Université de Montréal

Entanglement-assisted communication complexity and nonlocal games

par

Olivier Lalonde

Département d'informatique et de recherche opérationnelle
Faculté des arts et des sciences

Mémoire présenté en vue de l'obtention du grade de
Maître ès sciences (M.Sc.)
en informatique théorique et quantique

28 août 2023

© Olivier Lalonde, 2023

Université de Montréal

Faculté des arts et des sciences

Ce mémoire intitulé

Entanglement-assisted communication complexity and nonlocal games

présenté par

Olivier Lalonde

a été évalué par un jury composé des personnes suivantes :

Michel Boyer

(président-rapporteur)

Gilles Brassard

(directeur de recherche)

Frédéric Dupuis

(codirecteur)

William Slofstra

(membre du jury)

Résumé

Ce mémoire étudie le problème ancestral [1] de déterminer la puissance relative de l'intrication préalable en complexité de la communication comparée à la communication quantique. L'idée maîtresse du mémoire est d'opérer un rapprochement entre la complexité de la communication et la théorie des jeux non-locaux. Spécifiquement, nous contemplons une variété de manières de convertir des jeux non-locaux pour lesquels il est su que beaucoup d'intrication est requise en problèmes de complexité de la communication. Ce faisant, nous obtenons les problèmes de communications affichant les plus grands écarts connus à ce jour entre les deux modèles pour des problèmes fonctionnels.

Mots-clés: informatique quantique, intrication, jeux non-locaux

[1] Pour les standards de l'informatique quantique, qui n'est prise au sérieux que depuis quelques décennies.

Abstract

This thesis studies the age-old [2] problem of determining the relative power of shared prior entanglement in communication complexity compared to quantum communication. The central idea of the thesis is to build a connection between communication complexity and the well-developed theory of nonlocal games. To be more specific, we contemplate a variety of ways of converting nonlocal games into communication complexity problems for which it is known that a great deal of prior entanglement is required into communication complexity problems. In so doing, we obtain communication problems exhibiting the largest known gaps at the time of writing between the two models for functional problems.

Keywords: Quantum computing, entanglement, nonlocal games

[2] For the standards of quantum computing, at least.

Contents

List of symbols and abbreviations

RE	The class of recursively enumerable languages
coRE	The class of languages whose complements are recursively enumerable
$\chi(G)$	The chromatic number of G
$\xi(G)$	The orthogonal rank of G
$\omega(G)$	The clique number of G
$\alpha(G)$	The independence number of G, i.e. the clique number of its complement
$\chi_t(G)$	The quantum chromatic number of G in model t
$\chi_f(G)$	The fractional chromatic number of G
$\overline{\vartheta}(G)$	The Lovász number of the complement of G

Remerciements

J'aimerais remercier mes directeurs Gilles Brassard et Frédéric Dupuis pour leur temps, leur soutien infaillible ainsi que pour avoir valorisé mon autonomie en recherche. Je suis particulièrement reconnaissant à Gilles pour avoir bien voulu financer mon voyage au *workshop* organisé par l'Université McGill qui portait sur le résultat récent MIP*=RE et qui a eu un impact non-négligeable sur le contenu de ce mémoire, ainsi que pour avoir payé mon billet d'avion pour Amsterdam, m'y avoir acueilli et m'avoir invité au Concertgebouw. Je voudrais aussi remercier Fred pour les remarques qu'il a faites dans mes présentations au lab et qui ont eu un impact énorme sur la direction qu'a prise ma recherche. Si ce n'était de lui, le modèle de communication avec intrication préalable étudié dans ce mémoire serait avec communication quantique et la quasi-totalité des résultats du chapitre 4 n'auraient jamais été trouvés.

J'aimerais aussi remercier Louis Salvail pour sa présence toujours divertissante et pour le séjour au chalet ainsi que Michel Boyer pour avoir accepté la lourde responsabilité de présider mon jury. Je voudrais également saluer tous les autres étudiants et postdocs qui ont été de passage au LITQ ou dans le labo de quantique de McGill pendant mon séjour, soient Mohammed Barhoush, Sophie Berthelette, Mathieu Bérubé-Vallières, Julien Codsi, Nick Courtemanche, Samuel Ducharme, Shany Gazaille, Muxue Guo, Simon Hennessey-Patry, Philippe Lamontagne, Rémi Ligez, Rotem Liss, Ashutosh Marwah, Vincent Quirion, John Stuart, Louis-Charles Thibodeau, Samuel Whaite et Andrew Williams. Je voudrais spécialement saluer Julien Codsi pour m'avoir invité au JuliaCon et pour les très nombreuses conversations intéressantes, ainsi que Vincent Quirion pour avoir bien voulu collaborer avec moi un été de temps. Je voudrais également remercier Jean Laprés-Chartrand pour le paquet de cartes que j'ai gagné comme prix pour avoir gagné le tournoi de dix de l'association étudiante avec lui, une semaine après avoir appris les règles du dix. Je voudrais enfin saluer mes coéquipiers Henri Zhu, Hou Han Zhang, Raphaël Lima-Barbosa, Raphaël Nunez, Gabriel Tinica, Noela-Joyce Lomandong et Emmanuel Hebrard, avec qui j'ai eu l'honneur de participer à trois éditions du championnat universitaire pancanadien d'échecs avec un succès non-négligeable.

I would like to thank the universities of Amsterdam and Waterloo for their hospitality. It was certainly a pleasant experience to be able to discuss the intricacies of the Cleve-Buhrman model with Richard Cleve and Harry Buhrman. This thesis would be a great deal shorter if it hadn't been for the input of several people from these two places, to whom I am indebted. Subsection 4.2.8 about direct sums in one-way, zero-error communication complexity was inspired by conversations with Harry Buhrman and Ronald de Wolf during my stay in Amsterdam; section 4.1 as a whole simply would not exist had it not been for a conversation with Yuming Zhao during my stay in Waterloo, and Richard Cleve referred me to the result of Slofstra about the amount of entanglement required to play a XOR game near-optimally, without which the results of section 4.1 would not be very interesting; William Slofstra and Kieran Mastel referred me to the very recent result of Sam Harris ([**101**]), without which subsection 4.2.6, which probably contains the most interesting results of this thesis, would not have come into being; finally, Ashwin Nayak referred me to the results of [**76**] and also for having proofread 3.3.5. I would also like to thank Anand Natarajan and John Wright for interesting discussions about their MIP*=RE result; the presentation of nonlocality in chapter 2 owes much to them. I also thank David Roberson for helpful conversations about quantum graph theory, Travis Russell for useful conversations about his hierarchy, as well as Junqiao Lin for engaging discussions about nonlocal games back in Amsterdam which set me on the way to the results of chapter 4. I would also like to thank Ronald de Wolf for having sent me the master's thesis of his former student Jonas Kamminga, from which some parts of this thesis have drawn inspiration. Finally, I would like to thank William Slofstra very much for agreeing to be on my thesis committee.

J'aimerais également remercier mes parents pour leur soutien au fil des ans.

Chapter 1

Introduction

This work is devoted to the study of the role of entanglement in communication complexity. Communication complexity, first introduced by Abelson ([**6**]) and formalized by Yao ([**8**]), concerns itself with the following problem:

(1) Given a discrete-valued function $f(x,y)$ known to both parties (typically taken to be boolean), if Alice is given x and Bob is given y, how much do they need to communicate in order to compute $f(x,y)$?

At first, this task was studied in the classical setting, where communication was quantified in terms of bits. With the advent of quantum information theory, two proposals for a quantum analogue of the aforementioned classical setting were put forth. The first is due to Yao ([**19**]), and consists in allowing Alice and Bob to communicate with quantum resources, specifically qubits. The second is due to Cleve and Buhrman ([**26**]), and consists in allowing Alice and Bob to share arbitrary prior quantum entanglement while keeping the communication classical. The teleportation protocol of Bennett, Brassard, Crépeau, Jozsa, Peres and Wootters ([**17**]) implies that any protocol in the Yao model may be simulated in the Cleve-Buhrman model by increasing the amount of communication by a factor of two, so that the Cleve-Buhrman model is at least as powerful as the Yao model up to a constant factor.

Although this took more than a decade to establish, we now know that the Yao model (and, by extension, the Cleve-Buhrman model which subsumes it) can yield an enormous (exponential) reduction in the amount of required communication for some problems compared to the classical setting. However, the relative strength of the Cleve-Buhrman model compared to the Yao model is much more mysterious, despite these having been around for more than 25 years. In fact, while it is known that the Cleve-Buhrman model can be much more powerful than the Yao model for a multitude of related communication tasks, it is still unknown whether the Cleve-Buhrman model can be more powerful than the Yao model for tasks of the form (1), provided that shared randomness is considered free.

There are essentially two school of thoughts about the Yao vs Cleve-Buhrman problem. One might be led to think that they should be equivalent because, despite years of efforts, essentially no use for entanglement was known for tasks of the form (1) aside from generating classical shared randomness and teleporting qubits (using either the original teleportation protocol ([**17**]), which corresponds to a factor of two overhead, or a more efficient remote state preparation protocol (see, for example, [**39**])). In particular, no candidates at all were known that might possibly exhibit a separation between the Cleve-Buhrman and Yao models. In contrast, one might think that since entanglement is known to be an extremely powerful resource in so many relating settings, most prominently in the context of nonlocal games as exemplified in the recent MIP*=RE result ([**93**]), some of that power ought to drip off on communication complexity in a hitherto unknown manner.

1.1. Contributions of this thesis

The point of view taken on the Yao vs Cleve-Buhrman problem in this thesis is resolutely in favour of the second point of view. We make what we feel is reasonably significant headway towards showing this by showcasing more ways in which entanglement may be used in communication complexity. A number of these come from the theory of nonlocal games, which has been progressing at a rapid pace in recent years. While the idea that communication complexity and nonlocality are intertwined is in no way new (see, for example, [**81, 64**]), we are, we believe, the first to attempt to systematically turn nonlocal games into communication complexity problems.

1.1.1. What is entanglement good for in communication complexity?

We will study the problem of the amount of entanglement (measured in ebits) that may be needed by a communication protocol in the Cleve-Buhrman model. Up till now, provided that shared randomness is considered free, no scenario was known in which more shared prior entanglement was needed than communication. If no such scenario existed, we would trivially have that the Yao and Cleve-Buhrman models are roughly equivalent, as one could convert a Cleve-Buhrman protocol into a Yao one by having one party generate the prior entanglement locally, dispatch it to the other party and then carry on with the original protocol, thereby incurring a factor of two overhead in the communication.

We give two examples of such scenarios. First, based on a theorem of Slofstra, in section 4.1, we show that there are some problems for which any protocol using only one bit of communication whose success probability is close to being best possible must use an amount of shared prior entanglement that is exponential in the lengths of Alice and Bob's inputs, when measured in ebits. In section 4.2, we get a stronger result in the exact setting by making use of another theorem of Slofstra: namely, we show that there exists a sequence of functions $\{f_i\}$, where the inputs of function f_i are bit strings of length i, which can all be computed exactly with one trit of communication from Alice to Bob in the Cleve-Buhrman model, but which are such that the amount of shared prior entanglement required by any such protocol for the function f_i grows faster than any computable function of i.

While the above examples collapse if some breathing room in terms of success probabilities is permitted (meaning that in all cases, there will exist a protocol with the same amount of communication and very slightly worse success probability that uses far less entanglement), we also give a heuristic construction inspired by the interactive hashing protocol of Naor, Ostrovsky, Venkatesan and Yung ([**31**]) which, together with the recent MIP*=RE result of Ji, Natarajan, Vidick, Wright and Yuen ([**93**]), seems to suggest that no useful (e.g. computable) upper bound exists on the amount of entanglement that may be required to compute a given function with good probability and with a near-optimal amount of communication.

In another vein, in subsection 3.3.2, we give an example of an entanglement-assisted one-way protocol for the equality function which achieves the same success probability as the standard one using shared randomness for the same amount of communication but whose entanglement cost is reduced by a factor of two compared to that of any correct protocol using entanglement exclusively as a source of shared randomness. This may be seen as another way in which entanglement can be used in communication complexity which differs from teleportation or the generation of shared randomness, although not in a useful way from the standpoint of computing functions with as little communication as possible.

1.1.2. Small separations between the Cleve-Buhrman and the Yao models

Up till now, as we mentioned previously, no use was known for entanglement in communication complexity other than generating shared randomness and teleporting qubits. In particular, no problem was known for which n bits of communication with prior shared entanglement allowed one to do any better than n qubits of communication with shared randomness, for some n.

We give three examples of such problems. All three problems are explicit, and the first and the third are quite small. First, again based on the aforementioned theorem of Slofstra, in section 4.1, we give a collection of problems for which a protocol with one bit of communication and shared prior entanglement achieves a success probability that is slightly larger than that of any protocol using one qubit of communication and shared randomness. Also, in subsection 4.2.7, we give an example of a function derived from a result due to Mančinska and Roberson which admits an exact

protocol in the Cleve-Buhrman model using two bits of communication while no exact protocol using two qubits of communication and no prior entanglement exists. We also provide our own example of such a function, based on the concept of a vector clump which we invented for this express purpose, which has smaller inputs than Mančinska and Roberson's example.
More significantly, in subsection 3.3.5, inspired by the work of Hadiashar and Nayak ([**91**]), we introduce a new communication complexity problem, called the distance between subspaces problem, which appears to have the potential for yielding a large (possibly even exponential) separation between the Cleve-Buhrman and Yao models. As far as we know, this is the only known candidate for such a separation.

1.1.3. Separating the different flavours of entanglement in communication complexity

Although, until now, we have been speaking of 'the' Cleve-Buhrman model, in fact, there are multiple Cleve-Buhrman models, in view of the familiar fact from nonlocality theory that there are multiple ways of formalizing entanglement mathematically. To the best of our knowledge, before this thesis, shared entanglement in communication complexity has always been formalised in the standard finite-dimensional, tensor-product model. On the one hand, we show that many lower bounds that were shown to hold in the case of tensor-product entanglement hold for the stronger commuting operators model also. On the other hand, we show that a number of separations between the entanglement models which have been shown in recent years carry over to communication complexity. First, we show that there exists a function f for which, for every $\varepsilon > 0$, there exists a entanglement-assisted protocol in the tensor-product model with one trit of communication that achieves success probability $1-\varepsilon$, and yet not such protocol exists that achieves success probability one. This parallels a theorem of Slofstra (theorem 2.5.1) and is proven by reduction from it. We also show that there exists a function g which admits an exact entanglement-assisted protocol in the commuting operators model with one trit of communication, whereas any entanglement-assisted protocol for g in the tensor-product model with one trit of communication has success probability at most $1-\varepsilon$, for some (unknown and likely incredibly small) $\varepsilon > 0$. This parallels the MIP*=RE result ([**93**]) and is also proven by reduction from it.

1.2. Structure of this thesis

This thesis assumes that the reader is acquainted with the basic language of quantum information theory already. If this is not the case, we refer the reader to, for example, the excellent textbook of Watrous ([**86**]). However, no background knowledge about nonlocal games or communication complexity is assumed.

Chapter 2 presents the basics of nonlocality, notably the definitions of a correlation set and a nonlocal game, the CHSH game, pseudotelepathy, the sum-of-squares hierarchy as well as a discussion of a constellation of results that have been shown for the various models of entanglement in recent years, culminating in the recent MIP*=RE result and its implications.

Chapter 3 is devoted to the basics of communication complexity. Specifically, we present the *dramatis personæ* of the thesis, namely, the classical model, the Yao model and the Cleve-Buhrman model. While the discussions of the first two models present mostly standard results, the third contains a fair amount of new material, including a novel entanglement-assisted protocol for the equality function and the recasting in the commuting-operators model of well-known communication complexity lower bounds. We also discuss a possible limitation of the power of the Cleve-Buhrman model relative to the Yao model due to de Wolf and present a candidate for a large separation between the two models in the bounded-error setting.

Chapter 4 contains the bulk of the new material contained in this thesis. In the first section, we go over what we call one-time-pad problems, which are a generalisation of the distributed CHSH problem of [**38**] and which turn out to be very closely related to the theory of XOR games, the basic theory of which we sketch. From there, based on a theorem of Slofstra, we show that near-optimal protocols with one bit of communication for some one-time-pad problems require a great deal of prior entanglement in the Cleve-Buhrman model, and also exhibit small gaps between the Yao and Cleve-Buhrman models. We will also discuss a generalisation of XOR games and its relevance to communication complexity. In the second section, we build the theory of one-way, exact communication complexity, which, bizarrely enough, never seems to have been studied explicitly in the Cleve-Buhrman model despite having received a fair amount of attention in the Yao model. As it turns out, this subject is closely related to quantum graph theory, as was first noticed by Ronald de Wolf in his PhD thesis ([**37**]): we make this link even more explicit. After the presentation of the graph parameters which are of interest for the theory, namely, the communication and chromatic numbers of a graph, and the statement of some useful lemmas, subsection 4.2.6 studies the comparison between the communication and chromatic numbers: there, it is shown how to embed the chromatic numbers inside communication complexity. In conjunction with the results of Harris ([**101**]), this has a number of implications for communication complexity which were discussed in the previous section. We then discuss two problems for which the Cleve-Buhrman model is slightly more powerful than the Yao model with the equivalent amount of communication, one due to Mančinska and Roberson and the other due to us. We end by presenting some negative results on direct sums in one-way, exact communication complexity. Section 3 discusses a conjectured embedding of nonlocal games in communication complexity which, if correct, would imply a stronger impossibility result on showing a limitation of prior entanglement than the ones that could be established mathematically, and section 4 discusses how our results restrict the realm of possibilities for a hypothetical quantum Newman's theorem.

Chapter 5 summarises what has been studied in this thesis and mentions some possible future research directions.

Chapter 2

The theory of quantum nonlocality

This chapter covers the basics of the theory of quantum nonlocality[1]. Although the material that we present is in no way new, we are not aware of any source that covers it in a satisfactory manner.

After defining the notions of a correlation set and of a nonlocal game, we present the simplest nonlocal game that is interesting, namely, the classic CHSH game; we then introduce the notion of pseudotelepathy, discuss the sum-of-squares hierarchy and the commuting operators model, and finally talk about the recent progress that has been made on the complexity of certain properties of nonlocal games and the relative strengths of the various entanglement models. While the account of the theory of quantum nonlocality that is presented in this chapter is quite high-level, in chapter 4, a more in-depth presentation of the theory of certain classes of nonlocal games will be provided.

2.1. Correlation sets

Perhaps the most fundamental concept in the theory of nonlocality is that of a **correlation set**. Given finite **input sets** X,Y and finite **output sets** A,B, we have the following experiment in mind: Alice and Bob, who are disallowed to communicate but who are assumed to have unbounded computational power, are respectively given $x \in X$, $y \in Y$ (with neither knowing what the other party's input is) and are requested to respectively output $a \in A$, $b \in B$. We also allow them to make their outputs depend on a random variable with the distribution of their choice (independent of their inputs, of course), the value of which both parties are made aware of: this additional resource is known among physicists as a **local hidden variable**, and is more commonly referred to as a **public coin** or as **shared randomness** among computer scientists. Given the resources that the players are permitted to use, we then consider the **correlations** that their outputs might exhibit given their inputs. The collection of all such correlations is referred to as the **correlation set**

[1] It should be emphasised that this is questionable terminology considering that nonsignalling and locality have been argued to be equivalent by Raymond-Robichaud ([**84**]), so that what is known as quantum nonlocality is, in fact, arguably a perfectly local phenomenon. We stick to this term, *faute de mieux*.

corresponding to the allowed resources, and is denoted by C_t, where t is the model under scrutiny. Formally, an element of C_t is a collection of distributions on $A \times B$ indexed by $X \times Y$, and will be denoted by $\{p(a,b|x,y)\}$. Given a correlation in C_t, a particular means by which the parties can achieve that correlation is referred to as a **strategy**. Note that the fact that the players are allowed to share randomness has the effect of making correlation sets convex (viewing correlations as vectors in $\mathbb{R}^n$).

Arguably the smallest correlation set that is still interesting is the so-called class of **classical**[2] **correlations**, denoted by C_c, which corresponds to correlations that can be achieved by classical players. This set is readily described as the convex hull of **deterministic correlations**, namely those in which the output of each party is determined completely by their input. At the other extreme, we have the so-called class of **non-signalling correlations**, C_{ns}, namely, the class of correlations such that, for each of a given player's inputs, the probability distribution of the player's output is independent of the other player's input. Operationally, this means that whatever one of the players is doing cannot have any observable effect on what the other will observe: in particular, these are the correlations that do not enable telepathy (i.e. allowing the players to communicate without actually communicating physically). The correlation sets that will be considered in this thesis will all be subsets of C_{ns}. While there are very powerful reasons to think that this class is uninteresting as a model of what is physically possible in the world that we live in, it can sometimes nevertheless be a useful object of study as a more manageable relaxation of correlation classes that are physically relevant but whose structure is more complicated. For instance, membership in the class C_{ns} may be recast as a linear program and can therefore be tested very efficiently in practice: similarly, the non-signalling value of a nonlocal game, to be defined in the next section, is a linear program and is efficiently computable. This is in stark contrast to the case of most of the other correlation sets under consideration in this thesis, for which these two tasks are either known or thought to be intractable in general.

Sandwiched between C_c and C_{ns}, we have various correlation sets corresponding to models inspired by quantum mechanics. Perhaps the most physically interesting one is the set C_{qf}[3] of correlations that can be realised using the standard finite-dimensional, tensor-product formalization of entanglement. Formally, given a finite-dimensional Hilbert space $\mathscr{H}$, a strategy in this context is specified by an entangled state $|\psi\rangle$ on $\mathscr{H} \otimes \mathscr{H}$ and measurements $\{A_a^x\}$[4] and $\{B_b^y\}$ on $\mathscr{H}$ for

[2] Classical in the sense of special relativity, since nonrelativistic physics is nonlocal in the fullest sense of the word. This is sometimes referred to as the class of local correlations: refer to the previous footnote for why we dislike this terminology.

[3] Traditionally denoted C_q, where ‘q’ stands for quantum: we dislike this terminology because there are multiple other quantum correlation sets, and prefer ‘qf’, which stands for the more precise ‘quantum finite-dimensional‘.

[4] In the remainder of this thesis, a subscript will be understood as indexing the operators which make up a measurement, and a superscript will be understood as indexing over a collection of measurements.

Alice and Bob, respectively, which may be assumed to be projective measurements without loss of generality by Naimark's theorem. The correlation that is produced by this strategy is then given by:

$$p(a,b|x,y) = \langle \psi | A_a^x \otimes B_b^y | \psi \rangle$$

Note that when considering models in which the parties are allowed to share an arbitrary amount of entanglement, since entanglement can be used to generate shared randomness, it can be assumed that the parties do not share a public coin as we permitted them to do in general. We will also sometimes consider the smaller set $C_{qf:d}$, which corresponds to correlations that can be realised in the finite-dimensional model using an entangled state of local dimension at most d. Clearly,

$$C_{qf} = \bigcup_d C_{qf:d}$$

By dropping the requirement that $\mathscr{H}$ be finite-dimensional in the previous definition, we obtain a potentially larger set, called C_{qs}. One might also consider the closure of these two sets: indeed, while it is easy to show that C_c and C_{ns} are closed, it is not obvious that either of C_{qf} or C_{qs} is. Scholz and Werner ([**61**]) showed that the closures of these two sets are equal, and we will denote this closure by $C_{\overline{q}}$ [5].

We will also be studying a strengthening of the tensor-product model called the **commuting operators model**, first introduced by Tsirelson ([**10**]). This model is obtained by dropping the requirement that the entanglement be in tensor-product form altogether and only requiring that Alice and Bob's measurements commute. Formally, a strategy in this more exotic model is specified by a possibly infinite-dimensional Hilbert space $\mathscr{H}$, a state $|\psi\rangle \in \mathscr{H}$, and choices $\{A_a^x\}$ and $\{B_b^y\}$ of projective measurements for Alice and Bob, respectively, with the property that for all x,y,a,b, $[A_a^x, B_b^y] = 0$. The correlation in the corresponding correlation set C_{qc} is then given by

$$p(a,b|x,y) = \langle \psi | A_a^x B_b^y | \psi \rangle$$

It is immediate that $C_{qf}, C_{qs} \subseteq C_{qc}$. Also, as a corollary of the convergence of the sum-of-squares hierarchy, to be discussed in the next section, we have that C_{qc} is closed, so that $C_{\overline{q}} \subseteq C_{qc}$ as well. It is known that if one restricts $\mathscr{H}$ to be finite-dimensional in the previous definition, the resulting correlation set is equal to C_{qf}: see, for example, theorem 5.2.4 of Vos's master's thesis ([**88**]) for a short proof of this.

A few words should be said about why one might be interested in this new model. While this model was originally introduced by inspiration from algebraic quantum field theory, where spacelike separations are enforced by imposing commutativity relations, it would seem that the

[5] Traditionally denoted C_{qa}, where qa is supposed to stand for ‘quantum approximate‘. We prefer our notation, which emphasises that this set is a closure.

actual quantum field theories which we believe describe describe physical reality appear to be approximable by tensor products [6]. As far as we are concerned, the real reason why this model is interesting is the fact that the sum-of-squares hierarchy, which we introduce later in this chapter, converges to it.

In summary, we have that, for all d:

$$C_c \subseteq C_{qf:d} \subseteq C_{qf} \subseteq C_{qs} \subseteq C_{\overline{q}} \subseteq C_{qc} \subseteq C_{ns}$$

Provided that the input and output sets are all of size at least 2, it is straightforward in retrospect to show that $C_c \subsetneq C_{qf:2} \subseteq C_{qf}$, and this will be done in next section. The result that C_c and C_{qf} are different in general is known as **Bell's theorem**. As a corollary of the so called Tsirelson's bound, it will also follow that $C_{qc} \subsetneq C_{ns}$. All there remains to determine is how the various entanglement models compare. Tsirelson thought he had proved that they are all equal, but after it transpired that his proof was incorrect, this was downgraded as a conjecture, which became known as **Tsirelson's problem**. We will see at the end of this section that, in fact, every single inequality in the chain above is strict provided that the input and output sets are taken to be large enough, although the proofs of these separations are quite involved and will not be described.

2.2. Nonlocal games

While very general, the notion of a correlation set is somewhat difficult to work with directly. In order to compare correlation sets, it is more convenient to evaluate how well Alice and Bob can perform a certain task given that they are allowed to share a correlation of their choice from a given correlation set. The most fruitful such task is given by the notion of a nonlocal game, which we now define.

Definition 1. *A (two-player)* ***nonlocal game*** *(*[**44**]*) G is specified by finite input sets X, Y, finite output sets A, B, a probability distribution $\mu_{x,y}$ on $X \times Y$ and a predicate $V : X \times Y \times A \times B \rightarrow \{0,1\}$.*

Operationally, we think of a nonlocal game as encoding the data corresponding to the following scenario. Alice and Bob are as in the previous section. The inputs $x \in X$, $y \in Y$ are now sampled according to the distribution $\mu_{x,y}$, and a binary outcome is determined based on the inputs and the outputs, which is encoded by the predicate V. We think of this outcome as the players winning or losing the game. For a given correlation set C_t, we will be interested in the highest possible winning probability that can be achieved using correlations from that set: this will be called the

[6] The reader is referred to, for example, the discussion section of `https://scottaaronson.blog/?p=4512`. See also [**83**] for an argument for why infinite-dimensional Hilbert spaces are unphysical in the context of quantum gravity.

value of the game and will be denoted by $\omega_t(G)$. Formally,

$$\omega_t(G) = \sup_{\{p(a,b|x,y)\}\in C_t} \sum_{x\in X, y\in Y} \mu_{x,y} \sum_{\substack{a\in A, b\in B \\ V(x,y,a,b)=1}} p(a,b|x,y)$$

Note that since the above definition is framed as a weighted average and the weighted average of a collection of values is no greater than the maximum of the said collection, we can assume without loss of generality that the players do not make use of the shared randomness which we authorised them to use when defining the notion of a correlation set. Quantumly, it may be seen that the values of a game with respect to the correlation sets C_{qf}, C_{qs} and $C_{\overline{q}}$ are all equal because these sets all have the same closure. This value will simply be denoted by ω^* in all that follows. Note that, for any game G,

$$\omega^*(G) = \lim_{d\to\infty} \omega_{qf:d}(G)$$

2.3. Examples of nonlocal games

In this section, we discuss the classic CHSH game of Clauser, Horne, Shimony and Holt ([**2**]), as well as the so-called phenomenon of pseudo-telepathy, which, in the bipartite case, is most simply exhibited by the *magic square game*, whose history of discovery is somewhat difficult to unentangle but which is most commonly ascribed to Mermin and Peres.

2.3.1. The CHSH game

The most prominent example of a nonlocal game is the **CHSH game**, first introduced by [**2**] [7]. In this case, we take $X = Y = A = B = \{0,1\}$, we take the input distribution $p_{x,y}$ to be uniform and we set

$$V(x,y,a,b) = [(x \wedge y) = (a \oplus b)]$$

Checking all 16 possible deterministic strategies reveals that none of them wins the game with certainty. This implies that any deterministic strategy must fail on at least one input, and therefore that the success probability of any classical strategy is at most $\frac{3}{4}$. Conversely, the strategy in which both players output ‘0’ irrespective of their inputs can be seen to achieve this winning probability. We therefore have that $\omega_c(CHSH) = \frac{3}{4}$.

As it turns out, the sharing of entanglement enables the players to achieve a higher winning probability. This can be done using a strategy that employs the following maximally entangled state:

$$|\psi_{AB}\rangle = \frac{1}{\sqrt{2}}|0\rangle_A|0\rangle_B + \frac{1}{\sqrt{2}}|1\rangle_A|1\rangle_B$$

[7] This work, along with its experimental implementation, netted Clauser the 2022 Nobel prize for physics.

Letting R_θ correspond to the basis $\{\cos\theta\,|0\rangle+\sin\theta\,|1\rangle\,,-\sin\theta\,|0\rangle+\cos\theta\,|1\rangle\}$, Alice will measure her part of the state with R_0 if $x=0$ and $R_{\frac{\pi}{4}}$ otherwise, and Bob will measure his part of the state with $R_{\frac{\pi}{8}}$ if $y=0$ and $R_{\frac{-\pi}{8}}$ otherwise. If $x\wedge y=0$, the angle between the vectors corresponding to the same measurement outcome will be $\frac{\pi}{8}$, so the probability that Alice and Bob will output the same bit is $\cos(\frac{\pi}{8})^2=\frac{1}{2}+\frac{\sqrt{2}}{4}\approx 0.854$. If $x\wedge y=1$, this same angle is $\frac{3\pi}{8}$, so that the same probability is $\cos^2(\frac{3\pi}{8})=\frac{1}{2}-\frac{\sqrt{2}}{4}\approx 0.146$. In other words, for all inputs, the probability that the winning condition is met is exactly $\frac{1}{2}+\frac{\sqrt{2}}{4}\approx 0.854$, which is therefore the winning probability of the strategy. Since $\omega_c(CHSH)<\omega_{qf:2}(CHSH)$, it follows that provided that the input and output sets are all of cardinality two or greater, we have:

$$C_c \subsetneq C_{qf:2}$$

This fact is known as **Bell's theorem**. It can also be shown that the entangled strategy we derived above for the CHSH game is actually optimal, even in the stronger commuting operators model. This result is known as **Tsirelson's bound** ([**10**]), and can be readily shown using the first level of the sum-of-squares hierarchy that we will describe in the next section. Furthermore, the correlation such that no matter the inputs, a and b are both uniformly distributed and individually independent of both inputs and $a\oplus b=x\wedge y$ with probability one can be checked to be nonsignalling and wins the CHSH game with probability one on all inputs. This implies:

$$C_{qc} \subsetneq C_{ns}$$

Therefore, the only correlation sets that remain to be compared are the quantum ones.

We note that the CHSH game does a fair amount of heavy lifting despite its apparent simplicity: it already separated the classical, quantum and non-signalling models, and in section 1 of chapter 4, we will encounter no fewer than 3 different generalisations of this game, each with very interesting features.

2.3.2. Perfect strategies and pseudo-telepathy

Given a nonlocal game G and a model t, we will be interested in whether t admits a strategy which wins the game G with probability one. Such a strategy will be referred to as being **perfect**. If we are only interested in the existence of a perfect strategy in a given setting, the only useful information provided by the distribution $p_{x,y}$ is which input pairs have nonzero probability, i.e. if p and q are such that $p_{x,y}=0$ if and only if $q_{x,y}=0$, then G has perfect strategy under input distribution p if and only if it has a perfect strategy under input distribution q. In this context, therefore, we can simply forget about the input distribution and speak of a promise instead: namely, for some subset $P\subseteq X\times Y$, Alice and Bob are given inputs $x\in X, y\in Y$ with the promise that $(x,y)\in P$ and they are expected to output $a\in A, b\in B$ such that $V(x,y,a,b)=1$ with certainty. In this context, a strategy will be referred to as being **ε-perfect** if, for every possible choice of legal inputs, the probability that Alice and Bob win the game is at least $1-\varepsilon$. Clearly, if a given game has a perfect strategy in a given model, it will have value one in that model, but the converse is not true in general. However, it does hold if the underlying correlation set is closed (and hence compact), by the extreme value theorem.

An interesting feature of some nonlocal games is that of being **pseudo-telepathic**, meaning that the game has a perfect finite-dimensional quantum strategy but no perfect classical strategy. The most prominent (and simplest known in the bipartite setting [8]) example of a game exhibiting pseudo-telepathy is the **magic square game** of Mermin and Peres. We think of Alice and Bob as filling a 3×3 square with binary values. Alice is given an index $x \in \{1,2,3\}$ for a row, Bob is given an index $\{1,2,3\}$ for a column, and they are asked to reply with assignments to all three binary values which fill up their respective row and column in such a way that the assignments agree on the square on which the row and the column intersect, and also that the XOR of the values outputted by Alice be zero and that the XOR of the values outputted by Bob be one. It can easily be checked that no perfect classical strategy for this game exists: indeed, taking said strategy to be deterministic, it can be seen that this would amount to the existence of bits $\{b_{i,j}\}_{i,j\in[3]}$ such that:

$$b_{i,1} \oplus b_{i,2} \oplus b_{i,3} = 0,\ \forall i \in [3]$$

$$b_{1,j} \oplus b_{2,j} \oplus b_{3,j} = 1,\ \forall j \in [3]$$

But this can easily be seen to be impossible because the first equation implies that $\sum_{i,j} b_{i,j}$ is even and the second equation implies that it is odd. However, it turns out that this game does have a perfect entangled strategy which makes use of two EPR pairs. This strategy is derived in chapter 3 of ref. [**89**].

We note in passing that an interesting feature of a game exhibiting pseudo-telepathy is that it can be turned into a game with quantum value one and classical value arbitrarily close to zero in the following way. Given a game G and $n \in \mathbb{N}$, let us define the parallel repeated game G^n in the following way: Alice and Bob are each given n inputs $x_1,\dots,x_n$ and $y_1,\dots,y_n$, respectively, with each input pair sampled independently from the distribution associated to G, and they must provide outputs $a_1,\dots,a_n$ and $b_1,\dots,b_n$ such that $V(x_i,y_i,a_i,b_i)$ holds for all i. Clearly, if G has a perfect strategy in a given model, so will G^n. If not, one would expect that the games thus obtained get increasingly harder as n grows, with the value of G^n going to zero. Classically, this is indeed the case, thanks to the following deceptively difficult result due to Raz :

Theorem 2.3.1 ([**30**]). *Suppose that a game G has no perfect classical strategy. There exists a constant $C < 1$ which is efficiently computable from the description of the game such that, for all*

[8] It may be argued that the three-player GHZ game is even simpler.

n,

$$\omega_c(G^n) \leq C^n$$

This result was ported to the quantum setting by Bavarian, Vidick and Yuen in the following form:

Theorem 2.3.2 ([**75**]). *There exists an efficiently computable transformation (called* ***anchoring****) which, given a nonlocal game G, produces another nonlocal game G' such that, in all models, G' has value one if and only if G does. Furthermore, under the assumption that we somehow know an $\varepsilon > 0$ such that $\omega^*(G) < 1 - \varepsilon$, there exists a constant $C < 1$ which is efficiently computable from the description of the game such that, for all n,*

$$\omega^*(G'^n) \leq C^n$$

Finally, the following result, due to Cleve, Høyer, Toner and Watrous, will be useful to us later on:

Theorem 2.3.3 ([**44**]). *If a game G is such that $|A| = |B| = 2$, then G is not pseudo-telepathic.*

We note that the notion of pseudo-telepathy above was stated for the finite-dimensional model. An extended meaning of this notion would refer to games which have a perfect commuting-operators strategy but no perfect classical strategy. It can be shown that the above result still holds for this extended meaning with little modification to the proof.

2.4. The sum-of-squares hierarchy

In this section, we introduce a sequence $C_{qc:1}, C_{qc:2}, \ldots$ of correlation sets. These were first put forth in 2008 independently by Navascués, Pironio and Acín ([**57**]) and by Doherty, Liang, Toner and Wehner ([**58**]). This sequence of correlation sets is often called the NPA hierarchy, which we think is unfair considering that this hierarchy was discovered independently by another group of researchers. The $C_{qc:n}$ have the following properties:

(1) It holds that

$$C_{qc:1} \subseteq C_{ns}$$

Also, for every n,

$$C_{qc:n+1} \subseteq C_{qc:n}$$

In particular, for any nonlocal game G,

$$\omega_{qc:n+1}(G) \leq \omega_{qc:n}(G) \leq \omega_{ns}(G)$$

(2) We have that

$$C_{qc} = \bigcap_{n=1}^{\infty} C_{qc:n}$$

In particular, for any nonlocal game G,

$$\omega_{qc}(G) = \lim_{n \to \infty} \omega_{qc:n}(G)$$

(3) The $C_{qc:n}$ are all closed sets: this, together with the last property, implies that C_{qc} is also closed.

(4) For every n, $C_{qc:n}$ can be parameterised by a semidefinite program of size $O(A^n)$, where A depends polynomially on the size of the game. Also, this semidefinite program satisfies Slater's condition. It then follows from the work of Nesterov and Nemirovskii ([**18**]) that for a fixed n and given a convex distance metric D satisfying minor technical conditions (such as Kolmogorov distance, for example), there exists [9] a computer program which, given a correlation $p(a,b|x,y)$, can efficiently compute

$$\inf_{q(a,b|x,y) \in C_{qc:n}} D(p,q)$$

to within arbitrary precision (and, in particular, prove that $p(a,b|x,y) \notin C_{qc:n}$ if this indeed the case), or alternatively, given a nonlocal game G, computes $\omega_{qc:n}(G)$ efficiently to within arbitrary precision.

We will loosely follow the presentation of Watrous ([**95**]). We begin by defining $C_{qc:1}$: from there, defining the other $C_{qc:n}$ will be straightforward. $C_{qc:1}$ will be defined by imposing constraints on correlations that would necessarily have to be obeyed by a correlation that is actually in C_{qc}: this will give us that $C_{qc} \subseteq C_{qc:1}$. Suppose that we are given a correlation $p(a,b|x,y)$ that is in C_{qc}, so that, for some infinite-dimensional Hilbert space $\mathscr{H}$, there exists a state $|\psi\rangle \in \mathscr{H}$ and projective measurements $\{A^x_a\}$, $\{B^y_b\}$ for Alice and Bob, respectively, so that for all $x \in X, y \in Y, a \in A, b \in B$, $[A^x_a, B^y_b] = 0$ and:

$$p(a,b|x,y) = \langle \psi | A^x_a B^y_b | \psi \rangle$$

Writing $\tau(\cdot) = \langle \psi | \cdot \psi \rangle$ as a shorthand, we see that τ satisfies the following properties, where $E, E_1, \ldots, E_n$ and F stand for arbitrary operators:

(1)

$$\tau(\mathbb{I}) = 1$$

[9] We note that this is in principle. In practice, unless the dimensions of the inputs and the outputs are quite small, even the optimization problem corresponding to $n = 2$ is essentially intractable because of memory limitations.

(2)

$$\sum_{a\in A} \tau(EA^x_aF) = \tau(EF),\ \forall\, x \in X$$

(3)

$$\sum_{b\in B} \tau(EB^y_bF) = \tau(EF),\ \forall\, y \in Y$$

(4)

$$\tau(EA^x_aA^x_{a'}F) = \delta_{a,a'}\tau(EA^x_aF),\ \forall\, x \in X, a,a' \in A$$

(5)

$$\tau(EB^y_bB^y_{b'}F) = \delta_{b,b'}\tau(EB^y_bF),\ \forall\, y \in Y, b,b' \in B$$

(6)

$$\tau(EA^x_aB^y_bF) = \tau(EB^y_bA^x_aF)\ \forall\, x \in X, y \in Y, a \in A, b \in B$$

(7)

$$\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i\alpha_j\tau(E_i^\dagger E_j) \geq 0,\ \forall\ \alpha_1,\ldots,\alpha_n \in \mathbb{R}$$

Only the last property merits explanation, and can be seen to follow the fact that, setting $E = \sum_{i=1}^n \alpha_i E_i$,

$$\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i\alpha_j\tau(E_i^\dagger E_j) = \|E\,|\psi\rangle\,\|^2 \geq 0$$

Now, taking the $\{\tilde{A}^x_a\}$ and $\{\tilde{B}^y_b\}$ to be placeholders, letting Σ be the alphabet formed by these symbols and taking $\Sigma^{\leq n}$ be the collection of all words over Σ of length at most n, we take a pseudo-state $\tilde{\tau}$ to be a function $\tilde{\tau} : \Sigma^{\leq n} \to \mathbb{R}$. The tildes are there to emphasise that we think of these as mimicking genuine projective measurements and states. We then impose constraints on $\tilde{\tau}$ which would necessarily be satisfied if the $\tilde{A}^x_a$, $\tilde{B}^y_b$ were genuine measurement operators and $\tilde{\tau}$ was a genuine state acting on operators formed by monomials in the $\tilde{A}^x_a$ and $\tilde{B}^y_b$ of size at most two, based on the properties listed above. Explicitly, taking $E_1,\ldots,E_n$ to be an enumeration of Σ, these properties are:

(1)

$$\tilde{\tau}(\varepsilon) = 1$$

(2)

$$\sum_{a\in A} \tilde{\tau}(\tilde{A}^x_a) = 1,\ \forall\, x \in X$$

(3)

$$\sum_{b\in B} \tilde{\tau}(\tilde{B}^y_b) = 1,\ \forall\, y \in Y$$

(4)

$$\tilde{\tau}(\tilde{A}^x_a\tilde{A}^x_{a'}) = \delta_{a,a'}\tilde{\tau}(\tilde{A}^x_a),\ \forall\, x \in X, a,a' \in A$$

(5)

$$\tilde{\tau}(\tilde{B}^y_b \tilde{B}^y_{b'}) = \delta_{b,b'} \tilde{\tau}(\tilde{B}^y_b),\ \forall\, y \in Y, b, b' \in B$$

(6)

$$\sum_{a \in A} \tilde{\tau}(\tilde{A}^x_a \tilde{B}^y_b) = \tilde{\tau}(\tilde{B}^y_b),\ \forall\, x \in X, y \in Y, b \in B$$

(7)

$$\sum_{b \in B} \tilde{\tau}(\tilde{A}^x_a \tilde{B}^y_b) = \tilde{\tau}(\tilde{A}^x_a),\ \forall\, x \in X, y \in Y, a \in A$$

(8)

$$\tilde{\tau}(\tilde{A}^x_a \tilde{B}^y_b) = \tilde{\tau}(\tilde{B}^y_b \tilde{A}^x_a) \geq 0,\ \forall\, x \in X, y \in Y, a \in A, b \in B$$

(9)

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j \tilde{\tau}(E_i E_j) \geq 0,\ \forall\, \alpha_1, \ldots, \alpha_n \in \mathbb{R}$$

We define $C_{qc:1}$ to be the set of correlations $p(a,b|x,y)$ such that, for some pseudo-state $\tilde{\tau}$ satisfying all the above properties, it holds that:

$$p(a,b|x,y) = \tilde{\tau}(\tilde{A}^x_a \tilde{B}^y_b)$$

It is not hard to see that $C_{qc:1} \subseteq C_{ns}$: this is enforced by properties 6 and 7 above. Note also that $C_{qc:1}$ is a convex set because a convex combination of pseudo-states is also a pseudo-state, and also that it is closed. Also, since all the properties above are linear constraints on $\tilde{\tau}$ save the last one, which is a positive-semidefinite constraint, we see that the space of pseudo-states can be written as a rather small semidefinite program (featuring a positive semidefinite matrix of size $|X||A| + |Y||B| + 1$), so that $C_{qc:1}$ is a manageable set from a computational standpoint. We can think of $C_{qc:1}$ as a much better relaxation of the set C_{qf} than C_{ns} is: indeed, while C_{ns} and C_{qf} generally look very different, $C_{qc:1}$ is often found to be a reasonable approximation of C_{qf}.

The same game can be played in the same way for any $n \geq 1$, by taking $\tilde{\tau}$ to be a function $\Sigma^{\leq 2n} \to \mathbb{R}$ and deriving linear constraints for $\tilde{\tau}$ from the general properties of any legitimate state that were given previously. For each n, we get a closed correlation set $C_{qc:n}$, and it is immediate that it is contained in $C_{qc:n-1}$, for given a correlation in $C_{qc:n}$, restricting the corresponding pseudo-state $\tilde{\tau}$ to monomials of length at most $2n-2$ will yield a pseudo-state that will certify membership in $C_{qc:n-1}$. Furthermore, it is proven in [**95**] that:

$$C_{qc} = \bigcap_{n=1}^{\infty} C_{qc:n}$$

The proof also shows as a byproduct that the underlying Hilbert space of any commuting operators strategy may be assumed to be separable (i.e. have a basis of countable dimension).

In summary, we have that, given a nonlocal game G, for all n,

$$\omega_{qf:n}(G) \leq \omega^*(G) \leq \omega_{qc}(G) \leq \omega_{qc:n}(G)$$

And that, in the limit of n tending to infinity, $\omega_{qf:n}(G)$ goes to $\omega^*(G)$ and $\omega_{qc:n}(G)$ goes to $\omega_{qc}(G)$. We know that $C_{\overline{q}} \subseteq C_{qc}$, and if it held that this inclusion weren't proper, it would follow that $\omega^*(G)$ can be approximated arbitrarily well (in principle) by a computer program, which would calculate $\omega_{qf:n}(G)$ and $\omega_{qc:n}(G)$ for every n until it is found that $\omega_{qc:n}(G) - \omega_{qf:n}(G) < \varepsilon$, for some predetermined precision $\varepsilon > 0$. While we wouldn't know in advance how long it would take until such a n was found (implying that this program could be extremely inefficient), we would know that such a n does exist and hence that the program would eventually terminate.

2.5. Separations between the entanglement models and computability issues

We now address the problem of comparing the various models of entanglement that have been discussed thus far. This question turns out to be closely related to the computability of the various properties of nonlocal games that we have introduced. In particular, one may wonder about the hardness of the following decision problems for a given model t and a given game G:

(1) Does G have a perfect strategy?
(2) Given some fixed $0 < \varepsilon < 1$, under the promise that G either has a perfect strategy or value at most ε, is the former statement true?
(3) Does G have value one?

For $t \in \{c, \overline{q}, qc\}$, by compactness, problems 1 and 3 are equivalent. Classically, problem (1) is in NP because it is possible to verify the correctness of a perfect deterministic strategy in polynomial time, and its NP-hardness follows from the fact that it is possible to reduce 3-COL to it, as will be discussed in section 2 of chapter 4. In fact, the hardness-of-approximation point of view of the PCP theorem ([**15**]) implies that the second problem is also NP-complete classically.

On the quantum side of things, in the finite-dimensional setting, the first two problems can be seen to be in the class of recursively enumerable languages, denoted by RE, as they are both decidable if an upper bound is imposed on the local dimension of the shared entangled state, so that one may bruteforce every possible dimension until the existence of a perfect strategy is detected. Similarly, in the commuting operators model, the existence of the sum-of-squares hierarchy implies that the first two problems are in coRE.

The first to get the ball rolling was Slofstra ([**82**]), who showed the following [10]:

[10] Actually, Slofstra first showed in [**92**] that $C_{qs} \subsetneq C_{qc}$. The following theorem is a strengthening of this.

Theorem 2.5.1. *There is an explicit game G with $\omega^*(G) = 1$ but with no perfect tensor-product strategy. This game has input sets of size 184 and 235 and output sets of size 8 and 2. In particular, the tensor-product correlation sets are not closed, i.e. $C_{qf}, C_{qs} \subsetneq C_{\overline{q}}$ if the input and output sets are taken to be large enough.* [11]

This already refuted a strong form of Tsirelson's problem by showing that $C_{qf} \subsetneq C_{qc}$. Another separation in this vein was shown by Coladangelo and Stark, who could show (with a much shorter proof than Slofstra's):

Theorem 2.5.2. *We have that $C_{qf} \subsetneq C_{qs}$. In fact, there is an explicit correlation in C_{qs} with input sets of size 4 and 5 and output sets of size 3 that is not in C_{qf}. Since $C_{qs} \subseteq C_{\overline{q}}$, this gives another proof of the fact that C_{qf} is not closed.*

There remained to determine whether $C_{\overline{q}} = C_{qc}$: this was known as the weak Tsirelson's problem.

In addition to the aforementioned separation, Slofstra could also partially settle problem (1) for both quantum models:

Theorem 2.5.3 ([**92**], [**82**])**.** *For $t = qf, qs$, question (1) is undecidable. Also, for $t = qc$, problem (1) is coRE-complete, and hence also undecidable. Also, problem (3) is undecidable for $t = qf$.*

Note however that the above result has no direct bearing on the complexity of problem (2) in the quantum models, and it is unknown whether the exact complexity of problem (1) for $t = qf$ can be pinned down using Slofstra's methods. In the tensor-product setting, this was settled by Ji, Natarajan, Vidick, Wright and Yuen ([**93**]), who, in so doing, administered the final blow to Tsirelson's conjecture. A nonlocal game G is said to be *synchronous* if $X = Y$, $A = B$ and if, when presented with the same inputs, the players win if and only if their outputs are the same. They could show:

Theorem 2.5.4 ([**93**])**.** *Given any fixed $0 < \varepsilon < 1$, there exists an efficiently computable reduction from Turing machines M to (extremely large) synchronous nonlocal games G_M such that, if M*

[11] Slofstra's proof of this result is notoriously difficult. In the way of alternative proofs, there is the result of Coladangelo and Stark that we mention next: however, their separation is not framed as a nonlocal game. There is also the reasonably short proof of Dykema, Paulsen and Prakash ([**87**]) which also yields a small correlation in $C_{\overline{q}}$ that is not in C_{qf} and which also has very small input and output sets (of size 5 and 2, respectively), but their separation isn't quite framed as a nonlocal game either (although it's close enough). To the best of our knowledge, the only other paper which recovers Slofstra's result in full is the one by Mousavi, Nezhadi and Yuen ([**97**]) that is discussed below, but their proof can hardly be regarded as much more accessible than Slofstra's original proof, in addition to not being explicit. We should also mention that the non-closure of quantum correlations isn't a particularily exotic phenomenon: we expect a generic nonlocal game G with no perfect finite-dimensional strategy not to have a finite-dimensional strategy whose winning probability exactly attains $\omega^*(G)$, so this is a mathematical proof of something that was suspected to be true from numerical evidence.

halts, then G_M has a perfect finite-dimensional strategy, and $\omega^*(G_M) < \varepsilon$ otherwise. In particular, MIP=RE. As a corollary, for $t = qf$, problem (2) is RE-complete, and therefore undecidable.*

It follows that $C_{\overline{q}} \subsetneq C_{qc}$, because, as was discussed in the previous section, $C_{\overline{q}} = C_{qc}$ would imply the decidability of problem (2) for the tensor-product setting. Since problem (2) is no harder than problem (1) and since problem (1) is known to be contained in RE for $t = qf$, it also follows that problem (1) is RE-complete for $t = qf$.

Building upon the techniques introduced by [**93**], Mousavi, Nezhadi and Yuen ([**97**]) could show that problem (3) is even harder than the complexity class RE for $t = qf$, being complete for a higher level of the arithmetic hierarchy:

Theorem 2.5.5. *For $t = qf$, problem (3) is complete for Π_2.*

The most important remaining open problem in this area of study is the complexity of problem (2) in the commuting operators model. The MIP*=RE paper conjectured:

Conjecture 2.5.6. *Given any $0 < \varepsilon < 1$, there exists an efficiently computable reduction from Turing machines M to synchronous nonlocal games G_M such that, if M does not halt, then G_M has a perfect commuting operators strategy, while if it does hat, then $\omega_{qc}(G_M) < \varepsilon$. In particular, MIP^{co} = coRE. As a corollary, in the commuting operators model, problem (2) is complete for coRE.*

Finally, we note that the complexity of problem (1) for $t = qs$ is also unresolved. While Slofstra's results also show that problem (1) is undecidable for $t = qs$, it is not clear that it is contained in RE, or, for that matter, in any level of the arithmetic hierarchy.

Chapter 3

Communication complexity

Communication complexity concerns itself with the problem of determining how much communication is required between two (or more) parties in order to compute a given function of their inputs. This chapter discusses the three models that will be under consideration in this thesis, namely the classical model, the Yao model and the Cleve-Buhrman model.

3.1. Classical communication complexity

3.1.1. The model

Let X and Y be finite input sets (typically taken to be $X = Y = \{0,1\}^n$ for some n) and fix some boolean function $f : X \times Y \to \{0,1,\perp\}$. In an instance of the communication problem, the two parties, traditionally named Alice and Bob, are given $x \in X$, $y \in Y$, respectively, with the promise that $f(x,y) \neq \perp$, and they must communicate as few classical bits as possible in order to compute $f(x,y)$. The strategy that Alice and Bob are employing will be referred to as a **protocol**. One simple such protocol would be for Alice to simply send her input to Bob, from which Bob can compute the value of f, for a communication cost of $\lceil \log_2 |X| \rceil$. It should be stressed that unlike in complexity theory, no attention is paid to computational efficiency in communication complexity: we assume that Alice and Bob are all-powerful computationally and are only interested in the amount of exchanged communication.

A more formal picture of the above intuitive scenario is as follows. A protocol will be seen as being specified by a directed binary tree T with the following properties:

(1) Each node is labelled either as an Alice node or as a Bob node.
(2) Each non-leaf node in the tree has exactly two children, with one outgoing arc labelled as 0 and the other labelled as 1.

(3) To every Alice node v, there is an associated transition function $f_v(x) : X \to [0,1]$. Likewise, to every Bob node v, there is an associated transition function $f_v(y) : Y \to [0,1]$.

Operationally, we have the following scenario in mind. At each step of the protocol, the current state of the protocol is encoded by a node v in T, which is the root node initially. At every turn, whoever owns the current node v samples a bit b in such a way that the probability that one is obtained is the value of the transition function. If v is a non-leaf node, bit b is sent to the other party and the protocol transitions to the child of v corresponding to the bit b. If v is a leaf node, the protocol's answer for the value of $f(x,y)$ is b and the protocol halts. The **communication cost** of the protocol is measured by the depth of the tree, which essentially corresponds to the amount of communication of the protocol in the worst case [1].

It is necessary to specify exactly what it means for a protocol to compute f. This thesis will concern itself with the two most common meanings in the literature:

- The **exact communication complexity** of f: this corresponds to requiring that the protocol computes $f(x,y)$ with certainty. The smallest amount of communication for which this is possible will be denoted by $C_E(f)$. In this case, it may be seen that the protocol can be assumed to be deterministic, i.e. that the transition functions are all integer-valued.
- The **bounded-error communication complexity** of f: this corresponds to requiring that for some choice of $0 \leq \varepsilon < \frac{1}{2}$, for all legal inputs, the success probability of the protocol be at least $1-\varepsilon$. The smallest amount of communication for which this is possible will be denoted by $C_\varepsilon(f)$. As in complexity theory, it is customary to focus on the case $\varepsilon = \frac{1}{3}$.

Also, by analogy with the corresponding definition for nonlocal games, for a given amount of communication C, we will write $\omega_c(f,C)$ to mean the highest possible success probability in the worst case of a protocol for f with communication C.

An important extension of the standard communication model that we just described is the so-called **public-coin model**. Informally, a public coin protocol is one in which Alice and Bob are allowed to share a random variable with the distribution of their choice. Perhaps surprisingly, this can actually help reduce the amount of required communication in some cases, as we will see in the next subsection. By contrast, we will refer to protocols of the form given previously as private coin protocols. Formally, given a communication cost C and letting $\{\Pi_k\}$ be the set of all deterministic protocols for f with communication cost C, a **public coin protocol** for f with communication C is specified by a probability distribution p_k over the $\{\Pi_k\}$. At the outset, one such protocol is sampled independently of Alice and Bob's input according to the distribution

[1] Modulo trivial cases in which some nodes are unreachable.

p_k and is given to them, which they then proceed to run. It may be seen that this encompasses all private coin protocols, as the private randomness involved there can be packed inside the probability distribution p_k. We will denote the presence of a public coin by the superscript 'pub', so that we will write $\omega_c^{\text{pub}}(f,C)$ to denote the value of f with communication C in the public-coin model, and $C_\varepsilon^{\text{pub}}(f)$ to denote the bounded-error complexity of f in the public-coin model. We will also sometimes speak of the **public coin cost** of the protocol: this corresponds to the size of the support of the distribution p_k when measured in bits, i.e. $\lceil \log_2 \#\{k \mid p_k \neq 0\} \rceil$. Clearly, a public coin protocol for f with communication cost C and public coin cost P can be turned into an equivalent protocol with communication cost $C+P$ in which one party samples the public coin, sends it to the other and they carry on with the rest of the original protocol.

We will also often be interested in the context in which only Alice is allowed to communicate to Bob, which we will call the **one-way setting**. In this scenario, we will add a '1' in superscript to our notations to denote this restriction.

3.1.2. The equality function in the various settings

Perhaps the most prototypical problem in communication complexity is the equality function EQ_n, defined to be a function whose inputs are bit strings of length n and such that $EQ_n(x,y) = I[x=y]$. In this subsection, we will look at the communication complexity of the equality function in various models. Despite (or perhaps thanks to) its apparent simplicity, it turns out that this function is of substantial importance to the theory of communication complexity. Though the results we are about to discuss are very well-known, we interleave them with apparently new observations, mainly so as not to bore the knowledgeable reader (or ourselves).

3.1.2.1. The exact complexity of EQ_n. We begin by studying the complexity of EQ_n in the exact setting. The trivial protocol yields that $C_E(EQ_n) \leq n$, and it so happens that the trivial protocol is in fact optimal in this case. There are several ways to show this, but the one that we find simplest is, we believe, new:

Theorem 3.1.1.

$$C_E(EQ_n) = n$$

PROOF. We proceed by induction. The statement clearly holds for $n=1$ by the impossibility of telepathy. Suppose now that $n \geq 2$ and that the statement holds for $n-1$. Take an exact protocol for EQ_n with communication $m \geq 1$. As mentioned previously, we may assume without loss of generality that the protocol is deterministic, and, by symmetry, we can assume that the root node is an Alice node. For $b \in \{0,1\}$, let $X_b \subseteq \{0,1\}^n$ be the set of inputs that cause Alice to send bit b to Bob in the first round of the protocol. Clearly, there must be at least one b such that $|X_b| \geq 2^{n-1}$.

Hence, fixing some injection $g : \{0,1\}^{n-1} \to X_b$, we can hijack the original protocol to obtain a protocol for EQ_{n-1} where both sides apply g to their input strings x,y and proceed to run the original protocol where the first round is omitted and where the first bit that was communicated is assumed to have been b. The cost of this new protocol is $m-1$, which, by the induction hypothesis, is no smaller than $n-1$. This concludes the proof. □

We give another proof of the previous result that is perhaps less pleasing but is also substantially more general. Given a total boolean function f, the **communication matrix** of f, denoted by M_f, is the $|X| \times |Y|$ matrix such that ${M_f}_{x,y} = f(x,y)$, and $\operatorname{rk} M_f$ refers to the linear-algebraic rank of M_f over the reals.

Theorem 3.1.2 (Lemma 1.28 of [**24**]). *Any total function f satisfies*

$$C_E(f) \geq \log_2 \operatorname{rk} M_f$$

PROOF. Take an exact protocol for f with communication C, which we can assume to be deterministic, and let $\gamma = (v_0, v_1, v_2, \ldots, v_k)$ be some path down the communication tree of the protocol, with v_0 being the root node and v_k being a leaf node. Letting M^{v_k} be the $|X| \times |Y|$ boolean matrix such that $M^{v_k}_{x,y} = 1$ if, given that Alice and Bob's inputs were x and y, the protocol goes down the path γ and outputs 1, we have that M^{v_k} is of rank one. Indeed, letting $u \in \{0,1\}^{|X|}$ be such that $u_x = 1$ if and only if, for every Alice non-leaf node v_l in γ, the transition corresponding to x does end up in node v_{l+1}, and if v_k is an Alice node, Alice does answer with '1' given that her input is x, and letting $w \in \{0,1\}^{|Y|}$ be defined in the same way but for Bob, we get that

$$M^{v_k} = u^T w$$

Now, since, by assumption,

$$M_f = \sum_{v \text{ a leaf node}} M^v$$

and the number of leaf nodes is at most 2^C, the conclusion follows from the subadditivity of the rank (i.e. $\operatorname{rk}(A+B) \leq \operatorname{rk}(A) + \operatorname{rk}(B)$ for all matrices A,B). □

Since the communication matrix of EQ_n is the $2^n \times 2^n$ identity matrix, which has rank 2^n, we obtain another proof of theorem 3.1.1.

3.1.2.2. The bounded-error complexity of EQ_n. Turning to the bounded-error complexity of EQ_n, perhaps surprisingly, we can show that we obtain a dramatic reduction in complexity compared to the exact case:

Theorem 3.1.3.

$$C_{1/3}(EQ_n) \leq 2\log_2 n + O(1)$$

PROOF. We give the protocol due to Rabin and Yao. By the theorem of number theory known as Bertrand's postulate [2] , there exists some prime $3n < p < 6n$. Given $z \in \{0,1\}^n$, define the polynomial $q_z(X) \in (\mathbb{Z}/p\mathbb{Z})[X]$ by:

$$q_z(X) = \sum_{i=0}^{n-1} z_{i+1} X^i$$

The protocol runs as follows. Given inputs $x, y \in \{0,1\}$, Alice picks a uniformly random t in $\mathbb{Z}/p\mathbb{Z}$ and sends it to Bob, along with $c = q_x(t)$. This has communication cost $2\lceil \log_2 n \rceil + O(1)$. Bob then answers 1 if $c = q_y(t)$ and 0 otherwise. Clearly, if $x = y$, this will succeed with probability 1. If, on the other hand, $x \neq y$, since $q_x - q_y$ has degree at most p and is not the zero polynomial, for t chosen uniformly at random, the probability that $q_x(t) = q_y(t)$ is at most $\frac{1}{3}$. This proves the correctness of the protocol. □

Although this is unnecessary, we remark that the Rabin-Yao protocol can be improved upon slightly. This can be done because the protocol features a computation of the equality function as a subroutine, which could be done more efficiently by using the Rabin-Yao protocol itself: instead of having Alice send $q_x(t)$ to Bob in the above protocol, Alice and Bob could run the (vanilla) Rabin-Yao protocol another time on $q_x(t)$ and $q_y(t)$, viewed as bit strings of length $\log_2(n) + O(1)$. Taking a larger value for p than in the protocol above so that the total error probability remains at most $\frac{1}{3}$, this shows that $C_{1/3}(EQ_n) \leq \log_2 n + O(\log_2 \log_2 n)$. In fact, by optimizing the way in which the bit strings are mapped to polynomials, the reader can see for himself that we can show that $C_{1/3}(EQ_n) \leq \log_2 n + O(\log_2 \log_2 \ldots \log_2 n)$, where the number of imbricated $\log_2$ terms can be made as large as one wishes at the expense of increasing the constant coefficient. This leads one to suspect that $C_{1/3}(EQ_n) \leq \log_2 n + O(1)$, and this is indeed the case, as will be shown later.

3.1.2.3. The public-coin complexity of EQ_n. It turns out that the complexity of EQ_n is reduced even more when we allow Alice and Bob to share a public coin, in that it becomes constant. In view of the discussion in the previous subsection, this is not very surprising, as having t be provided by a public coin instead of being sampled by Alice and sent to Bob in the tower of Rabin-Yao protocols already implies that $C^{\text{pub}}_{1/3}(EQ_n) \leq O(\log_2 \log_2 \ldots \log_2 n)$, where, again, the number of imbricated logarithms can be made to be as large as one wishes, making $C^{\text{pub}}_{1/3}(EQ_n)$ constant for all practical purposes. There is however a much simpler public-coin protocol with constant communication for the equality function, which we now present.

[2] This result was stated as a conjecture in 1845 by Bertrand and proved in 1850 by Chebyshev. Mildly intriguing is the use of the term 'postulate', which has the same meaning as 'axiom' and is therefore clearly inapplicable when referring to a theorem. The introduction of this terminology appears to have been perpetrated by Chebyshev himself, who referred to the result as the 'postulatum connu de Bertrand'. The reader is referred to `https://math.stackexchange.com/questions/1920672/why-is-bertrands-postulate-called-a-postulate` for a discussion of the history behind this rather unfortunate name.

Theorem 3.1.4.

$$C^{pub}_{1/3}(EQ_n) \leq 2$$

PROOF. The protocol operates as follows. The two parties begin by picking $r_1, r_2 \in_R \{0,1\}^n$ independently. Given her output $x \in \{0,1\}^n$, Alice computes $c_1 = r_1 \cdot x$, $c_2 = r_2 \cdot y$ (where the dot product is taken modulo 2), and sends $c_1 c_2$ to Bob. Bob then checks if $c_1 = r_1 \cdot y$ and $c_2 = r_2 \cdot y$, and answers 1 if this the case, and 0 otherwise. Clearly, if $x = y$, the protocol will always answer 1, while if $x \neq y$, by the so-called random subsum principle (see, for example, claim A.31 of [**63**]), we have that $r_1 \cdot x = r_2 \cdot y$ with probability $\frac{1}{2}$, so the probability that the protocol outputs 1 is $\frac{1}{4} < \frac{1}{3}$. □

Note that this protocol has a public coin cost of $2n$, unlike the previous protocol based on the Rabin-Yao protocol, which has a public coin cost of $\log_2 n + O(\log \log_2 n)$. Newman's theorem, which we present in the next section, shows that $\log_2 n + O(1)$ shared random bits are sufficient.

3.1.3. Largest possible gaps between the settings

It can be shown that the separations between the complexities of the equality function in the various considered models are in fact as large as they can be in an asymptotic sense for any function. With regards to the separation between the exact and bounded-error complexities, we have:

Theorem 3.1.5 ([**24**], lemma 3.8)**.**

$$C^1_E(f) \leq (C_{1/3}(f) + 2)2^{C_{1/3}(f)}$$

PROOF. Take an optimal protocol for f with communication C that achieves a success probability of $\frac{2}{3}$. It isn't hard to see that there are positive real numbers $a_{\gamma,x}$ and $b_{\gamma,y}$ such that for every path γ from the root to a leaf node and for every Alice and Bob inputs x and y, the probability that the protocol goes down the path γ is $a_{\gamma,x} b_{\gamma,y}$. This means that the probability that the protocol outputs 1 is:

$$\sum_{\gamma} a_{\gamma,x} b_{\gamma,y}$$

The idea of the proof is to have Alice send discretizations of all the $a_{\gamma,x}$ that are sufficiently accurate so as to allow Bob to evaluate the above probability to sufficient precision, which will allow him to determine $f(x,y)$ with certainty. Note that the number of distinct paths is at most 2^C. Because $b_{\gamma,y} \leq 1$ for all γ, y, it follows that the discretizations need only be precise to within, say, $2^{-(C+3)}$ for the result to be accurate to within $\frac{1}{8}$, which is accurate enough because $\frac{2}{3} - \frac{1}{8} > \frac{1}{2}$. This will require $C + 2$ bits for each path, which corresponds to at most $(C+2)2^C$ bits in total, as desired. □

In the case of the equality function, this shows that

$$C_{1/3}(EQ_n) \geq \log_2 n - O(\log_2 \log_2 n)$$

So that the protocols we gave earlier are not very far off from being optimal. Actually, it is possible to show a $\log_2 n + O(1)$ lower bound on $C_{1/3}(EQ_n)$ ([**102**]), and, as we are about to show, this is is optimal up to a constant.

With regards to the separation between the private- and public-coin complexities, we have the so-called Newman's theorem:

Theorem 3.1.6 ([**13**]: [**24**], theorem 3.14)**.** *Given a function f with input sets $X = Y = \{0,1\}^n$ and given any $\varepsilon > 0$, a public coin protocol for f with success probability p can be converted into another public coin protocol for f with the same communication cost that has success probability at least $p - \varepsilon$ and public coin cost $\lceil 1 + \log_2 n + 2\log_2 \frac{1}{\varepsilon} \rceil$.*

PROOF. Take a public coin protocol for f with success probability p of the form discussed in the first subsection. Given a deterministic protocol Π and inputs x,y, we will write $S(x,y,\Pi)$ to denote the boolean variable which equals one if Π does produces $f(x,y)$ on inputs x,y and zero otherwise. By hypothesis, we have that for every legal input pair x,y, if Π is distributed according to the distribution given by the public coin protocol, we have that

$$\mathbb{E}[F(x,y,\Pi)] \geq p$$

For some $m \geq 1$ to be fixed later, generate protocols $\Pi_1, \dots, \Pi_m$ independently at random according to this distribution. Chernoff's bound gives that for any legal input pair x,y:

$$P\left(\frac{1}{m}\sum_{l=1}^{m} F(x,y,\Pi_l) \leq p - \varepsilon\right) < \exp\left(-2\varepsilon^2 m\right)$$

Taking $m = \lceil \frac{n}{\varepsilon^2} \rceil$, we get:

$$P\left(\frac{1}{m}\sum_{l=1}^{m} F(x,y,\Pi_l) \leq p - \varepsilon\right) < \exp(-2n) < 2^{-2n}$$

Since there are at most 2^{2n} legal input pairs (x,y), the union bound implies that with nonzero probability, we have that for every legal input pair x,y,

$$\frac{1}{m}\sum_{l=1}^{m} F(x,y,\Pi_l) > p - \varepsilon$$

Letting $\Pi_1, \dots, \Pi_m$ be some definite choice of protocols satisfying the above, we see that the new public coin protocol for f which picks some $l \in [m]$ uniformly at random, runs Π_l and outputs the result has success probability at least $p - \varepsilon$ for all legal input pairs. This completes the proof. □

It follows that $C_{1/3}(EQ_n) = \log_2 n + O(1)$. More generally, we have:

Corollary 3.1.7. *Any function f with input sets $X = Y = \{0,1\}^n$ satisfies*

$$C_{1/3}(f) \leq O(C^{pub}_{1/3}(f)) + \log_2 n + O(1)$$

In particular, one shouldn't expect a public coin to ever be of very great use compared to private randomness, unless the inputs are enormously larger than the communication.

3.1.4. Distributional complexity and Yao's principle

Previously, we defined the bounded-error complexity $C_\varepsilon(f)$ to be the smallest amount of communication such that a protocol exists which as a success probability of at least $1-\varepsilon$ in the worst case. We now introduce another conceptualization of this, namely, the **distributional complexity** of f: given a probability distribution $\mu_{x,y}$ on legal inputs, C^μ_ε is defined to be the smallest C such that a deterministic protocol exists which has a success probability of at least $1-\varepsilon$ given that the inputs were sampled according to μ. Similarly, we define $\omega^\mu_c(f,C)$ to be the highest success probability of any deterministic protocol with communication C for the previous definition of success probability. We clearly have, for every C,

$$\omega_c(f,C) \leq \omega^\mu_c(f,C)$$

And therefore:

$$C^\mu_\varepsilon(f) \leq C_\varepsilon(f)$$

We show that more can be said about the relationship between the distributional and public-coin complexities of a given function f. Given a communication cost C, let $\{\Pi_k\}_k$ be the list of all possible deterministic classical protocols for f with communication C. We let $V(x,y,\Pi)$ be a binary variable which equals one if the output of the protocol Π on inputs x,y is equal to $f(x,y)$. We see that $\omega^{\text{pub}}_c(f,C)$ can be encoded as the optimal value of the following linear program, where we take $R \subseteq X \times Y$ to be the set of legal input pairs:

$$\max q \tag{3.1.1}$$

$$\text{s.t.} \sum_k p_k V(x,y,\Pi_k) \geq q \quad \forall (x,y) \in R \tag{3.1.2}$$

$$\sum_k p_k = 1 \tag{3.1.3}$$

$$p_k \geq 0 \qquad \forall k \tag{3.1.4}$$

In the above, q corresponds to the success probability of a public coin protocol, while p_k is the probability distribution over protocols which specifies the public coin protocol. The first set of constraints encodes the fact that for any legal inputs, the success probability of the protocol for those inputs must be at least q, which was our definition of $\omega_c^{\mathrm{pub}}(f,C)$. Note that by elementary linear programming theory, the above program has an optimal solution that is an extreme point. We therefore have:

Proposition 3.1.8. *Given a function f with $X = Y = \{0,1\}^n$, there is no loss in generality in restricting ourselves to public coin protocols with public coin cost at most $2n$.*

Note the contrast between the above result and Newman's theorem. Newman's theorem says that a public coin cost of $\log_2 n + O(1)$ suffices for a near-optimal protocol using a given amount of communication, while the above result says that a public coin cost of $2n$ suffices for an optimal protocol.

The dual of the previous linear program is the following linear program:

$$\min q' \tag{3.1.5}$$

$$\text{s.t.} \sum_{(x,y)\in R} \mu_{x,y} V(x,y,\pi_k) \le q' \qquad \forall\, k \tag{3.1.6}$$

$$\sum_{(x,y)\in R} \mu_{x,y} = 1 \tag{3.1.7}$$

$$\mu_{x,y} \ge 0 \qquad \forall\, (x,y) \in R \tag{3.1.8}$$

This linear program visibly encodes the least q'^* and the corresponding probability distribution $\mu_{x,y}$ such that, for all deterministic protocols, the success probability of that protocol under the distribution μ is at most q'^*. By the strong duality theorem of linear programming, $q^* = q'^*$. In notation, this means:

Theorem 3.1.9 (Yao's principle)**.** *For any function f and any amount of communication C,*

$$\omega_c^{pub}(f,C) = \min_{\mu} \omega_c^{\mu}(f,C)$$

In particular, for every ε,

$$C_{\varepsilon}^{pub}(f) = \max_{\mu} C_{\varepsilon}^{\mu}(f)$$

3.2. Quantum communication complexity: the Yao model

In this section, we describe a quantum version of the classical model of communication complexity that was described in the previous section. This model was introduced in 1993 by one of the pioneers of communication complexity itself, namely Yao ([**19**]), and hence will be called the **Yao model**. In this setting, the participants are permitted to communicate using quantum states, but are not allowed to share prior entanglement. While the model was introduced by Yao, much of its early theory was developed by Kremer in his master's thesis ([**21**]).

3.2.1. The model

Allowing the parties to communicate using qubits complicates matters somewhat compared to subsection 3.1.1, as it is no longer possible to speak of a communication tree. The most common approach in the literature is to force the parties to communicate in turns, irrespective of the history of the communication up till now, at the possible cost of weakening the model. Following Briët, Buhrman, Leung, Piovesian and Speelman ([**74**]), we will assume that protocols in this new model are of the following form: the participants start out with the system AB, with A and B both initially of dimension 1. At round k, the participant whose turn it is to communicate tacks on a register formed of ℓ_k qubits to his system, applies a unitary depending on his input to his system and sends the register that was tacked on to the other party. At the end of the day, the participant who just received something measures the content of their register using a POVM depending on their input, the result of which is the output of the protocol. The total communication cost of the protocol is the sum of the ℓ_k.

We may consider the analogues of the complexity measures that have been studied in the previous section in the Yao model. Given a function f and an amount of communication C, we will write, for example, $\omega_{yao}(f,C)$ to mean the highest success probability achievable by a protocol for f with communication cost at most C. In terms of quantifying the amount of communication required to achieve something, we will use the letter 'Q' instead of the letter 'C' to refer to the complexity of a function in the Yao model instead of in the classical model. For example, we will write $Q_{1/3}^{\text{pub}}(f)$ to denote the public coin communication complexity of f in the Yao model. Note that since we are forcing the participants to communicate in turns, it might or might not be the case that the complexity of a function in the Yao model is always less than its complexity in the classical setting. However, it is not difficult to see that the worsening cannot be by more than a factor of two, so that this issue may be regarded as not very significant.

Several of the results that were mentioned in the previous section go through unchanged in this new setting. We still have Newman's theorem:

Theorem 3.2.1. *Given a function f with input sets $X = Y = \{0,1\}^n$ and given any $\delta > 0$, a public-coin protocol for f in the Yao model with error probability ε can be converted into another public-coin protocol for f in the Yao model with the same communication cost that has error probability at most $\varepsilon + \delta$ and public coin cost $1 + \log_2 n + 2\log_2 \frac{1}{\delta}$.*

The proof of Yao's principle (theorem 3.1.9) still goes through, although some caution is required because the space of possible protocols is no longer finite, but its compactness saves the day:

Theorem 3.2.2. *For any function f and any amount of communication C,*

$$\omega_{yao}^{pub}(f,C) = \min_{\mu} \omega_{yao}^{\mu}(f,C)$$

In particular, for every ε,

$$Q_{\varepsilon}^{pub}(f) = \max_{\mu} Q_{\varepsilon}^{\mu}(f)$$

Also, an analogue of theorem 3.1.5 still holds in the Yao model, which was originally shown by Kremer.

Theorem 3.2.3 (Theorem 4 of [**21**], theorem 1.1 of [**74**])**.**

$$C_E^1(f) \leq (Q_{1/3}(f) + O(1))2^{2Q_{1/3}(f)}$$

This is a slightly worse bound than the one in 3.1.5, but as we will see, this worsening is unavoidable.

3.2.2. The equality function in the Yao model

We briefly discuss the various complexities of the equality function in the Yao model. First, in the exact setting, Buhrman and de Wolf showed that the rank lower bound for the exact communication complexity of total functions still holds in the Yao model:

Theorem 3.2.4 ([**35**])**.** *Any total function f satisfies*

$$Q_E(f) \geq \log_2 \mathrm{rk} M_f$$

From which we get that $Q_E(EQ_n) = n$ for all n.

In the case of bounded-error communication complexity, we have seen that classically, $C_{1/3}(EQ_n) = \log_2 n + O(1)$, while theorem 3.2.3 gives the weaker bound

$$Q_{1/3}(EQ_n) \geq \frac{1}{2}\log_2 n + O(\log_2 \log_2 n)$$

This opens the door to a potential factor two savings compared to the classical case. As it turns out, it is in fact the case that $Q_{1/3}(EQ_n) = \frac{1}{2}\log_2 n + O(1)$ ([**102**]). This will follow from the results of subsection 3.3.2.

Finally, we of course have that $Q_{1/3}^{\text{pub}}(EQ_n) \leq C_{1/3}^{\text{pub}}(EQ_n) = O(1)$.

3.2.3. Separations between the classical and Yao models

In this subsection, we will look at some problems which are known to exhibit gaps between the Yao and the classical models. The earliest examples of such gaps came from the so called **BCW simulation** of Buhrman, Cleve and Wigderson ([**29**]), which allows one to turn quantum algorithms into Yao protocols for communication complexity problems. We describe a very slightly restricted version which will be sufficient for our purposes.

Theorem 3.2.5 ([**29**])**.** *For $n \in \mathbb{N}$, suppose that the function $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ to be computed is of the form $f(x,y) = g(x \,\Box\, y)$, where $\Box \in \{\wedge, \vee, \oplus\}$ stands for bitwise application. If there exists an exact/bounded-error quantum algorithm for evaluating g using t oracle calls, there exists an exact/bounded-error Yao protocol for f with communication $O(t \log_2 n)$.*

PROOF. The idea of the simulation is that Alice will be running the quantum algorithm for g on her side, and each oracle call will be executed by communicating with Bob. At a given step of the algorithm for g, suppose that the oracle call $|i\rangle\,|b\rangle \mapsto |i\rangle\,|b \oplus (x \,\Box\, y)_i\rangle$ needs to be applied. This is achieved in the following way: Alice first tacks on a new binary register in the state $|0\rangle$, applies the transformation $|i\rangle\,|b\rangle\,|0\rangle \mapsto |i\rangle\,|b\rangle\,|x_i\rangle$ and sends the state to Bob. Bob applies the transformation $|i\rangle\,|b\rangle\,|b'\rangle \mapsto |i\rangle\,|b \oplus (y_i \Box b')\rangle\,|b'\rangle$ to what he received from Alice and sends the result back to her. Alice then applies the transformation $|i\rangle\,|b\rangle\,|b'\rangle \mapsto |i\rangle\,|b\rangle\,|b' \oplus x_i\rangle$ to what she received from Bob. It is apparent that this will have achieved the required oracle call, and the communication required was $2\lceil \log_2 n \rceil + 4$ qubits. □

Although the above scheme yields efficient Yao protocols for a number of communication problems, two in particular stand out:

(1) An efficient exact protocol for a distributed version of the Deutsch-Jozsa problem ([**16**]): in this problem, for some $n \in \mathbb{N}$, Alice and Bob are each given strings $x,y \in \{0,1\}^{2^n}$ with the promise that either $x = y$ or $|x \oplus y| = 2^{n-1}$, and they must determine which is the case with certainty. This corresponds to setting $\Box = \oplus$ and $g(x) = |x|$ in the above simulation; the Deutsch-Jozsa algorithm then yields an exact Yao protocol for this problem with communication cost $2n$, while a hard combinatorial result of Frankl and Rödl ([**11**]]) shows that the amount of communication required by any exact classical algorithm for this problem must be exponential in n. A cleaner exposition of this separation, reframed in the language of quantum graph theory, will be given in section 4.2.4. Note however that this separation collapses in the bounded-error setting, since the function to be computed is a subset of the equality function EQ_{2^n} and therefore has a bounded-error classical protocol with communication $O(n)$.

(2) An efficient bounded-error protocol for the disjointness function: for $n \in \mathbb{N}$, define the **disjointness** function $DISJ_n$ on binary strings of length n by $DISJ_n(x,y) = 1$ if and only if, for some i, $x_i = 1$ and $y_i = 1$. This corresponds to setting $\square = \wedge$ and $g(x) = I[x \neq 0]$ in the above. By using a slightly modified version of Grover's algorithm ([**25**]), we get a bounded-error protocol for $DISJ_n$ with communication $O(\sqrt{n}\log n)$. This is a sizeable improvement over the classical case, where it was shown by Kalyanasundaram and Schnitger ([**12**]) that $C_{1/3}(DISJ_n) = \Theta(n)$. This was later improved slightly by Aaronson and Ambainis ([**49**]), who gave a bounded-error Yao protocol for $DISJ_n$ with communication $O(\sqrt{n})$, again based on Grover's algorithm. This matches the quantum lower bound of $\Omega(\sqrt{n})$ shown by Razborov ([**41**]).

While the disjointness function exhibits a quadratic separation between the Yao and classical models, this remains a far cry from the upper bound imposed by theorem 3.2.3, which leaves the door open to an exponential separation between the two. Another drawback of this separation is that the corresponding Yao protocols requires a large number of rounds. The first exponential separation between the Yao and the classical settings in the bounded-error case was given by the so-called **hidden matching problem** of Bar-Yossef, Jayram and Kerenidis [**45**]. While originally framed as a relational problem, this separation was soon afterwards recast into functional problems (with promise) independently by Gavinsky, Kempe and de Wolf ([**51**]) and by Kerenidis and Raz ([**55**]). While the two problems are similar, we will follow the presentation of [**51**]. It should be emphasized that we are only considering protocols in which Alice is communicating to Bob.

Let $n \in \mathbb{N}$ and take a real-valued parameter $\alpha \leq \frac{1}{2}$. Alice's input is a string $x \in \{0,1\}^n$, while Bob's input is a disjoint collection of pairs of elements of $\{1,\ldots,n\}$ $(i_1,j_1),(i_2,j_2),\ldots,(i_k,j_k)$, where $k = \lceil \alpha n \rceil$, as well as a string $u \in \{0,1\}^k$. The promise that Alice and Bob are given is that the value of $x_{i_l} \oplus x_{j_l} \oplus u_l$ is the same for every $\ell \in [k]$, say b, and their task is to compute b. In the Yao model, this can be accomplished with $\lceil \log_2 n \rceil$ qubits of communication using the following protocol:

(1) Alice sends Bob the state

$$\frac{1}{2^{\frac{n}{2}}}\sum_i (-1)^{x_i}|i\rangle$$

(2) Bob measures what he received from Alice using the projective measurement $\{P_l\}_{l\in[k+1]}$, where, for $l \in [k]$, P_l is the projector onto the subspace generated by $|i_l\rangle$ and $|j_l\rangle$ and where P_{k+1} is the projector onto the rest of the space.

(3) If outcome $k+1$ was obtained, which happens with probability roughly $1-2\alpha$, Bob declares his ignorance. Otherwise, if outcome $l \in [k]$ was obtained, the residual state is:

$$\frac{1}{\sqrt{2}}(-1)^{x_{i_l}}|i_l\rangle + \frac{1}{\sqrt{2}}(-1)^{x_{j_l}}|j_l\rangle$$

From here, Bob can recover the value of $x_{i_l} \oplus x_{j_l}$ (as the two possible values correspond to orthogonal states), from which he can recover b thanks to his knowledge of u_l.

It can be seen that the protocol's success probability is 2α, with a probability of declaring ignorance of $1-2\alpha$. This means that a success probability of $\frac{2}{3}$ can be achieved by running this protocol $O(\frac{1}{\alpha})$ times, which requires $O(\frac{\log_2 n}{\alpha})$ qubits of communication.

Classically, a $O(\sqrt{n})$ protocol with communication from Alice to Bob may be built by appealing to the birthday paradox. For this, we will allow Alice and Bob to share a public coin: Newman's theorem shows that the public coin might be removed at the cost of $O(\log n)$ bits of extra communication, which does not impact the complexity of the protocol. Alice and Bob pick a subset S of size $\frac{\sqrt{n}}{\alpha}$ of $[n]$, and Alice will send the values of x corresponding to these indices to Bob. The birthday paradox implies that with good probability, for some l, i_l and j_l will both be in S, which will let Bob recover the value of b. [**51**] show that if α is taken to be $\Theta\left(\frac{1}{\log n}\right)$, this protocol is asymptotically optimal classically when communication is restricted to flow from Alice to Bob, while there is a bounded-error quantum protocol with $O((\log_2 n)^2)$ qubits of communication from Alice to Bob.

It is however not hard to see that the above separation collapses if bidirectional communication is permitted, and one would ideally prefer to have a separation that is achieved by an exact Yao protocol. These two drawbacks were lifted by Klartag and Regev ([**65**]), who gave an example of a functional promise problem with an exponential gap between its exact, one-way Yao complexity and its bidirectional bounded-error complexity, and is hence essentially the strongest separation between the Yao and classical models that can be hoped for. This separation is based on the **vector in subspace problem**, which was first introduced by Kremer ([**21**]) and is defined as follows: for some even $n \in \mathbb{N}$, Alice's input is a unit vector $|\psi\rangle \in \mathbb{C}^n$, Bob's is a subspace V of $\mathbb{C}^n$ of dimension $\frac{n}{2}$, and under the promise that either $|\psi\rangle \in V$ or $|\psi\rangle \in V^{\perp}$, the problem is to determine which of the two is the case. It is quite straightforward to see that this can be done exactly in the Yao model with $O(\log n)$ qubits of communication, as Alice can send $|\psi\rangle$ to Bob and Bob can measure what he received according to the measurement $\{P_V, \mathbb{I}-P_V\}$ to determine which is the case with certainty, at the cost of $\lceil \log_2 n \rceil$ qubits of communication. [**65**] use sophisticated techniques from analysis and probability theory to show a $\Omega(n^{\frac{1}{3}})$ bound on the classical bounded-error communication complexity of the problem, with no restriction on the communication being one-way.

One should note that the above two exponential separations between the Yao and classical models are both based on promise problems and may therefore be regarded as artificial, with the second in particular being as tailored to the strengths of quantum computing as a quantum supremacy experiment. What the largest possible gap between the complexities of a total function in the Yao and classical models is remains unknown. For some time, the total function exhibiting the largest known gap between classical and quantum communication complexity was the disjointness function. The first superquadratic gap for total functions was shown by Anshu, Belovs, Ben-David, Göös, Jain, Kothari, Lee and Santha ([**80**]), who gave a family of total functions f_n with $C_{1/3}(f_n) = \Omega(Q_{1/3}(f_n)^{2.5})$. One might also wonder whether exact quantum communication can ever be stronger than classical communication for a total function. This was shown to be the case by Ambainis ([**72**]), who proved that there exists a family of total functions g_n for which $C_{1/3}(g_n) = \Omega(Q_E(g_n)^{1.15})$. This was slightly improved upon by [**80**], who showed the same result with exponent 1.5.

3.3. Entanglement-assisted communication complexity: the Cleve-Buhrman model

This section studies another plausible quantum analogue of the classical setting of communication complexity. This time, we will let the parties share an entangled state of their choice, while restricting the communication between them to be classical. This model was originally introduced by Cleve and Buhrman ([**26**]) and will be referred to from now on as the **Cleve-Buhrman** model.

Unlike the comparison between the quantum and classical communication models, which is comparatively well-understood, very little is known about how the Yao and Cleve-Buhrman models compare. By quantum teleportation ([**17**]), we know that a Yao protocol with communication n can be converted into a Cleve-Buhrman one with communication $2n$, so that the complexity of a function in the Cleve-Buhrman model is never more than twice its complexity in the Yao model. Since entanglement can be turned into a public coin through the act of measuring, we have that the bounded-error complexity of the equality function EQ_n in the entanglement-assisted model is constant, while we have seen that it is logarithmic in the private-coin quantum communication model. This means that in the absence of a public coin, we know that the Cleve-Buhrman model is stronger than the Yao model, although even that separation is not overly impressive. However, in the presence of a public coin, it is unknown if the two models are roughly equivalent.

We mention that there are in fact two commonly studied models of communication communication complexity which involve shared entanglement in the literature, namely the one in which the participants communicate using classical bits, which is the original proposal of Cleve and Buhrman, and the one in which the participants communicate using qubits, which is sometimes called the **hybrid model** and is arguably more popular. Clearly, the hybrid model is no weaker than the Cleve-Buhrman model, and the Cleve-Buhrman model can simulate the hybrid model with a factor two overhead in communication cost, again thanks to teleportation. While it can sometimes be more convenient to consider the hybrid model, in this thesis, we will restrict ourselves to the pure Cleve-Buhrman model with classical communication. There are two main reasons for this. First, we find that many results in the literature that are stated for the hybrid model admit simpler proofs in the Cleve-Buhrman model. The second reason is that taking the communication to be quantum would obscure several parallels that we wish to draw with the theory of nonlocal games: several of the communication problems that will be discussed in the next chapter appear to be impossibly hard to analyse in the hybrid model, while the analysis in the Cleve-Buhrman model is quite straightforward. As an added bonus, taking the communication to be classical allows us to explore a generalisation of the original Cleve-Buhrman model in which the parties are allowed to share entanglement in the commuting operators model: as far as we know, we are the first to study communication complexity in the presence of infinite dimensional entanglement. How this might be carried out in the hybrid model is not obvious to us.

3.3.1. The model

The communication model is mostly a straightforward adaptation of the classical model. Given a function $f : X \times Y \to \{0,1,\perp\}$, a protocol for f will be seen as being specified by a Hilbert space $\mathscr{H}$, a state $|\psi\rangle \in \mathscr{H}$ as well as by a directed binary tree T with the same properties as in the classical model. The twist is that instead of there being an associated transition function to every node, to every Alice node v and Alice output $x \in X$, there is an associated binary projective measurement $\{A_b^{v,x}\}_{b\in\{0,1\}}$, and to every Bob node v and Bob output $y \in Y$, there is an associated binary projective measurement $\{B_b^{v,y}\}_{b\in\{0,1\}}$. We further insist that any two Alice and Bob operators commute. At each step of the protocol, the party who currently owns the node performs a measurement on the shared state to determine their bit, which is either communicated to the other party of given as the protocol's answer depending on whether the node isn't or is a leaf node.

Take a path $\gamma = (v_0, v_1, v_2, \dots, v_k)$ down the tree, where v_0 is the root node and where v_k a leaf node, and let $(b_0, b_1, \dots, b_{k-1})$ be the corresponding communication transcript. Also, list the indices $i_1 < i_2 < \dots < i_{k_A}$ and $j_1 < j_2 < \dots < j_{k_B}$ corresponding to Alice's and Bob's nodes, respectively. Given a bit b, we see that if we set $b_k = b$ and define the positive operators

$$A_{\gamma,x,b} = A_{b_{i_1}}^{v_{i_1},x} A_{b_{i_2}}^{v_{i_2},x} \dots A_{b_{k_A}}^{v_{i_{k_A}},x} \dots A_{b_{i_2}}^{v_{i_2},x} A_{b_{i_1}}^{v_{i_1},x}$$

$$B_{\gamma,y,b} = B_{b_{j_1}}^{v_{j_1},y} B_{b_{j_2}}^{v_{j_2},y} \dots B_{b_{k_B}}^{v_{j_{k_B}},y} \dots B_{b_{j_2}}^{v_{j_2},y} B_{b_{j_1}}^{v_{j_1},y}$$

By commutativity and as per the Born rule, the probability that the protocol goes down γ and outputs b is given by

$$\langle\psi|A_{\gamma,x,b} B_{\gamma,y,b}|\psi\rangle$$

So that the probability that the protocol will output b is given by

$$\sum_{\gamma} \langle\psi|A_{\gamma,x,b} B_{\gamma,y,b}|\psi\rangle$$

Note that for each path, one of $A_{\gamma,x,b}$ and $B_{\gamma,y,b}$ does not actually depend on b, but making this explicit would only complicate the notation. Note also that:

$$\sum_{\gamma}(A_{\gamma,x,0}B_{\gamma,y,0}+A_{\gamma,x,1}B_{\gamma,y,1})=\mathbb{I} \tag{3.3.1}$$

If no further restrictions are set on $\mathscr{H}$, we obtain an analogue of the commuting operators model in communication complexity. Restricting $\mathscr{H}$ to be finite-dimensional and restricting the entanglement to be in tensor product form, we obtain an analogue of the finite-dimensional model[3]. In line with the notation we have been using in the previous chapters, we will write $\omega^*(f,C)$ to mean the supremum over the winning probabilities of tensor-product protocols for f with communication C, and we will write $\omega_{qc}(f,C)$ to mean the same thing in the commuting operators model. Also, for $t \in \{qf,qc\}$, we will write $C_E^t(f)$ to mean the least C such that f has an exact protocol in model t with C bits of communication; and, in the bounded-error case, for $\varepsilon \geq 0$, we will write $C_\varepsilon^t(f)$ to mean the least C such that $\omega_t(f,C) \geq 1-\varepsilon$.

A word of warning is in order here: in the finite-dimensional case, our definition of bounded-error communication complexity differs from the standard definition in the literature. Although this is never stated explicitly, $C_\varepsilon^*(f)$ is taken to mean the least C such that some protocol with communication C achieves error probability ε. What *our* definition says is rather that $C_\varepsilon^*(f)$ stands for the least C such that for every $\delta > 0$, some protocol with communication C achieves error $\varepsilon+\delta$. Those two are not the same, in general: in particular, C_0^{qf} (the **vanishing-error communication complexity**) and C_E^{qf} do not coincide. This can be seen as a communication complexity analogue of Slofstra's result that $C_{qf} \subset C_{\overline{q}}$, and will be explored in section 2 of chapter 4. It should however be stressed that these two complexity measures do coincide for all the other models under consideration in this thesis. We will use the term **zero-error communication complexity** to refer to the study of both the vanishing-error complexity and the exact complexity in the future.

The properties that are enjoyed by the commuting operators model in the case of correlation sets can all be shown to hold in communication complexity as well, and for the same reasons: given a tree of the form above, there is an analogue of the sum-of-squares hierarchy which converges from above to the best success probability of any protocol using that communication tree. Concretely, this means that there exists a computable sequence $p_1 \geq p_2 \geq \dots$ which converges to $\omega_{qc}(f,C)$ for any C, although this is clearly of very little use indeed in practice because the number of communication trees grows exponentially with C and because the resulting

[3] One might also consider the case where the entanglement is in tensor-product form but the space is infinite-dimensional, as was done for nonlocal games, obtaining the model *qs*: we will not do this.

semidefinite programs would be so large that they would be unsolvable in practice unless C and f's inputs are very small. It can also be shown that the commuting operators model reduces to the usual finite-dimensional model if we restrict $\mathscr{H}$ to be finite-dimensional, reflecting the analogous result in nonlocality theory. It is also possible to show that Yao's principle (3.1.9) still holds in the Cleve-Buhrman model for both entanglement models, although this requires some care in the case of the finite-dimensional model.

Regarding the finite-dimensional setting, it is natural to wonder about the type of entanglement that might be required by a protocol: as far as we know, this question was first asked by Gavinsky ([**52**]). In the case of general correlations, it is already known that to every entangled state there corresponds a correlation that can only be achieved using that state, up to local operations, which might lead one to think that there are certain communication problems for which sharing a maximally entangled state could be highly suboptimal as opposed to more general types of entanglement. This intuition was proven to be at least partially incorrect by Coudron and Harrow ([**90**]), who showed that maximal entanglement is sufficient for communication complexity, at least up to a multiplicative factor:

Theorem 3.3.1 ([**90**], restated)**.** *For any $\varepsilon > 0$, there exists a universal constant K such that any given entanglement-assisted protocol in model qf with communication $C \geq 30$ can be turned into an entanglement-assisted protocol using only a maximally entangled state which deviates from the previous protocol with probability at most ε on every input and which has communication at most*

$$K\left(\frac{C}{\varepsilon}+\frac{\log\frac{1}{\varepsilon}}{\varepsilon}\right)$$

Whether maximal entanglement is exactly sufficient for communication complexity, meaning that any protocol for a function f that uses a general entangled state can be replaced by a protocol for f with the same amount of communication which has the same success probability as the original protocol and which uses a maximally entangled state, is unknown at present.

One might wonder about a further generalisation of the above model, in which Alice and Bob could be permitted to share arbitrary nonsignalling correlations [4]. We will not study this generalisation, and for a very good reason, namely, that it causes communication complexity to become trivial, meaning that one can compute any function with a constant amount of communication. In the case that Alice and Bob are allowed to share arbitrary nonsignalling correlations, this is not very hard to establish, as the correlation which has the players' inputs x and y as inputs and produces uniformly random bits a,b as outputs such that $a \oplus b = f(x,y)$ is easily seen to be nonsignalling: the players can therefore compute f exactly with one bit of communication by sampling

[4] Allowing them to share signalling correlations would make the exact communication complexity of any function be zero bits, which would be a bit of a problem.

this correlation and having Alice send a to Bob. One ([**48**]) can also show that the above idea can be made resistant to noise: namely, any correlations which allow for winning the CHSH game with probability at least $\approx 90.8\%$ also allow for computing any boolean function with a constant amount of communication and with an arbitrarily small (but fixed) error probability. The reader is referred to [**104**] for some recent developments in this line of work.

3.3.2. The equality function in the Cleve-Buhrman model

We complete our survey of the equality function by looking at its complexity in the Cleve-Buhrman model. In the exact setting, it is not known that $C_E^{qc}(EQ_n) = n$, although this is known if the communication is restricted to be one-way: see lemma 4.2.19. We can however show that the rank lower bound does hold in the finite-dimensional setting, which implies that the finite-dimensional communication complexity of EQ_n is exactly n. This was originally shown by Buhrman and de Wolf [**35**]: our result is ever so slightly stronger than theirs because our communication model is somewhat more general (as they force Alice and Bob to communicate in turns), but not in any serious capacity. Our proof is essentially a recasting of theirs, which assumes that the communication is quantum, in the model with classical communication. The main point of presenting this proof is that while we could not make it work for the commuting operators model, we feel that trying to adapt the proof below remains the most promising way of trying to show that $C_E^{qc}(EQ_n) = n$.

We begin by a lemma:

Lemma 3.3.2. *Suppose that on inputs x,y, an entanglement-assisted protocol produces bit b with certainty. We have that*

$$\sum_{\gamma} A_{\gamma,x,b} B_{\gamma,y,b} \left|\psi\right\rangle = \left|\psi\right\rangle$$

PROOF. By hypothesis, we have that

$$\left\langle\psi\right| \left(\sum_{\gamma} A_{\gamma,x,\bar{b}} B_{\gamma,y,\bar{b}}\right) \left|\psi\right\rangle = 0$$

Since the operator in the above expression is a sum of positive operators and is therefore also positive, it follows that

$$\left(\sum_{\gamma} A_{x,\gamma,\bar{b}} B_{y,\gamma,\bar{b}}\right) \left|\psi\right\rangle = 0$$

The conclusion then follows from equation 3.3.1. □

We can now show that theorem 3.1.2 still holds for entanglement-assisted communication complexity in the qf model:

Theorem 3.3.3 ([35]). *Any total boolean function f satisfies*

$$C_E^{qf}(f) \geq \log_2 \mathrm{rk} M_f$$

PROOF. Take an exact entanglement-assisted protocol for f with communication C, and fix some $n \in \mathbb{N}$, which will represent a certain number of copies of the protocol. For Alice inputs $x_1, x_2, \ldots, x_n$, we define $|v_{x_1,x_2,\ldots,x_n}\rangle \in \mathbb{C}^{2^{nC}} \otimes H$ by:

$$|v_{x_1,x_2,\ldots,x_n}\rangle = \sum_{\gamma_1,\ldots,\gamma_n} |\gamma_1\rangle \otimes |\gamma_2\rangle \otimes \ldots \otimes |\gamma_n\rangle \otimes (A_{\gamma_n,x_n,1} \ldots A_{\gamma_2,x_2,1} \ldots A_{\gamma_1,x_1,1} |\psi\rangle)$$

Also, for Bob inputs $y_1, y_2, \ldots, y_n$, we define $|w_{y_1,y_2,\ldots,y_n}\rangle \in \mathbb{C}^{2^{nC}} \otimes H$ analogously by:

$$|w_{y_1,y_2,\ldots,y_n}\rangle = \sum_{\gamma_1,\ldots,\gamma_n} |\gamma_1\rangle \otimes |\gamma_2\rangle \otimes \ldots \otimes |\gamma_n\rangle \otimes (B_{\gamma_1,y_1,1} B_{\gamma_2,y_2,1} \ldots B_{\gamma_n,y_n,1} |\psi\rangle)$$

From the previous lemma, we see that:

$$\begin{aligned}
\langle v_{x_1,x_2,\ldots,x_n} | w_{y_1,y_2,\ldots,y_n}\rangle &= \sum_{\gamma_1,\gamma_2,\ldots,\gamma_n} \langle\psi| A_{\gamma_1,x_1,1} A_{\gamma_2,x_2,1} \ldots A_{\gamma_n,x_n,1} B_{\gamma_1,y_1,1} B_{\gamma_2,y_2,1} \ldots B_{\gamma_n,y_n,1} |\psi\rangle \\
&= \left(\sum_{\gamma_1,\gamma_2,\ldots,\gamma_{n-1}} \langle\psi| \left(A_{\gamma_1,x_1,1} A_{\gamma_2,x_2,1} \ldots B_{\gamma_1,y_1,1} B_{\gamma_2,y_2,1} \ldots \right) \left(\sum_{\gamma_n} A_{\gamma_n,x_n,1} B_{\gamma_n,y_n,1} |\psi\rangle \right) \\
&= f(x_n,y_n) \sum_{\gamma_1,\gamma_2,\ldots,\gamma_{n-1}} \langle\psi| A_{\gamma_1,x_1,1} A_{\gamma_2,x_2,1} \ldots A_{\gamma_{n-1},x_{n-1},1} B_{\gamma_1,y_1,1} B_{\gamma_2,y_2,1} \ldots B_{\gamma_{n-1},y_{n-1},1} |\psi\rangle \\
&= \ldots \\
&= \Pi_{i=1}^n f(x_i,y_i)
\end{aligned}$$

Taking the matrix V to have its rows composed of the $|v_{x_1,x_2,\ldots,x_n}\rangle$ and taking the matrix W to have its columns composed of the $|w_{y_1,y_2,\ldots,y_n}\rangle$, it follows from the previous calculation that $VW = {M_f}^{\otimes n}$. We therefore have that:

$$(\mathrm{rk} M_f)^n = \mathrm{rk} VW \leq \mathrm{rk} V \leq (2^{nC}) \dim H$$

Since the above equation holds for all n and since $\dim H$ is finite by assumption, it follows that $C \geq \log_2 \mathrm{rk} M_f$, as desired. □

Corollary 3.3.4. *We have that $C_E^{qf}(EQ_n) = n$ for all n.*

In the bounded-error setting, since, as mentioned previously, shared entanglement allows for the production of a public coin through measuring, we have that $C_{1/3}^{qf}(EQ_n) = O(1)$. However, we can say a bit more about the complexity of the equality function in the bounded error setting: namely, we can optimise the amount of shared entanglement in a protocol for the equality function. This is original work by the author. It follows from the previously shown fact that

$C_{1/3}(EQ_n) = \log_2 n + O(1)$ that the number of shared random bits in any correct public-coin classical protocol for EQ_n with constant communication must be at least $\log_2 n + O(1)$. In particular, any finite-dimensional protocol for EQ_n with error probability at most $\frac{1}{3}$ which uses entanglement exclusively as a source of shared randomness must use at least $\log_2 n + O(1)$ ebits. We now show that this can be improved upon: namely, we show that there exists an entanglement-assisted protocol which uses only $\frac{1}{2}\log_2 n + O(1)$ ebits. This is tight in view of the fact that $Q_{1/3}(EQ_n) = \frac{1}{2}\log_2 n + O(1)$. What follows is part of a paper written by the author jointly with Mande and de Wolf ([**102**]).

We will use the probabilistic method. In the following, $m \leq d$ are natural numbers to be determined later. We take the initial entangled state to be the maximally entangled state in $D = 2^d$ dimensions, i.e., d Bell states:

$$|\Psi_{AB}\rangle = \frac{1}{\sqrt{D}} \sum_{i\in\{0,1\}^d} |i\rangle_A |i\rangle_B .$$

For every $z \in \{0,1\}^n$, pick independently a Haar-random element $U_z = \{|\psi_{z,r}\rangle\}_{r\in\{0,1\}^d}$ of $SU(D)$. With respect to that choice, the protocol for EQ_n that we have in mind is as follows:

(1) Alice, on input $x \in \{0,1\}^n$, measures her part of $|\Psi\rangle$ in the basis U_x, obtaining $r^A \in \{0,1\}^d$. She then sends $b \equiv r^A_1 r^A_2 \dots r^A_m$ to Bob.

(2) Bob, on input $y \in \{0,1\}^n$, measures his part of $|\Psi\rangle$ in the conjugate basis of U_y, obtaining $r^B \in \{0,1\}^d$. He outputs 1 if $r^B_i = b_i$ for every $1 \leq i \leq m$ and 0 otherwise.

The one-way communication complexity of this protocol Π' is m bits. We proceed with its error analysis. We will show that provided that D and m were chosen properly, the resulting protocol is correct with nonzero probability: in particular, some choice of the U_z yields a correct protocol. At the end of step 1, the new joint state will be

$$|\Psi'\rangle = |\psi_{x,r^A}\rangle_A \otimes \overline{|\psi_{x,r^A}\rangle}_B$$

In particular, if $x = y$, then $r^A = r^B$ and the protocol is guaranteed to succeed. Suppose now that $x \neq y$. For $b \in \{0,1\}^m$, using the shorthand

$$R_b \equiv \{r \in \{0,1\}^d \mid r_i = b_i \; \forall i \in [m]\},$$

we find that the probability that the protocol fails (i.e., outputs 1) is given by

$$\frac{1}{D} \sum_{b\in\{0,1\}^m} \sum_{r^A, r^B \in R_b} |\langle\psi_{x,r^A}|\psi_{y,r^B}\rangle|^2.$$

Since R_b has cardinality 2^{d-m} and the expectation over the choice of U_z's of every term in the sum is 2^{-d}, we find that the expectation of the entire sum is 2^{-m}. We will now show that if d is taken to be large enough, with nonzero probability, this sum will be close to its expectation for every input pair, which will show that the protocol will have worst-case error probability at most $\approx 2^{-m}$. The rest of our analysis will rely on the following concentration inequality, which is derived in the third chapter of [**71**]:

Theorem 3.3.5 ([**71**])**.** *Let $F : SU(n) \to \mathbb{R}$ be a function with Lipschitz constant K with respect to the Frobenius norm, and let μ be the uniform distribution (Haar measure) on $SU(n)$. Then, for every $\delta > 0$,*

$$\Pr_{\mu}[|F(U) - \mathbb{E}_{\mu}[F]| > \delta] < 2\exp\left(-\frac{\delta^2 n}{4K^2}\right)$$

We show:

Theorem 3.3.6. *Let $\{\phi_r\}_{r\in\{0,1\}^d}$ be a fixed orthonormal basis of $\mathbb{C}^D$. Given $U = \{\psi_r\}_{r\in\{0,1\}^d} \in SU(D)$, define $F : SU(D) \to \mathbb{R}$ by*

$$F(U) = \sum_{b\in\{0,1\}^m} \sum_{r,r'\in R_b} |\langle\phi_r|\psi_{r'}\rangle|^2$$

Then $F(U)$ has Lipschitz constant $\sqrt{D}$.

Proof. Let $U = \{\psi_r\}_{r\in\{0,1\}^d}$ and $U' = \{\psi'_r\}_{r\in\{0,1\}^d}$ be two different elements of $SU(D)$. For $b \in \{0,1\}^m$, write

$$P_b = \sum_{r\in R_b} |\phi_r\rangle\langle\phi_r|, \quad Q_b = \sum_{r\in R_b} |\psi_r\rangle\langle\psi_r|, \quad Q'_b = \sum_{r\in R_b} |\psi'_r\rangle\langle\psi'_r|.$$

We see that

$$F(U) = \sum_{b\in\{0,1\}^m} \mathrm{tr}(P_b Q_b) \quad \text{and} \quad F(U') = \sum_{b\in\{0,1\}^m} \mathrm{tr}(P_b Q'_b)$$

Therefore

$$\begin{aligned} F(U) - F(U') &= \sum_{b\in\{0,1\}^m} \mathrm{tr}(P_b(Q_b - Q'_b)) \\ &\leq \sum_{b\in\{0,1\}^m} D_{\mathrm{tr}}(Q_b, Q'_b) \\ &\leq D \sum_{r\in\{0,1\}^d} \frac{1}{D}\sqrt{1 - |\langle\psi_r|\psi'_r\rangle|^2} \\ &\leq \sqrt{D^2 - \left(\sum_{r\in\{0,1\}^d} |\langle\psi_r|\psi'_r\rangle|\right)^2} \end{aligned}$$

Where the first inequality follows from the variational characterization of trace distance, the second from the convexity of trace distance, the fact that the R_b's partition $\{0,1\}^d$, and a well-known expression for the trace distance of two pure states, and the third inequality follows from the concavity of the function $\sqrt{1-z^2}$. On the other hand,

$$d(U,U') = \sqrt{\sum_{r\in\{0,1\}^d} \| \, |\psi_r\rangle - |\psi'_r\rangle \, \|^2} \geq \sqrt{2D - 2\sum_{r\in\{0,1\}^d} |\langle\psi_r\rangle \, \psi'_r|},$$

where the second inequality follows the fact that $\Re(z) \leq |z|$ for any complex number z. We find

$$\begin{aligned} \frac{|F(U)-F(U')|}{d(U,U')} &\leq \sqrt{\frac{D^2 - \left(\sum_{r\in\{0,1\}^d} |\langle\psi_r|\psi'_r\rangle|\right)^2}{2D - 2\sum_{r\in\{0,1\}^d} |\langle\psi_r|\psi'_r\rangle|}} \\ &= \sqrt{\frac{D + \sum_{r\in\{0,1\}^d} |\langle\psi_r|\psi'_r\rangle|}{2}} \leq \sqrt{D}, \end{aligned}$$

where the last inequality follows from Cauchy-Schwarz. □

For every pair of distinct inputs $x,y \in \{0,1\}^n$ and for every $\delta > 0$, it follows from the previous two results that the probability that the protocol's error probability on these inputs exceeds $2^{-m} + \delta$, is upper bounded by

$$2\exp\left(\frac{-\delta^2 D^2}{4}\right)$$

Setting $\delta = 2^{-m}$, $\varepsilon = 2^{-m+1}$ and $d = \lceil \frac{1}{2}\log_2 n + \log_2 \frac{1}{\varepsilon} + 4 \rceil$, by the union bound there is a positive probability that the resulting protocol has error probability at most ε for all input pairs. This implies the existence of the desired protocol, with $m = \lceil \log 1/\delta \rceil = \lceil \log 1/\varepsilon \rceil + 1$ bits of communication.

3.3.3. Lower bounds for bounded-error, entanglement-assisted communication complexity

In this subsection, we look at a number of known lower bounds for entanglement-assisted communication complexity and port them to the commuting operators model. The following result follows easily from our discussion in the first subsection of this section:

Proposition 3.3.7. *Given some protocol in the commuting operators model with C bits of communication between Alice and Bob, there exists a finite-dimensional Hilbert space $\mathscr{H}$ and collections of vectors $\{|v_x\rangle\}_{x\in X}$, $\{|w_y\rangle\}_{y\in Y}$ in $\mathscr{H}$, all of norm at most $2^{\frac{C}{2}}$, such that the probability that the protocol outputs* 1 *on input* (x,y) *is* $\langle v_x|w_y\rangle$.

PROOF. Take an entanglement-assisted protocol in the commuting operators model with communication C and initial state $|\psi\rangle$ in some Hilbert space $\mathscr{H}$. As seen previously, if Alice's input is x and Bob's input is y, the protocol's probability of outputting 1 is:

$$\sum_{\gamma} \langle\psi|A_{\gamma,x,1}B_{\gamma,y,1}|\psi\rangle$$

Take $\mathscr{H}'$ to be the finite-dimensional subspace of $\mathscr{H}$ spanned by the $A_{\gamma,x,1}|\psi\rangle$ and the $B_{\gamma,y,1}|\psi\rangle$. Taking n to be the total number of paths down the tree, for every x, we define $|v_x\rangle \in \mathbb{C}^n \otimes \mathscr{H}'$ by:

$$|v_x\rangle = \sum_{\gamma} |\gamma\rangle \otimes (A_{\gamma,x,1}|\psi\rangle)$$

Similarly, for every Bob input y, we define $|w_y\rangle \in \mathbb{C}^n \otimes \mathscr{H}'$ by:

$$|w_y\rangle = \sum_{\gamma} |\gamma\rangle \otimes (B_{\gamma,y,1}|\psi\rangle)$$

We can see that the norms of the above vectors are at most $2^{\frac{C}{2}}$ and that the protocol's probability of outputting 1 on valid inputs x,y is $\langle v_x|w_y\rangle$. □

Our first lower bound consists in a limitation on the separation between the commuting operators model and the classical model with shared randomness in the bounded-error case, and is a generalisation (and simplification) of a result first shown by [**59**] for the finite-dimensional case. As we have seen previously, it holds that $C^{qc,1}_{1/3}(EQ_n) \leq C^{\text{pub},1}_{1/3}(EQ_n) = O(1)$, while $C_{1/3}(EQ_n) = \log_2 n + O(1)$. This means that there can be no analogue of theorem 3.2.3 for the difference between the models with classical communication with and without prior entanglement if no shared randomness is present. If shared randomness is present however, such an analogue does exist, as was shown by Shi and Zhu ([**59**]). Our proof will be based on the following restatement of the discussion in section 2 of ([**59**]), which itself describes a protocol due to Kremer, Nisan and Ron ([**22**]):

Proposition 3.3.8. *For some fixed $C > 0$ and for some $n \in \mathbb{N}$, consider the following problem: Alice and Bob are given vectors $u, v \in \mathbb{C}^n$ such that $\langle u, v\rangle$ is real and such that $\|u\|_2, \|v\|_2 \leq C$, and they wish to approximate $\langle u, v\rangle$. There exists a universal constant K (independent of C and n) such that, for every ε, $\delta > 0$, there exists a one-way classical protocol with shared randomness and with at most $\frac{KC^4}{\delta^2} \log_2 \frac{1}{\varepsilon}$ bits of communication that produces an approximation of $\langle u, v\rangle$ that is correct up to an additive error of δ except with probability at most ε.*

PROOF SKETCH. We give an alternative to the protocol discussed in [**22**] using the classical Johnson-Lindenstrauss lemma. Note that at the cost of doubling n, we can assume that the components of u and v in the standard basis are all real. In our protocol, Alice and Bob use their shared randomness to pick a uniformly random subspace V of $\mathbb{R}^n$ of a given dimension d. Writing P_V to mean the orthogonal projection onto V, Alice then sends a sufficiently fine discretization of $P_V(u)$ to Bob, from which he can approximately evaluate $\frac{n}{d}\langle P_V(u), P_V(v)\rangle$, which will be a good approximation of $\langle u, v\rangle$ with high probability if d was taken to be large enough. How large d needs to be can be worked out using the Johnson-Lindenstrauss lemma: this does not depend in any way on n, but does depend on the norms of u and v. The amount of communication needed by the protocol can be worked out to be exactly as in the statement of the proposition. □

In combination with proposition 3.3.7, the previous result yields:

Theorem 3.3.9. *For any boolean function f, we have that $C^{pub,1}_{1/3}(f) = O\left(2^{2C^{qc}_{1/3}(f)}\right)$.*

PROOF. Given the $|v_x\rangle$ and $|w_y\rangle$ promised by proposition 3.3.7, on inputs x and y, Alice and Bob use the protocol promised by proposition 3.3.8 to calculate $\langle v_x|w_y\rangle$ sufficiently precisely so as to be able to determine the protocol's most probable output (and therefore to determine the value of $f(x,y)$) with good probability. □

Proposition 3.3.7 also allows us to extract lower bounds for specific communication problems. The heavy lifting in this regard has already been carried out by Linial and Shraibman ([**60**]): we show that their work also goes through for the commuting operators model. Proposition 3.3.7 can be reformulated in matrix language as follows (where the appropriate definitions can be found in [**60**]). This gives an analogue of lemma 12 of [**60**].

Proposition 3.3.10. *Given some protocol in the commuting operators model with input sets X and Y, with binary output and with C bits of communication between Alice and Bob, if the $|X| \times |Y|$ matrix P denotes the acceptance probabilities of the protocol, i.e. $P_{x,y}$ is the probability that the protocol outputs 1 on inputs (x,y), there exist matrices V and W such that $P = VW$ and*

$$\|V\|_{2\to\infty}, \|W\|_{1\to 2} \leq 2^{\frac{C}{2}}$$

From their lemma 12, [**60**] proceeds to recover most known lower bounds for entanglement-assisted communication complexity in the finite-dimensional case. Since lemma 12 still holds in the commuting operators model, all the lower bounds shown there carry over, most notably:

(1) $C^{qc}_{1/3}(\mathrm{DISJ}_n) = \Theta(\sqrt{n})$ (by reduction from Razborov's ([**41**]) results)
(2) $C^{qc}_{1/3}(\mathrm{IP}_n) = \Theta(n)$
(3) For a randomly chosen total function f with input strings of length n, with overwhelming probability, $C^{qc}_{1/3}(f) = \Theta(n)$, i.e. the trivial protocol is optimal up to a constant factor.

3.3.4. A limitation on the power of the Cleve-Buhrman model compared to the Yao model

We now turn to discussing a result of Ronald de Wolf which constrains the power of shared entanglement compared to quantum communication. This material comes from the master's thesis of Jonas Kamminga ([**100**]), which, as far as we know, is the only place in which the result appears in print. Given a function f, we write $f^{(k)}$ to denote the **direct sum** of f, meaning that Alice and Bob are given k instances of f and they must compute the value of f on each of the k instances. f is said to have the **direct sum property** with respect to a complexity measure D if it satisfies the reasonable-sounding condition that $\frac{1}{100}kD(f) \leq D(f^{(k)})$ for all k (where the $\frac{1}{100}$ factor is a mostly arbitrary constant chosen to be small enough so that the theorems we are about to discuss hold). Informally, this means that to compute the k copies of f, modulo a multiplicative factor, it is optimal to individually run the best protocol for f on each of the k copies. The following was proved by de Wolf, based on an idea of Tapp:

Theorem 3.3.11 ([**100**]). *Suppose that f has the direct-sum property for $D = Q_{1/3}$. There exists a universal constant K such that, for all d,*

$$Q_{1/3}(f) \leq K\, C^{qf:d}_{1/3}(f) \log_2 \log_2 d$$

Though this result has strictly no direct bearing on the original problem of whether the Yao and Cleve-Buhrman models are equivalent in the bounded-error case, it implies that for functions with the direct-sum property, if at all possible, a significant separation between the Yao and Cleve-Buhrman models in the bounded-error case could only be achieved with a Cleve-Buhrman protocol using an enormous amount of prior entanglement, which may be taken to mean that in a real-world setting, the two models are essentially equivalent for functions with the direct sum-property. Chapter 5 of [**100**] then proceeds to study what is known about direct sums in both the Yao and Cleve-Buhrman models. A theorem of Jain, Radhakrishnan and Sen is discussed to the effect that all functions do have the direct-sum property for $D = C^{*,1}_{1/3}$. Two examples of functions due to Rao and Sinha and Anshu, Touchette, Yao and Yu, respectively, are then presented which do not have this property in either quantum models in the context of bidirectional communication for distributional complexity. This strongly suggests the possibility that it is not true that all functions have the direct-sum property in the Yao model, and therefore that the above result might not be directly useful in showing that the Cleve-Buhrman and Yao models are equivalent for all functional problems.

3.3.5. The distance between subspaces problem

We end this section by proposing a first-of-its-kind candidate for a separation between the Yao and Cleve-Buhrman models. This problem generalises the vector in subspace problem that was encountered in the previous section and was inspired by a reading of a paper of Hadiashar and Nayak ([**91**]).

For some $n,k \in \mathbb{N}$ with nk even, we define the **distance between subspaces problem** in the following way. Alice and Bob are each given classical descriptions of subspaces $A, B \leq \mathbb{C}^{nk}$, respectively, with A of dimension n and B of dimension $\frac{nk}{2}$.Writing P_U to denote the orthogonal projector on subspace U, they are given the promise that if we write $r = \frac{1}{n}\operatorname{tr}(P_A P_B)$ (it can be seen that $r \in [0,1]$), then either $r \geq \frac{2}{3}$ or $r \leq \frac{1}{3}$ [5] , and they must determine which is the case with good probability and with one-way communication from Alice to Bob. We see that the $n = 1$ case reduces to a problem that is harder than the vector in subspace problem while still being easy quantumly (in the bounded-error setting), and so exhibits an exponential separation between the classical and quantum models.

We now show that there exists a good entanglement-assisted protocol for this problem with communication $\log_2 k + O(1)$, independently of the value of n. The protocol runs as follows. Alice and Bob begin by sharing m maximally entangled states $|\psi_1\rangle, ..., |\psi_m\rangle$ each with local dimension nk, for some m to be determined later. For every $i \in [m]$, Alice then measures her part of the state according to the projective measurement $\{\overline{P_A}, \mathbb{I} - \overline{P_A}\}$, getting m binary outcomes $b_1,...,b_m$, with each b_i being equal to 0 with probability $\frac{1}{k}$. If we take $m = \Theta(k)$, with overwhelming probability, there will exist some index $i \in [m]$ with $b_i = 0$, which Alice sends to Bob at the cost of $\log_2 k + O(1)$ bits of communication. Bob's part of the i'th state is now in state $\frac{1}{n}P_A$, which he measures according to the projective measurement $\{P_B, \mathbb{I} - P_B\}$ and reports the result as the answer of the protocol. Since the probability that 0 is obtained is exactly r, the protocol's success probability will be $\frac{2}{3}$ minus the small probability that no such i existed earlier.

On the other hand, we do not know of a good one-way Yao protocol for this problem. In fact, it can be shown that the most natural attempt to build such a protocol most likely cannot work. This would be a protocol in which Alice tries to communicate enough information to Bob so that he can construct a good enough approximation of the state $\rho_A = \frac{1}{n}P_A$ on his side, which he could then measure according to the measurement $\{P_B, \mathbb{I} - P_B\}$. Ref. [**91**] showed:

Theorem 3.3.12 ([**91**], heavily paraphrased)**.** *Suppose that during the phase of such a protocol that involves the construction of an approximation $\tilde{\rho}_A$ of ρ_A above, Bob does not look at B in any way. Given $\varepsilon > 0$ and k large enough, there exists a constant C such that for all large enough values of n, any protocol of the form above that communicates less than $\log_2 n + \log_2 k - C$ qubits is such that if A is generated uniformly at random according to Haar measure, we have that*

$$\mathbb{E}[D_{tr}(\tilde{\rho}_A, \rho_A)] > 1 - \varepsilon$$

In other words, modulo a constant amount of communication, any Yao protocol which allows Bob to approximate ρ_A to any kind of acceptable precision must essentially consist in Alice sending the whole of ρ_A to Bob. Note also that this holds even in the presence of arbitrary shared randomness.

[5] The original promise in the vector in subspace problem was that $r \in \{0,1\}$. In our case, this seems too strong because it might allow for circumventing the almost-impossibility result below. This is because $\tilde{\rho}_A$ would no longer need to be close to ρ_A in trace distance, as in our version, but only such that $\operatorname{supp} \tilde{\rho}_A \subseteq \operatorname{supp} \rho_A$, for example (refer to the statement of the result for an explanation of the notation).

This means that unless Bob's knowledge of B somehow provides enough information to enable the construction of a good approximation of ρ_A, attempting to construct a protocol of the form above is a dead end, as n is allowed to be arbitrarily larger than k. As far as we can tell, the only good one-way Yao protocol for this problem which readily suggests itself consists in running the Cleve-Buhrman protocol through the reduction of theorem 3.3.9, at the cost of an exponential increase in the communication. It would be interesting to either try to look for a better one-way Yao protocol, or to show that this is essentially the best that can be done for general n if k is large enough. The latter would clearly close the Yao vs. Cleve-Buhrman problem, at least in the one-way setting, but we do not know of a proof technique that could allow one to show this, as the argument of [**65**] seems very specific to the case of classical communication. It is also eminently possible that this problem could give an exponential separation between the two models even when bidirectional communication is allowed, although this is well less-grounded than the case of one-way communication as no equivalent of theorem 3.3.12 is known for the case of bidirectional communication.

Chapter 4

Turning nonlocal games into communication problems

4.1. One-time-pad problems and XOR games

This section describes a generalisation of a communication problem due to Buhrman, Cleve and van Dam ([**26**]), which was the first two-party communication problem exhibiting a quantum advantage to be discovered. The problem to be generalised is the following: Alice is given $x,c \in \{0,1\}$, Bob is given $y \in \{0,1\}$, and they must compute the function

$$f((x,c),y) = c \oplus (x \wedge y)$$

with one bit of communication from Alice to Bob. Classically, it can be shown that the best success probability achievable is $\frac{3}{4}$, simply by checking all possible classical deterministic protocols. In the Cleve-Buhrman model, however, we see that a success probability of $\frac{1}{2} + \frac{\sqrt{2}}{4}$ can be achieved based on the theory of the CHSH game which was covered in the second chapter, using the following protocol: Alice and Bob play the CHSH game with inputs x and y, thereby obtaining outputs a and b; Alice then sends $z = a \oplus c$ to Bob, who outputs $z \oplus b$. Since, no matter the inputs, the probability that $a \oplus b$ equals $x \wedge y$ is $\frac{1}{2} + \frac{\sqrt{2}}{4}$, this is also the success probability of the protocol.

In this section, we will generalise this problem to what we will call one-time-pad problems, viewing the bit c in the above problem as a one-time-pad that we force Alice to apply to her communication. As it turns out, these games are closely related to a class of nonlocal games called XOR games which generalise the CHSH game. In the first two subsections, we will sketch the theory of XOR games, after which we will go back to communication complexity and describe some implications of this theory to the entanglement cost of protocols and to the Cleve-Buhrman vs Yao problem. We will end by discussing a possible generalisation of XOR games.

4.1.1. XOR games and Tsirelson's theorem

A nonlocal game is said to be a **XOR game** if the output sets satisfy $|A| = |B| = \{0,1\}$ and if the game's predicate is of the form $V(a,b|x,y) = [(a \oplus b) = g(x,y)]$ for some function $g : X \times Y \to \{0,1\}$. The most notable such game is the CHSH game, where $|X| = |Y| = 2$, the distribution is taken to be uniform and where $g(x,y) = (x \wedge y)$. As was mentioned in the second chapter, we have that $\omega(\text{CHSH}) = \frac{3}{4}$ and $\omega^*(\text{CHSH}) = \omega_{qc}(\text{CHSH}) = \frac{\sqrt{2}}{4} \approx 0.854$. In this section, we will be exposing the basic theory of XOR games, loosely following the lecture nodes of Richard Cleve ([**89**]).

Unlike the case of general nonlocal games, it turns out that XOR games are very well-behaved in that the space of possible strategies can be described in a tractable way. We sketch the standard proof of this result, leading up to Tsirelson's theorem. Given a finite-dimensional quantum strategy for the game, which is specified by a composite system AB, a possibly entangled state $|\psi\rangle$ over AB and choices of measurements $\{A^x_a\}$, $\{B^y_b\}$ over A and B, respectively (which we may assume to be projective as the output sets are binary, thanks to a theorem of Cleve, Høyer, Toner and Watrous ([**44**])), it is convenient to consider the following matrices, which we refer to as **observables**:

$$A_x = A^x_0 - A^x_1$$
$$B_y = B^y_0 - B^y_1$$

The A_x and B_y have the property of being Hermitian and squaring to the identity. Conversely, the eigenvalues of any matrix satisfying these two properties must be contained in $\{1,-1\}$ and eigenvectors corresponding to different eigenvalues must be orthogonal, so that any choice of A_x and B_y satisfying these two properties is of the above form. For given inputs x,y, if we define $p_{x,y}$ to be the probability that $a \oplus b = 0$, we see that $\langle\psi|A_x \otimes B_y|\psi\rangle = 2p_{x,y} - 1$; this last quantity is referred to as the **bias** of the strategy on input (x,y), to be denoted $e_{x,y}$. Clearly, the biases specify completely the probabilities that $a \oplus b$ equals one for every input pair and therefore specify completely the winning probability provided that the function g and the distribution on the inputs has been fixed.

Now, for every x, set $|a_x\rangle = (A_x \otimes \mathbb{I})|\psi\rangle$, and for every y, set $|b_y\rangle = (\mathbb{I} \otimes B_y)|\psi\rangle$. These are all unit vectors because the A_x and B_y are Hermitian, square to one and $|\psi\rangle$ has norm one. We also see that $e_{x,y} = \langle a_x|b_y\rangle$. We refer to any such choice of vectors as a **vector system**. We have seen that any strategy gives rise to a vector system: it turns out that, in addition, every vector system also gives rise to a strategy. This is the content of Tsirelson's theorem, which we now present. We first state the following standard lemma:

Lemma 4.1.1. *Defining $|\psi\rangle \in \mathbb{C}^n \otimes \mathbb{C}^n$ by*

$$|\psi\rangle = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} |i\rangle\,|i\rangle$$

Any two matrices $A,B \in \mathbb{C}^{n\times n}$ satisfy

$$\langle\psi|A\otimes B|\psi\rangle = \frac{1}{n}\operatorname{tr}(AB^T)$$

We now show:

Theorem 4.1.2 (Tsirelson, [**10**]). *Given a vector system $\{|a_x\rangle\}_{x\in X}$, $\{|b_y\rangle\}_{y\in Y}$ of vectors in $\mathbb{R}^d$, there exists a choice of observables $\{A_x\}_{x\in X}, \{B_y\}_{y\in Y}$ on $\mathbb{C}^{2^d}$ such that, for every input pair,*

$$\langle a_x|b_y\rangle = \langle\psi|A_x\otimes B_y\,|\psi\rangle$$

where $|\psi\rangle$ is the maximally entangled state from the previous lemma.

PROOF. We write X and Z to denote the classical Pauli matrices

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

These matrices are Hermitian, square to the identity and anticommute, i.e. $XZ = -ZX$. For $1 \leq i \leq d$, define:

$$M_i = \left(\bigotimes_{j=1}^{i-1} Z\right) \otimes X \left(\bigotimes_{j=i+1}^{d} \mathbb{I}\right)$$

It is not difficult to see that the M_i are Hermitian, square to the identity and mutually anticommute, by the aforementioned anticommutativity property of the Pauli matrices. Given a unit vector $v \in \mathbb{R}^d$, define $\mathscr{O}(v)$ by

$$\mathscr{O}(v) = \sum_{i=1}^{d} v_i M_i$$

Clearly, $\mathscr{O}(v)$ is Hermitian. Also,

$$\begin{aligned}
\mathscr{O}(v)^2 &= \sum_{i=1}^{d} v_i^2 M_i^2 + \sum_{1\leq i<j\leq d} v_i v_j (M_i M_j + M_j M_j) \\
&= \sum_{i=1}^{d} v_i^2 \mathbb{I} \\
&= \mathbb{I}
\end{aligned}$$

So that $\mathscr{O}(v)$ is an observable. Finally, given unit vectors v,w, by making use of the previous lemma,

$$\begin{aligned}\langle\Psi|\,\mathscr{O}(v)\otimes\mathscr{O}(w)\,|\Psi\rangle &= \frac{1}{d}\operatorname{tr}(\mathscr{O}(v)\mathscr{O}(w))\\ &= \frac{1}{d}\sum_{i=1}^{d} v_i w_i + \sum_{1\leq i\neq j\leq d} v_i w_j \operatorname{tr}(M_i M_j)\end{aligned}$$

Since, for $i \neq j$,

$$\operatorname{tr}(M_i M_j) = \frac{\operatorname{tr}(M_i M_j) + \operatorname{tr}(M_j M_i)}{2} = \frac{\operatorname{tr}(M_i M_j + M_j M_i)}{2} = 0$$

The theorem is proved by setting $A_x = \mathscr{O}(|a_x\rangle)$ for every $x \in X$ and $B_y = \mathscr{O}(|b_y\rangle)$ for every $y \in Y$. □

The above result has a number of useful corollaries. Perhaps the most interesting one is the fact that it provides a means for approximating the entangled value of an arbitrary XOR game efficiently:

Corollary 4.1.3. *Given a XOR game G, there exists an efficient algorithm for approximating $\omega_{qf}(G)$ to arbitrary precision.*

PROOF. Set z^* to be equal to the optimum of the following semidefinite program (which we write using dot products of vectors, but which can straightforwardly be recast in the more standard form involving a semidefinite matrix by considering the vectors' Gram matrix):

$$\max \sum_{x\in X, y\in Y} \mu_{x,y}(-1)^{g(x,y)}\langle a_x|b_y\rangle \tag{4.1.1}$$

$$\text{s.t.}\ \langle a_x|a_x\rangle = 1 \qquad \forall x\in X \tag{4.1.2}$$

$$\langle b_y|b_y\rangle = 1 \qquad \forall y\in Y \tag{4.1.3}$$

$$\tag{4.1.4}$$

By the correspondence between vector systems and strategies, we see that $\omega_{qf}(G) = \frac{z^*+1}{2}$. Since semidefinite programs can be solved efficiently to within an arbitrary precision $\varepsilon > 0$, this proves the corollary. □

This is in contrast to the classical case, where it is known (by reduction from MAXCUT and by appealing to the PCP theorem) that even approximating the value of a XOR game to within a fixed given small enough precision is NP-hard.

Another important corollary of Tsirelson's theorem is that it restricts the power of entanglement for XOR games:

Corollary 4.1.4. *Given a XOR game G, there exists a finite-dimensional strategy for G with success probability exactly $\omega_{qf}(G)$ which makes use of $min(|X|,|Y|)+1$ EPR pairs. Also, we have that the semidefinite program mentioned in the previous corollary is isomorphic to the first level of the sum-of-squares hierarchy for G. In particular, $\omega^*(G) = \omega_{qc}(G)$.*

PROOF. For the first part, suppose that $|X| \leq |Y|$: the other case is symmetric. Given a vector system $\{|a_x\rangle\}$, $\{|b_y\rangle\}$, we may project the $|b_y\rangle$ onto the subspace generated by the $\{|a_x\rangle\}$ without changing the scalar products $\langle a_x|b_y\rangle$. Clearly, this subspace is of dimension at most $|X|$. To renormalise the $|b_y\rangle$ after this (which might now have norm strictly less than one), we can add another vector to the space that is orthogonal to the $|a_x\rangle$ and appropriately choose the scalar product of the $|b_y\rangle$ with this new vector. Also, since the semidefinite program in the proof of the above corollary is real, we can assume that said vector space is real. The second part is left as an exercise to the reader. □

Note that for general nonlocal games, none of the statements in the above corollary hold.

Finally, although this won't be of any direct importance to us, the following can be deduced from the above correspondence:

Corollary 4.1.5 ([89]). *Let G be a XOR game. There exists a universal constant $C \approx 1.1382$ such that $\omega^*(G) \leq C\omega_c(G)$. In particular, if $\omega_c(G)$ is close to 0, so is $\omega^*(G)$. Also, there exists a universal constant $D \approx 0.74202$ such that if $\omega_c(G) > D$, then $\omega^*(G) \leq \left(\sin\frac{\pi\omega_c(G)}{2}\right)^2$. In particular, if $\omega^*(G)$ is close to 1, so is $\omega_c(G)$.*

Note how tightly the CHSH game saturates the above corollary: the first part of corollary, along with the fact that the classical value of the CHSH game is 0.75, gives an upper bound on the entangled value of the CHSH game that is tight to within 10^{-4}, and the upper bound given by the second part matches the entangled value exactly. In particular, the above result shows that the quantum-classical gap exhibited by the CHSH game is pretty much optimal for XOR games, both additively and multiplicatively. It follows that one cannot hope for dazzling gaps between the classical and entangled values of a XOR game, as opposed to the case of general nonlocal games, where, as we have seen, the gap can be arbitrarily close to one.

4.1.2. The tightness of Tsirelson's theorem and the CHSH(n) games

As we have seen, given a XOR game G with input sets X,Y, Tsirelson's theorem allows us to build an exactly optimal entangled strategy for G. However, this strategy requires a surprisingly large amount of prior entanglement: if Alice and Bob's input were represented as bit strings of length n, the number of EPR pairs used by the constructed strategy would be exponential in n. In fact, it is known that if we are satisfied with strategies that are ε-optimal, it is possible to do much better than this:

Theorem 4.1.6 (Regev, unpublished)**.** *There exists a universal constant C such that, given a XOR game G and any $\varepsilon > 0$, there exists an ε-optimal finite-dimensional strategy for G which makes use of at most $\frac{C}{\varepsilon^2}$ EPR pairs.*

However, there is a sense in which the construction we gave is essentially optimal. Namely, Slofstra ([**66**]) showed that there are XOR games with the property that any strategy that is close enough to be optimal (where the definition of 'close enough' must necessarily depend on the game so as not to contradict the previous theorem) must use an amount of entanglement that morally matches the upper bound given by Tsirelson's theorem above. We now describe a family of games introduced by Slofstra in the aforementioned paper for which this holds [1], which are called the CHSH(n) games. For $n \geq 2$, the corresponding game is defined to have input sets

$$X = [n]$$

$$Y = \{(j_1,j_2) \mid j_1,j_2 \in [n],\ j_1 \neq j_2\}$$

The players are given the promise that if Alice's input is i and Bob's input is (j_1,j_2), then $i = j_1$ or $i = j_2$. The function g which specifies the winning condition of the game is

$$g(i,(j_1,j_2)) = (j_1 > j_2) \wedge (i = j_1)$$

The CHSH(n) games get their name from the fact that CHSH(2) can be seen to be the usual CHSH game, which was introduced in the second chapter, where $x = 0$ correspond to the input $i = 1$, $x = 1$ corresponds to $i = 2$, $y = 0$ corresponds to $(1,2)$ and $y = 1$ corresponds to $(2,1)$.

Taking the input distribution to be uniform over the legal input pairs, it can be shown that the quantum values of the CHSH(n) games all satisfy the same bound as the usual CHSH game, i.e. $\omega^*(\text{CHSH}(n)) \leq \frac{1}{2} + \frac{\sqrt{2}}{4}$. It is also straightforward to describe a strategy which achieves this in the worst case using the language of observables that was introduced in the previous section. Take Alice's observables $A_1,\ldots,A_n$ to be any collection of mutually anticommuting observables on a finite-dimensional Hilbert space $\mathscr{H}$ of dimension d, such as the ones used in the proof of Tsirelson's theorem, and, for $(j_1,j_2) \in Y$, define $B_{(j_1,j_2)}$ by

$$B_{(j_1,j_2)} = \frac{1}{\sqrt{2}}((-1)^{I[j_1>j_2]}A_{j_1}^T + A_{j_2}^T)$$

[1] It should be said that many corners are being cut in our presentation of Slofstra's results. The upper bound given previously on the amount of entanglement needed by an optimal strategy for a XOR game isn't actually optimal (being off by a multiplicative factor), and the games we are about to introduce saturate the best possible bound only in a moral sense. We feel that giving a completely precise account of these things would take us too far afield, and we refer the interested reader to Slofstra's paper.

A routine verification shows that the $B_{(j_1,j_2)}$ are also observables. Finally, taking $|\psi\rangle \in \mathscr{H} \otimes \mathscr{H}$ to be the maximally entangled state from lemma 4.1.1, we see that, because $\mathrm{tr}(A_iA_j) = d\delta_{i,j}$,

$$\langle\psi|A_iB_{(j_1,j_2)}|\psi\rangle = \frac{1}{d}\mathrm{tr}(A_iB^T_{(j_1,j_2)}) = \frac{(-1)^{I[j_1>j_2\wedge i=j_1]}}{\sqrt{2}}$$

Translating from the language of biases back into the language of success probabilities, we see that this means that, importantly for us, for all choices of legal inputs, the success probability of the players using this strategy is exactly $\frac{1}{2}+\frac{\sqrt{2}}{4}$. Using the M_i that were defined in the course of proving Tsirelson's theorem as Alice's observables, this strategy will take n ebits [2] .

Slofstra also shows:

Theorem 4.1.7 ([**66**], oversimplified)**.** *There exists a sequence $\{\varepsilon_n\}_{n\in\mathbb{N}}$ of strictly positive real numbers (necessarily converging to zero in view of theorem 4.1.6) such that, for every n, any ε_n-optimal entangled strategy for CHSH(n) must employ $\Omega(n)$ ebits.*

This will be of interest for us in the two following subsections.

4.1.3. The complexity of binary one-time-pad problems with classical communication

We now generalise the problem of [**26**] that was introduced at the beginning of this section to **one-time-pad problems**. Given input sets X,Y, a set $R \subseteq X\times Y$ of legal inputs and a function $g : X\times Y \to \{0,1\}$, Alice is given $x\in X, c\in\{0,1\}$ and Bob is given $y \in Y$ with $(x,y)\in R$, and they are requested to compute the function

$$f((x,c),y) = c\oplus g(x,y)$$

with one bit of communication from Alice to Bob. In notation, we are interested in the value $\omega_t^{\mathrm{pub}}(f,1)$ for all models t. We might also look at the value $\omega_t^{\mu'}(f,1)$ for some distribution μ' such that $\mu'_{(x,c),y} = 0$ whenever $(x,y)\notin R$, which upper bounds $\omega_t^{\mathrm{pub}}(f,1)$. Actually, it can be seen that the tightest upper bounds are obtained when c is uniformly distributed and independent of x and y: therefore, we only need to concern ourselves with the marginal distribution μ on $X\times Y$.

In the classical and Cleve-Buhrman models, the obvious thing to do would be for Alice and Bob to play the corresponding XOR game specified by μ and g with the best strategy available, obtaining outputs $a,b\in\{0,1\}$, after which Alice send $z = c\oplus a$ to Bob and Bob outputs $b\oplus z$. The success probability of this protocol will be exactly the same as the success probability of the strategy used for the XOR game. We now show that in the distributional setting, we may restrict

[2] This can be lowered to $\frac{n}{2}$ ebits.

ourselves to strategy of this form:

Theorem 4.1.8. *In all of the classical, finite-dimensional and commuting operators models, any protocol for the above distributional communication problem with success probability p can be converted into an equivalent (i.e., in the finite-dimensional setting, using the same entangled state) strategy for the corresponding XOR game with success probability no smaller than p.*

PROOF. We give the proof for qc: the other cases are similar. Take a protocol for f, specified by a Hilbert space H, a state $|\psi\rangle$ on H and projective measurements $A_z^{x,c}$, $B_r^{y,z}$ for Alice and Bob, respectively, where r is the output of the protocol. Writing the shorthand $\tau(\cdot)$ to mean $\langle\psi|\cdot|\psi\rangle$, for a given input pair (x,y), we see that the protocol's success probability on that input pair is:

$$\frac{1}{2}\left(\tau\left(A_0^{x,0}B_{g(x,y)}^{y,0}\right)+\tau\left(\left(\mathbb{I}-A_0^{x,0}\right)B_{g(x,y)}^{y,1}\right)+\tau\left(\left(\mathbb{I}-A_0^{x,0}\right)\left(\mathbb{I}-B_{g(x,y)}^{y,0}\right)\right)+\tau\left(A_0^{x,0}\left(\mathbb{I}-B_{g(x,y)}^{y,1}\right)\right)\right)$$

Rearranging, this gives

$$\frac{1}{2}+\frac{(-1)^{g(x,y)}}{2}\tau\left(\left(A_0^{x,0}-A_0^{x,1}\right)\left(B_0^{y,0}-B_0^{y,1}\right)\right)$$

For the rest of the proof, we freeze Alice's input to be some $x\in X$ with nonzero probability under μ. We show that it can be assumed that $A_0^{x,1}=\mathbb{I}-A_0^{x,0}$. This effectively means that we can think of Alice as performing one single binary measurement, irrespective of the value of c, and then sending the result XOR-ed with c to Bob.

Writing

$$B=\sum_y \mu_{y|x}(-1)^{g(x,y)}\left(B_0^{y,0}-B_0^{y,1}\right)$$

By linearity, we get that the protocol's success probability is:

$$\frac{1}{2}+\frac{1}{2}\tau\left(\left(A_0^{x,0}-A_0^{x,1}\right)B\right)$$

If we drop the old value of $A_0^{x,0}$ and take $A_0^{x,0}=\mathbb{I}-A_0^{x,1}$ instead, the success probability of the new protocol is:

$$\frac{1}{2}+\frac{1}{2}\tau\left(\left(\mathbb{I}-2A_0^{x,1}\right)B\right)$$

While if we instead drop the old value of $A_0^{x,1}$ and take $A_0^{x,1}=\mathbb{I}-A_0^{x,0}$, the success probability of the new protocol is:

$$\frac{1}{2}+\frac{1}{2}\tau\left(\left(2A_0^{x,0}-\mathbb{I}\right)B\right)$$

The average of the two previous success probabilities is equal to the success probability of the original protocol, so that one of the two modified protocols is at least as good as the original protocol. This completes the proof from the point of view of Alice. In the same way, we can show that without loss of generality, it holds that $B_0^{y,1}=\mathbb{I}-B_0^{y,0}$ for all y: this amounts to Bob making a measurement independently of Alice's communication, XOR-ing the result with Alice's communication and outputting the result, as desired. □

This shows that in the case of CHSH, the best success probability is $\frac{3}{4}$ classically and $\frac{1}{2}+\frac{\sqrt{2}}{4}$ in the entanglement-assisted scenario. Also, in combination with theorem 4.1.7, this yields an example of a communication problem for which a near-optimal protocol using one bit of communication from Alice to Bob requires a great deal of prior entanglement (i.e. exponential in the sizes of the inputs when viewed as bit strings).

4.1.4. The case of the Yao model

In this subsection, we investigate the complexities of one-time-pad problems in the Yao model. We start with an achievability result:

Proposition 4.1.9. *Given a XOR game G and a strategy for G making use of an entangled state of local dimension 2, there exists a private-coin protocol for f in the Yao model such that, for every input pair x,y and for every c, the protocol's success probability is equal to the success probability of the original strategy for G given that Alice and Bob's inputs were x and y. In particular,*

$$\omega^1_{qf:2}(f,1) \leq \omega^1_{yao}(f,1)$$

PROOF. Take the best strategy for G which makes use of an entangled state of local dimension 2. This is specified by a state $|\psi\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2$, projectors $\{A^x_a\}$ for Alice and projectors $\{B^y_b\}$ for Bob. We now describe a Yao protocol for f with the properties given in the statement of the proposition. Given Alice and Bob's input (x,c) and y, Alice begins by flipping a private coin $a \in \{0,1\}$ in such a way that the probability that $a = 1$ is the same as the probability of her obtaining 1 in the chosen strategy for G. Writing ρ_B to be the reduced density matrix on Bob's side in the above strategy for G given that Alice measured according to the projector $\{A^x_a\}$ and got result a, if $a \oplus c = 0$, Alice sends ρ_B to Bob, and if $a \oplus c = 1$, she sends $\mathbb{I} - \rho_B$ otherwise. Bob then applies his measurement from the original strategy for G to what he received from Alice and outputs the result. □

This means that the one-time-pad problem based on the CHSH game, which was introduced at the beginning of this section, also has a Yao protocol with one qubit of communication achieving a success probability of $\frac{1}{2}+\frac{\sqrt{2}}{4}$, and therefore does not feature a gap between the Yao and Cleve-Buhrman models.

We now show that this is optimal: indeed, we show that any Yao protocol with one qubit of communication for a one-time-pad problem can be converted into a Cleve-Buhrman protocol for the same one-time-pad problem with one bit of communication that is at least as good and which uses a constant amount of entanglement. Note that if the states that Alice is sending in that Yao protocol happened to all be real, this conversion would easily be achieved because real states in two dimensions can be teleported with one bit of communication and one shared EPR pair [3]. The general case where Alice's communication may be complex will require slightly more work.

Theorem 4.1.10. *Given the distribution μ on $X \times Y$ corresponding to G,*

$$\omega_{yao}^{1,\mu'}(f,1) \leq \omega_{qf:256}(G) = \omega_{qf:256}^{1,\mu'}(f,1)$$

PROOF. Take a Yao protocol for f specified by pure states $\{|\psi\rangle_{x,l}\} \in \mathbb{C}^2$ to be sent to Bob by Alice and measurements $\{E_r^y\}$ to be performed by Bob on the system he received from Alice, which may be assumed to be projective measurements by the proof of proposition 2 of [**44**]. Clearly, for a given y, E_0^y has rank either 0, 1 or 2. If E_0^y were to have rank 0 or 2, which would amount to Bob always outputting the same thing on input y irrespective of the communication, then because the bit c is picked uniformly at random, the success probability conditionally on Bob's input being y would be exactly $\frac{1}{2}$. This means that Bob could instead scramble the communicated state by performing one of $\mathbb{I}, X, Z, XZ$ at random and then measuring in the computational basis, which would also yield a success probability of $\frac{1}{2}$. It may therefore be assumed that E_0^y has rank 1 for all y, so that we can write $E_0^y = |\phi_y\rangle\langle\phi_y|$. Also, for the same reason as in the proof of theorem 4.1.8, we can assume that $|\psi_{x,0}\rangle \perp |\psi_{x,1}\rangle$ for all x. This means that for all input pairs (x,y), the success probability is the same for $l = 0$ and $l = 1$.

We now invoke Tsirelson's correspondence between entangled strategies and vector systems to turn the protocol into an entangled strategy for G. For $x \in X$, define

$$|v_x\rangle = \frac{1}{\sqrt{2}}|\psi_{x,0}\rangle \otimes \overline{|\psi_{x,0}\rangle} - \frac{1}{\sqrt{2}}|\psi_{x,1}\rangle \otimes \overline{|\psi_{x,1}\rangle}$$

And for $y \in Y$, define

$$|w_y\rangle = \frac{1}{\sqrt{2}}|\phi_{y,0}\rangle \otimes \overline{|\phi_{y,0}\rangle} - \frac{1}{\sqrt{2}}|\phi_{y,1}\rangle \otimes \overline{|\phi_{y,1}\rangle}$$

Where the overline means complex conjugation. Because $|\psi_{x,0}\rangle \perp |\psi_{x,1}\rangle$ and $|\phi_{y,0}\rangle \perp |\phi_{y,1}\rangle$, these vectors have unit norm: using the same fact, we see that if we write $p_{x,y} = |\langle\psi_{x,0}|\phi_{y,0}\rangle|^2$ to mean the probability that the protocol's output is 0 given that Alice and Bob's inputs were x and y, then

$$\langle v_x|w_y\rangle = 2p_{x,y} - 1$$

This means that running this vector system through Tsirelson's correspondence would yield a strategy such that, for every input pair (x,y), the probability that the outputs XOR to 0 is the same as the probability that Bob outputs 0 in the original Yao protocol. These vectors are in $\mathbb{C}^4$, so that we

[3] Rotating a real two-dimensional state by 90 degrees always yields an orthogonal state; this fact can easily be turned into a protocol for teleporting said state. No such transformation is possible for general complex states.

can embed them in $\mathbb{R}^8$ without changing their inner products. This gives a strategy for G using 8 EPR pairs with success probability equal to that of the original protocol. $\square$

Note that we know that for the CHSH(n) game, the corresponding one-time-pad problem f is such that $\omega_{qf}^{1,\text{pub}}(f,1) = \frac{1}{2} + \frac{\sqrt{2}}{4}$, while, taking μ to be uniform over the legal inputs, the amount of entanglement required by any near-optimal strategy is $\Omega(n)$. By the previous result, the corresponding one-time-pad problem must be such that $\omega_{\text{yao}}^{1,\text{pub}}(f,1) < \frac{1}{2} + \frac{\sqrt{2}}{4}$ if n is taken to be large enough. As far as we know, this gives the first ever example of a communication problem for which r bits of communication with prior entanglement can buy you more than r qubits of communication without prior entanglement but with arbitrary shared randomness, although the gap between the two values is obviously not very impressive. Taking $n = 5$, numerically, we find that the gap is about 2%. Another such problem will be encountered in the next section.

4.1.5. Group games and maximal entanglement in communication complexity

We end this section by mentioning a possible generalisation of XOR games which we do not believe has been studied in the literature. Given finite input sets X,Y, a distribution $p_{x,y}$ on $X \times Y$, a finite group H [4] and a function $g : X \times Y \to H$, we define the corresponding **group game** to be the game with both output sets equal to H and such that the winning condition is $a^{-1}b = g(x,y)$. Taking $H = \mathbb{Z}/2\mathbb{Z}$ gives standard XOR games. One might wonder about which properties of XOR games still hold in this generalised setting. Note that even if we take H to be the next-simplest non-trivial finite group, namely $H = \mathbb{Z}/3\mathbb{Z}$, the observable picture that we used to analyse XOR games earlier in this section is no longer available, and, indeed, some properties of XOR games that could be proven using this picture can be shown to be false.

Cui, Mehta, Mousavi and Nezhadi ([**99**]) study a different generalisation of the CHSH game from the one that was mentioned earlier in this section. Although this is not trivial to figure out from their presentation, the games they consider are as follows: for $n \geq 2$, the inputs are taken to be binary, the function g is the same as in the the original CHSH game, namely $g(x,y) = x \wedge y$, and we take $H = \mathbb{Z}/n\mathbb{Z}$ [5]. Clearly, $n = 2$ is just the usual CHSH game. For $n = 3$, they could already prove that in the entangled setting, the resulting game exhibits some features that are forbidden for XOR games: it is no longer the case that the bound given by the first level of the sum-of-squares hierarchy always matches the entangled value of the game, and it is no longer true that the optimal

[4] The letter G is rather overused in this thesis, as it is used for games, groups and graphs. We compensated for this earlier by using the letter ‘g’ for the function corresponding to a XOR game, which secretly stood for ‘group’.

[5] These games, being a natural enough generalisation of the CHSH game, had already (sort of) been studied by Dupuis, Gisin, Hassidim, Méthot and Pilpel ([**56**]), although from an angle that is orthogonal to the one taken by [**99**]. In particular, [**56**] does not consider the landscape of quantum strategies for these games.

winning probability can always be attained by a strategy which uses a maximally entangled state. [6]

This last point is interesting to us because it might yield an example of a communication problem for which general entanglement might allow one to do better than maximal entanglement; no communication problem with this property is known at present (theorem 3.3.1 implies that it is not possible to do much better with general entanglement than with maximal entanglement; here, we are restricting ourselves to protocols with one bit of communication only, on which theorem 3.3.1 has no bearing). One might be tempted to generalise one-time-pad problems to group games in the obvious way: namely, we give Alice $x \in X$ and $c \in G$, we give Bob $y \in Y$, and we request them to compute

$$f((x,c),y) = c\, g(x,y)$$

A strategy is then given by having them play the game, obtaining results $a,b \in H$, and having Alice send $z = ca^{-1}$ to Bob, who then replies with zb. If we could prove that theorem 4.1.8 still holds for this generalised class of communication problems, by the result of [**99**] mentioned earlier, we would have shown that there exists some communication problem for which a maximally entangled state isn't optimal. Unfortunately, it does not: letting our computer loose on the one-time-pad problem corresponding to the game of [**99**], we could find a protocol which has a noticeably higher success probability than the value of the game and that this protocol makes use of a maximally entangled state.

It would be interesting to look for a provably correct way of turning the game of [**99**] into a communication problem for which it is not true that maximally entangled states are complete. We also think that it would be interesting to try to further develop the theory of group games: For example, it would be interesting to see if some results which we know hold for XOR games could be proved for these in a weaker form, for example:

(1) Is the tensor-product value of a group game always attained by some finite-dimensional strategy? If so, is there a reasonable upper bound on the minimal dimension of the entangled state required to achieve this?

[6] As an aside, their question 1.1, which they solve using this game, asks: 'Are there states other than the maximally entangled state that can be self-tested by a nonlocal game? '. Actually, the tilted CHSH inequalities of [**67**], while typically framed as a means of self-testing 2×2 entangled states using correlations, can straightforwardly be recast as a traditional nonlocal game in the following way. For a given parameter $p \in [0,1]$, Alice is given $x \in \{0,1\}$, and Bob is given $y,c \in \{0,1\}$; x, y, c are all independent, x and y are uniformly distributed and c equals 1 with probability p. Their outputs a,b are to be binary. If $c = 1$, the winning condition is $a \oplus b = x \wedge y$; if $c = 0$, the winning condition is $x = 0$. $p = 1$ is the vanilla CHSH game, while $p = 0$ is a rather boring game; making p vary from 0 to 1, we obtain a self-test for every 2×2 entangled pure state. This is a much easier and satisfying answer to their question 1.1, as they could only exhibit one single non-maximally-entangled state which admits a self-test. It seems very likely that something like this can be done to show that every entangled state can be self-tested using a nonlocal game, as we know can be done using general correlations.

(2) Similarly, are the tensor-product and commuting operators models equivalent for such games?

(3) What is the computational complexity of the value of such games? Is there always some level of the sum-of-squares hierarchy which gives the value of the game exactly? Is there an upper bound on said level?

(4) Does a version of corollary 4.1.5 still hold for these games, or can the quantum-classical gap of such games be arbitrarily close to one provided that the group G is taken to be large/complicated enough?

4.2. One-way, zero-error communication complexity and quantum graph theory

In this section, we take a look at communication complexity in a rather restricted model, namely the one in which only Alice is allowed to communicate to Bob and in which the participants are required to obtain the right answer with certainty. As we shall see, this path leads us by force into graph theory. Take G to be a finite simple graph. Following section 8.5 of de Wolf's PhD thesis ([**37**]), we define the **promise equality problem** EQ_G as follows: Alice and Bob are given vertices of G as input, with the promise that their inputs are either the same or adjacent, and they must determine which of the two is the case. It is possible to show that problems of this form are complete for exact, one-way communication complexity, and also that the reduction from a general problem to a promise equality problem is rather straightforward (although the proof of the correctness of the reduction takes a bit of work). This is done in the next subsection.

In the remainder of this section and unless specified otherwise, we will be measuring the amount of communication in terms of **register size** and not in terms of number of bits sent. Namely, a protocol with communication k will correspond to Alice sending an element c of the register $\{1,\dots,k\}$ to Bob in the classical and Cleve-Buhrman models and a state $\rho \in D(\mathbb{C}^k)$ in the Yao model. We will also assume that the players do not make use of shared randomness as this is clearly useless in the exact setting.

We now mention some facts about one-way protocols in the Cleve-Buhrman model which will be needed in the course of this section. First, it should be noted that if the communication is classical, one-way communication complexity in the distributional setting can actually be regarded as a nonlocal game. Indeed, for a given amount of communication C, consider the game where Alice and Bob are given $x, (y,c)$ as inputs, where (x,y) is sampled according to some distribution on $X \times Y$ and where $c \in [C]$ is uniformly distributed and independent of x and y, where Alice outputs $a \in [C]$ and Bob outputs $b \in \{0,1\}$, and they win if either $c \neq a$ or if $f(x,y) = b$. It can easily be seen that $\omega_t(G) = \frac{C-1}{C} + \frac{1}{C}\omega_t^{\mu,1}(f,C)$, and in particular that G has a perfect strategy if and only if f has an exact one-way protocol with communication C. Next, we have that in the finite-dimensional settings, protocols can be put in **standard form**, to be defined in the statement of the following lemma. We will be writing $\rho_{x,c}$ to mean the reduced density matrix on Bob's side provided that Alice's input was x and she communicated c to Bob. The trick we use is due to [**53**].

Lemma 4.2.1. *At the cost of increasing the communication cost of the protocol, any one-way protocol in the Cleve-Buhrman model with finite-dimensional entanglement can be put into standard form. This means that if the protocol's communication is k, Alice's communication c is uniformly distributed over $[k]$, and the $\rho_{x,c}$ all have the same rank.*

PROOF. We may assume that in the original protocol, the shared entangled state has full Schmidt rank. In addition to the entanglement used in the original protocol, Alice and Bob share a maximally entangled state $|\psi\rangle$ over $\mathbb{C}^k \otimes \mathbb{C}^k$. Alice begins by proceeding as in the original protocol, obtaining c'. She then measures her share of $|\psi\rangle$ in the standard basis, obtaining a uniformly random $d \in [k]$, and sends $c = c' \oplus d$ to Bob. Bob then measures his part of the state to obtain d and recovers $c' = c \oplus (-d)$. Clearly, c is uniformly distributed, and it may be verified that the $\rho_{x,c}$ all have the same rank. □

4.2.1. The completeness of promise equality problems for one-way, zero-error communication complexity

In all that follows, we have some function $f : X \times Y \to Z \cup \{\perp\}$ in mind, where X, Y and Z are arbitrary finite sets, and we will construct a graph G_f such that the zero-error complexities of f and EQ_{G_f} are equal in all models [7]. This will show that promise equality problems are complete for exact, zero-error communication complexity. While the construction itself is simple, proving its correctness in all cases is somewhat tedious, but some of the lemmas we will prove to this end will be useful later on.

Let us first consider the classical case. For every $x \in X$, take $c_x \in \{1, \ldots, k\}$ to be Alice's message on input x. For any distinct $x,x' \in X$, if there exists $y \in Y$ such that $f(x,y) \neq f(x',y)$ and $f(x,y), f(x',y) \neq \Box$, we must necessarily have that $c_x \neq c_{x'}$, for otherwise, Bob would reply the same thing when given input y whether Alice was given input x or x'. Conversely, any choice of messages c_x with the aforementioned property yields a classical protocol for f: since the sets

$$C_r = \{c_x \mid x \in X, f(x,y) = r\}$$

[7] In the classical and quantum settings. They are not equal in the presence of general nonsignalling correlations, but as we have established previously, communication complexity in the presence of general nonsignalling correlations isn't wildly interesting.

are all disjoint, Bob can determine which $r \in R$ is such that C_r contains the message that he received from Alice and output it.

Let us then define an undirected graph G_f with vertex set X and such that $(x,x') \in E(G_f)$ if and only if there exists $y \in Y$ such that $f(x,y) \neq f(x',y)$ and $f(x,y), f(x',y) \neq \square$. It is apparent that there exists a protocol for EQ_{G_f} in the classical setting with communication k if and only if it is possible to assign $c_x \in \{1,\ldots,k\}$ to every vertex x in such a way that whenever $(x,x') \in E(G_f)$, we have that $c_x \neq c_{x'}$, by the previous discussion, so that f and EQ_{G_f} have the same one-way, exact communication complexities in the classical setting. We will show that the same is true in all settings:

Theorem 4.2.2. *In the one-way case, the exact and zero-error complexities of f and EQ_{G_f} are all equal. In particular, promise equality problems are complete for exact and zero-error one-way communication.*

In each case, the proof is quite similar to the one just given. Note that this contains the following result as a special case [8].

Corollary 4.2.3 ([**36**]). *Suppose that f is total, meaning that all input pairs are possible. We can assume that for every x,x', there exists some y with $f(x,y) \neq f(x',y)$, as otherwise, one of the two can be thrown out without changing the communication complexity. Then, G_f is just the complete graph on $|X|$ vertices, so that the one-way, zero-error complexities of f are equal in all models, by proposition 4.2.19 below.*

4.2.1.1. The proof for the Yao model. The proof for the Yao model is a direct adaptation of the argument for the classical case, and follows at once from the following result:

Proposition 4.2.4. *Let $\mathscr{H}$ be a finite-dimensional Hilbert space and let $C_1, C_2, \ldots, C_n$ be a collection of sets of density matrices on $\mathscr{H}$. The following are equivalent:*

(1) There exists a projective measurement $\{P_r\}_{r=1}^{n+1}$ on $\mathscr{H}$ such that for any $r \in \{1,\ldots,n\}$ and for any $\rho \in C_r$, $\operatorname{tr}(P_r\rho) = 1$.

(2) For any $1 \leq r < r' \leq n$, $\rho \in C_r$ and $\sigma \in C_{r'}$, we have that ρ and σ are perfectly distinguishable, or equivalently, $\operatorname{supp}\rho \perp \operatorname{supp}\sigma$.

PROOF. That (1) implies (2) is clear enough, as the first part provides a projective measurement to distinguish perfectly between ρ and σ. For the other direction, for every $r \in \{1,\ldots,n\}$, define the

[8] Actually, Klauck did not show this for the commuting operators model. This is forgivable considering that the commuting operators model had not yet been defined.

subspace V_r by

$$V_r = \bigoplus_{\rho \in C_r} \operatorname{supp} \rho$$

So that, for any $1 \leq r_1 < r_2 \leq n$, $V_{r_1} \perp V_{r_2}$. Taking P_r to be the orthogonal projector with image V_r and setting $P_{n+1} = \mathbb{I} - \sum_r P_r$, we see that the resulting collection of orthogonal projectors constitutes a projective measurement that does what we want. □

First take an exact, one-way protocol for f in the Yao model with communicated register $\mathscr{H}$, and take $y \in Y$. We can then derive an equivalent protocol for EQ_{G_f} where, on input y and upon the reception of ρ_x from Alice, Bob measures according to the projective measurement $\{P_{\operatorname{supp} \rho_y}, \mathbb{I} - P_{\operatorname{supp} \rho_y}\}$. Conversely, given an exact protocol for EQ_{G_f}, for every $r \in R$, we set:

$$C_r = \{\rho_x \mid x \in X, f(x,y) = r\}$$

The implication (2) → (1) in the above proposition then yields an exact protocol for f, since, for $r \neq r'$ and every $\rho \in C_r$, $\sigma \in C_{r'}$, ρ and σ are perfectly distinguishable, by the correctness of the original protocol.

4.2.1.2. The proof for the Cleve-Buhrman model in the exact case. We now turn to the Cleve-Buhrman model in the exact case: the approximate case will be dealt with in the next section. We will only consider the commuting operators model: the proof will also go through for the other entanglement models. We begin by showing:

Proposition 4.2.5. *Let H be a Hilbert space, and let V_1, V_2 be closed subspaces of $\mathscr{H}$. If P_1 and P_2 are the orthogonal projectors on V_1 and V_2, respectively, then P_1 and P_2 commute if and only if there exist (necessarily) orthogonal closed subspaces U_1, U_2 of $\mathscr{H}$ satisfying:*

(1) $U_1 \subseteq V_2$

(2) $U_2 \subseteq V_2^{\perp}$

(3) $V_1 = U_1 \oplus U_2$

PROOF. For the forward direction, set

$$U_1 = V_1 \cap V_2$$

$$U_2 = V_1 \cap V_2^{\perp}$$

These subspaces are both closed because V_1, V_2 and $V_2^{\perp}$ are. Clearly, (1) and (2) hold. We show that (3) also holds. The inclusion $U_1 \oplus U_2 \subseteq V_1$ is clear. For the other inclusion, let $v \in V_1$. Then

$$v = P_1(v) = P_2(P_1(v)) + (\mathbb{I} - P_2)(P_1(v))$$

By the commutativity of P_1 and P_2, we have that $P_2(P_1(v)) \in U_1$, and similarly, by the commutativity of P_1 and $(\mathbb{I} - P_2)$ and since $(\mathbb{I} - P_2)$ is the orthogonal projector on $V_2^{\perp}$, we have that

$(\mathbb{I}-P_2)(P_1(v)) \in U_2$. We therefore have that $v \in U_1 \oplus U_2$, which shows the claim.

For the backwards implication, write $U_3 = V_2 \cap U_1^{\perp}$ and $U_4 = (U_1 \oplus U_2 \oplus U_3)^{\perp}$, so that $H = U_1 \oplus U_2 \oplus U_3 \oplus U_4$ (where we've freely used the fact that the direct sum of orthogonal, closed subspaces is also closed as well as the fact that if $W \leq H$ is closed, then $H = W \oplus W^{\perp}$). Given a generic element $v \in \mathscr{H}$ written uniquely as $v = u_1 + u_2 + u_3 + u_4$ with $u_i \in U_i$, it can be seen that $P_1 P_2 v = P_2 P_1 v = u_1$, which completes the proof. □

Corollary 4.2.6. *Let $\mathscr{H}$ be a Hilbert space, and let $\{V_i\}_{i=1}^n$ be a collection of closed subspaces of $\mathscr{H}$. Set*

$$W_1 = \overline{\bigoplus_{i=1}^n V_i}^{\,9}$$

$$W_2 = \bigcap_{i=1}^n V_i$$

For $i \in [n]$, write P_i for the orthogonal projector on V_i. If Q is an orthogonal projector such that $[P_i, Q] = 0$ for all i, we also have that $[P_{W_1}, Q] = [P_{W_2}, Q] = 0$.

Proof. Suppose that $[P_i, Q] = 0$ holds for all i. We will write T to denote the closed subspace projected onto by Q. For every i, take $U_{i,1}$, $U_{i,2}$ as in the previous proposition. We then have

$$W_1 = \overline{\left(\bigoplus_{i=1}^n U_{i,1}\right)} \oplus \overline{\left(\bigoplus_{i=1}^n U_{i,2}\right)}$$

Since $U_{i,1} \subseteq T$ and $U_{i,2} \subseteq T^{\perp}$ for every i and since T and $T^{\perp}$ are both closed, the two summands in the above expression are subspaces of T and $T^{\perp}$, respectively, so that the previous proposition implies that $[P_{W_1}, Q] = 0$. Similarly, since

$$W_2 = \left(\bigcap_{i=1}^n U_{i,1}\right) \oplus \left(\bigcap_{i=1}^n U_{i,2}\right)$$

We get that $[P_{W_2}, Q] = 0$ for the same reason. □

Fixing the register size k, first suppose that we have a protocol in the commuting operators model for f, specified by a Hilbert space $\mathscr{H}$, some initial state $|\psi\rangle \in H$, measurement operators $\{A_c^x\}_{x \in X, c \in [k]}$ and $\{B_r^{y,c}\}_{y \in Y, c \in [k], r \in R}$ for Alice and Bob, respectively. We can turn this into a protocol for EQ_{G_f} where Alice does as in the previous protocol and where, given a vertex $x' \in X$ and Alice's message c, if $\{P_i\}_{i=1}^n$ are all the projectors of the form $B_{f(x',y)}^{y,c}$ with $f(x',y) \neq \square$, setting V_i to be the image of P_i and setting $W = \cap_{i=1}^n V_i$, Bob measures according to the projective measurement $\{P_W, \mathbb{I} - P_W\}$. By the corollary, this projective measurement will still commute with Alice's measurements. This shows that the communication complexity of EQ_{G_f} is no greater than

[9] Note that a closure is required here, as direct sums do not preserve closure in general in infinite dimensions.

that of f; that they are equal follows along the same lines, by using the first part in the above corollary.

4.2.1.3. The proof for the Cleve-Buhrman model in the vanishing-error case. In this subsection, we complete the proof of theorem 4.2.2 by showing the case of the zero-error case in the Cleve-Buhrman model. Again, the proof is conceptually identical to the one in the classical case, but we need to work a bit harder to deal with the fact that the protocols we are dealing with are no longer exact.

Lemma 4.2.7. *Let ρ be a mixed state on some finite-dimensional Hilbert space $\mathscr{H}$, and let U,V be subspaces of $\mathscr{H}$. It holds that*

$$\mathrm{tr}(P_{(U\cap V)^\perp}\rho) \leq \mathrm{tr}(P_{U^\perp}\rho) + \mathrm{tr}(P_{V^\perp}\rho)$$

PROOF. We prove the result in the case where ρ is pure: the general case then follows from the spectral theorem and the linearity of the trace. Suppose that $\rho = |\psi\rangle\langle\psi|$. $|\psi\rangle$ has a unique decomposition of the form:

$$|\psi\rangle = |\psi_{U\cap V}\rangle + |\psi_{U\cap V^\perp}\rangle + |\psi_{U^\perp\cap V}\rangle + |\psi_{U^\perp\cap V^\perp}\rangle$$

It can then be seen that

$$\begin{aligned}
\mathrm{tr}(P_{(U\cap V)^\perp}\rho) &= \|P_{(U\cap V)^\perp}|\psi\rangle\|^2 \\
&= \||\psi_{U\cap V^\perp}\rangle\|^2 + \||\psi_{U^\perp\cap V}\rangle\|^2 + \||\psi_{U^\perp\cap V^\perp}\rangle\|^2 \\
&\leq \||\psi_{U\cap V^\perp}\rangle\|^2 + \||\psi_{U^\perp\cap V}\rangle\|^2 + 2\||\psi_{U^\perp\cap V^\perp}\rangle\|^2 \\
&= \|P_{U^\perp}|\psi\rangle\|^2 + \|P_{V^\perp}|\psi\rangle\|^2 \\
&= \mathrm{tr}(P_{U^\perp}\rho) + \mathrm{tr}(P_{V^\perp}\rho)
\end{aligned}$$

As desired. □

Proposition 4.2.8. *Let H be a finite graph, let $\varepsilon \geq 0$, and suppose that $\{\rho_v\}_{v\in V(H)}$ are mixed states on some finite-dimensional Hilbert space $\mathscr{H}$ such that, for all $(v,w) \in E(H)$,*

$$D_{tr}(\rho_v, \rho_w) \geq 1 - \varepsilon$$

There exist subspaces $\{V_v\}_{v\in V(H)}$ of $\mathscr{H}$ such that, for all $(v,w) \in E(H)$, V_v and V_w are orthogonal and such that, for all v,

$$\mathrm{tr}(P_{V_v}\rho_v) \geq 1 - 2(\deg v)\varepsilon$$

PROOF. We identify $V(H)$ with $[|V(H)|]$. For every $e = (v,w) \in E(H)$ with $v < w$, take $\{P_0^e, P_1^e\}$ to be the Helstrom (projective) measurement for optimally distinguishing between ρ_v and ρ_w with a uniform prior, which, by hypothesis, has error probability at most ε. This means that:

$$\mathrm{tr}(P^r_1\rho_v), \mathrm{tr}(P^r_0\rho_w) \leq 2\varepsilon$$

Let V^e_b be the subspace projected on by P^e_b. Given a vertex $v \in V(H)$, set:

$$V_v = \bigcap_{\substack{e=(u,w)\in E(H)\\ u<w\\ v\in e}} V^e_{\delta_{v,u}}$$

For $(v,w) \in E(H)$ with $v < w$, $V_v \subseteq V^e_0$ and $V_w \subseteq V^e_1 = \left(V^e_0\right)^{\perp}$, so V_v and V_w are orthogonal. Also, iterating the previous lemma gives that

$$\mathrm{tr}(P_{V_v}\rho_v) = 1 - \mathrm{tr}(P_{V_v^{\perp}}\rho_v) \geq 1 - 2(\deg v)\varepsilon$$

As desired. □

We show:

Proposition 4.2.9. *An ε-approximate protocol for f with finite-dimensional entanglement and communication k can be turned into a $2|X||R|\varepsilon$-approximate protocol for EQ_{G_f} with finite-dimensional entanglement and communication k.*

PROOF. Take such a protocol, which we can assume to be in standard form. For $x \in X, c \in [k]$, let $\rho_{x,c} \in D(B)$ be the reduced density matrix on Bob's side if Alice's input was x and she sent communication c to Bob after the result of her measurement. Since the protocol is in standard form and is ε-approximate, for every $(x,x') \in E(G_f)$,

$$D_{\mathrm{tr}}(\rho_{x,c}, \rho_{x',c}) \geq 1 - |R|\varepsilon$$

For every $y \in X$, $c \in [k]$, by the previous proposition, there exists a subspace $V_{y,c}$ of B such that:

$$\mathrm{tr}(P_{V_{y,c}}\rho_{y,c}) \geq 1 - 2|R||X|\varepsilon$$

And, for every $x \in |X|$ with $(x,y) \in E(G_f)$,

$$\mathrm{tr}(P_{V_{y,c}}\rho_{x,c}) \leq 2|R||X|\varepsilon$$

This means that, if Alice does exactly as in the original protocol and if, on input $y \in X$ and on having received c from Alice, Bob measures according to the projective measurement $\{P_{V_{y,c}^{\perp}}, P_{V_{y,c}}\}$, the result will be a $2|R||X|\varepsilon$-approximate protocol for EQ_{G_f}, as desired. □

4.2.2. The complexities of promise equality problems

In this subsection, we look at the communication complexity of EQ_G in the various models for a generic graph G.

In the classical case, as we have seen, an exact protocol for G with communication k will exist if and only if it is possible to assign messages $c_x \in [k]$ to every $x \in V(G)$ in such a way that whenever $(x,y) \in E(G)$, we have that $c_x \neq c_y$. By definition, this is possible if and only if k is greater or equal to G's chromatic number $\chi(G)$.

The case of quantum communication and no prior entanglement is rather similar, where the messages are quantum states on a system of dimension k instead of integers and with the condition that states assigned to adjacent vertices must be perfectly distinguishable, or equivalently, orthogonal. By the spectral theorem, any mixed state may be realised as a probability distribution over some set of pure states, so that we may take the states sent by Alice to Bob to be pure. In other words, G will have a quantum protocol with communication k if and only if it is possible to assign unit vectors $|\psi_x\rangle \in \mathbb{C}^k$ to every $x \in V(G)$ so that $\langle\psi_x|\psi_y\rangle = 0$ whenever $(x,y) \in E(G)$. The smallest k for which this is possible is known as the **orthogonal rank** of G and is denoted by $\xi(G)$. Additionally, the representation is said to be **flat** if, for every x, the components of $|\psi_x\rangle$ in the standard basis all have modulus $\frac{1}{\sqrt{k}}$: the smallest k for which G has a flat orthogonal representation in $\mathbb{C}^k$ is called the **flat orthogonal rank** of G, and is denoted by $\xi'(G)$. While it is not obvious that the previous definition even makes sense because it is not immediately clear that all graphs even have a flat orthogonal representation, this is so, and we even have:

Lemma 4.2.10. *For all graphs G,*

$$\xi(G) \leq \xi'(G) \leq \chi(G)$$

PROOF. The first inequality is clear. Also, a k-colouring of G can be turned into a flat orthogonal representation of G in $\mathbb{C}^k$ by mapping the colours to the columns of the quantum Fourier transform acting on $\mathbb{C}^k$, i.e. letting ζ be a primitive k-th root of unity and given $c \in [k]$, we set

$$|\psi_c\rangle = \frac{1}{\sqrt{k}} \sum_{j \in [k]} \zeta^{cj} |j\rangle$$

And if colour c is assigned to $x \in V(G)$, we assign the vector $|\psi_c\rangle$ to x. □

In the other direction, we have the following, which follows from theorem 3.2.3:

Proposition 4.2.11 ([53]). *For any graph G, we have that* $\xi(G) = \Omega(\log \chi(G))$.

As for the Cleve-Buhrman model, for a given entanglement model $t \in \{qf, qc\}$, we will write $\kappa_t(G)$ (the communication number of the graph) to denote the smallest k for which G has an exact protocol in that model with communication k. We will refer to such a protocol as a k-protocol from now on. We will also write $\kappa_{\overline{q}}(G)$ to mean the smallest k for which, for every $\varepsilon > 0$, G has a finite-dimensional ε-exact protocol. Clearly,

$$\kappa_{qc}(G) \leq \kappa_{\overline{q}}(G) \leq \kappa_{qf}(G) \leq \chi(G)$$

As was seen in the previous section, a k-protocol is specified by a Hilbert state $\mathscr{H}$, a state $|\psi\rangle$, measurements $\{A_c^x\}_{x \in V(G), c \in [k]}$ and $\{B_b^{y,c}\}_{y \in V(G), c \in [k], b \in \{0,1\}}$ for Alice and Bob, respectively. We further restrict $\mathscr{H}$ to be finite-dimensional and $\mathscr{H}$ to be a composite system in the finite-dimensional setting. We will often only state our results for the general commuting operators model, as the fact that $\mathscr{H}$ is finite-dimensional and in tensor-product form is often irrelevant and the proof for the commuting operators model works just as well for the finite-dimensional setting.

With respect to the comparison between the communication numbers and the orthogonal rank, we have the following:

Proposition 4.2.12. *For any graph G, we have*

$$\kappa_{qf}(G) \leq \xi(G)^2$$

Proof. Given that G has an orthogonal representation $\{|\psi_x\rangle\}$ in $\mathbb{C}^k$, by the qudit teleportation protocol (see Khatri and Wilde ([**94**]), subsection 3.3.2), Alice can teleport $|\psi_x\rangle$ to Bob at the cost of communicating a register of size k^2, which Bob can then measure according to the measurement $\{|\psi_y\rangle\langle\psi_y|, \mathbb{I} - |\psi_y\rangle\langle\psi_y|\}$ to determine whether $x = y$. □

With respect to the flat orthogonal rank, we can do better than this. The following is a folklore construction, and is a straightforward extension of the work of Pati ([**34**]).

Proposition 4.2.13. *For any graph G, we have*

$$\kappa_{qf}(G) \leq \xi'(G)$$

Proof. Suppose that $k = \xi'(G)$, and let $\{|\psi_x\rangle\}$ be a flat orthogonal representation of G in $\mathbb{C}^k$. The protocol will be the same as in the proof of the previous proposition, except the teleportation will be performed more efficiently. Fix a particular Alice input x and write

$$|\psi_x\rangle = \sum_{i \in [k]} \alpha_i |i\rangle$$

Let ζ be a primitive k-th root of unity and, for $j \in [k]$, write

$$|\phi_j\rangle = \sum_{i \in [k]} \zeta^{ij} \alpha_i |i\rangle$$

It can be seen that because the α_i all have modulus $\frac{1}{\sqrt{k}}$, we have $\langle\phi_j|\phi_{j'}\rangle = \delta_{j,j'}$. The teleportation protocol runs as follows. Alice and Bob start out by sharing a maximally entangled state $|\Phi\rangle \in$

$\mathbb{C}^k \otimes \mathbb{C}^k$. Alice measures her share of the state according to the basis $\overline{|\phi_1\rangle}, ..., \overline{|\phi_k\rangle}$, getting $j \in [k]$. The reduced state on Bob's part is now $|\phi_j\rangle$. Alice sends j to Bob, and Bob performs the unitary transformation $|i\rangle \mapsto \zeta^{-ij}|i\rangle$ to get back $|\psi_x\rangle$. □

More will be said about the comparison between the two in subsection 4.2.7.

We also give the following easy result implies that for any graph, if one of the previously discussed parameters is equal to two, so are all the others:

Proposition 4.2.14. *Suppose that either $\kappa_{qc}(G) = 2$ or $\xi(G) = 2$. Then $\chi(G) = 2$, or equivalently, G is bipartite.*

PROOF. First suppose that $\kappa_{qc}(G) = 2$. Recasting the corresponding communication problem as a nonlocal game, we see that Alice and Bob's outputs are both binary. Since the game has a perfect quantum strategy, it also has a perfect classical strategy, by theorem 2.3.3, which gives that $\chi(G) = 2$.

Suppose next that $\xi(G) = 2$, and let $x_1, x_2, \dots, x_k$ be a cycle in G. If $|\psi\rangle_x$ is an orthogonal representation of G in $\mathbb{C}^2$, it is easy to see that $|\psi_{x_1}\rangle$, $|\psi_{x_3}\rangle, \dots$, must all be collinear and orthogonal to $|\psi_{x_2}\rangle$, $|\psi_{x_4}\rangle$, which must also all be collinear. This forces k to be even, meaning that G is bipartite. □

We end this subsection by stating useful lemmas. We begin by stating the following standard results:

Lemma 4.2.15. *Let $\mathscr{H}$ be a Hilbert space, $|\psi\rangle \in \mathscr{H}$ and A,B be positive operators on $\mathscr{H}$. The operator AB is positive if and only if A and B commute.*

Lemma 4.2.16. *Let $\mathscr{H}$ be a Hilbert space, $|\psi\rangle \in \mathscr{H}$ and C a positive operator on $\mathscr{H}$. Then, $\langle\psi|C|\psi\rangle = 0$ if and only if $C|\psi\rangle = 0$.*

We now show:

Lemma 4.2.17. *Take a commuting k-protocol for G. For all $c \in [k]$ and for all $(x,y) \in E(G)$, we have:*

$$\langle\psi|A_c^x A_c^y|\psi\rangle = 0$$

PROOF. By the exactness of the protocol, for all $c \in [k]$ and all $(x,y) \in E(G)$, we have that $\langle\psi|A_c^x B_0^{x,c}|\psi\rangle = \langle\psi|A_c^x B_1^{y,c}|\psi\rangle = 0$, since these quantities correspond to the probabilities of Alice sending c given that her input was x and Bob either replying with '0' given that his output was x

or replying with ‘1’ given that his output was y, respectively. Using lemmas 4.2.15 and 4.2.16, we get

$$\begin{aligned}\langle\psi|A_c^x A_c^y|\psi\rangle &= \langle\psi|A_c^x(B_0^{x,c}+B_1^{x,c})A_c^y|\psi\rangle \\ &= (\langle\psi|A_c^x B_0^{x,c})A_c^y|\psi\rangle + \langle\psi|A_c^x(B_1^{x,c}A_c^y|\psi\rangle) \\ &= 0\end{aligned}$$

As desired. □

Next, we pin down the communication complexity of complete graphs and prove a crucial property of protocols for complete graphs. We write K_k to mean the complete graph on k vertices, and we write $\omega(G)$ [10] to mean the clique number of G. Note that the following two lemmas also follow from proposition 4.2.32.

Lemma 4.2.18. *For all k, we have that $\xi(K_k) = k$. In particular, $\omega(G) \leq \xi(G)$ for all G.*

PROOF. By elementary linear algebra, any collection of nonzero, pairwise orthogonal vectors is linearly independent. The result follows. □

We now show the equivalent result for the Cleve-Buhrman model. In fact, the proof shows something slightly stronger which will be crucial for proving that the various entanglement models are all distinct in communication complexity. It should be noted that something very similar to this was already shown by [**76**] for general nonsignalling correlations.

Lemma 4.2.19. *For all k, we have that $\kappa_{qc}(K_k) = k$. In particular, $\omega(G) \leq \kappa_{qc}(G)$ for all graphs G. Furthermore, for all $x, c, c' \in [k]$ with $c \neq c'$, any commuting k-protocol for K_k satisfies the following:*

$$B_1^{x,c'}A_c^x|\psi\rangle = 0$$

PROOF. That $\kappa_{qc}(K_k) \leq k$ is clear, as Alice can simply send her input to Bob. We now show that $\kappa_{qc}(K_k) \geq k$. Take a commuting k'-protocol for K_k. Given $c \in [k']$, let $\{E_x^c\}_{x\in[k]}$ be the following measurement, to be applied by Bob upon the reception of c: for $x = 1, \ldots, k$, measure according to $\{B_0^{x,c}, B_1^{x,c}\}$; if 1 was obtained, output x; otherwise, move on to $x+1$. If 0 was obtained in all cases (which can never happen), output k. These operators will commute with Alice’s operators because Bob’s original measurement operators did, and, by the exactness of the protocol (the result of the measurement is certain each time, and so does not disturb the state), this procedure will recover Alice’s input with certainty, meaning that $\langle\psi|A_c^x E_{x'}^c|\psi\rangle = 0$ for all $x \neq x'$.

[10] Not to be confused with the value of a nonlocal game.

We compute:

$$\begin{aligned}
k &= \sum_{x\in[k],c\in[k']} \langle\psi|A_c^x|\psi\rangle \\
&= \sum_{x\in[k],c\in[k']} \langle\psi|A_c^x E_x^c|\psi\rangle \\
&\leq \sum_{x\in[k],c,c'\in[k']} \langle\psi|A_{c'}^x E_x^c|\psi\rangle \\
&= \sum_{x\in[k],c\in[k']} \langle\psi|E_x^c|\psi\rangle \\
&= \sum_{c\in[k']} \langle\psi|\psi\rangle \\
&= k'
\end{aligned}$$

From whence it follows that $k' \geq k$. Also, if $k = k'$, the third inequality must visibly be an equality, and since all the added cross-terms are nonnegative, they must all be zero, i.e. $\langle\psi|A_{c'}^x E_x^c|\psi\rangle = 0$ for all x,c,c' with $c \neq c'$. Since $E_c^1 = B_1^{1,c}$, the conclusion follows for $x = 1$ by invoking lemmas 4.2.15 and 4.2.16. Since there is nothing special about $x = 1$, the proof is complete. □

In the ε-exact case, using proposition 4.2.8 to produce the measurement operators E_c^x used in the above lemma, we can also show along the same lines:

Lemma 4.2.20. *We have that $\kappa_{\overline{q}}(K_k) = k$. Furthermore, any ε-exact k-protocol for K_k in standard form in the finite-dimensional setting satisfies, for all x,c,$c' \in [k]$ with $c \neq c'$,*

$$D_{tr}(\rho_{x,c},\rho_{x,c'}) \geq 1 - O(\varepsilon)$$

where the constant hidden in the O notation depends on k.

4.2.3. The quantum chromatic numbers and their properties

In this subsection, we define the quantum chromatic numbers and survey the known literature about them. It will turn out that these parameters are very closely tied to the communication numbers. Given a graph G and a parameter $k \in \mathbb{N}$, consider the following nonlocal game, which is referred to as a **colouring game** ([**53**]). Alice and Bob are each given vertices x,$y \in V(G)$ with the promise that they are either identical or adjacent, and they are requested to output a,$b \in [k]$ under the condition that $a = b$ if and only if $x = y$. For each of the models, we are interested in the smallest k such that this game has a perfect strategy. Classically, it is not too hard to see that perfect strategies are in direct correspondence with colourings of G, so that this k is simply the chromatic number of the graph $\chi(G)$. In the quantum realm, for each of the entanglement models t, the smallest such k is called the **quantum chromatic number** of the graph and will be denoted by $\chi_t(G)$. We will also consider the approximate quantum chromatic number $\chi_{\overline{q}}$, which corresponds to the smallest k for which, for every $\varepsilon > 0$, the k-colouring game corresponding to the graph has an ε-perfect strategy. By analogy, a perfect strategy for the colouring game with k colours of G is referred to as a *k-colouring* of G.

Colouring games are of interest to us because perfect strategies for colouring games give rise to exact protocols for promise equality problems in the following way: if, for some number of colours k and some entanglement flavour t, a given graph admits a k-colouring, then this k-colouring may be turned into an exact protocol with communication k in which Alice and Bob play the colouring game with the perfect strategy and Alice sends the colour she obtained to Bob, who then compares it with the colour he obtained to determine whether the inputs were identical or adjacent. We therefore have that $\kappa_t(G) \leq \chi_t(G)$ for all models. We will come back to the interplay between the two parameters later in this section.

We now describe the three most common mechanisms for constructing graphs whose corresponding colouring games are interesting quantumly, i.e. $\chi_{qc}(G) < \chi(G)$. The first one is due to [**53**] and uses a construction based on quaternions and octonions. While it is not useful for obtaining large separations between the quantum and classical chromatic numbers, it is useful for building small graphs with $\chi_{qf}(G) < \chi(G)$.

Theorem 4.2.21. *For $k \in \{2,4,8\}$, if a graph G has an orthogonal representation in $\mathbb{R}^k$, then $\chi_{qf}(G) \leq k$.*

The following result is also due to [**53**], although it is a rather straightforward generalisation of what is done in [**47**]:

Theorem 4.2.22.

$$\chi_{qf}(G) \leq \xi'(G)$$

PROOF. This follows at once from the combination of propositions 4.2.13 and 4.2.34. □

Finally, we have the following result by Harris ([**101**]), building upon the work of Ji ([**69**]). Recall that a nonlocal game is said to be synchronous if the input sets are the same, the output sets are the same, and when Alice and Bob are given the same input, they win if and only if their outputs are the same.

Theorem 4.2.23 ([**101**])**.** *There exists an efficiently computable mapping from synchronous nonlocal games R to graphs G such that:*

(1) R has a perfect finite-dimensional strategy if and only if $\chi_{qf}(G) = 3$
(2) R has tensor-product value one if and only if $\chi_{\overline{q}}(G) = 3$
(3) R has a perfect commuting operators strategy if and only if $\chi_{qc}(G) = 3$

In particular, the results which were discussed in section 2.5 have the following corollaries (though, in some cases, the games need to be synchronised):

Corollary 4.2.24. *Given a graph G and an entanglement model t, the problem of determining whether $\chi_t(G) = 3$ is:*

(1) $t = qf$: Complete for RE

(2) $t = \overline{q}$: Complete for Π_2

(3) $t = qc$: Complete for coRE

We also have:

Corollary 4.2.25. *The parameters $\chi_{qc}, \chi_{\overline{q}}, \chi_{qf}$ are all distinct.*

Finally, and importantly for us, corollary 4.2.24 implies:

Corollary 4.2.26. *There does not exist a computable bound on the amount of entanglement that might be required by a finite-dimensional 3-colouring of a given graph.*

One might wonder if the above theorem and corollaries still hold if 3 is replaced by k for some $k \geq 4$. For most graph parameters, the generalisation from $k = 3$ to arbitrary $k \geq 3$ is automatic thanks to what [**74**] calls the 'suspension operation' for graphs, which consists in adding a $(k-2)$-clique to the graph and making every vertex in that clique adjacent to every other vertex in the graph, thereby lifting the $k = 3$ case to the general case. Most regrettably, this is known not to work for the quantum chromatic numbers: Mančinska and Roberson ([**77**]) provide a counterexample which will be presented in section 4.2.7. Therefore, while the above results are surely also true for general k-colourings, this lacks a proof at present.

We end by mentioning an intriguing consequence of the above corollary. We make the following definition. Given a graph parameter f that is monotonous, i.e. $f(G) \leq f(H)$ if $G \subseteq H$ (this is verified by all the graphs parameters in this section), we say that f has the **de Bruijn-Erdős property** if, given an infinite graph G, there exists a finite subgraph G' of G such that $f(G) = f(G')$. The original de Bruijn-Erdős theorem states that the chromatic number has (what we call) the de Bruijn-Erdős property, and Gottschalk's ([**1**]) proof of this based on Tychonoff's theorem can straightforwardly be adapted to show that ξ does as well, by virtue of the compactness of the unit sphere in $\mathbb{C}^k$. It is however easy to show that χ_{qf} does not have the de Bruijn-Erdős property: indeed, for every n, letting G_n be some 3-colourable graph with the property that any 3-colouring of G_n requires an entangled state of local dimension at least n, which must exist by corollary 4.2.26, and taking G to be the union of all the G_n, we see that G cannot have a finite-dimensional 3-colouring, even though any finite subgraph of G clearly does. On the other hand, it strikes us as conceivable that $\chi_{\overline{q}}$ and χ_{qc} do have the de Bruijn-Erdős property, since the underlying correlation sets are closed, and the aforementioned proof of Gottschalk fundamentally relies on the fact that a finite set is compact. While it is unlikely that anything useful will come out of this, we leave this as a fun question for the reader to think about.

4.2.4. Separating the exact quantum and classical communication complexities: the distributed Deutsch-Jozsa problem

Proposition 4.2.11 implies that the separation between the chromatic number and the orthogonal rank can never be larger than exponential. It turns out that this is tight, in that an exponential separation does arise for some families of graphs, most notably the Hadamard graph. The resulting communication complexity problem was first introduced by [**29**] as a distributed version of the classical Deutsch-Jozsa problem ([**16**]): this was already discussed in subsection 3.2.3. Our presentation will be simpler and cleaner than theirs thanks to the benefits of hindsight. It should be emphasised that while none of what follows is new, to the best of our knowledge, no completely satisfactory exposition of these results exists, so it seems worthwhile to go over them.

For a given number n, let G_n be the graph with vertex set $\{0,1\}^n$ such that, for any $x,x' \in \{0,1\}^n$, $(x,x') \in E(G_n)$ if

$$\#\{i \in [n] \mid x_i = x'_i\} = \frac{n}{2}$$

The graphs G_n are known in the literature as the **Hadamard graphs**. Clearly, in the case that n is odd, the graphs obtained in this way are edgeless and therefore not very interesting. If $n \equiv 2 \mod 4$, for any cycle $x_1, x_2, \ldots, x_k$ of G_n, we have that

$$0 = \sum_{i=1}^{n}(x_{1i} - x_{2i}) + \sum_{i=1}^{n}(x_{2i} - x_{3i}) + \ldots + \sum_{i=1}^{n}(x_{ki} - x_{1i})$$

Since each of the k terms is odd, k must be even. This shows that in this case, G_n is bipartite and all the previously discussed graph parameters are equal to two. This means that the only G_n's that are interesting to us are the ones where n is a multiple of 4.

We may construct an orthogonal representation of G_n in $\mathbb{C}^n$ by setting, for every vertex v,

$$|\psi_v\rangle = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(-1)^{v_i}|i\rangle$$

By construction, we have that $\langle\psi_u|\psi_v\rangle = 0$ for all $(u,v) \in E(G)$. Notice also that this orthogonal representation is flat. Therefore, in combination with theorem 4.2.22, we have:

Proposition 4.2.27. *For all n, $\chi_{qf}(G_n), \xi(G) \leq n$.*

In the classical realm, we have the following hard combinatorial theorem due to Frankl and Rödl. A pairwise nonadjacent collection of vertices of a given graph G is said to be an independent set, and the independence number of G, denoted $\alpha(G)$, corresponds to the size of the largest independent set of G. Since, in a coloring of G, the subset of vertices that are assigned a given color forms an independent set, we have the bound $\chi(G)\alpha(G) \geq |V(G)|$.

Theorem 4.2.28 ([**11**])**.** *There exists some $\varepsilon > 0$ such that it holds that $\alpha(G_{4n}) \leq (2-\varepsilon)^{4n}$ for all n. In particular, for all n, $\chi(G_{4n}) \geq \frac{2^{4n}}{\alpha(G_{4n})} \geq \left(\frac{2}{2-\varepsilon}\right)^{4n}$.*

This means that the G_{4n} exhibit an exponential separation between the classical and quantum communication complexities. In particular, this implies that there exists some smallest n_0 such that for all $n \geq n_0$, $\chi(G_{4n}) > 4n \geq \xi(G_{4n}), \chi_q(G_{4n})$, which was worked out by [**46**] to be equal to 3.

From the point of view of communication complexity, this gives an exponential separation between the classical and quantum models in the one-way, exact setting. In fact, we have the following:

Proposition 4.2.29 ([**74**])**.** *For all graphs G, we have that $C_E(EQ_G) = C_E^1(EQ_G)$.*

This means that we also have an exponential separation between the classical and quantum models, even in the presence of bidirectional communication. Note however that this separation is nullified when some error is allowed.

It is worth pointing out that it is unknown at present whether $\kappa_{qc}(G_{4n}) = \xi(G_{4n}) = 4n$ for all n. However, this would follow from lemmas 4.2.18 and 4.2.19 combined with the following widely believed conjecture:

Conjecture 4.2.30 (Hadamard)**.** *For all n, it holds that $\omega(G_{4n}) = 4n$.*

This is already known to be true in many special cases, such as when $4n$ is a power of two. We refer the curious reader to the survey of Hedayat and Wallis ([**5**]) for more information about this conjecture.

4.2.5. The Lovász ϑ number and the communication numbers

The ϑ number is an efficiently computable, real-valued graph parameter that was originally introduced by Lovász [**7**] with the purpose of upper bounding the Shannon capacity of a graph. Its usefulness mainly stems from the fact that it bounds a number of graph parameters that are otherwise hard to compute or estimate. As is the case in most of the literature, what is of actual interest to us is the value of the theta number of the complement of the graph and not that of the graph itself, so that we will write $\overline{\vartheta}(G)$ to mean $\vartheta(\overline{G})$. We have:

Theorem 4.2.31. *(* [**9**], [**78**], [**70**], [**74**]*) For any graph G,*

$$\omega(G) \leq \overline{\vartheta}(G) \leq \chi(G), \chi_{qf}(G), \chi_{qc}(G), \xi(G)$$

We adapt the proof that $\overline{\vartheta}(G) \leq \xi(G)$ from [**74**] to show that the lower bound holds for κ_{qc} (and, by extension, for all chromatic and communication numbers) as well. To this end, we will use the following standard formulation of the ϑ number as a semidefinite program:

$$\begin{aligned} \overline{\vartheta}(G) = \quad & \min t \\ & \text{s.t. } M_{i,i} = t-1, && \text{for } i = 1, \dots, n \\ & M_{i,j} = -1, && \text{for } (i,j) \in E(G) \\ & M \in \mathbb{R}^{|V(G)| \times |V(G)|}, M \succeq 0 \\ & t \in \mathbb{R} \end{aligned}$$

We show:

Proposition 4.2.32.

$$\overline{\vartheta}(G) \leq \kappa_{qc}(G)$$

PROOF. Let $k = \kappa_{qc}(G)$, and fix a corresponding commuting k-protocol. Let $\mathscr{H}' \leq \mathscr{H}$ be the finite-dimensional subspace spanned by all the vectors of the form $A_c^x |\psi\rangle$ as well as $|\psi\rangle$ itself. For every $x \in V(G)$, define $|v_x\rangle \in \mathbb{C}^k \otimes \mathscr{H}'$ by:

$$|v_x\rangle = \sum_{c=1}^{k} |c\rangle \otimes \left(\sqrt{k} A_c^x |\psi\rangle - \frac{1}{\sqrt{k}} |\psi\rangle \right)$$

And define M' to be the Gram matrix of the $|v_x\rangle$. We see that for every vertex x:

$$M'_{x,x} = \langle v_x | v_x \rangle = \sum_{c=1}^{k} \left((k-2) \langle \psi | A_c^x | \psi \rangle + \frac{1}{k} \right) = k-1$$

Similarly, for all $(x,y) \in E(G)$,

$$M'_{x,y} = \langle v_x | v_y \rangle = \left(\sum_{c=1}^{k} \langle \psi | A_c^x A_c^y | \psi \rangle \right) - 2 + 1 = -1$$

Where we invoked 4.2.17. Taking M to be the real part of M', which is just as positive semidefinite as M', we have that (M,k) is a feasible solution for the semidefinite program for $\overline{\vartheta}(G)$ with corresponding objective value k, which implies that $\overline{\vartheta}(G) \leq k$, as desired. □

Note that the previous result subsumes most of theorem 4.2.31, as $\kappa_{qc}(G) \leq \chi_q(G), \chi_{qc}(G), \chi(G)$. It is also worth mentioning that the bulk of [**74**] consists in showing that theorem 4.2.29 fails badly in the Yao model. This is achieved by exhibiting a family of graphs $\{G_n\}$ with the property that the two-round exact quantum communication complexity of

EQ_{G_n} (when measured in bits) is exponentially smaller than $\log \overline{\vartheta}(G_n)$, which lower bounds the one-way, exact communication complexity of the corresponding equality problem in the Yao model (when measured in qubits). The above result shows that their conclusions hold in the Cleve-Buhrman model also.

4.2.6. The comparison between the quantum communication and chromatic numbers and separations between different flavours of entanglement in communication complexity

In this section, we revisit the problem of how the communication and chromatic numbers compare. Previously, we mentioned that, for all t, $\kappa_t(G) \leq \chi_t(G)$, and that no graphs are known that exhibit a separation between the two parameters. Actually, the only property of the communication numbers for which we have no proof for the chromatic numbers is proposition 4.2.12. In this section, we look at cases in which there is provably no separation. This will ultimately allow us to import the previously discussed constellation of results for general nonlocal games into communication complexity.

In this regard, we can show that if maximally entangled states and projective measurements are complete for one-way, exact communication complexity, then the parameters always coincide. Given a Hilbert space $\mathscr{H}$ and a C^*-algebra $\mathscr{C}$ of operators on $\mathscr{H}$, a state $|\psi\rangle \in \mathscr{H}$ is said to be **tracial** if, for any $E, F \in \mathscr{C}$, it holds that $\langle\psi|EF|\psi\rangle = \langle\psi|FE|\psi\rangle$. In the special case that H of the form $A \otimes B$ with A and B of the same dimension and if the operators in $\mathscr{C}$ are all of the form $E_A \otimes \mathbb{I}_B$, then maximally entangled states are all tracial (this can be seen from lemma 4.1.1), and if $\mathscr{C}$ is the collection of operators acting as the identity on B, then the tracial states are precisely the maximally entangled states. Therefore, in effect, the notion of a tracial state should be seen as a generalisation of that of a maximally entangled states beyond the tensor-product model.

We will need the following result:

Lemma 4.2.33 (Corollary 5.6 of [**70**], reformulated)**.** *A graph G has a commuting k-colouring if and only if there exists a Hilbert space $\mathscr{H}$, projective measurements $\{E_a^x\}_{x \in V(G), a \in [k]}$ and a state $|\psi\rangle \in \mathscr{H}$ that is tracial with respect to the $\{E_a^x\}$ such that, for all $(x,y) \in E(G)$ and all $a \in [k]$, $\langle\psi|E_a^x E_a^y|\psi\rangle = 0$. The same is true for finite-dimensional k-colourings if the Hilbert space is restricted to be finite-dimensional.*

In combination with the previous lemma, lemma 4.2.17 gives:

Proposition 4.2.34. *Given a graph G, let $k = \kappa_t(G)$, for $t \in \{qf, qc\}$. If G has a k-protocol in the setting t that uses a tracial state and projective measurements, then $\chi_t(G) = k$ also.*

It is a longstanding open problem to determine whether maximal entanglement and projective measurements are complete for pseudotelepathy, of which one-way, exact communication complexity is a special case (see, for example, [**79**] for a discussion of this problem). If this turned out to be the case, it would follow that the communication and chromatic numbers are equal in general [11]. Note that theorem 3.3.1 has no bearing on whether the hypothesis of the previous proposition is true always, as it is only relevant for the bounded-error setting.

We now show how to embed colouring games in one-way communication complexity without assuming that the above is true. Given graphs G and H, their *Cartesian product*, denoted $G\Box H$, is the graph with vertex set $V(G) \times V(H)$ and such that $(v_1, w_1) \sim (v_2, w_2)$ if one of the following holds:

(1) $v_1 = v_2$, and $w_1 \sim w_2$
(2) $v_1 \sim v_2$, and $w_1 = w_2$

Fixing a graph G and $k \in \mathbb{N}$ for the remainder of this section, we define $G' = G\Box K_k$. Since the communication and chromatic numbers of K_k are all k, by proposition 4.2.19, we see that the communication and chromatic numbers of G' are all no smaller than k. We have the following:

Proposition 4.2.35. *For any entanglement model t, $\chi_t(G) \leq k$ if and only if $\chi_t(G') = k$.*

PROOF. The backward direction is clear, since G' has G as a subgraph. For the forward direction, a k-colouring of G gives rise to a k-colouring of G' in the following way: given vertices (x,l) and (y,l') of G' as input, Alice and Bob use the k-colouring of G with inputs x,y to obtain colours $a',b' \in [k]$ and output $a = a' \oplus l$, $b = b' \oplus l'$, where $\oplus$ refers to a cyclic shift. If $x = y$ and $l = l'$, clearly, we will always have $a = b$. If $(x,l) \sim (y,l')$, either $x = y$ and $l \neq l'$, in which case $a' = b'$ and therefore $a \neq b$, or $x \sim y$ and $l = l'$, in which case $a' \neq b'$ and therefore $a \neq b$. In either case, the colouring condition is satisfied. □

The point of the above construction is that it allows us to embed the quantum chromatic numbers inside communication complexity. Specifically:

Theorem 4.2.36. *For all $t \in \{qf, \overline{q}, qc\}$,*

$$\chi_t(G) \leq k \iff \kappa_t(G') = k$$

In particular, corollaries 4.2.24, 4.2.25 and 4.2.26 hold true for the communication numbers as well.

[11] In an email exchange with Dan Stahlke, it transpired that this result was known to him, at least in the tensor-product setting.

We will break up the proof of this result in two parts. We begin by showing this in the exact case:

Theorem 4.2.37. *For $t \in \{qf, qc\}$, $\chi_t(G') = k$ if and only if $\kappa_t(G') = k$.*

PROOF. The forward direction follows from the fact that in every model, the communication number is upper bounded by the quantum chromatic number and the fact that both are lower bounded by k in this case. We now turn to the backward direction, which we will only show for $t = qc$, as the same argument works for $t = qf$ as well. Suppose then that $\kappa_{qf}(G') = k$, and take a corresponding exact k-protocol for G'. We will turn this protocol into a perfect k-colouring for G'. Alice's measurement operators in this colouring will be the same as in the k-protocol. We now construct Bob's measurement operators.

For every $y \in V(G)$, $l, c \in [k]$, define $V_{(y,l),c}$ as the subspace that is projected upon by $B_1^{(y,l),c}$, and set $\tilde{V}_{(y,l),c}$ to be the intersection of $V_{(y,l),c}$ and of all the $(V_{(y,l),c'})^{\perp}$ with $c \neq c'$. The correctness of the protocol gives that $A_c^{(y,l)} |\psi\rangle \in V_{(y,l),c}$. This in combination with lemma 4.2.19 gives that $A_c^{(y,l)} |\psi\rangle \in \tilde{V}_{(y,l),c}$. Setting $\tilde{B}_c^{(y,l)}$ to be the projection on $\tilde{V}_c^{(y,l)}$, corollary 4.2.6 implies that $\tilde{B}_c^{(y,l)}$ commutes with Alice's measurement operators. Since, for $c \neq c'$, $\tilde{B}_c^{(y,l)}$ and $\tilde{B}_{c'}^{(y,l)}$ are orthogonal and since the above construction gives that for all $x, y \in V(G), l, l', c, c' \in [k]$ such that either $x = y, l = l', c \neq c'$ or $(x,l) \sim (y,l'), c = c'$,

$$\langle\psi|A_c^{(x,l)} \tilde{B}_{c'}^{(y,l')}|\psi\rangle = 0$$

If, for every input (y,l), the $\tilde{B}_c^{(y,l)}$ summed to the identity, they would constitute a legitimate projective measurement and we would have a perfect k-colouring for G'. If not, pad them with an extra projector $\tilde{B}_{k+1}^{(y,l)}$ so that they do sum to the identity. If Alice and Bob's inputs were the same and if they measured according to the $\{A_c^{(y,l)}\}$ and the $\{\tilde{B}_c^{(y,l)}\}$, respectively, and if Alice obtained colour $c \in [k]$, since $A_c^{(y,l)} |\psi\rangle \in \tilde{V}_{(y,l),c}$, Bob would also obtain colour c with certainty. In particular, the outcome $k+1$ would never come up. By non-signalling, the distribution of Bob's result is independent of whatever Alice does, so that the outcome $k+1$ is impossible no matter what, and so we have built a perfect strategy for the k-colouring game for G'. □

The same argument can straightforwardly be adapted to the ε-exact case:

Theorem 4.2.38. *Any ε-exact finite-dimensional k-protocol for G' in standard form can be turned into an $O(\varepsilon)$-perfect finite-dimensional k-colouring for G'. In particular, we have that $\chi_{\overline{q}}(G') = k$ if and only if $\kappa_{\overline{q}}(G') = k$.*

PROOF. Again, the backward direction is the only one that needs a proof. As in the exact case, the new strategy for the k-colouring game will use the same measurement operators as the original protocol for Alice. We will now build Bob's measurement. By the correctness of the protocol and by lemma 4.2.20, for every $x,y \in V(G)$ and for every $l,l',c,c' \in [k]$ satisfying either of the following conditions:

(1) $(x,l) \sim (y,l')$ and $c = c'$
(2) $x = y, l = l', c \neq c'$

It holds that

$$D_{tr}(\rho_{(x,l),c}, \rho_{(y,l'),c'}) = 1 - O(\varepsilon)$$

Invoking proposition 4.2.8, we can cook up subspaces $V_{(x,l),c}$ of B such that, whenever (1) or (2) holds, $V_{(x,l),c}$ and $V_{(y,l'),c'}$ are orthogonal and such that, for every $x \in V(G), l,c \in [k]$,

$$\mathrm{tr}(P_{V_{(x,l),c}} \rho_{(x,l),c}) = 1 - O(\varepsilon)$$

For every $y \in V(G)$, $l,c \in [k]$, set $B_c^{(y,l)}$ to be the projector onto $V_{y,l,c}$. We see that, for every $x,y \in V(G)$ and every $l,l',c \in [k]$ with $(x,l) \sim (y,l')$, because $V_{(y,l'),c} \subseteq (V_{(x,l),c})^{\perp}$,

$$\mathrm{tr}(B_c^{(y,l')} \rho_c^{(x,l)}) \leq \mathrm{tr}((\mathbb{I} - B_c^{(y,l)}) \rho_c^{(x,l)}) = O(\varepsilon)$$

This means that, if $\{B_c^{(y,l)}\}_{c \in [k]}$ were to form a valid projective measurement (i.e. sum to the identity) for every input pair (y,l), we would have a $O(\varepsilon)$-perfect colouring for G'. If not, we could pad them with another projector so that they do sum to the identity, and the probability of Bob obtaining this last outcome is bounded by $O(\varepsilon)$, for the same reason as in the proof of theorem 4.2.37. Overall, we may bound the error probability of the resulting approximate colouring for G' with $O(\varepsilon)$ for any input pair. □

This completes the proof of theorem 4.2.36.

4.2.7. The relationship between the orthogonal rank and the quantum chromatic numbers

In this subsection, we look at how the orthogonal rank and the various chromatic numbers compare. While many graphs are known which exhibit a separation between the quantum and classical chromatic numbers, far fewer separations are known between the quantum chromatic number and the orthogonal rank. The reason for this is that almost all examples of graphs with $\chi_{qf}(G) < \chi(G)$ in the literature are obtained by finding a graph with $\xi(G) < \chi(G)$ and applying one of the first two results of subsection 4.2.3 to turn the orthogonal representation into a quantum colouring, so that the resulting graph often has $\chi_{qf}(G) = \xi(G)$ by construction. As it turns out, however, it can be shown that they are incomparable, meaning that $\chi_{qc}(G) > \xi(G)$ and $\chi_{qf}(G) < \xi(G)$ are both possible [12]. The first case was shown by Mančinska and Roberson ([**77**]):

[12] Though not at the same time! Also note that $\chi_{qf}(G) \neq \xi(G)$ for some graphs follows from the previously discussed fact that $\chi_q(G)$ is uncomputable while $\xi(G)$ is.

Theorem 4.2.39 ([77])**.** *There exists a graph on 13 vertices, G_{13}, with*

$$3 = \xi(G_{13}) < \chi_{qc}(G_{13}) = \chi(G_{13}) = 4$$

Furthermore, G_{13} has a real three-dimensional representation.

At present, no upper bound on $\chi_{qc}(G)$ is known with respect to $\xi(G)$, save the very weak one given by proposition 4.2.11. Proposition 4.2.12 gives that $\kappa_{qf}(G) \leq \xi(G)^2$, and it seems likely that the same holds for $\chi_{qf}(G)$, but this lacks a proof.

Consider the graph G_{14} obtained by tacking on a vertex to G_{13} and making it adjacent to every other vertex. It is not hard to see that $\xi(G_{14}) = 4$ and $\chi(G) = 5$, and theorem 4.2.21 gives a 4-colouring of G_{14}, so that $\chi_q(G_{14}) = 4$.

[77] also shows that the other direction can be obtained using Ji's reduction:

Theorem 4.2.40 ([77])**.** *Suppose that G was obtained from Ji's reduction. Then, $\chi(G) = 3$ if and only if $\xi(G) = 3$.*

Corollary 4.2.41. *There exists a graph G with $3 = \chi_{qf}(G) < \xi(G)$. Taking G to come from the magic square game, G has 57 vertices and the corresponding 3-colouring game can be won with two EPR pairs.*

Adding a vertex to the graph G from the previous corollary and making it adjacent to every other vertex, we obtain a graph G' with $\kappa_{qf}(G') \leq \chi_{qf}(G') \leq 4$ and $\xi(G') = 5$. This corresponds to a communication problem with an exact protocol using two bits of communication and prior entanglement but with no exact protocol using two qubits of communication and no prior entanglement.

We now describe a way to construct graphs exhibiting a separation between the quantum and classical chromatic numbers which does not rely on an orthogonal representation and is therefore suitable for obtaining graphs with $\chi_q f(G) < \xi(G)$. This material can also be found in [103]. We begin with the following definition:

Definition 2. *A (m,n)-**vector clump** is a collection of unit vectors $\{|\psi_{j,k}\rangle\}_{j\in[m],k\in[n]}$ in $\mathbb{C}^{mn}$ which are all pairwise orthogonal. Two (m,n)-clumps $\{|\psi_{j,k}\rangle\}$ and $\{|\psi'_{j,k}\rangle\}$ are said to be **orthogonal** if, for every $j, j' \in [m]$, it holds that*

$$\sum_{k\in[n]} \langle\psi_{j,k}|\psi'_{j',k}\rangle = 0$$

By analogy with the orthogonal rank, we define the **clump rank** of a graph $\xi_c(G)$ to be the smallest n for which, for some m, there exists an assignment of (m,n)-vector clumps to the vertices of G $\{|\psi^v_{j,k}\rangle\}_{v\in V(G),j\in[m],k\in[n]}$ in such a way that for all $(u,v)\in E(G)$, $\{|\psi^u_{j,k}\rangle\}$ and $\{|\psi^v_{j,k}\rangle\}$ are orthogonal.

We need the following lemma:

Lemma 4.2.42 ([**53**], reformulated)**.** *G has a finite-dimensional k-colouring if and only if, for some $n\geq 1$, there exist subspaces $\{V^v_l\}_{v\in V(G),l\in[k]}$ of $\mathbb{C}^{kn}$, all of dimension n, such that, for all $v,v'\in V(G)$, $l,l'\in[k]$, if either $v=v'$ and $l\neq l'$ or if $(v,v')\in E(G)$ and $l=l'$, we have that*

$$V^v_l \perp V^{v'}_{l'}$$

We can now show:

Proposition 4.2.43. *For any graph G,*

$$\chi_{qf}(G)\leq \xi_c(G)^2$$

Proof. Suppose that $\xi_c(G)=n$, and take $\{|\psi^v_{j,k}\rangle\}$ to be a corresponding assignment of clumps. Letting ζ be a primitive n-th root of unity, for every $v\in V(G), j\in[m], r,s\in[n]$, take $|\phi^v_{j,r,s}\rangle\in\mathbb{C}^{mn^2}$ to be

$$|\phi^v_{j,r,s}\rangle=\frac{1}{\sqrt{n}}\sum_{k=1}^{n}\zeta^{ks}\,|k\rangle\,|\psi^v_{j,k\oplus r}\rangle$$

For $j,j'\in[m]$ and $r,r',s,s'\in[n]$, we see that

$$\begin{aligned}
\langle\phi^v_{j,r,s}|\phi^v_{j',r',s'}\rangle &= \frac{1}{n}\sum_{k=1}^{n}\zeta^{k(s'-s)}\,\langle\psi^v_{j,k\oplus r}|\psi^v_{j',k\oplus r'}\rangle\\
&=\delta_{j,j'}\delta_{r,r'}\frac{1}{n}\sum_{k=1}^{n}\left(\zeta^{(s'-s)}\right)^k\\
&=\delta_{j,j'}\delta_{r,r'}\delta_{s,s'}
\end{aligned}$$

So that, for a fixed vertex v, the $|\phi^v_{j,r,s}\rangle$ form an orthonormal basis of $\mathbb{C}^{mn^2}$. For every $r,s\in[n]$, setting $V^v_{(r,s)}$ to be the subspace spanned by the $\{|\phi^v_{j,r,s}\rangle\}_{j\in[m]}$, we see that for $(u,v)\in V(G)$, for any choice of (r,s) and for every $j,j'\in[m]$,

$$\begin{aligned}
\langle\phi^u_{j,r,s}|\phi^v_{j',r,s}\rangle &= \frac{1}{n}\sum_{k=1}^{n}\langle\psi^u_{j,k\oplus r}|\psi^v_{j',k\oplus r}\rangle\\
&=\frac{1}{n}\sum_{k=1}^{n}\langle\psi^u_{j,k}|\psi^v_{j',k}\rangle\\
&=0
\end{aligned}$$

So that $V^{v}_{(r,s)} \perp V^{u}_{(r,s)}$. The conclusion follows from lemma 4.2.42. □

We now use the notion of a vector clump to build a graph G_{21} which exhibits a separation between the quantum and classical chromatic numbers. The graph G_{21} is obtained as the orthogonality graph of the 21 (2,2)-clumps listed in appendix A, i.e. G_{21} is the graph on 21 vertices such that two vertices are adjacent if and only if the corresponding (2,2)-clumps in the list are orthogonal. The list was generated by a computer using the following algorithm:

(1) List all (2,2)-clumps made up of vectors whose components in the standard basis are all either -1, 0 or 1 prior to normalization.

(2) Build the corresponding orthogonality graph G.

(3) Repeat a number of times:

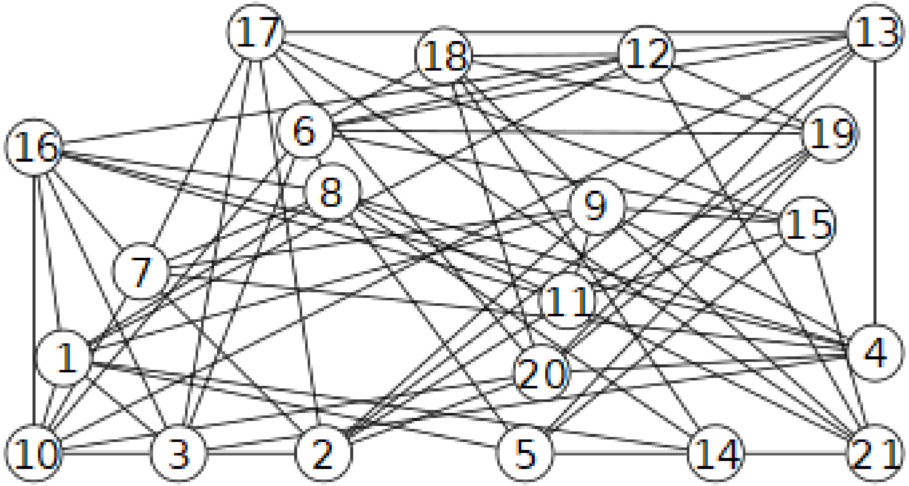
   (a) As long as it is possible to remove a vertex from G in such a way that the chromatic number of the resulting graph is greater than 4, do so, until a vertex-critical subgraph G' is obtained

   (b) Run G' through an optimiser to check if G' has an orthogonal representation in dimension 4: if it does, discard G', and if not, record it in memory.

(4) Output the clumps corresponding to the vertices of the smallest graph thereby obtained.

The graph G_{21} is depicted in figure 4.1.

Figure 4.1. A graphical representation of the graph G_{21}, drawn using the Julia package GraphPlot. The graph6 representation ([**105**]) of G_{21} is `TX_ac~QhaBO_TDaO@dDewW_gCd?WWI_c[?lg` .



Evidently, $\xi_c(G_{21}) = 2$, and therefore $\chi_{qf}(G_{21}) \leq 4$, by the previous proposition, and since the vertices 6, 12, 18, 19 form a clique, we have that $\chi_{qf}(G_{21}) = 4$. On the other hand, by the way G_{21} was built, we have $\chi(G_{21}) = 5$, and it is also true that $\xi(G_{21}) > 4$, so that $\xi(G_{21}) = 5$. A computer-assisted proof of this fact is described in [**103**]. In particular, G_{21} is by far the smallest graph known to exhibit a separation between the quantum and classical chromatic numbers which cannot be obtained by appealing to theorem 4.2.21. Conditionally on the truth of the above conjecture, since we know that the quantum and classical chromatic numbers coincide for all graphs on 13 vertices or less, G_{21} is probably not very far off from being the smallest possible example of a graph exhibiting the given separation.

We summarise the properties of G_{21} in the following table:

Table 4.1. Some properties of the graph G_{21}

Property	Value
# of vertices	21
# of edges	72
ω	4
χ_t, κ_t, $t \in \{qc, \overline{q}, qf\}$	4
$\overline{\vartheta}$	4
χ_f	4
χ	5
ξ	5

4.2.8. Direct sums in one-way, zero-error communication complexity

We end this section by discussing the problem of direct sums in one-way, zero-error communication complexity. Given a graph G which exhibits a gap between two models, one might be tempted to try to amplify this gap by having Alice and Bob solve multiple instances of EQ_G, in a way reminiscent of the parallel repetition of nonlocal games which was discussed in section 2. Namely, we could give Alice and Bob $(x_1,x_2,\dots,x_n)$ and $(y_1,y_2,\dots,y_n)$, respectively, where the x_i and the y_i are vertices of G, such that, for every i, either $x_i = y_i$ or $x_i \sim y_i$, and require Bob to output $b_1,\dots,b_n$ with $b_i = I[x_i = y_i]$ for every i. Looking at the graph G_{21} defined in the previous section, $2n$ bits of communication are sufficient to solve this problem with prior entanglement, while one could hope that roughly $(\log_2 5)n$ qubits of communication are required without prior entanglement. As we will see, this intuition proves to be incorrect.

Actually, the above direct sum problem, while the most natural, appears to be too difficult to analyse directly. We will instead define two other ways of direct summing a promise equality problem which sandwich it. In all three cases, we give Alice and Bob vertices $(x_1,\dots,x_n)$ and $(y_1,\dots,y_n)$ of G, respectively. We consider the following three problems:

(1) Alice and Bob are promised that x_i and y_i are either equal or adjacent for every i. Bob must determine whether there exists some i with $x_i \neq y_i$.

(2) Alice and Bob are promised that x_i and y_i are either equal or adjacent for every i. For every i, Bob must output $b_i = I[x_i = y_i]$. This is the original problem.

(3) Bob is given a $i \in [n]$ with the promise that x_i and y_i are either equal or adjacent and must output $b = I[x_i = y_i]$.

Problem (2) is clearly no easier than problem (1), as a protocol for problem (2) trivially gives a protocol for problem (1). Problem (3) is also no easier than problem (2) because given a protocol for (3) and given the promise of problem (2), Bob can collect the value of b_i for every i, as his measurements do not affect the state. Note that problem (3) (and, by extension, problem (1) and (2)) can be solved by running the best protocol for EQ_G n times, from which Bob can recover the value of b. As we will see, this isn't always optimal. We now relate problems (1) and (3) to graph products.

Given two graphs G and H, the **strong product** $G \boxtimes H$ is the graph with vertex set $V(G) \times V(H)$ such that $(x_1,x_2) \sim (y_1,y_2)$ if $(x_1,x_2) \neq (y_1,y_2)$ and if, for every i, either $x_i = y_i$ or $x_i \sim y_i$. Applying the mapping described earlier in this section, it can be seen that problem (1) corresponds to the promise equality problem with the n-th power of G under the strong product, denoted $G^{\boxtimes,n}$, as the underlying graph.

The **disjunctive product** $G \star H$ is the graph with vertex set $V(G) \times V(H)$ such that $(x_1,x_2) \sim (y_1,y_2)$ if, for some i, $x_i \sim y_i$. Likewise, it may be seen that problem (3) corresponds to the promise equality problem with $G^{\star,n}$ as the underlying graph.

In terms of lower bounds, Knuth showed:

Proposition 4.2.44 ([**20**]). *For any two graphs G,H,*

$$\overline{\vartheta}(G \boxtimes H) = \overline{\vartheta}(G)\overline{\vartheta}(H)$$

This means that the exact, one-way communication complexity of problem (1) (and, a fortiori, problems (2) and (3)) is at least $(\log_2 \overline{\vartheta}(G))n$ bits/qubits in all models. In particular, in the Cleve-Buhrman model, with base graph G_{21}, it follows from proposition 4.2.32 that the complexities of all three problems are precisely $2n$ bits.

We now turn to upper bounds. Given a graph G, let $R \subseteq \mathbb{Q}$ be the set of all positive rationals r such that, for some m,n with $r = \frac{m}{n}$, it is possible to assign subsets of $[m]$ of cardinality n $\{S_x\}_{x \in V(G)}$ to the vertices of G in such a way that for all $(x,y) \in E(G)$, $S_x \cap S_y = \emptyset$. Such an assignment is referred to as a fractional colouring of G. We then define the **fractional chromatic number** $\chi_f(G)$ by $\chi_f(G) = \inf R$. It is easy to see that the $n = 1$ case corresponds to vanilla colourings of G, so that $\chi_f(G) \leq \chi(G)$. It turns out that $\chi_f(G)$ can be recast as a linear program of size exponential in $|V(G)|$ ([**27**]), from which it can be seen that $\chi_f(G)$ is a computable quantity and is always rational.

We now show the rather important following property of the fractional chromatic number:

Proposition 4.2.45. *We have*

$$\chi_f(G) = \lim_{n\to\infty} \chi(G^{\star,n})^{1/n}$$

PROOF. Lovász ([**4**]) showed that, for any graph H,

$$\chi(H) \le \chi_f(H)(1+\log\alpha(H))$$

Feige ([**27**]) showed that $\chi_f(G^{\star,n}) = \chi_f(G)^n$ for all n. For all n, we therefore have

$$\chi_f(G) \le \chi(G^{\star,n})^{1/n} \le \chi_f(G)(1+n\log|V(G)|)^{1/n}$$

Elementary calculus gives the desired result. □

It may be established using a computer that $\chi_f(G_{21}) = 4$. This means that $2n+O(\log n)$ bits of communication are sufficient to solve problem (3) (and, in particular, problems (2) and (1)) for G_{21} in the classical setting. In particular, none of the three direct sums that we are studying can yield a very impressive gap between the exact complexities in the Yao and Cleve-Buhrman models with base graph G_{21}.

The **projective rank** ([**78**]) of G, denoted by $\xi_f(G)$, is a quantum generalisation of sorts of the fractional chromatic number. In this case, we let $R \subseteq \mathbb{Q}$ be the set of all positive rationals r such that, for some m,n with $r = \frac{m}{n}$, it is possible to assign n-dimensional subspaces of $\mathbb{C}^m$ $\{V_x\}_{x\in V(G)}$ to the vertices of G so that, for every $(x,y) \in E(G)$, $V_x \perp V_y$. Such an assignment is referred to as a projective representation of G. Again, we define $\xi_f(G)$ to be the infimum of R. The $n=1$ case corresponds to orthogonal representations of G, so that $\xi_f(G) \le \xi(G)$, and it may be seen that $\xi_f(G) \le \chi_f(G)$, as fractional colourings of G map to projective representations in a natural way by viewing colours as standard basis vectors. Note that theorem 4.2.28 together with the well-known fact that $\chi_f(G) \ge \frac{|V(G)|}{\alpha(G)}$ imply that the Hadamard graphs exhibit an exponential separation between the fractional chromatic number and the projective rank. Unlike the fractional chromatic number, it is unknown at present whether the projective rank is a computable quantity or whether it is always rational.

We show:

Proposition 4.2.46.

$$\xi_f(G) \le \kappa_{qf}(G)$$

PROOF. Take a perfect finite-dimensional protocol for G with communication $k = \kappa_{qf}(G)$, which we assume to be in standard form. Letting n be the rank of the $\rho_{x,c}$, for every x, we take $V_x \in \mathbb{C}^{nk}$ to be the support of $\rho_{x,1}$. Clearly, these subspaces satisfy $V_x \perp V_y$ for all $(x,y) \in E(G)$. This shows that $\xi_f(G) \le k$. □

We also have the following result:

Proposition 4.2.47 ([**73**]). *For all n,*

$$\xi_f(G^{\star,n}) = \xi_f(G)^n$$

By means of the somewhat unscientific procedure of replacing the χ's with the ξ's in proposition 4.2.45, one arrives at the following conjecture:

Conjecture 4.2.48. *We have*

$$\xi_f(G) = \lim_{n\to\infty} \xi(G^{\star,n})^{1/n}$$

Combined with proposition 4.2.47, the truth of the above conjecture would imply that direct sums can never be used to amplify the gap between the exact, one-way comunication complexities of the Yao and Cleve-Buhrman models in the style of problem (3) when the entanglement is restricted to be finite-dimensional:

Corollary 4.2.49. *We have*

$$\lim_{n\to\infty} \xi(G^{\star,n})^{1/n} \leq \lim_{n\to\infty} \kappa_{qf}(G^{\star,n})^{1/n}$$

Note however that even if the above conjecture were to be true, problems (1) and (2) might still turn out to be suitable for gap amplification.

4.3. A conjectured general embedding of nonlocal games inside communication complexity using interactive hashing

In the previous two sections, we showed how to turn certain restricted families of nonlocal games (namely, XOR games and colouring games) into communication complexity problems. In both cases, this was accomplished by twisting the players' arms into playing the game. In this section, we propose a more general way of doing this for more general nonlocal games which is loosely inspired by the interactive hashing protocol of Naor, Ostrovsky, Ventakesan and Yung ([**31**]). Although we could not prove that our reduction is correct [13], it seems to promise embeddings of nonlocal games into communication complexity that are much stronger than what the (provably correct) embeddings that were given in the previous two sections allow for, as the resulting separations and hardness results all vanish when some breathing room is permitted in terms of success probability. While what follows is stated in terms of an arbitrary nonlocal game,

[13] Doing so strikes us as extremely hard, even when the players are disallowed to share prior entanglement. The trouble is that while proofs of security of interactive hashing protocols like that of [**31**] can assume that one player is behaving honestly, this is not possible here since Alice and Bob are cooperating.

we specifically have the games coming from MIP*=RE in mind.

We begin by fixing notation. Given a finite set S and a parameter $N \in \mathbb{N}$, we define a (S,N)-*hash tree* T as being a binary tree of depth N where each vertex v has a function $h_v : S \to \{0,1\}$ associated to it and where each edge is labelled with either 0 or 1 in such a way that the two outgoing edges going out of any non-leaf vertex are labelled differently.

We now describe the proposed reduction. Take a nonlocal game G, with input sets X,Y, output sets A,B, input distribution $p_{x,y}$ and predicate $V : X \times Y \times A \times B \to \{0,1\}$. We make the hypothesis that for all x,y, there exist a,b such that $V(x,y,a,b) = 1$. Pick $N \approx K \log_2 \max(|A|,|B|)$ to be an even number, for some large enough constant K, say $K = 10$. In the corresponding communication problem, we give Alice and Bob inputs x,y sampled according to the input distribution of the nonlocal game: in addition, we give Alice a (A,N)-hash tree T_A, and we give Bob a (B,N)-hash tree T_B, both taken uniformly at random. Finally, we give Alice a function $f : \{0,1\}^{2N} \to \{0,1\}$ taken uniformly at random among the functions satisfying the promise to be described below.

We begin by describing what we expect Alice and Bob to do. First, they play the game with the best strategy available to them with the resources they are allowed to use, each obtaining outputs $a \in A, b \in B$. Next, they run their outputs through the hash trees in the following way. Let v_0 be the root of t_A, and let w_0 be the root of T_B. Alice begins by computing $b_1 = h_{v_0}(a)$ and sends b_1 to Bob. Let w_1 be the vertex such that the edge going from w_0 to w_1 is labelled by b_1: Bob then computes $b_2 = h_{w_1}(b)$ and sends it to Alice. Letting v_1 be the vertex such that the edge going from v_0 to v_1 is labelled by b_2, Alice computes $b_3 = h_{v_1}(a)$ and sends it to Bob, who performs the transition from w_1 corresponding to b_3, computes $b_4 = h_{w_1}(b)$, sends it to Alice and so forth. They keep on doing this until the bits $b_1,\dots,b_{2N}$ have been obtained, at which point Alice outputs $f(b_1 b_2 \dots b_{2N})$. We give Alice and Bob the promise that the map f is balanced and that there exists some bit d such that for all choices of outputs a and b such that $V(x,y,a,b) = 1$, if the above protocol is run from these choices of outputs, the value of $f(b_1 b_2 \dots b_{2N})$ will always be d, and d is the bit the players are supposed to compute. The choice of N above was made so that, with good probability, the string $b_1 b_2 \dots b_{2N}$ identifies uniquely the outputs the hashing protocol was run on: taking N to be too small, with good probability, the string $b_1 b_2 \dots b_{2N}$ Alice and Bob would obtain would correspond to a possible transcript coming from winning outputs, even if the outputs they ran the protocol on are not actually winning outputs. With our choice of N, if they ran the protocol from losing outputs, with high probability, the string they will obtain will result in Alice outputting a uniformly random bit.

We will be making the following hypothesis:

Conjecture 4.3.1. *Any communication protocol for the above problem which succeeds with high enough probability must essentially involve Alice and Bob figuring out outputs a,b and running them through the above interactive hashing scheme.*

In particular, if the communication is restricted to be at most $2N$ bits, in any given model, if p is the value of the nonlocal game G given the resources that we allow them to use, the best Alice and Bob seem to be able to do is play the game with a strategy achieving a success probability $\approx p$ and run the above interactive scheme, achieving a success probability of essentially $p + \frac{1-p}{2}$: with probability p, they obtain winning outputs and they win with certainty, and with probability $1 - p$, they obtain losing outputs and the bit they get is uniformly random. If the communication is not restricted to be at most $2N$ bits, we see that a classical protocol which solves the problem exactly consists in Alice sending Bob her input, from which he can determine winning outputs (a,b) and send a to Alice; they then proceed to run the hashing protocol, for a total communication cost of $2N + \log_2 |X| + \log_2 |A|$. We do not expect this to be optimal in general, but we do expect that if the value of G in a given model is very low, significantly more communication than $2N$ bits will be required in order to solve the problem with reasonable probability.

Of course, these considerations depend critically on the truth of 4.3.1. While we think that this is extremely likely to be true in the classical setting (which isn't of very great interest to us anyway), we admit that a certain leap of faith is required when the players are allowed to share entanglement. We would not be completely shocked if some Grover-style protocol of the style that was given for the disjointness function could be used for running the interactive scheme more efficiently, or even for breaking our scheme in an even more fundamental way.

4.4. The possible implications for a quantum Newman's theorem

We now discuss the possible implications of the previous constructions for the possibility of a version of theorem 3.1.6 for entanglement. By this we mean some hypothetical result which would imply that the amount of entanglement used by a protocol for a given function can always be assumed to be reasonably small at the expense of slightly worsening the success probability, as theorem 3.1.6 does for public coins. An analogue of 3.1.6 for entanglement-assisted communication complexity might look as follows:

Conjecture 4.4.1 (Hypothetical quantum Newman's theorem)**.** *There exists some sensible (at least computable) function $\mathscr{E}(n,\delta) : \mathbb{N} \times (0,\frac{1}{2}) \to \mathbb{R}$ (say, $\mathscr{E}(n,\delta) = \log_2 n + 2\log_2 \frac{1}{\delta} + O(1)$) such that, given a boolean function $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ and given an amount of communication C, there exists a protocol with communication cost C, entanglement cost $\mathscr{E}(n,\delta)$ and success probability at least $\omega^*(f,C) - \delta$.*

Just like Newman's theorem limits the power of sharing a public coin, such a result would limit the power of sharing prior entanglement, and, one might hope, help bridge the gap between the Yao and Cleve-Buhrman models. It should be mentioned that Jain, Radhakrishnan and Sen ([**62**]) prove a restriction on any such conjecture (in their result 8) by showing that a reasonable-sounding analogue for prior entanglement of the reduction used in the proof of Newman's original theorem (3.1.6) does not work. However, this is strictly a restriction on a method of proof and has no bearing on the statement of the conjecture itself.

The results in the previous two sections already restrict the realm of possibilities for the function N somewhat. For example, if the shared prior entanglement was restricted to be used as a public coin, one would have that $\mathscr{E}(n,\delta) = 2n$ would suffice, by proposition 3.1.8, while we have seen that this fails both for one-time-pad problems and for promise equality problems and in multiple ways. In general, we must have that, for all but a finite number of values of n,

$$\lim_{\delta \to 0^+} \mathscr{E}(n,\delta) = \infty$$

Also, the results of the previous section almost imply that conjecture 4.4.1 is false as stated, although not in a very significant way. As we discussed, there exists a reduction from general nonlocal games G to communication problems f such that, if G has a perfect quantum strategy, then f has a perfect qf-protocol using a trit of communication, while if $\omega_{qf}(G) < 1$, then any qf-protocol using a trit of communication must have error at least $\varepsilon > 0$, for some $\varepsilon > 0$. If, given a non-trivial upper bound on $\omega_{qf}(G) < 1$, say, and given a description of G, we could efficiently compute a lower bound for ε, the N from conjecture 4.4.1 would allow us to design an algorithm to determine which is the case between G having a perfect finite-dimensional strategy and $\omega_{qf}(G) < \frac{1}{2}$ under the promise that one of the two is the case, which is forbidden by the MIP*=RE result. The trouble is that the reduction from games to communication problems chains Harris' reduction and ours, and while our reduction is effective, Harris' isn't, so that no means is known of computing a lower bound for ε, although there surely is a way to make Harris' reduction effective. In any case, this would only imply that one would either have to slightly increase the communication cost of the protocol in the statement of the above conjecture or restrict to a fixed value of δ. Neither restriction is particularily important from the point of view of communication complexity, but this warns us away from trying to prove the above naive equivalent of Newman's theorem.

However, the conjectured reduction of the previous section presents a much more serious challenge to the possibility of showing something of the form 4.4.1. Indeed, MIP*=RE implies that there are nonlocal games which have a perfect strategy and which are such that the amount of prior entanglement required by any qf-strategy which achieves success probability at least ε, for some fixed choice of $0 < \varepsilon < 1$, is uncomputable in the size of the description of the game. In particular, conditionally on the truth of hypothesis 4.3.1, running the reduction described in the previous section on such games would yield communication problems which may be solved exactly with prior entanglement using some amount of communication C and any protocol with communication C and achieving a reasonable success probability must use massive amounts of prior entanglement, certainly much larger than the amount of communication, even if shared randomness is considered free, so that unless our construction can be broken, no such computable function $\mathscr{E}$ can exist, even for a fixed value of δ: and further, it even seems possible that any low-entanglement protocol which achieves a high winning probability for such problems must use at least αC bits of communication, for some universal $\alpha > 1$.

Chapter 5

Conclusion and open problems

In this thesis, we studied the power of entanglement in communication complexity. The most important results that we could obtain are the following:

(1) We gave an entanglement-assisted protocol for the equality function which matches the success probability of the standard one and uses half as much prior entanglement.

(2) We looked at the results of Hadiashar and Nayak ([**91**]) on state compression, and formulated a problem, called the density matrix in subspace problem, which has the potential for yielding a large separation between the Yao and Cleve-Buhrman models.

(3) We looked at one-time-pad-problems, made their relationship with XOR games explicit, and showed that in some cases, if the communication is restricted to be one bit from Alice to Bob, the entanglement cost of a near-optimal Cleve-Buhrman protocol is exponential in the sizes of the inputs. We could also show that Yao protocols can be turned into Cleve-Buhrman protocols which use a constant amount of prior entanglement, so that the Cleve-Buhrman model does slightly better than the Yao model when the communication is restricted to be one bit/qubit, even in the presence of shared randomness.

(4) We also looked at the communication complexity equivalent of colouring games, namely, promise equality problems. There, we could show that a number of separations and hardness results coming from the theory of nonlocal games also hold for communication complexity: namely, we could show that there are communication problems such that, for a fixed amount of communication C, there may be one-way protocols with communication C which achieve an arbitrarily small error probability, and yet no exact one-way protocol with communication C exists; we could also separate the tensor-product and commuting operators models in communication complexity, showing that there are some problems which admit an exact protocol making use of infinite-dimensional entangled state while any protocol using the same amount of communication and a finite-dimensional entangled state has success probability at most $1-\varepsilon$, for some $\varepsilon > 0$; finally, we showed that determining whether a communication problem has an exact protocol using a certain amount

of communication is an undecidable problem, implying that in the exact case, even if the amount of communication is fixed, the number of shared EPR pairs that might be needed by an exact communication protocol is unbounded. We also gave examples of problems with exact communication complexity smaller in the Cleve-Buhrman model than in the Yao model, introducing the notion of a vector clump in the process.

(5) Inspired by cryptography, we gave a proposed reduction from general nonlocal games to communication problems which, if correct, would yield problems for which the amount of prior entanglement required by a near-optimal bounded-error protocol cannot be upper bounded by a computable function in the description of the problem.

In our humble opinion, our work represents a serious improvement of our understanding of the role of shared entanglement in communication complexity. Our work leaves a number of research avenues open, most notably:

(1) It would be great to show that the distance between subspaces problem which we gave in subsection 3.3.5 is indeed hard in the Yao model, thereby separating the Cleve-Buhrman and Yao models in the one-way setting, or to show that the reduction we proposed in section 4.3 is correct. We believe that the former, while hard, could conceivably be pulled off, while the latter represents a monumental challenge.

(2) We think that the group games which we introduced in subsection 4.1.5 are worth investigating in more detail.

(3) The communication problems we gave are all instances of promise problems and might therefore be deemed artificial. It would be interesting to see if some of our results could be shown for total problems as well.

(4) While we could show separations between the various entanglement models in communication complexity in section 4.2, the separations we could obtain are all rather minute. For example, if we now know that there is a problem with an exact protocol in the commuting operators model with a trit of communication and that the success probability of any protocol in the tensor-product setting with a trit of communication is upper bounded by some $p < 1$, we do not know an explicit p with this property, and it seems likely that any such p would have to be comically close to one. It would be interesting to look for another provably correct embedding of the MIP*=RE result in communication complexity that would yield a more substantial separation. As far as we can see, nothing speaks against the possibility of an exponential separation between $C^{qc}_{1/3}$ and $C^{*}_{1/3}$. While exhibiting such a separation appears to be completely out of reach at the present time, we feel that this would be a far more dramatic separation between the commuting operators and tensor-product models than that which is given by the $MIP^* = RE$ result.

(5) While we know that the separation between the Cleve-Buhrman and classical public-coin models cannot be larger than exponential in the bounded-error setting, no such limitation is

known in the exact setting. In particular, there could be a sequence of graphs $\{G_i\}$ all with the same value of $\kappa_{qf}(G_i)$ but with unbounded $\chi(G_i)$, or equivalently, with unbounded $\xi(G_i)$. We think that it would be really interesting to try to prove that this can be the case, as it would separate the Cleve-Buhrman and Yao models in the one-way, exact settings. Similarly, such a sequence could conceivably exist with constant $\kappa_{qc}(G_i)$ and unbounded $\kappa_{qf}(G_i)$.

(6) We also think that our notion of a vector clump, which we used to build a small graph G_{21} with $\chi_{qf}(G_{21}) < \chi(G_{21})$, could well be of more general interest to quantum graph theorists, as it might enable constructing small graphs that have a richer structure from the point of view of the quantum chromatic number than the ones that can be found in the recent literature.

References

[1] W. Gottschalk (1951): *Choice functions and Tychonoff's theorem*. Proceedings of the American Mathematical Society, 2 (1)

[2] J. Clauser, M. Horne, A. Shimony and R. Holt (1969): *Proposed experiment to test local hidden-variable theories*. Physical Review Letters 23, no. 15, 880–884

[3] A. Holevo (1973): *Bounds for the quantity of information transmitted by a quantum communication channel*. Problems of Information Transmission. 9: 177–183.

[4] L. Lovász (1975): *On the ratio of optimal integral and fractional covers*. Discrete Math. 13:4, p. 960-981.

[5] A. Hedayat and W. Wallis (1978): *Hadamard Matrices and Their Applications*. Ann. Statist. 6(6), p. 1184-1238.

[6] H. Abelson (1978): *Lower bounds on information transfer in distributed computations*. Proceedings of the 19th Annual IEEE Symposium on Foundations of Computer Science, p. 151–158.

[7] L. Lovász (1979): *On the Shannon capacity of a graph*. IEEE Transactions on Information Theory, IT-25 (1): 1–7

[8] A. Yao (1979): *Some complexity questions related to distributive computing*. Proceedings of the 11th ACM Symposium on the Theory of Computing, pp. 209-213

[9] M. Grötschel, L. Lovász and A. Schrijver (1981): *The ellipsoid method and its consequences in combinatorial optimization*. Combinatorica 1 (1981), 169–197

[10] B. Tsirelson (1985): *Quantum analogues of the Bell inequalities. The case of two spatially separated domains.* Zapiski Nauchnykh Seminarov Leningradskogo Otdeleniya Matematicheskogo Instituta im. V. A. Steklova AN SSSR, Vol. 142, p. 174-179.

[11] P. Frankl and V. Rödl (1987): *Forbidden intersections*. Transactions of the American Mathematical Society, Vol. 300, No. 1 (Mar., 1987), pp. 259-286

[12] B. Kalyanasundaram and G. Schnitger (1987): *The probabilistic communication complexity of set intersection*. Second *Structure in Complexity Theory Conference*, p. 41–49.

[13] I. Newman (1991): *Private vs common random bits in communication complexity*. Information Processing Letters, 39(2):67–71, 1991.

[14] C. Bennett and S. Wiesner (1992): *Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states*. Physical Review Letters 69 (20): 2881–2884.

[15] S. Arora and S. Safra (1992): *Probabilistic Checking of Proofs: A New Characterization of NP*. Proceedings of the 33rd Annual Symposium on Foundations of Computer Science

[16] D. Deutsch and R. Jozsa (1992): *Rapid solutions of problems by quantum computation*. Proceedings of the Royal Society of London A. 439, p. 553-558.

[17] C. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres and W. Wootters (1993): *Teleporting an Unknown Quantum State via Dual Classical and Einstein-Podolsky-Rosen Channels*. Physical Review Letters. 70 (13): 1895–1899

[18] Y. Nesterov and A. Nemirovskii (1993): *Interior-Point Polynomial Algorithms in Convex Programming*. SIAM

[19] A. Yao (1993): *Quantum circuit complexity*. Proceedings of the 34th annual IEEE symposium on foundations of computer science, p. 352-361

[20] D. Knuth (1993): *The Sandwich Theorem*. arXiv:math/9312214v1

[21] I. Kremer (1995): *Quantum Communication*. Master's thesis, Hebrew University.

[22] I. Kremer, N. Nisan and D. Ron (1995): *On randomized one-round communication complexity*. Proceedings of the twenty-seventh annual ACM symposium on Theory of computing, pages 596–605

[23] D. Rohrlich and S. Popescu (1995): *Nonlocality as an axiom for quantum theory.* arXiv:quant-ph/9508009

[24] E. Kushilevitz and N. Nisan (1996): *Communication complexity*. Cambridge University Press

[25] L. Grover (1996): *A fast quantum mechanical algorithm for database search*. Procedures of the 28th STOC, p. 212–219.

[26] R. Cleve and H. Buhrman (1997): *Substituting quantum entanglement for communication*. arXiv:quant-ph/9704026

[27] U. Feige (1997): *Randomized graph products, chromatic numbers, and the Lovász ϑ-function*. Combinatorica, 17, p. 79-70.

[28] B. McKay (1998): *Isomorph-free exhaustive generation*. Journal of Algorithms 26(2), 306–324

[29] H. Buhrman, R. Cleve and A. Wigderson (1998): *Quantum vs. Classical Communication and Computation* arXiv:quant-ph/9802040

[30] R. Raz (1998): *A Parallel Repetition Theorem*. SIAM Journal on Computing, Vol. 27, Iss. 3

[31] M. Naor, R. Ostrovsky, R. Ventakesan and M. Yung (1998): *Perfect Zero-Knowledge Arguments for NP Using Any One-Way Permutation*. Journal of Cryptology, vol. 11, p. 87–108.

[32] G. Brassard, R. Cleve and A. Tapp (1999): *The cost of exactly simulating quantum entanglement with classical communication* arXiv:quant-ph/9901035

[33] A. Nayak (1999): *Optimal lower bounds for quantum automata and random access codes*. arXiv:quant-ph/9904093

[34] Pati (1999): *Minimum cbits for remote preperation and measurement of a qubit*. arXiv:quant-ph/9907022

[35] H. Buhrman and R. de Wolf (2000): *Communication Complexity Lower Bounds by Polynomials*. arXiv:cs/9910010

[36] H. Klauck (2000): *On quantum and probabilistic communication: Las Vegas and one-way protocols*. Proceedings of the 32nd Annual ACM Symposium on Theory of Computing (STOC), p. 644–651.

[37] R. de Wolf (2001): *Quantum computing and communication complexity*. PhD thesis, University of Amsterdam.

[38] H. Buhrman, R. Cleve and W. van Dam (2001): *Quantum entanglement and communication complexity*. arXiv:quant-ph/9705033

[39] C. Bennett, D. DiVincenzo, P. Shor, J. Smolin, B. Terhal, W. Wootters (2001): *Remote state preparation*. arXiv:quant-ph/0006044

[40] A. Nayak and J. Salzman (2002): *On communication over an entanglement-assisted quantum channel*. arXiv:quant-ph/0206122

[41] A. Razborov (2002): *Quantum communication complexity of symmetric predicates*. arXiv:quant-ph/0204025

[42] G. Brassard, A. Broadbent and A. Tapp (2003): *Multi-Party Pseudo-Telepathy*. arXiv:quant-ph/0306042

[43] R. Jain, J. Radhakrishnan and P. Sen (2003): *A direct sum theorem in communication complexity via message compression*. arXiv:cs/0304020

[44] R. Cleve, P. Høyer, B. Toner and J. Watrous (2004): *Consequences and limits of nolocal strategies*. arXiv:quant-ph/0404076

[45] Z. Bar-Yossef, T. Jayram and I. Kerenidis (2004): *Exponential separation of quantum and classical one-way communication complexity*. Proceedings of 36th ACM STOC, p. 128-137

[46] C. Godsil and M. Newman (2005): *Colouring an Orthogonality Graph*. arXiv:math/0509151

[47] D. Avis, J. Hasegawa, Y. Kikuchi and Y. Sasaki (2005): *A quantum protocol to win the graph colouring game on all Hadamard graphs* arXiv:quant-ph/0509047

[48] G. Brassard, H. Buhrman, N. Linden, A. Méthot, A. Tapp and F. Unger (2005): *A limit on nonlocality in any world in which communication complexity is not trivial*. arXiv:quant-ph/0508042

[49] S. Aaronson and A. Ambainis (2005): *Quantum search of spatial regions*. arXiv:quant-ph/0303041

[50] D. Gavinsky, J. Kempe, O. Regev and R. de Wolf (2005): *Bounded-error quantum state identification and exponential separations in communication complexity.* arXiv:quant-ph/0511013

[51] D. Gavinsky, J. Kempe and R. de Wolf (2006): *Exponential Separation of Quantum and Classical One-Way Communication Complexity for a Boolean Function*. arXiv:quant-ph/0607174

[52] D. Gavinsky (2006): *On the Role of Shared Entanglement*. arXiv:quant-ph/0604052

[53] P. Cameron, A. Montanaro, M. Newman, S. Severini and A. Winter (2007): *On the quantum chromatic number of a graph*. arXiv:quant-ph/0608016

[54] N. Gisin, A. Méthot and V. Scarani (2007): *Pseudo-telepathy: input cardinality and Bell-type inequalities*. arXiv:quant-ph/0610175

[55] I. Kerenidis and R. Raz (2007): *The one-way communication complexity of the Boolean Hidden Matching Problem*. arXiv:quant-ph/0607173

[56] F. Dupuis, N. Gisin, A. Hasidim, A. Méthot and H. Pilpel (2007): *No nonlocal box is universal*. arXiv:quant-ph/0701142

[57] M. Navascués, S. Pironio and A. Acín (2008): *A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations*. arXiv:0803.4290

[58] A. Doherty, Y. Liang, B. Toner and S. Wehner (2008): *The quantum moment problem and bounds on entangled multi-prover games*. arXiv:0803.4373

[59] Y. Shi and Y. Zhu (2008): *Tensor Norms and the Classical Communication Complexity of Nonlocal Quantum Measurement*. arXiv:quant-ph/0511071

[60] N. Linial and A. Shraibman (2008): *Lower bounds in communication complexity based on factorization norms*. Random Struct. Alg., 34: 368-394

[61] V. Scholz and R. Werner (2008): *Tsirelson's problem*. arXiv:0812.4305

[62] R. Jain, J. Radhakrishnan and P. Sen (2008): *Optimal Direct Sum and Privacy Trade-off Results for Quantum and Classical Communication Complexity*. arXiv:0807.1267

[63] S. Arora and B. Barak (2009): *Computational Complexity: A Modern Approach*. Cambridge University Press

[64] H. Buhrman, R. Cleve, S. Massar and R. de Wolf (2009): *Non-locality and communication complexity*. arXiv:0907.3584

[65] B. Klartag and O. Regev (2010): *Quantum One-Way Communication is Exponentially Stronger Than Classical Communication*. arXiv:1009.3640

[66] W. Slofstra (2011): *Lower bounds on the entanglement needed to play XOR non-local games*. arXiv:1007.2248

[67] A. Acín, S. Massar and S. Pironio (2012): *Randomness versus nonlocality and entanglement*. arXiv:1107.2754

[68] R. Cleve, W. van Dam, M. Nielsen and A. Tapp (2013): *Quantum entanglement and the communication complexity of the inner product function*. arXiv:quant-ph/9708019

[69] Z. Ji (2013): *Binary Constraint System Games and Locally Commutative Reductions*. arXiv:1310.3794

[70] V. Paulsen, S. Severini, D. Stahlke, I. Todorov and A. Winter (2014): *Estimating quantum chromatic numbers*. arXiv:1407.6918

[71] E. Meckes (2014): *Concentration of measure and the compact classical groups*. Lecture notes, available at `https://case.edu/artsci/math/esmeckes/Haar_notes.pdf`

[72] A. Ambainis (2014): *Superlinear Advantage for Exact Quantum Algorithms*. arXiv:1211.0721

[73] T. Cubitt, L. Mančinska, D. Roberson, S. Severini, D. Stahlke and A. Winter (2014): *Bounds on Entanglement Assisted Source-channel Coding via the Lovász ϑ Number and its Variants*. arXiv:1310.7120

[74] J. Briët, H. Buhrman, D. Leung, T. Piovesian and F. Speelman (2015): *Round elimination in exact communication complexity*. arXiv:1812.09290

[75] M. Bavarian, T. Vidick and H. Yuen (2015): *Anchored parallel repetition for nonlocal games*. arXiv:1509.07466

[76] X. Wu and H. Yuen (2015): *On the limits of communication with non-local resources*. `https://www.cs.umd.edu/~xwu/papers/ns_capacity_note.pdf`

[77] L. Mančinska and D. Roberson (2016): *Oddities of quantum colorings*. arXiv:1801.03542

[78] L. Mančinska and D. Roberson (2016): *Quantum Homomorphisms*. arXiv:1212.1724

[79] L. Mančinska, D. Roberson and A. Varvitsiotis (2016): *Deciding the existence of perfect entangled strategies for nonlocal games*. arXiv:1506.07429

[80] A. Anshu, A. Belovs, S. Ben-David, M. Göös, R. Jain, R. Kothari, T. Lee and M. Santha (2016): *Separations in Communication Complexity Using Cheat Sheets and Information Complexity*. arXiv:1605.01142

[81] H. Buhrman, L. Czekaj, A. Grudka, M. Horodecki, P. Horodecki, M. Markiewicz, F. Speelman and S. Strelchuk (2016): *Quantum communication complexity advantage implies violation of a Bell inequality*. arXiv:1502.01058

[82] W. Slofstra (2017): *The set of quantum correlations is not closed*. arXiv:1703.08618

[83] N. Bao, S. Carroll and A. Singh (2017): *The Hilbert Space of Quantum Gravity Is Locally Finite-Dimensional*. arXiv:1704.00066

[84] P. Raymond-Robichaud (2017): *L'équivalence entre le local-réalisme et le principe de non-signalement*. PhD thesis, Université de Montréal.

[85] A. Coladangelo and J. Stark (2018): *Unconditional separation of finite and infinite-dimensional quantum correlations*. arXiv:1804.05116

[86] J. Watrous (2018): *The Theory of Quantum Information*. Cambridge University Press.

[87] K. Dykema, V. Paulsen and J. Prakash (2018): *Non-closure of the set of quantum correlations via graphs*. arXiv:1709.05032

[88] G. Vos (2019): *Quantum correlation matrices and Tsirelson's problem: Previous work and three-player considerations*. Master's thesis. `http://resolver.tudelft.nl/uuid:fcdd6b06-ecb5-4d4f-9795-f4130c3ab7f6`

[89] R. Cleve (2019): *QIC 890: Entanglement and Nonlocal Effects: Lecture Notes*. `https://cleve.iqc.uwaterloo.ca/resources/Qic890LectureNotes2019Apr22(V22).pdf`

[90] M. Coudron and A. Harrow (2019): *Universality of EPR pairs in entanglement-assisted communication complexity, and the communication cost of state conversion*. arXiv:1902.07699

[91] S. Hadiashar and A. Nayak (2019): *On the Entanglement Cost of One-Shot Compression*. arXiv:1905.02110

[92] W. Slofstra (2020): *Tsirelson's problem and an embedding theorem for groups arising from non-local games*. arXiv:1606.03140

[93] Z. Ji, A. Natarajan, T. Vidick, J. Wright and H. Yuen (2020): MIP*=RE. arXiv:2001.04383

[94] S. Khatri and M. Wilde (2020): *Principles of Quantum Communication Theory: A Modern Approach*. arXiv:2011.04672

[95] J. Watrous (2020): *Advanced Topics in Quantum Information Theory: Lecture 8*. `https://cs.uwaterloo.ca/~watrous/QIT-notes/QIT-notes.08.pdf`

[96] T. Russell (2021): *A synchronous NPA hierarchy with applications*. arXiv:2105.01555

[97] H. Mousavi, S. Nezhadi and H. Yuen (2021): *Nonlocal Games, Compression Theorems, and the Arithmetical Hierarchy*. arXiv:2110.04651

[98] M. Gartska, M. Cannon and P. Goulart (2021): *COSMO: A Conic Operator Splitting Method for Convex Conic Problems*. Journal of Optimization Theory and Applications, vol. 190, p. 779-810

[99] D. Cui, A. Mehta, H. Mousavi and S. Nezhadi (2021): *A generalization of CHSH and the algebraic structure of optimal strategies*. arXiv:1911.01593

[100] J. Kamminga (2021): *Strengths and limitations of quantum entanglement*. Master's thesis, Radboud University. Private communication from Ronald de Wolf.

[101] S. Harris (2023): *Universality of graph homorphism games and the quantum colouring problem*. arXiv:2305.18116

[102] O. Lalonde, N. Mande and R. de Wolf (2023): *Tight Bounds for the Randomized and Quantum Communication Complexities of Equality with Small Error*. arXiv:2107.11806

[103] O. Lalonde (2023): *On the Quantum Chromatic Numbers of Small Graphs*. arXiv:2311.08194

[104] P. Botteron, A. Broadbent and M. Proulx (2024): *Extending the Known Region of Nonlocal Boxes that Collapse Communication Complexity*. arXiv:2302.00488

[105] B. McKay: *graph formats*. `http://users.cecs.anu.edu.au/~bdm/data/formats.html`

Appendix A

The clumps corresponding to the graph G_{21}

In this appendix, we give the (2,2)-clumps corresponding to the graph G_{21} given in subsection 4.2.7.

As in matrix notation, when specifying a clump $\{|\psi_{j,k}\rangle\}$, j runs from top to bottom and k runs from left to right.

(1)

$$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & -1 & 0 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & 1 & 0 \end{bmatrix}$$
$$\frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 1 & 0 & 1 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 1 & 0 & -1 \end{bmatrix}$$

(2)

$$\frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 0 & 1 & -1 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 & 0 & 0 \end{bmatrix}$$
$$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 & 0 & 0 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 0 & -1 & -1 \end{bmatrix}$$

(3)

$$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & 1 & 0 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} -1 & 0 & 1 & 0 \end{bmatrix}$$
$$\frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 1 & 0 & 1 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & -1 & 0 & 1 \end{bmatrix}$$

(4)

$$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & -1 & 0 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} -1 & 0 & -1 & 0 \end{bmatrix}$$
$$\frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 1 & 0 & 1 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 1 & 0 & -1 \end{bmatrix}$$

(5)

$$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 & 0 & 0 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 0 & -1 & -1 \end{bmatrix}$$
$$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 & 0 & 0 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 0 & -1 & 1 \end{bmatrix}$$

(6)

$$\frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 1 & 1 & 0 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 1 & -1 & 0 \end{bmatrix}$$
$$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & 0 & 1 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & 0 & -1 \end{bmatrix}$$

(7)

$$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & -1 & 0 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & 1 & 0 \end{bmatrix}$$
$$\frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 1 & 0 & 1 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & -1 & 0 & 1 \end{bmatrix}$$

(8)

$$\frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & -1 \end{bmatrix}, \frac{1}{2}\begin{bmatrix} -1 & -1 & 1 & -1 \end{bmatrix}$$
$$\frac{1}{2}\begin{bmatrix} 1 & -1 & 1 & 1 \end{bmatrix}, \frac{1}{2}\begin{bmatrix} -1 & 1 & 1 & 1 \end{bmatrix}$$

(9)

$$\frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & -1 \end{bmatrix}, \frac{1}{2}\begin{bmatrix} -1 & 1 & 1 & 1 \end{bmatrix}$$
$$\frac{1}{2}\begin{bmatrix} 1 & -1 & 1 & 1 \end{bmatrix}, \frac{1}{2}\begin{bmatrix} 1 & 1 & -1 & 1 \end{bmatrix}$$

(10)

$$\frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 1 & 0 & -1 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & -1 & 0 & -1 \end{bmatrix}$$
$$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & -1 & 0 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} -1 & 0 & -1 & 0 \end{bmatrix}$$

(11)

$$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & -1 & 0 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & -1 & 0 & 1 \end{bmatrix}$$
$$\frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 1 & 0 & 1 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & 1 & 0 \end{bmatrix}$$

(12)

$$\frac{1}{2}\begin{bmatrix} 1 & -1 & -1 & 1 \end{bmatrix}, \frac{1}{2}\begin{bmatrix} -1 & 1 & -1 & 1 \end{bmatrix}$$
$$\frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}, \frac{1}{2}\begin{bmatrix} -1 & -1 & 1 & 1 \end{bmatrix}$$

(13)

$$\frac{1}{2}\begin{bmatrix} 1 & 1 & -1 & -1 \end{bmatrix}, \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}$$
$$\frac{1}{2}\begin{bmatrix} 1 & -1 & 1 & -1 \end{bmatrix}, \frac{1}{2}\begin{bmatrix} -1 & 1 & 1 & -1 \end{bmatrix}$$

(14)

$$\frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 0 & 1 & -1 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 & 0 & 0 \end{bmatrix}$$
$$\frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 0 & 1 & 1 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 & 0 & 0 \end{bmatrix}$$

(15)

$$\begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -1 & 0 & 0 \end{bmatrix}$$
$$\begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 0 & 0 & 0 \end{bmatrix}$$

(16)

$$\frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 1 & 0 & -1 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 1 & 0 & 1 \end{bmatrix}$$
$$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & 1 & 0 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & -1 & 0 \end{bmatrix}$$

(17)

$$\frac{1}{2}\begin{bmatrix} 1 & -1 & -1 & -1 \end{bmatrix}, \frac{1}{2}\begin{bmatrix} -1 & -1 & -1 & 1 \end{bmatrix}$$
$$\frac{1}{2}\begin{bmatrix} 1 & -1 & 1 & 1 \end{bmatrix}, \frac{1}{2}\begin{bmatrix} 1 & 1 & -1 & 1 \end{bmatrix}$$

(18)

$$\frac{1}{2}\begin{bmatrix} 1 & 1 & -1 & -1 \end{bmatrix}, \frac{1}{2}\begin{bmatrix} 1 & -1 & -1 & 1 \end{bmatrix}$$
$$\frac{1}{2}\begin{bmatrix} 1 & -1 & 1 & -1 \end{bmatrix}, \frac{1}{2}\begin{bmatrix} -1 & -1 & -1 & -1 \end{bmatrix}$$

(19)

$$\frac{1}{2}\begin{bmatrix} 1 & 1 & -1 & -1 \end{bmatrix}, \frac{1}{2}\begin{bmatrix} -1 & 1 & 1 & -1 \end{bmatrix}$$
$$\frac{1}{2}\begin{bmatrix} 1 & -1 & 1 & -1 \end{bmatrix}, \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}$$

(20)

$$\frac{1}{\sqrt{2}}\left[\begin{array}{cccc} 0 & 1 & 1 & 0 \end{array}\right], \frac{1}{\sqrt{2}}\left[\begin{array}{cccc} -1 & 0 & 0 & 1 \end{array}\right]$$
$$\frac{1}{\sqrt{2}}\left[\begin{array}{cccc} 1 & 0 & 0 & 1 \end{array}\right], \frac{1}{\sqrt{2}}\left[\begin{array}{cccc} 0 & -1 & 1 & 0 \end{array}\right]$$

(21)

$$\left[\begin{array}{cccc} 1 & 0 & 0 & 0 \end{array}\right], \left[\begin{array}{cccc} 0 & 0 & 0 & -1 \end{array}\right]$$
$$\left[\begin{array}{cccc} 0 & 1 & 0 & 0 \end{array}\right], \left[\begin{array}{cccc} 0 & 0 & -1 & 0 \end{array}\right]$$