box.

In Fig. 5.3, we plot $h_N(x, 0)$ against the number of copies N. Notice that $h \geq 0$ means a distribution is local. For a single copy, $x = 0.3$ and $x = 0.5$ correspond to some local distribution, while the others are nonlocal. As we increase the number of copies, local points remain local; some nonlocal points (i.e. $x = 0.65$) becomes local when we combine six or more copies. The other three remain nonlocal, with the point $x = \frac{1}{\sqrt{2}}$ approaching local quickly. This is the first numerical evidence that the criterion of many-box locality may coincide with that of the quantum correlations.

Boundary of the MBL_N set

The points on the curve $h_N(x, y) = 0$ define the boundary of the MBL_N set on this slice. In Fig. 5.4, we show the boundaries MBL_N, by numerically solving $h_N(x, y) = 0$, for N from 2 to 10.

A feature worth noticing is the oscillatory behaviour with even and odd number of copies as shown in Fig. 5.3. In Fig. 5.4, on this slice, MBL_3 seemed to be contained inside MBL_2. For some distributions, odd number of copies display more nonlocality than even. This is possibly due to the way the course graining is applied. For even number of copies, say $N = 4$, the outcome in the centre ($\frac{N}{2}$ 0's and $\frac{N}{2}$ 1's) carries the largest probability. This implies a large overlap with the

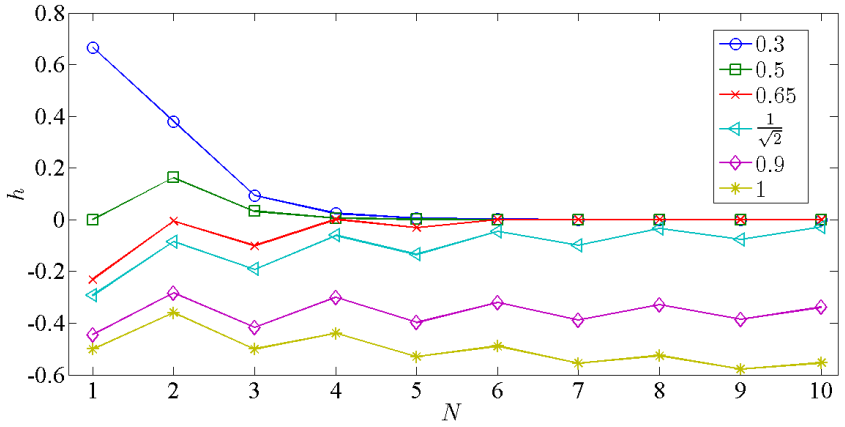


Figure 5.3: The plot the locality function h against N for different value of x. $h \leq 0$ means a distribution is local. Different curves represent a PR box mixing with different amount of white noise. Local boxes remain local as the number of copies N increases. Some non-local boxes become local after some N, some approaches local, while the others remain non-local.

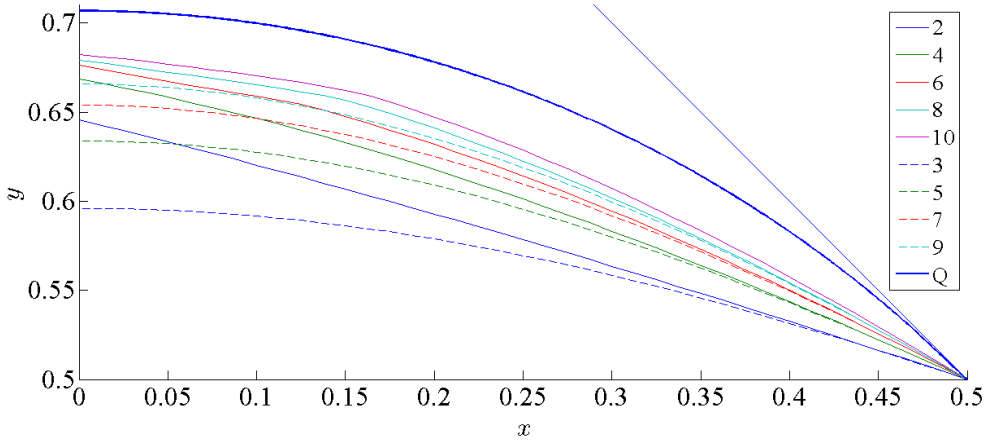


Figure 5.4: The boundary of the MBL_N sets for N from 2 to 10, in the slice parametrise according to (5.16), only the upper right sector is shown. Solid lines are used for the even N and dashed for odd N. Also shown is the boundary of the quantum set (thick solid curve) and the no-signalling set (outermost straight line).

local deterministic point which always output $a = \frac{N}{2}$ and $b = \frac{N}{2}$ for all the inputs, hence diminishing its nonlocality.

Once a point is N-box local, then it is also M-box local for any M that is a multiple of N. Moreover, if a box is both N_1-box local and N_2-box local, then it is (N_1+N_2)-box local. In particular, if a box $P \in MBL_2 \cap MBL_3$, then $P \in MBL_N$ for all $N \geq 2$.

We expect this even-odd separation will diminish as N tends to infinity. The effect of convoluting N-box distribution with another 1-box distribution, is not significant, the nonlocality of $(N+1)$-box distribution is close to that of the N-box distribution.

The MBL_N sets are not convex *a priori*. In fact, on another slice, they are not convex as we will see Sec. 5.5.

We would like to investigate how these sets evolve when number of copies N increasing and approach infinity. With the mathematical tools we introduce in Sec. 5.3, we will see that on this slice, the limit of MBL_N when $N \to \infty$, MBL_∞, coincide with the quantum set.

5.4.2 Analytical characterisation of many-box local set

Recall we can compute the many box distribution via the convolution theorem 5.7, on this symmetric slice, we have

$$\mathcal{F}[p(x,y)^{*N}] = \left(\mathcal{F}[p(x,y)]\right)^N . \tag{5.19}$$

All Bell inequalities $c_i \cdot p^{*N} \leq \beta_i$ must be satisfied for $p \in MBL_N$. Each Bell inequality defines a region in the slice, and the union of all of them is the (N)-box local set. Mathematically,

Proposition 5.10. *Let* $R_i^N := \left\{(x,y) | \tilde{c}_i \cdot \tilde{p}(x,y)^N \leq \beta_i\right\}$, *then* $MBL_N = \bigcap_{i\in I} R_i^N$, *where* I *is the set of all the* $d = N+1$ *outcome Bell inequalities.*

The boundary of MBL_N *are points that saturate one or more of the Bell inequalities,*

$$\tilde{c}_i \cdot \tilde{p}(x,y)^{N-1} = \beta_i. \tag{5.20}$$

Since Fourier transformation is linear, Eqn. (5.20) are nothing but polynomials in x and y. Each inequality (a hyperplane) in the d outcome space transforms into a hypersurface in the space of two outcome distributions, by the Fourier transform and raising to the power of N.

Liftings of CHSH

We now pick a special class of Bell inequalities in the d outcome space, namely liftings of CHSH inequality. A Bell inequality defined for a specific Bell scenario can be extended to a situation involving more parties, more inputs or more outcomes [138], becoming *liftings* of the original inequality. Here we are going to study the liftings of CHSH to more outcomes. Recall that the coefficient table for the CHSH inequality in Collins-Gisin form (also known as the Clauser-Horne inequality) is:

$$c_{CH} = \begin{array}{c|c|c} & \text{-1} & 0 \\ \hline\hline \text{-1} & 1 & 1 \\ \hline 0 & 1 & \text{-1} \end{array} \leq 0, \tag{5.21}$$

or equivalently in the full correlation form:

$$c_{CHSH} = \left(\begin{array}{cc|cc} 1 & \text{-1} & 1 & \text{-1} \\ \text{-1} & 1 & \text{-1} & 1 \\ \hline 1 & \text{-1} & \text{-1} & 1 \\ \text{-1} & 1 & 1 & \text{-1} \end{array}\right) \leq 2. \tag{5.22}$$

This original inequality is catered for two outcome experiments. For experiments with more than two outcomes, we may group several outcomes into one. In other words, the same Bell coefficient is assign to several probabilities. An example of CHSH lifted to three outcomes can be:

$$c_{lift1} = \left(\begin{array}{ccc|ccc} 1 & \text{-1} & \text{-1} & 1 & \text{-1} & \text{-1} \\ \text{-1} & 1 & 1 & \text{-1} & 1 & 1 \\ \text{-1} & 1 & 1 & \text{-1} & 1 & 1 \\ \hline 1 & \text{-1} & \text{-1} & \text{-1} & 1 & 1 \\ \text{-1} & 1 & 1 & 1 & \text{-1} & \text{-1} \\ \text{-1} & 1 & 1 & 1 & \text{-1} & \text{-1} \end{array}\right) \leq 2. \tag{5.23}$$

Here we group the second and third outcome of each measurement to one outcome, $\{0\} \to \bar{0}$ and $\{1,2\} \to \bar{1}$. Note that this grouping of outcomes may depend on

each party's input, but not the input of the other party. For example,

$$c_{lift2} = \left(\begin{array}{cccc|cc} 1 & \text{-1} & \text{-1} & 1 & \text{-1} & \text{-1} \\ \text{-1} & 1 & 1 & \text{-1} & 1 & 1 \\ \text{-1} & 1 & 1 & \text{-1} & 1 & 1 \\ \hline 1 & \text{-1} & \text{-1} & \text{-1} & 1 & 1 \\ 1 & \text{-1} & \text{-1} & \text{-1} & 1 & 1 \\ \text{-1} & 1 & 1 & 1 & \text{-1} & \text{-1} \end{array}\right) \leq 2. \tag{5.24}$$

Here, Alice set $\{0\} \to \bar{0}$ and $\{1,2\} \to \bar{1}$ when $X = 0$, while $\{0,1\} \to \bar{0}$ and $\{2\} \to \bar{1}$ when $X = 1$. Let us consider all the liftings including relabelling of inputs and outcomes, then the number of liftings of CHSH to d outcomes is $4 \cdot (2^d - 2)^4$, though not all of them are inequivalent. These are only the tip of the iceberg of all the Bell inequalities in the d outcome space. There are many more that are inequivalent to liftings of CHSH inequalities, for example the (liftings of) CGLMP inequalities.

Let us consider the following lifting of CHSH to d outcomes, where d is even: we group the first half of the outcome to 0, and the rest to 1, i.e. $\left\{0, 1, \cdots, \frac{d}{2} - 1\right\} \to \bar{0}$ and $\left\{\frac{d}{2}, \frac{d}{2} + 1, \cdots, d - 1\right\} \to \bar{1}$:

$$c = \left(\begin{array}{cc|cc} \mathbf{\text{-1}} & \mathbf{1} & \mathbf{1} & \mathbf{\text{-1}} \\ \mathbf{1} & \mathbf{\text{-1}} & \mathbf{\text{-1}} & \mathbf{1} \\ \hline \mathbf{1} & \mathbf{\text{-1}} & \mathbf{1} & \mathbf{\text{-1}} \\ \mathbf{\text{-1}} & \mathbf{1} & \mathbf{\text{-1}} & \mathbf{1} \end{array}\right) \leq 2, \tag{5.25}$$

where the boldface $\mathbf{1}$ represents an $d/2 \times d/2$ matrix filled with 1's. Notice that the block corresponding to ($X = 0$, $Y = 1$), ($X = 1$, $Y = 0$) and ($X = 1$, $Y = 1$) are identical, while the block ($X = 0$, $Y = 0$) has the opposite sign.

Examples of $N = 3$ and $N = 2$

We compute the curve corresponding to this family of Bell inequalities for small N according to Eqn. (5.20). Compare to the boundary of MBL_N numerically found by solving $h(x, y) = 0$, we found that for N equal to 3 and 5, the curve computed from this lifting matches exactly the boundary of the MBL_N set.

Case $N = 3$ When we combine $N = 3$ copies of the same box, we have $d = 4$ outcomes, and $\omega = e^{\frac{2\pi i}{4}} = i$.

Let us inverse Fourier transform the first block of the Bell coefficients, $(X = 0, Y = 0)$:

$$\begin{aligned}\tilde{c}_{kl00} &= \frac{1}{4^2}\sum_{a,b=0}^{3} c_{ab00}\omega^{-(ak+bl)} \\ &= -\frac{1}{4^2}\left[\left(\sum_{a=0,1}\omega^{-ak} - \sum_{a=2,3}\omega^{-ak}\right)\left(\sum_{b=0,1}\omega^{-bl} - \sum_{b=2,3}\omega^{-bl}\right)\right] \\ &= -\frac{1}{4^2}\left[(1+i^{-k})(1+i^{-l})(1-(-1)^k)(1-(-1)^l)\right].\end{aligned}$$

The other three blocks are the same but multiplied by -1. So in the table form:

$$\tilde{c}_{klXY} = \frac{1}{2}\left(\begin{array}{cccc|cccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & i & 0 & -1 & 0 & -i & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & -i & 0 & 1 & 0 & i \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -i & 0 & 1 & 0 & -i & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & i & 0 & 1 & 0 & i \end{array}\right). \tag{5.26}$$

For the probability (5.18), we apply the Fourier transformation. As an example, for $(X = 0, Y = 0)$:

$$\tilde{p}(k,l|0,0) = \frac{1}{4}\left[(1+x-y) + (1-x+y)i^k + (1-x+y)i^l + (1+x-y)i^{(k+l)}\right]. \tag{5.27}$$

Similarly for the other three blocks, so in table form:

$$\tilde{p}(x,y) = \frac{1}{2}\left(\begin{array}{c|c} m(x,y) & m(x,-y) \\ \hline m(x,-y) & m(-x,-y) \end{array}\right), \tag{5.28}$$

where

$$m(x,y) = \begin{pmatrix} 2 & 1+i & 0 & 1-i \\ 1+i & i(1-x+y) & (1-i)(x-y) & 1+x-y \\ 0 & (1-i)(x-y) & 2(x-y) & (1+i)(x-y) \\ (1-i) & 1+x-y & (1+i)(x-y) & -i(1-x+y) \end{pmatrix}. \tag{5.29}$$

We raise $\tilde{p}$ to the third power, by which we mean raising each element in $m(x, y)$ to the third power:

$$m^3(x,y) = \begin{pmatrix} 8 & -2+2i & 0 & -2-2i \\ -2+2i & -i(1-x+y)^3 & (-2-2i)(x-y)^3 & (1+x-y)^3 \\ 0 & (-2-2i)(x-y)^3 & 8(x-y)^3 & (-2+2i)(x-y)^3 \\ -2-2i & (1+x-y)^3 & (-2+2i)(x-y)^3 & i(1-x+y)^3 \end{pmatrix}. \tag{5.30}$$

So

$$\tilde{p}^3(x,y) = \frac{1}{8}\left(\begin{array}{c|c} m^3(x,y) & m^3(x,-y) \\ \hline m^3(x,-y) & m^3(-x,-y) \end{array}\right). \tag{5.31}$$

Following from Eqn. (5.20), we have

$$\tilde{c} \cdot \tilde{p}^3(x,y) = 2, \tag{5.32}$$

evaluating the inner product simplifies to

$$y(3 + 3x^2 + y^2) = 2. \tag{5.33}$$

This concludes the case for $N = 3$.

Case $N = 2$ When $N = 2$, we have $d = 3$ outcomes, so the lifting of the form (5.25) could not apply, since d is odd. Instead, curves from two liftings were required to recover the numerically computed boundary of the MBL_2 sets. The two liftings are:

$$c_1 = \left(\begin{array}{ccc|ccc} -1 & 1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 & 1 & 1 \\ \hline 1 & -1 & -1 & 1 & -1 & -1 \\ -1 & 1 & 1 & -1 & 1 & 1 \\ -1 & 1 & 1 & -1 & 1 & 1 \end{array}\right), \; c_2 = \left(\begin{array}{ccc|ccc} -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 & 1 & 1 \\ \hline 1 & 1 & -1 & 1 & -1 & -1 \\ -1 & -1 & 1 & -1 & 1 & 1 \\ -1 & -1 & 1 & -1 & 1 & 1 \end{array}\right), \tag{5.34}$$

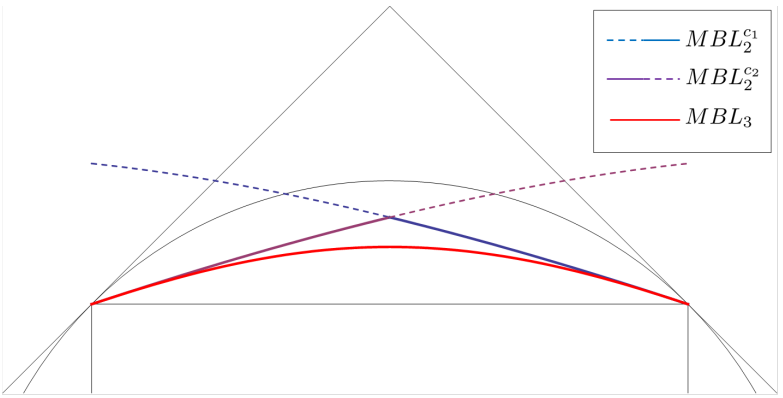


Figure 5.5: Analytic curve for the boundary of MBL_2 and MBL_3 on the symmetric slice (only the upper sector is shown). The NS polytope, local polytope and the quantum set are shown as before. The analytical curves corresponding to Eqn. (5.35) (blue), Eqn. (5.36) (purple) and Eqn. (5.33) (red) are shown. The boundary of MBL_2 (thick blue and purple segment) consists of two curves originated from two liftings of CHSH, while the boundary of MBL_3 is given by the lifting described in Eqn. (5.25).

resulting in the polynomials:

$$\frac{1}{2}(1+4y+(x+y)^2)=2 \tag{5.35}$$

$$\frac{1}{2}(1+4y+(x-y)^2)=2 \tag{5.36}$$

In this case the boundary of the MBL_2 set is the concatenation of these two curves, as shown in Fig. 5.5.

5.4.3 Solution for any number of copies

Numerical evidence showed that for odd N, the only relevant Bell inequality is the lifting of the form (5.25). We will limit ourselves to only odd number of N, so that we have an even d number of outcomes and the lifting of the form (5.25) can apply. This is not a loss of generality for the reason we have mentioned earlier: when N is large, the difference between the $(N+1)$-box distribution and N-box distribution diminishes. Now we shall derive the analytic solution of the boundary of MBL_N for all odd N.

First let us define

$$\begin{aligned} f(\eta) &= \sum_{k,l} \tilde{p}_{k,l}^{d-1}\ \tilde{c}_{k,l} \\ &= \sum_{k,l} \left\{ \frac{1}{4} \left[(1+\eta) + (1-\eta)\omega^k + (1-\eta)\omega^l + (1+\eta)\omega^{k+l}) \right] \right\}^{d-1} \times \\ &\qquad \frac{(1-\omega^{-\frac{dk}{2}})^2(1-\omega^{-\frac{dl}{2}})^2}{d^2(1-\omega^{-k})(1-\omega^{-l})}, \end{aligned}$$

where as before $\omega = e^{\frac{2\pi i 1}{d}}$, $d = N+1$, and k, l runs from 0 to $d-1$. $\tilde{p}_{k,l}$ and $\tilde{c}_{k,l}$ refers to the Fourier transformed probability and the inverse Fourier transformed Bell coefficient for a certain input combination. η may take value $u \equiv x - y$ or $v \equiv x + y$ depending on the input X and Y. On the other hand, $\tilde{c}_{k,l}$ can be computed from

$$\tilde{c}_{k,l} = \frac{1}{d^2} \sum_{ab} c_{ab} \omega^{-(ak+bl)}. \tag{5.37}$$

Take for example the top right block of (5.25), where $c_{ab} = 1$ when $a \in \{0, \cdots, d/2-1\}$ and $b \in \{0, \cdots, d/2-1\}$ or $a \in \{d/2, \cdots, d-1\}$ and $b \in \{d/2, \cdots, d-1\}$, and -1 otherwise. The sum then becomes

$$\begin{aligned} \tilde{c}_{k,l} &= \sum_{a=0}^{d/2-1} \sum_{b=0}^{d/2-1} \omega^{-(ak+bl)} + \sum_{a=d/2}^{d-1} \sum_{b=d/2}^{d-1} \omega^{-(ak+bl)} - \sum_{a=0}^{d/2-1} \sum_{b=d/2}^{d-1} \omega^{-(ak+bl)} - \sum_{a=d/2}^{d-1} \sum_{b=d/2}^{d-1} \omega^{-(ak+bl)} \\ &= \left(\sum_{a=0}^{d/2-1} \omega^{-ak} - \sum_{a=d/2}^{d-1} \omega^{-ak} \right) \left(\sum_{b=0}^{d/2-1} \omega^{-bl} - \sum_{b=d/2}^{d-1} \omega^{-bl} \right) \\ &= (1-\omega^{-\frac{dk}{2}})(1-\omega^{-\frac{dl}{2}}) \left(\sum_{a=0}^{d/2-1} \omega^{-ak} \right) \left(\sum_{b=0}^{d/2-1} \omega^{-bl} \right) \\ &= \frac{(1-\omega^{-\frac{dk}{2}})^2(1-\omega^{-\frac{dl}{2}})^2}{d^2(1-\omega^{-k})(1-\omega^{-l})}. \end{aligned}$$

It is not difficult to see that the violation of the lifted CHSH inequality of a point (x, y) is:

$$-f(u) + f(v) + f(v) + f(-u).$$

Notice the fact that f is odd in its argument (because $d-1$ is odd), the violation is in fact $2[f(u) - f(v)]$. So,

Proposition 5.11. *On this symmetric slice of the two-party two-input two-outcome no-signalling polytope, the boundary of the MBL_N sets are given by*

$$2f(u) - 2f(v) = 2.$$

Writing f as a polynomial

We aim to simplify $f(\eta)$ by carrying out the sum over k and l, hence expressing the violation in terms of polynomials in terms of η (eventually in terms of x and y).

First we simplify the term of $\tilde{p}_{k,l}$. Group the terms with η

$$\begin{aligned}
\tilde{p}_{k,l}^{d-1} &= \left\{\frac{1}{4}[(1+\omega^k+\omega^l\omega^{k+l}) + (1-\omega^k-\omega^l+\omega^{k+l})\eta]\right\}^{d-1} \\
&= \left\{\frac{1}{4}[(1+\omega^k)(1+\omega^l) + (1-\omega^k)(1-\omega^l)\eta]\right\}^{d-1} \\
&= \frac{1}{4^{d-1}}\sum_{m=0}^{d-1}\binom{d-1}{m}\eta^m(1-\omega^k)^m(1-\omega^l)^m(1+\omega^k)^{d-m-1}(1+\omega^l)^{d-m-1},
\end{aligned}$$

where we used the binomial expansion in the third line. k and l can be separated, this will simplify the sum over k and l later.

Now let us move to the $\tilde{c}_{k,l}$ term. Note that

$$(1-\omega^{-\frac{dk}{2}}) = \begin{cases} 0 & \text{if } k \text{ is even,} \\ 2 & \text{if } k \text{ is odd,} \end{cases}$$

and

$$\frac{1}{1-\omega^{-k}} = -\frac{\omega}{1-\omega^k}.$$

Combining these with the $\tilde{p}_{k,l}$ term, we have

$$\begin{aligned}
f(\eta) = \frac{16}{d^2 4^{d-1}} \sum_{k,l=\text{odd}} \sum_{m=0}^{d-1} \binom{d-1}{m}\eta^m \times \\
(1-\omega^k)^m(1-\omega^l)^m(1+\omega^k)^{d-m-1}(1+\omega^l)^{d-m-1}\frac{\omega^k}{1-\omega^k}\frac{\omega^l}{1-\omega^l}. \quad (5.38)
\end{aligned}$$

Now exchange the order of the sums and separate k from l, realizing that the sum over k is identical to the sum over l, we have

$$f(\eta) = \frac{16}{d^2 4^{d-1}} \sum_{m=0}^{d-1} \binom{d-1}{m} \eta^m \underbrace{\left(\sum_{k=\text{odd}}^{d-1} (1-\omega^k)^m (1+\omega^k)^{d-m-1} \frac{\omega^k}{1-\omega^k} \right)^2}_{S^2(d,m)}. \quad (5.39)$$

This is a polynomial in η, with the coefficients involving $S^2(d,m)$ to be further simplified.

Simplifying $S(d,m)$

We digress a little to introduce a small mathematical trick that would be useful:

Proposition 5.12. *For an even integer d, and integer $1 \leq r \leq d$, we have*

1.
$$\sum_{l=\text{odd}}^{d-1} (\omega^l)^r = \frac{d}{2}(\delta_{r,0} - \delta_{r,\frac{d}{2}}) \quad (5.40)$$

where $\delta_{r,s}$ is the Kronecker delta function, such that $\delta(r,s) = 1$ if $r = s$ and 0 otherwise.

2.
$$g(r) = \sum_{l=\text{odd}}^{d-1} \frac{(\omega^l)^r}{1-\omega^l} = \begin{cases} -\frac{d}{4}, & 1 \leq r \leq \frac{d}{2}, \\ \frac{d}{4}, & \frac{d}{2}+1 \leq r \leq d. \end{cases} \quad (5.41)$$

Proof. If $r \not\equiv 0 \bmod \frac{d}{2}$, then

$$\sum_{l=\text{odd}}^{d-1} (\omega^l)^r = \frac{1-(\omega^{2r})^{\frac{d}{2}}}{1-\omega^{2r}} = \frac{1-1}{1-\omega^{2r}} = 0.$$

When $r \equiv 0 \bmod d$, $\sum_{l=\text{odd}}^{d-1} (\omega^l)^r = \sum_{l=\text{odd}}^{d-1} (+1) = \frac{d}{2}$ and when $r \equiv \frac{d}{2} \bmod d$, $\sum_{l=\text{odd}}^{d-1} (\omega^l)^r = \sum_{l=\text{odd}}^{d-1} (-1) = -\frac{d}{2}$. This proves the part 1.

For the part 2., by factoring out one $(w^l)^{r-1}$,

$$\begin{aligned} g(r) &= \sum_{l=\text{odd}}^{d-1} (\omega^l)^{r-1} \frac{w^l}{1-\omega^l} = \sum_{l=\text{odd}}^{d-1} (\omega^l)^{r-1} \left(\frac{1}{1-\omega^l} - 1 \right) \\ &= g(r-1) - \sum_{l=\text{odd}}^{d-1} (\omega^l)^{r-1}. \end{aligned}$$

Then by induction, we have

$$g(r) = g(0) - \sum_{s=1}^{r} \sum_{l=\text{odd}}^{d-1} (\omega^l)^{r-s}.$$

The base case $g(0)$ can be evaluated as

$$\begin{aligned} g(0) = \sum_{l=\text{odd}}^{d-1} \frac{1}{1-\omega^l} &= \frac{1}{2} \sum_{l=\text{odd}}^{d-1} \frac{1}{1-\omega^l} + \frac{1}{1-\omega^{-l}} \\ &= \frac{1}{2} \sum_{l=\text{odd}}^{d-1} \frac{1}{1-\omega^l} - \frac{\omega^l}{1-\omega^l} = \frac{d}{4} \end{aligned}$$

For the summation, we invoke 1., so we have

$$g(r) = \frac{d}{4} - \sum_{s=1}^{r} \left[\delta_{r,s} \frac{d}{2} + \delta_{r,s+\frac{d}{2}} \left(-\frac{d}{2} \right) \right],$$

Notice the second δ can be 1 only if $r \geq \frac{d}{2} + 1$, while the first δ can always be satisfied. □

With this trick we can further simplify $f(\eta)$. For $m = 0$,

$$S(d, 0) = \sum_{k=\text{odd}} (1+\omega^k)^{d-1} \frac{\omega^k}{1-\omega^k}.$$

By binomial expansion,

$$S(d,0) = \sum_{k=\text{odd}}^{d-1} \sum_{\alpha=0}^{d-1} \binom{d-1}{\alpha} (\omega^k)^\alpha \frac{\omega^k}{1-\omega^k} = \sum_{\alpha=0}^{d-1} \binom{d-1}{\alpha} g(\alpha+1) = 0,$$

due to Prop. 5.12(2) and the symmetry of the binomial coefficients, $\sum_{\alpha=0}^{d/2-1} \binom{d-1}{a} - \sum_{\alpha=d/2}^{d-1} \binom{d-1}{a} = 0$. This is in line with the fact that $f(\eta)$ is an odd function, so the constant term in its power expansion must be 0.

For $m \geq 1$. By binomial expansion, we have

$$\begin{aligned} S(d,m) &= \sum_{k=\text{odd}}^{d-1} (1-\omega^k)^{m-1} (1+\omega^k)^{d-m-1} \omega^k \\ &= \sum_{k=\text{odd}}^{d-1} \sum_{\alpha=0}^{m-1} \sum_{\beta=0}^{d-m-1} \binom{m-1}{\alpha} \binom{d-m-1}{\beta} (-\omega^k)^\alpha (\omega^k)^\beta \omega^k \end{aligned}$$

Exchange the order of the sums, and apply Prop. 5.12(1),

$$\begin{aligned} S(d,m) &= \sum_{\alpha=0}^{m-1}\sum_{\beta=0}^{d-m-1}\binom{m-1}{\alpha}\binom{d-m-1}{\beta}(-1)^{\alpha}\sum_{k=\text{odd}}^{d-1}(\omega^k)^{\alpha+\beta+1} \\ &= \sum_{\alpha=0}^{m-1}\sum_{\beta=0}^{d-m-1}\binom{m-1}{\alpha}\binom{d-m-1}{\beta}(-1)^{\alpha}\frac{d}{2}\left(\delta_{\alpha+\beta+1,0}-\delta_{\beta,\frac{d}{2}-\alpha-1}\right) \\ &= \frac{d}{2}\sum_{\alpha=0}^{m-1}\binom{m-1}{\alpha}\binom{d-m-1}{\frac{d}{2}-\alpha-1}(-1)^{\alpha+1}, \end{aligned}$$

where $\binom{m-1}{\alpha}\binom{d-m-1}{\beta}\delta_{\alpha+\beta+1,0}=0$ in the second line, since $1\le\alpha+\beta+1\le d-1$. Now the summation over α can be evaluated (with *Mathematica*):

$$S(d,m) = -\frac{d}{2}\frac{2^{d-2}\Gamma(\frac{d-m}{2})}{\Gamma(\frac{d}{2})\Gamma(1-\frac{m}{2})}, \tag{5.42}$$

where the $\Gamma(\cdot)$ is the Gamma function [139]. When m is even, the denominator diverges to infinity, so $S(d,m)=0$ for even m. This confirms the fact that f is an odd function, who should not contain even power in its power expansion.

We invoke property of the Gamma function to state the expression in the more familiar factorial form. Note that, for non-negative integer μ:

$$\Gamma(\mu+1)=\mu!, \qquad \Gamma(\frac{1}{2}+\mu)=\frac{(2\mu)!}{4^{\mu}\mu!}\sqrt{\pi}, \qquad \Gamma(\frac{1}{2}-\mu)=\frac{(-4)^{\mu}\mu!}{(2\mu)!}\sqrt{\pi}.$$

Let us denote $m=2\mu+1$ for a non-negative integer $0\le\mu\le\frac{d}{2}-1$, then

$$\begin{aligned} S(d,m) &= -\frac{d\,2^{d-2}}{2}\frac{\Gamma(\frac{1}{2}+(\frac{d}{2}-\mu-1))}{\Gamma(\frac{d}{2})\Gamma(\frac{1}{2}-\mu)} \\ &= -\frac{d\,2^{d-2}}{2}\cdot\frac{(2(\frac{d}{2}-\mu-1))!}{4^{\frac{d}{2}-\mu-1}(\frac{d}{2}-\mu-1)!}\cdot\frac{1}{(\frac{d}{2}-1)!}\cdot\frac{(2\mu)!}{(-4)^{\mu}\mu!} \\ &= (-1)^{\mu+1}\frac{d}{2}\cdot\frac{1}{(\frac{d}{2}-1)!}\cdot\frac{(2(\frac{d}{2}-\mu-1))!}{(\frac{d}{2}-\mu-1)!}\cdot\frac{(2\mu)!}{\mu!} \end{aligned}$$

Substitute $S(d,m)$ to Eqn. (5.39), and recall $m = 2\mu + 1$:

$$\begin{aligned} f(\eta) &= \frac{16}{d^2 4^{d-1}} \sum_{m=\text{odd}}^{d-1} \binom{d-1}{m} \eta^m \left(\frac{d}{2} \cdot \frac{1}{(\frac{d}{2}-1)!} \cdot \frac{(2(\frac{d}{2}-\mu-1))!}{(\frac{d}{2}-\mu-1)!} \cdot \frac{(2\mu)!}{\mu!} \right)^2 \\ &= \frac{16}{4^d} \sum_{m=\text{odd}}^{d-1} \eta^m \frac{(d-1)!}{m!(d-m-1)!} \cdot \left(\frac{1}{(\frac{d}{2}-1)!} \cdot \frac{(2(\frac{d}{2}-\mu-1))!}{(\frac{d}{2}-\mu-1)!} \cdot \frac{(2\mu)!}{\mu!} \right)^2 \\ &= \frac{16}{4^d} \sum_{m=\text{odd}}^{d-1} \eta^m \frac{(d-1)!}{m!(d-m-1)!} \cdot \left(\frac{1}{(\frac{d}{2}-1)!} \cdot \frac{(d-m-1)!}{(\frac{d-m-1}{2})!} \cdot \frac{(m-1)!}{(\frac{m-1}{2})!} \right)^2 \end{aligned} \tag{5.43}$$

$$=: \sum_{m=\text{odd}}^{d-1} C(d,m)\eta^m, \tag{5.44}$$

where we write $C(d,m)$ in terms from Gamma functions:

$$C(d,m) := \frac{16}{4^d} \frac{\Gamma(d)}{\Gamma(m+1)\Gamma(d-m)} \left(\frac{\Gamma(d-m)\Gamma(m)}{\Gamma(\frac{d}{2})\Gamma(\frac{d-m+1}{2})\Gamma(\frac{m+1}{2})} \right)^2. \tag{5.45}$$

As an example, in Table 5.1, we list the analytical expression for the MBL_N set for $N = 1, 3, 5$ ($d = 2, 4, 6$), recall Prop. 5.11, where $u = x - y$ and $v = x + y$. Compare with the numerical result shown in Fig. 5.4, we are convinced that this

Table 5.1: Examples of the boundary of the MBL_N sets, for $N = 1, 3, 5$. $N = 1$ is simply the CHSH inequality.

N	expression
1	$4y = 2$
3	$y(3 + 3x^2 + y^2) = 2$
5	$\frac{1}{16}y(45 + 45x^4 + 10y^2 + 9y^4 + 30x^2(1 + 3y^2)) = 2$

lifting of the CHSH inequality is the only relevant inequality characterise the MBL_N sets on this slice. In the next section, we will take the limit N goes to infinity and study the behaviour of the boundary of MBL_∞.

5.4.4 Solution for infinite copies

One of the motivation to characterize these sets is the question of whether the set of the quantum distributions, Q coincides with those that become local when we combine infinitely many copies, MBL_∞. In this section, we will show that, on this particular slice, the two sets coincide.

The boundary of the quantum set on this slice can be derived from the Tsirelson--Landau-Masanes inequalities [140–142]:

$$|\sum_{X,Y} \arcsin(E_{XY}) - 2\arcsin(E_{X'Y'})| \leq \pi, \quad \forall X', Y', \tag{5.46}$$

where $E_{XY} = P(a = b|X, Y) - P(a \neq b|X, Y)$.

In particular, the following

$$-\arcsin(E_{00}) + \arcsin(E_{01}) + \arcsin(E_{10}) + \arcsin(E_{11}) \leq \pi, \tag{5.47}$$

is suitable for the top sector of this slice. Using our parametrization, it is not difficult to see that $E_{00} = -E_{11} = u$ and $E_{01} = E_{10} = v$. In terms of u and v, we can rewrite the above as:

$$\frac{4}{\pi}\arcsin(v) - \frac{4}{\pi}\arcsin(u) = 2, \tag{5.48}$$

with the right hand side purposely chosen to compare with the bound of the CHSH inequality. Note that the Taylor expansion of $\arcsin(\eta)$ is:

$$\arcsin(\eta) = \sum_{m=\text{odd}}^{\infty} \frac{1}{m}\frac{(m-2)!!}{(m-1)!!}\eta^m, \tag{5.49}$$

where $(\cdot)!!$ is the double factorial, defined as

$$n!! = \begin{cases} n \cdot (n-2) \cdots 5 \cdot 3 \cdot 1, & \text{when } n \text{ is odd,} \\ n \cdot (n-2) \cdots 6 \cdot 4 \cdot 2, & \text{when } n \text{ is even,} \\ 1, & n = 0 \text{ or } -1. \end{cases}$$

We now turn to the family of curves defined by the lifting of CHSH. Let us introduce some properties of Gamma function. First is what is called the duplication formula:

$$\Gamma(\mu)\Gamma(\mu + \frac{1}{2}) = 2^{1-2\mu}\sqrt{\pi}\Gamma(2\mu). \tag{5.50}$$

Second is a limit:

$$\lim_{n\to\infty} \frac{\Gamma(n+\alpha)}{\Gamma(n)n^\alpha} = 1. \tag{5.51}$$

Consider the coefficients of the power expansion, $C(d,m)$, when d tends to infinity:

$$C(\infty,m) := \lim_{d\to\infty} C(d,m) = \lim_{d\to\infty} \frac{16}{4^d}\frac{\Gamma(d)\Gamma(d-m)}{\Gamma^2(\frac{d}{2})\Gamma^2(\frac{d-m+1}{2})}\cdot\frac{\Gamma^2(m)}{\Gamma(m+1)\Gamma^2(\frac{m+1}{2})}.$$

By the duplication formula,

$$\begin{aligned} C(\infty,m) &= \frac{\Gamma^2(m)}{\Gamma(m+1)\Gamma^2(\frac{m+1}{2})}\cdot\lim_{d\to\infty}\frac{16}{4^d}\frac{2^{d-1}}{\sqrt{\pi}}\frac{\Gamma(\frac{d}{2}+\frac{1}{2})}{\Gamma(\frac{d}{2})}\cdot\frac{2^{d-m-1}}{\sqrt{\pi}}\frac{\Gamma(\frac{d-m}{2})}{\Gamma(\frac{d-m}{2}+\frac{1}{2})} \\ &= \frac{16}{2^{m+2}\pi}\cdot\frac{\Gamma^2(m)}{\Gamma(m+1)\Gamma^2(\frac{m+1}{2})}\lim_{d\to\infty}\frac{\Gamma(\frac{d}{2}+\frac{1}{2})}{\Gamma(\frac{d}{2})}\cdot\frac{\Gamma(\frac{d-m}{2})}{\Gamma(\frac{d-m}{2}+\frac{1}{2})} \\ &= \frac{16}{2^{m+2}\pi}\cdot\frac{\Gamma^2(m)}{\Gamma(m+1)\Gamma^2(\frac{m+1}{2})}, \end{aligned}$$

where the limit evaluates to 1. Note that $\Gamma(m+1) = m\Gamma(m)$, and applying the duplication formula,

$$C(\infty,m) = \frac{16}{2^{m+2}\pi}\cdot\frac{1}{m}\frac{2^{m-1}}{\sqrt{\pi}}\frac{\Gamma(\frac{m}{2})}{\Gamma(\frac{m+1}{2})} = \frac{2}{m\pi^{3/2}}\frac{\Gamma(\frac{m}{2})}{\Gamma(\frac{m+1}{2})}.$$

Finally, the double factorial and the Gamma function is connected by the following. Recall that $m = 2\mu+1$ and for integer μ,

$$\Gamma(\frac{m}{2}) = \Gamma(\mu+\frac{1}{2}) = \frac{(2\mu-1)!!}{2^\mu}\sqrt{\pi} = \frac{(m-2)!!}{2^\mu}\sqrt{\pi},$$

and

$$\Gamma(\frac{m+1}{2}) = \Gamma(\mu+1) = \mu! = \frac{(2\mu)!!}{2^\mu} = \frac{(m-1)!!}{2^\mu}.$$

Hence

$$C(\infty,m) = \frac{2}{\pi}\frac{1}{m}\frac{(m-2)!!}{(m-1)!!}, \tag{5.52}$$

which is exactly the Taylor coefficient of $\frac{2}{\pi}\arcsin(\eta)$ (see Eqn. (5.7)). Recall Prop. 5.11, we have

Theorem 5.13. *On the symmetric slice of the two-party, two-input, two-outcome no-signalling polytope, the boundary of the MBL_∞ set coincides with the boundary of the quantum set and is given by:*

$$\frac{4}{\pi}\arcsin(u) - \frac{4}{\pi}\arcsin(v) = 2.$$

In other words, on this symmetric slice, the quantum set Q is defined by the principle of MBL_N with N taken to infinity. It would be remarkable if this is the case for the entire no-signalling polytope because that would imply MBL is the principle that defines the quantum set (at least in this $(2,2;2,2)$ scenario).

Unfortunately however, on this slice, the quantum set Q coincides with Q_1 (the hierarchy collapses, $Q = \cdots = Q_2 = Q_1$). Thus the principle of ML identifies the quantum set as well as the principle of MBL. We can only conclude is that, on this slice, MBL does not rule out any quantum distribution; all the quantum distributions on this slice satisfy MBL_∞. In order to see if MBL defines a smaller set of correlation than ML, one has to go to another slice of the no-signalling polytope.

5.5 On another slice

We turn to another slice of the $(2,2;2,2)$ no-signalling polytope that passes through a PR-box, a local deterministic point and the completely mixed distribution. Let us denote any point on this two dimensional slice as (x,y), with the parametrization:

$$p(x,y) := x\,(P_{LD} - P_{mix}) + y\,(P_{PR} - P_{mix}) + P_{mix}, \tag{5.53}$$

where

$$P_{LD} = \left(\begin{array}{c|c|c} 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \end{array}\right),\ P_{PR} = \left(\begin{array}{c|c|c} 1 & \frac{1}{2} & \frac{1}{2} \\ \hline \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \hline \frac{1}{2} & \frac{1}{2} & 0 \end{array}\right),\ P_{mix} = \left(\begin{array}{c|c|c} 1 & \frac{1}{2} & \frac{1}{2} \\ \hline \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \hline \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \end{array}\right), \tag{5.54}$$

and $x, y \in [0,1]$, or in the full correlation form:

$$p(x,y) = \frac{1}{4}\left(\begin{array}{cc|cc} 1+3x+y & 1-x-y & 1+3x+y & 1-x-y \\ 1-x-y & 1-x+y & 1-x-y & 1-x+y \\ \hline 1+3x+y & 1-x-y & 1+3x-y & 1-x+y \\ 1-x-y & 1-x+y & 1-x+y & 1-x-y \end{array}\right). \tag{5.55}$$

With the SDP hierarchy, one can plot the set Q_1 and higher levels of the hierarchy in this slice. Already with local level one, we see that there is a separation between Q_1' and Q_1, that is the hierarchy does not collapse in this slice, see Fig. 5.6. Clearly Q_2' is strictly smaller than Q_1 and this implies that the true quantum set Q is strictly smaller than Q_1 one this slice. Note that both end points, $(0, 1/\sqrt{2})$

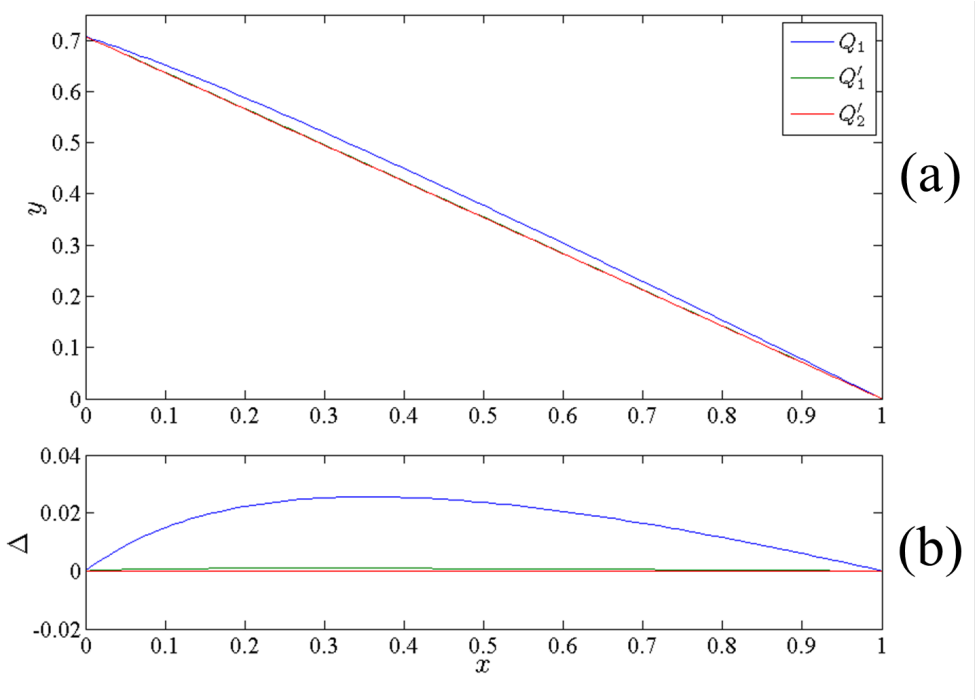


Figure 5.6: (a) The boundaries of three set of correlations, Q_1 (blue), Q_1' (green) and Q_2' (red), shown on the slice of the NS polytope. Q_1 correspond to the first level of the NPA hierarchy, while Q_1' and Q_2' correspond to the first and second of the Moroder hierarchy. Q_2' is a straight line up to numerical uncertainty. Q_1' and Q_2' are close to each other though they do not overlap as we can see the difference between the three curves and Q_2' as shown in (b).

and $(1, 0)$, in Fig. 5.6(a) are quantum, so are points on the straight line connecting them. With the upper bound Q_2' being a straight line up to numerical uncertainty, it is suggested that the boundary of quantum set Q on this slice is simply the straight line connecting the two extremal points.

As we have done for the symmetric slice before, we can numerically determine the boundary of the MBL_N set for small N's. In Fig. 5.7, we have shown the numerical determined boundary for MBL_2 and MBL_3 on this slice.

There are some remarkable qualitative feature of these sets. First, the even-odd oscillation is also present in this slice. Moreover, MBL_3 is not contained in MBL_2 as in the symmetric slice. Second, the sets are not convex. That is to say, given $P_1, P_2 \in MBL_N$, one cannot conclude that any convex combination $vP_1+(1-v)P_2$ also belongs to that set. More rigorously, we can show that:

Proposition 5.14. *MBL_2 are not convex, i.e. $\exists P_1, P_2 \in MBL_2$ such that $(1-\epsilon)P_1 + \epsilon P_2 \notin MBL_2$.*

Proof. Since we aim to show the existence, we can choose $P_1 = P$ such that $P \in MBL_2$ but $P \notin L$, and $P_2 = P_{LD}$ a local deterministic point.

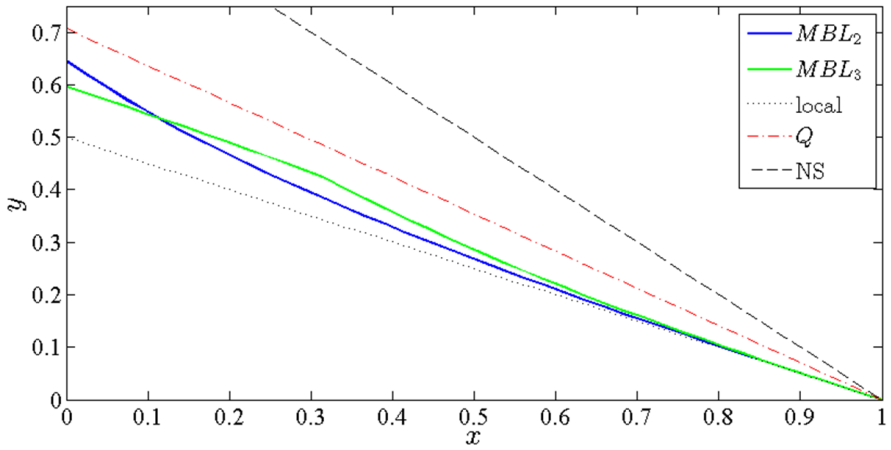


Figure 5.7: MBL_2 and MBL_3 on the other slice that passes through a local deterministic point. The dotted line represents the CHSH facet, dash-dot line the quantum set and dashed the no-signalling set.

Notice the fact that, when we convolute P with a local deterministic point, it is equivalent to apply a local relabelling of outcomes to P. For example, if the local deterministic point is the one that always output zero, then P remains the same, $P * P_{LD}(a,b) = P(a,b)$; if it is the one that always output one, then all the outcomes of P are shifted by $P * P_{LD}(a,b) = P(a-1,b-1)$. Clearly local relabelling of outcomes does not change the amount of nonlocality in a distribution. Convolution of local deterministic points P_{LD}^{*k} is also a deterministic point.

Consider the following distribution

$$\Big((1-\epsilon)P + \epsilon P_{LD}\Big)^{*2} = (1-\epsilon)^2 P^{*2} + 2\epsilon(1-\epsilon)P * P_{LD} + \epsilon^2 P_{LD}^{*2} \tag{5.56}$$

Clearly, the first term of the sum $(1-\epsilon)^2 P^2$ and the last term $\epsilon^2 P_{LN}^{*2}$ are local. However, the second term is nonlocal since P is nonlocal. For suitable choice of ϵ, $\Big((1-\epsilon)P + \epsilon P_{LD}\Big)^{*2} \notin L$, hence the set MBL_N is not convex. □

Nonetheless, we conjecture that the MBL_N sets will be convex in the limit of N tends to infinity.

Inequality and analytic result

One may aim to find analytic expression for the boundary of MBL_N set in the slice as we have done in the symmetric slice. Unfortunately, there is not a single inequality whose Fourier transform gives us the curve. Recall that Proposition 5.10 implies that the boundary maybe consist of concatenating several Fourier transformed Bell inequalities by taking the minimum. Indeed, this is the case for MBL_3 in this slice. The boundary of MBL_3 is made of two Fourier transformed CHSH

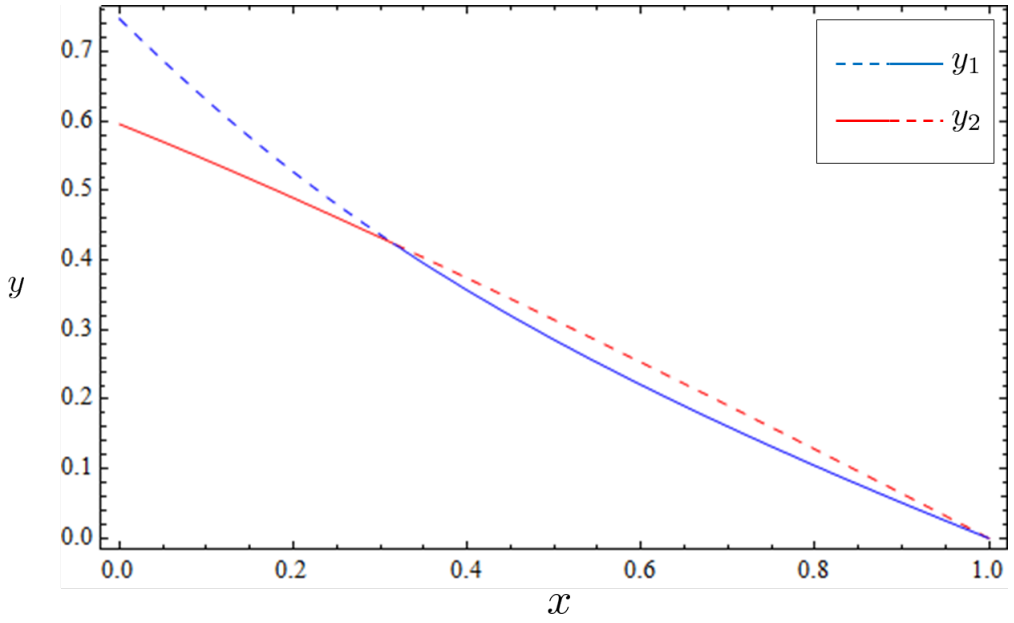


Figure 5.8: Analytic expression for MBL_3. Two curves $y_1(x)$ and $y_2(x)$ correspond to two liftings of CHSH c_1 and c_2. The boundary of MBL_3 is by taking the minimum of these two curves, shown in solid line.

liftings to $d = 4$ outcomes:

$$c_1 = \left(\begin{array}{cc|cc} -1_1 & 1_{1,3} & 1_1 & -1_{1,3} \\ 1_{3,1} & -1_3 & -1_{3,1} & 1_3 \\ \hline 1_1 & -1_{1,3} & 1_1 & -1_{1,3} \\ -1_{3,1} & 1_3 & -1_{3,1} & 1_3 \end{array}\right) \leq 2,$$

$$c_2 = \left(\begin{array}{cc|cc} -1_2 & 1_2 & 1_2 & -1_2 \\ 1_2 & -1_2 & -1_2 & 1_2 \\ \hline 1_2 & -1_2 & 1_2 & -1_2 \\ -1_2 & 1_2 & -1_2 & 1_2 \end{array}\right) \leq 2,$$

where $1_{m,n}$ is an $m \times n$ matrix filled with ones. c_1 corresponds to grouping the outcomes as $\{0\} \to \bar{0}$, $\{1,2,3\} \to \bar{1}$, and c_2 corresponds to $\{0,1\} \to \bar{1}$, $\{2,3\} \to \bar{1}$. By carrying out the Fourier transformation solving Eqn (5.20), one finds c_1 and c_2 correspond to:

$$\frac{1}{8}(9 + 19x^3 + 6y + 3y^2 + 2y^3 + x^2(3 + 54y) + 3x(-5 + 12y + 3y^2)) = 2,$$
$$\frac{1}{2}(-5x^3 - 6x^2(-1 + y) + 3x(1 + y^2) + 2y(3 + y^2)) = 2.$$

The analytic result for general N remains an open problem. Evidence suggests that it requires more than one inequality to compute the boundary; increasing

number of inequalities are needed as N increase. More techniques are needed to further study these sets in order to compare with the quantum set on this slice.

5.6 Conclusion

We have presented the notion of a new physical principle called many-box locality. The N-box local set, MBL_N, are examined numerically as well as analytically with novel tools based on Fourier transformation of probabilities and Bell coefficients. With that we analyse these sets on two slices of the $(2, 2; 2, 2)$ no-signalling polytope.

Let us discuss some the implications of possible relations between the quantum set Q, the almost quantum set Q_1' and the MBL_N sets. A priori, we know that $MBL_\infty \subseteq ML = Q_1$ and also $Q \subseteq Q_1' \subseteq Q_1$. However there is no obvious relation between the quantum set Q,the almost quantum set Q_1' and the MBL_∞ set. The even-odd oscillation aside, the MBL_N sets seem to be monotonically increasing for both even and odd N. These sets are bounded from above by the NS polytope, whose facets are positivity constraints. Hence we expect the MBL_N sets to converge. There are a few possible relations between Q and the MBL_N sets. If $Q \not\subseteq MBL_\infty$, it means there exists a quantum distribution whose N-box distribution is nonlocal for any N. In other words, nonlocality can be macroscopic, we may not observe this due to the limit of our measurement precision. If $MBL_\infty \subset Q_1'$, then MBL defines a set that is smaller than the almost quantum set, which satisfies all the previous physical principles. The most welcome relation would be that $MBL_\infty = Q$, in which case, MBL will be a first example of a physical principle that defines the quantum set.

On the symmetric slice, we derived an general expression for the boundary of MBL_N for all odd N; in the limit of $N \to \infty$, the principle of MBL recover the boundary of the quantum set. On another slice, analytic solution for general N is not directly available.

The relation between the set MBL_∞ and Q remains open. Many-box locality is still a promising principle that may define a set of distributions closer to the quantum set than the previously proposed.

CHAPTER 6

CONCLUSION AND OUTLOOK

In this thesis, we have explored two aspects of quantum physics that are in contrast with classical physics: complexity and nonlocality. We discussed three quantum size measures: tree size as a complexity measure, the dimension of a quantum systems, and a physical principle called many-box locality.

6.1 Tree size complexity

In Chapter 2, we studied tree size as a complexity measure for multiqubit states. We identified the most complex few qubit states according to this measure. We classified families of many-qubit states into simple and complex states depending on whether tree size scales polynomially or superpolynomially in the number of qubits. Raz's theorem allows us to show superpolynomial lower bound on tree size. A few classes of complex states were demonstrated, including the 2D cluster state. We also presented a complexity witness that can be efficiently measured.

The relation between tree size complexity and quantum computing is also discussed. In the measurement-based quantum computation, a resource state with superpolynomial tree size is necessary for universal quantum computation. In the circuit model, the treeBQP conjecture states that if tree size is polynomial at each step of the computation, then the computation can be efficiently simulated with classical computers. We were only able to show a weaker version: if at each step of the computation and every bipartition, there is a polynomial tree representation that is separating with respect to the bipartition, then the computation can be efficiently simulated.

Tree size has the advantage of being in principle computable and being able to show superpolynomial lower bounds. However, the operational interpretation

of tree size remains unclear. If treeBQP=BPP, then polynomial tree size implies efficient classical simulation, and hence large tree size is a necessary condition for quantum speed up.

Another limitation of tree size is the notion of individual qubits. To talk about the minimal tree representation, we allow any invertible local operations. In this definition, we have unconsciously chosen a preferred tensor structure of the Hilbert space. This maybe natural from a computer science perspective, since (quantum) computers after all are made of basic units of (qu)bit. However, in a physical system, there may not be such a preferred tensor structure. One may search for a complexity measure that is independent of the choice of a preferred basis. There are some state that are intrinsically multimode, for example $(a^\dagger)^2 + (b^\dagger)^2 |vac\rangle = |2,0\rangle + |0,2\rangle \neq (C^\dagger)^2 |vac\rangle$ cannot be written as a single mode state. One may try to use this minimum number of modes as a complexity measure.

Quantum mechanics do not put any restriction on tree size. One may try to construct physical models that restricts the complexity of quantum states and study what modifications to quantum mechanics would lead to such a restriction, see for example the toy model of [143]. Other than mass, macroscopicity, complexity, along what other axis can one test the limit of quantum mechanics?

6.2 Dimension witness

In Chapter 4, we studied a dimension witness based on a Bell inequality, called CGLMP_4 inequality. The dimension witness we studied is device independent in the sense that no assumptions are made on the state and measurements, and based only on the violation of the CGLMP_4 inequality. The maximal quantum violation, I_4^*, and the corresponding maximal violation state is found. Then an upper bound I_N on the maximal violation of CGLMP_4 inequality with qutrit systems is derived, based on negativity with an SDP, where as a lower bound on the maximal violation with qutrits, I_3^* is found with a non-linear optimization procedure. Any violation larger than I_N certifies the presence of a quantum system with dimension at least four. Finally, we discussed a feature of this dimension witness: it can be violated by sequential measurements of two pairs of maximally entangled qubits. Due to this feature, with this dimension witness, we can certify the generation of four dimensional entangled system (but trivially equivalent to two pairs of entangled qubits) and coherent manipulation on qubit systems only.

Assuming the CGLMP family of inequalities constitute dimension witnesses for higher dimension, it is still an open question whether this feature is present in

the whole family of CGLMP inequalities. One may hope to find another device independent dimension witness for qudit, such that its violation will certify not only the generation of qudits but also coherent manipulations over the d dimensional Hilbert space.

6.3 Many-box locality

In Chapter 5, we proposed a physical principle called many-box locality, in hope of defining the quantum set of distributions. It is similar to macroscopic locality but two assumptions are modified, and defines a possibly smaller set of distribution. The N-box local set, MBL_N, is studied numerically and analytically with a novel method based on Fourier transformation of probabilities and Bell coefficients. On the symmetric slice of the $(2, 2; 2, 2)$ no-signalling polytope, MBL_∞ coincides with the quantum set Q. On another slice, an analytic expression of the boundary of the MBL_N of general N is not directly available, neither is the limit MBL_∞. Thus we could not compare MBL_∞ and Q on that slice.

More tools is needed to analyse the MBL_N sets and their limit MBL_∞. Investigation can also be naturally extended beyond the simplest $(2, 2; 2, 2)$ scenario.

The tool of Fourier transformation of probabilities and Bell coefficients is interesting on its own. One may look for applications of this technique in other Bell inequality studies.

6.4 Quantum–classical boundary

Let us discuss how these size measures shed light on the quantum–classical boundary.

The quantum–classical boundary in terms of tree size complexity can seen at the superpolynomial-polynomial separation of tree size. This is not to say that any states with polynomial tree size is classical, because a Bell violation does not require a state with superpolynomial tree size. Rather we are saying that from a complexity point of view, states with polynomial tree size can be simulated with a classical computer. Hence, superpolynomial tree size can be seen as a quantum signature of complex states.

The quantum–classical boundary in the context of nonlocality has several meanings. The violation of Bell inequality *per se* is definitely a quantum–classical boundary. This boundary is blurred by for example our inability to resolve the outcomes in a Bell experiment. Mildly nonlocal distributions maybe become local hence compatible with classical physics, if we are unable to identify the N pairs

of particles. Some maximally nonlocal quantum boxes approach local as N tends to infinity. Our study of many-box locality suggests that the quantum-classical transition of the many-box distribution may define the boundary of the quantum set. The quantum–classical boundary of the many-box distribution may be related to the supraquantum–quantum boundary of the original microscopic distribution.

In the context of dimension witness, the quantum–classical boundary lies in our ability to manipulate the system. Every physical system can be described by quantum mechanics, so we cannot label a system as classical or quantum. The notion of classical dimension or quantum dimension arises due to the limitation of how one can prepare and manipulate the system. In the prepare-and-measure scenario, when one can only prepare states that commutes with each other, the system appears to have classical dimension d. In the Bell scenario, once a Bell inequality is violated, we can certify that we have a quantum system. Violating a dimension witness in a Bell scenario will certify a lower bound on the quantum dimension.

Complexity and nonlocality only captures a small part of quantum weirdness. Progress in the study of non-classicality of quantum phenomena will help us better understand the quantum Nature. Hopefully identifying and harnessing the true quantum power will lead to breakthrough in technology.

APPENDIX A

BIG O NOTATION

In computer science, big O notation is useful in the analysis of algorithms. One compares different algorithms by how their processing time or working space, respond to changes in the input size. Similarly, in the context of tree size, we are interested in the scaling behaviour of tree size (or other parameters) as a function of the number of qubits.

There are several related notations, o, O, Ω and Θ, defined as follows:

- $f(n) = o(g(n))$ or f is dominated by g asymptotically, means $\forall \epsilon > 0$, $\exists n_0$, such that $\forall n > n_0$, $|f(n)| \leq \epsilon \cdot |g(n)|$.
- $f(n) = O(g(n))$ or f is bounded above by g asymptotically, means $\exists k > 0$ and n_0 such that $\forall n > n_0$, $f(n) \leq k \cdot g(n)$.
- $f(n) = \Omega(g(n))$ or f is bounded below by g asymptotically, means $\exists k > 0$ and n_0 such that $\forall n > n_0$, $f(n) \geq k \cdot g(n)$.
- $f(n) = \Theta(g(n))$ or f is bounded above and below by g, means $\exists k_1 > 0, k_2 > 0$ and n_0 such that $\forall n > n_0$, $k_1 \cdot g(n) \leq f(n) \leq k_2 \cdot g(n)$.

BIBLIOGRAPHY

[1] A. Einstein, B. Podolsky, and N. Rosen. "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" *Phys. Rev.* **47** 10 (1935), pp. 777–780.

[2] J. S. Bell. "On the Einstein-Podolsky-Rosen paradox". *Physics* **1** 3 (1964), pp. 195–200.

[3] S. J. Freedman and J. F. Clauser. "Experimental Test of Local Hidden-Variable Theories". *Phys. Rev. Lett.* **28** 14 (1972), pp. 938–941.

[4] A. Aspect, P. Grangier, and G. Roger. "Experimental Tests of Realistic Local Theories via Bell's Theorem". *Phys. Rev. Lett.* **47** 7 (1981), pp. 460–463.

[5] A. Aspect, J. Dalibard, and G. Roger. "Experimental Test of Bell's Inequalities Using Time- Varying Analyzers". *Phys. Rev. Lett.* **49** 25 (1982), pp. 1804–1807.

[6] W. Tittel, J. Brendel, B. Gisin, T. Herzog, H. Zbinden, and N. Gisin. "Experimental demonstration of quantum correlations over more than 10 km". *Phys. Rev. A* **57** 5 (1998), pp. 3229–3232.

[7] W. Tittel, J. Brendel, H. Zbinden, and N. Gisin. "Violation of Bell Inequalities by Photons More Than 10 km Apart". *Phys. Rev. Lett.* **81** 17 (1998), pp. 3563–3566.

[8] M. A. Rowe, D. Kielpinski, V. Meyer, C. A. Sackett, W. M. Itano, C. Monroe, and D. J. Wineland. "Experimental violation of a Bell's inequality with efficient detection". *Nature* **409** 6822 (2001), pp. 791–794.

[9] J.-W. Pan, D. Bouwmeester, M. Daniell, H. Weinfurter, and A. Zeilinger. "Experimental test of quantum nonlocality in three-photon Greenberger–Horne–Zeilinger entanglement". *Nature* **403** 6769 (2000), pp. 515–519.

[10] B. G. Christensen, K. T. McCusker, J. B. Altepeter, B. Calkins, T. Gerrits, A. E. Lita, A. Miller, L. K. Shalm, Y. Zhang, S. W. Nam, N. Brunner, C. C. W. Lim, N. Gisin, and P. G. Kwiat. "Detection-Loophole-Free Test of Quantum Nonlocality, and Applications". *Phys. Rev. Lett.* **111** 13 (2013), p. 130406.

[11] M. Giustina, A. Mech, S. Ramelow, B. Wittmann, J. Kofler, J. Beyer, A. Lita, B. Calkins, T. Gerrits, S. W. Nam, R. Ursin, and A. Zeilinger. "Bell violation using entangled photons without the fair-sampling assumption". *Nature* **497** 7448 (2013), pp. 227–230.

[12] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. L. Vermeulen, R. N. Schouten, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, M. Markham, D. J. Twitchen, D. Elkouss, S. Wehner, T. H. Taminiau, and R. Hanson. "Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres". *Nature* **advance online publication** (2015).

[13] E. Schrödinger. "Die gegenwärtige Situation in der Quantenmechanik". *Naturwissenschaften* **23** 48 (1935), pp. 807–812.

[14] E. Schrödinger. *The present situation in quantum mechanics: a translation of Schrödinger's "cat paradox" paper*. Trans. by J. D. Trimmer.

[15] F. Fröwis and W. Dür. "Measures of macroscopicity for quantum spin systems". *New J. Phys.* **14** 9 (2012), p. 093039.

[16] S. Nimmrichter and K. Hornberger. "Macroscopicity of Mechanical Quantum Superposition States". *Phys. Rev. Lett.* **110** 16 (2013), p. 160403.

[17] B. Yadin and V. Vedral. "A general framework for quantum macroscopicity in terms of coherence". *arXiv:1505.03792 [quant-ph]* (2015). arXiv: 1505.03792.

[18] M. Schlosshauer. "Decoherence, the measurement problem, and interpretations of quantum mechanics". *Rev. Mod. Phys.* **76** 4 (2005), pp. 1267–1305.

[19] P. Pearle. "Reduction of the state vector by a nonlinear Schr\"odinger equation". *Phys. Rev. D* **13** 4 (1976), pp. 857–868.

[20] G. C. Ghirardi, A. Rimini, and T. Weber. "Unified dynamics for microscopic and macroscopic systems". *Phys. Rev. D* **34** 2 (1986), pp. 470–491.

[21] C. Monroe, D. M. Meekhof, B. E. King, and D. J. Wineland. "A "Schrodinger Cat" Superposition State of an Atom". *Science* **272** 5265 (1996), pp. 1131–1136.

[22] M. Aspelmeyer, T. J. Kippenberg, and F. Marquardt. "Cavity optomechanics". *Rev. Mod. Phys.* **86** 4 (2014), pp. 1391–1452.

[23] J. R. Friedman, V. Patel, W. Chen, S. K. Tolpygo, and J. E. Lukens. "Quantum superposition of distinct macroscopic states". *Nature* **406** 6791 (2000), pp. 43–46.

[24] M. Arndt, O. Nairz, J. Vos-Andreae, C. Keller, G. van der Zouw, and A. Zeilinger. "Waveparticle duality of C60 molecules". *Nature* **401** 6754 (1999), pp. 680–682.

[25] R. Landauer. "The physical nature of information". *Physics Letters A* **217** 45 (1996), pp. 188–193.

[26] S. Goldstein, T. Norsen, D. Tausk, and N. Zanghi. "Bell's theorem". *Scholarpedia* **6** 10 (2011), p. 8378.

[27] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner. "Bell nonlocality". *Rev. Mod. Phys.* **86** 2 (2014), pp. 419–478.

[28] A. K. Ekert. "Quantum cryptography based on Bell's theorem". *Phys. Rev. Lett.* **67** 6 (1991), pp. 661–663.

[29] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani. "Device-Independent Security of Quantum Cryptography against Collective Attacks". *Phys. Rev. Lett.* **98** 23 (2007), p. 230501.

[30] S. Pironio, A. Acín, S. Massar, A. B. de la Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe. "Random numbers certified by Bells theorem". *Nature* **464** 7291 (2010), pp. 1021–1024.

[31] A. Acín, S. Massar, and S. Pironio. "Randomness versus Nonlocality and Entanglement". *Phys. Rev. Lett.* **108** 10 (2012), p. 100402.

[32] J.-D. Bancal, N. Gisin, Y.-C. Liang, and S. Pironio. "Device-Independent Witnesses of Genuine Multipartite Entanglement". *Phys. Rev. Lett.* **106** 25 (2011), p. 250404.

[33] N. Brunner, M. Navascués, and T. Vértesi. "Dimension Witnesses and Quantum State Discrimination". *Phys. Rev. Lett.* **110** 15 (2013), p. 150501.

[34] D. Mayers and A. Yao. "Self testing quantum apparatus". *arXiv:quant-ph/0307205* (2003). arXiv: quant-ph/0307205.

[35] M. McKague, T. H. Yang, and V. Scarani. "Robust self-testing of the singlet". *J. Phys. A: Math. Theor.* **45** 45 (2012), p. 455304.

[36] B. W. Reichardt, F. Unger, and U. Vazirani. "Classical command of quantum systems". *Nature* **496** 7446 (2013), pp. 456–460.

[37] T. H. Yang and M. Navascués. "Robust self-testing of unknown quantum systems into any entangled two-qubit states". *Phys. Rev. A* **87** 5 (2013), p. 050102.

[38] H. N. Le, Y. Cai, X. Wu, and V. Scarani. "Tree-size complexity of multiqubit states". *Phys. Rev. A* **88** 1 (2013), p. 012321.

[39] H. N. Le, Y. Cai, X. Wu, R. Rabelo, and V. Scarani. "Maximal tree size of few-qubit states". *Phys. Rev. A* **89** 6 (2014), p. 062333.

[40] Y. Cai, H. N. Le, and V. Scarani. "State complexity and quantum computation". *Ann. Phys.* **527** 9 (2015), 684–700.

[41] K. Hornberger, S. Gerlich, P. Haslinger, S. Nimmrichter, and M. Arndt. "Colloquium: Quantum interference of clusters and molecules". *Rev. Mod. Phys.* **84** 1 (2012), pp. 157–173.

[42] A. J. Leggett. "Testing the limits of quantum mechanics: motivation, state of play, prospects". *J. Phys.: Condens. Matter* **14** 15 (2002), R415.

[43] W. Dür, C. Simon, and J. I. Cirac. "Effective Size of Certain Macroscopic Quantum Superpositions". *Phys. Rev. Lett.* **89** 21 (2002), p. 210402.

[44] G. Björk and P. G. L. Mana. "A size criterion for macroscopic superposition states". *J. Opt. B: Quantum Semiclass. Opt.* **6** 11 (2004), p. 429.

[45] J. I. Korsbakken, K. B. Whaley, J. Dubois, and J. I. Cirac. "Measurement-based measure of the size of macroscopic quantum superpositions". *Phys. Rev. A* **75** 4 (2007), p. 042106.

[46] F. Marquardt, B. Abel, and J. von Delft. "Measuring the size of a quantum superposition of many-body states". *Phys. Rev. A* **78** 1 (2008), p. 012109.

[47] C.-W. Lee and H. Jeong. "Quantification of Macroscopic Quantum Superpositions within Phase Space". *Phys. Rev. Lett.* **106** 22 (2011), p. 220401.

[48] M. Arndt and K. Hornberger. "Testing the limits of quantum mechanical superpositions". *Nat Phys* **10** 4 (2014), pp. 271–277.

[49] S. Aaronson. "Multilinear Formulas and Skepticism of Quantum Computing". In: *In Proc. ACM STOC*. ACM Press, 2004, pp. 118–127.

[50] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[51] D. Braun and B. Georgeot. "Quantitative measure of interference". *Phys. Rev. A* **73** 2 (2006), p. 022314.

[52] D. Braun and B. Georgeot. "Interference versus success probability in quantum algorithms with imperfections". *Phys. Rev. A* **77** 2 (2008), p. 022318.

[53] G. Vidal. "Efficient Classical Simulation of Slightly Entangled Quantum Computations". *Phys. Rev. Lett.* **91** 14 (2003), p. 147902.

[54] M. Van den Nest. "Universal Quantum Computation with Little Entanglement". *Phys. Rev. Lett.* **110** 6 (2013), p. 060504.

[55] D. Gross, S. T. Flammia, and J. Eisert. "Most Quantum States Are Too Entangled To Be Useful As Computational Resources". *Phys. Rev. Lett.* **102** 19 (2009), p. 190501.

[56] R. Jozsa and N. Linden. "On the role of entanglement in quantum-computational speed-up". *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **459** 2036 (2003), pp. 2011–2032.

[57] D. Gottesman. "The Heisenberg representation of quantum computers". *arXiv preprint quant-ph/9807006* (1998).

[58] A. Datta, S. T. Flammia, and C. M. Caves. "Entanglement and the power of one qubit". *Phys. Rev. A* **72** 4 (2005), p. 042316.

[59] A. Datta, A. Shaji, and C. M. Caves. "Quantum Discord and the Power of One Qubit". *Phys. Rev. Lett.* **100** 5 (2008), p. 050502.

[60] W. Weaver. "Science and complexity". In: *Facets of Systems Science*. Springer, 1991, pp. 449–456.

[61] A. N. Kolmogorov. "On tables of random numbers". *Theoretical Computer Science* **207** 2 (1998), pp. 387–395.

[62] M. Müller. "On the quantum Kolmogorov complexity of classical strings". *Int. J. Quantum Inform.* **07** 04 (2009), pp. 701–711.

[63] P. M. B. Vitanyi. "Quantum Kolmogorov Complexity Based on Classical Descriptions". *IEEE Transactions on Information Theory* **47** 6 (2001). arXiv: quant-ph/0102108.

[64] A. Berthiaume, W. van Dam, and S. Laplante. "Quantum Kolmogorov Complexity". *Journal of Computer and System Sciences* **63** 2 (2001), pp. 201–221.

[65] A. C.-C. Yao. "Quantum circuit complexity". In: *, 34th Annual Symposium on Foundations of Computer Science, 1993. Proceedings.* 1993, pp. 352–361.

[66] A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. A. Smolin, and H. Weinfurter. "Elementary gates for quantum computation". *Physical Review A* **52** 5 (1995), p. 3457.

[67] M. Möttönen, J. Vartiainen, V. Bergholm, and M. Salomaa. "Quantum Circuits for General Multiqubit Gates". *Physical Review Letters* **93** 13 (2004).

[68] G. Vidal and C. M. Dawson. "Universal quantum circuit for two-qubit transformations with three controlled-NOT gates". *Phys. Rev. A* **69** 1 (2004), p. 010301.

[69] F. Vatan and C. Williams. "Optimal quantum circuits for general two-qubit gates". *Phys. Rev. A* **69** 3 (2004), p. 032315.

[70] V. V. Shende and I. L. Markov. "On the CNOT-cost of TOFFOLI gates". *arXiv preprint arXiv:0803.2316* (2008).

[71] S. Aaronson and D. Gottesman. "Improved simulation of stabilizer circuits". *Phys. Rev. A* **70** 5 (2004), p. 052328.

[72] A. Acín, A. Andrianov, L. Costa, E. Jané, J. I. Latorre, and R. Tarrach. "Generalized Schmidt Decomposition and Classification of Three-Quantum-Bit States". *Phys. Rev. Lett.* **85** 7 (2000), pp. 1560–1563.

[73] H. A. Carteret, A. Higuchi, and A. Sudbery. "Multipartite generalization of the Schmidt decomposition". *Journal of Mathematical Physics* **41** 12 (2000), pp. 7932–7939.

[74] W. Dür, G. Vidal, and J. I. Cirac. "Three qubits can be entangled in two inequivalent ways". *Physical Review A* **62** 6 (2000), p. 062314.

[75] L. Lamata, J. León, D. Salgado, and E. Solano. "Inductive classification of multipartite entanglement under stochastic local operations and classical communication". *Phys. Rev. A* **74** 5 (2006), p. 052336.

[76] L. Lamata, J. León, D. Salgado, and E. Solano. "Inductive entanglement classification of four qubits under stochastic local operations and classical communication". *Phys. Rev. A* **75** 2 (2007), p. 022318.

[77] M. Bourennane, M. Eibl, C. Kurtsiefer, S. Gaertner, H. Weinfurter, O. Gühne, P. Hyllus, D. BruSS, M. Lewenstein, and A. Sanpera. "Experimental Detection of Multipartite Entanglement using Witness Operators". *Physical Review Letters* **92** 8 (2004).

[78] M. Eibl, S. Gaertner, M. Bourennane, C. Kurtsiefer, M. Żukowski, and H. Weinfurter. "Experimental Observation of Four-Photon Entanglement from Parametric Down-Conversion". *Phys. Rev. Lett.* **90** 20 (2003), p. 200403.

[79] F. Verstraete and J. I. Cirac. "Matrix product states represent ground states faithfully". *Phys. Rev. B* **73** 9 (2006), p. 094423.

[80] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac. "Matrix product state representations". *arXiv preprint quant-ph/0608197* (2006).

[81] R. Orús. "A practical introduction to tensor networks: Matrix product states and projected entangled pair states". *Annals of Physics* **349** (2014), pp. 117–158.

[82] R. Raz. "Multi-linear formulas for permanent and determinant are of super-polynomial size". *J. ACM* **56** 2 (2009), 8:1–8:17.

[83] E. W. Weisstein. *Ryser Formula.*

[84] L. G. Valiant. "The complexity of computing the permanent". *Theoretical computer science* **8** 2 (1979), pp. 189–201.

[85] S. Arora and B. Barak. *Computational Complexity: A Modern Approach*. IT Pro. Cambridge University Press, 2009.

[86] R. S. Bird. "A simple division-free algorithm for computing determinants". *Information Processing Letters* **111** 2122 (2011), pp. 1072 –1074.

[87] D. Deutsch and R. Jozsa. "Rapid Solution of Problems by Quantum Computation". *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **439** 1907 (1992), pp. 553–558.

[88] P. W. Shor. "Polynomial time algorithms for prime factorization and discrete logarithms on a quantum computer". *SIAM J. Sci. Statist. Comput.* **26** (1997), p. 1484.

[89] J. P. Buhler, H. W. Lenstra Jr, and C. Pomerance. "Factoring integers with the number field sieve". In: *The development of the number field sieve*. Springer, 1993, pp. 50–94.

[90] E. Assmus and J. Key. *Designs and Their Codes*. Cambridge Tracts in Mathematics. Cambridge University Press, 1992.

[91] H. J. Briegel and R. Raussendorf. "Persistent Entanglement in Arrays of Interacting Particles". *Phys. Rev. Lett.* **86** 5 (2001), pp. 910–913.

[92] R. Raussendorf and H. J. Briegel. "A One-Way Quantum Computer". *Phys. Rev. Lett.* **86** 22 (2001), pp. 5188–5191.

[93] R. Raussendorf, D. E. Browne, and H. J. Briegel. "Measurement-based quantum computation on cluster states". *Phys. Rev. A* **68** 2 (2003), p. 022312.

[94] C. A. Sackett, D. Kielpinski, B. E. King, C. Langer, V. Meyer, C. J. Myatt, M. Rowe, Q. A. Turchette, W. M. Itano, D. J. Wineland, and C. Monroe. "Experimental entanglement of four particles". *Nature* **404** 6775 (2000), pp. 256–259.

[95] O. Gühne, P. Hyllus, D. Bruß, A. Ekert, M. Lewenstein, C. Macchiavello, and A. Sanpera. "Detection of entanglement with few local measurements". *Phys. Rev. A* **66** 6 (2002), p. 062305.

[96] G. Tóth and O. Gühne. "Detecting Genuine Multipartite Entanglement with Two Local Measurements". *Physical Review Letters* **94** 6 (2005).

[97] S. Anders and H. J. Briegel. "Fast simulation of stabilizer circuits using a graph-state representation". *Phys. Rev. A* **73** 2 (2006), p. 022334.

[98] I. Pitowsky. *Quantum probability-quantum logic*. Lecture notes in physics. Springer-Verlag, 1989.

[99] S. Popescu and D. Rohrlich. "Quantum nonlocality as an axiom". *Found Phys* **24** 3 (1994), pp. 379–385.

[100] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt. "Proposed Experiment to Test Local Hidden-Variable Theories". *Phys. Rev. Lett.* **23** 15 (1969), pp. 880–884.

[101] B. S. Cirel'son. "Quantum generalizations of Bell's inequality". *Letters in Mathematical Physics* **4** 2 (1980), pp. 93–100.

[102] M. Navascués, S. Pironio, and A. Acín. "A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations". *New J. Phys.* **10** 7 (2008), p. 073013.

[103] T. Moroder, J.-D. Bancal, Y.-C. Liang, M. Hofmann, and O. Gühne. "Device-Independent Entanglement Quantification and Related Applications". *Phys. Rev. Lett.* **111** 3 (2013), p. 030501.

[104] N. Brunner, S. Pironio, A. Acin, N. Gisin, A. A. Méthot, and V. Scarani. "Testing the Dimension of Hilbert Spaces". *Phys. Rev. Lett.* **100** 21 (2008), p. 210503.

[105] R. Gallego, N. Brunner, C. Hadley, and A. Acín. "Device-Independent Tests of Classical and Quantum Dimensions". *Phys. Rev. Lett.* **105** 23 (2010), p. 230501.

[106] H. Buhrman, R. Cleve, S. Massar, and R. de Wolf. "Nonlocality and communication complexity". *Rev. Mod. Phys.* **82** 1 (2010), pp. 665–698.

[107] S. Wehner, M. Christandl, and A. C. Doherty. "Lower bound on the dimension of a quantum system given measured data". *Phys. Rev. A* **78** 6 (2008), p. 062112.

[108] O. Gühne, C. Budroni, A. Cabello, M. Kleinmann, and J.-Å. Larsson. "Bounding the quantum dimension with contextuality". *Phys. Rev. A* **89** 6 (2014), p. 062107.

[109] M. M. Wolf and D. Perez-Garcia. "Assessing Quantum Dimensionality from Observable Dynamics". *Phys. Rev. Lett.* **102** 19 (2009), p. 190504.

[110] A. Mukherjee, A. Roy, S. S. Bhattacharya, S. Das, M. R. Gazi, and M. Banik. "Device independent Schmidt rank witness by using Hardy paradox". *arXiv:1407.2146 [quant-ph]* (2014). arXiv: 1407.2146.

[111] M. Hendrych, R. Gallego, M. Miuda, N. Brunner, A. Acín, and J. P. Torres. "Experimental estimation of the dimension of classical and quantum systems". *Nat Phys* **8** 8 (2012), pp. 588–591.

[112] J. Ahrens, P. Badzig, A. Cabello, and M. Bourennane. "Experimental device-independent tests of classical and quantum dimensions". *Nat. Phys.* **8** 8 (2012), pp. 592–595.

[113] A. C. Dada, J. Leach, G. S. Buller, M. J. Padgett, and E. Andersson. "Experimental high-dimensional two-photon entanglement and violations of generalized Bell inequalities". *Nature Physics* **7** 9 (2011), pp. 677–680.

[114] M. Krenn, M. Huber, R. Fickler, R. Lapkiewicz, S. Ramelow, and A. Zeilinger. "Generation and confirmation of a (100 ×100)-dimensional entangled quantum system". *PNAS* **111** 17 (2014), pp. 6243–6247.

[115] D. Collins, N. Gisin, N. Linden, S. Massar, and S. Popescu. "Bell Inequalities for Arbitrarily High-Dimensional Systems". *Phys. Rev. Lett.* **88** 4 (2002), p. 040404.

[116] D. Collins and N. Gisin. "A relevant two qubit Bell inequality inequivalent to the CHSH inequality". *J. Phys. A: Math. Gen.* **37** 5 (2004), p. 1775.

[117] V. Scarani, N. Gisin, N. Brunner, L. Masanes, S. Pino, and A. Acín. "Secrecy extraction from no-signaling correlations". *Phys. Rev. A* **74** 4 (2006), p. 042339.

[118] A. Acín, T. Durt, N. Gisin, and J. I. Latorre. "Quantum nonlocality in two three-level systems". *Phys. Rev. A* **65** 5 (2002), p. 052325.

[119] S. Zohren and R. D. Gill. "Maximal Violation of the Collins-Gisin-Linden-Massar-Popescu Inequality for Infinite Dimensional States". *Phys. Rev. Lett.* **100** 12 (2008), p. 120406.

[120] T. H. Yang, T. Vértesi, J.-D. Bancal, V. Scarani, and M. Navascués. "Robust and Versatile Black-Box Certification of Quantum Devices". *Phys. Rev. Lett.* **113** 4 (2014), p. 040401.

[121] Y.-C. Liang, C.-W. Lim, and D.-L. Deng. "Reexamination of a multisetting Bell inequality for qudits". *Phys. Rev. A* **80** 5 (2009), p. 052116.

[122] K. F. Pál and T. Vértesi. "Maximal violation of a bipartite three-setting, two-outcome Bell inequality using infinite-dimensional quantum systems". *Phys. Rev. A* **82** 2 (2010), p. 022116.

[123] L. Hardy. "Quantum Theory From Five Reasonable Axioms". *arXiv:quant-ph/0101012* (2001). arXiv: quant-ph/0101012.

[124] G. Ludwig. "Attempt of an axiomatic foundation of quantum mechanics and more general theories. II". *Comm. Math. Phys.* 5 (1967), pp. 331–348.

[125] G. Ludwig. "An improved formulation of some theorems and axioms in the axiomatic foundation of the Hilbert space structure of quantum mechanics". *Comm. Math. Phys.* 1 (1972), pp. 78–86.

[126] C. Piron. *Axiomatique quantique*. pt. 2. impr. Birkhäuser, 1964.

[127] G. Chiribella, G. M. D'Ariano, and P. Perinotti. "Informational derivation of quantum theory". *Phys. Rev. A* **84** 1 (2011), p. 012311.

[128] G. Chiribella, G. M. D'Ariano, and P. Perinotti. "Quantum from principles". *arXiv:1506.00398 [math-ph, physics:quant-ph]* (2015). arXiv: 1506.00398.

[129] L. Masanes and M. P. Müller. "A derivation of quantum theory from physical requirements". *New J. Phys.* **13** 6 (2011), p. 063001.

[130] B. Dakic and C. Brukner. "Quantum Theory and Beyond: Is Entanglement Special?" *arXiv:0911.0695 [quant-ph]* (2009). arXiv: 0911.0695.

[131] G. Brassard, H. Buhrman, N. Linden, A. A. Méthot, A. Tapp, and F. Unger. "Limit on Nonlocality in Any World in Which Communication Complexity Is Not Trivial". *Phys. Rev. Lett.* **96** 25 (2006), p. 250401.

[132] N. Linden, S. Popescu, A. J. Short, and A. Winter. "Quantum Nonlocality and Beyond: Limits from Nonlocal Computation". *Phys. Rev. Lett.* **99** 18 (2007), p. 180502.

[133] M. Pawłowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Żukowski. "Information causality as a physical principle". *Nature* **461** 7267 (2009), pp. 1101–1104.

[134] M. Navascués and H. Wunderlich. "A glance beyond the quantum model". *Proc. R. Soc. A* **466** 2115 (2010), pp. 881–890.

[135] T. Fritz, A. B. Sainz, R. Augusiak, J. B. Brask, R. Chaves, A. Leverrier, and A. Acín. "Local orthogonality as a multipartite principle for quantum correlations". *Nat Commun* **4** (2013).

[136] M. Navascués, Y. Guryanova, M. J. Hoban, and A. Acín. "Almost quantum correlations". *Nat Commun* **6** (2015).

[137] J.-D. Bancal, C. Branciard, N. Brunner, N. Gisin, S. Popescu, and C. Simon. "Testing a Bell inequality in multipair scenarios". *Phys. Rev. A* **78** 6 (2008), p. 062110.

[138] S. Pironio. "Lifting Bell inequalities". *Journal of Mathematical Physics* **46** 6 (2005), p. 062112.

[139] E. W. Weisstein. *Gamma function.*

[140] B. S. Tsirel'son. "Quantum analogues of the Bell inequalities. The case of two spatially separated domains". *J Math Sci* **36** 4 (1987), pp. 557–570.

[141] L. J. Landau. "Empirical two-point correlation functions". *Found Phys* **18** 4 (1988), pp. 449–460.

[142] L. Masanes. "Necessary and sufficient condition for quantum-generated correlations". *arXiv:quant-ph/0309137* (2003). arXiv: quant-ph/0309137.

[143] V. Scarani. "Limiting the complexity of quantum states: a toy theory". *arXiv:1503.08545 [quant-ph]* (2015). arXiv: 1503.08545.