Nonlocality of symmetric states and its applications in quantum information

Zizhu Wang

ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY

TELECOM ParisTech

2013-ENST-0015

EDITE - ED 130

Doctorat ParisTech

T H È S E

pour obtenir le grade de docteur délivré par

TELECOM ParisTech

Spécialité « Informatique et Réseaux »

présentée et soutenue publiquement par

Zizhu WANG

le 28 mars 2013

Non-localité des Etats Symétriques et ses Applications en Informatique Quantique

Directeur de thèse : **Gérard COHEN**
Co-encadrement de la thèse : **Damian MARKHAM**

Jury

M. Samson ABRAMSKY, Professeur, University of Oxford	Rapporteur
M. Gilles DOWEK, Directeur de recherches, INRIA	Rapporteur
M. Pablo ARRIGHI, Maître de conférences HDR, Université de Grenoble	Examinateur
M. Daniel BROWNE, Maître de conférences, University College London	Examinateur
M. Terence RUDOLPH, Professeur, Imperial College London	Examinateur
M. Gilles ZÉMOR, Professeur, Université Bordeaux 1	Examinateur
M. Gérard COHEN, Professeur, Télécom ParisTech	Directeur de thèse
M. Damian MARKHAM, Chargé de recherche, CNRS LTCI	Directeur de thèse

TELECOM ParisTech
école de l'Institut Mines-Télécom - membre de ParisTech

T H È S E

Acknowledgements

First I would like to thank my supervisors, Damian Markham and Gerard Cohen, for providing support on this project. A very special thanks to Damian Markham, who not only showed me the English (i.e. relaxed, informal, cordial, lazy, etc.) way to do research, but also become like a friend and mentor to me. When I started this thesis, I was a hard core computer scientist and Damian was a hard core physicist. I think we mutually adapted in these 3 years, and now I can proudly call myself a fake physicist. I know sometimes I'm lazy and annoying, thanks for not kicking my ass!

Life in the quantum group is, for the lack of a better word, nonchalant. So I would like to thank everyone in our group, Eleni Diamanti, Romain Alleaume, Isabelle Zaquine, Tom Lawson, Marc Kaplan, Rupesh Kumar, Anne Marin, Anna Pappa, Hao Qin, Adel Sohbi, Paul Jouguet and Sebastien Kunz-Jacques for all the memorable moments of group life from the quantum lunches to the Christmas dinners.

Stimulating discussions with Pablo Arrighi and Elham Kashefi are often very beneficial. I would also like to give Pablo Arrighi a special thanks for being the first person to entrust me with a research problem.

I would also like to thank Samson Abramsky, Gilles Dowek, Dan Browne, Gilles Zemor and Terry Rudolph for agreeing to be part of my jury. I also want to thank Terry for a wonderful summer at Imperial. From these three months I learned invaluable lessons from him on how to be a good scientist.

A special thanks to Feng Yan and Ping Yan, for the countless afternoon chats about academia, the past and the future and the countless dinners together.

Also, I wish to give my most sincere thanks to my parents. As part of the generation who lived through the cultural revolution, they value education above all else. If it weren't for their dedication and hard work I would not have had the chance to come to France and become a researcher.

Lastly, I wish to thank my better half, Wei, for her continued support and encouragement. We have been through the toughest time of our lives together, and the future can only be brighter for us and our family!

Abstract

This thesis is about the nonlocal properties of permutation symmetric states and the potential usefulness of such properties in quantum information processing. The nonlocality of almost all symmetric states, except Dicke states, is shown by constructing an n-party Hardy paradox. With the help of the Majorana representation, suitable measurement settings can be chosen for these symmetric states which satisfy the paradox. An extended CH inequality can be derived from the probabilistic conditions of the paradox. This inequality is shown to be violated by all symmetric states. The nonlocality properties and entanglement properties of symmetric states are also discussed and compared, notably with respect to persistency and monogamy. It is shown that the degeneracy of some symmetric states is linked to the persistency, which provides a way to use device independent tests to separate nonlocality classes. It is also shown that the inequalities used to show the nonlocality of all symmetric states are not strictly monogamous. A new inequality for Dicke states is shown to be monogamous when the number of parties goes to infinity. But all these inequalities can not detect genuine nonlocality. Applications of nonlocality to communication complexity and Bayesian game theory are also discussed.

Resume

Le sujet de cette these est sur les proprietes non-locales des etats symetriques invariant sous les permutations des systemes et les usages potentiels de ces etats dans le domaine de traitement d'information quantique. La non-localite de presque tous les etats symetriques, hors les etats de Dicke, est etablie par une version etendue du paradoxe de Hardy. Grace a la representation de Majorana pour les etats symetriques, des parametres de mesure avec lesquels toutes les conditions du paradoxe sont satisfaites peuvent etre trouves. Une version etendue de l'inegalite de CH peut etre derivee a partir des conditions probabilistes de ce paradoxe. Cette inegalite est violee par tous les etats symetriques. Les proprietes de la non-localite et les proprietes de l'intrication sont aussi discutees et comparees, notamment par rapport a la persistance et la monogamie. Des resultats indiquent que la degenerescence de certains etats symetriques est liee a la persistance, qui donne une façon d'inventer des tests qui sont independants des dispositifs vise pour separer les differentes classes de non-localite. Il est aussi montre que l'inegalite utilisee pour demontrer la non-localite des tous les etats symetriques n'est pas monogame dans le sens strict. Neanmoins, une nouvelle inegalite pour les etats de Dicke est proposee, qui est monogame quand le nombre de participants tends vers l'infinite. Malheureusement, toutes ces inegalites sont incapables de detecter la non-localite authentique. Des applications de la non-localite a la complexite de communication et aux jeux bayesiens sont discutees.

Contents

Acknowledgements i

Abstract iii

Résumé v

Publications xii

0 Sommaire de la these 1
- 0.1 Resume des chapitres 1
- 0.2 La non-localite des etats symetriques 5
- 0.3 La non-localite et les classes d'intrication 19
- 0.4 Une inegalite monogame pour les etats de Dicke 23

1 Introduction 27

2 Background 30
- 2.1 Essential Algebra 30
- 2.2 Postulates of Quantum Mechanics 36
- 2.3 Scratching the Surface of Quantum Entanglement 40
- 2.4 The Facets of Nonlocality 45
  - 2.4.1 A Little Bit of History 45
  - 2.4.2 Bell's Inequality 47
  - 2.4.3 Nonlocality from Correlations 48
  - 2.4.4 Probability-free Nonlocality & Mermin Inequality 51
  - 2.4.5 Almost Probability-free Nonlocality & Hardy Paradox 55
  - 2.4.6 Unification of Different Approaches to Nonlocality 57
- 2.5 Semidefinite Programming 60

3 The Majorana Representation of Symmetric States 62
3.1 Geometry of Complex Numbers 63
3.1.1 The Complex Plane and the Riemann Sphere 63
3.1.2 The Möbius Transformation 65
3.2 From Complex Geometry to the Majorana Representation 72
3.3 Physical Interpretations 76
3.4 The Majorana Representation and Entanglement 79

4 Nonlocality of Symmetric States 83
4.1 Bipartite Hardy Paradox Revisited 83
4.2 Multipartite Hardy Paradox and the Inequality P^n 88
4.3 Violation of P^n By Almost All Symmetric States 91
4.4 Violation of P^n By All Symmetric States 96

5 Degeneracy and its Consequences 101
5.1 Degeneracy and Persistency of Nonlocality 102
5.2 Device Independent Classification of States 104

6 Analysis of Nonlocal Properties for Symmetric States 109
6.1 Large n Results for $|W_n\rangle$ and $|GHZ_n\rangle$ 109
6.2 Monogamy of Entanglement and Monogamy of Correlations 112
6.3 Monogamy and Genuine Nonlocality of Dicke States 116

7 Applications 120
7.1 Application to Communication Complexity 120
7.2 Application to Bayesian Games 124

8 Summary 127
8.1 New Results in This Thesis 127
8.2 Recent Progress on Related Topics 128
8.3 Outlooks 128

Bibliography 129

List of Figures

1 L'etat tetraedre (a) et l'etat $|000+\rangle$ (b) dans la representation de Majorana. 20

2 L'etat tetraedre (a) et l'etat GHZ a 4 qubits $|GHZ_4\rangle$ (b) dans la representation de Majorana. 21

3 Comparaison de la violation de L (75) par $|S(n, \frac{n}{2})\rangle$ (n) et $|W_n\rangle$ (l) en fonction de n (le nombre de parties). 24

3.1 The complex plane with the number a + bi. 63

3.2 The stereographic projection of the complex plane to the unit sphere. . . 64

3.3 The Riemann sphere with spherical coorinates □ and □. 65

3.4 On the complex plane, a point P inside the unit circle (blue) gets inverted to P' outside the unit circle (red). 68

3.5 Effects of complex inversions on generalized circles. 70

3.6 The Bloch sphere. 72

3.7 The Tetrahedron. $|T\rangle = \sqrt{\frac{1}{3}}|S(4,0)\rangle + \sqrt{\frac{2}{3}}|S(4,3)\rangle$. 80

3.8 The Octahedron. $|O\rangle = \frac{1}{\sqrt{2}}(|S(6,1)\rangle + |S(6,5)\rangle)$. 81

3.9 The Cube. $|C\rangle = \frac{1}{2\sqrt{6}}(\sqrt{5}|S(8,0)\rangle + \sqrt{14}|S(8,4)\rangle + \sqrt{5}|S(8,8)\rangle)$. . . . 81

3.10 The Icosahedron. $|I\rangle = \frac{1}{5}(\sqrt{7}|S(12,1)\rangle$ □ $\sqrt{11}|S(12,6)\rangle$ □ $\sqrt{7}|S(12,11)\rangle)$. 81

3.11 The Dodecahedron. $|D\rangle = \frac{1}{25\sqrt{3}}(\sqrt{187}|S(20,0)\rangle + \sqrt{627}|S(20,5)\rangle + \sqrt{247}|S(20,10)\rangle$ □ $\sqrt{627}|S(20,15)\rangle + \sqrt{187}|S(20,20)\rangle)$. 82

4.1 The three symmetric Bell states in the Majorana representation. 87

5.1 The tetrahedron state (a) and the state $|000+\rangle$ (b) in the Majorana representation. 106

5.2 The tetrahedron state (a) and the 4-qubit GHZ state (b) in the Majorana representation. 107

6.1 Violations of P^n by the state $|GHZ_n\rangle$ (n), with the numerical violations of P^n (s) and Q^n_{n-1} (t) of $|W_n\rangle$ as a function of n (number of parties), comparing to $\frac{1}{2^{Eg(|W_n\rangle)}}$ (l) and $\frac{1}{2^{Eg(|GHZ_n\rangle)}}$(u). 111

6.2 A comparison of the violation of L (6.14) for the states $|S(n, \frac{n}{2})\rangle$ (n) and $|W_n\rangle$ (l) as a function of n (the number of parties). 117

List of Tables

1 Les bornes de violations maximales de P^4 et Q_3^4 pour $|T_i$ et $|000+i$. Une violation de Q_3^4 implique que l'etat n'est pas dans la classe LUP de $|T_i$. . 20

2 Les bornes de violations maximales de P^4 et Q_3^4 pour $|T_i$ et $|GHZ_4i$. Une violation de $P^4 > 0.1241$ implique que l'etat n'est pas dans la classe LUP de $|GHZ_4i$, et une violation de Q_3^4 implique l'etat n'est pas dans la classe LUP de $|T_i$. 21

3.1 The Platonic solids. 80

4.1 Translation between the two different notations for bipartite Hardy paradox. 84

5.1 SDP bounds on the maximum violation of P^4 and Q_3^4 for $|T_i$ and $|000+i$. Because of computational difficulties, the values for $|000+i$ assume that all parties measure in the same basis (a numerical optimization over the four Euler angles in the two measurement settings indicate this is still optimal). We thus have that a violation of Q_3^4 implies the state is not in the LU class of $|T_i$. 106

5.2 SDP bounds on the maximum violation of P^4 and Q_3^4 for $|T_i$ and $|GHZ_4i$. We thus have that a violation of $P^4 > 0.1241$ implies the state is not in the LU class of $|GHZ_4i$, and a violation of Q_3^4 implies the state is not in the LU class of $|T_i$. 107

List of Publications

Z. Wang, D. Markham. Nonlocality of Symmetric States. Physical Review Letters, 108, 210407 (2012). preprint: arXiv:1112.3695.

(contains material from Chapter 3, 4 and 5)

Z. Wang, D. Markham. Nonlocality and entanglement for symmetric states. Physical Review A, 87, 012104 (2013). preprint: arXiv:1210.1754.

(contains material from Chapter 3, 5 and 6)

Chapitre 0

Sommaire de la thèse

0.1 Résumé des chapitres

L'objectif de cette thèse, indiqué par son titre, est d'examiner les caractéristiques non-locales des états symétriques afin de mesurer l'utilité de ces états dans le traitement de l'information quantique. Le sujet est abordé d'une manière pédagogique, avec un pré-requis minimum d'informatique quantique et de fondement de la mécanique quantique.

Le chapitre 2 donne tous les information nécessaires pour comprendre la mécanique quantique de base et la notion de non-localité. La section 2.1 est un rappel de l'algèbre linéaire essentielle pour la compréhension des structures linéaires utilisées dans la mécanique quantique : l'espace de Hilbert, l'opérateur hermitique, etc. Les postulats de la mécanique quantique sont introduits dans la section 2.2. Les différentes notions mathématiques de la section précédente acquièrent alors des sens physiques. Les postulats présentés ici sont des variations des postulats courants de la communauté informatique quantique, avec notamment l'inclusion de POVM dans le postulat de mesure. La section prochaine donne un rappel de la théorie d'intrication. Sont introduits la notion de mesure d'intrication et le régime de l'opération locale et la communication classique. Les axiomes dont les différentes mesures d'intrication satisfont souvent sont aussi énoncés. Un rappel du développement historique de non-localité, dont l'origine date de l'enfance de la mécanique quantique, est donné par la section 2.4. Après cette introduction historique, les trois façons courantes pour montrer la non-localité sont expliquées : les fonctions de corrélation, l'inégalité de MABK avec les corrélations parfaites et le paradoxe de Hardy avec l'inégalité de CH. Cette section se termine avec une discussion de quelques résultats récents qui tentent de réunir les différentes façons de montrer la non-localité : soit par la géométrie convexe, soit par la théorie des catégories ou par les hyper-graphes. La dernière section de ce chapitre donne une introduction à la technique numérique utilisée pour obtenir les résultats numériques dans les chapitres prochains : la programmation semi-définie.

L'aspect géométrique des états symétriques est abordé dans le chapitre 3, sous la forme de la représentation de Majorana. La section 3.1 montre le lien intime entre les nombres complexes et la géométrie. Les nombres complexes, en outre de la représentation géométrique normale sur la plaine complexe, admettent une autre représentation plus compacte sur la surface de la sphère de Riemann. Les automorphismes de cette sphère, connus sous le nom de transformations de Möbius, ont une interprétation directe dans la théorie d'intrication des états symétriques comme les visualisations géométrique des classes SLOCC d'intrication. La section 3.2 présente l'outil principal pour l'étude de la non-localité et l'intrication des états symétriques : la représentation de Majorana. La représentation géométrique d'un état symétrique est présentée avec la procédure constructive de décomposer un état symétrique en ses points de Majorana et de construire un état symétrique à partir d'un ensemble de points de Majorana. Quelques propriétés intéressantes sont aussi dérivées à partir de ces procédures de décomposition/ reconstruction. Ces propriétés seront utilisées dans les chapitres ultérieurs pour démontrer la non-localité des états symétriques. La section 3.3 est un rappel de l'aspect physique de la représentation de Majorana et ses significations historiques, notamment le lien entre la géométrie complexe et le spin physique. Le lien entre le spin et la relativité restreinte n'est pas démontré explicitement, néanmoins quelque références sont données [PR84, PR86]. La dernière section de ce chapitre est une revue de résultats principaux d'une thèse précédente sur l'intrication d'états symétriques [Aul11a]. Le lien entre les transformations de Möbius et les classes SLOCC d'intrication est donné d'une façon explicite. Les états symétriques dont les points de Majorana forment les solides de Platon ont une signification particulière grâce à la dualité entre le solide de Platon formé par les points de Majorana et le solide de Platon formé par les points plus proches en terme de la mesure géométrique d'intrication.

Le chapitre 4 contient principalement les résultats techniques. La première section est un rappel du paradoxe de Hardy en deux parties, avec la procédure pour retrouver les bases de mesure qui satisfont les conditions du paradoxe. La prochaine section donne l'extension du paradoxe de deux parties en n parties, ainsi que l'inégalité étendue, qui est appelée P^n dans cette thèse. La section 4.3 donne la preuve du théorème principal qui montre que tous les états symétriques sauf les états de Dicke satisfont le paradoxe pour n parties et par conséquent violent P^n. Une partie de cette preuve montre comment choisir les bases de mesure qui satisfont le paradoxe étendu. La dernière section donne la preuve de l'existence des bases de mesure qui permettent les états de Dicke de violer P^n, même si ils ne satisfont pas le paradoxe de Hardy étendu.

Le chapitre 5 porte sur un aspect spécial des états symétriques, qui se manifeste sous la représentation de Majorana : la dégénérescence. La dégénérescence dénote le phénomène que plusieurs points de Majorana d'un état coïncident. Il vient des racines dégénérées d'un polynôme. En utilisant la dégénérescence, il est possible de montrer

que l'intrication et la non-localité persistent dans les sous-systèmes. Cette persistance est mis en évidence par l'addition de plusieurs conditions au paradoxe de Hardy et à l'inégalité P^n pour donner une nouvelle inégalité Q_d^n. Ces conditions peuvent être satisfaites parfaitement seulement par un état correctement dégénéré. Ainsi, ayant le correct degré de dégénérescence garantie la violation de Q_d^n. L'utilisation de la programmation semi-définie permet de classifier les états symétriques en classes non-locales sous régimes d'opérations locales ou permutations des systèmes (LUP). Les états ainsi classifiés appartiennent aussi aux différentes classes SLOCC, grâce au lien entre les opérations SLOCC et les transformations de Möbius.

Le chapitre 6 donne davantage d'analyse numérique des propriétés d'intrication et de la non-localité des états symétriques. La première section analyse la violation de P^n par les états de GHZ et les états W quand le nombre de parties augmente. La violation est toujours bornée par la mesure géométrique d'intrication à cause de la structure de P^n. Malheureusement, à cause de cette limite, P^n et Q_d^n ne sont pas monogames. La notion de monogamie, définie dans les contextes d'intrication et de non-localité, est abordée en section 6.2. Bien que P^n n'est pas monogame à proprement parler, il existe quand même de la monogamie dans un sens élargi. La dernière section présente une nouvelle inégalité pour les états de Dicke, inspirée par les résultats de Heaney, Cabello, Santos et Vedral. Cette inégalité est strictement monogame lorsque le nombre de participants tend vers l'infini. Néanmoins, ni cette nouvelle inégalité ni P^n (ou bien Q_d^n) ne peut détecter la non-localité authentique.

Le chapitre 7 est dédié aux deux applications de résultats obtenus dans les chapitres précédents : la complexité de communication et la théorie des jeux bayésiens. Il est connu depuis plus de dix ans que la non-localité a des applications en la complexité de communication. Dans le cas le plus extrême, avec une boîte de PR (qui représente des corrélations plus fortes que permis par la mécanique quantique), la complexité de communication pour les fonctions booléennes est rendue triviale. Cette section sur la complexité de communication commence par un rappel du modèle pour deux parties, et comment réduire la complexité par des données aléatoires, privées ou partagées. Ensuite un avantage quantique est donné pour certains fonctions qui n'ont pas d'avantage classique avec des données aléatoires classiques. Cependant, plus de travaill est nécessaire pour développer un modèle multi-partie avec relations au lieu de fonctions, parce que les conditions du paradoxe de Hardy étendu forment une relation. L'autre application abordée dans ce chapitre est les jeux bayésiens, qui représentent les jeux avec des informations incomplètes. Des procédures de transformations d'une inégalité de Bell en un jeux bayésien existent déjà. Le paradoxe de Hardy a été déjà traduit en un jeux bayésien avec une matrice de gains explicite. Un résultat récent de Brunner et Linden [BL12] donne un lien plus profond entre les jeux bayésiens et la non-localité.

Le dernier chapitre résume les nouveaux résultats dans cette thèse, signale quelque

nouveaux developpements dans les sujets similaires. Il envisage aussi des nouvelles directions de recherche dans le futur et pose des questions ouvertes.

0.2 La non-localité des états symétriques

La non-localite est une caracteristique fondamentale de la mecanique quantique qui est en train d'être reconnue comme une ressource cle dans la theorie du traitement de l'information quantique, ayant des applications dans la securite dispositif-independant [BHK05, Col06, PAM+10, MPA11], la complexite de communication [BCMdW10] et le calcul quantique par mesure [RB01, AB09a]. Une autre caracteristique liee mais differente est la notion de l'intrication. Dans un cadre multi-partie, l'intrication se manifeste d'une façon compliquee : il existe differentes classes d'intrication , chacune joue un rôle potentiellement different comme une ressource. Actuellement très peu est connu si, et comment, les richesses de l'intrication multi-partie sont reflechis dans les caracteristiques non-locales.

Les caracteristiques non-locales des etats symetrique de qubits seront explorees. Ces etats forment un ensemble utile pour diverses taches de traitement d'information quantique grâce à ses occurrences naturelles comme des etats fondamentaux dans certains modèles de Bose-Hubbard. Ils sont aussi parmi les etats plus accessibles experimentalement. Jusqu'à present, relativement peu est connu sur la non-localite de ces etats, avec des resultats concernant majoritairement les etats de W ou GHZ [Mer95, Cab02, HCSV11].

Une comprehension approfondie de la non-localite de ces etats nous permet non seulement de mieux realiser leur potentiels comme des ressources d'informatique quantique, mais aussi de mieux comprendre la relation subtile entre la non-localite et l'intrication d'etats multi-parties. Recemment, l'intrication d'etats symetriques est etudie à l'aide d'un outil mathematique : la representation de Majorana [Maj32, BKM+09, Aul11b, RM11, Mar11, AMM10]. Ici le même outil sera utilise pour etudier la non-localite de ces etats, qui permet ainsi une comparaison facile.

Commençons par n parties, indexee par i, chacune choisit une position de mesure M_i et obtient un resultat r_i. Il n'existe que deux positions possibles, avec deux resultats possibles par position. Une loi de probabilite est locale ou admet un modèle à variables cachées locales (LHV, local hidden variable model) si la loi de probabilite à plusieurs variables peut être ecrit comme le produit de probabilites individuelles, etant donnee la valeur d'une variable cachee λ :

$$P(r_1,\ldots,r_n|M_1,\ldots,M_n) = \int \rho(\lambda) \prod_i P(r_i|M_i,\lambda)\, d\lambda, \qquad (1)$$

où $P(r_i|M_i,\lambda)$ est la probabilite pour la partie i d'obtenir le resultat r_i à la position M_i quand la valeur de la variable cachee est λ. $\rho(\lambda)$ est la loi de probabilite de λ. $P(r_1,\ldots,r_n|M_1,\ldots,M_n)$ est la loi de probabilite pour toutes les parties n quand ils mesurent aux position $M_1,\ldots,M_n$ et obtiennent les resultats $r_1,\ldots,r_n$. Les souscrits seront ignores quand leur sens est clair. Il est evident que tous les etats-produits admettent

un LHV par des mesures locales. Néanmoins, la non-localité ne suit pas directement de l'intrication [HHHH09].

Le paradoxe de Hardy a été proposé comme un test «quasiment sans probabilité» de la non-localité de presque tous les états intriqués de deux parties [Har93, Har94]. Les sections suivantes servent à montrer que tous les états symétrique de n parties peuvent satisfaire les conditions du paradoxe de Hardy étendu et l'inégalité associée. Bien qu'il existe des résultats précédents en généralisant le paradoxe de Hardy à n parties [GR10], une procédure constructive est donnée ici pour calculer les paramètres de mesure.

Le paradoxe original de Hardy est constitué de quatre conditions probabilistes qui sont imposées sur les résultats d'une expérience avec deux parties [Har93, Har94]. Ces conditions sont compatibles individuellement avec la définition de LHV donnée par (33). Mais une contradiction logique est obtenue si les quatre conditions sont imposées simultanément. Il est montré par Hardy que pour tous les états intriqués de deux parties, il existe des positions de mesures qui satisfont toutes les conditions, démontrent ainsi l'incompatibilité du modèle LHV et la mécanique quantique. La seule exception est les états maximalement intriqués. Heureusement, la non-localité des états maximalement intriqués est connue auparavant [Bel64, Han98].

L'extension du paradoxe de Hardy à plusieurs parties peut être construit comme suit : d'abord supposons qu'il y a n parties qui participent à l'expérience. Chaque partie peut choisir librement sa position de mesure parmi les deux choix donnés, qui sont notés par 0 et 1. La partie obtient un résultat après la mesure, qui est aussi parmi les deux résultats possibles, 0 ou 1. La première condition qu'on s'impose est que quand tout le monde choisit la position 0, alors il est possible qu'ils aient tous obtenus le résultat 0 :

$$P(00\ldots00|00\ldots00) > 0. \tag{2}$$

Les n conditions suivantes sont les mêmes que la condition ci-dessus pour $n \square 1$ parties, mais quand la partie n choisit la position 1 au lieu de 0, les n parties ne obtiennent jamais toutes le résultat 0.

$$P(00\ldots00|\square(00\ldots01)) = 0, \tag{3}$$

où le symbole $\square$ signifie la permutation d'une chaîne de bits avec un seul 1 et $n \square 1$ zéros. Le modèle de LHV (33) et la condition (2) implique qu'il existe au moins une valeur de la variable cachée $\square$ tel que $\forall i, P(0_i|0_i\square) > 0$. Alors, pour cette valeur de $\square$, on peut déduire qu'à partir de (3), $\forall i, P(0_i|1_i\square) = 0$. Comme il n'existe que deux résultats possible par position de mesure, (3) implique que pour cette valeur de $\square$, si tout le monde choisissait la position 1, ils obtiendraient tous le résultat 1 avec certitude.

La dernière condition imposée est une contradiction à la conclusion ci-dessus. Si tout

le monde mesurait la position 1, alors ils ne obtiendraient jamais tous le resultat 1 :

$$P(11\ldots1|11\ldots11) = 0. \tag{4}$$

Clairement (2), (3) et (4) sont incompatibles avec LHV. Dans le cas où $n = 2$, on retrouve le paradoxe de Hardy original [Har93, Har94].

Neanmoins, une procedure constructive sera expliquee qui permet de calculer les positions 0 et 1 pour presque tous les etats symetriques telles que les conditions (2) a (4) sont toutes satisfaites.

Un etat symetrique de qubits invariant sous la permutation peut etre decompose comme $|\psi\rangle = \sum_{k=0}^{n} c_k |S(n,k)\rangle$, où les $|S(n,k)\rangle = \binom{n}{k}^{-\frac{1}{2}} \sum_{perm} |\underbrace{0\ldots0}_{n-k}\underbrace{1\ldots1}_{k}\rangle$ sont les etats de Dicke.

Dans la representation de Majorana, l'etat $|\psi\rangle$ est ecrit comme une somme de permutations des produits tensoriels de n qubits $\{|\alpha_1\rangle\ldots|\alpha_n\rangle\}$, qui sont appeles les points de Majorana (MPs) de l'etat $|\psi\rangle$:

$$|\psi\rangle = K \sum_{perm} |\alpha_1\ldots\alpha_n\rangle. \tag{5}$$

K est une constante de normalisation qui depend de distances entre les MPs. Dans cette representation, les operations unitaires locales de la forme $U^{\otimes n}$ tournent tous les points de Majorana simultanement, et sont donc equivalentes a une rotation de la sphere de Bloch.

L'invariance sous les permutations persiste aussi dans les sous-espaces. Soit $|\psi\rangle$ un etat symetrique de n qubits, alors pour un qubit arbitraire $|\alpha\rangle$ (non-normalise), l'etat de $n-1$ qubits $\langle\alpha|\psi\rangle$ est aussi invariant sous les permutations :

$$\langle\alpha|\psi\rangle = \sum_{i=1}^{n} C_i \sum_{perm} |\underbrace{\alpha_1\ldots\alpha_n}_{\{1,\ldots,n\}-i}\rangle, \tag{6}$$

où $C_i = \langle\alpha|\alpha_i\rangle$ et $\{1,\ldots,n\}-i$ signifie que le MP $|\alpha_i\rangle$ est jete.

L'equation suivante est vraie si et seulement si $|\alpha_i\rangle$ est un MP de $|\psi\rangle$, en plus le point $|\alpha_i^{\perp}\rangle$ est son point diametralement oppose sur la sphere de Bloch :

$$(\langle\alpha_i^{\perp}|)^{\otimes n}|\psi\rangle = 0. \tag{7}$$

Les conditions (2) a (4) peuvent etre satisfaites par presque tous les etats symetriques pour les positions de mesure obtenues par la procedure suivante. D'abord, (7) peut etre interprete comme l'amplitude de probabilite qui donne la condition (4), si les mesures sont projectives et la position 1 pour toutes les parties consistent en $\{|\alpha_i\rangle, |\alpha_i^{\perp}\rangle\}$.

Pour les conditions (3), si une d'entre elles est satisfaite alors par l'invariance de

l'etat toutes les autres sont satisfaites automatiquement. La projection de l'etat $|\psi\rangle$ sur l'un de ses MPs $|\phi_i\rangle$ donne un nouveau etat symetrique de $n-1$ qubits :

$$|\psi'\rangle = \langle\phi_i|\psi\rangle = \sum_{j=1}^{n} C_j \sum_{perm} |\underbrace{\phi_1 \ldots \phi_n}_{\{1,\ldots,n\}-\{j\}}\rangle, \tag{8}$$

avec $C_j = \langle\phi_i|\phi_j\rangle$.

L'etat $|\psi'\rangle$ se decompose en $n-1$ points de Majorana, probablement differents de MPs de $|\psi\rangle$. En fait, le theoreme ci-dessous montre qu'il existe toujours au moins un MP de $|\psi'\rangle$ qui est different de tous les MPs de $|\psi\rangle$.

Théorème 1. Soit $S := \{|\phi_1\rangle, |\phi_2\rangle, \ldots, |\phi_n\rangle\}$ l'ensemble de MPs de l'état $|\psi\rangle$. Soit $S_i := \{|\phi_1\rangle, |\phi_2\rangle, \ldots, |\phi_{n-1}\rangle\}$ l'ensemble de MPs de l'état $|\psi_i\rangle = \langle\phi_i|\psi\rangle$. Alors $S_i \subset S$ ssi $|\psi\rangle$ est un état de Dicke ou un état équivalent par rotations de la sphere de Bloch.

Démonstration. D'abord nous allons demontrer un lemme important pour le reste de cette demonstration. Le but de ce lemme est de montrer que les conditions d'orthogonalite telles que (7) et (72) sont satisfaites seulement si le produit tensoriel (la partie bra) est compose d'un point Majorana de l'etat symetrique dans la partie ket, et le degre du produit tensoriel est borne par le degre de degenerescence de ce MP.

Lemme 2. Soit $|\psi\rangle$ un état symétrique de n qubits avec MPs (distincts) $\{|\phi_1\rangle, |\phi_2\rangle, \ldots, |\phi_l\rangle\}$ ayant degrés de dégénérescence $\{d_1, d_2, \ldots, d_l\}$, alors

$$(\langle\phi|)^{\otimes c}|\psi\rangle = 0 \tag{9}$$

si et seulement si $|\phi\rangle = |\phi_i^{\perp}\rangle$, $\langle\phi_i^{\perp}|\phi_i\rangle = 0$ pour un certain i, et $c \geq n - d_i + 1$ (ou de façon équivalente, $d_i \geq n - c + 1$).

Démonstration. La direction si se suit directement de l'expansion de $|\psi\rangle$ en utilisant (5),(71) et la condition sur les MPs dans la definition de la representation de Majorana. Nous allons nous concentrer sur la direction seulement si

D'abord notons que $(\langle\phi|)^{\otimes c}|\psi\rangle = 0$ avec $c \geq n$, alors $(\langle\phi|)^{\otimes n}|\psi\rangle = 0$. L'explication ci dessus implique $(\langle\phi|)^{\otimes n}|\psi\rangle = 0$ n'est possible que si $|\phi\rangle$ est un point diametralement oppose a un MP de $|\psi\rangle$. Donc $|\phi\rangle = |\phi_i^{\perp}\rangle$, $\langle\phi_i^{\perp}|\phi_i\rangle = 0$ pour un certain i.

Maintenant nous aimerons montrer que $(\langle\phi_i^{\perp}|)^{\otimes c}|\psi\rangle = 0$ implique $c \geq n - d_i + 1$. Mais nous allons montrer un enonce equivalent : $c < n - d_i + 1$ (ou de façon equivalente $c \leq n - d_i$) implique $(\langle\phi_i^{\perp}|)^{\otimes c}|\psi\rangle \neq 0$.

Si $c = n - d_i$, alors

$$(h\Pi_i^{\perp}|)^{\otimes n-d_i}|\psi\rangle \tag{10}$$

$$= (n-d_i)!\,(h\Pi_i^{\perp}|)^{\otimes n-d_i}|\underbrace{1\ldots1}_{\{1\ldots n\}\setminus i}\rangle\,(|\Pi_i\rangle)^{\otimes d_i} \neq 0. \tag{11}$$

Pour aller de (10) à (11), nous avons utilisé le fait que quand $c = n - d_i$, si nous développons $|\psi\rangle$ avec (5),(71) tous les autres termes disparaissent. Ce terme n'est pas zéro car par hypothèse tous les autres MPs sont différents de $|\Pi_i\rangle$.

Pour le cas $c < n - d_i$, si $(h\Pi_i^{\perp}|)^{\otimes c}|\psi\rangle = 0$, alors il implique $(h\Pi_i^{\perp}|)^{\otimes n-d_i}|\psi\rangle = 0$, qui est en contradiction avec l'argument ci-dessus. Donc quand $c - n - d_i$, $(h\Pi_i^{\perp}|)^{\otimes c}|\psi\rangle \neq 0$. □

L'énoncé de ce théorème fait référence à l'état $|\psi_1\rangle = h\Pi_1|\psi\rangle$. Ni l'énoncé ni cette démonstration ne dépendent du choix de $|\Pi_1\rangle$. En fait nous allons prouver que pour n'importe quel $|\Pi_i\rangle$, pour satisfaire l'énoncé du théorème tous les autres MPs de $|\psi\rangle$ sont soient orthogonal à lui soit coïncident avec lui, ce qui donne un état de Dicke (potentiellement tourné). Pour la simplicité, nous allons fixer $|\Pi_i\rangle$ à $|\Pi_1\rangle$, donc $|\psi_1\rangle$ est fixé à $|\psi_1\rangle$.

L'état $|\psi\rangle$ se décompose en MPs $|\Pi_i^{\perp}\rangle$ comme dans (71). De la même façon, $|\psi_1\rangle = h\Pi_1|\psi\rangle$ se décompose en

$$|\psi_1\rangle = K\sum_{\text{perm}}|\Pi_1^{m_1}\Pi_2^{m_2}\ldots\Pi_k^{m_k}\rangle, \tag{12}$$

$$\forall i \neq j, |\Pi_i\rangle \neq |\Pi_j\rangle, \sum_{i=1}^{k} m_i = n - 1. \tag{13}$$

En plus de $|\psi_1\rangle$, il y a un autre état symétrique de $(n-1)$ qubits qui nous intéresse : l'état qui se compose de tous les MPs de $|\psi\rangle$ sauf $|\Pi_i\rangle$.

$$|\psi_{\neq i}\rangle := K\sum_{\text{perm}}|\underbrace{\Pi_1\ldots\Pi_1}_{\{1\ldots n\}\setminus i}\rangle. \tag{14}$$

Par le lemme 2, nous allons démontrer deux corollaires concernant les états $|\psi_1\rangle$, $|\psi_{\neq i}\rangle$ et les degrés de dégénérescence de leur MPs :

Corollaire 3. $(h\psi|)^{\otimes c}|\psi_1\rangle = 0$ si et seulement si $|\psi\rangle = |\Pi_i^{\perp}\rangle$, $h\Pi_i^{\perp}|\Pi_i\rangle = 0$ pour un certain i et $c - n - m_i$ (ou de façon équivalente, $m_i - n - c$).

Démonstration. La démonstration suit directement de 2 en notant que $|\psi_1\rangle$ est un état symétrique de $n-1$ qubits. □

Corollaire 4. $(h\Box j)^{\Box c} j\ _{=i}i = 0$ si et seulement si $j\Box i = j\Box_j^? i$, $h\Box_j^? j\Box_j i = 0$ pour un certain j, avec 1). Si $j \neq i$, alors $c \Box n \Box d_j$ (ou de façon équivalente, $d_j \Box n \Box c$) ; 2). Si $j = i$, alors $d_j > 1$ et $c \Box n \Box d_j + 1$ (ou de façon équivalente, $d_j \Box n \Box c + 1$).

Démonstration. Pour le cas $j \neq i$, la demonstration suit du lemme 2, comme le lemme precedent. Quand $j = i$, la degenerescence de $j\Box_j i$ dans la decomposition de $j\ _{=i}i$ est de $d_j \Box 1$, alors par lemme 2 il est zero pour $d_i > 1$ et par consequent $d_j \Box 1 \Box n \Box c$ \$ $d_j \Box n \Box c + 1$. □

Nous allons demontrer les 3 lemme principaux du theoreme. Pour la clarte, nous gardons les notations $j\Box_i i$ est un MP de $j\ i$ avec degre de degenerescence d_i, et $j\Box_i i$ est un MP de $j\ _1 i$ avec degre de degenerescence m_i. Nous allons aussi utiliser K pour denoter une constante globale de normalisation.

Lemme 5. $\forall i, j\Box_i i = j\Box_i i$, avec $m_i \Box d_i \Box 1$.

Démonstration. D'abord, si $d_i = 1$, alors l'enonce est toujours vrai, donc nous nous concentrons sur le cas $d_i > 1$.

D'apres le corollaire 3, $(h\Box_i^? j)^{\Box n \Box d_i + 1} j\ _1 i = 0$ si et seulement si $j\Box_i i = j\Box_i i$ et $m_i \Box n \Box (n \Box d_i + 1) = d_i \Box 1$. Donc il suffit de demontrer que $(h\Box_i^? j)^{\Box n \Box d_i + 1} j\ _1 i = 0$.

$j\ _1 i$ se decompose en

$$j\ _1 i = K \sum_l h\Box_1 j\Box_l i\, j\ _{=l} i\,. \tag{15}$$

Avec cette decomposition, nous obtenons

$$\begin{aligned}
&(h\Box_i^? j)^{\Box n \Box d_i + 1} j\ _1 i && (16)\\
&= K (h\Box_i^? j)^{\Box n \Box d_i + 1} \sum_l h\Box_1 j\Box_l i\, j\ _{=l} i && (17)\\
&= K \sum_l h\Box_1 j\Box_l i\, (h\Box_i^? j)^{\Box n \Box d_i + 1} j\ _{=l} i && (18)\\
&= 0, && (19)
\end{aligned}$$

ou nous avons utilise les deux cas du corollaire 4 pour aller de (18) a (19). □

Lemme 6. $\forall i$, si $j\Box_i i = j\Box_i i$ et $m_i \Box d_i$, alors $h\Box_1 j\Box_i i = 0$.

Démonstration. D'apres le corollaire 3, si $j\Box_i i = j\Box_i i$ et $m_i \Box d_i$, alors $(h\Box_i^? j)^{\Box n \Box d_i} j\ _1 i = 0$.

En utilisant la même décomposition de $|\psi_1\rangle$ que dans la démonstration du lemme précédent, nous avons :

$$(h\Box_i^? j)^{\Box n \Box d_i} |\psi_1\rangle \tag{20}$$
$$= K \sum_j h\Box_1 j \Box_j i \, (h\Box_i^? j)^{\Box n \Box d_i} |\psi_{=j}\rangle . \tag{21}$$

Depuis le corollaire 4, nous pouvons déduire que le seul terme qui ne disparait pas est quand $j = i$, alors nous avons :

$$(h\Box_i^? j)^{\Box n \Box d_i} |\psi_1\rangle \tag{22}$$
$$= K h\Box_1 j \Box_i i \, \underbrace{(h\Box_i^? j)^{\Box n \Box d_i} |\psi_{=i}\rangle}_{\Box} \tag{23}$$
$$= 0. \tag{24}$$

Comme ni K ni $\Box$ dans (23) sont 0, nous pouvons conclure que $h\Box_1 j \Box_i i = 0$.

□

Lemme 7. $\forall i$, si $j\Box_i i = j\Box_i i$ alors $m_i \geq d_i$.

Démonstration. Si $m_i > d_i$, alors d'après lemme 6 $h\Box_1 j \Box_i i = 0$. Nous pouvons noter $j\Box_i^? i$ comme $j\Box_1 i$.

Développons partiellement $|\psi\rangle$ dans la base $\{j\Box_1 i, j\Box_1^? i\}$ donne

$$|\psi\rangle = K(j\Box_1 i \, |\psi_1\rangle + j\Box_1^? i \, |\psi^\perp_1\rangle) \tag{25}$$

Pour $(h\Box_i^? j)^{\Box n \Box d_i} |\psi\rangle$:

$$(h\Box_i^? j)^{\Box n \Box d_i} |\psi\rangle \tag{26}$$
$$= (h\Box_i^? j)^{\Box n \Box d_i} K(j\Box_1 i \, |\psi_1\rangle + j\Box_1^? i \, |\psi^\perp_1\rangle) \tag{27}$$
$$= K(h\Box_i^? j)^{\Box n \Box d_i \Box 1} |\psi_1\rangle \tag{28}$$
$$= 0, \tag{29}$$

où nous avons noté $j\Box_i^? i = j\Box_1 i$ pour aller de (27) à (28). Nous avons aussi utilisé le corollaire 3 et notre hypothèse $m_i > d_i$ pour aller de (28) à (29).

Clairement, la conclusion que $(h\Box_i^? j)^{\Box n \Box d_i} |\psi\rangle = 0$ est en contradiction avec le lemme 2, donc $m_i \geq d_i$. □

En combinant les lemme 5, 6 et 7, nous obtenons le corollaire principal qui va conclure la démonstration.

Corollaire 8. $\forall i$, $|\Box_i\rangle = |\Box_i\rangle$ avec degré de dégénérescence tel que soit 1). $m_i = d_i$, $\langle\Box_1|\Box_i\rangle = 0$ ou 2). $m_i = d_i - 1$, $\langle\Box_1|\Box_i\rangle \neq 0$.

Pour conclure cette demonstration, notons que comme $\sum_k m_k = n - 1$, et $n = \sum d_i$, nous avons

$$\sum_k m_k = n - 1 = \Big(\sum_k d_k\Big) - 1 \tag{30}$$

$$= \underbrace{(d_i - 1)}_{\text{group 1}} + \underbrace{\sum_{j\neq i} d_j}_{\text{group 2}}, \tag{31}$$

ou nous avons separe les d_i en deux groupes, groupe un, avec le -1 associe a un d_i particulier, et groupe deux, avec les d_j qui reste. Qu'implique la condition $S_i = S$ dans l'enonce du theoreme - i.e., ou tous les MPs $|\Box_i\rangle$ coincident avec $|\Box_i\rangle$? Il est clair depuis le corollaire 8 et (30) que ce n'est possible que si le groupe un et donne par d_1, et le groupe deux est donne par un MP qui est l'antipode de $|\Box_1\rangle$. Du coup il n'y a que deux MPs de $|\ \rangle$, $|\Box_1\rangle$ et $|\Box_2\rangle = |\Box_1^{\perp}\rangle$. Cet etat est exactement un etat de Dicke (possiblement tourne). Le raisonnement reste valide si on substitue un autre $|\Box_i\rangle$ pour $|\Box_1\rangle$ au debut.

□

Soit $|\Box_i\rangle$ un MP de l'etat $|\ ^{0}\rangle$, defini dans (8), qui est different de tous les MPs de $|\ \rangle$. Alors a partir de (7) on obtient

$$(\langle\Box_i^{\perp}|)^{\otimes n-1}|\ ^{0}\rangle = \langle\Box_i|\underbrace{\Box_i^{\perp}\ldots\Box_i^{\perp}}_{n-1}|\ \rangle = 0. \tag{32}$$

Par les choix de $\{|\Box_i\rangle, |\Box_i^{\perp}\rangle\}$ comme la position 0 pour toutes les parties, l'amplitude (32) implique la satisfaction de (3) par la symetrie. $\underbrace{\langle\Box_i^{\perp}\ldots\Box_i^{\perp}|\ \rangle}_{n} \neq 0$ parce que $|\Box_i\rangle$ n'est pas un MP de $|\ \rangle$. Par consequent, (2) est satisfait automatiquement. Le theoreme ci-dessus montre que cette procedure s'applique a tous les etats symetriques sauf les etats de Dicke.

Pourtant, le paradoxe ne peut pas etre teste directement a cause de ses conditions exactes, qui ne sont jamais satisfaites quand le bruit et l'inexactitude de l'environnement sont pris en compte. Pour rendre le paradoxe resistant au bruit, une inegalite est necessaire. La bornee obtenue sous LHV est violee par (2) quand la procedure ci-dessus est utilisee.

Théorème 9. Le polynôme de Bell de n systèmes

$$P^n := P(0\ldots0|00\ldots00) \\ \sum_{\square} P(00\ldots00|\square(00\ldots01)) \\ \square \\ \square\, P(1\ldots1|11\ldots11)$$

est borné sous LHV par $P^n \square 0$.

Démonstration. Supposons que

$$P(r_1,\ldots,r_n|M_1,\ldots,M_n) = \int \square(\square) \prod_i P(r_i|M_i,\square)\, d\square, \tag{33}$$

la probabilité jointe est un produit de probabilité de chaque partie (nous allons omettre $\square$ car nous intégrons toujours sur $d\square$). Le polynôme de Bell P^n devient :

$$\begin{aligned} & P_1(0)P_2(0)\ldots P_n(0) \\ \square\, & P_1(1)P_2(1)\ldots P_n(1) \\ \square\, & (1 \square P_1(1))P_2(0)\ldots P_n(0) \\ & \vdots \\ \square\, & P_1(0)\ldots P_{n\square1}(0)(1 \square P_n(1)), \end{aligned} \tag{34}$$

en notant

$$\begin{aligned} P_i(0) &= P_i(0|0), \\ P_i(1) &= P_i(1|1). \end{aligned}$$

Une expansion donne :

$$P_1(0)P_2(0)\ldots P_n(0) \square P_1(1)P_2(1)\ldots P_n(1) \tag{35}$$
$$+ P_1(1)P_2(0)\ldots P_n(0) \square P_2(0)P_3(0)\ldots P_n(0) \tag{36}$$
$$+ P_1(0)P_2(1)\ldots P_n(0) \square P_1(0)P_3(0)\ldots P_n(0) \tag{37}$$
$$\vdots$$
$$+ P_1(0)P_2(0)\ldots P_n(1) \square P_1(0)P_2(0)\ldots P_{n\square1}(0) \tag{38}$$

Notons que les rangs (36) à (38) sont tous $\square 0$, car pour tous $0 \square P_i \square 1$,

$$\prod_{i=1}^{n} P_i \square \prod_{i=1}^{n\square1} P_i. \tag{39}$$

En factorisant le second terme de (35) et le premier terme de (36), nous avons :

$$P_1(1)(P_2(0)\dots P_n(0) \otimes P_2(1)\dots P_n(1)). \tag{40}$$

Si $P_2(0)\dots P_n(0) \otimes P_2(1)\dots P_n(1)$, alors (40) $\otimes$ 0 (par (39) dans le premier terme de (35)et le second terme de (36), mais aussi les rangs (37) à (38)), donc la démonstration est terminée. Sinon

$$\begin{aligned} &P_1(1)(P_2(0)\dots P_n(0) \otimes P_2(1)\dots P_n(1)) \\ &\otimes(P_2(0)\dots P_n(0) \otimes P_2(1)\dots P_n(1)), \end{aligned} \tag{41}$$

ce qui implique que (34) est plus petit de ou égal à

$$P_1(0)P_2(0)\dots P_n(0) \otimes P_2(1)P_3(1)\dots P_n(1) \tag{42}$$
$$+ P_1(0)P_2(1)\dots P_n(0) \otimes P_1(0)P_3(0)\dots P_n(0) \tag{43}$$
$$\vdots$$
$$+ P_1(0)P_2(0)\dots P_n(1) \otimes P_1(0)P_2(0)\dots P_{n\otimes1}(0). \tag{44}$$

Répétons cette procédure n fois, en supposant $\prod_k P_k(0) > \prod_k P_k(1)$ chaque fois (sinon la démonstration est terminée). Après ces n étapes, il est démontré que (34) $\otimes$ $P_1(0)P_2(0)\dots P_n(0) \otimes P_1(0)P_2(0)\dots P_{n\otimes1}(0)$, où le côté droit est $\otimes$ 0 d'après (39) (Voir aussi [GR10] pour une démonstration alternative).

□

Même si la procédure ci-dessus pour calculer les bases de mesure ne marche pas pour les états de Dicke, l'inégalité est toujours violée par les états de Dicke dans un scénario à deux positions / deux résultats :

Théorème 10. Il existe un angle $0 < \otimes < \otimes$ tel que tous les états de Dicke $|S(n,k)\rangle$ ($\{k,n\} \in \mathbb{N}, 1 < k < n$) violent l'inégalité du Théorème 9 si la base 0 est définie par $\{|+\rangle, |\otimes\rangle\}$ et la base 1 est définie par $\{\cos\frac{\otimes}{2}|0\rangle \otimes \sin\frac{\otimes}{2}|1\rangle, \sin\frac{\otimes}{2}|0\rangle + \cos\frac{\otimes}{2}|1\rangle\}$.

Démonstration. D'abord nous réécrivons l'inégalité en fonction de n, k et □.

$$
\begin{aligned}
&P(0\ldots0|0\ldots0)\\
&= |\langle +,\ldots,+|S(n,k)\rangle|^2\\
&= ((\tfrac{1}{\sqrt{2}})^n \sqrt{\tbinom{n}{k}\tbinom{n}{k}})^2, \qquad (45)\\
&P(0\ldots0|1\ldots0)\\
&= |(\cos\tfrac{\theta}{2}\langle 0| - \sin\tfrac{\theta}{2}\langle 1|)\otimes\langle +,\ldots,+|S(n,k)\rangle|^2\\
&= ((\tfrac{1}{\sqrt{2}})^{n-1} \sqrt{\tbinom{n}{k}\tbinom{n-1}{k}}(\cos\tfrac{\theta}{2}\sqrt{\tbinom{n}{k}\tbinom{n-1}{k-1}} - \sin\tfrac{\theta}{2}\sqrt{\tbinom{n-1}{k-1}\tbinom{n-1}{k}}))^2, \qquad (46)\\
&P(1\ldots1|1\ldots1)\\
&= |\langle -|^{\otimes n}(\sin\tfrac{\theta}{2}\langle 0| + \cos\tfrac{\theta}{2}\langle 1|)^{\otimes n}|S(n,k)\rangle|^2\\
&= (\sqrt{\tbinom{n}{k}\tbinom{n}{k}}(\cos\tfrac{\theta}{2})^k(\sin\tfrac{\theta}{2})^{n-k})^2. \qquad (47)
\end{aligned}
$$

Pour simplifier notre calcul, nous divisons (45) - (47) par $\sqrt{\binom{n}{k}}$. La propriété de positivité ne sera pas changée par cette remise en échelle.

$$
\begin{aligned}
&P'(0\ldots0|0\ldots0)\\
&= (\tfrac{1}{2})^n, \qquad (48)\\
&P'(0\ldots0|1\ldots0)\\
&= (\tfrac{1}{2})^{n-1}\sqrt{\tbinom{n}{k}\tbinom{n-1}{k}}(\cos\tfrac{\theta}{2}\sqrt{\tbinom{n}{k}\tbinom{n-1}{k-1}} - \sin\tfrac{\theta}{2}\sqrt{\tbinom{n-1}{k-1}\tbinom{n-1}{k}})^2, \qquad (49)\\
&P'(1\ldots1|1\ldots1)\\
&= (\cos\tfrac{\theta}{2})^{2k}(\sin\tfrac{\theta}{2})^{2n-2k}. \qquad (50)
\end{aligned}
$$

A cause de la symétrie de $|S(n,k)\rangle$, les n probabilités $P'(0\ldots0|1\ldots0)$ à $P'(0\ldots0|0\ldots1)$ sont toutes égales. Après des simplifications, le polynôme de Bell devient :

$$
\begin{aligned}
P'(n,k,\theta) = &(\tfrac{1}{2})^n\\
&- n(\tfrac{1}{2})^{n-1}(\frac{n-k}{n}\cos\tfrac{\theta}{2} - \frac{k}{n}\sin\tfrac{\theta}{2})^2\\
&- (\cos\tfrac{\theta}{2})^{2k}(\sin\tfrac{\theta}{2})^{2n-2k}. \qquad (51)
\end{aligned}
$$

Notons d'abord quelques propriétés de (49) et (50). Pour (49), il peut atteindre 0 pour tout n, k quand $\tan\frac{\theta}{2} = \frac{n-k}{k}$. (50) est 0 quand $\theta = 0$ et $\theta = \pi$, il atteint son

maximum quand $(\cos\frac{\theta}{2})^{2k} = (\sin\frac{\theta}{2})^{2n-2k}$. Si nous fixons θ et n, sa derivee par rapport a k est

$$2(\cos\frac{\theta}{2})^{2k}(\sin\frac{\theta}{2})^{2n-2k}(\log\cos\frac{\theta}{2} - \log\sin\frac{\theta}{2}), \tag{52}$$

ce qui signifie que pour un θ et n fixe, quand $\theta < \frac{\pi}{2}$, (50) est strictement decroissant par rapport a k ; quand $\theta > \frac{\pi}{2}$, (50) il est strictement croissant par rapport a k ; et quand $\theta = \frac{\pi}{2}$, (50) est independant de k.

Nous allons considerer l'equation $n \cdot (49) = (\frac{1}{2})^{n+1}$:

$$n(\frac{1}{2})^{n-1}(\frac{n-k}{n}\cos\frac{\theta}{2} + \frac{k}{n}\sin\frac{\theta}{2})^2 = (\frac{1}{2})^{n+1}, \tag{53}$$

$$\Leftrightarrow \quad (\frac{n-k}{n}\cos\frac{\theta}{2} + \frac{k}{n}\sin\frac{\theta}{2})^2 = \frac{1}{4n}. \tag{54}$$

En prenant la racine carree de chaque cote, regroupant les termes avec $\sin\frac{\theta}{2}$ a un cote, substituant $\sin\frac{\theta}{2}$ avec $\sqrt{1-\cos^2\frac{\theta}{2}}$ et prenant le carre de chaque cote, (54) devient une equation quadratique ayant $\cos\frac{\theta}{2}$ comme l'inconnu. Elle peut avoir 0, 1 ou 2 racines selon la valeur de θ dans l'intervalle $[0, \pi]$. Si elle n'a pas de racine, alors $n \cdot (49) < (\frac{1}{2})^{n+1}$ pour tout θ (car (49) peut toujours atteindre 0). Notons aussi que (50) est zero quand $\theta = 0$ et $\theta = \pi$, donc si (54) n'a pas de racine alors (51) > 0 quand $\theta = 0$ et $\theta = \pi$. De meme facon, si (54) a une racine, nous pouvons demontrer que (51) > 0 quand $\theta = 0$ ou $\theta = \pi$, selon la valeur de la racine : quand la racine est plus petite que $\frac{\pi}{2}$, nous prenons $\theta = \pi$, sinon nous prenons $\theta = 0$.

Nous allons considerer le cas ou (54) a deux racines. Nous obtenons les formes fermees des racine dans l'intervalle $[0, \pi]$ en resolvant l'equation quadratique, nous allons noter les racines θ_+ et θ_- :

$$\cos\frac{\theta_+}{2} = \frac{k\sqrt{n} - n\sqrt{n} + \sqrt{8k^4 - k^2n - 8k^3n + 4k^2n^2}}{2(2k^2 - 2kn + n^2)} \tag{55}$$

$$\cos\frac{\theta_-}{2} = \frac{k\sqrt{n} - n\sqrt{n} - \sqrt{8k^4 - k^2n - 8k^3n + 4k^2n^2}}{2(2k^2 - 2kn + n^2)} \tag{56}$$

Nous allons demontrer que pour un n fixe, $(\cos\frac{\theta_+}{2})^{2k}(\sin\frac{\theta_+}{2})^{2n-2k} < (\frac{1}{2})^{n+1}$ et $(\cos\frac{\theta_-}{2})^{2k}(\sin\frac{\theta_-}{2})^{2n-2k} < (\frac{1}{2})^{n+1}$ pour tout k. A cause de la monotonicite de (50), il suffit de demontrer que l'inegalite est vraie pour $k = \frac{n}{2}$. Pour le moment nous allons prendre n comme un nombre pair. Nous ne considerons que la racine positive. Le raisonnement est symetrique et s'applique a la racine negative avec quelque changements.

Quand n est pair et $k = \frac{n}{2}$, $(\cos\frac{\theta_+}{2})^{2k}(\sin\frac{\theta_+}{2})^{2n-2k}$ se simplifie a :

$$(\frac{1}{32})^k(\underbrace{\frac{\sqrt{k}+\sqrt{k(4k-1)}}{k}}_{z})^{2k}(\underbrace{2+\frac{\sqrt{(4k-1)}}{k}}_{z})^k. \tag{57}$$

Pour trouver une borne superieure de (57), nous reecrivons (-) et (--) :

$$(-) = ((\frac{\sqrt{k}+\sqrt{k(4k-1)}}{k})^2)^k = (4-\frac{2\sqrt{4k-1}}{k})^k \tag{58}$$

$$= (4-\frac{4\sqrt{k-\frac{1}{4}}}{k})^k = (4(1-\frac{\sqrt{k-\frac{1}{4}}}{k}))^k, \tag{59}$$

$$(--) = (2+\frac{2\sqrt{k-\frac{1}{4}}}{k})^k = (2(1+\frac{\sqrt{k-\frac{1}{4}}}{k}))^k. \tag{60}$$

En substituant (59) et (60) dans (57), nous obtenons

$$(\frac{1}{32}-4-2-(1-\frac{k-\frac{1}{4}}{k^2}))^k = (\frac{1}{2})^{2k}(1-\frac{k-\frac{1}{4}}{k^2})^k. \tag{61}$$

Notons que $\lim_{n\to\infty}(1-\frac{1}{x})^x = e^{-1}$, et il atteint cette limite par le bas. Donc nous avons

$$(1-\frac{k-\frac{1}{4}}{k^2})^k < (1-\frac{1}{k})^k - e^{-1} < \frac{1}{2} \tag{62}$$

Depuis (61) et (62), nous obtenons la borne desiree

$$\begin{aligned}&(\cos\frac{\theta_+}{2})^{2k}(\sin\frac{\theta_+}{2})^{2n-2k}\\ &< (\frac{1}{2})^{2k}(1-\frac{1}{k})^k < (\frac{1}{2})^{2k+1}\end{aligned} \tag{63}$$

Pour demontrer que le theoreme reste valide pour n impair, supposons $n = 2k+1$ pour un certain k, alors (51) devient

$$(\frac{1}{2})^{2k+1}-(2k+1)(\frac{1}{2})^{2k}(\frac{k+1}{2k+1}\cos\frac{\theta}{2}-\frac{k}{2k+1}\sin\frac{\theta}{2})^2 \tag{64}$$

$$-(\cos\frac{\theta}{2})^{2k}(\sin\frac{\theta}{2})^{2k+2} \tag{65}$$

$$= \frac{1}{2}((\frac{1}{2})^{2k} \tag{66}$$

$$-(2k+1)(\frac{1}{2})^{2k-1}(\frac{k+1}{2k+1}\cos\frac{\theta}{2}-\frac{k}{2k+1}\sin\frac{\theta}{2})^2 \tag{67}$$

$$-2(\sin\frac{\theta}{2})^2(\cos\frac{\theta}{2})^{2k}(\sin\frac{\theta}{2})^{2k}) \tag{68}$$

Il est evident que $\frac{k+1}{2k+1} > \frac{1}{2}$ et $\frac{k}{2k+1} > \frac{1}{2}$, donc (67) peut etre borne par

$$(67) < 2k(\frac{1}{2})^{2k-1}(\frac{1}{2}\cos\frac{\theta}{2} - \frac{1}{2}\sin\frac{\theta}{2})^2. \tag{69}$$

Supposons que $\cos\frac{\theta}{2} = \frac{1}{\sqrt{2}}$ (l'autre cas sera discute ci-dessous), ce qui signifie que $2(\sin\frac{\theta}{2})^2 = 1$, nous pouvons borner (68) par :

$$(68) < (\cos\frac{\theta}{2})^{2k}(\sin\frac{\theta}{2})^{2k}. \tag{70}$$

En substituant (69) et (70) dans (67) et (68), nous avons

$$\begin{aligned}
&(\frac{1}{2})^{2k+1} \\
&< (2k+1)(\frac{1}{2})^{2k}(\frac{k+1}{2k+1}\cos\frac{\theta}{2} - \frac{k}{2k+1}\sin\frac{\theta}{2})^2 \\
&< (\cos\frac{\theta}{2})^{2k}(\sin\frac{\theta}{2})^{2k+2} \\
&< \frac{1}{2}((\frac{1}{2})^{2k} \\
&< 2k(\frac{1}{2})^{2k-1}(\frac{1}{2}\cos\frac{\theta}{2} - \frac{1}{2}\sin\frac{\theta}{2})^2 \\
&< (\cos\frac{\theta}{2})^{2k}(\sin\frac{\theta}{2})^{2k}),
\end{aligned}$$

qui est effectivement deux fois l'expression pour $n = 2k$. Si $\cos\frac{\theta}{2} = \frac{1}{\sqrt{2}}$, nous pouvons prendre $n = 2k - 1$, et avoir un argument similaire. Pour la racine negative, le meme raisonnement s'ensuit.

En conclusion, pour n pair, si $n \cdot (49) = (\frac{1}{2})^{n+1}$, alors $(50) < (\frac{1}{2})^{n+1}$, ce qui signifie que $(51) > 0$. Pour n impair, selon la valeur de $\cos\frac{\theta}{2}$ ou $\cos\frac{-\theta}{2}$, nous pouvons toujours borner (51) en considerant le n pair le plus proche. Donc (51) peut etre positif pour certaines valeurs de θ. □

0.3 La non-localité et les classes d'intrication

Dans la représentation de Majorana, il est possible que les MPs $\{|\alpha_1\rangle, \ldots, |\alpha_n\rangle\}$ d'un état symétrique ne soient pas tous distincts. Dans une nouvelle notation qui intègre la dégénérescence, d_i indique le degré de dégénérescence du MP $|\alpha_i\rangle$. Alors (5) devient

$$|\psi\rangle = K \sum_{perm} |\alpha_1^{d_1} \alpha_2^{d_2} \ldots \alpha_l^{d_l}\rangle, \qquad (71)$$
$$\forall i \neq j, |\alpha_i\rangle \neq |\alpha_j\rangle, \sum_{i=1}^{l} d_i = n.$$

En plus, (7) devient

$$(\langle \alpha_i^{\perp} |)^{\otimes k} |\psi\rangle = 0. \qquad (72)$$

avec $(n - d_i) < k \leq n$. Les opérations locales et communications classiques ne changent pas la dégénérescence, même de la manière stochastique (SLOCC) [BKM+09] (voir aussi [RM11, Aul11b]). Par conséquent, différents degrés de dégénérescence correspondent à différentes classes d'intrication.

Quand la dégénérescence est prise en compte, le paradoxe peut être étendu en considérant un sous-ensemble de parties. La version probabiliste de (72) montre que la corrélation persiste dans le sous-ensemble

$$P(\underbrace{1|\ldots|1}_{k}|\underbrace{1|\ldots|1}_{k}) = 0, \qquad (73)$$

pour $(n - d_i) < k \leq n$.

L'inégalité du théorème 9 est aussi étendue à

$$Q_d^n := P^n - P(\underbrace{1|\ldots|1}_{n-1}|\underbrace{1|\ldots|1}_{n-1}) - \ldots - P(\underbrace{1|\ldots|1}_{n-d+1}|\underbrace{1|\ldots|1}_{n-d+1}) \geq 0. \qquad (74)$$

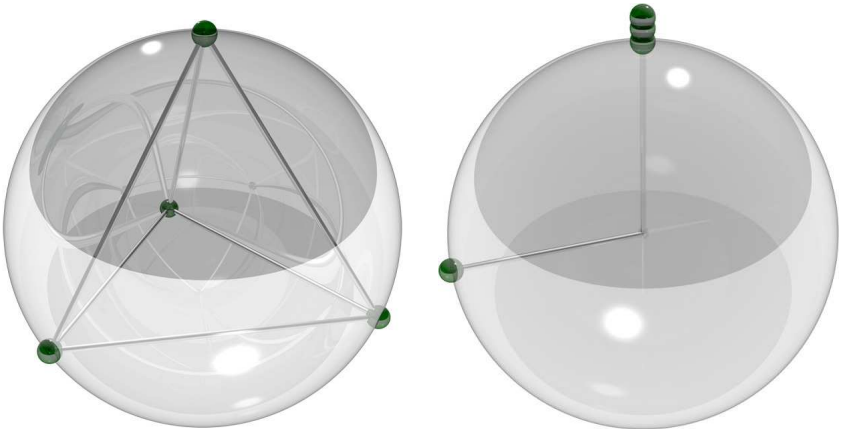
La borne supérieure de LHV n'est pas changée parce que les termes soustraits sont non-positifs.

Les inégalités P^n et Q_d^n peuvent être utilisées pour la classification de SLOCC. Pour un exemple, nous allons considérer trois états : l'état tétraèdre $|T\rangle = \sqrt{\frac{1}{3}}|S(4,0)\rangle + \sqrt{\frac{2}{3}}|S(4,3)\rangle$, l'état de GHZ à 4 qubits $|GHZ_4\rangle = \frac{1}{\sqrt{2}}(|0000\rangle + |1111\rangle)$ et état $|000+\rangle = K_{perm}|000+\rangle = \sqrt{\frac{2}{5}}|0000\rangle + \sqrt{\frac{1}{5}}|S(4,1)\rangle$. Chaque état appartiens à une classe SLOCC différente (cf. chapitre 3, section 3.1.2). Ils sont groupés en deux groupes : un groupe contient $|T\rangle$ et $|000+\rangle$, avec deux degrés de dégénérescence différents ; l'autre groupe

contient $|T\rangle$ et $|GHZ_4\rangle$, avec le meme degre de degenerescence. Les deux groupes sont illustres dans Fig. 1 et 2, respectivement. Une classification independante des dispositifs est realisee a partir de bornes calculees avec SDP.

L'equivalence des etats symetriques sous LUP est donnee simplement par la distribution de ses points de Majorana, parce qu'une operation unitaire locale qui mene un etat symetrique a un etat symetrique est en effet une rotation de la sphere de Bloch [BKM+09, MKG+10] (en plus, la permutation ne change bien sur pas un etat symetrique). Alors, chaque etat qu'on utilise ici est in-equivalent sous LU ou LUP. Le fait que la violation trouvee est sur toutes les mesures possibles garantit que les bornes soient valides pour tout etat equivalent sous LU et LUP.

Pour le premier groupe illustre dans Fig. 1, les bornes se trouvent en Table 1.



(a) (b)

FIGURE 1: L'etat tetraedre (a) et l'etat $|000+\rangle$ (b) dans la representation de Majorana.

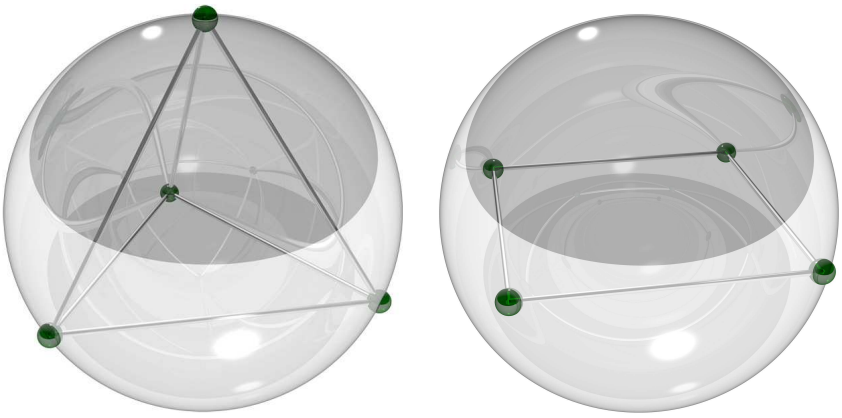
Etat	P^4	Q_3^4
$\|T\rangle$	0.1745	-0.0609
$\|000+\rangle$	0.0142	0.0141

TABLE 1: Les bornes de violations maximales de P^4 et Q_3^4 pour $|T\rangle$ et $|000+\rangle$. Une violation de Q_3^4 implique que l'etat n'est pas dans la classe LUP de $|T\rangle$.

La Table. 2 montre les bornes de P^4 et Q_3^4 pour le deuxieme groupe, illustre par Fig. 2. Les bornes sont calculees en utilisant la programmation semi-definie de la section 2.5.

A partir de ces tables, il est facile d'envisager un test independant des dispositifs qui permet de discriminer les classes LUP dans chaque groupe.

Pour le premier groupe, le test est plus faible a cause des contraintes sur les bases de mesures. Malgre notre verification numerique et l'hypothese raisonnable de restreindre les bases de mesures, il n'y a pas de garantie qu'un etat ne soit pas dans la classe $|000+\rangle$



(a) (b)

FIGURE 2: L'état tétraèdre (a) et l'état GHZ à 4 qubits $|GHZ_4\rangle$ (b) dans la représentation de Majorana.

Etat	P^4	Q_3^4
$\|T\rangle$	0.1745	-0.0609
$\|GHZ_4\rangle$	0.1241	0.0563

TABLE 2: Les bornes de violations maximales de P^4 et Q_3^4 pour $|T\rangle$ et $|GHZ_4\rangle$. Une violation de $P^4 > 0.1241$ implique que l'état n'est pas dans la classe LUP de $|GHZ_4\rangle$, et une violation de Q_3^4 implique l'état n'est pas dans la classe LUP de $|T\rangle$.

s'il viole P^4 de plus de 0.0142. Néanmoins, on peut conclure que si un état viole Q_3^4, alors il n'est pas dans la classe $|T\rangle$, mais doit être dans la classe $|000+\rangle$.

Dans le second groupe, si le test P^4 donne une violation $\square$ 0.1241, alors l'état ne doit pas être dans la classe de $|GHZ_4\rangle$, mais dans la classe de $|T\rangle$. De même, si le test Q_3^4 donne une violation, l'état est dans la classe de $|GHZ_4\rangle$. Dans ce cas, même s'il n'y a pas de dégénérescence, une séparation de classe est possible par le test Q_3^4.

Les états choisis sont dans différentes classes SLOCC. Aller au-delà de 4 qubits est difficile à cause de l'augmentation exponentielle de paramètres à optimiser. Néanmoins un simple test d'optimisation de bases pour les états W et GHZ indique qu'il est possible de les séparer en utilisant la technique ci-dessus. Ceci fait avancer la discussion de comment interpréter la classification d'intrication par les caractères non-locaux. D'un côté la dégénérescence des MPs garantit la persistance de corrélations [WM12], ce qui est vrai pour tout état. Certains états «représentatifs» comme $|000+\rangle$ et $|W\rangle$ peuvent être séparés d'états moins dégénérés en utilisant cette propriété. Il est similaire à la force de non-localité à cause de perte de systèmes [BSV12] [BV12]. D'un autre côté, l'exemple ci-dessus montre qu'il est possible d'utiliser Q_d^n pour séparer certains états avec le même degré de dégénérescence. Ces états se trouvent naturellement dans l'espace des phases

de certains condensats [MS07]. L'idee ci-dessus peut servir comme un temoin de phases.

0.4 Une inégalité monogame pour les états de Dicke

Une nouvelle inegalite est introduite dans le but de montrer la monogame stricte des etats de Dicke dans la limite de n tendent vers l'infinite. Cette inegalite est basee sur un resultat recent de Heaney, Cabello, Santos et Vedral [HCSV11], ou ils montrent que pour les etats W il est possible de construire une inegalite «maximale» dont la violation atteint le maximum algebrique quand n ! 1 , et ainsi reproduire la correlation parfaite des etats de stabilisateurs et l'inegalite de Mermin (cf. chapitre 2, section 2.4.4). Cette inegalite de Heaney, Cabello, Santos et Vedral (desormais denommee l'inegalite de HCSV) est construite pour les etats W mais il est possible de l'etendre a tous les etats de Dicke.

En suivant et etendant les raisonnements dans [HCSV11] pour les etats W, si toutes les n parties mesurent σ_z sur un etat de Dicke $|S(n,k)\rangle$, $n - k$ d'entre elles vont obtenir le resultat 0 et les autres k parties vont obtenir le resultat 1 avec certitude (il faut noter qu'il n'est pas possible de savoir qui a obtenu quel resultat). Imaginons que $n - k - 1$ parties ont obtenu 0 et $k - 1$ ont obtenu 1. Les deux autre parties ont decider de mesurer dans la base σ_x, et elles obtiennent toujours le meme resultat. Le modele de LHV stipule que le resultat d'une partie est independant de la base de mesure des autres, et par consequent si n'importe quel ensemble de deux parties a decide de mesurer σ_x, ils vont toujours obtenir le meme resultat. Si les resultat sont donnes par LHV, le raisonnement ci-dessus stipule que tout le monde va obtenir le meme resultat s'ils ont mesure σ_x. Mais un simple calcul montre que ce n'est pas le cas pour tous les etats de Dicke.

L'inegalite de Bell associee est

$$\begin{aligned} L = & \sum P(\sigma(0\underbrace{\ldots}_{n-k}0\,1\underbrace{\ldots}_{k}1)|0\ldots0) \\ & - \sum P(\sigma(0\underbrace{\ldots}_{n-k-1}0\,1\underbrace{\ldots}_{k-1}1\,0\,1)|\sigma(0\underbrace{\ldots}_{n-2}0\,1\,1)) \\ & - P(0\ldots0|1\ldots1) - P(1\ldots1|1\ldots1) - 0, \end{aligned} \tag{75}$$

ou les permutations dans la deuxieme et troisieme lignes sont sur les parties qui fixent les bases de mesure et le resultat, comme dans l'inegalite P^n. Pour montrer qu'elle est compatible avec LHV, il suffit de noter qu'elle est compatible avec toutes les strategies deterministes, comme les autres strategies sont des melanges probabilistes de strategies deterministes) [WW01]. Il n'est pas difficile de voir que mettre n'importe quel terme $P(\sigma(0\underbrace{\ldots}_{n-k}0\,1\underbrace{\ldots}_{k}1)|0\ldots0)$ a 1 n'est pas compatible avec garder tous les termes negatifs a zeros. En plus, il n'est pas possible d'avoir plus d'un terme positif a 1, donc l'expression est toujours negative et une violation est incompatible avec LHV. Pour un etat de Dicke, $|S(n,k)\rangle$, L est violee par $1 - \frac{\binom{n}{k}}{2^{n-1}}$. Comme les etats W dans [HCSV11], l'inegalite atteint son maximum quand n tend vers l'infini.

Le graphe ci-dessous en Fig. 3 visualise la violation de L par $|S(n, \frac{n}{2})\rangle$ et $|W_n\rangle$. Les etats W atteignent 1 plus rapidement, a cause de ses faibles intrications.

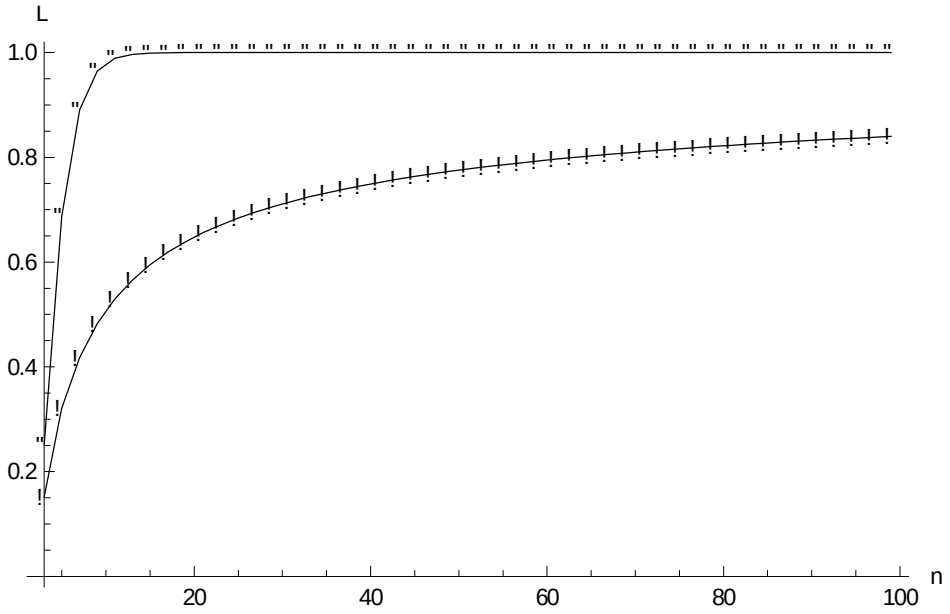


FIGURE 3: Comparaison de la violation de L (75) par $|S(n, \frac{n}{2})\rangle$ (n) et $|W_n\rangle$ (l) en fonction de n (le nombre de parties).

En plus de la monogamie, une autre propriete utile pour la cryptographie est la non-localite authentique. Une correlation est authentique si elle ne peut pas être atteinte par un sous-ensemble de parties. Les inegalites de type Svetlichny [Sve87] ont pour but de detecter la non-localite authentique : seules les correlations authentiques de n parties peuvent violer ces inegalites. Malheureusement toutes les inegalites presentees ne peuvent pas detecter la non-localite authentique. Le theoreme ci-dessous montre cela explicitement pour L.

Théorème 11. *L'inégalité L ne peut pas détecter la non-localité authentique.*

Démonstration. Pour montrer que L ne peut pas detecter la non-localite authentique, nous allons grouper les deux premieres parties et montrer que L = 1 dans un modele LHV partiellement non-local (ou les deux premieres parties sont traitees comme une seule partie). En langage mathematique, un modele LHV signifie

$$P(a_1 \dots a_n | A_1 \dots A_n) = \int \rho(\lambda) \prod_{1 \le i \le n} P_i(a_i | A_i, \lambda)\, d\lambda, \qquad (76)$$

ou les indices en bas denotent les parties.

Pourtant, dans un modele LHV partiellement non-local nous permettons un ensemble de parties de se groupe et se comporte comme une seule partie. En particulier, cela

signifie que

$$P(a_1 \dots a_n | A_1 \dots A_n) = \int \rho(\lambda) P_{12}(a_1 a_2 | A_1 A_2, \lambda) \prod_{3 \leq i \leq n} P_i(a_i | A_i, \lambda) \, d\lambda. \quad (77)$$

Nous donnons un modèle explicite qui donne les valeurs 0 ou 1 à tous les termes de L. Il implique qu'un seul terme dans la somme $\sum P(\lambda(0\dots01\dots1)|0\dots0)$ (avec $n-k$ et k répétitions) est 1, tous les autre termes sont 0. Supposons, sans perte de généralité, que

$$P(0\dots01\dots1|0\dots0) = 1. \quad (78)$$

($n-k$ et k répétitions)

Ceci implique

$$P_{12}(00|00) = 1 \quad (79)$$

$$P_3(0|0) = 1, \dots, P_{n-k}(0|0) = 1 \quad (80)$$

$$P_{n-k+1}(1|0) = 1, \dots, P_n(1|0) = 1, \quad (81)$$

d'où on déduit

$$P_{12}(01|00) = P_{12}(10|00) = P_{12}(11|00) = 0 \quad (82)$$

$$P_3(1|0) = 0, \dots, P_{n-k}(1|0) = 0 \quad (83)$$

$$P_{n-k+1}(0|0) = 0, \dots, P_n(0|0) = 0. \quad (84)$$

Pour les termes $\sum P(\lambda(0\dots01\dots101)|\lambda(0\dots011))$ ($n-k-1$ et $k-1$ répétitions, $n-2$ répétitions), nous allons les fixer à 0, en utilisant (82) à (84) avec quelques conditions supplémentaires sur les probabilités, sans donner des inconsistances.

Pour voir comment fixer tous les termes à 0, nous allons diviser les termes dans la somme en trois cas (a, b sont des bits, $\bar{a}, \bar{b}$ sont leur négations logiques) :

1. $P(ab|\lambda(0\dots01\dots01)|00|\lambda(0\dots011))$.

   Dans ce cas, si a et b ne sont pas tout les deux 0, alors d'après (82), la probabilité est 0. Sinon, nous fixons $P_i(0|1) = 0$, où $i \in 1, 2$.

2. $P(ab|\lambda(0\dots01\dots\bar{b})|01|\lambda(0\dots001))$, $P(ab|\lambda(0\dots01\dots\bar{a})|10|\lambda(0\dots001))$.

   Dans ce cas, si $a = b = 1$, alors il existe $P_i(0|1)$ où $i \in 1, 2$. Donc nous avons $P_i(0|1) = 0$ et $P_{12}(11|01) = 1, P_{12}(11|10) = 1$, sans avoir d'inconsistance avec le cas précédent. Les deux dernières affectations impliquent aussi que si a et b ne sont pas tout les deux 1, alors $P_{12}(ab|01) = P_{12}(ab|10) = 0$.

3. $P(a\bar{a}\square(0\ldots01\ldots01)j11\square(0\ldots0))$.

   Dans ce cas, la probabilite est toujours 0. La conclusion vient du principe des tiroirs : il y a $n \square k \square 1$ resultats 0 quand les parties 3 a n mesurent dans la position 0, donc au moins une partie de $n \square k+1$ a n va obtenir le resultat 0 dans la position 0. D'apres (84) la probabilite est 0.

Dans le dernier cas, comme la probabilite est toujours 0 independant des affectations de deux premieres parties, nous pouvons fixer $P_{12}(00j11) = 0$ et $P_{12}(11j11) = 0$ sans avoir d'inconsistance. Ces affectations garantissent les deux dernieres probabilites dans L : $P(0\ldots0j1\ldots1)$ et $P(1\ldots1j1\ldots1)$ sont 0.

Donc nous pouvons fixer les probabilites d'une façon consistante tel que tous les terme negatifs de L sont 0 et la somme de tous les termes positifs sont 1, c-a-d $L = 1$. Ca signifie que L ne peut pas detecter la non-localite authentique.

Un argument similaire peut être avance pour P^n ou Q_d^n. □

En conclusion il semble qu'on doive faire un choix equilibre entre l'utilite d'une inegalite et la circonstance. Il est montre que les inegalites P^n et Q_d^n sont utiles pour separer les classes. En effet tous les etats intriques violent P^n [YCZ+12]. Mais leur violations ne sont jamais assez fortes pour garantir la monogame. Ces inegalites sont aussi incapables de detecter la non-localite authentique (même pas L, avec ses nombreux termes positifs). D'un autre côte, les inegalites en esperances mathematiques (qui ont forcement de nombreux termes positifs) peuvent avoir une violation maximale pour n'importe quel n. Mais ils ne peuvent pas detecter la non-localite de tous les etats - il existe des etats qui ne violent pas n'importe quelle inegalite en esperances mais qui violent P^n [ZBLW02]. D'une façon similaire a la situation en theorie d'intrication, il semble qu'il n'est pas possible de capturer toutes les caracteristiques de non-localite par une seule inegalite.

Chapter 1

Introduction

This thesis, as its name suggests, investigates the nonlocal features of permutation symmetric states, in order to gauge the potential usefulness of these states in quantum information processing. The approach taken is a very pedagogical one, assuming little prior knowledge of quantum information or foundations of quantum mechanics.

Chapter 2 gives all the background information necessary to understand basic quantum mechanics and nonlocality. Section 2.1 is a review of essential linear algebra to understand the linear structures used in quantum mechanics: Hilbert space, Hermitian operators, etc. Section 2.2 introduces the postulates of quantum mechanics, showing how different concepts introduced in the previous section acquire physical meaning. The postulates chosen are variations of the ones commonly used in quantum information, especially the measurement postulate, which includes POVMs. The next section, Section 2.3, briefly reviews the concepts and results in entanglement theory, beginning from the definition of entanglement. The central notion of entanglement measures, local operations and classical communication, is also introduced together with common axioms for entanglement measures to satisfy. Moving from entanglement, Section 2.4 reviews the history and developments of nonlocality, a concept dating back to the early days of quantum mechanics. After giving brief historical information, three common ways to show nonlocality are shown: nonlocality from correlation functions, perfect correlations with MABK inequality and the Hardy paradox with the CH inequality. This section ends by discussing recent attempts to unify different approaches to nonlocality, either by using convex geometry, or category theory, or hyper graphs. The last section of this chapter introduces the main numerical technique used to obtain numerical results in later chapters: semidefinite programming.

Chapter 3 introduces the geometrical aspects of symmetric states, in the form of the Majorana representation. Section 3.1 shows that complex numbers and geometry are intimately connected. Complex numbers, in addition to the usual geometrical representation on the complex plane, has another, compact representation by the Riemann sphere.

The automorphisms of the Riemann sphere, called Möbius transformations, have a direct interpretation in entanglement of symmetry states as the geometrical manifestation of SLOCC operations. Section 3.2 explains the main tool used by the study of both the entanglement and the nonlocality of symmetric states: the Majorana representation. The geometric representation of a symmetric state given by the Majorana representation is shown, together with constructive procedures to decompose a symmetric state into its Majorana points and reconstruct a symmetric state from a set of Majorana points. Several useful properties derived from the decomposition/ reconstruction are also shown in this section, which will be used later to prove the nonlocality of symmetric states. Section 3.3 reviews the physical aspects of the Majorana representation and its historical significance. The connection between complex geometry and physical spin is shown. What is not shown explicitly but nevertheless mentioned is the connection between spinors and special relativity [PR84, PR86]. The last section of this chapter reviews the main results of an earlier thesis on the entanglement properties of symmetric states [Aul11a]. The connection between the Möbius transformation and SLOCC classification is made explicit. It also is shown that the symmetric states whose Majorana points form the Platonic solids are special because their closest product states form the dual Platonic solid.

Chapter 4 mainly contains mathematical proofs of several main theorems used in the thesis. The first section reviews the original result by Hardy, including how to find suitable measurement settings to satisfy the bipartite Hardy paradox. Then the next section extends the bipartite Hardy paradox to n party, together with a generalized CH inequality, which will be called P^n. Section 4.3 proves the main theorem showing almost all symmetric states satisfy the n-party Hardy paradox, and as a result also violate P^n. As part of the proof, it is also shown how to choose the measurement settings to satisfy the paradox. The last section proves the existence of measurement settings which allow Dicke states to also violate P^n, although they do not satisfy the n-party paradox.

Chapter 5 focuses on a special feature of symmetric states, made evident by the Majorana representation: degeneracy. Degeneracy denotes the fact that not all Majorana points of a symmetric states are distinct. It comes from the degeneracy of polynomial roots. Using degeneracy, it is possible to show that entanglement and nonlocality may persist to subsystems. To show the persistency of nonlocality, new conditions are added to the n-party Hardy paradox and a new inequality Q_d^n is defined. The conditions, old plus new, can only be perfectly satisfied by a state with enough degeneracy. Thus having enough degeneracy also guarantees the violation of Q_d^n. The relationship between degeneracy and entanglement is known. Semidefinite programming techniques allow the device independent classification of states into different nonlocal classes using local unitaries and permutation of systems (LUP) classification scheme. The classified states also sit in different SLOCC classes, using the link between SLOCC operations and the Möbius transformation.

Chapter 6 performs additional analysis (mainly numerical) of the nonlocality and entanglement properties of symmetric states. The first section performs numerical analysis on the strength of violation of P^n by GHZ and W states, as the number of parties grow. Because of the structure of P^n, the violation is always bounded by the geometric measure of entanglement of the state. Unfortunately, because of this bound, P^n and Q_d^n are not monogamous, which is a useful property in cryptography. Monogamy is defined both in the context of entanglement and in the context of nonlocality. Section 6.2 introduces these two concepts of monogamy, and shows that even though P^n is not monogamous in the strict sense, there is still some monogamy in the broad sense. The last section of the chapter shows a new inequality for Dicke states, based on recent results of Heaney, Cabello, Santos and Vedral, which is monogamous in the strict sense when the number of parties goes to infinity. But neither this new inequality nor P^n (or Q_d^n) can detect genuine nonlocality.

Chapter 7 is dedicated to two potential applications of the results obtained in previous chapters: communication complexity and Bayesian game theory. The fact that nonlocality is beneficial to communication complexity has been known for more than ten years. In the most extreme case, the usage of a PR box, which represents stronger than quantum correlations, renders the communication complexity of all Boolean functions trivial. This section first reviews the model of bipartite communication complexity, showing how classical randomness, both private and shared, can sometimes reduce the communication complexity of computing certain functions. Then a quantum advantage is shown to exist for some functions for which classical randomness does not help. However, the probabilities in the n-party Hardy paradox do not form a function, instead they form a relation. To the best of my knowledge, the theory of multiparty communication complexity of relations is not as well-developed as multiparty communication complexity of functions or the bipartite communication complexity of relations. So further research into this area is needed. The other application of nonlocality is to Bayesian games, which represent games with incomplete information. Models already exist to translate an inequality into a Bayesian game. The Hardy paradox has already been translated into a Bayesian game, with an explicit payoff matrix. A recent result by Brunner and Linden [BL12] has pointed out the broad link between nonlocality and Bayesian game theory.

The last chapter summarizes the new results in this thesis, points out a few new developments in related areas and envisages a few future directions of research and open questions.

Chapter 2

Background

This chapter introduces most of the mathematical language of quantum mechanics and nonlocality used by the rest of the thesis. The type of mathematics in this chapter is rather general, leaving more specialized concepts such as complex geometry and the Majorana representation to be treated in later chapters. Because nonlocality is considered as part of the foundation of quantum mechanics, the approach taken in this chapter is very pedagogical: starting from the basics of linear algebra and build up the mathematical foundations of quantum mechanics.

2.1 Essential Algebra

According to one of the postulates of quantum mechanics, the "stage" for the quantum mechanical show is a complex Hilbert space: a complex vector space which is complete in the norm induced by the inner product. In what follows each word in this definition will be properly defined, leading to a mathematically rigorous yet self-contained understanding of the concept of a complex Hilbert space.

The first important word in the definition is complex. It is in fact an abbreviation for two related concepts: the set of complex numbers and the field of complex numbers. A field is a set F with two binary operations, denoted + and □, which must satisfy these axioms:

- Closure under + and □. $\forall a, b \in F$, $a + b \in F$ and $a \square b \in F$. This is just another way of saying + and □ are binary operations.

- Associativity of + and □. $\forall a, b, c \in F$, $a + (b + c) = (a + b) + c$ and $a \square (b \square c) = (a \square b) \square c$.

- Commutativity of + and □. $\forall a, b \in F$, $a + b = b + a$ and $a \square b = b \square a$.

- Distributivity of $\cdot$ over $+$. $\forall a, b, c \in F$, $a \cdot (b + c) = a \cdot b + a \cdot c$ and $(a + b) \cdot c = a \cdot c + b \cdot c$.
- Existence of identity elements for $+$ and $\cdot$. $0 \in F$ such that $\forall a \in F$, $0 + a = a + 0 = a$. $1 \in F$ such that $\forall a \in F$, $1 \cdot a = a \cdot 1 = a$. Also, $0 \neq 1$.
- Existence of inverses for $+$ and $\cdot$. $\forall a \in F$, $\exists -a \in F$ such that $a + (-a) = 0$. $\forall a \neq 0 \in F$, $\exists a^{-1} \in F$ such that $a \cdot a^{-1} = 1$.

The set of real numbers, together with addition and multiplication, satisfies all the axioms, thus forming a field R. The set of complex numbers is a superset of the set of real numbers, having the form $a + b \cdot i$, where $a, b \in R, i^2 = -1\}$. The field of complex numbers is given by the set of complex numbers with addition and multiplication. When there is no need to distinguish the two, the symbol C is used to denote both the set and the field of complex numbers.

The second and perhaps the most important concept in the definition, is the notion of a vector space. A vector space is a mathematical structure built on top a field. The elements of the field are called scalars in the context of vector spaces. A vector space V is an nonempty set of "things", called vectors, together with the field of scalars F and two operations: $+ : V \times V \to V$ and $\cdot : F \times V \in V$, which also satisfy the following axioms:

- Associativity of $+$. $\forall u, v, w \in V$, $u + (v + w) = (u + v) + w$.
- Commutativity of $+$. $\forall u, v \in V$, $u + v = v + u$.
- Existence of the zero vector. $0 \in V$ such that $\forall u \in V$, $0 + u = u$.
- Existence of inverses for $+$. $\forall u \in V$, $\exists -u \in V$ such that $u + (-u) = 0$.
- Distributivity of $\cdot$. $\forall a \in F, u, v \in V$, $a \cdot (u + v) = a \cdot u + a \cdot v$. $\forall a, b \in F, u \in V$, $(a + b) \cdot u = a \cdot u + b \cdot u$.
- Compatibility of $\cdot$ with scalar operation $\cdot$. $\forall a, b \in F, u \in V$, $(a \cdot b) \cdot u = a \cdot (b \cdot u)$.
- Existence of the identity element of $\cdot$. $1 \in F$ such that 1 is the identity element of $\cdot$ and $\forall u \in V$, $1 \cdot u = u$.

A complex vector space is a vector space whose scalar field is C. From now on, unless mentioned explicitly, all vector spaces (including Hilbert spaces) are complex.

The scalars offer richer structures in a vector space by defining additional operations on it. The next concept in the definition we need to know, the inner product $\langle \cdot, \cdot \rangle : V \times V \to C$, is the most common additional operation. To qualify as an inner product, $\langle \cdot, \cdot \rangle$ must satisfy the following properties:

- Positive definite. $\forall u \in V$, $\langle v, v \rangle \geq 0$, with $\langle v, v \rangle = 0$ iff $v = 0$.

- Conjugate symmetric. $\forall u, v \in V, \langle u, v\rangle = \overline{\langle v, u\rangle}$.
- Linearity in the first argument. $\forall u, v, w \in V, a, b \in C, \langle au + bv, w\rangle = a\langle u, w\rangle + b\langle v, w\rangle$.

The last two properties imply that the inner product is sesquilinear for complex vector spaces: $\forall u, v, w \in V, a, b \in C, \langle w, au + bv\rangle = \bar{a}\langle w, u\rangle + \bar{b}\langle w, v\rangle$. Two vectors are said to be orthogonal if their inner product is 0.

In an Euclidean space, the inner product of two vectors is related to their "length" and the angle between them. More precisely, it is the product of each of their lengths times the cosine of the angle between them. However, there is something missing in this intuition: the "length" of a vector is not defined so far. To make things worse, there is more than one way to define what "length" is in a vector space! As the definitions above show, concepts tend to be defined axiomatically in algebra, and the concept of "length" is no exception. Except it is not called "length", it is called norm, the abstract generalization of "length". The norm, $\|\cdot\|: V \to R$, is defined axiomatically as any function satisfying the following:

- Only the zero vector has norm zero. $\|v\| = 0$ iff $v = 0$.
- Compatibility with the norm of the scalar field. $\forall a \in C, v \in V, \|av\| = |a|\|v\|$, where $|\cdot|$ is the modulus.
- Triangle inequality. $\forall v, w \in V, \|v + w\| \leq \|v\| + \|w\|$.

For any inner product space, there is always a norm "induced" by the inner product:

$$\forall v \in V, \|v\| = \sqrt{\langle v, v\rangle}.$$

The geometrical intuition given above for the inner product function can be also applied to this norm: since the angle of a vector with itself is 0,the cosine is 1. The the square root of the product of its "length" with itself (or in the complex case, with its conjugate) gives the absolute value (or the modulus) of the "length" of the vector.

The last significant word in the definition is the notion of completeness. However, to define completeness in a vector space rigorously, the knowledge of a Cauchy sequence is needed. A Cauchy sequence in a vector space V is comprised of a sequence of vectors $\{x_n\}$, such that $\forall \varepsilon > 0, \exists \delta \in N$ which makes $\|x_m - x_n\| < \varepsilon$ for all $m, n > \delta$. For a vector space to be complete, then every Cauchy sequence in it needs to converge to a vector in it. Luckily, all (normed) finite dimensional vector spaces are complete, the proof of which can be found in most functional analysis textbooks.

Although by now a Hilbert space can be rigorously defined, it is still too abstract. The basic elements of a vector space, the vectors, have not been properly introduced. Vectors are supposed to be abstract entities satisfying the axioms, but this does not mean

they cannot be visualized or written down. To write down a vector, a basis is needed. Before defining what a basis is, it is more useful to define linear independence. A set of vectors v_i is called linearly independent if for $a_i \in C$, the only way that the expression $a_1v_1 + a_2v_2 + \ldots + a_nv_n = 0$ holds is when all $a_i = 0$. If a set of vectors is not linearly independent, they are linearly dependent. Linearly independent vectors can serve as "starting points" to reach other vectors by scalar multiplication and vector addition. The vector space spanned by a set of linearly independent vectors is the vector space where the vectors are all possible combinations of scalar multiplications and vector additions. If a set of vectors are linearly independent and spans a space V, then this set is called a basis of V. The number of elements in this set is called the dimension of the vector space. If in addition to being linearly independent, they are also mutually orthogonal (i.e. having inner product zero), then the basis is called an orthogonal basis. If each vector has norm 1, then the basis is called an orthonormal basis. Not all vector spaces can have orthonormal bases, but with its inner product and induced norm, an orthonormal basis can be always defined in a Hilbert space.

Given two or more vector spaces, one natural task is to construct a larger vector space whose vectors are constructed from the vectors of the original spaces in the least restrictive manner possible. This leads to the tensor product of two (or more) vector spaces.

To construct the tensor product of two vector spaces U of dimension m V of dimension n, take any vector $u \in U$ and $v \in V$, then the vector (u, v) of dimension $m \cdot n$ is a member of a larger vector space $U \otimes V$, which is called the tensor product of U and V. To see why the dimension of the tensor product space is $m \cdot n$ instead of $m + n$, take any basis of U, $\{u_1, \ldots, u_m\}$, and a basis of V, $\{v_1, \ldots, v_n\}$, then any combination (u_i, v_j) is basis of $U \otimes V$.

Although taking the tensor product of two vector spaces produces a much bigger space, the new space is nevertheless a vector space, so basic properties concerning vector addition and scalar multiplication should still apply in a tensor product space, albeit in a modified form. Instead of requiring the vector addition to be linear, it should be bilinear in a tensor product space, i.e. linear in both the u and v part of (u, v). And the scalar multiplication should be compatible with with both parts as well. The additional axioms for the vectors (u, v) of $U \otimes V$ are:

- Bilinearity in both U and V. $\forall u, u_1, u_2 \in U$ and $\forall v, v_1, v_2 \in V$, $(u_1 + u_2, v) = (u_1, v) + (u_2, v)$, $(u, v_1 + v_2) = (u, v_1) + (u, v_2)$.
- Compatibility with scalar multiplication in both U and V. $\forall z \in C$, $(c \cdot u, v) = c \cdot (u, v) = (u, c \cdot v)$.

Note that the definition of the tensor product given here is meant to be intuitive instead of mathematically rigorous. For a mathematically rigorous treatment of the

tensor product, either in terms of the free product or in terms of universal pairs in category theory, see [Sze04, Rom08]

Knowing what a Hilbert space is only reveals half of the picture. Vectors are static objects in a vector space, and they are meant to be moved around, transformed into other vectors (possibly in a different vector space) or in general, be manipulated. The way this manipulation is done should respect the vector space structure, to guarantee that the manipulated vectors still remain vectors. The most natural transformations which fulfill this requirement are linear transformations. A linear transformation from a vector space V to W is defined as any function $\Box : V \to W$ such that $\Box(au + bv) = a\Box(u) + b\Box(v)$, $\forall a, b \in C$ and $u, v \in V$. In the case that $V = W$, the linear transformation is called a linear operator.

There is also a special case when W is the scalar field of V, considered as a vector space whose scalar field is itself. In this case the function $\Box$ takes any vector in V to a scalar in its scalar field. Such a function $\Box$ is called a linear functional over V. The linear functionals also form a vector space themselves, it is called the dual space of V, denoted by $V^{\Box}$.

Being a normed vector space, linear transformations on a Hilbert space have some extra properties. The most useful linear transformations in quantum mechanics are bounded linear transformations, defined as any linear transformation $\Box$ such that $\exists a > 0, \forall v \in V, \|\Box(v)\| \leq a\|v\|$. Bounded linear functionals of a Hilbert space are very special because the following well-known theorem, whose proof can be found in [Sze04] [Rom08] and [DM05], states that the dual space is essentially the same as (called isomorphic to) the original space:

Theorem 1. (Riesz representation theorem) Let $\Box$ be a bounded linear functional on a Hilbert space H, there exists a unique vector $x' \in H$ such that $\Box(x) = \langle x, x' \rangle$, $\forall x \in H$. Also, $\|\Box\| = \|x'\|$.

Everything above is general enough to cover both finite and infinite dimensional Hilbert spaces. However, infinite dimensional Hilbert spaces are not needed in this thesis. From now on only finite dimensional Hilbert spaces will be considered. All linear operators are bounded in a finite dimensional Hilbert space,

In a finite dimensional Hilbert space H, every linear operator $\Box$ has an adjoint $\Box^{\Box}$ defined as $\langle \Box(x), y \rangle = \langle x, \Box^{\Box}(y) \rangle$, $\forall x, y \in H$. If $\Box = \Box^{\Box}$, then it is called self-adjoint. If an operator commutes with its adjoint, $\Box\Box^{\Box} = \Box^{\Box}\Box$, then it is called a normal operator. If there exists an operator $\Box^{-1}$ such that $\Box^{-1}(\Box(x)) = x$, $\forall x \in H$, then $\Box^{-1}$ is called the inverse of $\Box$.

When a linear operator acts on a vector, it is usually transformed into a different vector. But it is possible that the operator merely scales the vector by a factor $\lambda \in C$: $\Box(v) = \lambda v$. When this happens, the number λ is called an eigenvalue of the operator $\Box$,

and the vector u is called the eigenvector of Ω associated with the eigenvalue ω. The eigenvalues of self-adjoint operators are special because they are always real numbers:

Theorem 2. Let Ω be a self-adjoint linear operator on a Hilbert space H, with ω one of its eigenvalues, then $\omega \in \mathbb{R}$

Proof. Suppose the eigenvector associated with ω is v, then $\Omega(v) = \omega v$. Which implies $\langle v, \Omega(v)\rangle = \langle v, \omega v\rangle = \omega\|v\|^2$. Because Ω is self-adjoint, $\langle \Omega(v), v\rangle = \langle \omega v, v\rangle = \langle v, \overline{\omega} v\rangle = \overline{\omega}\|v\|^2$. If $\|v\| \neq 0$, $\overline{\omega} = \omega$, which means $\omega \in \mathbb{R}$. □

The inner product structure of a Hilbert space has special significance in physics, so it is important to consider linear operators that preserve this structure. Effectively, a linear operator Ω preserves the inner product (called an isometry) if $\langle \Omega(x), \Omega(y)\rangle = \langle x, y\rangle$, $\forall x, y \in H$. If the isometry is bijective, then it can be shown that $\Omega^{-1} = \Omega^{\dagger}$. Such operators are called unitary operators.

2.2 Postulates of Quantum Mechanics

Having set the stage for the postulates of quantum mechanics, it is time to actually introduce them. The postulates are a set of statements "pairing" physical objects and processes with abstract notions of Hilbert spaces, giving an explicit recipe for doing physical calculations in Hilbert spaces. Different textbooks give different numbers of postulates, either 4 [NC00] [Sha94] or 5 [DM05] or 6 [CTDL97]. Some textbooks do not list the postulates explicitly at all, mixing the mathematical background with physical concepts from the beginning [Sze04] [SN10] [Wei12]. Even though the number of postulates differ from textbook to textbook, all the postulates, either explicitly listed or implicitly blended into the background, cover three aspects of a physical theory: what a physical system is, how to probe a physical system to extract information, and what is the dynamics governing the evolution of a physical system. The postulates listed here will mainly follow the standard textbook of quantum information [NC00], while only working with pure states.

Postulate 1. The state space. A physical system S is represented by a (complex) Hilbert space H. A state of the system at a given instant is given by a vector in this Hilbert space. Such a state vector is called a ket, denoted by $|\ i\rangle$. A composite physical system is represented by the tensor product of its components. Kets of the composite system correspond to linear combinations of tensor products of its components.

There are a few subtleties in this postulate. First of all, states of the physical system and vectors in the corresponding Hilbert space are not mapped one-to-one. A vector, together with all its scalar multiples, corresponds to the same state. To reduce such arbitrariness, all state vectors are supposed to be normalized: $\| |\ i\| = 1$. The second subtlety arises from this normalization requirement. Only requiring the norm of a ket to be 1 does not completely remove the arbitrariness, because any complex number with norm 1 multiplied with a normalized ket will still give a normalized ket. Because of this, a complex number with norm 1 is usually called a phase, and two kets different only by a global phase are undistinguishable.

By Theorem 1, to every ket $|\ i\rangle$ there is a corresponding linear functional, called a bra, denoted by $\langle\ |$. The normalization requirement can be restated as $|\langle\ |\ i\rangle|^2 = 1$.

Normalized kets which are also orthogonal are ideal tools to translate abstract notions in a Hilbert space into concrete terms by forming an orthonormal basis. In quantum information, a qubit is a 2 dimensional Hilbert space with basis kets $|0\rangle$ and $|1\rangle$, where $\langle 0|1\rangle = 0$. Any vector in the qubit space can be written as $|\ i = \alpha|0\rangle + \beta|1\rangle$, with $|\alpha|^2 + |\beta|^2 = 1$, to keep $|\ i$ normalized. The qubit is said to be in superposition of the state $|0\rangle$ and $|1\rangle$, because the coefficients α and β contain information about the qubit, which can only be accessed by the measurement procedure described in the next postulate. Historically, in physics, the qubit describes the state space of a spin-$\frac{1}{2}$ particle,

and is sometimes called a 2-spinor. The connection between the qubit, complex geometry and the spin will be explored in Chatpter 3.

Postulate 2. Measurements. Measurements of a physical system are described by a set of measurement operators M_i, acting on the state space of the system. Each measurement operator represents a possible outcome of the measurement, indexed by i. For a state $|\psi\rangle$, the probability $p(i)$ that the outcome i occurs is given by $p(i) = \langle\psi|M_i^\dagger M_i|\psi\rangle$. After the outcome i has been observed, the state $|\psi\rangle$ is transformed into $\frac{M_i|\psi\rangle}{\sqrt{\langle\psi|M_i^\dagger M_i|\psi\rangle}}$.

The measurement postulate has profound implications in both physics and philosophy. It acts as the bridge between the macroscopic observed world and the tiny quantum mechanical systems. The probabilistic interpretation of measurement results gives most of the non-classical features of quantum mechanics, such as nonlocality and contextuality. Recently, there are theories which take a probabilistic interpretation of the observed world as an axiom, combining it with other physically motivated axioms, to construct a whole family of physical theories (called generalized probabilistic theories), some even go beyond the normal quantum theory defined by these postulates [Har01, Spe07, Bar07, PPK+09].

Returning to the measurement postulate, there are a few hidden assumptions. First of all, since measurement outcomes are only observed probabilistically, the probability of observing one outcome, from the complete set of possible outcomes, when a measurement is performed must be 1. This implies

$$\sum_i p(i) = \sum_i \langle\psi|M_i^\dagger M_i|\psi\rangle = 1, \tag{2.1}$$

$$\sum_i M_i^\dagger M_i = I, \tag{2.2}$$

The second assumption comes from the fact that states are rays in the state space, meaning $|\psi\rangle$ and $c|\psi\rangle$ represent the same physical state. This may pose a problem for (2.1), because if the state is not normalized, the probability may be bigger than 1. This is one of the reasons why only normalized states are used.

The measurement postulate looks different from the usual way it is written in physics textbooks because it uses positive operator-valued measure (POVM). A POVM is given by a set of POVM elements $\{E_i\}$, defined by $E_i = M_i^\dagger M_i$. The only requirements for a POVM measurement is a set of positive operators, the POVM elements, with the property that $\sum_i E_i = I$.

The traditional formalism to model quantum measurements are called projective measurements, which can be seen as a special case of POVM measurements. A projective measurement is described by a Hermitian operator on the state space, called an observable. In addition to being Hermitian, an observable M is a sum of projectors P_m:

$M = \sum_m p_m P_m$, where $p_m > 0$, $\sum_m p_m = 1$. For an observable M, the Hermiticity means that $M^\dagger M = M^2$, and projectors P_m satisfy $P_m^2 = P_m$. So the measurement postulate can be simplified when the measurement is projective: the POVM elements of a projective measurement are the projectors, the projector P_m projects the state $|\ \rangle$ onto its the eigenspace with probability p_m, yielding an eigenvalue as the outcome of the measurement, with the state after measurement being the eigenvector associated with the eigenvalue. According to Theorem 2, the eigenvalues of Hermitian operators are always real. The POVM elements of a projective measurement are P_m themselves. There is one property of projective measurements not possessed by a general POVM measurement: the eigenvectors of Hermitian operators are orthogonal, projective measurements model the "either-or" situation: if there are 2 possible outcomes (represented by 2 mutually orthogonal eigenvectors of some Hermitian operator), each having a non-zero probability to occur, then if one if them is not observed by a measurement, the other one must have been observed. This is a very important property for the Hardy paradox, to be introduced later in this chapter and generalized in Chapter 4.

Because after a projective measurement, the state is transformed into an eigenvector of the projector, it is sometimes more convenient to directly use the eigenvector of a projector to model a projective measurement. For example, for a state $|\ \rangle$ and an observable $M = |\square\rangle\langle\square| - |\square^\perp\rangle\langle\square^\perp|$, the probability of obtaining the result $+1$ is given by

$$p_{+1} = \langle\ |\square\rangle\langle\square|\ \rangle = |\langle\square|\ \rangle|^2. \tag{2.3}$$

In projective measurements, it is also useful to know the expectation value of an observable instead of the probability of obtaining each of its outcomes. The expectation value of an observable M for the state $|\ \rangle$ is given by

$$E(M) = \langle M\rangle = \langle\ |M|\ \rangle. \tag{2.4}$$

Postulate 3. The evolution of a closed system. The evolution of a closed quantum system is described by a unitary transformation U. The state $\rho(t)$ at time t evolves to $\rho(t') = U\rho(t)U^\dagger$ at time t', with U depending on t and t' only.

This postulate, although extensively used by the quantum information community, is only half of the picture: it only deals with discrete time steps. In continuous time, the Schrödinger equation

$$i\hbar\frac{d|\ \rangle}{dt} = H|\ \rangle \tag{2.5}$$

governs the evolution of the system. The object H is called the Hamiltonian of the system, which completely determines the dynamics of the system. The Hamiltonian is a

Hermitian operator, whose eigenvectors are called the energy eigenstates of the system, with the associated eigenvalues called the energy. The Hamiltonian and the unitary operator in the discrete time case are related through an exponentiation function:

$$U(t, t') = e^{-\frac{iH(t'-t)}{\hbar}}. \tag{2.6}$$

2.3 Scratching the Surface of Quantum Entanglement

The linearity of quantum states has many important consequences. The most important of which is entanglement. The term, coined by Schrödinger, was originally used to describe the kind of quantum states used in the EPR thought experiment (cf. the next section). In the modern language, entanglement characterizes the ability to factorize the Hilbert space of a multiparty state into tensor products of smaller Hilbert spaces [GT09, HHHH09]. For bipartite pure states, being entangled means that a state $|\psi\rangle \in \mathcal{H}$ can not be decomposed as

$$|\psi\rangle = |\phi\rangle_1 \otimes |\phi\rangle_2, \tag{2.7}$$

where $|\psi\rangle_1 \in \mathcal{H}_1$, $|\psi\rangle_2 \in \mathcal{H}_2$ and $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. When a state can be written as (2.7), it is called a product or separable state.

For mixed states the definition is similar but slightly more subtle. Because mixed states allow convex combinations of terms like (2.7), the direct analogy of (2.7) leads to the definition of a mixed product state:

$$\rho = \rho_1 \otimes \rho_2, \tag{2.8}$$

while the convex combinations of (2.8) give the definition of a mixed separable state:

$$\rho = \sum_i p_i \rho_1^i \otimes \rho_2^i, \tag{2.9}$$

with $\sum_i p_i = 1$.

The simple forms of the definitions above can be misleading. For bipartite pure states, a simple procedure called the Schmidt decomposition can be used to determine whether a state is entangled or not [Per93, NC00]. Any bipartite pure state $|\psi\rangle$ can be decomposed as:

$$|\psi\rangle = \sum_i s_i |\phi_1^i\rangle |\phi_2^i\rangle, \tag{2.10}$$

where $|\phi_1^i\rangle$ and $|\phi_2^i\rangle$ are orthonormal bases of the first and second subsystem. The state $|\psi\rangle$ is a product state iff there is only one s_i, i.e. it can always be written as the tensor product of two vectors. Simple as it is, this procedure only works for pure states and does not generalize to more than two parties.

For bipartite mixed states, determining whether a state is entangled or not is a very hard problem. Although the number of p_i in (2.9) can be bounded by the Caratheodory theorem [Car11, HHH97, VP98], they bear no obvious relationship to the eigenvalues of ρ^i. Several criteria have been proposed to test the separability of bipartite mixed states.

These include the PPT (positive partial transpose) criterion [Per96], the not completely positive map criterion [HHH96] and the realignment criterion [CW03, Rud05] (for more details, see [GT09, HHHH09]).

For multipartite states, both pure or mixed, entanglement gets more complicated because of the larger Hilbert space and the possibility of having partial entanglement, where not all n parties are entangled.

Despite these difficulties in detecting entanglement, entangled states are a cornerstone of quantum information. All truly quantum protocols or algorithms use entangled states, and different entangled states have different properties, thus suitable for different tasks. For quantum information processing purposes, entanglement can be measured and manipulated, yielding quantitative results relating specific states and information processing tasks such as communication or teleportation. The study of how to quantify entanglement yields different entanglement measures (for a review see [PV07]).

Before talking about entanglement measures, it is instructive to consider a related question first: how to model a general information processing task? Although according to Postulate 3, all quantum states must undergo unitary evolution, the "states" considered maybe those of a composite system. This composite system may consist of two parts: the actual system on which the information processing task is performed, and the environment, which may be entangled with the actual system. When considered this way, the evolution of the actual system, when studied alone, may no longer be unitary (although the system+ the environment always go through unitary evolution). The evolution of the system of interest, after discarding the environment, is called a quantum operation. Mathematically, quantum operations are modeled by completely positive (CP) maps of two kinds, depending on the desired outcome of the quantum operation:

- If after the quantum operation, the state of the system is changed deterministically into another state, the quantum operation is characterized by a completely positive trace-preserving (CPTP) map. The CPTP map is also called a quantum channel.

- If after the quantum operation, the state of the system is changed probabilistically into another state, the quantum operation is characterized by a completely positive trace-nonincreasing map.

In both cases, the quantum operation is modeled by Kraus operators K_i, satisfying the conditions below:

- For CPTP maps, $\sum_i K_i^\dagger K_i = I$.
- For CP trace-nonincreasing maps, $\sum_i K_i^\dagger K_i \leq I$.

In both cases, for a state ρ, the quantum operation denoted by Kraus operators K_i takes it to

$$\rho \to \rho' = \sum_i \frac{K_i \rho K_i^\dagger}{\mathrm{Tr}(K_i \rho K_i^\dagger)}. \tag{2.11}$$

The model local operations and classical communication (LOCC) was proposed to measure the usefulness of entangled states. In this model, take the bipartite case as an example, Alice and Bob each possesses half of an entangled state. Then they are free to perform any quantum operation on their own state. In LOCC, the only restriction on the quantum operation is that it must be deterministic (i.e. a quantum channel). Then they are allowed to send each other classical messages, and depending on received messages, perform more quantum operations (which must also be deterministic), and so on. Mathematically, each round of communication-operation can be modeled by (assuming Alice sends a message to Bob first)

$$I_A \otimes K_B^{a_1}((K_A \otimes I_B)\rho(K_A^\dagger \otimes I_B))I_A \otimes K_B^{\dagger a_1}, \tag{2.12}$$

where $K_B^{a_1}$ is the Kraus operator of the quantum operation of Bob upon receiving the message a_1 from Alice.

From (2.12) it can be seen that the mathematical structure of a multi-round LOCC operation is quite complicated. Luckily, it is necessary to consider Kraus operators which can be decomposed into a product form, with each component acting only on Alice and Bob.

Because LOCC models a general quantum information processing task, an interesting problem arises when pairing specific quantum states to specific tasks. For some tasks such as teleportation, it is known that some states can perform this task. Is it possible that other states can be transformed, through LOCC, to these known states, thus being able to perform the task as well? This is the motivation for studying LOCC equivalence of entangled states. LOCC equivalence means that a state ρ can be turned into another state σ and back deterministically, by applying a series of local quantum operations which also depend on the classical message received from other parties. It turns out that the requirement two states can be turned into each other via LOCC is very strong, making LOCC equivalence the same as local unitary equivalence for pure states [BPR+00]: two n party pure states (of qubits) $|\psi\rangle$ and $|\phi\rangle$ are LOCC equivalent iff there exist $A_1^{a_1}, \ldots, A_n^{a_n} \in SU(2)$ such that

$$|\phi\rangle = A_1^{a_1} \otimes \ldots \otimes A_n^{a_n} |\psi\rangle. \tag{2.13}$$

The LOCC equivalence requires deterministic interconversion. For pure states, this requirement can be relaxed to probabilistic interconversion, corresponding to CP trace-

nonincreasing quantum operations, which allows a more coarse-grained classification of states [BPR+00, DVC00]. The only difference in this stochastic LOCC (SLOCC) scheme is instead of local unitary equivalence, all invertible local operations are allowed:

$$|\psi\rangle = A_1^{a_1} \otimes \ldots \otimes A_n^{a_n} |\phi\rangle, \tag{2.14}$$

where $A_1^{a_1}, \ldots, A_n^{a_n} \in SL(2, \mathbb{C})$. The $A_i^{a_i}$ are called invertible local operations (ILOs). Invertibility guarantees the states can be converted to each other via local operations and classical communication.

One of the first applications of SLOCC classification was to show that all 3-qubit pure states belong to two SLOCC classes: one class having the GHZ state as its representative, the other having the W state as its representative [DVC00]. Unfortunately, simple argument counting by the authors of [DVC00] reveal that starting from 4 qubits, the number of SLOCC classes will be infinite. Fortunately, for symmetric states, the situation is better understood and SLOCC classification has a direct geometric meaning (see Chapter 3).

With LOCC, entanglement no longer just represents vague quantum correlations, it is now a resource, being able to assist communication and computational tasks. It is possible to quantify entanglement with the help of LOCC and a few axioms, giving rise to various entanglement measures $E(\rho) : H \to \mathbb{R}^+$, assigning a positive real number to represent the amount of entanglement of a state ρ.

Although different entanglement measures obey different axioms, the most common axioms are:

1. Separable states have zero entanglement: $E(\rho) = 0 \Leftrightarrow \rho \in H_{Sep}$.
2. Local unitaries do not change entanglement: $\rho = U_1 \otimes \ldots \otimes U_n \rho U_1^\dagger \otimes \ldots \otimes U_n^\dagger \Rightarrow E(\rho) = E(\sigma)$.
3. LOCC can not increase entanglement: if $\rho \xrightarrow{LOCC} \sum_i p_i \sigma_i$, $E(\rho) \geq \sum_i p_i E(\sigma_i)$.
4. $E(\rho)$ is convex: $E(\sum_i p_i \sigma_i) \leq \sum_i p_i E(\sigma_i)$.
5. $E(\rho)$ is additive: $E(\rho^{\otimes n}) = nE(\rho)$.
6. $E(\rho)$ is strongly additive: $E(\rho \otimes \sigma) = E(\rho) + E(\sigma)$.

The most interesting entanglement measure for symmetric states is the geometric measure, defined as

$$E_G(|\psi\rangle) = -\log_2(\langle\phi|\psi\rangle^2), \tag{2.15}$$

where $|\phi\rangle$ is the closest product state, with $|\phi\rangle \in H_{Sep}$ and $|\phi\rangle$ maximizes $\langle\phi|\psi\rangle^2$.

It is easy to see that the geometric measure satisfies the first 4 axioms above, although in general it is not additive [WH02]. For symmetric states the geometric measure is interesting because it can be computed efficiently, tanks to a property that it is sufficient to only consider symmetric $|\Box\rangle$, meaning $|\Box\rangle = |\Box\rangle \otimes \ldots \otimes |\Box\rangle$, where $|\Box\rangle$ is a qubit [HKW+09]. The entanglement of symmetric states in terms of the geometric measure has been well studied [HKW+09, AMM10, KWK+10, AMM11, Aul11b, Mar11, RM11], with a recent thesis which can serve as a review of the topic [Aul11a].

2.4 The Facets of Nonlocality

Nonlocality is arguably the most striking manifestation of the quantum mechanical world. The first question about the locality of quantum mechanics was asked in 1935, then the first breakthrough came almost 30 years later in 1964. Since then, the study of nonlocality flourished both theoretically and experimentally. The theoretical studies of nonlocality involves many subtle concepts leading to many different approaches to the problem. This section will give an overview of the three most common approaches to nonlocality and their reconciliation through a common framework, after motivating the problem from a historical perspective.

2.4.1 A Little Bit of History

Radical changes are characteristics of revolutions, and the quantum revolution certainly brought more radical changes than one is willing to immediately accept, even to great minds like Albert Einstein. Using superpositions of states, making probabilistic predictions and assigning a Hermitian matrix rather than a simple real number to any physical quantity, thus making some simultaneous measurements inaccurate have implications not only on the "behind the scenes" mathematical machinery governing the (microscopic) world, but also on the results of macroscopic experiments.

Not long after the mathematical foundations of quantum mechanics have been laid, Einstein, Podolsky and Rosen (EPR) noticed a counterintuitive feature coming from the tensor product structure of composite systems [EPR35]. Because the tensor product is the most general product (the mathematical word is free product) between the Hilbert spaces of each subsystem, it not only contains information about each subsystem, but also information pertaining to both subsystems. EPR considered a two-particle system, whose wave function encodes the distance between them and their total momentum. These two kinds of information can be precisely known simultaneously because their observables commute. The fact that their value can be precisely known, i.e. predicted with certainty (EPR called such quantities elements of reality), also means that their values should represent some real underlying physical quantity. However, if a measurement is made on the position of one particle, the position of the second particle can be deduced immediately without it being disturbed, no matter how far away their are. The same inference can be made if the momentum of one particle is measured. This seemingly blatant violation of special relativity made EPR suggest that quantum mechanics is not complete. Einstein later called this phenomenon spooky action at a distance.

The EPR paper drew little attention from working physicists at the time, most of whom are adherents of the "shut up and calculate" doctrine. According to Google Scholar, of the 11400 citations the paper have today, only about 100 occurred between 1935 and 1970. However, the paper did have one famous critic: Niels Bohr. In a paper

whose title is exactly the same as the EPR paper [Boh35], Bohr introduced the notion of complementarity: some observables are not meant to be measured together. In the EPR scenario, if the position measurement is made on one particle, there is no way to decide which measurement on the second particle would give precise results unless the choice of the first measurement is known. If a wrong choice is made on the second particle, then the first measurement introduces a disturbance, making the second measurement no longer precise. But without communication, neither party knows a wrong choice has been made, thus making the disturbance undetectable.

Einstein did not buy this complementarity argument. To him, space-like separated particles behave independently. Then the debate seemed to fade away. Almost 30 years had passed when another disgruntled quantum mechanic started questioning its foundations. John Bell revisited the EPR paradox and showed that the EPR scenario is incompatible with quantum mechanical predictions [Bel64]. Instead of reasoning about the position and momentum of two particles, Bell, following Bohm and Aharonov, used two spin half particles and restricted the measurements to components of their spin. To model the spooky action, Bell introduced λ, the local hidden varible (LHV).

It does not matter what λ is. It can be a discrete set of values, or a set of functions, or a continuous set of values or functions. It is meant to complete quantum mechanics to explain the correlations which arise from measurements of space-like separated parties. Bell extended the EPR "elements of reality" measurements to general measurements, allowing each party to use λ as a common source of randomness. The local in local hidden variable addresses Einstein's concern that space-like separated parties should behave independently. If the parties are space-like separated, their measurements can still depend on λ. Denote the probability of obtaining the result a when the observable A is measured by $P(a|A)$, and similarly define $P(b|B)$, a local hidden variable theory will predict that the joint probability $P(ab|AB)$ should be the product of $P(a|A)$ and $P(b|B)$, taking λ into account, and averaged over the distribution of λ:

$$P(ab|AB) = \int P(a|A,\lambda)\,P(b|B,\lambda)\,\rho(\lambda)\,d\lambda, \tag{2.16}$$

where $\rho(\lambda)$ is a probability density function of λ, assuming λ is continuous, with the property that $\int \rho(\lambda)\,d\lambda = 1$ (if λ is discrete, it is sufficient to replace the integral with a sum, and $\rho(\lambda)$ becomes a probability mass function).

The definition above can be easily generalized to n parties, measuring observables M_1 to M_n, obtaining results m_1 to m_n:

$$P(m_1,\ldots,m_n|M_1,\ldots,M_n) = \int P(m_1|M_1,\lambda)\ldots P(m_n|M_n,\lambda)\,\rho(\lambda)\,d\lambda. \tag{2.17}$$

There are two assumptions on (2.16) and (2.17):

1. If the parties are space-like separated, the measurement outcomes of one of the observables does not depend on the measurement settings of the other one.

2. The physical quantities represented by these observables have well-defined values, revealed by measuring the corresponding observable.

The first assumption is called the locality assumption, because the parties being measured can be space-like separated, making all actions, including measurements, local. The second assumption is called the realism assumption, because as Einstein put it "I like to think the moon is there even if I am not looking at it". Together they are called local realism. A physical theory assuming the validity of local realism is called a local hidden variable theory (or LHV theory for short).

Bell showed that while for some states the joint probability can be obtained by making local measurements and assuming a hidden variable exists between the measurement outcomes, there are states which when certain observables are measured, give results that are incompatible with (2.16) and (2.17). This incompatibility means quantum mechanics is not locally realistic, at least one assumption in local realism should be abandoned.

The first experiment to test local realism was performed 8 years after the publication of [Bel64], and quantum mechanics is vindicated [FC72]. However, even before the first experiment has been performed, the possibility of experimental loopholes, allowing local realistic explanations of quantum mechanical results, was raised [Pea70]. There are two kinds of loopholes in an experimental test of local realism: the locality loophole arises when the "separate parties" in the experiment are not far enough (i.e. space-like separated); the detection loophole arises when the efficiency of the measurement device is below certain threshold to rule out any LHV theory. The most famous tests of LHV vs. quantum mechanics are performed 9 years after the first test [AGR81, AGR82, ADR82]. The authors are well aware of the potential loopholes in these tests, and tried to address one of them, the locality loophole, in [ADR82] by varying measurement setting quick enough so that no signal about the setting of one party can reach the other party. However, as a result of technological limitations, the detector efficiency is not high enough to close the detection loophole. The locality loophole is not completely closed either because the settings change in a fast but predictable way. Subsequent tests have either closed the locality loophole by using random settings [WJS+98], or closed the detection loophole by using ions instead of photons [RKM+01]. But no experiment to date has closed both loopholes.

2.4.2 Bell's Inequality

The popularity of Bell's work is partly due to the way he proposed to test local realism. In the EPR paper, the whole scenario is a thought experiment. Because position and

momentum are both continuous variables, the EPR thought experiment is difficult to carry out. Bohm reformulated the EPR experiment by using two entangled spin-$\frac{1}{2}$ particles and the quantities measured are x, y, z components of the spin, made by a Stern-Gerlach apparatus [Boh51]. Later, Bohm and Aharonov proposed an EPR test by measuring polarization of photons [BA57]. Although these proposed experiments use discrete variables, there is no clear criterion with which a definitive statement separating quantum mechanical predictions and LHV theories can be made solely from interpreting the experimental data. The pioneering work of Bell, who used the same physical setup made of two entangled spin-$\frac{1}{2}$ particles as Bohm, proposed using an inequality as the criterion to separate quantum mechanical predictions and any LHV theory.

Why an inequality? The main reason is that an inequality allows a theory independent formulation of the LHV condition (2.16) and (2.17) using probabilities or expectation values, which would produce a bound, called the local realistic bound, as the limit of any local realistic theory, without using any quantum mechanical assumption. Then using the postulates of quantum mechanics, these probabilities or expectation values can be associated with various quantum measurements performed on some chosen states, which allow the direct comparison of quantum mechanical predictions with the local realistic bound. If the bound is violated, then these states/ measurements serve as witnesses of the nonlocality of quantum mechanics. The inequality allows direct experimental test because it tolerates errors (whether in state preparation or measurements). As long as the error is not very big the local realistic bound can still be violated.

Different types of inequalities exist for different states/ measurements, using different types of statistical data (probabilities or expectation values). Different inequalities offer different resistance to errors and show different strengths of violation. In what follows three types of approaches to nonlocality are explained in detail. All three approaches use different types of inequalities. In fact, the last two approaches also allow nonlocality proofs with inequalities. But these proofs without inequalities are not resistance to any experimental error, thus the only way of testing nonlocality is still via the inequalities. Because Bell first used inequalities to test LHV theory, these different types of inequalities are all called Bell inequalities.

2.4.3 Nonlocality from Correlations

This is the most famous approach to nonlocality in quantum mechanics. The idea itself comes from the reformulation of the EPR argument by Bohm [Boh51] [BA57]. The original inequality by Bell can be formulated as follows.

Let $a, b, c \in \{+1, -1\}$, then it is easy to verify the following holds:

$$ab - ac = -(1 - bc). \tag{2.18}$$

Also note that in (2.18), if $b = c$, then both sides are 0, otherwise both sides equal to -2.

Now imagine a, b, c are results of tosses of three coins, controlled by some hidden mechanism which may bias the tosses. This means that the probability that a, b, c get $+1$ (heads) or -1 (tails) are not necessarily $\frac{1}{2}$. Nevertheless, let $\langle ab\rangle$ denote the expectation value of the pair of coins (as random variables) which produce results a, b, and similarly define $\langle ac\rangle, \langle bc\rangle$, then it can be shown that

$$|\langle ab\rangle - \langle ac\rangle| \leq 1 + \langle bc\rangle. \tag{2.19}$$

The expression (2.19) is the original Bell inequality. To show (2.19) can be violated by quantum states with quantum measurements, consider the state

$$|\psi\rangle = \frac{|01\rangle + |10\rangle}{\sqrt{2}}, \tag{2.20}$$

which may represent a pair of entangled photons. Now, let a, b, c be the measurement outcomes of these 3 observables:

$$a \to \hat{m}_a = \frac{-\sigma_z - \sigma_x}{\sqrt{2}}, \quad b \to \hat{m}_b = \sigma_z, \quad c \to \hat{m}_c = \sigma_x, \tag{2.21}$$

where the $\sigma_{x,y,z}$ are Pauli matrices:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.22}$$

It can be easily verified that $\hat{m}_a, \hat{m}_b$ and $\hat{m}_c$ are all Hermitian, having eigenvalues $+1$ and -1 with orthogonal eigenvectors. They represent projective measurements. So indeed a, b, c can be model as the outcomes when these observables are measured. According to Postulate 2 of quantum mechanics, the expectation value for projective measurements are given by

$$\langle \hat{m}_a \hat{m}_b\rangle = \langle \psi | \hat{m}_a \hat{m}_b | \psi \rangle, \langle \hat{m}_a \hat{m}_c\rangle = \langle \psi | \hat{m}_a \hat{m}_c | \psi \rangle, \langle \hat{m}_b \hat{m}_c\rangle = \langle \psi | \hat{m}_b \hat{m}_c | \psi \rangle, \tag{2.23}$$

where $\hat{m}_a \hat{m}_b$ means that the measurement $\hat{m}_a$ is performed on the first photon and the measurement $\hat{m}_b$ is performed on the second photon, etc.

Simple calculation shows that

$$\langle \hat{m}_a \hat{m}_b\rangle = \frac{1}{\sqrt{2}}, \quad \langle \hat{m}_a \hat{m}_c\rangle = -\frac{1}{\sqrt{2}}, \quad \langle \hat{m}_b \hat{m}_c\rangle = 0. \tag{2.24}$$

So the LHS of (2.19) is $\sqrt{2}$, while the RHS is 1, a clear violation.

In the reasoning above, the observable $\hat{m}_c$ is shared by both parties. In reality this is very difficult to achieve. In fact, a whole body of research has since been carried out on

how reference frames can be shared by distant parties [BRS07]. Not long after [Bel64] was published, Clauser Horne Shimony and Holt gave an improved version of (2.19), allowing both parties to choose from two independent measurement settings [CHSH69]. Their version, dubbed the CHSH inequality, is the most famous Bell inequality in quantum information.

The CHSH inequality can be derive from an argument similar to the one given above. Instead of having a, b, c, now suppose there are four numbers $a, b, c, d \in \{+1, -1\}$. Instead of (2.18), now consider the expression

$$ac + bc + bd - ad. \tag{2.25}$$

It can be easily shown that (2.25) $= -2$. Using $\hat{m}_a$ to $\hat{m}_d$ to model the measurements, the CHSH inequality is

$$|\langle \hat{m}_a \hat{m}_c \rangle + \langle \hat{m}_b \hat{m}_c \rangle + \langle \hat{m}_b \hat{m}_d \rangle - \langle \hat{m}_a \hat{m}_d \rangle| \leq 2. \tag{2.26}$$

To show (2.26) can be violated by a quantum state with appropriate measurement, again consider the state $|\ \rangle$ given above. The measurement are now given by

$$a \to \hat{m}_a = \frac{-\sigma_z - \sigma_x}{\sqrt{2}}, \quad b \to \hat{m}_b = \frac{-\sigma_z + \sigma_x}{\sqrt{2}}, \quad c \to \hat{m}_c = -\sigma_x, \quad d \to \hat{m}_d = -\sigma_z. \tag{2.27}$$

Again using Postulate 2, the LHS of (2.26) is $|-2\sqrt{2}| = 2\sqrt{2}$, which is bigger than 2.

It turns out the the value $2\sqrt{2}$ is the best value that can be achieved by quantum mechanics in this CHSH scenario, and this value is called the Tsirelson bound after the author who proved its optimality [Cir80]. Although the Tsirelson bound represents what quantum mechanics can do, it is clearly not the maximum possible value for the LHS of the CHSH inequality. If $\langle \hat{m}_a \hat{m}_c \rangle = \langle \hat{m}_b \hat{m}_c \rangle = \langle \hat{m}_b \hat{m}_d \rangle = 1$ and $\langle \hat{m}_a \hat{m}_d \rangle = -1$, then the LHS of (2.26) would be 4. Although this much higher bound can not be achieved by quantum mechanics, there are probability distributions which can achieve this value, without violating special relativity and signaling faster than light [PR94]. The distribution is given by assigning probability $\frac{1}{2}$ whenever the AND of the measurement settings, represented by bits, equals to the XOR of the outcomes, again represented by

bits. All other probabilities are zero. Explicitly, the nonzero probabilities are:

$$P(00|00) = P(11|00) = \frac{1}{2}, \tag{2.28}$$

$$P(00|01) = P(11|01) = \frac{1}{2}, \tag{2.29}$$

$$P(00|10) = P(11|10) = \frac{1}{2}, \tag{2.30}$$

$$P(01|11) = P(10|11) = \frac{1}{2}. \tag{2.31}$$

Such probability distributions are called PR boxes. Since quantum mechanics does not signal faster than light, why is there a gap between the two values? Why such PR boxes do not exist in nature? Is there any physical principle imposed by quantum mechanics that restricts its nonlocality? Such questions lead to results such as information causality [PPK+09], and the realization that any stronger than quantum correlation will make communication complexity trivial [BBL+06] (also see 7.1).

The reason quantum mechanics violates (2.26) stems from the correlations exhibited by the measurements, and the inequality itself consists of sums of correlation terms. For this reason inequalities of the CHSH type are sometimes referred to as inequalities of correlation functions. The interest in the CHSH inequality exploded after the second experiment by Aspect et al., where the CHSH inequality is shown to be experimentally violated [AGR82]. Soon it was shown that all entangled bipartite pure states violate the CHSH inequality (a result known as Gisin's theorem) [Gis91], and all entangled multipartite pure states violate different versions of it [PR92]. Then, the CHSH inequality was generalized to more than two parties [WW01] and more than two measurement settings [Per99]. Although for pure states these results show that nonlocality is very common in quantum mechanics, it is known from early on that mixed states present a greater challenge: there are entangled mixed states which have an LHV model [Wer89]. There are also pure states which do not violate any multiparty extension of CHSH inequality, when choosing from two settings and two outcomes per setting (called having two dichotomic observables per party) [ZBLW02]. However, all pure states, including these states which do not violate any inequality of correlation functions, do violate the inequality P^n presented in later chapters [YCZ+12].

2.4.4 Probability-free Nonlocality & Mermin Inequality

The second kind of nonlocality concerns a special class of states called stabilizer states [Got96, Got97]. These states have the special property that certain combinations of local Pauli operators do not change the state (i.e. the state is stabilized by these

operators). For example, it can be easily verified that the GHZ state,

$$|GHZ\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle), \tag{2.32}$$

is stabilized by

$$g_1 = \sigma_x \otimes \sigma_x \otimes \sigma_x, \tag{2.33}$$
$$g_2 = \sigma_z \otimes \sigma_z \otimes I, \tag{2.34}$$
$$g_3 = I \otimes \sigma_z \otimes \sigma_z, \tag{2.35}$$

meaning $\langle GHZ| g_i |GHZ\rangle = 1$, for $i = 1, 2, 3$. The g_i are not all the operators that stabilize $|GHZ\rangle$, they are called the generators of the stabilizer subgroup. Other elements of the group are obtained by multiplying the generators. For $|GHZ\rangle$, there are 8 elements in the stabilizer group. In addition to the three generators and the identity $I = I \otimes I \otimes I$, the other four elements are:

$$g_1 g_2 = \sigma_y \otimes \sigma_y \otimes \sigma_x, \tag{2.36}$$
$$g_1 g_3 = \sigma_x \otimes \sigma_y \otimes \sigma_y, \tag{2.37}$$
$$g_2 g_3 = \sigma_z \otimes I \otimes \sigma_z, \tag{2.38}$$
$$g_1 g_2 g_3 = \sigma_y \otimes \sigma_x \otimes \sigma_y. \tag{2.39}$$

More generally, a n qubit state $|\ \rangle$ is stabilized by a set of 2^n stabilizers s_i if for all s_i, s_j $|\ \rangle = |\ \rangle$. The stabilizers of interest in quantum information are tensor products of Pauli operators, and the state $|\ \rangle$ can be uniquely fixed by its stabilizers s_i (up to global phase) if the stabilizers commute and the state is not the zero vector ($-I$ is not a stabilizer) [NC00].

The stabilizer formalism is an elegant and powerful tool, with applications in error correction [Got97], measurement based quantum computing [RB01, RBB03, BKMP07, DKP07], entanglement theory [HEB04, HDE+06], secret sharing [MS08] and as will be explained below, nonlocality [Mer90, GTHB05, CGR08].

In the definition of local realism, there is the realism assumption that all physical quantities should have well-defined values, revealed by measurement of the corresponding observables. Greenberger, Horne and Zeilinger found a counterexample using the state named after them [GHZ89]. Suppose there are three players, each can measure three physical quantities, called X, Y, Z, which when measured give values $+1$ or -1,

and the measurements satisfy the four equations below:

$$X_1 \otimes X_2 \otimes X_3 = 1, \tag{2.40}$$

$$X_1 \otimes Y_2 \otimes Y_3 = -1, \tag{2.41}$$

$$Y_1 \otimes X_2 \otimes Y_3 = -1, \tag{2.42}$$

$$Y_1 \otimes Y_2 \otimes X_3 = -1, \tag{2.43}$$

where subscripts denote players.

It is not hard to see that the four equations can not be satisfied all at once: take the product of last three equations yields

$$Y_1^2 \otimes X_1 \otimes Y_2^2 \otimes X_2 \otimes Y_3^2 \otimes X_3 = -1. \tag{2.44}$$

Since $X_1 \otimes X_2 \otimes X_3 = 1$, this implies $Y_1^2 \otimes Y_2^2 \otimes Y_3^2 = -1$, which is impossible because none of them is complex.

From here, it is clear that $|GHZ\rangle$, with some of its stabilizers, can satisfy all four equations if instead of assuming X, Y, Z are numbers equal to $+1$ or -1, observables which are a Pauli matrices, are used. Because the stabilizers all commute, their measurements can be made simultaneously, thus making the measurement outcomes elements of reality.

To see how to turn this nonlocality without probability into an inequality, a similar trick used in the CHSH inequality can be employed: instead of talking about products of numbers taking discrete values, consider the average of these products over the distribution of the hidden variable, for which a local realistic bound can be calculated. Using $\langle X_1 \otimes X_2 \otimes X_3 \rangle$ to denote the expectation value of the product $X_1 \otimes X_2 \otimes X_3$ over the hidden distribution (and similarly for the other products), the following inequality can be established

$$\langle X_1 \otimes X_2 \otimes X_3 \rangle - \langle X_1 \otimes Y_2 \otimes Y_3 \rangle - \langle Y_1 \otimes Y_2 \otimes X_3 \rangle - \langle Y_1 \otimes X_2 \otimes Y_3 \rangle \tag{2.45}$$

$$- X_1 \otimes X_2 \otimes X_3 - X_1 \otimes Y_2 \otimes Y_3 - Y_1 \otimes X_2 \otimes Y_3 - Y_1 \otimes Y_2 \otimes X_3 \tag{2.46}$$

$$\leq 3, \tag{2.47}$$

where the number 3 is calculated assuming $X, Y, Z \in \{\pm 1\}$.

The GHZ state can violate this bound when Postulate 2 is used to compute the

expectation values:

$$\begin{cases} \langle X_1 \otimes X_2 \otimes X_3 \rangle \to \langle GHZ | \sigma_x \otimes \sigma_x \otimes \sigma_x | GHZ \rangle = 1, \\ \langle X_1 \otimes Y_2 \otimes Y_3 \rangle \to \langle GHZ | \sigma_x \otimes \sigma_y \otimes \sigma_y | GHZ \rangle = -1, \\ \langle Y_1 \otimes Y_2 \otimes X_3 \rangle \to \langle GHZ | \sigma_y \otimes \sigma_y \otimes \sigma_x | GHZ \rangle = -1, \\ \langle Y_1 \otimes X_2 \otimes Y_3 \rangle \to \langle GHZ | \sigma_y \otimes \sigma_x \otimes \sigma_y | GHZ \rangle = -1, \end{cases} \tag{2.48}$$

$$\Rightarrow \langle X_1 \otimes X_2 \otimes X_3 \rangle - \langle X_1 \otimes Y_2 \otimes Y_3 \rangle - \langle Y_1 \otimes Y_2 \otimes X_3 \rangle - \langle Y_1 \otimes X_2 \otimes Y_3 \rangle = 4. \tag{2.49}$$

Both the reasoning and the inequality above can be directly generalized to n party GHZ states

$$|GHZ_n\rangle = \frac{1}{\sqrt{2}}(|0\ldots 0\rangle + |1\ldots 1\rangle), \tag{2.50}$$

whose generators of of the stabilizer subgroup are analogues to the generators of $|GHZ\rangle$, with $g_1 = \sigma_x \otimes \sigma_x \otimes \ldots \otimes \sigma_x$, the rest being tensor products of two σ_z and the identity on all other parties. The generalized n party inequality, called the Mermin inequality [Mer90], whose LHS is given by:

$$\langle X_1 \otimes \ldots \otimes X_n \rangle \leq \sum \langle Y_1 \otimes Y_2 \otimes X_3 \otimes \ldots \otimes X_n \rangle + \sum \langle Y_1 \otimes Y_2 \otimes Y_3 \otimes Y_4 \otimes X_5 \otimes \ldots \otimes X_n \rangle \ldots, \tag{2.51}$$

where each term except the first one contains an even number of Y and are permutations of other terms with the same number of Y, with the sign changes depending on the number of Y is a multiple of 2 or a multiple of 4. The local realistic bound for n even is $2^{\frac{n}{2}}$, while for odd n it is $2^{\frac{n-1}{2}}$. The violation of the Mermin inequality grows exponentially with n, because the number of terms in (2.51) is 2^{n-1}, with all of them contributing +1 to the sum.

Although (2.51) seems to be tailor-made for the GHZ state, giving an exponential violation, it is still not the optimal, because there are "only" 2^{n-1} terms in the sum, if the number of terms goes to 2^n there should be an even bigger violation. It turns out that it is possible to increase the number of terms to 2^n, if instead of 2 settings, 3 settings are used, with the first party measures observables that are the ones used for the CHSH inequality $\frac{\sigma_x - \sigma_y}{\sqrt{2}}$ [Ard92, BK93]. Because the measured observables are longer stabilizers of the GHZ state, each term in the sum equals to $\frac{1}{\sqrt{2}}$ instead of 1, which means this inequality would give a violation of $(2^n - \frac{1}{\sqrt{2}}) > 2^n - \frac{1}{2} = 2^{n-1}$. This type of inequality, is called the MABK (Mermin-Ardehali-Belinskii-Klyshko) inequality.

The most direct generalization of probability-free nonlocality is to graph states (which are equivalent to stabilizer states up to local unitaries), with 3 settings (the 3 Pauli matrices in the stabilzers) and two outcomes per setting (± 1) [GTHB05, GC08]. The Mermin inequality, which only uses 2 settings, has been generalized to all graph

states [TGB06], but only some graph states show an exponential violation [CGR08]. Experiments have been performed to test these inequalities [GYX+10] [CVDM+09]. But because of the perfect correlation exhibited by the measurement in these inequalities, detector efficiency in these experiments must be higher than $\frac{n}{2n-2}$ to rule out any LHV theory [CRV08].

2.4.5 Almost Probability-free Nonlocality & Hardy Paradox

Nonlocality from correlations and probability-free nonlocality represent two extremes: expectation value containing correlations must be built from statistics of many experimental runs, while probability-free only use perfect correlations. The third type, which can be seen as a mixture of the other two approaches, was proposed by Hardy in early 90s as a logical paradox first, then put into an inequality [Har93, Har94]. This inequality is the first inequality using probabilities instead of expectation values. It is first derived by the first half of CHSH [CH74], and experimentally tested by the first and third (the most famous) experiments of Aspect et al [AGR81, ADR82].

To see what the Hardy paradox is and why it is a logical paradox at all, consider two people, Alice and Bob, each can choose from two types of sealed boxes given to them individually from a referee, Charlie. One type of box contains food, upon opening, it will reveal either a baguette or a bowl of noodles. The other type contains drinks, upon opening, either a cup of coffee or a cup of tea can be found. Charlie made the following promises to Alice and Bob with regards to the type of boxes they choose to open and the contents they will find inside:

1. If they both choose the food, then it is possible that they both get baguettes.
2. If Alice chooses the food box and gets a baguette, then if Bob chooses the drink box he will never get a cup of tea.
3. If Bob chooses the food box and gets a baguette, then if Alice chooses the drink box she will never get a cup of tea.
4. If they both choose drinks, they will never both get coffee.

After the rules and promises are made clear, Alice and Bob proceed to open the food boxes and found two baguettes inside. Now what can they deduce from this discovery?

Promise number 1 allows both of them to have baguettes, so it is not broken. From promise 2 Alice can deduce that since she has a baguette, had Bob chose to open the drink box he would have found a cup of coffee for sure. Bob makes a similar inference from promise 3: had Alice chose to open the drink box, she would have found a cup of coffee for sure. Together they conclude if they had opened the drink boxes, both of them would had a cup of coffee, which promise 4 says is impossible.

Why does the paradox occur? Partly because in this example, the boxes are supposed to contain real things, things that exist even if the boxes are sealed. And because the boxes must contain one of two things, it is possible to infer with certainty what is not in the box, given the thing in the box. The promises can be translated into probabilities by using upper case letters to represent types of boxes (F= food, D= drink) and lower case letters as contents of boxes (b= baguette, n= noodles, c= coffee, t= tea):

$$P(b, b|F, F) > 0 \tag{2.52}$$

$$P(b, t|F, D) = 0 \tag{2.53}$$

$$P(t, b|D, F) = 0 \tag{2.54}$$

$$P(c, c|D, D) = 0. \tag{2.55}$$

Hardy showed [Har93] that almost all bipartite entangled states can satisfy all four probabilities together, given the right choice of measurements. The only states for which there are no measurements to satisfy all four probabilities are the maximally entangled states. But these states maximally violate the CHSH inequality.

The four probabilities are exactly the ones found in the inequality proposed by Clauser and Horne in [CH74] (known as the CH inequality):

$$P(b, b|F, F) \square P(b, t|F, D) \square P(t, b|D, F) \square P(c, c|D, D) \square 0. \tag{2.56}$$

The violation in quantum mechanics is immediate since almost all quantum states, with appropriate measurements, can satisfy all four probabilities (2.52) to (2.55), thus making the LHS of (2.56) positive.

It is also clear that the algebraic maximum of (2.56) is 1, achieved when $P(b, b|F, F) = 1$ and all others 0. Can a quantum state achieve this? If not, can any nonsignaling distribution achieve this? Mermin gave negative answers to both questions in [Mer95], where he showed that achieving this algebraic maximum leads to faster than light signaling, so no nonsignaling theory, not quantum mechanics, not even PR boxes, can achieve the maximum.

In later chapters, both the Hardy paradox and the CH inequality will be extended to n parties. The geometrical meaning of the Hardy paradox, both in this two party case and the later n party extension, will be explored as well.

2.4.6 Unification of Different Approaches to Nonlocality

Unlike entanglement, which stems from the intrinsic properties of quantum states, nonlocality is a phenomenological property. In principle, the definition (2.17) does not refer to any explicit physical theory. As long as a theory provides a means to assign probabilities to phenomenological events, together with basic rules of probability, nonlocality can be defined. In fact, this is the approach taken by Popescu and Rohrlich in the invention of PR boxes: take nonlocality as an axiom, then use the resulting probability distributions to see if a physical theory can arise from these distributions. In the PR box case, even though the distribution does not allow faster than light signaling, quantum mechanics can never achieve this distribution. On the other hand, as shown earlier in this section, quantum mechanics provides many ways to nonlocality, with each approach suitable for some state/ measurement combinations. Thus a unified approach to nonlocality is useful both as a tool to study nonlocality in quantum mechanics, and as part of an investigation of general, possibly post-quantum, physical theories which can be constructed from a few axioms and share some characteristics with quantum theory [Har01, MAG06, Spe07, Bar07].

The most natural way to a unified approach to nonlocality is to consider probability distributions as vectors in a real Euclidean space. The idea itself dates back to the 80s [Pit89, Pit94], but went through a recent revival when convex geometry was used to facilitate the study of such vector spaces [BLM+05].

The most basic object in convex geometry is the convex set K, which is a subset of R^n with the property that the closed line segment defined by an arbitrary pair of distinct points lies completely within the set [Zie95, Grü03]. Equivalently, a set $K \subset R^n$ is convex iff $\forall a, b \in K$, if $0 \leq \lambda \leq 1$, then $\lambda a + (1-\lambda) b \in K$. From this definition it is easy to see that the empty set $\emptyset$ and R^n itself are convex. The intersection of convex sets is also convex. Furthermore, for any $a_i \in K$ and $\lambda_i \in R^+$, $i = 1, \ldots, n$, if $\sum_i \lambda_i = 1$, then $A = \sum_i \lambda_i a_i \in K$. The element A is called the convex combination of a_i. Because convex geometry will be used to model vectors of probabilities, it is simpler to only consider closed convex sets.

The intuition for a convex set is a geometric figure which does not have any dent. The vertices of a convex set, which do not lie within any open segment in the convex set,are called its extreme points. More rigorously, a point x is an extreme point of a convex set K if $\forall a, b \in K$ and $0 < \lambda < 1$, $x = \lambda a + (1-\lambda) b$ implies $x = a = b$. For any subset S of R^n, it is useful to consider the smallest convex set which contains S, called the convex hull of S. The convex hull K of S is defined as the intersection of all convex subsets of R^n which contain S. The convex hull of a finite number of points in R^n is called a polytope.

In addition to defining a polytope as the convex hull of its vertices, an equivalent definition can be given by using linear inequalities. A linear equation $u \cdot x = c$, where

$u \in R^n$, $c \in R$ defines a hyperplane in R^n, cutting the whole space in half. The inequality $u \cdot x \leq c$ then defines a half space. A polytope in this picture is the bounded intersection of a finite number of half spaces, described by linear inequalities. An inequality of the form $u \cdot x \leq c$ is valid for a polytope P if it is satisfied for all $u \in P$. A face of P is defined to be the intersection of P with any linear equation $u \cdot x = c$ where $u \cdot x \leq c$ is valid for P. Faces may have different dimensions, defined by the dimension of the half space of the linear equation. For a polytope of dimension n, 1 dimensional faces are the points in the convex hull definition, so they are still called vertices, while 2 dimensional faces are called edges and $n-1$ dimensional faces are called facets.

The probability of obtaining measurement outcomes $m_1, \ldots, m_n$ when observables $M_1, \ldots, M_n$ are measured, denoted by $P(m_1, \ldots, m_n | M_1)$, can be treated as a vector of dimension 2^{2n}, having components $(m_1, \ldots, m_n, M_1, \ldots, M_n)$. But normally these 2^{2n} components are not all independent, because to model different physical situations, different constraints apply.

The constraint modeling LHV theory is just (2.17). A weaker constraint, called nonsignaling, is used to model the physical limitations imposed by special relativity: information must travel at a finite speed, thus distant parties can not influence the measurement choices and outcomes of each other. In terms of probabilities, this means that

$$\begin{aligned} & \sum_{m_i} P(m_1, \ldots, m_i, \ldots, m_n | M_1, \ldots, M_i, \ldots, M_n) \\ = & \sum_{m_i} P(m_1, \ldots, m_i, \ldots, m_n | M_1, \ldots, M_i', \ldots, M_n) \\ = & P(m_1, \ldots, m_{i-1}, m_{i+1}, \ldots, m_n | M_1, \ldots, M_{i-1}, M_{i+1}, \ldots, M_n), \end{aligned} \tag{2.57}$$

for all choices M_i and M_i' for all parties i.

With these constraints, the probabilities $P(m_1, \ldots, m_n | M_1, \ldots, M_n)$ form polytopes. If the probabilities satisfy the locality constraint (2.17), then the polytope is called the local polytope, if they satisfy the nonsignaling constraint (2.57), the polytope is called the nonsignaling polytope. Because the nonsignaling constraint is weaker, i.e. all local probability distributions are nonsignaling, the local polytope sits inside the nonsignaling polytope. The dimensions of these polytopes have to be decided by the particular probability distribution under consideration. For example, the distribution which gives rise to the CHSH inequality can be represented by a 8 dimensional polytope [BLM+05].

Why does the use of polytopes allow a unified framework to study nonlocality? Because every facet of the local polytope correspond to a linear equation whose constant is the local realistic bound. In other words, the half space defined by the facet which contains the polytope itself corresponds to a Bell's inequality. The vertices of the polytope are extremes of its defining probability distribution.

But quantum mechanical measurements, even with fixed number of settings/ outcomes, still do not have a finite number of extreme points, a probability distribution arising from quantum mechanical measurements form a convex set, but not a polytope. Also because quantum mechanics violates Bell's inequalities, the convex set is a strict superset of the local polytope. But quantum mechanics can not reach the PR box violation of the CHSH inequality, so it is a proper subset of the nonsignaling polytope.

In recent years, several new techniques to unify different approaches to nonlocality are proposed. These techniques not only unifies different approaches to nonlocality, they also put nonlocality and contextuality (which can be seen as a more general notion than nonlocality) in the same framework.

The technique used by Abramsky and Brandenburger [AB11] puts nonlocality and contextuality into the same categorical language using sheaf theory. The conditions of nonlocality and contextuality can be rephrased in terms of obstruction to the existence of global sections. They also established a hierarchy of different approaches to nonlocality. Using a concept they call strong contextuality, they showed that Mermin> Hardy> Bell.

Another recent technique, although also covered briefly by the two authors above, uses hypergraphs to characterize nonlocality and contextuality [CSW10, FLS12]. In this framework it is easy to write the proofs of contextuality as conditions on a graph. And the questions about nonlocality and contextuality can be phrased in terms of graph properties. For example, it is shown that the Lovasz □-function, which was originally proposed to bound the Shannon capacity of a graph, actually provides a way to compute the maximum quantum violation of contextuality inequalities.

2.5 Semidefinite Programming

Semidefinite programming was developed in the 90s as a tool to study convex optimization problems [VB96]. The method has been adapted in early 2000s as a way to numerically find the global extrema of a real-valued polynomial [Las01]. Also around this time, the study of multiparty nonlocality produced increasingly complex results, which made it hard to obtain analytical properties about various multiparty Bell inequalities. As a result, numerical studies about the optimality and violations of these inequalities began to emerge [WW01] [PS01]. In 2006, Wehner [Weh06] used SDP as an analytical tool to both prove the original Tsirelson bound for the CHSH inequality and to find new bounds for the generalized CHSH inequality with n settings and 2 outcomes per setting. Since then, SDP has been employed as a numerical tool to study various aspects of multiparty entanglement and features of multiparty nonlocality, for example in [NPA07] [BPA+08]. A recent paper [BSV12] used SDP to show that one can distinguish two different classes of entangled states based on violations of Bell inequalities.

For our purposes, we employ a similar technique to the one used in [BSV12]. Since, without loss of generality, we only use projective measurements [WW01] and probabilities instead of expectation values, the measurement operator we use is different. Suppose each player i can measure either one of two bases and obtain either one of two possible outcomes. We model these four different situations by four measurement operators:

$$\begin{aligned}
M_i^{00} &= \frac{1}{2}(I_i + \square_{i0}\square_x^i + \square_{i0}\square_y^i + \square_{i0}\square_z^i) \\
M_i^{01} &= \frac{1}{2}(I_i \square \square_{i0}\square_x^i \square \square_{i0}\square_y^i \square \square_{i0}\square_z^i) \\
M_i^{10} &= \frac{1}{2}(I_i + \square_{i1}\square_x^i + \square_{i1}\square_y^i + \square_{i1}\square_z^i) \\
M_i^{11} &= \frac{1}{2}(I_i \square \square_{i1}\square_x^i \square \square_{i1}\square_y^i \square \square_{i1}\square_z^i),
\end{aligned} \tag{2.58}$$

where M_i^{jk} denotes the player i chooses to measure in basis j and obtains the outcome k, and $v_i^0 = (\square_{i0}, \square_{i0}, \square_{i0})$, $v_i^1 = (\square_{i1}, \square_{i1}, \square_{i1})$ are two unit vectors in R^3.

Now we can write the probabilities in P_n using these single-qubit measurement

operators:

$$\begin{aligned}
P(0\ldots0|0\ldots0) &= \mathrm{Tr}(\rho M_1^{00} \otimes \ldots \otimes M_n^{00}) \\
P(0\ldots0|0\ldots1) &= \mathrm{Tr}(\rho M_1^{00} \otimes \ldots \otimes M_n^{10}) \\
&\vdots \\
P(0\ldots0|1\ldots0) &= \mathrm{Tr}(\rho M_1^{10} \otimes \ldots \otimes M_n^{00}) \\
P(1\ldots1|1\ldots1) &= \mathrm{Tr}(\rho M_1^{11} \otimes \ldots \otimes M_n^{11}),
\end{aligned} \tag{2.59}$$

where $\rho = |\psi\rangle\langle\psi|$ is the density matrix of a n-qubit permutation symmetric state $|\psi\rangle$. Rewriting P_n this way results in a vector polynomial of 2n variables (v_i^0 and v_i^1 for each i)

$$\begin{aligned}
&V(v_1^0, v_1^1, \ldots, v_n^0, v_n^1) = \\
&\mathrm{Tr}(\rho M_1^{00} \otimes \ldots \otimes M_n^{00}) \\
&+\mathrm{Tr}(\rho M_1^{00} \otimes \ldots \otimes M_n^{10}) \\
&\vdots \\
&+\mathrm{Tr}(\rho M_1^{10} \otimes \ldots \otimes M_n^{00}) \\
&+\mathrm{Tr}(\rho M_1^{11} \otimes \ldots \otimes M_n^{11}).
\end{aligned} \tag{2.60}$$

The goal of an SDP program is to maximize $V(v_1^0, v_1^1, \ldots, v_n^0, v_n^1)$, subject to the constraint that the Gram matrix formed by the vectors v_i^0 and v_i^1 is positive semidefinite [HJ12].

Chapter 3

The Majorana Representation of Symmetric States

In 1932, a few years before the EPR paper was published, a genius Italian physicist, Ettore Majorana, published a paper on the effect of inhomogeneous magnetic fields on atoms [Maj32] (see [Maj06] for a retyped version and an English translation, also see [BR45] for a more in depth review). Then six years later he disappeared at sea, with suicide the most prevalent explanation. In his short life, he made great contributions to the study of symmetry in physics. His most famous contribution is perhaps the Majorana fermion, a fermion who is its own antiparticle [Maj37], which very recently has been observed in a nanowire [MZF+12]. Although the Majorana representation, which came out of his 1932 paper, is less well-known than the fermion named after him, it has quietly influenced many areas of modern quantum information.

On the more theoretical physics side, the Majorana representation has been used to study the phase space of quantum dynamics [Leb99] and the geometric phase [Han98]. It also found applications in many body physics [RVM08] and condensed matter physics [MS07] [BTD07]

The first application of the elegant framework on foundations of quantum mechanics was by Roger Penrose, whose research record uses complex geometry as a common thread [ZP93, Pen00]. In this work, the Majorana was used representation to construct orthogonal states which can be used to give a proof of the Kochen-Specker theorem on the contextuality of quantum mechanics [KS67]. Later, the Majorana representation found its use in entanglement theory [BKM+09, AMM10, AMM11, Aul11b, RM11, Aul11a, Mar11], where several results will be discussed in this chapter.

Until recently, the only multipartite states which also has an elegant mathematical tool which permits a systematic study of their nonlocal properties are stabilizer states, by using the MABK inequality (cf. Chapter. 2). Now the Majorana representation has found

another application in nonlocality which allows the systematic study of the nonlocality of all symmetric states. This chapter introduces the necessary background to the Majorana representation, starting from its geometrical aspects. At the end of the chapter, the main results of applying the Majorana representation to study the entanglement of symmetric states will also be discussed.

3.1 Geometry of Complex Numbers

3.1.1 The Complex Plane and the Riemann Sphere

In Chapter 2, complex numbers are defined algebraically, as a field and a vector space over itself. In addition to its algebraic structures, complex numbers have very elegant geometrical representations [Sch80, Nee99, Pen04, Bea05, Sin05].

The most common way to visualize a complex number $z = a + bi$ is through complex plane (Fig. 3.1), where the horizontal axis represents the real part and the vertical axis represents the imaginary part.

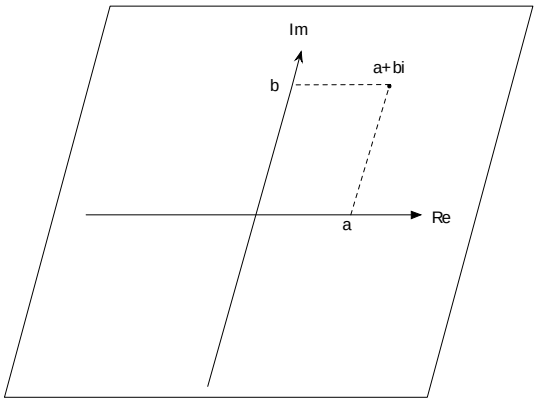


Figure 3.1: The complex plane with the number $a + bi$.

The complex plane is easy to use but there are drawbacks. For example, to visualize two complex numbers x, y with $kx - yk \ll 1$, the plane needs to be scaled. If another point z close to y is added, then the scaling of the plane may make it hard to visualize all three points, with y and z lumped together. This inconvenience comes from the fact that the complex plane can not be visualized in a finite region. Luckily, there is a more compact (both topologically and aesthetically speaking) way to represent complex numbers in the form of a unit sphere, called the Riemann sphere.

To map a complex number from the plane to the sphere, stereographic projection is used (Figure.(3.2)[1]). The idea is simple: make the equatorial plane of the sphere coincide with the complex plane, with the center of the sphere coincide with the origin of

[1] Source: Wikipedia

the complex plane. This way the equator of the sphere represents the complex numbers of modulus 1. It is easy to see that the line passing through the north pole and a complex number on the plane will intersect the sphere at one point. This point on the sphere is the image of the point on the plane.

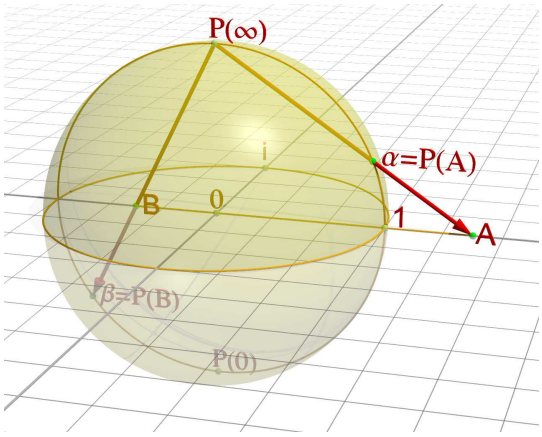


Figure 3.2: The stereographic projection of the complex plane to the unit sphere.

This procedure can map every point on the plane to the sphere. But not all points on the sphere correspond to a point on the plane: the north pole itself can not be mapped. There are several ways to remedy this. The easiest way is to use an extension of the complex numbers, called the extended complex numbers, which has a special number, 1, that is mapped to the north pole of the sphere. The number 1 can be thought to be close to very big complex numbers, which surrounds the complex plane, with the property that $z + 1 = 1$, $z \cdot 1 = 1$. So in an geometric sense it is not a point on the plane. The extended complex numbers, denoted by C_1, allows division by zero to be defined: $\frac{z}{0} = 1$, $\frac{z}{1} = 0$.

Another to resolve the north pole problem is to use two complex planes, glued back to back, with each plane providing the projection of almost every point on the sphere except at one pole. Each point (except the poles) on the sphere will be mapped to two complex numbers, y and z, with the conversion formula $y = \frac{1}{z}$ and $z = \frac{1}{y}$. The north pole is mapped to the origin of the lower plane, and the south pole is mapped to the origin of the upper plane, with both origins playing the role of 1 as in C_1.

Gluing two complex planes back to back has another geometrical interpretation: the complex projective line, CP^1. The complex projective line is defined as the rays of C^2: given two complex numbers, not both zero, $(y, z) \in C^2$, then the rays of C^2 correspond to (ay, az), where $a \in C$, $a \neq 0$. The two copies of the complex plane can be found as $(1, \frac{z}{y})$ and $(\frac{y}{z}, 1)$, which the same conversion formula $y = \frac{1}{z}$ provides the transition between the two copies.

For convenience, the sphere and the plane may use different coordinate systems.

While for the plane the most natural coordinate system is still the Cartesian system, with axes labeled by real and complex parts of a complex number, it is easier to work with the spherical coordinate system for the sphere. The spherical coordinate system uses two angles $0 \leq \theta \leq \pi$ and $0 \leq \phi \leq 2\pi$ (Figure 3.3). Simple trigonometric calculations yield the following conversion formula:

$$x + iy = \cot\frac{\theta}{2} e^{i\phi}. \tag{3.1}$$

Note that (3.1) corresponds to the situation where the sphere minus the north pole is mapped to one of the glued planes. For the other plane, the formula becomes

$$x - iy = \tan\frac{\theta}{2} e^{-i\phi}. \tag{3.2}$$

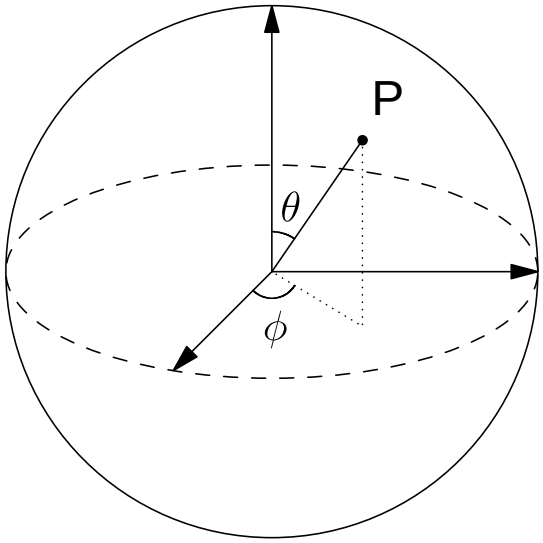


Figure 3.3: The Riemann sphere with spherical coorinates θ and ϕ.

3.1.2 The Möbius Transformation

Having a compact geometrical representation of complex numbers allows the study of transformations of complex numbers visually. Given one complex number, i.e. a point on the Riemann sphere, the mapping to any other complex number can be completely characterized by the rotation of the sphere. But given a set of points on the Riemann sphere, what are the continuous bijective transformations that change their locations? A simple algebraic form called the Möbius transformation turns out to completely characterize these bijections from the Riemann sphere to itself (called automorphisms of the Riemann sphere).

A Möbius transformation $f(z)$ is a function of one complex variable $z \in C_1$ given by

$$f(z) = \frac{az+b}{cz+d}, \tag{3.3}$$

where a, b, c, d are complex numbers, and $ad - bc \neq 0$.

The requirement $ad - bc \neq 0$ guarantees means that the Möbius transformation is not a constant map. Because if $ad - bc = 0$, then

$$f(y) - f(z) = \frac{ay+b}{cy+d} - \frac{az+b}{cz+d} \tag{3.4}$$

$$= \frac{(ad-bc)(y-z)}{(cy+d)(cz+d)} \tag{3.5}$$

$$= 0. \tag{3.6}$$

If the coefficients are written in a matrix form

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \tag{3.7}$$

then the requirement that $ad - bc \neq 0$ means that the determinant of M is nonzero, i.e. the matrix is invertible.

For any matrix of the form (3.7), it is possible to multiply its coefficients by $\pm\frac{1}{\sqrt{ad-bc}}$ to obtain a matrix with determinant 1. The matrices M with $|M| = 1$ are called normalized. With normalization, the arbitrariness of the coefficients can be reduced considerably. In fact there are two matrices with opposite sign which correspond to a normalized Möbius transformation [Nee99].

A few other subtleties arise because the Riemann sphere contains ∞. The way ∞ is defined above, by setting $\infty = \frac{z}{0}$, and the intuitions that ∞ is close to very big complex numbers, make it possible to define $f(\infty)$ at ∞. First of all, when $z \to \infty$ and $c \neq 0$, then $\frac{az+b}{cz+d} \to \frac{a}{c}$, so it is sensible to define

$$f(\infty) = \frac{a}{c}, \tag{3.8}$$

when $c \neq 0$. Secondly, if $cz + d = 0$, then the value of f should be ∞. So if $c \neq 0$, then

$$f(-\frac{d}{c}) = \infty. \tag{3.9}$$

This only leaves the case where $c = 0$. If $c = 0$, then

$$f(\infty) = \infty, \tag{3.10}$$

which is also consistent with both definitions above.

The Möbius transformation is defined algebraically above. But it is intrinsically geometric concept. Thus it is more helpful to move to a geometry-centric picture first and visualize a Möbius transformation, moving its algebraic properties to a secondary position.

It can be shown easily that a Möbius transformation $f(z)$ is a composition of these four more elementary transformations:

1. Translation by $\frac{d}{c}$: $z \mapsto z + \frac{d}{c}$.
2. Complex inversion (also called reciprocal): $z \mapsto \frac{1}{z}$.
3. Expansion and rotation: $z \mapsto \frac{ad-bc}{c^2} z$.
4. Translation by $\frac{a}{c}$: $z \mapsto z + \frac{a}{c}$.

Of these four steps, the second one, complex inversion, is more complicated thus merits a closer look.

Algebraically, the complex inversion for one complex number is very simple: for $z = r e^{i\theta}$, the inversion takes it to

$$\frac{1}{z} = \frac{1}{r e^{i\theta}} = \frac{1}{r} e^{-i\theta}. \tag{3.11}$$

On the complex plane, the inversion can be visualized in Fig. 3.4. The circle in Fig. 3.4 is the circle representing complex numbers of modulus 1. If $r < 1$, meaning it lies within the circle, then the inversion takes it to a point outside the circle. If $r > 1$ then it gets inverted to a point inside the circle. The second effect of inversion reflects the phase angle θ with respect of the real axis. Complex numbers of unit modulus stay on the circle, but with reflected phase angles.

It is more interesting to study the effect of complex inversion when the object being inverted is not a complex number, but rather a circle. Why is this interesting? Because under stereographic projection, a line in the complex plane traces out a circle passing through the north pole on the Riemann sphere. Circles in the complex plane trace out circles on the Riemann sphere as well, with straight lines being circles of infinite radius (hence they pass through the north pole). Thus circles on the Riemann sphere give a unified object for studying circles and straight lines in the complex plane. This unified object can be represented algebraically as generalized circles [Sch80].

To derive the expression for a generalized circle, first consider what constitutes a circle. There must be an origin o, a radius r and an equation defining the circle whose solutions are the set of points at the distance r away from the origin. To meet these requirements, an arbitrary complex number o may be chosen as the origin, and the

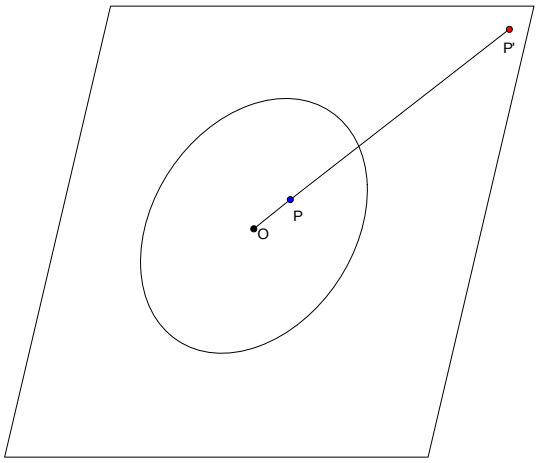


Figure 3.4: On the complex plane, a point P inside the unit circle (blue) gets inverted to P^0 outside the unit circle (red).

radius r is any real number, with the circle z defined by the equation

$$|z - o|^2 = r^2, \tag{3.12}$$

$$(z - o)(\overline{z - o}) = r^2, \tag{3.13}$$

$$z\bar{z} - z\bar{o} - \bar{z}o + o\bar{o} - r^2 = 0. \tag{3.14}$$

Now multiply the last line by some real constant:

$$\gamma z\bar{z} + \beta z + \bar{\beta}\bar{z} + \delta = 0, \tag{3.15}$$

with γ, δ real, β complex. It is easier to write the coefficients as a Hermitian matrix

$$C = \begin{pmatrix} \gamma & \beta \\ \bar{\beta} & \delta \end{pmatrix}. \tag{3.16}$$

To differentiate various types of circles represented by (3.15), the determinant of C, $|C| = \gamma\delta - |\beta|^2$ can be used:

- When $\gamma \neq 0$ and $|C| < 0$, then C represents a normal circle.
- When $\gamma \neq 0$ and $|C| = 0$, then C represents a point.
- If $\gamma = 0$, then then C represents a straight line.

Two matrices C and C^0 represent the same generalized circle if there is a non-zero real number r such that $C = rC^0$.

From (3.15), it is easy to see the characteristics of complex inversion. If $y = \frac{1}{z}$, then

$$\alpha\frac{1}{y}\frac{1}{\bar{y}} + \beta\frac{1}{y} + \bar{\beta}\frac{1}{\bar{y}} + \gamma = 0, \tag{3.17}$$

$$\alpha y\bar{y} + \bar{\beta}y + \beta\bar{y} + \gamma = 0. \tag{3.18}$$

Together with the properties of $|C|$, the effects of complex inversion on generalized circles can be summarized as follows:

- Straight lines through the origin ($\alpha = 0$ and $\gamma = 0$) are mapped to straight lines through the origin ($\alpha = 0$ and $\gamma = 0$). Fig(3.5a)).
- Straight lines not through the origin ($\alpha \neq 0$ and $\gamma = 0$) are mapped to circles through the origin ($\alpha = 0$ and $\gamma \neq 0$) and vice versa. Fig(3.5b).
- Circles not through the origin ($\alpha \neq 0$ and $\gamma \neq 0$) are mapped to circles not through the origin ($\alpha \neq 0$ and $\gamma \neq 0$). Fig(3.5c).

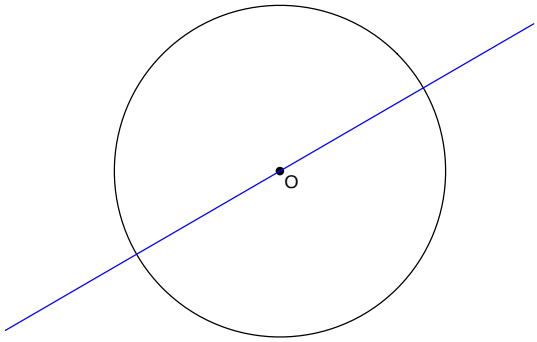
From Fig. 3.5, it is clear that circles on the Riemann sphere are again mapped to circles on the Riemann sphere.

Given two sets of points, determining whether there is a Möbius transformation connecting them can be tricky. The most obvious criterion is that the two sets should contain the same number of points. In general, the coefficients of a Möbius transformation are not unique. Intuitively, in order to specify a unique Möbius transformation, four complex numbers are needed. Geometrically speaking, the images of four points on the Riemann sphere are needed. But it turns out that actually only three numbers are necessary to specify a unique Möbius transformation. Geometrically speaking, there is a unique Möbius transformation connecting any two sets of three distinct points. If a Möbius transformation leaves some points unchanged, they are called fixed points of the transformation. If a Möbius transformation has more than two fixed points, then it is the constant map.

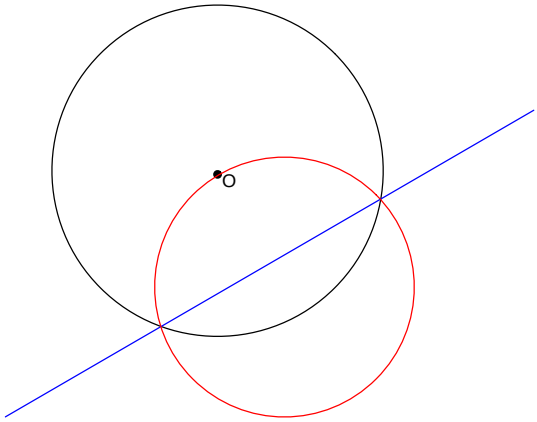
Although for more than three points there is in general no unique Möbius transformation mapping them to another set of points. But for four (distinct) points, the Möbius transformation does preserve the cross ratio, defined by

$$[z_1, z_2, z_3, z_4] := \frac{(z_1 - z_3)(z_2 - z_4)}{(z_1 - z_2)(z_3 - z_4)}. \tag{3.19}$$

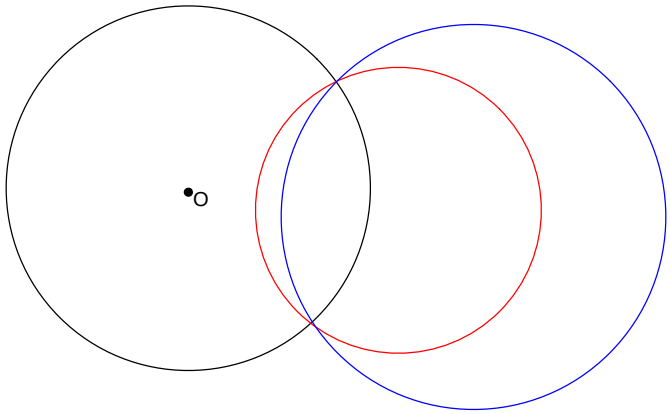
Then a necessary and sufficient condition for the existence of a Möbius transformation connecting two sets of four distinct points $\{y_1, y_2, y_3, y_4\}$ and $\{z_1, z_2, z_3, z_4\}$ is $[y_1, y_2, y_3, y_4] = [z_1, z_2, z_3, z_4]$. A useful result about cross ratios is that any four points lie on the same circle (on the complex plane) iff their cross ratio is real.



(a) A straight line through the origin (blue) is inverted to itself, but with different identification of points.



(b) A straight line not through the origin (blue) is inverted to a circle through the origin (red) and vice versa.



(c) A circle not through the origin (blue) is inverted to a circle not through the origin (red) and vice versa.

Figure 3.5: Effects of complex inversions on generalized circles.

The last remarkable property of the Möbius transformation comes from the algebraic depiction of the rotation of the Riemann sphere. It can be shown that for a Möbius transformation $f(z)$ to be a rotation, it needs to be of the form

$$f(z) = \frac{az + b}{-\bar{b}z + \bar{a}}, \tag{3.20}$$

with $|a|^2 + |b|^2 = 1$. In the matrix form, rotations of the sphere correspond to matrices of the form

$$U = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}, |a|^2 + |b|^2 = 1. \tag{3.21}$$

Thus Möbius transformations which are rotations of the sphere correspond to elements of SU(2).

3.2 From Complex Geometry to the Majorana Representation

The Riemann sphere has a direct physical counterpart: the Bloch sphere [Blo46, NC00, Sin05], which represents the state space of a spin-$\frac{1}{2}$ particle, i.e. a qubit (Figure.(3.6)). This is not a coincidence. Just like the representation of the Riemann sphere as CP^1, the state space of a qubit can also be identified with CP^1, because all states are rays in Hilbert space (i.e. they are scalar multiples of each other). In the Bloch sphere, the north and south poles are identified with the states $|0\rangle$ and $|1\rangle$, respectively. An arbitrary normalized qubit can be written as

$$|\psi\rangle = \cos(\frac{\theta}{2})|0\rangle + e^{i\phi}\sin(\frac{\theta}{2})|1\rangle, \tag{3.22}$$

with θ, ϕ the same spherical coordinates as used on the Riemann sphere. Conversely, given a qubit $|\psi\rangle = \cos(\frac{\theta}{2})|0\rangle + e^{i\phi}\sin(\frac{\theta}{2})|1\rangle$, there is a complex number $z = x + iy$ where (3.1) holds. The inner product between any two qubits can be easily calculated in terms of θ, ϕ. One particular interesting geometric aspect of the inner product is that two orthogonal qubits are antipodal points on the Bloch sphere.

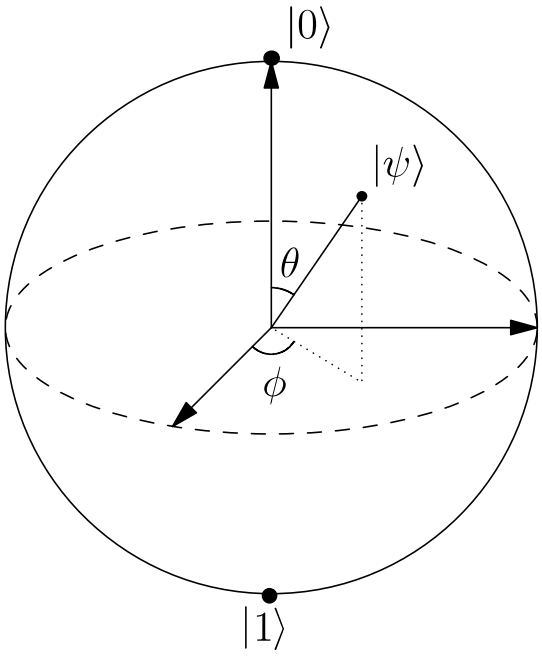


Figure 3.6: The Bloch sphere.

The Majorana representation is a way of visualizing any permutation symmetric state of n qubits as a set of n unordered points on the Bloch sphere.

In order to understand the Majorana representation, it is instructive to first look at the similarities between single-variable polynomials with complex coefficients and the

symmetric states of qubits.

A single-variable polynomial with complex coefficients (hereafter referred to simply as a polynomial) has the form

$$p(X) = a_n X^n + a_{n-1} X^{n-1} + \ldots + a_1 X + a_0, \tag{3.23}$$

where $X \in \mathbb{C}$ and $\forall i, a_i \in \mathbb{C}$.

The simplest pure permutation symmetric states of qubits (hereafter referred to simply as symmetric states) are Dicke states, of the form

$$|S(n,k)\rangle = \frac{1}{\sqrt{\binom{n}{k}}} \sum_{perm} |\underbrace{0\ldots0}_{n-k}\underbrace{1\ldots1}_{k}\rangle, \tag{3.24}$$

where the sum is taken over all permutations of the string with k ones and $n-k$ zeros.

In fact, the Dicke states $|S(n,k)\rangle$, where $k = 0, 1, \ldots, n$ form an orthonormal basis of the symmetric subspace of the Hilbert space of n qubit symmetric states. That is to say, any n qubit symmetric state $|\psi\rangle$ can be written as a linear combination of the Dicke states:

$$|\psi\rangle = K(\alpha_n |S(n,n)\rangle + \alpha_{n-1} |S(n,n-1)\rangle + \ldots + \alpha_1 |S(n,1)\rangle + \alpha_0 |S(n,0)\rangle), \tag{3.25}$$

where K is some normalization constant.

Now the similarity between (3.23) and (3.25) is obvious: if $\{X, X^2, \ldots, X^n\}$ are seen as a basis in the vector space of polynomials (which is also a Hilbert space with an appropriate definition of the inner product), then the coefficients $\{a_0, \ldots, a_n\}$ and $\{\alpha_0, \ldots, \alpha_n\}$ play the same role. This is in essence how the Majorana representation works: if a symmetric state of n qubits is treated as a polynomial of degree n, then by the fundamental theorem of algebra, this polynomial must have n roots, which can be represented by n points on the Bloch sphere. The state $|\psi\rangle$ is uniquely determined, up to a global phase, by these n points. The points obtained in the Majorana representation are called, without surprise, Majorana points (MPs). Because the surface of the Bloch sphere corresponds to the state space of a spin-$\frac{1}{2}$ particle, the Majorana points are all normalized qubits.

Given this definition, it is now instructive to see how the Majorana representation works operationally. There are two directions in which the Majorana representation can be used:

1. Given any symmetric state, use the Majorana representation to find its MPs.

2. Given a set of MPs, use the Majorana representation to reconstruct the corresponding symmetric state.

The first direction can be achieved by coherent state decomposition: let $|\alpha\rangle$ equals to $|0\rangle + \alpha|1\rangle$, with $\alpha \in \mathbb{C}$ as an unknown (this is just writing a qubit as an element of $\mathbb{CP}^1$: $(1, \alpha)$), then the overlap

$$(\langle\alpha|)^{\otimes n}|\psi\rangle = K(\bar{c}_n \alpha^n |S(n,n)\rangle + \ldots + \bar{c}_0 \alpha^0 |S(n,0)\rangle), \tag{3.26}$$

where c_i is the complex coefficient of $|S(n,i)\rangle$ in the expansion of $|\psi\rangle$ in the Dicke basis.

(3.26) can be seen as a polynomial of degree n with one complex variable, so it has n roots. There is a subtle point here: because it is possible that some c_i are zero, the polynomial may not be of full degree. There are two ways to solve this problem: either allow ∞ to count as a root, or to use $|\alpha\rangle = \alpha|0\rangle + |1\rangle$ to get another polynomial, then use $\alpha = \frac{1}{\alpha}$ to identify the roots. The n complex roots are then translated to n qubits, $|\alpha_1^?\rangle, \ldots, |\alpha_n^?\rangle$. The n antipodal points of these qubit on the Bloch sphere, denoted by $|\alpha_1\rangle, \ldots, |\alpha_n\rangle$, are the MPs of $|\psi\rangle$.

The second direction, reconstructing the state from its MPs, is even simpler. Given a set of MPs $\{|\alpha_1\rangle, \ldots, |\alpha_n\rangle\}$, the state $|\psi\rangle$ can be recovered by

$$|\psi\rangle = K(\sum_{perm} (|\alpha_1\rangle \otimes |\alpha_2\rangle \otimes \ldots \otimes |\alpha_n\rangle)), \tag{3.27}$$

which is the sum of all possible permutations of tensor products of the MPs.

From the definition of the Majorana representation and (3.27), a few observations can be made immediately

1. A local unitary will be a rotation of the sphere.
2. A symmetric state is a product state iff all its MPs coincide.
3. The antipodal points of the MPs define the product states orthogonal to the original state.

The first statement comes from (3.21). The second statement is true from (3.27): if all the $|\alpha_i\rangle$ are the same, then the permutation will not change the tensor product, so the state is a product state. If more than one MPs are the same (geometrically represented by stacking the MPs at the same location), then the MP is called degenerate. If $|\psi\rangle$ has degenerate MPs, the notation in (3.27) will be slightly altered to

$$|\psi\rangle = K\sum_{perm} |\alpha_1^{d_1} \alpha_2^{d_2} \ldots \alpha_l^{d_l}\rangle, \tag{3.28}$$
$$\forall i \neq j, |\alpha_i\rangle \neq |\alpha_j\rangle, \sum_{i=1}^{l} d_i = n.$$

The third statement is true because for any $|\alpha_i^?\rangle$ orthogonal to an MP $|\alpha_i\rangle$ of the

state $|\ \rangle$, the state $|\Box_i^{?}\rangle^{\Box n}$ is orthogonal to $|\ \rangle$

$$\langle \underbrace{\Box_i^{?}| \Box \ldots \Box \langle \Box_i^{?}|}_{n} \ \rangle = 0, \tag{3.29}$$

because in (3.27), every term in the sum contains $|\Box_i\rangle$ in the tensor product. The number of $|\Box_i^{?}\rangle$ can be reduced if $|\Box_i\rangle$ is degenerate. If $|\Box_i\rangle$ has degeneracy d (i.e. d MPs are the same), then (3.29) becomes

$$\langle \underbrace{\Box_i^{?}| \Box \ldots \Box \langle \Box_i^{?}|}_{n\Box d+1} \ \rangle = 0. \tag{3.30}$$

Degeneracy already plays an important role in the SLOCC classification of symmetric states because SLOCC operations can not change the degeneracy of the MPs of the symmetric states (see Section 3.4). In later chapters, degeneracy will be shown to be linked with the persistency of nonlocality and entanglement into subsets of the parties.

3.3 Physical Interpretations

Mathematically, the qubit shares a similar structure to another object: the 2-spinor [Car13, Dir28, PR84, PR86], which is usually a column vector with 2 complex components:

$$= \begin{pmatrix} 1 \\ 2 \end{pmatrix} \tag{3.31}$$

The spin is a characteristic of the angular momentum of a particle. Like linear momentum, angular momentum should be conserved in particles. Noether's Theorem [Noe18, KS11] states that for every conserved quantity, there is an associated symmetry of the physical system. For example, energy/ mass conservation can be associated with time translation symmetry while linear momentum conservation can be associated with space translation symmetry. In a similar way, the spin, which arises from angular momentum conservation, is associated with rotational symmetry.

This is where the Riemann/ Bloch sphere picture comes in: they provide an elegant geometric picture to study the rotational symmetry of simple quantum mechanical systems. For a particle in 3-dimensional space, a rotation can be represented by a 3×3 real orthogonal matrix R having the property $R^T R = I$. It can be easily checked that the determinant of R equals to ± 1. The set of all orthogonal matrices form a group by using matrix multiplication as group operation, in the case of rotations in 3-dimensional space, the group is called the orthogonal group of dimension 3, or O(3), and the half of the group with $|R| = 1$ (which contains the identity) is called special orthogonal group of dimension 3, or SO(3). The orthogonal matrices are essentially real counterparts of unitary matrices.

From (3.21), rotations of the Riemann sphere correspond to Möbius transformations whose matrices are elements of SU(2). Then what is the relationship between SU(2) and rotations? To see the connection, consider the 3-dimensional vector space V with vectors $v = (x, y, z)$. A 2×2 Hermitian matrix can be associated to each vector, by using the Pauli matrices:

$$A = x\sigma_x + y\sigma_y + z\sigma_z. \tag{3.32}$$

An inner product on the matrix space can be defined by

$$\langle A, B \rangle = \frac{1}{2}\mathrm{Tr}(AB). \tag{3.33}$$

It can also be easily checked that

$$|A| = \pm\|v\|. \tag{3.34}$$

Take any element U of SU(2), it can be checked that the mapping $\rho_U(A) = UAU^\dagger$ with A representing a vector in V is an orthogonal transformation in V. From the definition above, V can be seen as R^3, thus the orthogonal transformations using elements of SU(2) can be mapped to elements of SO(3). Also, it is easy to see that $\rho_{-U}(A) = \rho_U(A)$, so in fact ρ takes two elements of SU(2) to the same element of SO(3) (called a double cover of SO(3)).

Having established SU(2) as the representation of rotations in 3-dimensional space, it is easy to see how it acts on a spinor: just use matrix multiplication. It is even more interesting to take a closer look at the Pauli matrices, since they can be seen as the basis of R^3. In a typical spin measurement experiment, called the Stern-Gerlach experiment [GS22], a stream of particles are sent through a variable magnetic field, then through interaction with the magnetic field, their trajectories are changed. The change of trajectories is a manifestation of the spin of the particles, and the magnetic field acts as a measurement device (since the experimenter knows the direction of the magnetic field). For particles like electrons, there are only two possible direction which the trajectories can take (contrasting the classical prediction with an continuous range of directions). If the Pauli matrices are related to spin, then a measurement of one of them should project the spin along the direction of one of its eigenvectors. It is easy to see that the eigenvectors of the Pauli matrices, written as spinors with superscripts denoting the corresponding eigenvalue, are

$$\sigma_x^{+1} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}, \quad \sigma_x^{-1} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}, \tag{3.35}$$

$$\sigma_x^{+1} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\i\end{pmatrix}, \quad \sigma_x^{-1} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-i\end{pmatrix}, \tag{3.36}$$

$$\sigma_z^{+1} = \begin{pmatrix}1\\0\end{pmatrix}, \quad \sigma_z^{-1} = \begin{pmatrix}0\\1\end{pmatrix}. \tag{3.37}$$

With this notation, the qubit $|0\rangle$ is just the $+1$ eigenstate of σ_z, and the qubit $|1\rangle$ is the -1 eigenstate, and the unitaries of SU(2) are rotations of the Bloch sphere. The Pauli matrices represent measurements of the spin along x, y, z axes.

A single 2-spinor corresponds to a particle of spin-$\frac{1}{2}$. For physical and historical reasons (that have something to do with $\hbar$), the values of spin are always integer multiples of $\frac{1}{2}$. The Majorana representation, physically speaking, allows the decomposition of any spin $\frac{n}{2}$ as a set of n spin-$\frac{1}{2}$ points on the Bloch sphere [Maj32, BR45]. If the two directions of an electron after going through a Stern-Gerlach experiment are labeled up $|\uparrow\rangle$ and down $|\downarrow\rangle$, a particle with spin-1, after being sent though a Stern-Gerlach device,

can have 3 different directions, (in the Majorana representation):

$$|\uparrow,\uparrow\rangle, \quad |\downarrow,\uparrow\rangle, \quad |\downarrow,\downarrow\rangle. \tag{3.38}$$

Note that the these vectors are all symmetric. In particular, $|\downarrow,\uparrow\rangle = |\uparrow,\downarrow\rangle$. For a particle of spin $\frac{n}{2}$, there are $n+1$ directions possible.

All these happened long before the dawn of quantum information. Quantum information, with an axiomatic bottoms-up approach inspired by computer science, works with abstract objects which do not need to have obvious physical counterparts. So instead of a 2-spinor, people prefer to work with qubits. And instead of studying particles with high spin, people study symmetric states of qubits, which using the Majorana representation capture the essence of particles with high spin: in the modern notation, the different directions a particle of spin $\frac{n}{2}$ can take are just the $n+1$ Dicke states $|S(n,0)\rangle$ to $|S(n,n)\rangle$.

3.4 The Majorana Representation and Entanglement

This section presents several results obtained by applying the Majorana representation to the study of multiparty entanglement. These results can be found in [BKM+09, AMM11, Aul11b, Mar11, RM11, Mar11], where the relationship between complex geometry and multiparty entanglement are explored in more depth.

In the previous chapter, entanglement classification was introduced as part of the process to quantify entanglement. Although entanglement classification for general pure and mixed states is a hard problem, it does have direct geometrical implications for symmetric states.

The most fundamental connection between entanglement classes of symmetric states and complex geometry is can be summarized in the following theorem [Aul11b, RM11, Aul11a]:

Theorem 3. Two symmetric states $|\psi\rangle$ and $|\phi\rangle$ are SLOCC equivalent iff there is a Möbius transformation connecting their Majorana points.

The proof is obvious from the definitions of the normalized Möbius transformation (3.7 and paragraph below) and the SLOCC ILO (2.14).

From this theorem, two corollaries are also easy to prove:

Corollary 4 ([BKM+09]). Two symmetric states $|\psi\rangle$ and $|\phi\rangle$ whose MPs have different degeneracies are in different SLOCC classes.

This can be seen from the fact that a Möbius transformation can not change the degeneracy of the MPs.

Corollary 5 ([RM11]). Two symmetric states $|\psi\rangle$ and $|\phi\rangle$ are LOCC equivalent iff they have the same MPs up to rotation of the Bloch sphere.

This can be deduced from (3.21).

To actually quantify entanglement, a suitable entanglement measure is indispensable. As mentioned in the previous chapter, the geometric measure, thanks to its symmetric state-friendly properties, has been used to extensively study the entanglement of symmetric states.

In the definition of the geometric measure $E_G(|\psi\rangle)$ (2.15), the closest product state $|\phi\rangle$ maximizes the quantity $|\langle\phi|\psi\rangle|^2$, thus can be seen as close to the state $|\psi\rangle$. For symmetric states, a class of states whose MPs form the Platonic solids are very interesting because of the relationship between the states themselves and their closest product states (which, when illustrated by points on the Bloch sphere, are called closest product points or CPPs).

The following table summarizes all the Platonic solids. The symmetric states whose MPs are these solids are given in the captions of the figures, written in the Dicke basis.

The geometric measure of entanglement E_G, defined in (2.15), of these Platonic solid states are also included in the table.

Name	Vertices	Figure	Dual	E_G
Tetrahedron	4	3.7	Tetrahedron	$\log_2 3 \approx 1.585$
Octahedron	6	3.8	Cube	$\log_2 \frac{9}{2} \approx 2.167$
Cube	8	3.9	Octahedron	$\log_2 \frac{24}{5} \approx 2.263$
Icosahedron	12	3.10	Dodecahedron	$\log_2 \frac{243}{28} \approx 3.117$
Dodecahedron	20	3.11	Icosahedron	$\log_2 \frac{1875}{187} \approx 3.326$

Table 3.1: The Platonic solids.

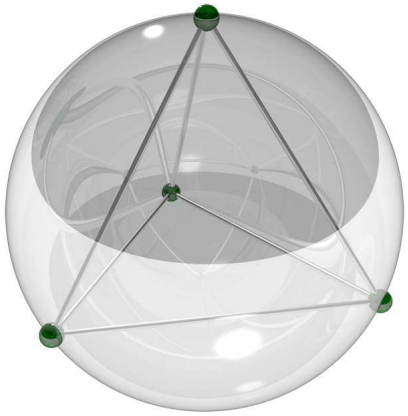


Figure 3.7: The Tetrahedron. $|T\rangle = \sqrt{\frac{1}{3}}|S(4,0)\rangle + \sqrt{\frac{2}{3}}|S(4,3)\rangle$.

One interesting fact about Platonic solids is that the dual polyhedron of a Platonic solid is also a Platonic solid (cf. last column of Table.3.1). In the case of symmetric states, the CPPs of a symmetric state whose MPs are vertices of a Platonic solid coincide with the vertices of its dual [Aul11a].

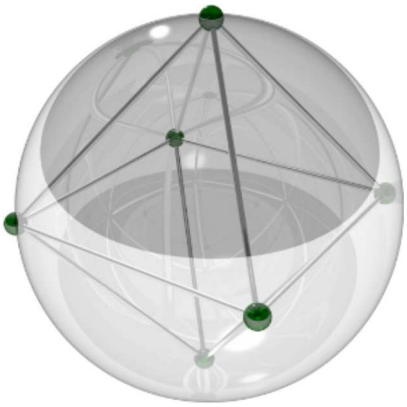


Figure 3.8: The Octahedron. $|O\rangle = \frac{1}{\sqrt{2}}(|S(6,1)\rangle + |S(6,5)\rangle)$.

Figure 3.9: The Cube. $|C\rangle = \frac{1}{2\sqrt{6}}(\sqrt{5}|S(8,0)\rangle + \sqrt{14}|S(8,4)\rangle + \sqrt{5}|S(8,8)\rangle)$.

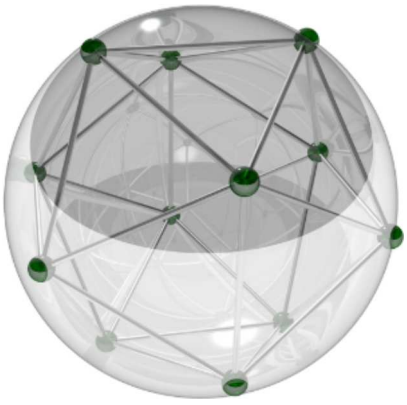


Figure 3.10: The Icosahedron. $|I\rangle = \frac{1}{5}(\sqrt{7}|S(12,1)\rangle - \sqrt{11}|S(12,6)\rangle - \sqrt{7}|S(12,11)\rangle)$.

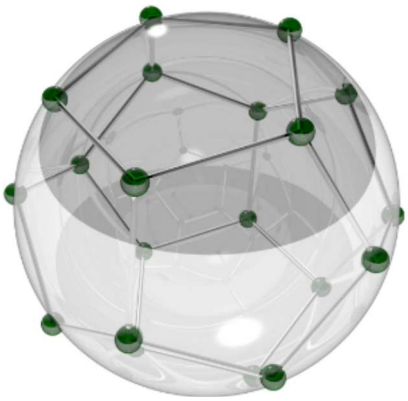


Figure 3.11: The Dodecahedron. $|D\rangle = \frac{1}{25\sqrt{3}}(\sqrt{187}|S(20,0)\rangle + \sqrt{627}|S(20,5)\rangle + \sqrt{247}|S(20,10)\rangle - \sqrt{627}|S(20,15)\rangle + \sqrt{187}|S(20,20)\rangle)$.

Chapter 4

Nonlocality of Symmetric States

After a lengthy introduction of background information interspersed with necessary mathematical tools to work with nonlocality and symmetric states, this chapter does the technical work to show that all symmetric states violate the same inequality and almost all symmetric states satisfy the extended n-party Hardy paradox.

After first reviewing the technical aspects of the original (bipartite) Hardy paradox, including how to construct the measurement bases for almost all bipartite entangled states, the generalization of the Hardy paradox and the constructive procedure to find the measurement settings for almost all symmetric states will be shown and proved. The final section shows that even though Dicke states do not satisfy the extended Hardy paradox, they still violate the generalized n-party CH inequality, which will be called P^n.

One common theme in this chapter is the role of symmetry breaking in determining the measurement settings, and how some symmetry persist into groups of subsystems. The symmetry breaking aspect has already been identified by Hardy, while the persistency of symmetry is a nice feature of symmetric states which also has a direct geometrical meaning in terms of the Majorana representation. The persistence of symmetry will be discussed more in detail in the next chapter.

4.1 Bipartite Hardy Paradox Revisited

The bipartite Hardy paradox has been explained verbally and mathematically in Section 2.4.5, together with the CH inequality, whose violation is guaranteed when the four probabilities defining the paradox (2.52) to (2.55) are satisfied. The missing link, which will be explained in this section as a warmup to the multipartite case, is how to choose measurement settings to actually satisfy (2.52) to (2.55), for almost all entangled pure states. This section essentially presents the result in [Har93] while slightly updating the notation using modern quantum information conventions.

In Section 2.3, a very useful mathematical tool was introduced to characterize the

entanglement of bipartite pure states: the Schmidt decomposition (2.10). Using the Schmidt decomposition, an arbitrary bipartite pure state can be decomposed using two sets of orthonormal bases, each belonging to one party. The Schmidt decomposition works for bipartite pure states of arbitrary dimension, and the two parties may even differ in dimension [NC00]. For the current problem, however, only qubits will be considered.

If only pure states of qubits are considered, then it is possible to use $|0\rangle$ and $|1\rangle$ as the orthonormal basis for both parties. Using the Schmidt decomposition, an arbitrary bipartite state $|\psi\rangle$ can be decomposed as

$$|\psi\rangle = a|00\rangle - b|11\rangle, \quad a, b \geq 0, \quad a^2 + b^2 = 1. \tag{4.1}$$

Note that (4.1) differs slightly from the usual definition of a Schmidt decomposition, where the Schmidt coefficients a, b are nonnegative. The negative sign in front of b is chosen to facilitate later calculations, so b is assumed to be nonpositive. Given the state $|\psi\rangle$ in another basis, the coefficients a and b can be calculated using singular value decomposition [NC00].

The most important step in bipartite Hardy paradox is the choice of measurement settings. The nature of the paradox, where an inference is made on one outcome while the other outcome is observed, calls for projective measurements (see Section 2.2, Postulate 2). The paradox uses two measurement settings, which will be labeled 0 and 1, corresponding to the box F and D in Section 2.4.5. The outcomes of each measurement will also be labeled 0 and 1. The translation between the notation used in (2.52) to (2.55) and the new notation is summarized in Table 4.1.

Content \ Box	0	1
0	Food, baguette	Drink, tea
1	Food, noodles	Drink, coffee

Table 4.1: Translation between the two different notations for bipartite Hardy paradox.

In the new notation, (2.52) to (2.55) are translated to

$$P(0,0|0,0) > 0 \tag{4.2}$$

$$P(0,0|0,1) = 0 \tag{4.3}$$

$$P(0,0|1,0) = 0 \tag{4.4}$$

$$P(1,1|1,1) = 0. \tag{4.5}$$

To calculate these probabilities, two orthonormal bases are used to construct the projectors of each observable corresponding to the settings 0 and 1. The basis used

by setting 0 will be labeled $\{|\psi^0\rangle, |\psi^1\rangle\}$ and the basis used by setting 1 will be labeled $\{|\phi^0\rangle, |\phi^1\rangle\}$, with superscripts denoting the outcomes. Using these two bases, (4.2) to (4.5) can be written as

$$P(0,0|0,0) = |\langle\phi^0| \otimes \langle\phi^0|)|\psi\rangle|^2, \tag{4.6}$$

$$P(0,0|0,1) = |\langle\phi^0| \otimes \langle\phi^0|)|\psi\rangle|^2, \tag{4.7}$$

$$P(0,0|1,0) = |\langle\phi^0| \otimes \langle\phi^0|)|\psi\rangle|^2, \tag{4.8}$$

$$P(1,1|1,1) = |\langle\phi^1| \otimes \langle\phi^1|)|\psi\rangle|^2. \tag{4.9}$$

In [Har93], Hardy proposed the following two bases, using a and b from (4.1):

$$|\psi^0\rangle = \frac{1}{\sqrt{|a|^3 + |b|^3}}(\sqrt{b^3}|0\rangle - \sqrt{a^3}|1\rangle), \tag{4.10}$$

$$|\psi^1\rangle = \frac{1}{\sqrt{|a|^3 + |b|^3}}(\sqrt{a^3}|0\rangle + \sqrt{b^3}|1\rangle), \tag{4.11}$$

$$|\phi^0\rangle = \frac{1}{\sqrt{|a| + |b|}}(\sqrt{\bar{a}}|0\rangle - \sqrt{\bar{b}}|1\rangle), \tag{4.12}$$

$$|\phi^1\rangle = \frac{1}{\sqrt{|a| + |b|}}(\sqrt{\bar{b}}|0\rangle + \sqrt{\bar{a}}|1\rangle). \tag{4.13}$$

Substituting (4.10) to (4.13) into (4.6) to (4.9):

$$P(0,0|0,0) \tag{4.14}$$

$$= |\langle\phi^0| \otimes \langle\phi^0|)|\psi\rangle|^2 \tag{4.15}$$

$$= \frac{(\sqrt{b^3}\langle 0| - \sqrt{a^3}\langle 1|) \otimes (\sqrt{b^3}\langle 0| - \sqrt{a^3}\langle 1|)(a|00\rangle - b|11\rangle)}{|a|^3 + |b|^3} \tag{4.16}$$

$$= \frac{1}{|a|^3 + |b|^3}(a\,|b^3 - a^3\,|b). \tag{4.17}$$

$$P(0,0|0,1) \tag{4.18}$$

$$= |\langle\phi^0| \otimes \langle\phi^1|)|\psi\rangle|^2 \tag{4.19}$$

$$= \frac{(\sqrt{b^3}\langle 0| - \sqrt{a^3}\langle 1|) \otimes (\sqrt{\bar{a}}\langle 0| - \sqrt{\bar{b}}\langle 1|)(a|00\rangle - b|11\rangle)}{\sqrt{(|a|^3 + |b|^3)(|a| + |b|)}} \tag{4.20}$$

$$= \frac{1}{\sqrt{(|a|^3 + |b|^3)(|a| + |b|)}}(a\sqrt{a}\,|b^3 - b\sqrt{a^3}\,|b) \tag{4.21}$$

$$= 0. \tag{4.22}$$

$$P(0,0|1,0) \tag{4.23}$$
$$= |(\langle\Box^1| \otimes \langle\Box^0|)|\ \ i\rangle|^2 \tag{4.24}$$
$$= \frac{(\sqrt{a}\langle 0| \otimes \sqrt{b}\langle 1|) \otimes (\sqrt{b^3}\langle 0| \otimes \sqrt{a^3}\langle 1|)(a|00\rangle \otimes b|11\rangle)}{\sqrt{(|a|^3 + |b|^3)(|a| + |b|)}} \tag{4.25}$$
$$= \frac{1}{\sqrt{(|a|^3 + |b|^3)(|a| + |b|)}}(a\sqrt{a}\sqrt{b^3} \otimes b\sqrt{a^3}\sqrt{b}) \tag{4.26}$$
$$= 0. \tag{4.27}$$

$$P(1,1|1,1) \tag{4.28}$$
$$= |(\langle\Box^1| \otimes \langle\Box^1|)|\ \ i\rangle|^2 \tag{4.29}$$
$$= \frac{(\sqrt{b}\langle 0| + \sqrt{a}\langle 1|) \otimes (\sqrt{b}\langle 0| + \sqrt{a}\langle 1|)(a|00\rangle \otimes b|11\rangle)}{(|a| + |b|)} \tag{4.30}$$
$$= \frac{1}{(|a| + |b|)}(a\sqrt{b} \otimes b\sqrt{a}) \tag{4.31}$$
$$= 0. \tag{4.32}$$

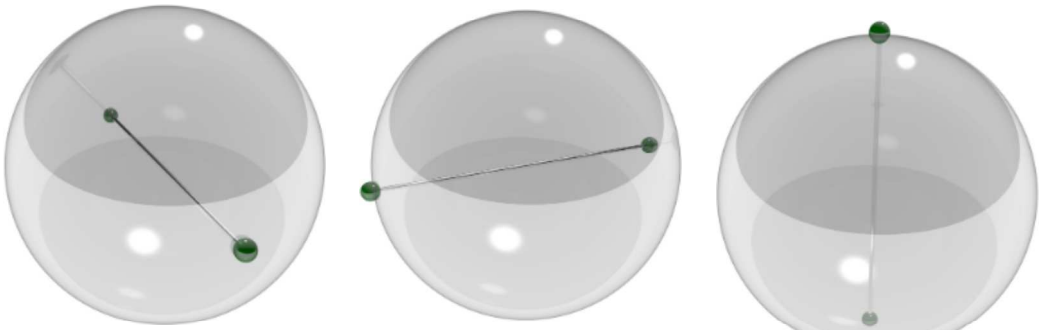
From these calculations it can be seen that the last three probabilities are always 0. Also, $P(0,0|0,0) = 0$ when $a = \pm b = \frac{1}{\sqrt{2}}$, which corresponds to the maximally entangled states, or either $a = 0$ or $b = 0$, which corresponds to product states. So the bipartite Hardy paradox works for all entangled pure states except the maximally entangled states.

Hardy hinted that maximally entangled states are too symmetric ("the proof relies on a certain lack of symmetry that is not available in the case of a maximally entangled state." [Har93]). The maximally entangled states of two qubits, also known as Bell states, are a set of four states including three symmetric states and one antisymmetric state:

$$|\Box^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \tag{4.33}$$
$$|\Box^-\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle), \tag{4.34}$$
$$|\Box^+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle), \tag{4.35}$$
$$|\Box^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle). \tag{4.36}$$

In the Majorana representation, the three symmetric Bell states $|\Box^+\rangle, |\Box^-\rangle, |\Box^+\rangle$ all

have two Majorana points which are antipodal (Fig. 4.1). The antisymmetric Bell state, also known as the singlet state, can be brought to a symmetric state by performing local unitaries ($\sigma_z \otimes I$, for example, takes it to $|\Psi^+\rangle$). Symmetric states with only two antipodal MPs correspond to Dicke states in the Majorana representation. The multipartite Hardy paradox also does not work for Dicke states, as will be shown below. The lack of symmetry, as pointed out by Hardy, will also be given a geometrical meaning in the multipartite paradox.



(a) $|\Psi^+\rangle$, with MPs $\frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle)$, $\frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle)$.

(b) $|\Phi^-\rangle$, with MPs $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$.

(c) $|\Psi^+\rangle$, with MPs $|0\rangle$, $|1\rangle$.

Figure 4.1: The three symmetric Bell states in the Majorana representation.

4.2 Multipartite Hardy Paradox and the Inequality P^n

The multiparty extension of the Hardy paradox is very straightforward. Following the bipartite paradox, a verbal argument will be given first, followed by a set of probabilistic conditions defining the paradox, then an inequality using these probabilities will be shown to hold for LHV model and violated by these probabilistic conditions.

Using the same terminology as in Section 2.4.5, but suppose there are n people, each given two sealed boxes from Charlie, one containing a food item and the other a drink, while only one box can be opened. The promises from Charlie are

- If everyone opens their food boxes, then it is possible that everyone gets baguettes.
- If everyone except one person opens their food boxes and gets baguettes, then if that person opens the drink box, he/ she will not get tea.
- If everyone opens their drink boxes, they will not all get coffee.

Now, just as in the bipartite case, if everyone opens their food boxes and indeed all find baguettes inside, what inference can be made on the unopened drink box? By the second promise above, everyone can assume that if he/ she is the only person opening the drink box, then a cup of coffee will be found inside for sure. Because everyone can make an inference this way, then should everyone chose to open the drink box instead of the food box in the beginning, all of them should find coffee inside. But this contradicts the last promise.

Again using Table. 4.1, these promises can be written as probabilities:

$$P(0\ldots0|0\ldots0) > 0, \tag{4.37}$$

$$P(0\ldots0|0\ldots1) = 0, \tag{4.38}$$

$$\vdots$$

$$P(0\ldots0|1\ldots0) = 0, \tag{4.39}$$

$$P(1\ldots1|1\ldots1) = 0. \tag{4.40}$$

The lines from (4.38) to (4.39) represent n probabilities, whose settings are permutations of a n-bit string with a single one and $n-1$ zeros, which will also be written succinctly as $P(0\ldots0|\square(0\ldots1))$.

In the bipartite case, the four probabilities can be put into the CH inequality to be tested experimentally. In the multipartite case, a multiparty version of the CH inequality also exists, although it has been rediscovered many times [Cer04, GR10, WM12]. Because this inequality plays an important role later, it is given a name P^n:

$$P^n := P(0\ldots0|0\ldots0) - \sum_{\square} P(0\ldots0|\square(0\ldots1)) \leq P(1\ldots1|1\ldots1). \tag{4.41}$$

The local bound is established by the theorem below.

Theorem 6. The Bell polynomial for n systems

$$P^n := P(0\ldots0|0\ldots0) - \sum_{\vec{x}} P(0\ldots0|\vec{x}(0\ldots1)) - P(1\ldots1|1\ldots1)$$

is bounded under LHV as $P^n \geq 0$.

Proof. Assuming (2.17), joint probabilities are products of probabilities of each party (here we dropped λ because we are always integrating over $d\lambda$). The Bell polynomial P^n can be rewritten as:

$$\begin{aligned}
&P_1(0)P_2(0)\ldots P_n(0)\\
&-P_1(1)P_2(1)\ldots P_n(1)\\
&-(1-P_1(1))P_2(0)\ldots P_n(0)\\
&\vdots\\
&-P_1(0)\ldots P_{n-1}(0)(1-P_n(1)),
\end{aligned} \tag{4.42}$$

by noting

$$\begin{aligned}
P_i(0) &= P_i(0|0),\\
P_i(1) &= P_i(1|1).
\end{aligned}$$

The subscripts denote different parties and essentially show that probabilities of obtaining the same result by different parties are independent. An expansion gives:

$$P_1(0)P_2(0)\ldots P_n(0) - P_1(1)P_2(1)\ldots P_n(1) \tag{4.43}$$
$$+ P_1(1)P_2(0)\ldots P_n(0) - P_2(0)P_3(0)\ldots P_n(0) \tag{4.44}$$
$$+ P_1(0)P_2(1)\ldots P_n(0) - P_1(0)P_3(0)\ldots P_n(0) \tag{4.45}$$
$$\vdots$$
$$+ P_1(0)P_2(0)\ldots P_n(1) - P_1(0)P_2(0)\ldots P_{n-1}(0) \tag{4.46}$$

Note that the rows (4.44) to (4.46) are all less than or equal to 0, because for all $0 \leq P_i \leq 1$,

$$\prod_{i=1}^{n} P_i - \prod_{i=1}^{n-1} P_i. \tag{4.47}$$

Factoring the second term in (4.43) and the first term in (4.44), we have:

$$P_1(1)(P_2(0)\dots P_n(0) \square P_2(1)\dots P_n(1)). \tag{4.48}$$

If $P_2(0)\dots P_n(0) \square P_2(1)\dots P_n(1)$, then (4.48) $\square$ 0 (by using (4.47) on the first term in (4.43) and the second term in (4.44), and on rows (4.45) to (4.46)), and there is nothing more to prove. Otherwise

$$\begin{aligned} &P_1(1)(P_2(0)\dots P_n(0) \square P_2(1)\dots P_n(1)) \\ &\square(P_2(0)\dots P_n(0) \square P_2(1)\dots P_n(1)), \end{aligned} \tag{4.49}$$

which means (4.42) is smaller than or equals to

$$P_1(0)P_2(0)\dots P_n(0) \square P_2(1)P_3(1)\dots P_n(1) \tag{4.50}$$

$$+ P_1(0)P_2(1)\dots P_n(0) \square P_1(0)P_3(0)\dots P_n(0) \tag{4.51}$$

$$\vdots$$

$$+ P_1(0)P_2(0)\dots P_n(1) \square P_1(0)P_2(0)\dots P_{n\square1}(0). \tag{4.52}$$

Repeat the same procedure n times, each time assuming $\prod_k P_k(0) > \prod_k P_k(1)$ (otherwise we can terminate the proof). What has been shown after these n steps is (4.42) $\square$ $P_1(0)P_2(0)\dots P_n(0) \square P_1(0)P_2(0)\dots P_{n\square1}(0)$, whose right hand side is less than or equal to 0 by (4.47) [1]. $\square$

[1] An alternative proof is given in [GR10]

4.3 Violation of P^n By Almost All Symmetric States

It is time to show that almost all entangled symmetric states, with the exception of Dicke states, can satisfy (4.37) to (4.40), thus also violating P^n(4.41).

Consider any entangled permutation symmetric state (which is not a Dicke state) $|\psi\rangle$ with MPs $\{|\alpha_1\rangle, \ldots, |\alpha_n\rangle\}$ (degeneracy will be considered in the next chapter, here it does not matter if some of the MPs are degenerate). From the definition of the Majorana representation, it is not hard to see that (4.40) can be achieved by using the antipodal point of any MP as the projector of setting 1 outcome 1. If $|\alpha_i\rangle$ is an MP, $|\alpha_i\rangle^\perp$ its antipodal point, then by (3.29),

$$P(1\ldots1|1\ldots1) = |\langle\alpha_i^\perp|\otimes\underbrace{\ldots\otimes\langle\alpha_i^\perp|}_{n}\psi\rangle|^2 = 0. \tag{4.53}$$

Using $\{|\alpha_i\rangle, |\alpha_i\rangle^\perp\}$ as setting 1 also fixes the 0 outcome of setting 1 in (4.38) to (4.39). Also, because of the permutation symmetry of $|\psi\rangle$, (4.38) to (4.39) are either all satisfied or none of them is satisfied. The result of using $|\alpha_i\rangle$ as outcome 0 of setting 1 is that projecting $|\psi\rangle$ onto it, the resulting $(n-1)$-qubit state is also symmetric:

$$|\psi_i\rangle = \langle\alpha_i|\psi\rangle. \tag{4.54}$$

To satisfy the n probabilities (4.38) to (4.39), it is sufficient to use one of the MPs $|\alpha_j\rangle$ of $|\psi_i\rangle$ and its antipodal point $|\alpha_j\rangle^\perp$ as setting 0:

$$P(0\ldots0|(0\ldots1)) = |\langle(\langle\alpha_j^\perp|\otimes\underbrace{\ldots\otimes\langle\alpha_j^\perp|}_{n-1}\otimes\langle\alpha_i|)|\psi\rangle|^2 = 0. \tag{4.55}$$

But there is a catch: unless $|\alpha_j\rangle$ is different from all the MPs of $|\psi\rangle$, the first probability (4.37) will not hold. Luckily, the theorem below shows that unless $|\psi\rangle$ is a Dicke state, there is always a way of choosing $|\alpha_i\rangle$ such that at least one MP of $|\psi_i\rangle$ is different of all the MPs of $|\psi\rangle$. By using such a choice, all $n+2$ probabilities (4.37) to (4.40) are all satisfied, thus also violating P^n.

Theorem 7. Let $S := \{|\alpha_1\rangle, |\alpha_2\rangle, \ldots, |\alpha_n\rangle\}$ be the set of MPs of the state $|\psi\rangle$. Let $S_i := \{|\alpha_1\rangle, |\alpha_2\rangle, \ldots, |\alpha_{n-1}\rangle\}$ be the set of MPs of the state $|\psi_i\rangle = \langle\alpha_i|\psi\rangle$. Then $S_i \subseteq S$ iff $|\psi\rangle$ is a Dicke state up to rotations of the Bloch Sphere.

First we prove an important lemma upon which the rest of the proof will rely. The purpose of this lemma is to show that orthogonality conditions like (3.29) and (3.30) are only true if the state which we take the tensor product of (the bra half) is an MP of the permutation symmetric state (the ket half), and the order of tensor products depends on the degeneracy of the MP.

Lemma 8. If $|\psi\rangle$ is a permutation symmetric state of n qubits with (distinct) MPs $\{|\alpha_1\rangle, |\alpha_2\rangle, \ldots, |\alpha_l\rangle\}$, each having degeneracy $\{d_1, d_2, \ldots, d_l\}$, then

$$(\langle\alpha|)^{\otimes c}|\psi\rangle = 0 \tag{4.56}$$

if and only if $|\alpha\rangle = |\alpha_i^{\perp}\rangle$, $\langle\alpha_i^{\perp}|\alpha_i\rangle = 0$ for some i, and $c \geq n - d_i + 1$ (or equivalently, $d_i \geq n - c + 1$).

Proof. The if direction follows simply from expanding $|\psi\rangle$ using (3.27),(3.28) and the condition on MPs in the definition above. We will now focus on the only if direction.

First notice that if $(\langle\alpha|)^{\otimes c}|\psi\rangle = 0$ with $c \geq n$, then $(\langle\alpha|)^{\otimes n}|\psi\rangle = 0$. As explained above, $(\langle\alpha|)^{\otimes n}|\psi\rangle = 0$ is only possible if $|\alpha\rangle$ is an antipodal point of some MP of $|\psi\rangle$. Therefore $|\alpha\rangle = |\alpha_i^{\perp}\rangle$, $\langle\alpha_i^{\perp}|\alpha_i\rangle = 0$ for some i.

Now we want to show that $(\langle\alpha_i^{\perp}|)^{\otimes c}|\psi\rangle = 0$ implies $c \geq n - d_i + 1$. Instead we will show the equivalent statement: $c < n - d_i + 1$ (or equivalently $c \leq n - d_i$) implies $(\langle\alpha_i^{\perp}|)^{\otimes c}|\psi\rangle \neq 0$.

If $c = n - d_i$, then

$$(\langle\alpha_i^{\perp}|)^{\otimes n-d_i}|\psi\rangle \tag{4.57}$$

$$= (n-d_i)!(\langle\alpha_i^{\perp}|)^{\otimes n-d_i}|\underbrace{\alpha_1\alpha_2\ldots\alpha_i}_{\{1\ldots l\}\setminus\{i\}}\rangle(|\alpha_i\rangle)^{\otimes d_i} \neq 0. \tag{4.58}$$

To get from (4.57) to (4.58), we used that fact that when $c = n - d_i$, if we expand $|\psi\rangle$ using (3.27),(3.28), all other terms disappear. This term cannot be zero since by assumption no other MPs are equal to $|\alpha_i\rangle$.

When $c < n - d_i$, if $(\langle\alpha_i^{\perp}|)^{\otimes c}|\psi\rangle = 0$, then this would imply $(\langle\alpha_i^{\perp}|)^{\otimes n-d_i}|\psi\rangle = 0$, which contradicts the argument just given. Thus when $c \leq n - d_i$, $(\langle\alpha_i^{\perp}|)^{\otimes c}|\psi\rangle \neq 0$. □

The statement of Theorem 7 refers to the state $|\psi_i\rangle = \langle\alpha_i|\psi\rangle$. Neither the statement nor this proof depend on the choice of $|\alpha_i\rangle$. In fact what will be shown in this proof is that for any choice of $|\alpha_i\rangle$, for the conditions in Proposition 1 to be satisfied, all other MPs of $|\psi\rangle$ are either orthogonal to it or lie on top of it, which is the definition of a (possibly rotated) Dicke state. For convenience, we now fix this $|\alpha_i\rangle$ to be $|\alpha_1\rangle$, so $|\psi_i\rangle$ will now be fixed to be $|\psi_1\rangle$.

The state $|\psi\rangle$ can be decomposed into MPs $|\alpha_i^{d_i}\rangle$ as in (3.28). Similarly, the state $|\psi_1\rangle = \langle\alpha_1|\psi\rangle$ can be decomposed into

$$|\psi_1\rangle = K\sum_{\text{perm}}|\alpha_1^{m_1}\alpha_2^{m_2}\ldots\alpha_k^{m_k}\rangle, \tag{4.59}$$

$$\forall i \neq j, |\alpha_i\rangle \neq |\alpha_j\rangle, \sum_{i=1}^{k} m_i = n - 1. \tag{4.60}$$

Apart from $|\Psi_1\rangle$, there is another $(n-1)$-qubit permutation symmetric state useful to us, which is the state composed of all MPs of $|\Psi\rangle$ except $|\Phi_i\rangle$:

$$|\Psi_{\neq i}\rangle := K \sum_{perm} |\underbrace{\Phi_1 \ldots \Phi_n}_{\{1\ldots n\}-\{i\}}\rangle . \tag{4.61}$$

By using Lemma 8, we can prove two corollaries about the states $|\Psi_1\rangle$, $|\Psi_{\neq i}\rangle$ and the degeneracies of their MPs:

Corollary 9. $(\langle\Phi|)^{\otimes c}|\Psi_1\rangle = 0$ if and only if $|\Phi\rangle = |\Phi_i^{\perp}\rangle$, $\langle\Phi_i^{\perp}|\Phi_i\rangle = 0$ for some i and $c \leq n - m_i$ (or equivalently, $m_i \leq n - c$).

Proof. Follows directly from Lemma 8 by noticing that $|\Psi_1\rangle$ is a permutation symmetric state of $n-1$ qubits. □

Corollary 10. $(\langle\Phi|)^{\otimes c}|\Psi_{\neq i}\rangle = 0$ if and only if $|\Phi\rangle = |\Phi_j^{\perp}\rangle$, $\langle\Phi_j^{\perp}|\Phi_j\rangle = 0$ for some j, and 1). if $j \neq i$, then $c \leq n - d_j$ (or equivalently, $d_j \leq n - c$); 2). if $j = i$, then $d_j > 1$ and $c \leq n - d_j + 1$ (or equivalently, $d_j \leq n - c + 1$).

Proof. When $j \neq i$, the proof follows directly from Lemma 8, in the same way as the proof of the previous corollary. When $j = i$, the degeneracy of $|\Phi_j\rangle$ in $|\Psi_{\neq i}\rangle$ is $d_j - 1$, then by Lemma 8 it can only be zero for $d_i > 1$ and then $d_j - 1 \leq n - c \Leftrightarrow d_j \leq n - c + 1$. □

We now proceed to the three main lemmas of the proof. For clarity, we continue to use the notation that $|\Phi_i\rangle$ is an MP of $|\Psi\rangle$ with degeneracy d_i, and $|\Phi_i\rangle$ is an MP of $|\Psi_1\rangle$ with degeneracy m_i. We also use K to denote the global normalization constant.

Lemma 11. $\forall i, |\Phi_i\rangle = |\Phi_i\rangle$, with $m_i = d_i - 1$.

Proof. First of all, if $d_i = 1$, then the statement is always true, so now we will focus on the case when $d_i > 1$.

From Corollary 9, $(\langle\Phi_i^{\perp}|)^{\otimes n-d_i+1}|\Psi_1\rangle = 0$ if and only if $|\Phi_i\rangle = |\Phi_i\rangle$ and $m_i \leq n - (n - d_i + 1) = d_i - 1$. So it suffices to show that $(\langle\Phi_i^{\perp}|)^{\otimes n-d_i+1}|\Psi_1\rangle = 0$.

$|\Psi_1\rangle$ can be decomposed into

$$|\Psi_1\rangle = K \sum_l \langle\Phi_1|\Phi_l\rangle |\Psi_{\neq l}\rangle . \tag{4.62}$$

Using this decomposition, we have

$$(\langle\Phi_i^{\perp}|)^{\otimes n-d_i+1}|\Psi_1\rangle \tag{4.63}$$
$$= K(\langle\Phi_i^{\perp}|)^{\otimes n-d_i+1} \sum_l \langle\Phi_1|\Phi_l\rangle |\Psi_{\neq l}\rangle \tag{4.64}$$
$$= K \sum_l \langle\Phi_1|\Phi_l\rangle (\langle\Phi_i^{\perp}|)^{\otimes n-d_i+1}|\Psi_{\neq l}\rangle \tag{4.65}$$
$$= 0, \tag{4.66}$$

where we used both cases in Corollary 10 to get from (4.65) to (4.66). □

Lemma 12. $\forall i$, if $|\Box_i\rangle = |\Box_i\rangle$ and $m_i \neq d_i$, then $\langle\Box_1|\Box_i\rangle = 0$.

Proof. By Corollary 9, if $|\Box_i\rangle = |\Box_i\rangle$ and $m_i \neq d_i$, then $(\langle\Box_i^?|)^{\otimes n-d_i}|\ _1\rangle = 0$.

Using the same decomposition of $|\ _1\rangle$ as in the proof of the previous lemma, we have:

$$(\langle\Box_i^?|)^{\otimes n-d_i}|\ _1\rangle \tag{4.67}$$
$$= K\sum_j \langle\Box_1|\Box_j\rangle(\langle\Box_i^?|)^{\otimes n-d_i}|\ _{=j}\rangle\,. \tag{4.68}$$

From Corollary 10, we can deduce that the only term which does not vanish in the sum is when $j = i$, thus we have:

$$(\langle\Box_i^?|)^{\otimes n-d_i}|\ _1\rangle \tag{4.69}$$
$$= K\langle\Box_1|\Box_i\rangle\underbrace{(\langle\Box_i^?|)^{\otimes n-d_i}|\ _{=i}\rangle}_{\in\mathbb{C}} \tag{4.70}$$
$$= 0. \tag{4.71}$$

Since neither K nor $\Box$ in (4.70) is 0, we can conclude that $\langle\Box_1|\Box_i\rangle = 0$.

□

Lemma 13. $\forall i$, if $|\Box_i\rangle = |\Box_i\rangle$ then $m_i \neq d_i$.

Proof. If $m_i > d_i$, then by Lemma 12 $\langle\Box_1|\Box_i\rangle = 0$. We can write $|\Box_i^?\rangle$ as $|\Box_1\rangle$.

Partially expanding $|\ \rangle$ in the $\{|\Box_1\rangle, |\Box_1^?\rangle\}$ basis gives

$$|\ \rangle = K(|\Box_1\rangle|\ _1\rangle + |\Box_1^?\rangle|\ ^\perp{}_1\rangle) \tag{4.72}$$

Now consider $(\langle\Box_i^?|)^{\otimes n-d_i}|\ \rangle$:

$$(\langle\Box_i^?|)^{\otimes n-d_i}|\ \rangle \tag{4.73}$$
$$= (\langle\Box_i^?|)^{\otimes n-d_i}K(|\Box_1\rangle|\ _1\rangle + |\Box_1^?\rangle|\ ^\perp{}_1\rangle) \tag{4.74}$$
$$= K(\langle\Box_i^?|)^{\otimes n-d_i-1}|\ _1\rangle \tag{4.75}$$
$$= 0, \tag{4.76}$$

where we used the fact that $|\Box_i^?\rangle = |\Box_1\rangle$ to go from (4.74) to (4.75). We used Corollary 9 and our assumption that $m_i > d_i$ to go from (4.75) to (4.76).

Clearly, the conclusion that $(\langle\Box_i^?|)^{\otimes n-d_i}|\ \rangle = 0$ contradicts Lemma 8, so $m_i \neq d_i$. □

Combining Lemmas 11, 12, 13, we get the main corollary which will prove the theorem.

Corollary 14. $\forall i, |\Omega_i\rangle = |\Omega_i\rangle$ with degeneracies such that either 1). $m_i = d_i$, $\langle\Omega_1|\Omega_i\rangle = 0$ or 2). $m_i = d_i - 1$, $\langle\Omega_1|\Omega_i\rangle \neq 0$.

To finish the proof, we first observe that since $\sum_k m_k = n - 1$, and $n = \sum d_i$, we have

$$\sum_k m_k = n - 1 = \Big(\sum_k d_k\Big) - 1 \tag{4.77}$$

$$= \underbrace{(d_i - 1)}_{\text{group 1}} + \underbrace{\sum_{j\neq i} d_j}_{\text{group 2}}, \tag{4.78}$$

where we have separated the d_i into two groups, group one, where the minus one is associated a particular d_i, and group two, the remaining d_j. Next we consider what is implied by the condition $S_i = S$ in Proposition 1 - i.e., where all Majorana points $|\Omega_i\rangle$ coincide with $|\Omega_i\rangle$. It is clear from Corollary 14 and (4.77) that this can only be done if group one is given by d_1, and group two comes from one Majorana point which is orthogonal to $|\Omega_1\rangle$. Thus there are only two Majorana points of $|\ \rangle$, $|\Omega_1\rangle$ and $|\Omega_2\rangle = |\Omega_1^{\perp}\rangle$. This is exactly a Dicke state up to rotation of the Majorana sphere. Substitute any $|\Omega_i\rangle$ for $|\Omega_1\rangle$ in the beginning, the reasoning above is still valid.

4.4 Violation of P^n By All Symmetric States

For Dicke states, although (4.37) to (4.40) can not be satisfied because Theorem 7 fails, meaning the state $|\psi_i\rangle = \langle\Box_i| \; \rangle$ is also a Dicke state (this can be verified easily since for a Dicke state, all its MPs are either $|0\rangle$ or $|1\rangle$). But the theorem below show that it is nevertheless possible to violate P^n by a relatively simple choice of settings. But unlike the case for non-Dicke states, this theorem is not constructive: for non-Dicke states, calculating the setting 0 only involves finding the roots of a polynomial, while the theorem below only shows the existence of a setting 0 which can violate P^n.

Theorem 15. There exists an angle $0 < \Box < \Box$ such that all Dicke states $|S(n,k)\rangle$ ($\{k,n\} \in \mathbb{N}, 1 < k < n$) violate the inequality in Theorem 6 when using the measurement setting $\{|+\rangle, |\Box\rangle\}$ as setting 0 and $\{\cos\frac{\Box}{2}|0\rangle \Box \sin\frac{\Box}{2}|1\rangle, \sin\frac{\Box}{2}|0\rangle + \cos\frac{\Box}{2}|1\rangle\}$ as setting 1.

Proof. First we will write the inequality in Theorem 6 as a function of n, k and $\Box$.

$$
\begin{aligned}
&P(0\ldots0|0\ldots0) \\
&= |\langle +,\ldots,+|S(n,k)\rangle|^2 \\
&= \left((\tfrac{1}{\sqrt{2}})^n \sqrt{\binom{n}{k}} \Box \tfrac{1}{2} \sqrt{\binom{n}{k}}\right)^2, \qquad (4.79)\\
&P(0\ldots0|1\ldots0) \\
&= |(\cos\tfrac{\Box}{2}\langle 0| \Box \sin\tfrac{\Box}{2}\langle 1|) \Box \langle +,\ldots,+|S(n,k)\rangle|^2 \\
&= \left((\tfrac{1}{\sqrt{2}})^{n\Box 1} \sqrt{\binom{n}{k} \Box \tfrac{1}{2}} \left(\cos\tfrac{\Box}{2}\sqrt{\binom{n}{k} \Box \binom{n\Box 1}{k}} \Box \sin\tfrac{\Box}{2}\sqrt{\binom{n\Box 1}{k\Box 1}}\right)\right)^2, \qquad (4.80)\\
&P(1\ldots1|1\ldots1) \\
&= |\langle -|^{n\Box}(\sin\tfrac{\Box}{2}\langle 0| + \cos\tfrac{\Box}{2}\langle 1|)|S(n,k)\rangle|^2 \\
&= \left(\sqrt{\binom{n}{k} \Box \tfrac{1}{2}} \sqrt{\binom{n}{k}} (\cos\tfrac{\Box}{2})^k (\sin\tfrac{\Box}{2})^{n\Box k}\right)^2. \qquad (4.81)
\end{aligned}
$$

To simplify our calculations, we divide each probability (4.79) - (4.81) by $\sqrt{\binom{n}{k}}$. Doing this rescales the Bell polynomial in Theorem 6, which will not change its positivity

property.

$$P^0(0\ldots0j0\ldots0) = (\frac{1}{2})^n, \tag{4.82}$$

$$P^0(0\ldots0j1\ldots0) = (\frac{1}{2})^{n-1}\binom{n-2}{k}(\cos\frac{\theta}{2}^{n-1}\sin\frac{\theta}{2}^{k-1})^2, \tag{4.83}$$

$$P^0(1\ldots1j1\ldots1) = (\cos\frac{\theta}{2})^{2k}(\sin\frac{\theta}{2})^{2n-2k}. \tag{4.84}$$

Because of the permutation symmetry of $|S(n,k)\rangle$, the n probabilities $P^0(0\ldots0j1\ldots0)$ to $P^0(0\ldots0j0\ldots1)$ are all equal. After some simplification, the rescaled Bell operator becomes:

$$\begin{aligned} P^0(n,k,\theta) = {} & (\frac{1}{2})^n \\ & - n(\frac{1}{2})^{n-1}(\frac{n-k}{n}\cos\frac{\theta}{2} - \frac{k}{n}\sin\frac{\theta}{2})^2 \\ & - (\cos\frac{\theta}{2})^{2k}(\sin\frac{\theta}{2})^{2n-2k}. \end{aligned} \tag{4.85}$$

First we note some properties of (4.83) and (4.84). For (4.83), it can always reach 0 for all n,k when $\tan\frac{\theta}{2} = \frac{n-k}{k}$. (4.84) is 0 when $\theta = 0$ and $\theta = \pi$, and it reaches its maximum when $(\cos\frac{\theta}{2})^{2k} = (\sin\frac{\theta}{2})^{2n-2k}$. Also, if we fix θ and n, its derivative with respect to k is

$$2(\cos\frac{\theta}{2})^{2k}(\sin\frac{\theta}{2})^{2n-2k}(\log\cos\frac{\theta}{2} - \log\sin\frac{\theta}{2}), \tag{4.86}$$

which means that for fixed θ and n, when $\theta < \frac{\pi}{2}$, (4.84) is monotonically increasing with respect to k, when $\theta > \frac{\pi}{2}$, (4.84) is monotonically decreasing with respect to k, and when $\theta = \frac{\pi}{2}$, (4.84) is independent of k.

Now consider the equation $n \cdot (4.83) = (\frac{1}{2})^{n+1}$:

$$n(\frac{1}{2})^{n-1}(\frac{n-k}{n}\cos\frac{\theta}{2} - \frac{k}{n}\sin\frac{\theta}{2})^2 = (\frac{1}{2})^{n+1}, \tag{4.87}$$

$$\Rightarrow \quad (\frac{n-k}{n}\cos\frac{\theta}{2} - \frac{k}{n}\sin\frac{\theta}{2})^2 = \frac{1}{4n}. \tag{4.88}$$

After taking the square root of both sides, moving the term with $\sin\frac{\theta}{2}$ to one side, substuting $\sin\frac{\theta}{2}$ with $\sqrt{1-\cos^2\frac{\theta}{2}}$ and square both sides again, (4.88) can be seen as a quadratic equation having $\cos\frac{\theta}{2}$ as unknown. It may have 0, 1 or 2 roots when θ takes

any value in the interval $[0, \pi]$. If it has no root, then we know $n \cdot (4.83) < (\frac{1}{2})^{n+1}$ for all values of θ (because (4.83) can always reach 0). Also note that (4.84) is zero when $\theta = 0$ and $\theta = \pi$, so if (4.88) has no root then (4.85) > 0 when $\theta = 0$ and $\theta = \pi$. Similarly, if (4.88) has one root, we can show that (4.85) > 0 when $\theta = 0$ or $\theta = \pi$, depending on the location of the root: when the root is smaller than $\frac{\pi}{2}$, we take $\theta = \pi$, otherwise we take $\theta = 0$.

We now proceed to the case when (4.88) has two roots. By solving the quadratic equation we can have closed forms of the two roots in $[0, \pi]$, which we call θ_+ and θ_-:

$$\cos\frac{\theta_+}{2} = \frac{k\sqrt{n} - n\sqrt{n} + \sqrt{8k^4 - k^2n - 8k^3n + 4k^2n^2}}{2(2k^2 - 2kn + n^2)} \tag{4.89}$$

$$\cos\frac{\theta_-}{2} = \frac{k\sqrt{n} - n\sqrt{n} - \sqrt{8k^4 - k^2n - 8k^3n + 4k^2n^2}}{2(2k^2 - 2kn + n^2)} \tag{4.90}$$

What we want to show now is that for any fixed n, $(\cos\frac{\theta_+}{2})^{2k}(\sin\frac{\theta_+}{2})^{2n-2k} < (\frac{1}{2})^{n+1}$ and $(\cos\frac{\theta_-}{2})^{2k}(\sin\frac{\theta_-}{2})^{2n-2k} < (\frac{1}{2})^{n+1}$ for all k. From the monotonicity of (4.84), it suffices to show that if the inequalities hold for $k = \frac{n}{2}$, then they must hold for all k. So for now we will only consider the case when n is even. We also restrict ourselves to the positive root. The argument below is symmetric, with appropriate sign/ monotonicity changes it also applies to the negative root.

When n is even and $k = \frac{n}{2}$, $(\cos\frac{\theta_+}{2})^{2k}(\sin\frac{\theta_+}{2})^{2n-2k}$ simplify to:

$$(\frac{1}{32})^k \underbrace{(\frac{\sqrt{k} + \sqrt{k(4k-1)}}{k})^{2k}}_{(\alpha)} \underbrace{(2 + \frac{\sqrt{(4k-1)}}{k})^k}_{(\alpha\alpha)}. \tag{4.91}$$

To get the upper bound on (4.91), we first rewrite (α) and ($\alpha\alpha$):

$$(\alpha) = ((\frac{\sqrt{k} + \sqrt{k(4k-1)}}{k})^2)^k = (4 \cdot \frac{2\sqrt{4k-1}}{k})^k \tag{4.92}$$

$$= (4 \cdot \frac{4\sqrt{k - \frac{1}{4}}}{k})^k = (4(1 \cdot \frac{\sqrt{k - \frac{1}{4}}}{k}))^k, \tag{4.93}$$

$$(\alpha\alpha) = (2 + \frac{2\sqrt{k - \frac{1}{4}}}{k})^k = (2(1 + \frac{\sqrt{k - \frac{1}{4}}}{k}))^k. \tag{4.94}$$

Substitute (4.93) and (4.94) into (4.91), we get

$$(\frac{1}{32} \cdot 4 \cdot 2 \cdot (1 \cdot \frac{k - \frac{1}{4}}{k^2}))^k = (\frac{1}{2})^{2k}(1 \cdot \frac{k - \frac{1}{4}}{k^2})^k. \tag{4.95}$$

Note that $\lim_{n\to\infty}(1 - \frac{1}{x})^x = e^{-1}$, and it approaches this limit from below. We thus

have

$$(1 - \frac{k - \frac{1}{4}}{k^2})^k < (1 - \frac{1}{k})^k \cdot e^{-1} < \frac{1}{2} \tag{4.96}$$

From (4.95) and (4.96), we get the desired bound

$$\begin{aligned}&(\cos\frac{\pi}{2})^{2k}(\sin\frac{\pi}{2})^{2n-2k}\\ &< (\frac{1}{2})^{2k}(1 - \frac{1}{k})^k < (\frac{1}{2})^{2k+1}\end{aligned} \tag{4.97}$$

To show the theorem holds for odd n, we let $n = 2k + 1$ for some k, then (4.85) becomes

$$(\frac{1}{2})^{2k+1} \cdot (2k+1)(\frac{1}{2})^{2k}(\frac{k+1}{2k+1}\cos\frac{\theta}{2} \cdot \frac{k}{2k+1}\sin\frac{\theta}{2})^2 \tag{4.98}$$

$$\cdot(\cos\frac{\theta}{2})^{2k}(\sin\frac{\theta}{2})^{2k+2} \tag{4.99}$$

$$= \frac{1}{2}((\frac{1}{2})^{2k} \tag{4.100}$$

$$\cdot(2k+1)(\frac{1}{2})^{2k-1}(\frac{k+1}{2k+1}\cos\frac{\theta}{2} \cdot \frac{k}{2k+1}\sin\frac{\theta}{2})^2 \tag{4.101}$$

$$\cdot 2(\sin\frac{\theta}{2})^2(\cos\frac{\theta}{2})^{2k}(\sin\frac{\theta}{2})^{2k}) \tag{4.102}$$

It is easy to see that $\frac{k+1}{2k+1} \cdot \frac{1}{2}$ and $\frac{k}{2k+1} \cdot \frac{1}{2}$, so (4.101) can be bounded by

$$(4.101) \cdot 2k(\frac{1}{2})^{2k-1}(\frac{1}{2}\cos\frac{\theta}{2} \cdot \frac{1}{2}\sin\frac{\theta}{2})^2. \tag{4.103}$$

Assume $\cos\frac{\pi}{2} \cdot \pm\frac{1}{\sqrt{2}}$ (the other case will be discussed below), which means $2(\sin\frac{\theta}{2})^2 \cdot 1$, we can bound (4.102) by:

$$(4.102) \cdot (\cos\frac{\theta}{2})^{2k}(\sin\frac{\theta}{2})^{2k}. \tag{4.104}$$

Substitute (4.103) and (4.104) into (4.101) and (4.102), we have

$$
\begin{aligned}
&(\frac{1}{2})^{2k+1} \\
&\quad (2k+1)(\frac{1}{2})^{2k}(\frac{k+1}{2k+1}\cos\frac{\theta}{2} \quad \frac{k}{2k+1}\sin\frac{\theta}{2})^2 \\
&\quad (\cos\frac{\theta}{2})^{2k}(\sin\frac{\theta}{2})^{2k+2} \\
&\frac{1}{2}((\frac{1}{2})^{2k} \\
&\quad 2k(\frac{1}{2})^{2k-1}(\frac{1}{2}\cos\frac{\theta}{2} \quad \frac{1}{2}\sin\frac{\theta}{2})^2 \\
&\quad (\cos\frac{\theta}{2})^{2k}(\sin\frac{\theta}{2})^{2k}),
\end{aligned}
$$

which is just two times the expression for when $n = 2k$. If $\cos\frac{\theta_+}{2} \quad \pm\frac{1}{2}$, we can let $n = 2k \quad 1$, and get a similar argument. For the negative root, the the same reasoning follows.

To sum up, for even n, if we let $n \quad (4.83) = (\frac{1}{2})^{n+1}$, then $(4.84) < (\frac{1}{2})^{n+1}$, which means $(4.85) > 0$. For odd n, depending on the value of $\cos\frac{\theta_+}{2}$ or $\cos\frac{\theta_-}{2}$, we can always bound (4.85) by considering the closest even n. So (4.85) can always be positive for some value of θ.

□

Chapter 5

Degeneracy and its Consequences

Multipartite states are an important resource for many areas of quantum information. Understanding the features that give rise to their usefulness is still a question under much investigation. Entanglement has become recognized as a key feature. However, the question becomes involved in the multipartite settings with different classes of entanglement [HHHH09] having potentially different roles in recognizing good resources. Intimately related to entanglement is notion of nonlocality [EPR35, Bel64], though it is known they are not the same [Wer89].

As well as the general interest in exploring the texture of multipartite state space, there is some practical interest in understanding the relationship between entanglement and nonlocality. Using the entanglement or nonlocal properties of multipartite states in the real world poses many experimental challenges. Unavoidable experimental inaccuracies like misalignment, noise and detector inefficiencies can render the outcome of an experiment meaningless. In quantum cryptography, for example, the presence of noise and detector inefficiencies can mask effective effective attacks on the security of the key distribution protocol [QFLM07, XQL10, LWW+10]. In entanglement theory, misalignment when trying to witness entanglement can lead to mistakes [BGLP11]. The answer from theorists is to make tangible claims without any assumptions about the measurement device, hence the name device independent. There is a natural connection to discussions of non-locality since Bell type arguments do not rely on any statements about measurements, only their statistics. Device independent proofs and tests have already been used extensively in quantum cryptography and secure communications [ABG+07, PAB+09, GBHA10, CK11, MPA11, PAM+10], and device independent entanglement witnesses [BGLP11] have been proposed. Recent results have shown device independent tests which are able to discriminate states that are inequivalent under local unitaries and permutation of systems (LUP) [BSV12].

This chapter and the next chapter furthers the study of nonlocality of symmetric states, studying deeper how the nonlocality exposed is related to entanglement classes

and the usefulness of the states. In this chapter, the focus will be on degeneracy, with consequences from persistency of nonlocality to device independent classification of states. These results also show the intricate relationship between multipartite nonlocality and multipartite entanglement.

5.1 Degeneracy and Persistency of Nonlocality

In Chapter 3, if a symmetric state $|\psi\rangle$ has degenerate MPs, then (3.30) shows that even a subset of players can make the probability $P(1\ldots1|1\ldots1)$ zero. In fact, if there is an MP of $|\psi\rangle$ with degeneracy d, then in addition to (4.37) - (4.40)

$$P(\underbrace{1\ldots1}_{n\geq 1}|\underbrace{1\ldots1}_{n\geq 1}) = 0, \tag{5.1}$$

$$P(\underbrace{1\ldots1}_{n\geq 2}|\underbrace{1\ldots1}_{n\geq 2}) = 0, \tag{5.2}$$

$$\vdots$$

$$P(\underbrace{1\ldots1}_{n\geq d+1}|\underbrace{1\ldots1}_{n\geq d+1}) = 0. \tag{5.3}$$

The probabilities (5.1) to (5.3) are computed by first tracing out 1, 2, . . . , $(d-1)$ players (because of the symmetry of the state $|\psi\rangle$, it does not matter which players to trace out), then performing a projective measurement using the MP basis and the reduced state. It should be noted that because the Majorana representation requires that any $|\phi\rangle$ which satisfies

$$(\underbrace{h^{\dagger}|\phi\rangle\otimes\ldots\otimes h^{\dagger}|\phi\rangle}_{n})|\psi\rangle = 0, \tag{5.4}$$

for a symmetric $|\psi\rangle$ to be an MP of $|\psi\rangle$, and only an MP $|\phi\rangle$ with degeneracy d satisfies

$$(\underbrace{h^{\dagger}|\phi\rangle\otimes\ldots\otimes h^{\dagger}|\phi\rangle}_{n\geq 1})|\psi\rangle = 0, \tag{5.5}$$

$$\vdots$$

$$(\underbrace{h^{\dagger}|\phi\rangle\otimes\ldots\otimes h^{\dagger}|\phi\rangle}_{n\geq d+1})|\psi\rangle = 0, \tag{5.6}$$

if all players measure trust their measurement devices and can guarantee that they measure in the same bases, then the only way to satisfy (4.37) - (4.40) and (5.1) - (5.3) for a symmetric state is by using a symmetric state where at least one MP has degeneracy d.

It is also obvious that by subtracting (5.1) - (5.3) from P^n a new inequality can be

defined:

$$Q_d^n := P^n \otimes P(\underbrace{1,\ldots,1}_{n-1}|\underbrace{1,\ldots,1}_{n-1}) \otimes \ldots \otimes P(\underbrace{1,\ldots,1}_{n-d+1}|\underbrace{1,\ldots,1}_{n-d+1}) \otimes 0. \tag{5.7}$$

The new inequality, Q_d^n, still keeps 0 as its local realistic bound because only nonnegative numbers are subtracted from P^n. Some persistency of nonlocality into subsets of players can be seen from the violation of Q_d^n: having a symmetric state with the right degeneracy guarantees its violation.

In addition to the persistency of nonlocality, the connection shown in Chapter 3 between the degeneracy of MPs and the SLOCC classification of states gives another use of degeneracy, as a way to classify symmetric states: two symmetric states with different degeneracy of MPs necessarily belong to different SLOCC classes because SLOCC operations (i.e. Möbius transformations) can not change the degeneracy of MPs.

However, in reality, the notion of degeneracy is not as strong as one might like. For example, moving one degenerate MP by an arbitrarily small distance give a symmetric state almost exactly like the original state. But since the degeneracy of its MPs has now been changed, the probabilities (5.1) to (5.3) can no longer be satisfied perfectly, although Q_d^n may still be violated, with a violation almost identical to the one obtained from the original state. Also, it is impossible to guarantee that all players measure in the same bases in practice, due to inevitable experimental imprecision. One may append to the list of conditions (5.1) to (5.3) some extra conditions which effectively imply the state is symmetric with respect to the basis given by measurement setting 1. However, whilst one may be able to define Hardy type paradoxes which can be used to identify classes of entanglement in this way, these conditions would be difficult to fit into an inequality, and further it is not clear that it would be possible at all that such inequalities would also strictly separate degeneracy classes.

5.2 Device Independent Classification of States

Although there are clear connections between the violation of Q_d^n and SLOCC entanglement classes through degeneracy of MPs - degeneracy d guarantees always violation of Q_d^n - the relationship is not as clear as we might like. An immediate question is that even though the violation of Q_d^n is guaranteed by the prescription using the Majorana representation, what about the maximal violation? Can we say that degeneracy guarantees that the level of violation stays high? Although we no longer have the analytic tools for general violation, we will see that numerics indicate this is the case, at least for W states. A deeper question though is what we can really understand from this. We would really like to know if it is possible to use these ideas and results to separate classes of states - so that different classes can really be differentiated by their nonlocal properties. This would lead to new ways of searching for new applications of states, as well as ways of probing the texture of multipartite states. To answer this, we will first go more into the subtle questions surrounding the classification of states, and then we will see some examples of how some separation of classes can be made.

On a practical level, it seems clear that different multipartite entangled states have different entanglement and locality properties. Famously GHZ states are highly nonlocal, but are highly sensitive to loss of systems - losing even one system takes them to a separable (hence 'local' state), whereas W states do not have the same extreme nonlocality [Cab01], but losing systems does not destroy the entanglement. In turn, different types of states may have different uses for quantum information.

The question of how to classify states in terms of entanglement and locality is a difficult one, particularly when we want to talk about how different classes might be meaningful either for different quantum information tasks, or their potential roles in many-body physics. Within entanglement theory, the most standard approach is to use the SLOCC classification (cf. Chapter 2, Section 2.3). Intuitively this classification is appealing since it separates states which cannot be reached from each other in the distributed setting, even with the aid of classical communication.

In terms of how one might classify states with respect to locality, there are several approaches. The standard setting for locality questions is one in which parties are not allowed to communicate classically - at least not after they have been told what bases to measure in, they may do before hand, for example to share classical randomness. Several options arise. In [HSG+11] it is proposed that a reasonable classification is to consider equivalence under local unitaries and permutation of systems (we denote this LUP). One may also consider states equivalent under local operations, which is in turn equivalent to local unitaries (we donate this LU). When considering correlations alone, without necessarily taking recourse to quantum states, in [GWAN12] a classification is presented called wiring and classical communication prior to inputs (WCCPI) - the

wiring is essentially the idea of using multiple copies of the resource (which could be a quantum state or 'box' giving a certain probability distribution) and allowing different ways of combining them. We do not consider the WCCPI classification further here, and rather focus on single copy classifications.

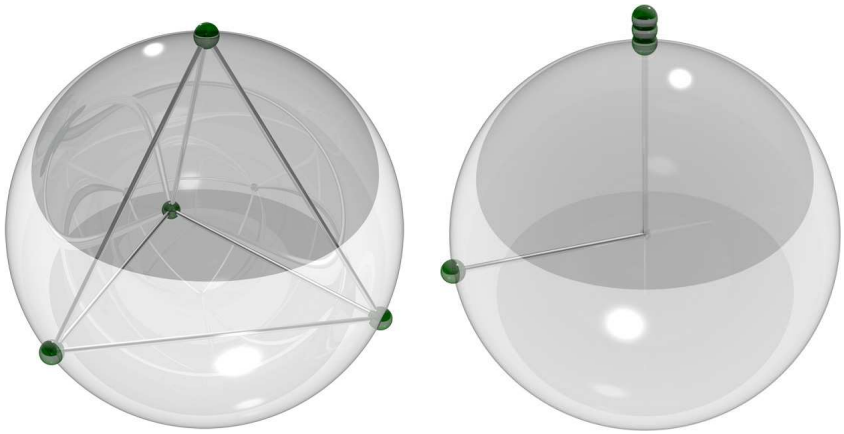
For all the classifications mentioned above, however, several difficulties emerge, which seem to limit their usefulness. First of all, there can be an infinite continuum of classes (for LUP and LU this is already true for two qubits, for SLOCC it is true for four or more [DVC00]). Second, and related to this, it is possible to have two states which are arbitrarily close to each other which are in different classes. This means that two states, which behave in almost exactly the same way for all possible experiments, can be in different classes. It is clear then that it is not possible to separate all classes of states in terms of their physical properties and in turn that the physical properties cannot be sensitive to all these classifications. Nevertheless, there does seem to be some difference between states, which can be identified through these classifications. For example, as we saw earlier, states of certain classes guarantee resistance of correlations to loss of systems, for both the LUP [BSV12] and the SLOCC [WM12] classifications (through the degeneracy of MPs as mentioned earlier). In [BSV12] this was used to separate two LUP classes in a device independent way.

Here we will use our inequalities to identify different sets of LUP classes of states of four qubits, hence also, in a device independent way. The LUP and LU classifications are well suited to discriminate via inequality violation because the maximum violation of an inequality is searched for over all measurement bases - which is equivalent to searching over all local unitaries. Thus, if we can say that a particular state cannot violate an inequality more than a certain amount (using SDP techniques for example, as we do here), this means that no state in the same LUP class can either. If the state is symmetric it also means no state in the same LUP state can either. The states we choose are also in different SLOCC classes (note, however, that the fact that no LU or LUP equivalent state can violate more than the amount we state does not necessarily mean that there does not exist an SLOCC equivalent state which can). Since this is done via violation of Bell-like inequalities - which makes no recourse to what measurements are made, this classification is done in a device independent way.

For the classification, we will consider three states: the tetrahedron state $|T\rangle = \frac{1}{3}|S(4,0)\rangle + \frac{2}{3}|S(4,3)\rangle$, the 4-qubit GHZ state $|GHZ_4\rangle = \frac{1}{\sqrt{2}}(|0000\rangle + |1111\rangle)$, and the state $|000+\rangle = K\sum_{perm}|000+\rangle = \frac{2}{\sqrt{5}}|0000\rangle + \frac{1}{\sqrt{5}}|S(4,1)\rangle$, which are all SLOCC-inequivalent (cf. Chapter 3, Section 3.1.2). We will consider them in two groups: one group consists of $|T\rangle$ and $|000+\rangle$, with differing degeneracy, the other group consists of $|T\rangle$ and $|GHZ_4\rangle$, with the same degeneracy. These are represented in Fig. 5.1 and 5.2 respectively. We will use numerical maximum violation of P^4 and Q_3^4 obtained from SDP to discriminate the states in a device independent way in each group.

The equivalence of symmetric states under LU and LUP is given simply by the MP distribution up to rotation of the sphere. This is because any local unitary taking a symmetric state to a symmetric state can be understood as a rotation of the sphere [BKM+09, MKG+10] (and that permutation obviously do not change a symmetric state). Thus each of the states we study here are LU and LUP inequivalent. As mentioned, the fact that we search for violation of inequalities over all measurements means that the bounds we present hold for all LU and LUP equivalent states.

For the first group, shown in Fig. 5.1, the results are shown in Table 5.1. Note that although we do not restrict the measurement bases for $|T\rangle$, as the degeneracy of the state $|000+\rangle$ is very high, we need to restrict the bases to get realistic SDP bounds.



(a) (b)

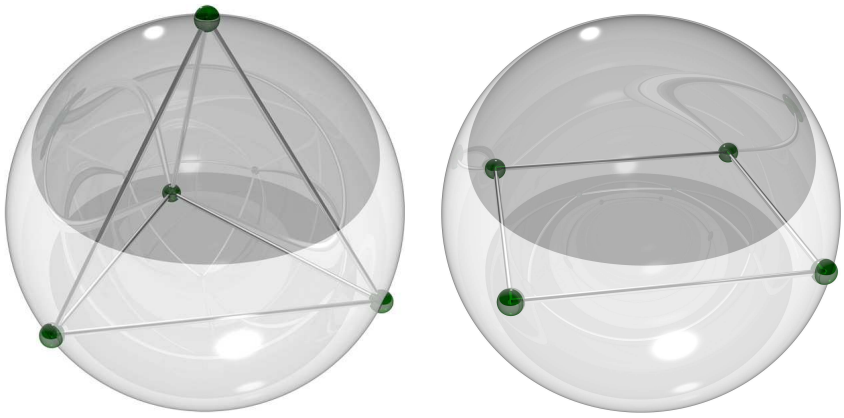
Figure 5.1: The tetrahedron state (a) and the state $|000+\rangle$ (b) in the Majorana representation.

State	P^4	Q_3^4
$\|T\rangle$	0.1745	-0.0609
$\|000+\rangle$	0.0142	0.0141

Table 5.1: SDP bounds on the maximum violation of P^4 and Q_3^4 for $|T\rangle$ and $|000+\rangle$. Because of computational difficulties, the values for $|000+\rangle$ assume that all parties measure in the same basis (a numerical optimization over the four Euler angles in the two measurement settings indicate this is still optimal). We thus have that a violation of Q_3^4 implies the state is not in the LU class of $|T\rangle$.

Table. 5.2 shows the bounds for P^4 and Q_3^4 for the second group, shown in Fig. 5.2, obtained using semidefinite programming techniques described in Section 2.5, without restricting the measurement bases of parties.

From these tables, one can easily envisage device independent tests to discriminate



(a) (b)

Figure 5.2: The tetrahedron state (a) and the 4-qubit GHZ state (b) in the Majorana representation.

State	P^4	Q_3^4
$\|T\rangle$	0.1745	-0.0609
$\|GHZ_4\rangle$	0.1241	0.0563

Table 5.2: SDP bounds on the maximum violation of P^4 and Q_3^4 for $|T\rangle$ and $|GHZ_4\rangle$. We thus have that a violation of $P^4 > 0.1241$ implies the state is not in the LU class of $|GHZ_4\rangle$, and a violation of Q_3^4 implies the state is not in the LU class of $|T\rangle$.

the LUP classes in each group.

For the first group, because of the restriction on measurement bases, we have a weaker test. Despite our best numerical checks and the seemingly reasonable assumption on the restriction of measurement of bases, we cannot guarantee that if a state has a violation of P^4 greater than 0.0142, it is not in the LUP $|000+\rangle$ class. However, we can still conclude that if a state violates Q_3^4 then it cannot be in the $|T\rangle$ class, but must be in the $|000+\rangle$ LUP class.

In the second group, if the P^4 test gives a violation 0.1241, then the state must not be in the $|GHZ_4\rangle$ LUP class, so must be in the $|T\rangle$ class. Similarly, if the Q_3^4 gives any violation at all, the state cannot be in the $|T\rangle$ LUP class and must be in the $|GHZ_4\rangle$ class. In this case, even though there is no degeneracy, separation can be seen using Q_3^4.

The example states chosen sit in different SLOCC classes. This was done by bounding the possible violation of inequalities using SDP techniques. Going above four qubits seems difficult as the numerics quickly get difficult with more parties, though simple basis checking numerics indicate that the W and GHZ states may be separated in this way. This furthers the discussion about how entanglement classifications can be interpreted using nonlocal features. On the one hand we have the general statement that degeneracy

of MPs guarantees persistency of correlations [WM12] to subsystems. This is true for all states, not just specific examples such as those expended upon here. We see that certain "example" states such as the $|000+\rangle$ and $|W\rangle$ states may be separated from less degenerate states using this fact. This can be compared to the robustness of nonlocality under system loss [BSV12] [BV12]. On the other hand, we also saw an example with the GHZ and T states where Q_d^n can be used to discriminate different classes, not related to degeneracy ($|T\rangle$ and $|GHZ\rangle$, both with degeneracy one). Intriguingly, we also remark that these states naturally appear in the phase space of spinor condensates [MS07], pointing to a potential interest of these ideas in many-body physics, for example to witness different phases of matter where standard order parameters fail. Existing connections between entanglement classes and symmetry could further be useful in this direction [Mar11].

Chapter 6

Analysis of Nonlocal Properties for Symmetric States

This chapter, as a continuation of the previous chapter, also aims to explore the similarities and differences between multipartite nonlocality and multipartite entanglement by using the tools developed in earlier chapters.

The focus of this chapter will be monogamy, a useful property for quantum cryptography.

The chapter starts by showing the trends of violations of P^n for W and GHZ states as n gets large. In terms of monogamy and other applications of nonlocal features (for example communication complexity gains, cf. Section 7.1), the higher the value of the violation the better. It is interesting to know how violation scales with n. Then the monogamy of entanglement and correlations will be defined. Although the inequalities considered so far are not monogamous, a new inequality L for Dicke states, based on a recent result on the W state, will be presented, and shown to be monogamous in the limit of large n. The notion of genuine nonlocality will also be discussed, although neither P^n, Q_d^n nor the new inequality L can detect genuine nonlocality.

6.1 Large n Results for $|W_n\rangle$ and $|GHZ_n\rangle$

While the use of SDP allows us to study the nonlocality of symmetric states with a few parties, the computational resources required to run the SDP program increase exponentially with the number of parties, which makes it impractical to obtain results for states with more than 4 parties. Luckily, for two commonly studied symmetric states, the W states

$$|W_n\rangle = |S(n,1)\rangle = \frac{1}{\sqrt{n}}\left(\sum_{perm} |0\underbrace{1\ldots1}_{n-1}\rangle\right), \tag{6.1}$$

and the GHZ states

$$|GHZ_n\rangle = \frac{1}{\sqrt{2}}(|0\ldots0\rangle + |1\ldots1\rangle), \tag{6.2}$$

it is possible to calculate analytically the violation of P^n if the measurement bases are those prescribed in Chapter 4. This allows us to give bounds on the maximum violation possible and see trends. We will use a combination of this and numerics to approximate the best violation.

For the W state, using the bases $\{|+\rangle, |-\rangle\}, \{|0\rangle, |1\rangle\}$ as settings 0 and 1 in P^n, we get the violation

$$v_w(n) = \frac{n-2}{2^n}. \tag{6.3}$$

This algebraic violation, while works for all $|W_n\rangle$, is not the optimum violation. By optimizing over the four Euler angles in the two bases, we obtained close to optimal numerical violations of P^n (s in Fig. 6.1) and Q^n_{n-1} (t in Fig. 6.1) for W states. It can be seen from the plot that the violations of P^n is close to the upper bound derived from the geometrical measure of entanglement, $\frac{1}{2^{E_g(|W_n\rangle)}}$.

For GHZ states, we can follow the procedure given in Chapter 4 to find the bases. Note that the MPs of GHZ states with an even number of parties and an odd number of parties are different. For example, $|+\rangle$ is an MP of $|GHZ_n\rangle$ when n is odd, but not when n is even. Nevertheless, the MPs in both cases are all equally distributed along the equator of the Bloch sphere, allowing us to have a single expression for the bases as a function of n. The basis 1, which consists an MP and its antipodal point, is $\{\frac{1}{\sqrt{2}}(|0\rangle - e^{-i\frac{\pi}{n}}|1\rangle), \frac{1}{\sqrt{2}}(|1\rangle + e^{i\frac{\pi}{n}}|0\rangle)\}$, and the basis 0 is $\{\frac{1}{\sqrt{2}}(|0\rangle - e^{i\frac{(2n-1)\pi}{n(n-1)}}|1\rangle), \frac{1}{\sqrt{2}}(|1\rangle + e^{-i\frac{(2n-1)\pi}{n(n-1)}}|0\rangle)\}$. Calculating the violation as a function of n (which is just the probability $P(0\ldots0|0\ldots0)$), we have (the n line in Fig. 6.1)

$$v_g(n) = \frac{1}{2^n}\left(1 + \cos\frac{(2n-1)\pi}{n-1}\right). \tag{6.4}$$

This violation agrees with the best found by numerics.

From Fig. 6.1, we can see that as n increases, the violations are always well below $\frac{1}{2^{E_g}}$, which follows the trend we noticed in the earlier SDP examples. We also numerically optimized the value of Q^n_{n-1}, which is always negative for GHZ states. This is in stark contrast to the situation for W states, where the violation of Q^n_{n-1} stays slightly below the violation of P^n. One interpretation of this phenomenon is that Q^n_d is closely related to the degeneracy of the state, and can be used as a 'witness' of degeneracy for these states.

One obvious statement we can get from this with relation to entanglement was that

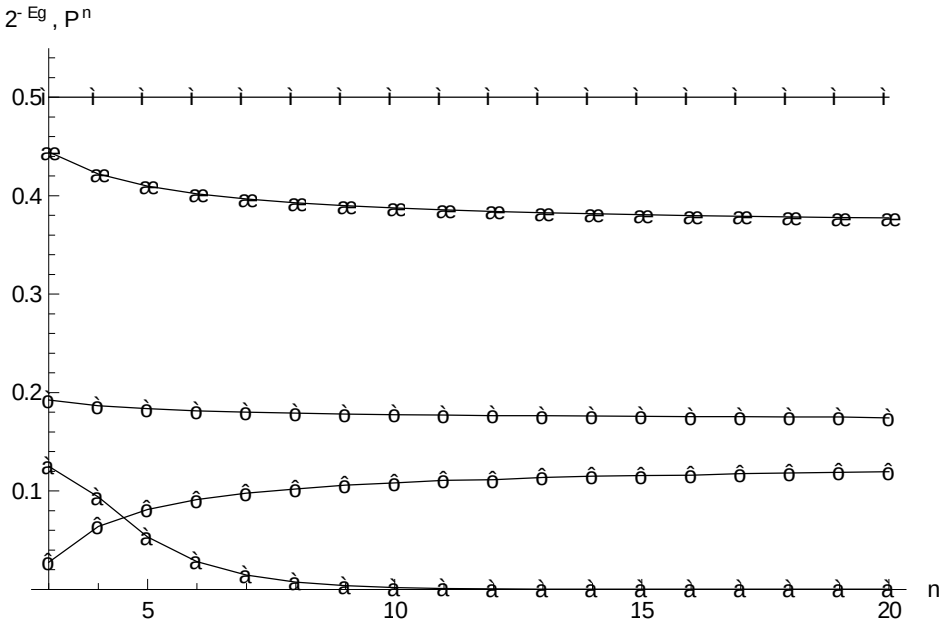


Figure 6.1: Violations of P^n by the state $jGHZ_n i$ (n), with the numerical violations of P^n (s) and Q^n_{n-1} (t) of $jW_n i$ as a function of n (number of parties), comparing to $\frac{1}{2^{Eg(jW_n i)}}$ (l) and $\frac{1}{2^{Eg(jGHZ_n i)}}$ (u).

the higher the entanglement is, the lower any possible violation of P^n and Q^n_d can be. At first this seems counterintuitive, but really it seems to stem from the simple fact that there is only one positive term - we later introduced larger inequalities with more positive terms based on the HCSV inequality [HCSV11], where the violation reaches its algebraic limit for all Dicke states in the high n limit. We looked at how the violation of inequalities scale with n for GHZ and W states. We see that W states fair much better for our inequalities, in contrast to the typical Mermin like inequalities where GHZ fairs better. We also look at the trends of the inequality violation with entanglement and see that this can be different. For W states and the $j000+i$ state the violation increases with entanglement so that it gets closer to the upper bound ($\frac{1}{2^{E_g}}$), where as for GHZ states it goes down for higher n.

6.2 Monogamy of Entanglement and Monogamy of Correlations

From earlier chapters, we have seen that it is possible to define similar concepts for nonlocality and entanglement to study similar properties in each context. For example, to study how different types can arise in entanglement, SLOCC classification is used, while in nonlocality the LUP classification can be used. There is another property, defined for both contexts, that highlights yet another interesting aspect of multiparty nonlocality and multiparty entanglement, that is the concept of monogamy [CKW00] [Ton09] (for a review see [See10]).

As its name suggests, monogamy measures the exclusiveness of entanglement or correlations, that is, how well they can be shared. For example if two parties share a maximally entangled state or a maximally correlated Popescu-Rohrlich (PR) box [PR94], the entangled systems or PR box cannot be entangled or correlated to anything else. In recent years it has been recognized as a key ingredient to the usefulness of states for example in security and device independent security scenarios [BKP06, PAM+10, CK11, AGCA12]. The idea being that if the correlations cannot be shared, that means that the eavesdropper is uncorrelated with the honest parties, so the information they share will not be leaked to the eavesdropper.

Monogamy of entanglement is a property of a particular quantum state. It measures the intra-subgroup entanglement tradeoff with respect to a suitably chosen entanglement measure. The most famous such measure is the tangle τ introduced in [CKW00], which measures the entanglement across a bipartition. The CKW inequality, proposed in [CKW00] as a conjecture and proved recently in [OV06], states that for all pure entangled states, the sum of all bipartite tangles between one party A and n parties $\{B_1, \ldots, B_n\}$ is less than or equal to the tangle between A and all B_i considered as a whole:

$$\tau(\tau_{AB_1}) + \tau(\tau_{AB_2}) + \ldots + \tau(\tau_{AB_n}) \leq \tau(\tau_{A(B_1 \ldots B_n)}). \tag{6.5}$$

Although it is known that symmetric states like the W state can saturate this inequality, not all states which saturate this inequality are symmetric.

The monogamy of 3-qubit symmetric states have been studied recently [SUDR12], using a different measure of quantum correlations, called the quantum deficiency (related to quantum discord [OZ01]). It was shown that SLOCC equivalent states do not necessarily have the same monogamy relation with respect to this measure. Here we focus on correlations of the measurement results directly (which we call simply monogamy of correlations).

Monogamy of correlations is normally defined in the context of correlations aris-

ing from probability distributions, without explicitly referring to quantum states and measurements. Intuitively, monogamy says that strong correlations cannot be shared. In a strict sense, we say an n-partite distribution, $P(a_1,\ldots,a_n|A_1,\ldots,A_n)$, is monogamous [BLM+05] [AGCA12], if the only nonsignaling extension to $n+1$ parties $P(a_1,\ldots,a_n,a_{n+1}|A_1,\ldots,A_n,A_{n+1})$ is the trivial one, i.e. such that

$$\begin{aligned} &P(a_1,\ldots,a_n,a_{n+1}|A_1,\ldots,A_n,A_{n+1}) \\ &= P(a_1,\ldots,a_n|A_1,\ldots,A_n)\,P(a_{n+1}|A_{n+1}). \end{aligned} \tag{6.6}$$

For all possible measurement settings A_k and A_k^0 for party k, the nonsignaling condition can be stated as

$$\begin{aligned} &P(a_1,\ldots,a_{k-1},a_{k+1}\ldots,a_n|A_1,\ldots,A_{k-1},A_{k+1},\ldots,A_n) \\ &= \sum_{a_k} P(a_1,\ldots,a_k,\ldots,a_n|A_1,\ldots,A_k,\ldots,A_n) \\ &= \sum_{a_k} P(a_1,\ldots,a_k,\ldots,a_n|A_1,\ldots,A_k^0,\ldots,A_n). \end{aligned} \tag{6.7}$$

That is, when tracing out one system, k, to get the marginal distributions, it does not matter which measurement setting A_k is used.

This strict sense of monogamy is guaranteed if an inequality reaches its algebraic maximum [BKP06]. Indeed, this fact is used to show monogamy for several states via several inequalities including GHZ states [BKP06] [AGCA12] [Ton09]. However, the inequalities P^n and Q_d^n here cannot show strict monogamy in this way, simply because no quantum state can ever achieve the algebraic bound, as the bound is given by the entanglement. In the following subsection we will develop another set of inequalities for which this idea does work.

Even if not demanding strict monogamy of correlations, it is possible to bound how well correlations can be shared. In [PB09], a bound is presented covering general nonsignaling theories by demanding tradeoffs of correlations in a multipartite setting, analogous to the monogamy of multipartite entanglement. To apply these results to our inequality, we will follow the prescription given in [PB09]. First we rewrite our

inequality to make all terms positive:

$$\begin{aligned} P^n = {} & P(0\ldots0|0\ldots0) \\ & - \sum_{a_1,\ldots,a_n \neq 0\ldots0} P(a_1,\ldots,a_n|0\ldots1)) \\ & \vdots \\ & - \sum_{a_1,\ldots,a_n \neq 0\ldots0} P(a_1,\ldots,a_n|1\ldots0)) \\ & - \sum_{a_1,\ldots,a_n \neq 1\ldots1} P(a_1,\ldots,a_n|1\ldots1)). \end{aligned} \tag{6.8}$$

By keeping all the probabilities on the left hand side and moving everything else to the right hand side, we define the inequality

$$\begin{aligned} P^{n'} = {} & P(0\ldots0|0\ldots0) \\ & + \sum_{a_1,\ldots,a_n \neq 0\ldots0} P(a_1,\ldots,a_n|0\ldots1) \\ & \vdots \\ & + \sum_{a_1,\ldots,a_n \neq 0\ldots0} P(a_1,\ldots,a_n|1\ldots0) \\ & + \sum_{a_1,\ldots,a_n \neq 1\ldots1} P(a_1,\ldots,a_n|1\ldots1) \\ & \leq n+1. \end{aligned} \tag{6.9}$$

Now we can partition the parties into two groups: group A with k parties and group B with $n-k$ parties. Consider a single group A which is possibly correlated with multiple identical B^i. The multiparty monogamy relation of [PB09] tells us that for any nonsignalling probability distribution for $n > 2$

$$\sum_{i=1}^{n^{k+2}} P^{n'}(A, B^i) \leq (n-k+2)(n+1), \tag{6.10}$$

where i runs over the possible combinations of measurement settings of $n-k$ parties that make up each B^i.

For Q^n_d, we can treat the $d-1$ extra probabilities as marginals of probabilities involving n parties:

$$P(\underbrace{1\ldots1}_{n-d+1}|\underbrace{1\ldots1}_{n-d+1}) = \sum_{b_1,\ldots,b_{d-1}} P(\underbrace{1\ldots1}_{n-d+1}\underbrace{b_1\ldots b_{d-1}}_{d-1}|\underbrace{1\ldots1}_{n}), \tag{6.11}$$

which leads to the inequality for $Q_d^{n'}$:

$$Q_d^{n'} = P^{n'} + \sum_{a_1,\ldots,a_{n-1}=1\ldots 1} P(a_1,\ldots,a_{n-1},b_1|\underbrace{1,\ldots,1}_{n}) \\ \vdots \\ + \sum_{a_1,\ldots,a_{n-d+1}=1\ldots 1} P(a_1,\ldots,a_{n-d+1},b_1,\ldots,b_{d-1}|\underbrace{1,\ldots,1}_{n}) \\ \leq n+d. \tag{6.12}$$

Because the expression for $Q_d^{n'}$ does not increase the number of settings for B, we have the monogamy inequality for $Q_d^{n'}$ similar to (6.10):

$$\sum_{i=1}^{n^{k+2}} Q_d^{n'}(A,B^i) \leq (n-k+2)(n+d). \tag{6.13}$$

We then looked at what can be said about the monogamy of the correlations exposed by our inequalities and chosen measurement settings. First, we see that P^n and Q_d^n are not suited to showing strict monogamy (that is, we cannot say violation at the level achieved by quantum states implies no correlations are shared with another party), since, by the fact that entanglement bounds the violation, any quantum violation cannot reach the algebraic limit. This may indicate that these inequalities are not so useful for device independent security for example, although bounds on correlation sharing less than these strict ones may be of interest. To this end, using techniques from [PB09] we bound how much correlations can be shared with the inequalities. We then define new inequalities based on the HCSV inequality, where we see that all Dicke states are strictly monogamous in the limit of high n, as has been seen before for W states [HCSV11]. In this sense the extreme nonlocality of GHZ and stabilizer states seems to be replicated by Dicke states in the large n limit. It remains open how general this is for all symmetric states.

6.3 Monogamy and Genuine Nonlocality of Dicke States

We now introduce a set of inequalities which can show strict monogamy of Dicke states in the high n limit. These are based on recent work by Heaney, Cabello, Santos and Vedral [HCSV11] where they show that for the W state, it is possible to construct nonlocality tests and inequalities that are maximal in some sense, i.e. the violation of the inequality goes to the algebraic maximum in the $n \to \infty$ limit, thus mimicking perfect correlations of stabilizer states and the Mermin inequality (cf. Chapter 2, Section 2.4.4). The inequality introduced in [HCSV11] by Heaney, Cabello, Santos and Vedral (hereafter referred to as the HCSV inequality), has the property that the larger n is, the higher the violation becomes. Although the original HCSV inequality only works for W states, it can be extended as follows to cover all Dicke states.

Following and extending the reasoning in [HCSV11] for W state, if all n parties measure in the σ_z basis on a Dicke state $|S(n,k)\rangle$, $n - k$ of them will get result 0 and the other k will get result 1 with certainty (though it is impossible to know who gets what). Now imagine that when $n - k - 1$ parties get 0 and the other $k - 1$ parties get 1, the remaining two decide instead to measure σ_x. In this case they will always get the same result. Since under LHV the results of one party should not depend on other parties' settings, this means that should any two chose to measure in σ_x, they would get the same result. If these results are given by an LHV distribution, this would mean that if all parties were to measure in σ_x in the beginning, they should all get the same result. Since everything above occurs with certainty, we should always see, under LHV, that if all parties measure σ_x they get the same result. However, simple calculation shows that this is not the case for all Dicke states.

The associated Bell inequality is

$$\begin{aligned} L = & \sum_{\text{perms}} P(\sigma(\underbrace{0\ldots0}_{n-k}\underbrace{1\ldots1}_{k})|0\ldots0) \\ & - \sum_{\text{perms}} P(\sigma(\underbrace{0\ldots0}_{n-k-1}\underbrace{1\ldots1}_{k-1}01)|\sigma(\underbrace{0\ldots0}_{n-2}11)) \\ & - P(0\ldots0|1\ldots1) - P(1\ldots1|1\ldots1) \geq 0, \end{aligned} \tag{6.14}$$

where the permutations in the second and third lines are over parties fixing the relationship between measurement settings and results, as with P^n. To see that this cannot be violated under LHV it is sufficient to see that it cannot be violated for any deterministic strategy (i.e. taking marginal probabilities to be zero or one) [WW01], since all LHV distributions can be considered as probabilistic mixtures of deterministic ones. It is not difficult to see that taking any one of the $P(\sigma(\underbrace{0\ldots0}_{n-k}\underbrace{1\ldots1}_{k})|0\ldots0)$ to be one cannot be compatible with keeping all the negative terms zero. Since these are the only possible

positive terms, and at most only one can be equal to one, for all deterministic local strategies the expression is non-positive and a violation is incompatible with LHV. For a Dicke state $|S(n,k)\rangle$, L is violated by $1 - \frac{\binom{n}{k}}{2^{n-1}}$. As for the W state considered in [HCSV11], this achieves the algebraic maximum in the limit of large n, imitating perfect correlations of GHZ and other stabilizer states. This also implies strict monogamy for the limit in n.

We plot the violation of L for $|S(n,\frac{n}{2})\rangle$ and $|W_n\rangle$ in Fig. 6.2. We see that the W state reaches one more quickly, in keeping with its lower entanglement.

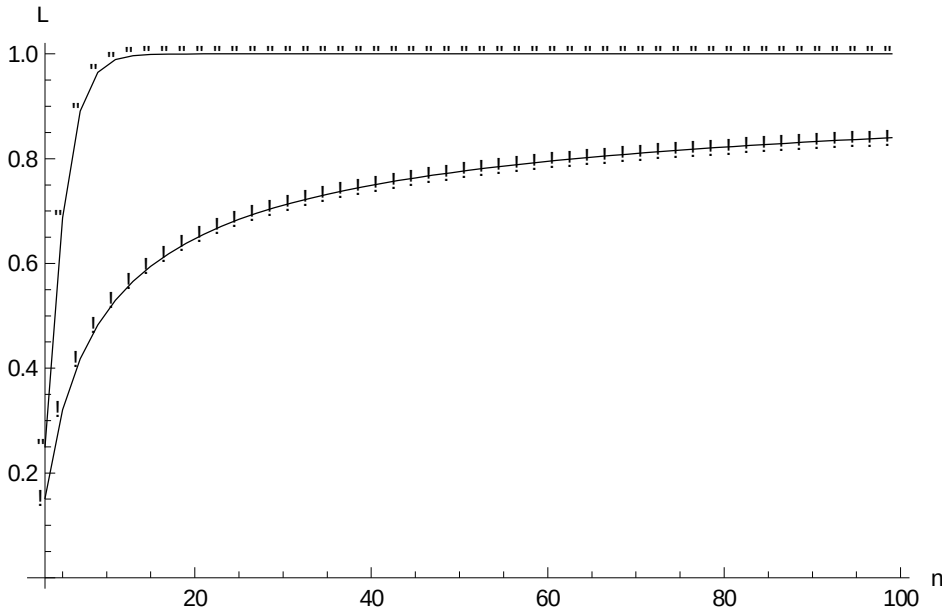


Figure 6.2: A comparison of the violation of L (6.14) for the states $|S(n,\frac{n}{2})\rangle$ (n) and $|W_n\rangle$ (l) as a function of n (the number of parties).

One can also ask what other nonlocal properties can be inspected by inequalities P^n and Q_d^n. Another property of multiparty correlations which is of interest, is whether it can be said to be "genuine" or not - that is, whether the correlations at hand could be achieved by grouping the n into subgroups or not. If not, we would say the correlations are genuinely n party. The Svetlichny type inequalities [Sve87] endeavor to identify this property - they should only be violated by genuinely n party correlated states. Unfortunately it is not to hard to see that all the inequalities we use in this work do not have this property - it is possible to group parties together such that local states with respect to the new groupings can violate the inequalities. This can be easily seen by grouping the first $n-2$ parties and construct an LHV model by only using deterministic probabilities (probabilities equal to 0 or 1). The grouping makes it possible to set all negative terms to 0, and (one of) the positive term to 1. A stronger statement can be made by only grouping the first two terms - so that the weakest grouping still allows nonlocal correlations to violate all our inequalities. This is shown explicitly for L below.

Theorem 16. The inequality L can not detect genuine nonlocality.

Proof. To show that L cannot detect genuine nonlocality, we will group the first two parties and show that L = 1 under partially nonlocal LHV (where the first two parties are considered as one). Mathematically, an LHV model means that we can write

$$P(a_1 \dots a_n | A_1 \dots A_n) = \int \rho(\lambda) \prod_{1 \le i \le n} P_i(a_i | A_i, \lambda)\, d\lambda, \tag{6.15}$$

where subscripts denote the parties.

Meanwhile, a partially nonlocal LHV means that we allow a subset of parties to be grouped together as a single (possibly nonlocal) party. In this proof, it means that

$$P(a_1 \dots a_n | A_1 \dots A_n) = \int \rho(\lambda) P_{12}(a_1 a_2 | A_1 A_2, \lambda) \prod_{3 \le i \le n} P_i(a_i | A_i, \lambda)\, d\lambda. \tag{6.16}$$

Below we give an explicit LHV model by setting all probabilities in L to equal to either 0 or 1. This implies only one term in the sum $\sum P(\lambda(\underbrace{0 \dots 0}_{n-k} \underbrace{1 \dots 1}_{k}) | 0 \dots 0)$ equals to 1, all other terms will be 0. Let us suppose, without loss of generality,

$$P(\underbrace{0 \dots 0}_{n-k} \underbrace{1 \dots 1}_{k} | 0 \dots 0) = 1. \tag{6.17}$$

This implies

$$P_{12}(00|00) = 1 \tag{6.18}$$

$$P_3(0|0) = 1, \dots, P_{n-k}(0|0) = 1 \tag{6.19}$$

$$P_{n-k+1}(1|0) = 1, \dots, P_n(1|0) = 1, \tag{6.20}$$

from which we can deduce

$$P_{12}(01|00) = P_{12}(10|00) = P_{12}(11|00) = 0 \tag{6.21}$$

$$P_3(1|0) = 0, \dots, P_{n-k}(1|0) = 0 \tag{6.22}$$

$$P_{n-k+1}(0|0) = 0, \dots, P_n(0|0) = 0. \tag{6.23}$$

For the terms $\sum P(\lambda(\underbrace{0 \dots 0}_{n-k-1} \underbrace{1 \dots 1}_{k-1} 01) | \lambda(\underbrace{0 \dots 0}_{n-2} 11))$, we will try to set all of them to 0, using (6.21) to (6.23) with some extra probability assignments, without causing inconsistencies.

To see how we can set all terms to 0, first we divide the terms in the sum into three different cases (a, b are both bits, $\bar{a}, \bar{b}$ denote their logical flip):

1. $P(ab\,|\,(0\ldots01\ldots01)|00\,|\,(0\ldots011))$.

   In this case, if a and b are not both 0,then by (6.21), the probability is 0. Otherwise, we can set $P_i(0|1) = 0$, where $i \in 1,2$.

2. $P(ab\,|\,(0\ldots01\ldots\bar{b})|01\,|\,(0\ldots001))$, $P(ab\,|\,(0\ldots01\ldots\bar{a})|10\,|\,(0\ldots001))$.

   In this case, if $a = b = 1$, then there exists $P_i(0|1)$ where $i \in 1,2$. Thus we can have $P_i(0|1) = 0$ and $P_{12}(11|01) = 1, P_{12}(11|10) = 1$, without causing any inconsistency with the previous case. The latter two assignments also imply that if a and b are not both 1, then $P_{12}(ab|01) = P_{12}(ab|10) = 0$.

3. $P(a\bar{a}\,|\,(0\ldots01\ldots01)|11\,|\,(0\ldots0))$.

   In this case, the probability is always 0. This can be deduced from the pigeonhole principle: there are $n - k - 1$ zero outcomes when parties 3 to n all measure in the 0 basis, so at least one party from $n - k + 1$ to n will get outcome 0 when measuring in the 0 basis. By (6.23) the probability is 0.

In the last case, because the probability is always 0 regardless of the probability assignments of the first two parties, we can set $P_{12}(00|11) = 0$ and $P_{12}(11|11) = 0$ without causing any inconsistency. These assignments guarantee that the last two probabilities in L: $P(0\ldots0|1\ldots1)$ and $P(1\ldots1|1\ldots1)$, are 0.

Thus we can consistently assign probabilities such that all negative terms in L are 0 and the sum of all positive terms are 1, so $L = 1$, violating the inequality under partially nonlocal LHV. This shows that L cannot detect genuine nonlocality.

A similar argument can be made for P^n and Q^n_d. □

In conclusion it seems that one must make a balanced choice over which inequalities will be useful depending on circumstances. We have seen that P^n and Q^n_d are interesting in terms of separating classes of states, and indeed it is known to be true that all entangled pure states will show some violation P^n [YCZ+12]. However, their violation can never be high enough to make the strongest statements we would like about monogamy. They also do not say whether correlations are "genuine" or not (even L, with its many positive terms, does not show genuine nonlocality or be maximally violated for finite n). On the other hand inequalities based only on expectation values (which necessarily have many positive terms) can have maximal violation for any n, but they cannot see the nonlocality of all states - there are entangled states which do not violate any inequality based on expectation values, which do violate P^n [ZBLW02]. In a similar situation to the role of different entanglement measures in entanglement theory, it seems unlikely that any single inequality will be able to capture all the nonlocal properties we might be interested in.

Chapter 7

Applications

7.1 Application to Communication Complexity

The early interests in studying quantum computing mainly come from the computational aspects. From the 1982 paper of Feynman [Fey82] which pointed out the simulation of a quantum system by a classical computer as the Achilles' heel of classical computation, to the 1996 paper by Grover [Gro96] and the 1997 paper by Shor [Sho97] demonstrating possible big gaps between quantum and classical algorithms, the early excitement of quantum computation comes mainly from quantum algorithms.

But nine years before Feynman and over twenty years before Grover or Shor, an important result on quantum communication had already been proven [Hol73]. In this paper, Holevo proved what is now known as the Holevo bound, which shows the number of qubits needed in transmitting a classical message from Alice to Bob is the same as the number of classical bits. This may seem discouraging for any further study of quantum communication, but it turns out that if one combines the ancient (in the quantum information world) notion of nonlocality with communication complexity [KN97, AB09b], which uses a different way to model the cost of communication as the one used by Holevo, there can be surprising results [BCMdW10].

Communication complexity has been well-studied in the classical case [KN97, AB09b]. The basic model is given by Yao [Yao79], who also made great contributions to quantum computing by initiating the research on quantum circuit complexity and quantum communication complexity in the same paper [Yao93]. The model involves two parties, Alice and Bob, each having an n-bit string, x_A and x_B, respectively. Their goal is to compute some Boolean function f of x_A and x_B, $f(x_A, x_B) \in \{0, 1\}$, using as little communication (as measured in total number of bits communicated) as possible. For this they may use a fixed protocol, which is divided into several rounds, and in each round they one party sends a bit to the other party. The protocol terminates when at least one party has the correct value of $f(x_A, x_B)$. One extreme situation arises when f is independent of x_A or

x_B. In this case no communication is required for one party to compute f, so it is not very interesting. The other extreme situation involves Alice (or Bob) sending x_a (or x_B) directly to the other party. The communication complexity in this worst case scenario is n bits. So for any Boolean function f, the communication complexity is always between 0 and n bits.

Just like in computability, allowing randomness in communication sometimes reduces its complexity. To introduce randomness in the model above, suppose Alice and Bob each possesses, in addition to x_A and x_B, a random string r_A and r_B. So and they are allowed to only give the correct value of f with some positive probability. If r_A and r_B are drawn from two independent random variables, the model is called private coin model. If r_A and r_B are drawn from a shared random variable (i.e. $r_A = r_B$), the model is called a public coin model.

To show the effect of having randomness in communication complexity, consider a few well-studied functions f:

- Equality. $f_{EQ}(x_A, x_B) = 1$ iff $x_A = x_B$. It is not hard to see that in the deterministic case, f_{EQ} needs n bits of communication: either Alice or Bob must send all her/his bits. But using private coins, the communication complexity can be reduced to $\Theta(\log_2 n)$ [KN97, Example 3.9]. With a public coin, the communication complexity is constant [KN97, Example 3.13].

- Disjointness or Intersection. $f_{DJ}(x_A, x_B) = 1$ if x_A and x_B agree on at least one bit, i.e. $x_A \cap x_B \neq \emptyset$. In the deterministic case, the lower bound is $\Omega(n)$ [KN97, Example 1.23, Exercise 1.26]. But this time randomness does not help much, the lower bound is still $\Omega(n)$ [KN97, Example 3.22] [Raz92, KS92].

- Inner Product. $f_{IP}(x_A, x_B)$ equals the XOR of the bitwise AND of x_A and x_B: $f_{IP} = \bigoplus_{i=1}^{n} x_A^i \wedge x_B^i$. Just like f_{DJ}, f_{IP} needs n bits of communication in both the deterministic and the randomized case [BCMdW10].

Using entanglement and nonlocality, the same function may enjoy a speedup.

The equality function already has O(1) complexity in the public coin model, so there is no interesting speedup in the quantum case.

The disjointness function, on the other hand, has a quadratic speedup [BCW98] to $O(\sqrt{n})$ thanks to the Grover search algorithm [Gro96].

The most interesting result comes from the inner product function. Although using quantum resources still can not change the lower bound, if a PR box (2.28) is used, the communication complexity of all Boolean functions become trivial [vD99, BCMdW10]. As a result, if one believes communication complexity should not be trivial, then PR boxes can not exist in nature. Later results by Brassard et al. [BBL+06] showed that even with noisy PR boxes which work with probability $\approx \frac{3+\sqrt{6}}{6} \approx 90.8\%$, the probabilistic communication complexity of Boolean functions is still trivial.

The model and results above all come from bipartite situations. In classical communication complexity, the generalization of the bipartite model to multipartite is not unique [KN97, Chapter 6]. The most obvious generalization to n party is by allowing f to have n arguments $x_1, \ldots, x_n$ while each party i only knows x_i. This model is weak in the sense that most lower bounds in this model can be derived using results in the bipartite model. For example, the n party equality function, defined as $f^n_{EQ}(x_1, \ldots, x_n) = 1$ iff all the x_i are equal, has complexity $\Omega(n)$ in this model. A more interesting model for n party communication complexity is the number on the forehead model, where each player i knows the input of all other players except x_i. In this model the complexity of f^n_{EQ} is 2: the player i broadcasts a bit indicating whether all input strings except x_i are equal, then a different player j broadcasts one bit indicating whether x_i is equal to any x_k where $k \neq j$.

In an attempt to show the advantages of multiparty nonlocality in multiparty communication complexity, Brukner et al. proposed a model to turn any Bell inequality of correlation functions into a communication complexity problem [idZB02]. The model only considers probabilistic communication complexity, where all players compute the right value of f with nonzero probability. But it is also known [ZBLW02] that there are states which do not violate any Bell inequality of correlation functions, and these states are symmetric hence they violate P^n. This suggests that something more is needed to show a communication complexity advantage from the violation of P^n.

Because almost all symmetric states not only violates P^n, but also satisfy the extended Hardy paradox (4.37) to (4.40), it is worthwhile to look at the structure of these (n + 2) probabilities.

By treating the measurement settings as inputs and the outcomes as outputs, (4.37) to (4.40) give (n + 2) conditions on possible input/ output pairs. While the first condition is straightforward: given the input 0...0, the output 0...0 sometimes occurs, the next (n + 1) conditions only specify forbidden input/ output pairs while saying nothing about the allowed ones. As a result of this ambiguity, the outputs in the extended Hardy paradox generally is not a function of the inputs. Because a function must be at least injective, i.e. mapping an input to a definitive output. Input/ output pairs in this case form something more general than a function: a relation.

Formally, a relation R is a subset of the Cartesian product of the sets of inputs with the set of possible outputs: $R \subseteq I_1 \times I_2 \times \ldots \times I_n \times O$. In the Hardy paradox relation the inputs are bits: $I_j \in \{0, 1\}$, and the output is a n-bit string $O \in \{0, 1\}^n$. Functions are special cases of relations when for every combination of inputs $x_1 \in I_1$, $x_2 \in I_2$ etc., there is a unique $z \in O$. Take the Hardy paradox example, the (n + 1)-tuple (0, 0, 0, 000) is in the Hardy relation R_H because of (4.37), while (0, 0, 1, 000) is not in R_H, because of (4.38). Depending on the full distribution, (0, 0, 1, 001) may or may not be in R_H.

Classically, there are some interesting results on communication complexity of rela-

tions of two parties, where each party possesses an n-bit string x_A and x_B and they are trying to find some z such that (x_A, x_B, z) is in the relation. In this case, randomization may reduce the lower bound of the communication complexity of some relations. The Example 5.5 given in [KN97] shows that while the disjointness function requires $\Omega(n)$ in both deterministic and randomized settings, and $O(\sqrt{n})$ with quantum resources, changing the function into a relation which estimates the size of the intersection (which is the relation variant of a function that outputs the size of the intersection, which must be at least as hard as the original disjointness function) is $O(\log_2 n)$ in the private coin model and $O(1)$ in the public coin model.

Although a rigorous result needs to be proven showing that the Hardy paradox offers an advantage in terms of the communication complexity of the relation R_H, intuition shows that the advantage is likely to be exponentially small. This is because the number of $(n+1)$-tuples in the relation is exponential in n, while the Hardy paradox only excludes $(n+1)$ of them. This is also consistent with the fact that as n grows, it is exponentially harder to observe the event that when all players measure in the setting 0, they all obtain the outcome 0 (4.37).

7.2 Application to Bayesian Games

In communication complexity, all participating parties try to cooperate to achieve the same goal: compute the value of the function correctly. However, in the real world, people have very legitimate reasons to try to hide information from someone else, or try to sabotage each other instead of just cooperating. In economics, where human behaviors are often modeled as games [FT91, OR94, Mye97, Ras07, vNM07], such situations correspond to Bayesian games, or games with incomplete information, pioneered by Harsanyi in the late 60s [Har67, Har68a, Har68b].

As an example, consider the following game with two players. Alice and Bob are friends. They kind of like each other and want to become romantically involved, but each person is not sure whether the other person thinks the same way. One day they decide to watch a movie together, maybe try to ask the other person if he/she feels the same. The cinema is currently showing two movies: the action-packed Expendables 2 and the classic musical Les Miserables. Alice strongly prefers to watch Les Miserables, while Bob prefers the Expendables 2. While they have different preferences, they prefer watching the same movie (albeit only enjoyed by one) than each go to his/her favorite movie (they can not ask the other person if he/she wants to be her/his girl/boyfriend if they go to different movies). Having an important question to ask only makes the movie watching experience more stressful. If Alice wants to be Bob's girlfriend but Bob refuses to be her boyfriend, this will ruin the movie for Alice, especially if the movie is Les Miserables. Same for Bob, especially if the movie is the Expendables 2. If they both said no to become romantically involved, it makes the situation slightly different: Alice feels less strongly about Les Miserables being ruined, while Bob feels more strongly about agreeing to watch this girly musical where a tone-deaf Australian is pretending to be French. Similarly, if they both said no, Bob feels the Expendables 2 is ruined less while Alice is disgusted by a 60 year olds who took so much steroid and botox that his facial expression remains constant throughout the movie. Since they do not know how the other person will answer the question, what is their average payoff in this game (which may have happened already somewhere in the world)?

To model this game, it is necessary to introduce some game theory vocabulary. In game theory parlance, a Bayesian game usually consists of [Osb03, BL12]

- A set of players.
- A set of states (also called the states of nature).
- For each player, a set of possible types.
- For each player, a set of possible actions.
- For each player, a deterministic signal function assigning a type to each state.

- For each player, a payoff function assigning a numerical value according to the state and the actions taken by the player.

For the game defined above, which is a Bayesian version of the Battle of the Sexes (BoS) game [Osb03, Cl08], the players are obviously Alice and Bob, with the states of nature being all four possibilities that if they want to become boy/girlfriend or not (yes/yes, yes/no, no/yes, no/no). The type of each player is his/her private feeling about the other person, the actions for each player are watching one of the two movies. The signals simply map the states of nature to the type of each player. The payoffs are given by the following matrix [Cl08]:

$$\left(\begin{array}{cc|cc} (3,1) & (0,0) & (-3,-1) & (0,0) \\ (0,0) & (1,3) & (0,0) & (-1,-3) \\ \hline (-3,-1) & (0,0) & (-1,-3) & (0,0) \\ (0,0) & (-1,-3) & (0,0) & (-3,-1) \end{array}\right). \tag{7.1}$$

The four quadratures of the matrix represent the four states of nature: NW= yes/yes, NE= yes/no, SW= no/yes, SE= no/no. The numbers in the parentheses are payoffs to Alice and Bob, respectively. Within each of the four smaller matrices, the four entries define which movie Alice and Bob goes to see (M= Les Miserables, E= The Expendables 2): NW= M/M, NE= M/E, SW= E/M, SE= E/E. To facilitate later studies (and make the connection with nonlocality more apparent), a labeling can be applied to the matrix. For the big matrix, each quadrature (or small matrix) can be labeled using two bits: one bit represents the state of nature of Alice, one bit for the state of nature of Bob. Using this labeling, the four small matrices are: NW= 00, NE= 01, SW= 10, SE= 11. Similarly, the four entries in each small matrix can be labeled using the same procedure. This way, each event for Alice and Bob, with its pair of payoffs, is assigned a label. For example, the event that Alice and Bob both agree to become romantically involved and watch the Expendables 2 together is assigned the label (11|00), where the two bits after the vertical line represent the state of nature (small matrix) and the two bits before the vertical line denote the event. For the event (11|00), Alice receive payoff 1 and Bob receive 3 (the entry (1,3) in the small matrix at the upper left corner of the big matrix). The probability that event (11|00) occurs is denoted by P(11|00).

Using the labeling above, the average payoff to Alice and Bob can be written as a function of probabilities of all the possible events, assuming that the four states of nature

occur with equal probability ($\frac{1}{4}$) [CI08]:

$$\begin{aligned}\square_A = &\frac{3}{4}(P(00|00) \square P(00|01) \square P(00|10) \square P(11|11)) \\ &+ \frac{1}{4}(P(11|00) \square P(11|01) \square P(11|10) \square P(00|11)), \qquad (7.2)\\ \square_B = &\frac{1}{4}(P(00|00) \square P(00|01) \square P(00|10) \square P(11|11)) \\ &+ \frac{3}{4}(P(11|00) \square P(11|01) \square P(11|10) \square P(00|11)). \qquad (7.3)\end{aligned}$$

Now it is obvious that the average payoffs for Alice and Bob are just the CH inequality (2.56) and one of its variants. To show that a quantum strategy has an advantage, there needs to be something akin to a local hidden variable in this Bayesian game setting. Curiously, there is still one thing missing from the discussion above: the function of the signal. Traditionally, a signal is supposed to function like an advise, originated from an outside advisor. It is very important that the advisor does not know the type of each player, otherwise the advisor can allow signaling among the players [Bra10, BL12]. Upon receiving the signal, each player can choose independently his/ her strategy, using the advice and the type, which is known only to the player. Enforcing these two conditions can be seen as enforcing a local hidden variable model in physics, and such a signal is called a classical signal [Bra10, BL12]. It can be shown that using a classical signal the average payoffs for Alice and Bob above are both 0, which is also a Bayesian equilibrium (no one can unilaterally do better than this average payoff). Not surprisingly, using a quantum signal (which comes from sharing an entangled state and making suitable measurements just as done in a nonlocality test), the average payoffs for Alice and Bob are both positive [CI08].

Although it is easy to tailor make games and payoff matrices which correspond to the n party Hardy paradox (4.37) to (4.40) [CI08], a deeper connection between Bayesian games and nonlocality exists [BL12]. Early results on quantum strategies of classical games usually focus on games with perfect information [Mey99, EWL99]. But to show quantum strategies have an advantage over classical strategies, assumptions have to be made, and these assumptions lead to controversy [BH01, vEP02]. As pointed out in [BL12], Bayesian games can be always modeled to show a quantum advantage, which means the payoffs need not be tailor made to a specific inequality, such as in the example above.

Chapter 8

Summary

8.1 New Results in This Thesis

The new results in this thesis can be roughly split up into two categories:

1. The results related to the nonlocality of all symmetric states.

2. The results related to the analysis of various nonlocal properties of symmetric states and their relationship to entanglement of symmetric states.

The first category mainly concerns Chapter 4, in which it is shown that almost all symmetric states satisfy the n-party Hardy paradox and violate the inequality P^n. A constructive procedure is given to determine the measurement bases to show the paradox and violation. It is also shown that although Dicke states do not satisfy the n-party Hardy paradox, they still violate P^n.

The second category mainly concerns Chapter 5 and 6. Degeneracy, as a feature of symmetric states and the Majorana representation, has known application in the entanglement theory of symmetric states. By exploiting degeneracy, new applications in nonlocality is shown, where degeneracy is related to the persistency of nonlocality into subsets. By using degeneracy as a common tool, device independent classifications of nonlocality is performed, with states sitting in different entanglement classes as well. Then it is shown that the geometric measure of entanglement bounds the violation of P^n, so it can not be monogamous in the strict sense (although it does exhibit some broad monogamy). Based on a recent work, a new inequality for Dicke states which is monogamous in the strict sense is shown. Although neither this new inequality nor the two original ones can detect genuine nonlocality.

8.2 Recent Progress on Related Topics

Since the publication of [WM12], there have been several new developments on related topics. The most interesting one is by Yu et al. [YCZ+12], which shows that all pure entangled states can violate P^n. P^n can be seen as a universal inequality which can be used to detect the nonlocality of all pure entangled states.

Another interesting development is on the persistency of nonlocality and entanglement into subsets [BV12]. In this work the notion of persistency is rigorously defined for entanglement and nonlocality, with analysis of different types of states having different properties. Although symmetric states do not feature prominently in this work, there are several results concerning symmetric states, such as the W state exhibits maximum persistency both in terms of entanglement and in terms of nonlocality, and a proposal to test the symmetry of a state in a device independent way.

8.3 Outlooks

Future works extending the results in this thesis can be divided into two directions: theoretical and experimental.

On the theoretical front, the two topics raised in Chatper 7 still need more research, especially on the communication complexity of relations. Another interesting question to ask is how to incorporate degeneracy (or persistency) into games. Will this represent some coalition?

Another possible theoretical direction is to use symmetric states to perform some contextuality test. As mentioned in Chapter 2, recent techniques using category theory or hyper graphs provide unified frameworks to work with both nonlocality and contextuality. It is interesting to see if symmetric states and the Majorana representation can be used as a tool in contextuality tests, especially when the stabilizer formalism is known to work both as a tool to show nonlocality and as a tool to show contextuality. Recent progress on the computational power of linear optics [AA11][1] may provide a surprising answer to this question.

Experimentally, recent results suggest that all the symmetric states used in this thesis can be generated [LLL+12]. It is interesting to see a violation of P^n for some states in an experiment. It is more interesting to see a violation of Q_d^n for a state with some degeneracy, to observe the persistency of nonlocality in a real experiment. As a first step towards this goal, Adel Sohbi in our group is currently studying the effect of noise on the violations of P^n and Q_d^n.

[1] Also by Terry Rudolph through private communication

Bibliography

[AA11] Scott Aaronson and Alex Arkhipov, The computational complexity of linear optics, Proceedings of the 43rd annual ACM symposium on Theory of computing (New York, NY, USA), STOC '11, ACM, 2011, pp. 333–342.

[AB09a] Janet Anders and Dan E. Browne, Computational Power of Correlations, Phys. Rev. Lett. 102 (2009), 050502.

[AB09b] Sanjeev Arora and Boaz Barak, Computational Complexity: A Modern Approach, Cambridge University Press, 2009.

[AB11] Samson Abramsky and Adam Brandenburger, The sheaf-theoretic structure of non-locality and contextuality, New Journal of Physics 13 (2011), no. 11, 113036.

[ABG+07] Antonio Acín, Nicolas Brunner, Nicolas Gisin, Serge Massar, Stefano Pironio, and Valerio Scarani, Device-Independent Security of Quantum Cryptography against Collective Attacks, Phys. Rev. Lett. 98 (2007), no. 23, 230501.

[ADR82] Alain Aspect, Jean Dalibard, and Gerard Roger, Experimental Test of Bell's Inequalities Using Time-Varying Analyzers, Phys. Rev. Lett. 49 (1982), no. 25, 1804–1807.

[AGCA12] Leandro Aolita, Rodrigo Gallego, Adan Cabello, and Antonio Acín, Fully Nonlocal, Monogamous, and Random Genuinely Multipartite Quantum Correlations, Phys. Rev. Lett. 108 (2012), 100401.

[AGR81] Alain Aspect, Philippe Grangier, and Gerard Roger, Experimental Tests of Realistic Local Theories via Bell's Theorem, Phys. Rev. Lett. 47 (1981), 460–463.

[AGR82] ———, Experimental Realization of Einstein-Podolsky-Rosen-Bohm Gedankenexperiment : A New Violation of Bell's Inequalities, Phys. Rev. Lett. 49 (1982), 91–94.

[AMM10] Martin Aulbach, Damian Markham, and Mio Murao, The maximally entangled symmetric state in terms of the geometric measure, New Journal of Physics 12 (2010), no. 7, 073025.

[AMM11] Martin Aulbach, Damian Markham, and Mio Murao, Geometric Entanglement of Symmetric States and the Majorana Representation, Theory of Quantum Computation, Communication, and Cryptography (Wim van Dam, Vivien Kendon, and Simone Severini, eds.), Lecture Notes in Computer Science, vol. 6519, Springer Berlin Heidelberg, 2011, pp. 141–158.

[Ard92] Mohammad Ardehali, Bell inequalities with a magnitude of violation that grows exponentially with the number of particles, Phys. Rev. A 46 (1992), 5375–5378.

[Aul11a] Martin Aulbach, Classification of Entanglement in Symmetric States, Ph.D. thesis, University of Leeds, July 2011.

[Aul11b] ———, Symmetric entanglement classes for n qubits, Arxiv preprint arXiv:1103.0271 (2011).

[BA57] David Bohm and Yakir Aharonov, Discussion of Experimental Proof for the Paradox of Einstein, Rosen, and Podolsky, Phys. Rev. 108 (1957), 1070–1076.

[Bar07] Jonathan Barrett, Information processing in generalized probabilistic theories, Phys. Rev. A 75 (2007), 032304.

[BBL+06] Gilles Brassard, Harry Buhrman, Noah Linden, Andre Allan Methot, Alain Tapp, and Falk Unger Limit on Nonlocality in Any World in Which Communication Complexity Is Not Trivial, Phys. Rev. Lett. 96 (2006), 250401.

[BCMdW10] Harry Buhrman, Richard Cleve, Serge Massar, and Ronald de Wolf, Nonlocality and communication complexity, Rev. Mod. Phys. 82 (2010), no. 1, 665–698.

[BCW98] Harry Buhrman, Richard Cleve, and Avi Wigderson, Quantum vs. classical communication and computation, Proceedings of the thirtieth annual ACM symposium on Theory of computing (New York, NY, USA), STOC '98, ACM, 1998, pp. 63–68.

[Bea05] Alan F. Beardon, Algebra and Geometry, Cambridge University Press, Cambridge, 2005.

[Bel64] John Stewart Bell, On the Einstein-Podolsky-Rosen paradox, Physics 1 (1964), no. 3, 195–200.

[BGLP11] Jean-Daniel Bancal, Nicolas Gisin, Yeong-Cherng Liang, and Stefano Pironio, Device-Independent Witnesses of Genuine Multipartite Entanglement, Phys. Rev. Lett. 106 (2011), 250404.

[BH01] Simon C. Benjamin and Patrick M. Hayden, Comment on "Quantum Games and Quantum Strategies", Phys. Rev. Lett. 87 (2001), 069801.

[BHK05] Jonathan Barrett, Lucien Hardy, and Adrian Kent, No Signaling and Quantum Key Distribution, Phys. Rev. Lett. 95 (2005), 010503.

[BK93] Aleksandr Vital'evich Belinskiĭ and David Nikolaevich Klyshko, Interference of light and Bell's theorem, Physics-Uspekhi 36 (1993), no. 8, 653.

[BKM+09] Thierry Bastin, Stephanie Krins, Pierre Mathonet, Michel Godefroid, Lucas Lamata, and Enrique Solano, Operational Families of Entanglement Classes for Symmetric N-Qubit States, Phys. Rev. Lett. 103 (2009), 070503.

[BKMP07] Daniel E Browne, Elham Kashefi, Mehdi Mhalla, and Simon Perdrix, Generalized flow and determinism in measurement-based quantum computation, New Journal of Physics 9 (2007), no. 8, 250.

[BKP06] Jonathan Barrett, Adrian Kent, and Stefano Pironio, Maximally Nonlocal and Monogamous Quantum Correlations, Phys. Rev. Lett. 97 (2006), 170409.

[BL12] Nicolas Brunner and Noah Linden, Bell nonlocality and Bayesian game theory, arXiv preprint arXiv:1210.1173 (2012).

[BLM+05] Jonathan Barrett, Noah Linden, Serge Massar, Stefano Pironio, Sandu Popescu, and David Roberts, Nonlocal correlations as an information-theoretic resource, Phys. Rev. A 71 (2005), no. 2, 022101.

[Blo46] Felix Bloch, Nuclear Induction, Phys. Rev. 70 (1946), 460–474.

[Boh35] Niels Bohr, Can Quantum-Mechanical Description of Physical Reality be Considered Complete?, Phys. Rev. 48 (1935), 696–702.

[Boh51] David Bohm, Quantum Theory, Prentice-Hall, New York, 1951.

[BPA+08] Nicolas Brunner, Stefano Pironio, Antonio Acín, Nicolas Gisin, André Allan Méthot, and Valerio Scarani, Testing the Dimension of Hilbert Spaces, Phys. Rev. Lett. 100 (2008), 210503.

[BPR+00] Charles H. Bennett, Sandu Popescu, Daniel Rohrlich, John A. Smolin, and Ashish V. Thapliyal, Exact and asymptotic measures of multipartite pure-state entanglement, Phys. Rev. A 63 (2000), 012307.

[BR45] Felix Bloch and Isidor Isaac Rabi, Atoms in Variable Magnetic Fields, Rev. Mod. Phys. 17 (1945), 237–244.

[Bra10] Adam Brandenburger, The relationship between quantum and classical correlation in games, Games and Economic Behavior 69 (2010), no. 1, 175 – 183.

[BRS07] Stephen D. Bartlett, Terry Rudolph, and Robert W. Spekkens, Reference frames, superselection rules, and quantum information, Rev. Mod. Phys. 79 (2007), 555–609.

[BSV12] Nicolas Brunner, James Sharam, and Tamas Vertesi, Testing the Structure of Multipartite Entanglement with Bell Inequalities, Phys. Rev. Lett. 108 (2012), 110501.

[BTD07] Ryan Barnett, Ari Turner, and Eugene Demler, Classifying vortices in S = 3 Bose-Einstein condensates, Phys. Rev. A 76 (2007), 013605.

[BV12] Nicolas Brunner and Tamas Vertesi, Persistency of entanglement and nonlocality in multipartite quantum systems, Phys. Rev. A 86 (2012), 042113.

[Cab01] Adan Cabello, Multiparty multilevel Greenberger-Horne-Zeilinger states, Phys. Rev. A 63 (2001), 022104.

[Cab02] ———, Bell's theorem with and without inequalities for the three-qubit Greenberger-Horne-Zeilinger and W states, Phys. Rev. A 65 (2002), no. 3, 032108.

[Car11] Constantin Caratheodory, Über den Variabilitätsbereich der Fourier'schen Konstanten von positiven harmonischen Funktionen, Rendiconti del Circolo Matematico di Palermo (1884-1940) 32 (1911), no. 1, 193–217.

[Car13] Elie Cartan, Les groupes projectifs qui ne laissent invariante aucune multiplicité plane, Bull. Soc. Math. France 41 (1913), no. 1, 53–96.

[Cer04] Jose L. Cereceda, Hardy's nonlocality for generalized n-particle GHZ states, Physics Letters A 327 (2004), 433–437.

[CGR08] Adan Cabello, Otfried Gühne, and David Rodríguez, Mermin inequalities for perfect correlations, Phys. Rev. A 77 (2008), 062106.

[CH74] John F. Clauser and Michael A. Horne, Experimental consequences of objective local theories, Phys. Rev. D 10 (1974), no. 2, 526–535.

[CHSH69] John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt, Proposed Experiment to Test Local Hidden-Variable Theories, Phys. Rev. Lett. 23 (1969), no. 15, 880–884.

[CI08] Taksu Cheon and Azhar Iqbal, Bayesian Nash Equilibria and Bell Inequalities, Journal of the Physical Society of Japan 77 (2008), no. 2, 024801.

[Cir80] Boris S. Cirel'son, Quantum generalizations of Bell's inequality, Letters in Mathematical Physics 4 (1980), 93–100.

[CK11] Roger Colbeck and Adrian Kent, Private randomness expansion with untrusted devices, Journal of Physics A: Mathematical and Theoretical 44 (2011), no. 9, 095305.

[CKW00] Valerie Coffman, Joydip Kundu, and William K. Wootters, Distributed entanglement, Phys. Rev. A 61 (2000), 052306.

[Col06] Roger Colbeck, Quantum and Relativistic Protocols for Secure Multi-party Computation, Ph.D. thesis, University of Cambridge, 2006.

[CRV08] Adan Cabello, David Rodríguez, and Ignacio Villanueva, Necessary and Sufficient Detection Efficiency for the Mermin Inequalities, Phys. Rev. Lett. 101 (2008), 120402.

[CSW10] Adan Cabello, Simone Severini, and Andreas Winter, (Non-) Contextuality of Physical Theories as an Axiom, arXiv preprint arXiv:1010.2163 (2010).

[CTDL97] Claude Cohen-Tannoudji, Bernard Diu, and Franck Laloë, Mécanique quantique, 2nd ed., Hermann, 1997.

[CVDM+09] Raino Ceccarelli, Giuseppe Vallone, Francesco De Martini, Paolo Mataloni, and Adan Cabello, Experimental Entanglement and Nonlocality of a Two-Photon Six-Qubit Cluster State, Phys. Rev. Lett. 103 (2009), 160401.

[CW03] Kai Chen and Ling-An Wu, A matrix realignment method for recognizing entanglement, Quantum Information & Computation 3 (2003), no. 3, 193–202.

[Dir28] Paul Adrien Maurice Dirac, The Quantum Theory of the Electron, Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character 117 (1928), no. 778, pp. 610–624.

[DKP07] Vincent Danos, Elham Kashefi, and Prakash Panangaden, The measurement calculus, J. ACM 54 (2007), no. 2, 8.

[DM05] Lokenath Debnath and Piotr Mikusiński, Introduction to Hilbert Spaces with Applications, 3rd ed., Academic press, Burlington, MA, 2005.

[DVC00] Wolfgang Dür, Guifre Vidal, and J. Ignacio Cirac, Three qubits can be entangled in two inequivalent ways, Phys. Rev. A 62 (2000), 062314.

[EPR35] Albert Einstein, Boris Podolsky, and Nathan Rosen, Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?, Phys. Rev. 47 (1935), 777–780.

[EWL99] Jens Eisert, Martin Wilkens, and Maciej Lewenstein, Quantum Games and Quantum Strategies, Phys. Rev. Lett. 83 (1999), 3077–3080.

[FC72] Stuart J. Freedman and John F. Clauser, Experimental Test of Local Hidden-Variable Theories, Phys. Rev. Lett. 28 (1972), 938–941.

[Fey82] Richard Phillips Feynman, Simulating physics with computers, International Journal of Theoretical Physics 21 (1982), 467–488 (English).

[FLS12] Tobias Fritz, Anthony Leverrier, and Ana Belen Sainz, A Combinatorial Approach to Nonlocality and Contextuality, arXiv preprint arXiv:1212.4084 (2012).

[FT91] Drew Fudenberg and Jean Tirole, Game Theory, The MIT Press, 1991.

[GBHA10] Rodrigo Gallego, Nicolas Brunner, Christopher Hadley, and Antonio Acín, Device-Independent Tests of Classical and Quantum Dimensions, Phys. Rev. Lett. 105 (2010), 230501.

[GC08] Otfried Gühne and Adan Cabello, Generalized Ardehali-Bell inequalities for graph states, Phys. Rev. A 77 (2008), 032108.

[GHZ89] Daniel M. Greenberger, Michael A. Horne, and Anton Zeilinger, Going beyond Bell's theorem, Bell's theorem, quantum theory, and conceptions of the universe 37 (1989), 69–72.

[Gis91] Nicolas Gisin, Bell's inequality holds for all non-product states, Physics Letters A 154 (1991), no. 5–6, 201 – 202.

[Got96] Daniel Gottesman, Class of quantum error-correcting codes saturating the quantum Hamming bound, Phys. Rev. A 54 (1996), 1862–1868.

[Got97] ———, Stabilizer Codes and Quantum Error Correction, Ph.D. thesis, California Institute of Technology, May 1997.

[GR10] Sibasish Ghosh and Shasanka Mohan Roy, Chain of Hardy-type local reality constraints for n qubits, Journal of Mathematical Physics 51 (2010), no. 12, 122204.

[Gro96] Lov K Grover, A fast quantum mechanical algorithm for database search, Proceedings of the twenty-eighth annual ACM symposium on Theory of computing, ACM, 1996, pp. 212–219.

[Grü03] Branko Grünbaum, Convex Polytopes, 2nd ed., Graduate Texts in Mathematics, vol. 221, Springer, New York, 2003.

[GS22] Walther Gerlach and Otto Stern, Das magnetische Moment des Silberatoms, Zeitschrift für Physik A Hadrons and Nuclei 9 (1922), no. 1, 353–355.

[GT09] Otfried Gühne and Geza Tóth, Entanglement detection, Physics Reports 474 (2009), no. 1–6, 1 – 75.

[GTHB05] Otfried Gühne, Geza Tóth, Philipp Hyllus, and Hans J. Briegel, Bell Inequalities for Graph States, Phys. Rev. Lett. 95 (2005), 120405.

[GWAN12] Rodrigo Gallego, Lars Erik Würflinger, Antonio Acín, and Miguel Navascués, Operational Framework for Nonlocality, Phys. Rev. Lett. 109 (2012), 070401.

[GYX+10] Wei-Bo Gao, Xing-Can Yao, Ping Xu, He Lu, Otfried Gühne, Adan Cabello, Chao-Yang Lu, Tao Yang, Zeng-Bing Chen, and Jian-Wei Pan, Bell inequality tests of four-photon six-qubit graph states, Phys. Rev. A 82 (2010), 042334.

[Han98] John H. Hannay, The Berry phase for spin in the Majorana representation, Journal of Physics A: Mathematical and General 31 (1998), no. 2, L53.

[Har67] John Charles Harsanyi, Games with Incomplete Information Played by "Bayesian" Players, I-III. Part I. The Basic Model, Management Science 14 (1967), no. 3, pp. 159–182 (English).

[Har68a] ———, Games with Incomplete Information Played by "Bayesian" Players, I-III. Part II. Bayesian Equilibrium Points, Management Science 14 (1968), no. 5, pp. 320–334 (English).

[Har68b] ———, Games with Incomplete Information Played by "Bayesian" Players, I-III. Part III. The Basic Probability Distribution of the Game, Management Science 14 (1968), no. 7, pp. 486–502 (English).

[Har93] Lucien Hardy, Nonlocality for two particles without inequalities for almost all entangled states, Phys. Rev. Lett. 71 (1993), 1665–1668.

[Har94] ———, Nonlocality of a Single Photon Revisited, Phys. Rev. Lett. 73 (1994), 2279–2283.

[Har01] ———, Quantum theory from five reasonable axioms, arXiv preprint quant-ph/0101012 (2001).

[HCSV11] Libby Heaney, Adán Cabello, Marcelo França Santos, and Vlatko Vedral, Extreme nonlocality with one photon, New Journal of Physics 13 (2011), no. 5, 053054.

[HDE+06] Marc Hein, Wolfgang Dür, Jens Eisert, Robert Raussendorf, Maarten van den Nest, and Hans J. Briegel, Entanglement in graph states and its applications, Quantum Computers, Algorithms, and Chaos (G. Casati, D.L. Shepelyansky, P. Zoller, G. Benenti, and Societa italiana di fisica, eds.), Proceedings of the International School of Physics "Enrico Fermi", IOS Press, 2006.

[HEB04] Marc Hein, Jens Eisert, and Hans J. Briegel, Multiparty entanglement in graph states, Phys. Rev. A 69 (2004), 062311.

[HHH96] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki, Separability of mixed states: necessary and sufficient conditions, Physics Letters A 223 (1996), no. 1–2, 1 – 8.

[HHH97] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki, Inseparable Two Spin-$\frac{1}{2}$ Density Matrices Can Be Distilled to a Singlet Form, Phys. Rev. Lett. 78 (1997), 574–577.

[HHHH09] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki, Quantum entanglement, Rev. Mod. Phys. 81 (2009), no. 2, 865–942.

[HJ12] Roger Alan Horn and Charles Royal Johnson, Matrix Analysis, 2nd ed., Cambridge University Press, Cambridge, October 2012.

[HKW+09] Robert Hübener, Matthias Kleinmann, Tzu-Chieh Wei, Carlos González-Guillén, and Otfried Gühne, Geometric measure of entanglement for symmetric states, Phys. Rev. A 80 (2009), 032324.

[Hol73] Alexander Semenovich Holevo, Bounds for the quantity of information transmitted by a quantum communication channel, Problemy Peredachi Informatsii 9 (1973), no. 3, 3–11.

[HSG+11] Marcus Huber, Hans Schimpf, Andreas Gabriel, Christoph Spengler, Dagmar Bruß, and Beatrix C. Hiesmayr, Experimentally implementable criteria revealing substructures of genuine multipartite entanglement, Phys. Rev. A 83 (2011), 022328.

[idZB02] Marek Żukowski and Časlav Brukner, Bell's Theorem for General N -Qubit States, Phys. Rev. Lett. 88 (2002), 210401.

[KN97] Eyal Kushilevitz and Noam Nisan, Communication Complexity, Cambridge University Press, 1997.

[KS67] Simon Bernhard Kochen and Ernst Paul Specker, The Problem of Hidden Variables in Quantum Mechanics, J. Math. Mech. 17 (1967), no. 1, 59–87.

[KS92] Bala Kalyanasundaram and Georg Schintger, The Probabilistic Communication Complexity of Set Intersection, SIAM Journal on Discrete Mathematics 5 (1992), no. 4, 545–557.

[KS11] Yvette Kosmann-Schwarzbach, The Noether Theorems: Invariance and Conservation Laws in the Twentieth Century, Sources and Studies in the History of Mathematics and Physical Sciences, Springer, New York, 2011.

[KWK+10] Nikolai Kiesel, Witlef Wieczorek, Stephanie Krins, Thierry Bastin, Harald Weinfurter, and Enrique Solano, Operational multipartite entanglement classes for symmetric photonic qubit states, Phys. Rev. A 81 (2010), no. 3, 032316.

[Las01] Jean-Bernard Lasserre, Global optimization with polynomials and the problem of moments, SIAM Journal on Optimization 11 (2001), no. 3, 796–817.

[Leb99] Patricio Leboeuf, Phase space approach to quantum dynamics, Journal of Physics A: Mathematical and General 24 (1999), no. 19, 4575.

[LLL+12] Lucas Lamata, Carlos E. Lopez, Benjamin Lanyon, Thierry Bastin, Juan Carlos Retamal, and Enrique Solano, Deterministic Generation of Arbitrary Symmetric States and Entanglement Classes, arXiv preprint arXiv:1211.0404 (2012).

[LWW+10] Lars Lydersen, Carlos Wiechers, Christoffer Wittmann, Dominique Elser, Johannes Skaar, and Vadim Makarov, Hacking commercial quantum cryptography systems by tailored bright illumination, Nat Photon 4 (2010), no. 10, 686–689.

[MAG06] Lluis Masanes, Antonio Acín, and Nicolas Gisin, General properties of nonsignaling theories, Phys. Rev. A 73 (2006), 012112.

[Maj32] Ettore Majorana, Atomi orientati in campo magnetico variabile, Il Nuovo Cimento 9 (1932), no. 2, 43–50.

[Maj37] ———, Teoria simmetrica dell'elettrone e del positrone, Il Nuovo Cimento 14 (1937), no. 4, 171–184.

[Maj06] Ettore Majorana, Ettore Majorana: Scientific Papers, Springer Berlin Heidelberg, 2006.

[Mar11] Damian Markham, Entanglement and symmetry in permutation-symmetric states, Phys. Rev. A 83 (2011), no. 4, 042332.

[Mer90] Nathaniel David Mermin, Extreme quantum entanglement in a superposition of macroscopically distinct states, Phys. Rev. Lett. 65 (1990), no. 15, 1838–1840.

[Mer95] ———, The Best Version of Bell's Theorem, Annals of the New York Academy of Sciences 755 (1995), no. 1, 616–623.

[Mey99] David A. Meyer, Quantum Strategies, Phys. Rev. Lett. 82 (1999), 1052–1055.

[MKG+10] Pierre Mathonet, Stephanie Krins, Michel Godefroid, Lucas Lamata, Enrique Solano, and Thierry Bastin, Entanglement equivalence of N-qubit symmetric states, Phys. Rev. A 81 (2010), 052315.

[MPA11] Lluis Masanes, Stefano Pironio, and Antonio Acín, Secure device-independent quantum key distribution with causally independent measurement devices, Nature Communications 2 (2011), 238.

[MS07] H. Mäkelä and K.-A. Suominen, Inert States of Spin-S Systems, Phys. Rev. Lett. 99 (2007), 190408.

[MS08] Damian Markham and Barry C. Sanders, Graph states for quantum secret sharing, Phys. Rev. A 78 (2008), 042309.

[Mye97] Roger Bruce Myerson, Game Theory: Analysis of Conflict, Harvard Univ Press, 1997.

[MZF+12] Vincent Mourik, Kun Zuo, Sergey M. Frolov, Sebastien R. Plissard, Erik P. A. M. Bakkers, and Leo P. Kouwenhoven, Signatures of Majorana Fermions in Hybrid Superconductor-Semiconductor Nanowire Devices, Science 336 (2012), no. 6084, 1003–1007.

[NC00] Michael A. Nielsen and Isaac L. Chuang, Quantum Computation and Quantum Information, Cambridge University Press, Cambridge, 2000.

[Nee99] Tristan Needham, Visual Complex Analysis, Oxford University Press, Oxford, 1999.

[Noe18] Emmy Noether, Invariante Variationsprobleme, Nachr. D. König. Gesellsch. D. Wiss. Zu Göttingen Math.-Phys. Kl. II (1918), 235–257.

[NPA07] Miguel Navascues, Stefano Pironio, and Antonio Acín, Bounding the Set of Quantum Correlations, Phys. Rev. Lett. 98 (2007), 010401.

[OR94] Martin J. Osborne and Ariel Rubinstein, A Course in Game Theory, The MIT press, 1994.

[Osb03] Martin J. Osborne, An Introduction to Game Theory, Oxford University Press, New York, NY, USA, 2003.

[OV06] Tobias J. Osborne and Frank Verstraete, General Monogamy Inequality for Bipartite Qubit Entanglement, Phys. Rev. Lett. 96 (2006), 220503.

[OZ01] Harold Ollivier and Wojciech Hubert Zurek, Quantum Discord: A Measure of the Quantumness of Correlations, Phys. Rev. Lett. 88 (2001), 017901.

[PAB+09] Stefano Pironio, Antonio Acín, Nicolas Brunner, Nicolas Gisin, Serge Massar, and Valerio Scarani, Device-independent quantum key distribution secure against collective attacks, New Journal of Physics 11 (2009), no. 4, 045021.

[PAM+10] Stefano Pironio, Antonio Acín, Serge Massar, Antoine Boyer de la Giroday, Dzmitry N. Matsukevich, Peter Maunz, Steven Olmschenk, David Hayes, Le Luo, T. Andrew Manning, and Christopher R. Monroe, Random numbers certified by Bell's theorem, Nature 464 (2010), no. 7291, 1021–1024.

[PB09] Marcin Pawłowski and Časlav Brukner, Monogamy of Bell's Inequality Violations in Nonsignaling Theories, Phys. Rev. Lett. 102 (2009), 030403.

[Pea70] Philip M. Pearle, Hidden-Variable Example Based upon Data Rejection, Phys. Rev. D 2 (1970), 1418–1425.

[Pen00] Roger Penrose, On Bell non-locality without probabilities: some curious geometry, Quantum Reflections (Cambridge) (John Ellis and Daniele Amati, eds.), Cambridge University Press, October 2000, p. 1.

[Pen04] ———, The Road to Reality, Jonathan Cape, 2004.

[Per93] Asher Peres, Quantum Theory: Concepts and Methods, Fundamental Theories of Physics, vol. 57, Kluwer Academic Publishers, 1993.

[Per96] ———, Separability Criterion for Density Matrices, Phys. Rev. Lett. 77 (1996), 1413–1415.

[Per99] ———, All the Bell inequalities, Foundations of Physics 29 (1999), no. 4, 589–614.

[Pit89] Itamar Pitowsky, Quantum Probability, Quantum Logic, Lecture Notes in Physics, vol. 321, Springer, 1989.

[Pit94] ———, George Boole's 'Conditions of Possible Experience' and the Quantum Puzzle, The British Journal for the Philosophy of Science 45 (1994), no. 1, 95–125.

[PPK+09] Marcin Pawłowski, Tomasz Paterek, Dagomir Kaszlikowski, Valerio Scarani, Andreas Winter, and Marek Żukowski, Information causality as a physical principle, Nature 461 (2009), no. 7267, 1101–1104.

[PR84] Roger Penrose and Wolfgang Rindler, Spinors and Space-time: Two Spinor Calculus and Relativistic Fields, vol. 1, Cambridge University Press, Cambridge, 1984.

[PR86] ———, Spinors and Space-time: Spinor and Twistor Methods in Space-time Geometry, vol. 2, Cambridge University Press, Cambridge, 1986.

[PR92] Sandu Popescu and Daniel Rohrlich, Generic quantum nonlocality, Physics Letters A 166 (1992), no. 5, 293–297.

[PR94] ———, Quantum nonlocality as an axiom, Foundations of Physics 24 (1994), no. 3, 379–385.

[PS01] Itamar Pitowsky and Karl Svozil, Optimal tests of quantum nonlocality, Phys. Rev. A 64 (2001), 014102.

[PV07] Martin B. Plenio and Shashank Virmani, An introduction to entanglement measures, Quantum Information & Computation 7 (2007), no. 1, 1–51.

[QFLM07] Bing Qi, Chi-Hang Fred Fung, Hoi-Kwong Lo, and Xiongfeng Ma, Time-shift attack in practical quantum cryptosystems, Quantum Inf. and Comp. 7 (2007), no. 1, 073–082.

[Ras07] Eric Rasmusen, Games and Information: An Introduction to Game Theory, 3rd ed., Wiley-Blackwell, 2007.

[Raz92] Alexander A. Razborov, On the distributional complexity of disjointness, Theoretical Computer Science 106 (1992), no. 2, 385–390.

[RB01] Robert Raussendorf and Hans J. Briegel, A One-Way Quantum Computer, Phys. Rev. Lett. 86 (2001), 5188–5191.

[RBB03] Robert Raussendorf, Daniel E. Browne, and Hans J. Briegel, Measurement-based quantum computation on cluster states, Phys. Rev. A 68 (2003), 022312.

[RKM+01] Mary A. Rowe, David Kielpinski, Volker Meyer, Cass A. Sackett, Wayne M. Itano, Christopher R. Monroe, and David Jeffrey Wineland, Experimental violation of a Bell's inequality with efficient detection, Nature 409 (2001), no. 6822, 791–794.

[RM11] Pedro Ribeiro and Remy Mosseri, Entanglement in the Symmetric Sector of n Qubits, Phys. Rev. Lett. 106 (2011), no. 18, 180502.

[Rom08] Steven Roman, Advanced Linear Algebra, 3rd ed., Graduate Texts in Mathematics, Springer, New York, 2008.

[Rud05] Oliver Rudolph, Further Results on the Cross Norm Criterion for Separability, Quantum Information Processing 4 (2005), 219–239.

[RVM08] Pedro Ribeiro, Julien Vidal, and Remy Mosseri, Exact spectrum of the Lipkin-Meshkov-Glick model in the thermodynamic limit and finite-size corrections, Phys. Rev. E 78 (2008), 021106.

[Sch80] Hans Schwerdtfeger, Geometry of complex numbers, Dover Publications, 1980.

[See10] Michael P. Seevinck, Monogamy of correlations versus monogamy of entanglement, Quantum Information Processing 9 (2010), 273–294.

[Sha94] Ramamurti Shankar, Principles of Quantum Mechanics, 2nd ed., Plenum Press, New York, NY (United States), 1994.

[Sho97] Peter Williston Shor, Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer, SIAM Journal on Computing 26 (1997), no. 5, 1484–1509.

[Sin05] Stephanie Frank Singer, Linearity, Symmetry, and Prediction in the Hydrogen Atom, Undergraduate Texts in Mathematics, Springer, New York, 2005.

[SN10] Jun John Sakurai and Jim J. Napolitano, Modern Quantum Mechanics, 2nd ed., Addison-Wesley, 2010.

[Spe07] Robert W. Spekkens, Evidence for the epistemic view of quantum states: A toy theory, Phys. Rev. A 75 (2007), 032110.

[SUDR12] Sudha, A. R. Usha Devi, and Attipat K. Rajagopal, Monogamy of quantum correlations in three-qubit pure states, Phys. Rev. A 85 (2012), 012103.

[Sve87] George Svetlichny, Distinguishing three-body from two-body nonseparability by a Bell-type inequality, Phys. Rev. D 35 (1987), 3066–3069.

[Sze04] Peter Szekeres, A Course in Modern Mathematical Physics: Groups, Hilbert Space and Differential Geometry, Cambridge University Press, Cambridge, 2004.

[TGB06] Geza Toth, Otfried Gühne, and Hans J. Briegel, Two-setting Bell inequalities for graph states, Phys. Rev. A 73 (2006), 022303.

[Ton09] Ben Toner, Monogamy of non-local quantum correlations, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science 465 (2009), no. 2101, 59–69.

[VB96] Lieven Vandenberghe and Stephen Boyd, Semidefinite programming, SIAM Review 38 (1996), no. 1, 49–95.

[vD99] Wim van Dam, Nonlocality & communication complexity, Ph.D. thesis, University of Oxford, 1999.

[vEP02] Steven J. van Enk and Rob Pike, Classical rules in quantum games, Phys. Rev. A 66 (2002), 024306.

[vNM07] John von Neumann and Oskar Morgenstern, Theory of Games and Economic Behavior, 60th anniversary ed., Princeton University Press, Princeton, N.J., March 2007.

[VP98] Vlatko Vedral and Martin B. Plenio, Entanglement measures and purification procedures, Phys. Rev. A 57 (1998), 1619–1633.

[Weh06] Stephanie Wehner, Tsirelson bounds for generalized Clauser-Horne-Shimony-Holt inequalities, Phys. Rev. A 73 (2006), 022110.

[Wei12] Steven Weinberg, Lectures on Quantum Mechanics, Cambridge University Press, Cambridge, November 2012.

[Wer89] Reinhard F. Werner, Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model, Phys. Rev. A 40 (1989), 4277–4281.

[WH02] Reinhard F. Werner and Alexander Semenovich Holevo, Counterexample to an additivity conjecture for output purity of quantum channels, Journal of Mathematical Physics 43 (2002), no. 9, 4353–4357.

[WJS+98] Gregor Weihs, Thomas Jennewein, Christoph Simon, Harald Weinfurter, and Anton Zeilinger, Violation of Bell's Inequality under Strict Einstein Locality Conditions, Phys. Rev. Lett. 81 (1998), 5039–5043.

[WM12] Zizhu Wang and Damian Markham, Nonlocality of Symmetric States, Phys. Rev. Lett. 108 (2012), 210407.

[WW01] Reinhard F. Werner and Michael M. Wolf, All-multipartite Bell-correlation inequalities for two dichotomic observables per site, Phys. Rev. A 64 (2001), 032112.

[XQL10] Feihu Xu, Bing Qi, and Hoi-Kwong Lo, Experimental demonstration of phase-remapping attack in a practical quantum key distribution system, New Journal of Physics 12 (2010), no. 11, 113026.

[Yao79] Andrew Chi-Chih Yao, Some complexity questions related to distributive computing(Preliminary Report), Proceedings of the eleventh annual ACM symposium on Theory of computing (New York, NY, USA), STOC '79, ACM, 1979, pp. 209–213.

[Yao93] ———, Quantum Circuit Complexity, Proceedings of the 34th Annual Symposium on Foundations of Computer Science, IEEE Computer Society, 1993, p. 352.

[YCZ+12] Sixia Yu, Qing Chen, Chengjie Zhang, Choy Heng Lai, and Choo Hiap Oh, All Entangled Pure States Violate a Single Bell's Inequality, Phys. Rev. Lett. 109 (2012), 120402.

[ZBLW02] Marek Żukowski, Časlav Brukner, Wiesław Laskowski, and Marcin Wieśniak, Do All Pure Entangled States Violate Bell's Inequalities for Correlation Functions?, Phys. Rev. Lett. 88 (2002), 210402.

[Zie95] Günter Matthias Ziegler, Lectures on Polytopes, Graduate Texts in Mathematics, vol. 152, Springer, New York, 1995.

[ZP93] Jason Zimba and Roger Penrose, On Bell non-locality without probabilities: More curious geometry, Studies In History and Philosophy of Science Part A 24 (1993), no. 5, 697 – 720.