Quantum Entanglement in Non-local Games, Graph Parameters and Zero-error Information Theory

Giannicola Scarpa

Quantum Entanglement in Non-local Games, Graph Parameters and Zero-error Information Theory

ILLC Dissertation Series DS-2013-03

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

For further information about ILLC-publications, please contact

Institute for Logic, Language and Computation
Universiteit van Amsterdam
Science Park 107
1098 XG Amsterdam
phone: +31-20-525 6051
e-mail: illc@uva.nl
homepage: http://www.illc.uva.nl/

Quantum Entanglement in Non-local Games, Graph Parameters and Zero-error Information Theory

Academisch Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof.dr. D.C. van den Boom
ten overstaan van een door het college voor
promoties ingestelde commissie, in het openbaar
te verdedigen in de Agnietenkapel
op woensdag 27 november 2013, te 10.00 uur

door

Giannicola Scarpa

geboren te Salerno, Italië.

Promotor: Prof. Dr. R.M. de Wolf

Overige leden: Prof. Dr. H. Buhrman
Prof. Dr. J.S. Caux
Prof. Dr. M. Laurent
Dr. S. Severini
Dr. L. Torenvliet

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

The investigations were performed at the Centrum Wiskunde & Informatica (CWI) and were supported by Vidi grant 639.072.803 from the the Netherlands Organization for Scientific Research (NWO).

Printed and bound by Ipskamp Drukkers

ISBN: 978-90-6196-566-4

to Carminella,

my epic grandmother

Contents

Acknowledgments

You are now able to read this dissertation thanks to the careful guidance of Ronald de Wolf. In these four years in Amsterdam, I have seen him as a fatherly figure and an excellent teacher of both science and life. He has been especially good in dropping his usual rigorous mathematical approach, and trusting me at the very right moment.

Simone Severini is another person to praise in case you like this thesis. Thanks to him, I discovered my passion for graph theory and for working with a smile. I do not know how many hours I spent with him talking about chromatic numbers and noisy channels (but I can give an upper bound of 35064).

And how can I forget *the quantum killer*, Harry Buhrman? When casually passing by my office, he usually figures out a one-liner that destroys the conjecture written on my board. He is the man who distributes happiness and wisdom at CWI, and one of the best actors I ever directed.

I am grateful to the other members of my PhD committee, Jean-Sébastien Caux, Monique Laurent and Leen Torenvliet, for their enthusiasm and their kind and helpful comments that improved my work. Talking about improvements: a big thanks goes to Emma Cook, Ignacio Cascudo, Christian Schaffner and Teresa Piovesan, who patiently proofread my drafts, and to Berta Cillero and Nadia Sabbetta, who accepted to be at my side during my defense (they know karate). Tom Sterkenburg is the one who wrote my beautiful *samenvatting*, which I then translated to English (or the other way around).

I am immensely indebted to my coauthors and the colleagues who inspired me with beautiful scientific discussions (there are detailed acknowledgments at the beginning and at the end of each chapter). I would also like to mention some other people that are responsible for my love for science, in order of appearance: Walter Cosenza, Carlo Blundo, Roger Penrose, Giuseppe Persiano, Vincenzo Auletta, Fabrizio Illuminati and Andreas Winter.

Finally, I enjoyed working in Amsterdam thanks to the support from my loved one, my family and my friends. Serenity is the best way to be productive!

Giannicola Scarpa
October 2013, Amsterdam

Chapter 1

Introduction

1.1 Is the world quantum?

Quantum mechanics is so counter-intuitive that it struggled to be accepted by its own creators. Niels Bohr once said [Hei71, page 206]:

> *"Those who are not shocked when they first come across quantum theory cannot possibly have understood it".*

Albert Einstein was convinced that quantum mechanics is wrong (or at least incomplete). He once asked to Abraham Pais [Pai79, page 907]:

> *"Do you really believe the moon exists only when you look at it?".*

Why was Bohr shocked and why was Einstein worried about the moon? The concern was about issues with *locality* and *realism.* These are two properties of a physical theory that any of us would find obvious and essential. Locality means that an action in a certain place cannot influence the state of a distant object instantaneously, *i.e.*, the effects of an action travel at a finite speed. Realism means that all properties of a physical system have a specific, definite value, even before they are observed (measured). Such things seem so clearly true when looking at "real life". One cannot call a waiter in a restaurant and immediately get his attention: the sound has to get there and it travels at a finite speed. Also, I am overweight even when I am not standing on the scales!

Einstein, Podolsky and Rosen, in their famous 1935 article [EPR35], observed that quantum physics, in addition to being a theory that explained several classically unexplained phenomena, also predicts the existence of objects with a weird behavior. They exhibited an example of such objects, a bipartite state now known as *the EPR pair*. This was an example of a large class of systems that are now called *entangled*. They thought that their example could be used to prove quantum physics wrong or incomplete. Instead, it created a new branch of research.

Let us give an intuitive description of their argument. Since quantum mechanics seems to be consistent with experimental results, it is often assumed to be a complete theory. A physical theory is *complete* when it perfectly predicts the outcome of a measurement of each definite property of a system. They argue, with the following thought experiment, that such an assumption leads to a contradiction. First, they consider Heisenberg uncertainty principle of quantum mechanics: the values of some pairs of properties, such as the position and the momentum of a particle cannot *both* be predicted perfectly.[1] Second, they introduce the above-mentioned EPR pair: a pair of particles that are allowed to interact up to a certain moment and then separated and not allowed to interact anymore. This bipartite quantum system is designed such that if one measures the position of the first particle, then one knows perfectly the position of the second particle without observing it. Similarly, measuring the momentum of the first particle removes any uncertainty on the momentum of the second particle. Therefore, it seems that for the second particle both the position and the momentum have a definite value, even prior to observation. Here comes a "paradox": if one assumes that quantum theory is complete, then the position and the momentum cannot both have definite values in real life (otherwise there would be no uncertainty principle). On the other hand, quantum mechanics allows to define an EPR pair, which shows that both these quantities have a definite value for the second particle prior to its observation.

There are two explanations of this apparent contradiction. One is that quantum mechanics is wrong, or incomplete. The other is that, somehow, in an EPR pair observing a property of the first particle also determines the value of the property of the second particle, instantaneously. Einstein disliked the second explanation and later called the phenomenon "spooky actions at a distance" [EBB71, page 158]. He believed there is a theory that explains nature while respecting both locality and realism.

The discussion remained open for many years, until Bell in 1964 [Bel64] proposed an experiment that made it possible to test if nature respects locality and realism. He designed a set of measurements of the properties of a bipartite physical system, such that if locality and realism are respected, then the measurements outcomes could not exceed a certain value. This is known as a *Bell inequality*. The non-locality of the EPR pair in quantum physics allowed to *violate a Bell inequality* and obtain measurement outcomes that are larger than predicted by classical physics or any other local realistic theory. Experiments of this kind have been done many times, starting from Aspect *et al.* in 1981 [AGR81]. The results are surprising: nature violates a Bell inequality! The tests behaved according to the prediction of quantum physics. But hold your enthusiasm: this does not prove that quantum physics is the correct theory, it just *dis*proves that nature

[1] For example, if the position of a particle is certain, then quantum theory can only describe the result of a measurement of its momentum probabilistically.

behaves according to classical physics or any local realistic theory.

Moreover, there are some *loopholes*. The problem with these Bell inequality violations is that the amount of violation is very small and this fact allows for (very contrived) classical explanations of the experimental results. These explanations often rely on unlikely combinations of things like noise, malfunction in measurement devices and incorrect post-selection of the results. Researchers have tried to get rid of loopholes in various ways, one of which is to design *stronger* Bell inequalities. The ideal situation is to get a huge, thus easily noticeable, difference between classical and quantum behavior using very small physical systems in the experiments. Unfortunately, Junge *et al.* [JPP$^+$10] recently proved that on a bipartite system of local dimension n, the maximum Bell inequality violation is of order[2] n. Therefore, to obtain large Bell inequality violations one must work on large systems.

Our 2 cents Our contribution to the study of non-locality is 3-fold. In Chapter 2 of this thesis, we achieve the goal of finding very large Bell inequality violations. We design two experiments in the form of non-local games, and we prove that players relying on classical physics are outperformed by quantum players by a factor that is very close to the optimal violation proved by Junge and coauthors. In Chapter 3 and 4 we make use of non-locality and entanglement of physical systems in different settings. We use non-local games to approach mathematical problems on graphs, and we use entanglement to assist and improve communication in some problems of information theory. We proceed now to introduce each chapter individually.

1.2 Non-local games

The setting of "games" has been widely used in computer science. Sometimes we face complicated-looking problems, and we seek for a more intuitive equivalent way to describe and work with them. *Non-local games* are a good example of this: they make the concept of non-locality and Bell inequalities more accessible.

The setting we consider in this thesis is the following, illustrated in Figure 1.1. This is a game between two players, Alice and Bob, and a referee. A referee asks Alice a question x and asks Bob a question y. He chooses the pair of questions x, y according to a probability distribution that is known to Alice and Bob. He receives back an answer a from Alice and an answer b from Bob. After that, he decides according to the rules if the players win or lose.[3] During the game Alice and Bob are space-like separated: they are put so far away that information, which

[2] The local dimension is the dimension of one part of the bipartite system. The term "order" should be interpreted intuitively as "up to a constant factor".

[3] Notice that this setting is different from the usual setting used in game theory. Alice and Bob are not competing, but are collaborating to both win the game.

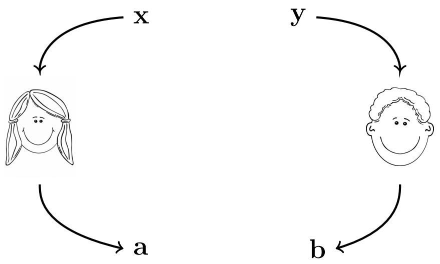


Figure 1.1: Setup of a non-local game.

travels at finite speed, cannot be exchanged between them until they produce the answers.

The *value* of the game is the maximum winning probability of the players. A Bell inequality as described above, in this setting is an upper bound on the classical value of the game. This means that whatever strategy classical players arrange before the game, they cannot exceed that probability of winning. Sometimes, however, there are quantum strategies that make smart use of entanglement and exceed the upper bound. Therefore, each game has a classical and quantum value, and for some games the quantum value is strictly larger than the classical. In this case we have a Bell inequality violation, and we quantify it by the ratio

$$\frac{\text{quantum value}}{\text{classical value}}.$$

Here is an example, due to Peres [Per91]. It is not the simplest and most common example, but it is very intuitive.[4] Consider the *magic square* from Figure 1.2. It is a 3×3 grid that has to be filled in with natural numbers, such that the sum of each row is an even number, and the sum of each column is a odd number. A little thought shows that it is impossible to completely fill in the square.

Now let us base a non-local game on it. The referee lets Alice and Bob arrange a strategy, then separates them. He chooses randomly two numbers x, y from 1 to 3. Then he asks Alice to give him the numbers a_1, a_2, a_3 of the x-th row, and Bob to give him the numbers b_1, b_2, b_3 of the y-th column. The players lose if any of the following happens:

- The intersection of the column and the row is different, *i.e.*, if $a_y \neq b_x$,
- The sum of Alice's numbers is not even,
- The sum of Bob's numbers is not odd.

There are 9 possible question pairs x, y to this game. The classical players can easily win in 8 out of 9 cases by preparing a square like in Figure 1.2 and

[4] The classic example is the CHSH game, explained in Section 2.1.

1	2	1
2	2	2
2	1	?

Figure 1.2: A 3×3 magic square from [Per91], with an incomplete filling.

by completing the missing entry in two different ways: Alice with a 1 and Bob with a 2. They will win all the question pairs except the questions $x = 3, y = 3$. Therefore, since the referee selects the questions at random, the classical value of the game is at least 8/9. It is actually equal to 8/9, because a little thought shows that from a classical strategy that wins in all cases one can derive a completion of the whole square, which is impossible.

Surprisingly, there is a way to win in all 9 cases using entanglement. The players can share a carefully designed entangled state and perform appropriate operations on it to use the non-classical correlations to their advantage. They would always answer correctly even while the referee is sure that they are not communicating (because of their huge distance in space). Therefore in the quantum case the value of this game is 1.

Here is a non-technical intuition on what happens. Alice and Bob share an entangled quantum system before the game starts. During the game, Alice measures her part of the state in a way that depends on her input x, obtaining 3 numbers that sum to an even number. Alice's actions instantly modify the quantum state on Bob's side at a distance (spooky!). Bob's measurement depends on y and is designed such that when he reads out the numbers, the intersection with Alice's row is always the correct number, and together with the other two (random) numbers sums to an odd number.

Basically, non-local correlations are used to "generate on-the-fly" a valid answer to any question pair, without any information being transmitted between Alice and Bob. Such a strategy allows to win the game but not to fill-in a whole magic square, which would be impossible. It is also important to notice that no information about Alice's input is transferred to Bob during the game, and vice versa. The only thing that Bob knows by measuring his part of the state is that one of his three random numbers is equal to one of Alice's random numbers, and it is the one in the right position. He knows that it is a wise choice to output such numbers.

Our contribution As argued above, the problem with Bell inequality violations, and with the known non-local games, is that the difference between the quantum and classical values is quite small. On the other hand, Junge *et al.* proved that the violation cannot be larger than the local dimension of the shared quantum state.

Our contribution is to exhibit two non-local games that exhibit a violation very close to optimal. They are formally discussed in Chapter 2. We informally introduce them here.

The first one is called *Hidden Matching*. The game is inspired by a communication problem in [BYJK08]. It goes as follows. Alice gets a random string of n bits (with n power of 2), and Bob gets a random matching on the n indices.[5] Alice outputs a bit string, Bob outputs a pair of indices in his matching and a bit. The winning condition is a bit technical, but a rough intuition is that an analysis on Alice and Bob's outputs should suggest that Bob knew the sum of Alice's bits in the positions he declared.

The game is designed to be very difficult to classical players (they can win with probability roughly $\frac{1}{2} + \frac{1}{\sqrt{n}}$) but perfectly playable by Alice and Bob that share a quantum state of local dimension n (they can win with probability 1). We obtain a very large separation by considering the deviation from $\frac{1}{2}$. When quantum players share an entangled state of local dimension n, the deviation from $\frac{1}{2}$ is roughly $\sqrt{n}$ times larger than in the classical case.

The second non-local game exhibits a larger separation than the other, but is not perfectly won by quantum players. It is called the *Khot-Vishnoi* game, and it is inspired by an example used by Khot and Vishnoi in [KV05].

Alice and Bob each are given as inputs a set of bit-strings. The two sets are not totally random: Alice's receives a set of n strings respecting some properties; Bob receives n strings that are obtained from Alice's by adding a random noise string z, which only the referee knows. Alice outputs a string a from her set and Bob outputs a string b from his set. They win the game if Bob's string is the noisy version of Alice's string, *i.e.*, $b = a + z$.

This turns out to be an extremely hard problem for classical players. Since for any output of Alice a there is only one good answer for Bob, by playing a random answer their winning probability is $1/n$. Basically, their best strategy is very close in winning probability to giving a random answer, thus the classical value is roughly $1/n$. It can be shown that quantum players, instead, by using n EPR pairs, have a winning probability close to *constant* (but far from perfect). Therefore, for n large enough, the quantum players sharing n EPR pairs outperform the classical ones by a factor of order roughly n.

Thus, this game almost matches the upper bound given by Junge and coauthors. To date, it is the non-local game that shows the largest quantum advantage.

1.3 Graph parameters

Our first application of non-locality to another area is the study of quantum graph parameters in Chapter 3. In order to explain our contribution, it is essential to define graphs are some important graph parameters.

[5] A matching on n indices is a set of $n/2$ distinct pairs (i, j). For example, for the numbers between 1 and 8 a possible matching is $\{(1, 3), (2, 4), (5, 8), (6, 7)\}$.

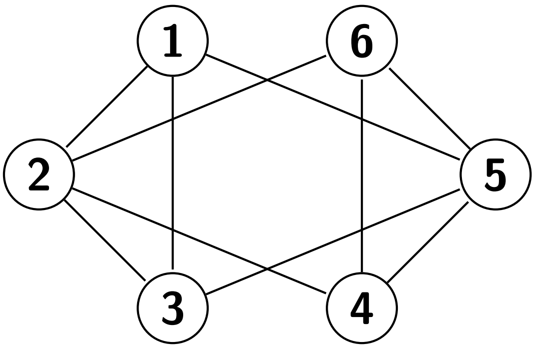


Figure 1.3: A graph on 6 nodes.

A *graph* is a simple yet important mathematical structure. It consists of a set of *nodes* (also called vertices), and a set of pairs of nodes called *edges.* Two nodes that form an edge are *adjacent* or *neighbors.* In the typical depiction of a graph the nodes are drawn as circles and the edges are lines connecting pairs of them. We can see an example in Figure 1.3. A *graph parameter* is a quantity that is associated to the graph and usually describes one of its properties. For example, basic graph parameters are its number of nodes and its number of edges. We talk about more interesting ones below.

Why are graphs so important? Basically, because *many* things can be represented with graphs. For example, we could represent a map, with the nodes being cities and the edges being roads connecting them. But also our family tree, a computer network, the structure of a molecule, the dependencies of software packages, and so on. It is surprising how many problems can be solved by constructing a graph and calculating one of its parameters.

Let us informally explain two of the most studied graph parameters (the formal definitions are in Section 3.2.1).

Chromatic number Suppose we are given a big piece of paper with the drawing of a graph. Then we are asked to color each circle using pencils, but *never* to fill in two adjacent nodes with the same color. Of course, one easy solution is to use one color per node, with a very artistic result. But now suppose each color pencil is very expensive, so we must use as few colors as possible. What is the minimum number of colors that we need to use? This is a graph parameter called the *chromatic number.* For example, the graph of Figure 1.3 has chromatic number 3, and Figure 1.4 shows an optimal coloring.

It turns out that computing the chromatic number of a graph is hard (it is one of the so-called NP-Hard problems. All known computer programs that solve this problem in general take a huge amount of time in proportion to the graph size. If the graph has n nodes, the best known program will have to perform roughly $2^n n$ elementary operations. This number can be enormous. Even if our computer performs millions of operations per second, this will not prevent us to

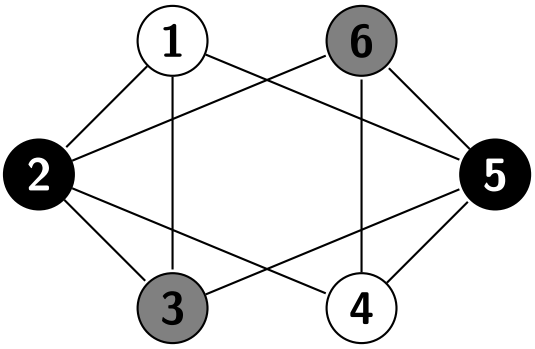


Figure 1.4: A coloring for the graph of Figure 1.3.

have to wait for ages as soon as we want to color 1000 nodes. One of the most important open problems in mathematics, the P vs. NP problem, could be solved if we find an efficient way to color a graph, or prove that there is none. It is hard even just to decide in general if a given graph has chromatic number equal to 3.

Independence number Suppose we have a wolf, a goat and a cabbage. The goat wants to eat the cabbage, and the wolf wants to eat the goat. How many of these can we keep at home? Well, bringing the wolf and the cabbage is optimal in this case. That was an easy instance of the problem. But now suppose we have a big aquarium that we want to populate, and we go into a shop to buy fish. The shop has hundreds of species, but not all species of fish can live with each other. The shopkeeper gives us a graph where each node is a species of fish, and two nodes form an edge if the two species are incompatible. How many different species of fish can we put safely in our aquarium?

This problem is equivalent to finding a subset of the nodes such that no two nodes are adjacent. Such a subset is called an *independent set*, and the size of the largest independent set is a graph parameter called the *independence number*. For example, the largest independent set in the graph of Figure 1.4 has 2 nodes.[6] Computing the independence number of a graph is also a hard problem, so the same discussion as for the chromatic number applies.

Lovász ϑ number Many famous mathematicians worked on the chromatic number or the independence number problems (for example Erdös, Lovász, Knuth and Schrijver).

A new twist on the problem came when the ϑ number of a graph was defined in [Lov79]. This is a graph parameter efficiently computable with a semidefinite program,[7] and it was introduced for solving a open problem of Shannon in infor-

[6] Notice that each pair of nodes with the same color form an independent set: this is not a coincidence, a coloring is a partitioning into independent sets!

[7] In fact, it started the whole field of semidefinite programming, a subfield of optimization.

mation theory: the zero-error capacity of the 5-cycle (we will discuss this kind of problems in Section 1.4). The Lovász ϑ number lies "sandwiched" between the independence number and the chromatic number of the complement graph,[8] *i.e.*, for all graphs G and their complement $\bar{G}$ it is true that

$$\text{independence number of } G \leq \vartheta \text{ number of } G \leq \text{chromatic number of } \bar{G}.$$

The bound works well for some classes of graphs (*e.g.*, the so-called perfect graphs, where the inequalities above are equalities), while the gap is large for other graphs. The quantity is very interesting for computer science, and Knuth dedicated an excellent survey to it [KD93].

Non-local games based on graph parameters Researchers have tried to find better bounds for the chromatic and independence numbers. One line of research, formally started by [CMN$^+$07] and [CLMW10] but implicitly present in previous works, uses quantum mechanics to define *quantum graph parameters.* The main tool we use to define quantum graph parameters is to consider non-local games based on the graph coloring and independent set problems.

We start with the *coloring game* on a graph, illustrated in Figure 1.5a. Suppose Alice and Bob claim that they have a coloring of a graph G that uses c colors, and a referee wants to test this claim. The referee can play the coloring game on G with c colors with Alice and Bob. He lets the players arrange a strategy and then he separates them, such that they cannot communicate with each other anymore. He gives Alice as input a node x of the graph chosen randomly and asks her to color it with one of the c possible colors. Alice gives him the color a as output. The referee gives Bob as input a node y of the graph chosen randomly, and asks him to color with one of the c colors. Bob gives him the color b as output. The referee then checks for the consistency of their outputs. If the players had the same node as input ($x = y$), then they must have given the same color as output ($a = b$). If players had adjacent nodes as input (x, y form an edge), then they must have given different colors as output ($a \neq b$). If the referee finds an inconsistency, the players lose the game.

Can the classical players always win the coloring game on G with c colors? It depends on c. If c is equal to or larger than the chromatic number of G, then the players can arrange a winning strategy that consists of a correct coloring. On the other hand, if c is strictly smaller than the chromatic number, then the players must lose for at least one input pair. Here is a proof, slightly technical, of the claim. Suppose, towards a contradiction, that players can win the coloring game on G with strictly fewer colors than the chromatic number. Then, we can construct a coloring for G as follows. Note that since for all $x = y$ we have $a = b$, Alice and Bob must follow an equivalent strategy. This implies that Alice must

[8] The complement of a graph is a new graph with the same set of nodes, but with an edge between two nodes if and only if there is no edge in the original graph.

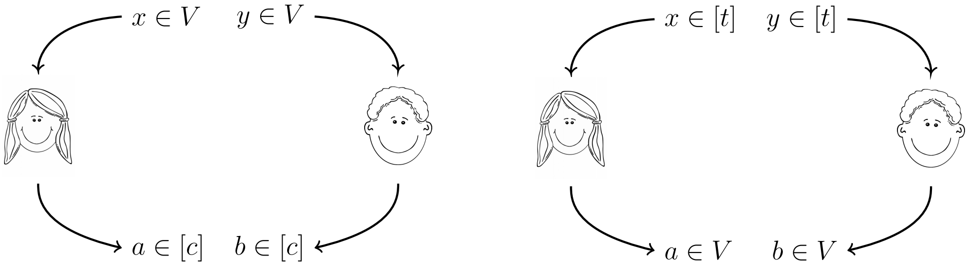
output different colors when asked adjacent nodes. Therefore, a proper coloring can be constructed by observing what color Alice outputs for each input.[9] The existence of such a coloring with c colors contradicts the assumption that the chromatic number is strictly larger than c. Therefore, the chromatic number of the graph G is the smallest number c such that Alice and Bob can always win the coloring game on G with c colors.

Cameron *et al.* and Avis *et al.* [CMN$^+$07, AHKS06] have found that if the players are allowed to share quantum entanglement, then for some graphs they can win the coloring game using fewer colors than the chromatic number! They defined the *quantum chromatic number* of a graph as the minimum number of colors such that players that share entanglement can always win the coloring game on G.

Winning the game by using entanglement is not equivalent to produce a valid coloring! Indeed, player produce their answers at random following a correlated probability distribution. How does this work? Here is an intuition (the technical definition is in Section 3.3). The players share an entangled state that, as explained above, exhibits non-classical correlations. Alice performs a measurement on her part of the state that depends on her input x and outputs the outcome. This measurement is designed to produce a random color from the set of allowed colors as outcome. Alice's actions influence instantaneously the state on Bob's side (again, spooky!). Bob performs a measurement on his part of the state that depends on his input y and outputs the measurement outcome. The measurements are carefully made to give the same (random) outcome for Alice and Bob if $x = y$ and to give different (random) outcomes if x, y form an edge. Therefore, Alice and Bob will always win the game, and wisely designed measurements that make use the non-classical correlation allow, for some graphs. to win with a number of colors strictly less than the chromatic number.

A similar discussion can be done for the *independence number* (see Figure 1.5b). Consider the same non-local setting, where Alice and Bob now are trying to prove that the graph G contains an independent set of size t. The referee lets them arrange a strategy, then separates them. He gives to Alice the input x a random number between 1 and t. Alice gives as output a node a, which is meant to correspond to the x-th element of the independent set. Similarly, the referee gives to Bob as input number y between 1 and t, and receives a node b as output. He then checks for consistency: if $x = y$, then he must have received outputs $a = b$, and if $x \neq y$, then he must have received outputs from an independent set, therefore a, b should not form an edge. A similar analysis as with the coloring game shows that the independence number of a graph is the maximum t such that the classical players can always win the independent set game on the graph with

[9] A technical note: since the players must always win, we can rule out any use of randomness. Any randomness used to make choices in the strategy must result in winning outputs, therefore the players can just fix the random outcomes and play an equivalent deterministic strategy.



(a) Coloring game: Alice and Bob each receive an input from the vertex set V of the graph, and produce one out of c colors as output.

(b) Independent set game: Alice and Bob each receive as input a number from 1 to t, and produce an output from the vertex set V.

Figure 1.5: Setup of non-local games based on graph parameters.

size-parameter t. Cubitt *et al.* [CLMW10] initiated the study of the quantum version of this game in the context of zero-error information theory (which is the topic of the next section). A more graph-parametric version of the problem has been given in [RM12]. Similarly to the coloring game, also for the independent set game, players that share entanglement have an advantage. For some graphs, they can always win the independent set game with a size-parameter t strictly larger than the independence number. The *quantum independence number* of a graph is the maximum t such that the quantum players can always win the independent set game on the graph with size-parameter t.

These quantum parameters are natural bounds on the classical ones, because the presence of entanglement cannot be a disadvantage for the players. It turns out that the efficiently-computable ϑ number is "sandwiched" also in between these quantum parameters. Therefore, the quantum parameters are potentially tighter bounds to the classical counterparts than the ϑ number. However, it is an open problem to determine their computational complexity.

Our contribution Chapter 3 contains our work on quantum graph parameters. We prove properties of the quantum chromatic number, especially in comparison with other graph quantities. We then find a surprising relationship between graphs exhibiting a separation between quantum and classical chromatic number and objects from the foundation of quantum theory known as *Kochen-Specker sets.* These are sets of vectors with some specific properties that were used to prove the Kochen-Specker theorem, a no-go result that excludes some underlying explanations for quantum mechanical behavior. Kochen and Specker exhibited a finite set of measurements (that are made of the vectors mentioned above) such that it is impossible to choose a pre-determined outcome to each measurement. In short, this construction proves that if there is a classical model based on variables describing the state of quantum system, then those variables cannot be *non-contextual*, *i.e.*, independent of the particular measurement. This result is similar to Bell inequalities but without a bipartite setting.

We defined a generalization of these important objects, called *projective Kochen-Specker sets.* We find that to win the coloring game with fewer colors than the chromatic number, Alice and Bob's strategies must consist of projective Kochen-Specker sets.

Our other contribution is about the independence number. We find several classes of graphs for which there is a separation between quantum and classical independence numbers. One class is based on our projective Kochen-Specker sets, one is based on the quantum chromatic number and the last one is based on *graph states* (well-known quantum systems). Moreover, we find that starting from any non-local game, we can construct a graph based on the game description and calculate a bound on the quantum value of the game from the quantum independence number of the graph. Computing this graph parameter appears to be a useful tool for studying non-locality in general.

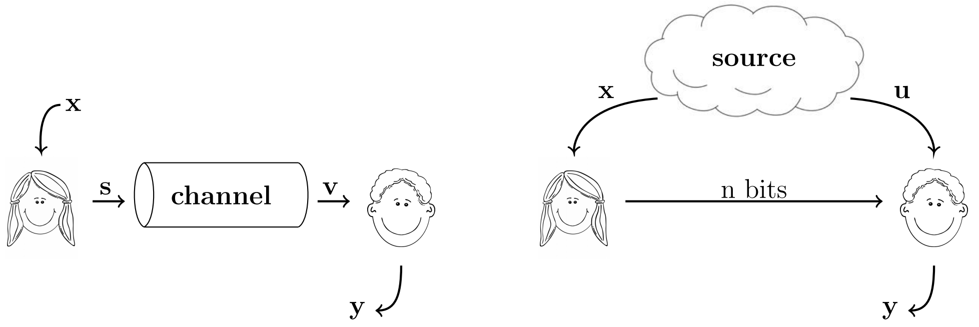
1.4 Zero-error information theory

Our second application of non-locality is to consider the advantages of entanglement in some information theory problems. The problems are about sending a message from Alice to Bob efficiently in presence of various kinds of noise. Namely, we study the *channel* problem, the *(dual) source* problem and their generalization, the *source-channel* problem. We focus on the zero-error setting, namely the setting where all the information must be transmitted without any possibility of errors. The main reasons are two: first, it has been shown that entanglement is not beneficial in the asymptotic *bounded-error* setting[10] [BSST02], and second, the techniques for the zero-error framework come from graph theory.

Channel problem The seminal paper by Shannon [Sha56] introduced the concept of zero-error capacity of a noisy channel. This setting is represented in Figure 1.6a.

Imagine that Alice and Bob communicate through a noisy channel that has an input set S and an output set V. Each element $s \in \mathsf{S}$ is associated to $\mathsf{V}(s)$, a subset of the output set. Such subsets for distinct $s, t \in \mathsf{S}$ may overlap. The channel is noisy in the sense that each element $s \in \mathsf{S}$, when sent into the channel, can produce as output any element of $\mathsf{V}(s)$. Therefore, there are some outputs of the channel on Bob's side that do not allow Bob to unequivocally distinguish between two inputs on Alice's side. Two inputs $s, t \in \mathsf{S}$ are *confusable* if there exists an output v such that both of them can produce v as output, *i.e.*, their output sets intersect.

[10] Where we allow the protocols to make a small probability of mistake while decoding the message.



(a) Channel problem **(b)** Source problem

Figure 1.6: Setup of zero-error information theory problems.

Let us provide a simple example of such a channel. Imagine that the input set consists of the numbers from 1 to 6, the output set is {odd, even, > 3, ≤ 3} and the obvious rule determines the output subsets. For example, we have that:

- If Alice sends "2" , Bob receives either "even" or "≤ 3",
- If Alice sends "3", Bob receives either "odd" or "≤ 3",
- If Alice sends "5", Bob receives either "odd" or "> 3".

Therefore, 2 and 3 are confusable, 3 and 5 are confusable but 2 and 5 are not.

The zero-error channel problem asks the following question:

> Alice communicates to Bob through a noisy channel with input set S, output set V and output subsets $\{\mathsf{V}(s)\}_{s\in\mathsf{S}}$. What is the maximum number of bits of information that Alice can send to Bob on average per use of the channel, without any chance of a mistake?

In this introduction, for simplicity, we illustrate the simpler problem of a single use of the channel. To solve this problem, it is useful to introduce the *confusability graph* of the channel. This is a graph that has node set S and where $s, t \in \mathsf{S}$ are adjacent if they are confusable. The confusability graph of our simple example is the graph in Figure 1.3. To maximize the number of bits sent with a single use of the channel, Alice and Bob may select an independent set $\mathsf{I} \subseteq \mathsf{S}$ of the confusability graph and restrict the inputs to that set. This way, the output subsets of the elements of I are disjoint, and Bob, upon receiving an output from the channel, can identify without error which input has been sent by Alice. How many bits have been sent? If I has size t, then sending one out of t possible messages transfers $\log_2(t)$ bits of information.[11] Therefore, the *channel capacity*

[11] When measuring the amount of information, it is possible to have a non-integer number of bits.

with a single use of the channel is given by the logarithm of the independence number of the confusability graph.

The solution for the original question, the average number of bits sent with many uses of the channel, is known as zero-error Shannon capacity. An intuition for how to solve this problem is the following. Multiple uses of the channel correspond to a larger channel whose confusability graph is given by a *graph power* (*i.e.*, a bigger graph resulting from a sequence of graph multiplications). For n that goes from one to infinity, let c_n be the maximum number of bits sent per each use of the channel, after n uses. The maximum average capacity of the channel is the maximum[12] of the set $\{c_1, c_2, \dots\}$. It is not known if the Shannon capacity is computable at all, although the already mentioned ϑ number is a useful upper bound.

Sometimes, if Alice and Bob have the additional resource of quantum entanglement, they can exceed the Shannon capacity for a channel. The maximum average number of bits sent in this case is called the *entanglement-assisted* zero-error Shannon capacity. This quantity was first studied by Cubitt *et al.* in [CLMW10].

Source problem We now consider a different scenario, studied by Witsenhausen [Wit76] in the zero-error setting. This setting is represented in Figure 1.6b.

Alice and Bob are connected to a *dual source*, that sends an input to Alice from a set X and some side information to Bob from a set U. Every $x \in \mathsf{X}$ has associated a set $\mathsf{U}(x) \subseteq \mathsf{U}$. These are all the elements of U that could be sent to Bob when x is sent to Alice. Of course, such sets may intersect. Two inputs x, y in X are *not uniquely identifiable* if their side information subsets intersect.

Consider the simple example we gave above for the channel, in a dual source form. The input set for Alice consists of the numbers from 1 to 6, the side information set for Bob is $\{\text{odd, even, } > 3, \leq 3\}$ and the obvious rule is used to determine the side information subsets. We have, for example, that:

- If Alice gets “2” , Bob gets either “even” or “≤ 3”,
- If Alice gets “3”, Bob gets either “odd” or “≤ 3”,
- If Alice gets “5”, Bob gets either “odd” or “> 3”.

Therefore, on Bob's side 2 and 3 are not uniquely identifiable, 3 and 5 are not uniquely identifiable but 2 and 5 are uniquely identifiable.

But the player's goal is for Alice to send her input to Bob, taking advantage of the side information. So this time the question is:

> Alice and Bob have a dual source with Alice's input set X, Bob's side information set U and side information subsets $\{\mathsf{U}(x)\}_{x\in\mathsf{X}}$. They have access to a perfect one-way channel from Alice to Bob. What is

[12] More precisely, the supremum.

> the minimum number of bits of information that Alice can send on average to Bob per source input, in order for him to recover Alice's inputs without chance of a mistake?

As with the channel, here we will restrict our discussion to the single source-input and we will use a graph parameter to answer the question.

Define the *characteristic graph* of the source as the graph with node set X and edges between $x, y \in \mathsf{X}$ that are not uniquely identifiable. It turns out that this graph is the same as Figure 1.3. This time, however, the solution to the problem is different. Suppose Alice and Bob agree on a coloring of the graph. When Bob receives a $u \in \mathsf{U}$, he can identify the subset $C \subseteq \mathsf{X}$ of Alice's inputs that could be associated with u. This subset is a *clique* of the graph, a set of mutually adjacent nodes (the complementary notion of an independent set). Therefore, in a coloring each of the elements of C must have a different color. Alice can send the color of her input x to Bob, and Bob will have no more doubts. How many bits of information are sufficient for Alice to send? If the chromatic number is c, then to communicate one out of c colors, Alice can send $\log_2(c)$ bits.

The discussion for multiple uses of the source is very similar to the one for the channel problem. To the best of our knowledge, the entangled version of this problem has not been discussed before.

Source-channel problem Consider the dual source problem, with the additional constraint that Alice and Bob communicate through a noisy channel. Source and channel are defined as before, with related characteristic and confusability graphs. This setting is represented in Figure 1.7. The question is the following:

> Alice and Bob have a dual source with Alice's input set X, Bob's side information set U and side information subsets $\{\mathsf{U}(x)\}_{x \in \mathsf{X}}$. They have access to a noisy channel from Alice to Bob, with input set S, output set V and output subsets $\{\mathsf{V}(s)\}_{s \in \mathsf{S}}$. What is the minimum ratio number of channel uses/number of source inputs, in order for Bob to recover Alice's inputs without chance of a mistake?

In the simplified single-source inputs scenario, the setting is as follows. Alice receives an input x, and sends a sequence of inputs $s_1, \ldots, s_k$ to Bob through the channel. Bob uses his side information u together with the channel outputs $v_1, \ldots, v_k$ to recover x without errors.

The answer to the question above is the source-channel rate, a function of the characteristic graph of the source and the confusability graph of the channel. This time, however, the solution does not depend on graph parameters but on *graph homomorphisms*. A homomorphism between two graphs is an edge-preserving map from the node set of the first graph to the node set of the second graph:

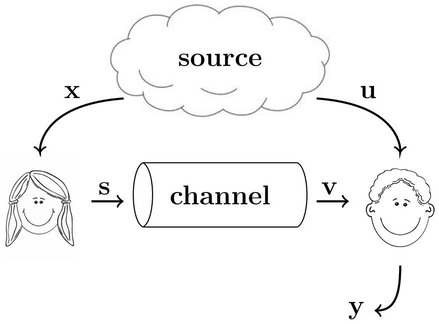


Figure 1.7: Setup of the source-channel problem.

adjacent nodes of the first graph are mapped to adjacent nodes of the second graph. When solving the source-channel problem, Alice and Bob find a homomorphism from the characteristic graph of the source to the complement of the confusability graph of (multiple uses of) the channel. An intuition for this is that they want to map non-uniquely-identifiable source inputs (edges in the source graph) to non-confusable channel inputs (non-edges in the channel graph).

Notice that, as intuition suggests, the source problem and channel problem can be seen as special cases of the source-channel problem. The source problem is the source-channel problem with a non-noisy binary channel (therefore one channel use per bit); the channel problem is the source-channel problem with a completely uncorrelated source (therefore Bob has no side information).

To the best of our knowledge, the entangled version of this problem has not been discussed before.

Our contribution In Chapter 4, we define the entangled version of the source-channel problem. We allow Alice and Bob to share entanglement and we give a rigorous mathematical description of the new setting. We prove many properties of our problem, including a relation with the ϑ number and with quantum graph parameters of Chapter 3. We also show that the entangled channel problem of [CLMW10] is a special case of our setup.

Then, we construct source-channel pairs where quantum players have a large advantage over classical players by extending a known class of graphs.

Remarkably, we use the celebrated *quantum teleportation* scheme [BB84] in the proofs of some of our results. To the best of our knowledge, this is the first time quantum teleportation was used as a tool in the zero-error channel capacity setting.

1.5 Basics of quantum theory

We now move to the technical part of this chapter, a brief introduction to quantum mechanics. We assume familiarity with the basic concepts of linear algebra, which are summarized in [NC00, Section 2.1]. Since this thesis is focused on non-locality and its applications, we will not discuss quantum circuits or quantum computing in detail. We will instead concentrate on the concepts of quantum state, evolution and measurement, especially in a bipartite setting. These notions are crucial to understand non-locality presented in Chapter 2 and its applications presented in Chapters 3 and 4.

State vectors A quantum system of finite dimension d lives in a d-dimensional complex inner product space, denoted by $\mathbb{C}^d$. Its quantum state is described by a unit vector in $\mathbb{C}^d$, called *state vector* (or *pure state*).

Let us illustrate these concepts with an example. A simple yet important quantum system is the *quantum bit*, frequently abbreviated as qubit. The quantum bit is the basic unit of quantum information, in the same way as the bit is the unit of classical information.

A classical bit can take value 0 or 1 and can be implemented by any classical physical system that can be in two distinct states, *e.g.*, the presence/absence of voltage in a wire, the reflection/absorption of light by the surface of a laser disc, the orientation of the magnetic field in hard drives. Analogously, a quantum bit can take values 0 and 1 and can be represented by any quantum physical system that can be in two distinct classical states. For example, consider the polarization of a photon: we can associate the value 0 to vertical polarization and 1 to horizontal polarization. These are two classical states, but a quantum system can be in a *superposition* of both.

In quantum mechanics, we can describe a quantum bit with a unit vector in $\mathbb{C}^2$. Let us define the two classical states as $|0\rangle = [1,0]^T$ and $|1\rangle = [0,1]^T$, following Dirac's notation.[13] These two orthonormal vectors are often called the *standard basis* or *computational basis*. The quantum bit can be in any state of the form

$$|\psi\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle = \begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix},$$

where α_0, α_1 are complex numbers satisfying $|\alpha_0|^2 + |\alpha_1|^2 = 1$. Thus, the state $|\psi\rangle$ of the quantum bit is a linear combination of the possible classical values, where the coefficients α_0, α_1 are called *amplitudes* and intuitively are "weights" indicating how much the qubit is towards 0 or 1.

[13] Dirac notation is standard in quantum information theory. A column vector ψ is written in the "ket" notation $|\psi\rangle$, while its complex conjugate transpose is written in the "bra" notation $\langle\psi|$. At first it may seem confusing, but it turns out to be very convenient in formulas. For example, given any two vectors ψ, ϕ of the same dimension, their "bra-ket" is their inner product $\langle\psi|\phi\rangle$ (a scalar) and their "ket-bra" is their outer product $|\psi\rangle\langle\phi|$ (a matrix).

Now consider two quantum systems, the first living in $\mathbb{C}^d$ and the second living in $\mathbb{C}^{d'}$. We can describe these two systems together as a larger quantum system living in the tensor space $\mathbb{C}^d \otimes \mathbb{C}^{d'}$. For example, consider two qubits. Their state can be described as a superposition of the basis states $\{|0\rangle \otimes |0\rangle, |0\rangle \otimes |1\rangle, |1\rangle \otimes |0\rangle, |1\rangle \otimes |1\rangle\}$, where $\otimes$ denotes the tensor product. They are often abbreviated as $\{|0\rangle|0\rangle, |0\rangle|1\rangle, |1\rangle|0\rangle, |1\rangle|1\rangle\}$ or even more concisely as $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$. Thus, two qubits can be in any state of the form

$$\alpha_{00}|00\rangle + \alpha_{01}|01\rangle + \alpha_{10}|10\rangle + \alpha_{11}|11\rangle,$$

with $|\alpha_{00}|^2 + |\alpha_{01}|^2 + |\alpha_{10}|^2 + |\alpha_{11}|^2 = 1$. In general, an n-qubit system is a unit vector that is a superposition of the 2^n basis states:

$$\sum_{i \in \{0,1\}^n} \alpha_i |i\rangle.$$

Evolution The evolution of d-dimensional quantum systems is described by *unitary matrices* in $\mathbb{C}^{d \times d}$. These are the transformations that preserve the norm of the vectors. Intuitively, these are the allowed operators because they map a state vector to another state vector, *i.e.*, preserve the norm 1.

Let us go back to our simple example to illustrate this new concept. While the only allowed operations on a classical bit are to flip the bit or to leave it untouched, there are infinitely many transformations that one can do on a quantum bit. We will see later a very interesting transformation. For now, let us just show the unitary transformation equivalent to a bit-flip, that is

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

We have that

$$X|\psi\rangle = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_0 \end{bmatrix}.$$

Thus, applying the X unitary on a quantum bit swaps its amplitudes for $|0\rangle$ and $|1\rangle$. Intuitively this means that if the qubit was "weighted more towards zero" it will now be "weighted more towards one".

Other important unitary matrices that we use later are the *Hadamard matrix* and the *Pauli matrices.* The (2×2) Hadamard matrix is defined as

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

We can easily construct Hadamard matrices that act on systems whose dimension is a power of two, for example n qubits. One can obtain such a Hadamard matrix by tensor products of 2×2 Hadamard matrices.[14] The matrix $H^{\otimes n}$ acts as follows on basis states:

$$H^{\otimes n}|i\rangle = \frac{1}{\sqrt{2^n}} \sum_{j \in \{0,1\}^n} (-1)^{i \cdot j} |j\rangle,$$

where $i \cdot j$ is the bitwise inner product of i and j. Hadamard matrices play an important role in Chapters 2 and 4. The Pauli matrices are the 2×2 matrices X (which we have already seen), the identity matrix I, together with

$$Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \text{ and } Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

They are important in Section 3.5 mainly because of the so-called Pauli group. This is a group of 16 matrices (I, X, Y, Z with coefficients ± 1 and $\pm i$) under matrix multiplication.

Measurement Unfortunately, we cannot directly access the state vector of a unknown quantum system. It would be nice if we could: for example, we could have extremely efficient communication by encoding a huge amount of data in the amplitudes of a single quantum bit. One of the most widely known counter-intuitive features of quantum mechanics is that "observation changes the state of the system". Let us see how it works, starting with our qubit example.

Measurements describe an experiment on a quantum state. Suppose we have a qubit in the state $|\psi\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle$ and we measure it *according to the standard basis*, *i.e.*, we check with the aid of a measurement device if this qubit is in the classical state $|0\rangle$ or $|1\rangle$. Quantum mechanics predicts that we observe outcome "0" with probability $|\alpha_0|^2$ and "1" with probability $|\alpha_1|^2$. That is why we need the state vector to be a unit vector: the squares of the amplitudes are probabilities, and they must sum to one! Quantum mechanics says that the state, after the measurement outcome "i", *collapses* to the classical state $|i\rangle$, therefore losing all the information contained in the amplitudes.

For most quantum algorithms, this is the end of the story. All we need is to manipulate the amplitudes of the qubits in a smart way and measure according to the standard basis in order to obtain the desired results. In this thesis, however, it is often crucial to use a more general form of measurement, known as *Positive Operator-Valued Measure (POVM)*. A t-outcome POVM is a collection

$$\mathsf{M} = \{E_i \in \mathbb{C}^{d \times d} : i \in [t]\}$$

[14] In general, a *Hadamard matrix* is a square matrix $A \in \{-1, 1\}^{\ell \times \ell}$ that satisfies $AA^{\mathsf{T}} = \ell I$. It is conjectured that there is a Hadamard matrix acting on spaces of every dimension that is multiple of 4. We do not know constructions for all such cases. (See Section 4.2.2.) Also, note that in the original definition, Hadamard matrices are ± 1-valued matrices. They have a normalization coefficient when used as quantum operations, to make them unitary.

of positive semidefinite[15] matrices $E_i \equiv M_i^\dagger M_i$ that satisfy $\sum_{i=1}^t E_i = I$, where I is the identity matrix. If we perform a t-outcome measurement M on a d-dimensional system with state vector $|\phi\rangle$, then we observe a random variable λ over the set $[t]$ whose probability distribution is given by

$$\Pr[\lambda = i] = \langle\phi|M_i^\dagger M_i|\phi\rangle.$$

In the event that $\lambda = i$, we say that we get measurement outcome i.

Before giving an example, let us define a simpler and important kind of measurement: the *projective measurement*. This is a special case of POVM where all elements M_i are projectors.[16] Therefore, for all i we have $E_i = M_i^\dagger M_i = M_i^2 = M_i$. Moreover, in a projective measurement all the elements are pairwise orthogonal, *i.e.*, for all $i \neq j$, $M_i M_j = 0$. We will use this kind of measurement and its properties extensively in Chapter 3. After a projective measurement gave outcome i, the state collapses to a new state

$$\frac{M_i|\phi\rangle}{||M_i|\phi\rangle||},$$

where $||\cdot||$ denotes the ℓ_2 norm of a vector. Notice that in this case, the state vector need not be a classical state, but it can still be in a quantum superposition.

Now we go back to our example, and illustrate how the measurement according to the standard basis can be seen as a projective measurement. Consider a qubit in state $|\psi\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle$ with real amplitudes. Let $\mathsf{M} = \{|0\rangle\langle 0|, |1\rangle\langle 1|\}$. Then, the probability of observing "0" can be written as

$$\langle\psi|(|0\rangle\langle 0|)^\dagger|0\rangle\langle 0||\psi\rangle = [\alpha_0, \alpha_1]\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix} = |\alpha_0|^2,$$

as explained before. It also turns out that the state after the measurement is[17]

$$\frac{|0\rangle\langle 0||\psi\rangle}{|||0\rangle\langle 0||\psi\rangle||} = \frac{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix}}{\left|\left|\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix}\right|\right|} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = |0\rangle.$$

Bipartite systems and entanglement In this thesis we often consider bipartite settings, where we usually have two players, called Alice and Bob, that share a quantum system. If Alice has a d-dimensional system and Bob has a d'-dimensional system, they share a system in $\mathbb{C}^d \otimes \mathbb{C}^{d'}$. Each party has direct

[15] A matrix is positive semidefinite if and only if all its eigenvalues are non-negative.

[16] Projectors are Hermitian matrices that satisfy $M^2 = M$.

[17] Up to a global phase. If two state vectors differ only by a scalar of absolute value 1, then they are physically indistinguishable.

access to his/her part of the state only. Therefore, if the parties are separated and unable to communicate, they can only perform *local operations*, *i.e.*, the unitary matrices are in tensor product form $A \otimes B$, where A acts on $\mathbb{C}^d$ and B acts on $\mathbb{C}^{d'}$. Moreover, if Alice performs a measurement $\{E_i\}_{i\in[k]}$ and Bob $\{F_j\}_{j\in\ell}$ on a shared system in state $|\phi\rangle$, the probability of outcome "i" for Alice and "j" for Bob is $\langle\phi|E_i \otimes F_j|\phi\rangle$. After a projective measurement, the state collapses as explained above.

The state itself need not be in a product form. Indeed, there exist states in $\mathbb{C}^d \otimes \mathbb{C}^{d'}$ that cannot be written as $|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$. These states are called *entangled*. One famous example, which we mentioned in Section 1.1, is the EPR pair, named after the authors of [EPR35]:

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle).$$

States of the form $|\psi\rangle^{\otimes k}$, *i.e.*, k EPR pairs, are very important for the results of Chapters 2 and 3.[18]

Density matrices There are two equivalent ways of representing and working with quantum states. The one we have just seen is based on state vectors, while the second one is based on *density matrices*. Both formulations have their own merits. The first is arguably simpler and deals very well with describing quantum algorithms, while the second is almost essential in multipartite settings, in order to easily describe sub-systems and mixtures of quantum states.[19] As many authors in the field, we will switch between the two formalism when convenient. That is why we introduce both of them.

The set of possible *states* of a d-dimensional quantum system is formed by the $d \times d$ complex positive semidefinite matrices whose trace equals 1. These

[18] At this point, let us clarify a common abuse of notation regarding EPR pairs. Strictly speaking, it is not correct that two EPR pairs are equal to the state $|\psi\rangle^{\otimes 2}$. One must take into account who possesses which qubit and rearrange the qubits after performing the tensor product. Let us denote by $|i\rangle_A$ a basis state of Alice's system, and by $|i\rangle_B$ a basis state of Bob's system. Then we have

$$\begin{aligned}|\psi\rangle \otimes |\psi\rangle =& \frac{1}{\sqrt{2}}(|0\rangle_A \otimes |0\rangle_B + |1\rangle_A \otimes |1\rangle_B) \otimes \frac{1}{\sqrt{2}}(|0\rangle_A \otimes |0\rangle_B + |1\rangle_A \otimes |1\rangle_B)\\ =& \frac{1}{2}[(|0\rangle_A \otimes |0\rangle_B \otimes |0\rangle_A \otimes |0\rangle_B)) + (|0\rangle_A \otimes |0\rangle_B \otimes |1\rangle_A \otimes |1\rangle_B)\\ &+ (|1\rangle_A \otimes |1\rangle_B \otimes |0\rangle_A \otimes |0\rangle_B) + (|1\rangle_A \otimes |1\rangle_B \otimes |1\rangle_A \otimes |1\rangle_B)]\\ =& \frac{1}{2}(|00\rangle_A|00\rangle_B + |01\rangle_A|01\rangle_B + |10\rangle_A|10\rangle_B + |11\rangle_A|11\rangle_B),\end{aligned}$$

where the last equality follows from the rearrangement of the qubits.

[19] Mixtures are simply probability distributions on quantum states. Any similarity with "mixed strategies" from game theory should not come as a surprise. Indeed, the term was introduced by von Neumann, who worked on both fields.

are called *density matrices.* A pure state $|\psi\rangle$ has density matrix $|\psi\rangle\langle\psi|$. A mixed state that is in pure state $|\psi_i\rangle$ with probability p_i has density matrix $\sum_i p_i|\psi_i\rangle\langle\psi_i|$. Distinct mixtures of pure states may have the same density matrix.

An operation on $\rho \in \mathbb{C}^{d\times d}$ is a mapping $\rho \mapsto U\rho U^\dagger$ induced by the unitary U. A t-outcome measurement M on a system in state ρ is a collection of positive semidefinite matrices $E_i = M_i^\dagger M_i$, as described above. By performing a measurement, one observes a random variable λ over the set $[t]$ whose probability distribution is given by

$$\Pr[\lambda = i] = \mathrm{Tr}(E_i\rho) = \mathrm{Tr}(M_i\rho M_i^\dagger).$$

After the measurement with outcome i, the state collapses to

$$\frac{M_i\rho M_i^\dagger}{\mathrm{Tr}(M_i\rho M_i^\dagger)}.$$

The possible states of a *pair* of quantum systems $(\mathcal{A},\mathcal{B})$ of dimensions d and d' respectively are the trace-1 positive semidefinite matrices in $\mathbb{C}^{d\times d}\otimes\mathbb{C}^{d'\times d'}$. A useful way to describe a part of the system is with a linear operator called the *partial trace*, which we introduce now. For matrices $A \in \mathbb{C}^{d\times d}$ and $B \in \mathbb{C}^{d'\times d'}$ define $\mathrm{Tr}_{\mathcal{A}}(A\otimes B) = \mathrm{Tr}(A)B$ and $\mathrm{Tr}_{\mathcal{B}}(A\otimes B) = A\mathrm{Tr}(B)$, and extend these definitions linearly to all matrices of $\mathbb{C}^{d\times d}\otimes\mathbb{C}^{d'\times d'}$.

The pair of systems $(\mathcal{A},\mathcal{B})$ is said to be in an *entangled* state if it is in a state ρ which is *not* a convex combination of states of the form $\rho_A\otimes\rho_B$. A pure state ρ in $\mathbb{C}^{d\times d}\otimes\mathbb{C}^{d\times d}$ is said to be *maximally entangled* if $\mathrm{Tr}_{\mathcal{A}}(\rho) = \mathrm{Tr}_{\mathcal{B}}(\rho) = \frac{1}{d}I$.

Now we consider the setting where Alice and Bob hold (possibly entangled) quantum systems $\mathcal{A}$ and $\mathcal{B}$, respectively, and they each perform a measurement.

Suppose that the pair $(\mathcal{A},\mathcal{B})$ is in the state ρ and that Alice performs a t-outcome measurement $\mathsf{M} = \{E_i\}_{i\in[t]}$ on $\mathcal{A}$. Then, the probability that Alice gets measurement outcome i equals $p_i = \mathrm{Tr}\big((E_i\otimes I)\rho\big)$. Moreover, in the event that Alice gets measurement outcome i, Bob's system $\mathcal{B}$ is left in the state $\rho_i = \mathrm{Tr}_{\mathcal{A}}\big((E_i\otimes I)\rho\big)/p_i$. If Bob now performs an r-outcome measurement $\mathsf{M}' = \{F_j\}_{j\in[r]}$ on $\mathcal{B}$, then the probability that he gets outcome $j\in[r]$ equals $\mathrm{Tr}(F_j\rho_i)$.

Chapter 2

Strong Bell inequality violations

This chapter is based on the paper *"Near-optimal and explicit Bell inequality violations"*, by H. Buhrman, O. Regev, the author and R. de Wolf.

The paper was first presented at the Quantum Information Processing conference in January 2011 as a featured talk. Then, in June 2011 it was presented at the IEEE Conference on Computational Complexity and a shorter version was published in the conference proceedings. Finally, the full paper was published in the journal "Theory of Computing" in December 2012.

2.1 Introduction

One of the most striking features of quantum mechanics is the fact that *entangled* particles can exhibit correlations that cannot be reproduced or explained by classical physics, or more precisely, by "local hidden-variable theories." This was first noted by Bell [Bel64] in response to Einstein, Podolsky, and Rosen's challenge to the completeness of quantum mechanics [EPR35]. Experimental realization of such correlations is the strongest proof we have that nature does not behave according to classical physics: nature cannot simultaneously be "local" (meaning that information does not travel faster than the speed of light) and "realistic" (meaning that measurable properties of particles such as its spin always have a definite—if possibly unknown—value). Many such experiments have been done. All behave in accordance with the predictions of quantum mechanics, though so far none has closed all "loopholes" that would allow some (usually very contrived) classical explanation of the observations based on imperfect behavior of, for instance, the photon detectors used.

Here we study quantitatively *how much* such correlations obtained from entangled quantum systems can deviate from what is achievable classically. It will be convenient to describe our results in terms of two-player *games*, which are described as follows. Two non-communicating parties, called Alice and Bob, receive inputs x and y according to some fixed and known probability distribution π, and are required to produce outputs a and b, respectively. There is a predicate specifying which outputs a, b are correct on inputs x, y. The definition of a game G consists of this predicate and the distribution π. The goal is to design games where entangled strategies have much higher winning probability than the best classical strategy. While this setting is used to study non-locality in physics, the same set-up is also used extensively to study the power of entanglement in computer science contexts like multi-prover interactive proofs [KKM$^+$11, KKMV09], parallel repetition [CSUU07, KRT10], and cryptography.

Entangled strategies start out with an arbitrary fixed entangled state. No communication takes place between Alice and Bob. For each input x, Alice has a measurement, and for each input y, Bob has a measurement. They apply the measurements corresponding to x and y to their halves of the entangled state, producing classical outputs a and b, respectively. Their goal is to maximize the winning probability. The *entangled value* $\omega^*(G)$ of the game is the supremum of the winning probability, taken over all entangled strategies. When restricting to strategies that use entanglement of local dimension n, the value is denoted $\omega^*_n(G)$. This should be contrasted with the *classical value* $\omega(G) = \omega^*_1(G)$ of the game, which is the maximum among all classical, non-entangled strategies. Shared randomness between the two parties is allowed, but is easily seen not to be beneficial.

The remarkable fact, alluded to above, that some entanglement-based correlations cannot be simulated classically, corresponds to the fact that there are games G for which the entangled value $\omega^*(G)$ is strictly larger than the classical value $\omega(G)$. The CHSH game is one particularly famous example [CHSH69]. Here, the inputs $x \in \{0,1\}$ and $y \in \{0,1\}$ are uniformly distributed, and Alice and Bob win the game if their respective outputs $a \in \{0,1\}$ and $b \in \{0,1\}$ satisfy $a \oplus b = x \wedge y$; in other words, a should equal b unless $x = y = 1$. The classical value of this game is easily seen to be $\omega(G) = 3/4$, while the entangled value is known to be $\omega^*(G) = 1/2 + 1/(2\sqrt{2}) \approx 0.85$. The entangled value is achieved already with 2-dimensional entanglement (*i.e.*, one EPR pair), so $\omega^*(G) = \omega^*_2(G)$ for this game [Tsi87].

One common figure of merit of a game is that of the *Bell inequality violation* exhibited by a game, which is defined as the ratio of entangled and classical values. More generally, we allow to replace the *values* (which are the maximum winning probabilities) by *biases* around some arbitrary "center" $p \in [0,1]$, where by bias we mean the maximum distance of the winning probability from p. For instance, by using $p = 1/2$ as the center, one can see that the CHSH game above exhibits a Bell inequality violation of $\sqrt{2}$.[1] In Section 2.2.2 we explain the origin of the term "Bell inequality violation," define it more formally, and explain the close relationship between games and Bell inequalities.

[1] Notice that this requires that the winning probability of non-entangled players is always between 1/4 and 3/4, which is clearly the case.

In two recent papers, Junge *et al.* [JPP$^+$10, JP11] studied how large a Bell inequality violation one can obtain. In terms of upper bounds, [JPP$^+$10] proved that the maximum Bell inequality violation $\omega_n^*(G)/\omega(G)$ obtainable with entangled strategies of local dimension n, is at most $O(n)$, and [JP11, Theorem 6.8] proved that if Alice and Bob have at most k possible outputs each, then the violation $\omega^*(G)/\omega(G)$ is at most $O(k)$, irrespective of the amount of entanglement they can use. (This improved an earlier $O(k^2)$ upper bound due to Degorre *et al.* [DKLR09], and was also obtained for the special case of games by Dukaric [Duk10, Theorem 4].)

In terms of lower bounds, [JPP$^+$10] showed the existence of a Bell inequality violation of $\Omega(\sqrt{n}/(\log n)^2)$, where n is both the entanglement dimension and the number of outputs of Alice and Bob. This was improved to $\sqrt{n}/\log n$ in [JP11]. Both constructions are probabilistic, and the proofs show that with high probability the constructed games exhibit a large violation, yielding the existence of such games, without giving an explicit formulation. Their proofs are heavily based on connections to the mathematically beautiful areas of Banach spaces and operator spaces, but as a result are arguably somewhat inaccessible to those unfamiliar with these areas, and it is difficult to get a good intuition for them. (It is actually possible to analyze their game and reprove many of their results—often with improved parameters—using elementary probabilistic techniques [Reg12].)

Our main result in this chapter is to exhibit two fully explicit games with strong Bell inequality violations. The first achieves the same violation as [JP11], namely $\sqrt{n}/\log n$, where n is both the number of possible outputs and the dimension of entanglement. The second achieves the much stronger violation of $n/(\log n)^2$, which is optimal up to a polylogarithmic factor by the results of [JPP$^+$10, JP11]. Even though the second game gives a much stronger violation, the first one still has some merit; for instance, entanglement allows the players to win it with certainty. Interestingly, although addressing a question in mathematical physics, both games are inspired by earlier work in theoretical computer science (communication complexity and unique games, respectively), and so is their analysis. In the remainder of this introduction we provide an overview of the two games.

2.1.1 The Hidden Matching game

The "Hidden Matching" problem was introduced in quantum communication complexity by Bar-Yossef *et al.* [BYJK08], and many variants of it were subsequently studied [GKRdW09, GKK$^+$08, Gav09]. The original version is as follows, where it should be noted that now we allow communication, in contrast to the setting of non-local games. Let n be a power of 2. Alice is given input $x \in \{0,1\}^n$ and Bob is given a perfect matching M, *i.e.*, a partition of the set $[n] = \{1, \dots, n\}$

into $n/2$ disjoint pairs $\{i, j\}$. Both inputs are uniformly distributed.[2] We allow one-way communication from Alice to Bob, and Bob is required to output a pair $\{i, j\} \in M$ and a bit $v \in \{0, 1\}$. They win if $v = x_i \oplus x_j$.

In Section 2.3.1 we show that if Alice sends Bob a c-bit message, then their optimal winning probability is $\frac{1}{2} + \Theta(\frac{c}{\sqrt{n}})$. Bar-Yossef *et al.* [BYJK08] earlier proved this for $c = \Theta(\sqrt{n})$, using information theory. However, their tools seem unable to give good bounds on the success probability for much smaller c. Instead, the main mathematical tool we use in our analysis is the so-called "KKL inequality" [KKL88] from Fourier analysis of Boolean functions. Roughly speaking, this inequality implies that if the message that Alice sends about x is short, then Bob will not be able to predict the parity $x_i \oplus x_j$ well for many $\{i, j\}$ pairs. His matching M is uniformly distributed, independent of x, and contains only $n/2$ of all $\binom{n}{2}$ possible $\{i, j\}$ pairs. Hence it is unlikely that he can predict any one of those $n/2$ parities well. The KKL inequality was used before to analyze another variant of Hidden Matching in [GKK$^+$08], though their analysis is different and more complicated because their variant of Hidden Matching is a promise problem with a non-product input distribution.

The following non-local version of the Hidden Matching problem (and the entangled strategy for it) is originally due to Buhrman, and related problems were studied in [GKRdW09, Gav09].

Definition 2.1.1 (Non-Local Hidden Matching Game ($\mathrm{HM}_n^{\mathrm{NL}}$))**.** *Let n be a power of 2 and $\mathcal{M}_n$ be the set of all perfect matchings on the set $[n]$. Alice is given $x \in \{0, 1\}^n$ and Bob is given $M \in \mathcal{M}_n$, both distributed according to the uniform distribution. Alice and Bob do not communicate. Alice's output is a string $a \in \{0, 1\}^{\log n}$ and Bob's output is an $\{i, j\} \in M$ and $d \in \{0, 1\}$. They win the game if and only if*

$$(a \cdot (i \oplus j)) \oplus d = x_i \oplus x_j, \tag{2.1}$$

where the dot indicates inner product (modulo 2) of two $\log n$-bit strings.

Observe that Alice has n possible outputs a and Bob has $2 \cdot n/2 = n$ possible outputs $(\{i, j\}, d)$ given his matching.

A classical strategy that wins this game induces a protocol for the original Hidden Matching problem with communication $c = \log n$ bits and the same winning probability p, as follows. Alice sends Bob the $\log n$-bit output a from the non-local strategy, Bob computes $v = (a \cdot (i \oplus j)) \oplus d$ and outputs $(\{i, j\}, v)$. We

[2] All our results also hold with minor modifications for the case that Bob's matching is chosen uniformly from the set $\{M_k \mid k \in \{0, \ldots, n/2-1\}\}$, where the matching M_k consists of the pairs $\{i, j\}$ where $i \leq n/2$ and $j = n/2 + 1 + (i + k - 1 \bmod n/2)$. This has the advantage of lowering the number of possible inputs to Bob to $n/2$. The main thing is to notice that Eq. (2.4) on page 33 still holds with respect to this new distribution on Bob's matching if we replace the right-hand side by $2/n$, and the rest of the proof goes through.

have that $v = x_i \oplus x_j$ with probability p. Hence, our bound for the original communication problem implies that no classical strategy can win with probability that differs from 1/2 by more than $O(\frac{\log n}{\sqrt{n}})$, as explained in Section 2.3.5.

In contrast, there is a strategy that wins with probability 1 using $\log n$ EPR pairs, which shows $\omega_n^*(G) = 1$.[3] This game therefore exhibits a Bell violation of $\Omega(\sqrt{n}/\log n)$ (by measuring the maximal deviation of the winning probability from 1/2). This order is the same as that obtained by Junge *et al.* [JPP$^+$10, JP11], but our game is fully explicit and arguably simpler (which would help any future experimental realization). One might feel that our reduction to a communication complexity lower bound is responsible for losing the $\log n$ factor; however in Theorem 2.3.8 we exhibit a classical strategy with winning probability $1/2 + \Omega(\sqrt{\log(n)/n})$. This shows that at least the square root of the log-factor is really necessary.

2.1.2 The Khot-Vishnoi game

Our second non-local game derives from the work of Khot and Vishnoi [KV05] on the famous *Unique Games Conjecture* (UGC), introduced by Khot [Kho02]. Although not necessary for the rest of this chapter, we now provide some background and motivation. Roughly speaking, the UGC says that approximating the classical value of so-called unique games is a hard problem, even if we are only interested in a very rough approximation that can tell the difference between value less than ε and value more than $1 - \varepsilon$. This conjecture implies many other hardness-of-approximation results that do not seem obtainable using the more standard techniques based on the PCP theorem. Khot and Vishnoi considered the standard semidefinite programming (SDP) relaxation of the classical value and showed that there are games for which it provides a very poor approximation, in the sense that the classical value is close to 0, yet the SDP relaxation is close to 1. This so-called integrality gap demonstrates that the standard SDP relaxation, which can be computed efficiently, does not lead to an algorithm contradicting the UGC.

Kempe, Regev, and Toner [KRT10] already observed that they could combine their "quantum rounding" technique with the game of [KV05] to get a game with n possible outputs exhibiting a Bell inequality violation of n^ε for some small constant $\varepsilon > 0$, using entanglement dimension n. Our main contribution in the second part of this chapter is a refined (and simpler) analysis of both the Khot-Vishnoi game and of the quantum rounding technique. We show that, somewhat surprisingly, nearly optimal violations can be obtained using this method.

[3] The reader might be a bit confused by the seeming overloading of the meaning of 'n'. Formally, 'n' is a parameter in the specification of the game. As it happens, for both of our games it is also the number of possible outputs for each player, *and* the local dimension of the entangled state that our strategy uses (though we do not claim that this entanglement-dimension n is necessary to achieve the best-possible entangled value).

We first give a precise definition of the Khot-Vishnoi (KV) game.

Definition 2.1.2 (Khot-Vishnoi Game (KV_n))**.** *The game is parametrized by an integer n, which we assume to be a power of 2, and a "noise parameter" $\eta \in [0, 1/2]$. Consider the group $(\{0,1\}^n, \oplus)$ of n-bit strings together with bitwise addition mod 2, and let H be the subgroup containing the n Hadamard codewords.*[4] *This subgroup partitions $\{0,1\}^n$ into $2^n/n$ cosets of n elements each. Alice receives a uniformly random coset x as input, which we can think of as $u \oplus H$ for uniformly random $u \in \{0,1\}^n$. Bob receives a coset y obtained from Alice's by adding a string of low Hamming weight, namely $y = x \oplus z = u \oplus z \oplus H$, where each bit of $z \in \{0,1\}^n$ is set to 1 with probability η, independently of the other bits. Addition of z gives a natural bijection between the two cosets, mapping each element of the first coset to a relatively nearby element of the second coset; namely, the distance between the two elements is the Hamming weight of z, which is typically around ηn. Each player is supposed to output one element from its coset, and their goal is for their elements to match under the bijection. In other words, Alice outputs an element $a \in x$, Bob outputs $b \in y$, and they win the game if and only if $a \oplus b = z$.*[5]

Notice that the number of possible inputs to each player is $2^n/n$ and the number of possible outputs for each player is n.

Based on the integrality gap analysis of Khot and Vishnoi, in Section 2.4 we show that no classical strategy can win this game with probability greater than $1/n^{\eta/(1-\eta)}$. We also sketch a classical strategy that achieves this winning probability up to lower order terms. In contrast, using a simplified version of the "quantum rounding" technique of [KRT10], we exhibit an entangled strategy that uses the n-dimensional maximally entangled state and wins with probability at least $(1-2\eta)^2$. This strategy follows from the observation that each coset of H defines an orthonormal basis of $\mathbb{R}^n$ in which we can do a measurement. Summarizing, we have entangled value $\omega_n^*(G) \geq (1-2\eta)^2$ and classical value $\omega(G) \leq 1/n^{\eta/(1-\eta)}$ for this game. Setting the noise-rate to $\eta = 1/2 - 1/\log n$, the entangled value is $\Omega(1/(\log n)^2)$ while the classical value is $O(1/n)$, leading to a Bell inequality violation $\omega_n^*(G)/\omega(G) = \Omega(n/(\log n)^2)$. Up to the polylogarithmic factor, this is optimal both in terms of the local dimension, and in terms of the number of possible outputs.

[4] For $a \in \{0,1\}^{\log n}$, the corresponding n-bit Hadamard codeword is defined as $h(a) := (a \cdot j)_{j \in \{0,1\}^{\log n}}$.

[5] Note that the winning condition for this game is a "randomized predicate," as there are n possible ways to obtain the same y from x, hence there are n possible winning predicates (one for each $z \in x \oplus y$) corresponding to each pair of inputs x, y. Strictly speaking, this requires a slightly more general definition of a game than the one given in the introduction; see the definition in Section 2.2.2. Although not relevant for any of our applications, we mention that one can modify the game in a straightforward manner, making it a game with a deterministic predicate. The thing to observe is that with very high probability exactly one of the n predicates dominates, namely the one corresponding to a z of Hamming weight around ηn.

Palazuelos [Pal12b] recently used our result to prove an interesting *super-activation* result. He identified a constant-dimensional quantum state (a mixture of the maximally entangled state of local dimension 8 with the completely mixed state) that cannot be used to violate any Bell inequality, but a sufficiently large number of copies of which *can* be used to violate a Bell inequality—namely the one associated with our KV game.

2.2 Preliminaries

2.2.1 Fourier analysis

The crucial technical tool used in the analysis of both of our games is Fourier analysis on the Boolean cube. We will just introduce what we need here, referring to [O'D08, Wol08] for more details and references. Unless mentioned otherwise, expectations and probabilities are taken over a uniformly random $x \in \{0,1\}^n$. Define the inner product between functions $f, g : \{0,1\}^n \to \mathbb{R}$ as

$$\langle f, g\rangle = \frac{1}{2^n} \sum_{x\in\{0,1\}^n} f(x)g(x) = \mathbb{E}[f(x)\cdot g(x)].$$

For $S \subseteq [n]$, the function $\chi_S(x) = (-1)^{\sum_{i\in S} x_i}$ is the parity (represented as a sign) of the variables indexed in S. These 2^n functions (one for each S) form an orthonormal basis for the space of all real-valued functions on the Boolean cube. The *Fourier coefficients* of f are $\widehat{f}(S) = \langle f, \chi_S\rangle$, and we can write f in its Fourier decomposition

$$f = \sum_{S\subseteq[n]} \widehat{f}(S)\chi_S.$$

Since the χ_S form an orthonormal basis, one can show

$$\langle f, g\rangle = \sum_S \widehat{f}(S)\widehat{g}(S). \tag{2.2}$$

For $p \geq 1$, the *p-norm* of f is defined as

$$\|f\|_p = \mathbb{E}[|f(x)|^p]^{1/p}.$$

This is monotone non-decreasing in p. For $p = 2$, Eq. (2.2) with $g = f$ gives Parseval's identity:

$$\|f\|_2^2 = \sum_S \widehat{f}(S)^2.$$

For $\rho \in [-1, 1]$, the *noise-operator* T_ρ adds "η-noise" to each of the input bits, where $\eta = (1-\rho)/2$. More precisely, the function $T_{1-2\eta}f$ is defined as

$$(T_{1-2\eta}f)(x) = \mathbb{E}_z[f(x \oplus z)],$$

where $z \in \{0,1\}^n$ is an "η-biased noise string," *i.e.*, each bit z_i is set to 1 with probability η, independently of the other bits. The linear operator T_ρ is diagonal in the Fourier basis: it just multiplies each χ_S by the factor $\rho^{|S|}$. Equivalently,

$$\widehat{T_\rho f}(S) = \rho^{|S|}\widehat{f}(S). \tag{2.3}$$

Since $T_\rho f$ is a convex combination of shifted copies $f_z(x) = f(x \oplus z)$ of f, by the triangle inequality we see that T_ρ is a contraction: $\|T_\rho f\|_p \leq \|f\|_p$ for all $p \geq 1$, $\rho \in [-1,1]$, and f. The *Bonami-Beckner hypercontractive inequality* implies $\|T_\rho f\|_q \leq \|f\|_p$ even for q somewhat bigger than p.

Theorem 2.2.1 (Bonami-Beckner). *For $f : \{0,1\}^n \to \mathbb{R}$, $1 \leq p \leq q$ and $\rho^2 \leq (p-1)/(q-1)$, we have*

$$\|T_\rho f\|_q \leq \|f\|_p.$$

An important consequence of the hypercontractive inequality is the so-called KKL inequality [KKL88].

Theorem 2.2.2 (KKL). *For $f : \{0,1\}^n \to \{-1,0,+1\}$ and $\delta \in [0,1]$, we have*

$$\sum_S \delta^{|S|}\widehat{f}(S)^2 \leq \Pr[f(x) \neq 0]^{2/(1+\delta)}.$$

Proof. Applying Theorem 2.2.1 with $q=2$, $p=1+\delta$, and $\rho=\sqrt{\delta}$ gives $\|T_\rho f\|_2 \leq \|f\|_p$. Note that because of the range of f, the right-hand side equals $\Pr[f(x) \neq 0]^{1/(1+\delta)}$. Squaring both sides and rewriting the left-hand side using Parseval's identity completes the proof. □

2.2.2 A more formal look at Bell violations

Before we analyze the two games mentioned above, let us first say something more about the mathematical treatment of general Bell inequalities. Readers who are content with the above (more concrete) approach in terms of winning probabilities of games, may safely skip this section.

Consider a game with n possible inputs to each player and k possible outputs. The behavior of the players (irrespective of whether they use a classical or an entangled strategy) can be summarized in terms of n^2 probability distributions, each on the set $[k] \times [k]$. We denote by $P(ab|xy)$ the probability of producing outputs a and b when given inputs x and y, with respect to a fixed strategy. As described in the introduction, a game is defined by a probability distribution π on the input set $[n] \times [n]$, as well as a (possibly randomized) predicate on $[k] \times [k]$ for each input pair (x,y). The winning probability of the players can be written as

$$\langle M, P\rangle = \sum_{abxy} M^{ab}_{xy} P(ab|xy).$$

where M_{xy}^{ab} is defined as the probability of the input pair (x, y) multiplied by the probability that the output pair (a, b) is accepted on this input pair. We call $M = (M_{xy}^{ab})$ the *Bell functional corresponding to the game.* More generally, a *Bell functional* is an arbitrary tensor $M = (M_{xy}^{ab})$ containing n^2k^2 real numbers.

We define the *classical value* of a Bell functional M as

$$\omega(M) = \sup_P |\langle M, P\rangle|,$$

where the supremum is over all distributions P representing classical strategies. Similarly, the *entangled value* of M is defined as

$$\omega^*(M) = \sup_P |\langle M, P\rangle|,$$

where the supremum now is over all entangled strategies (using an entangled state of arbitrary dimension). If the entangled state is restricted to local dimension n, the value is denoted $\omega_n^*(M)$. We note that if M is the Bell functional corresponding to a game, then these definitions coincide with our definitions from the introduction, and in this case the absolute value is unnecessary since M is non-negative.

A *Bell inequality* is an upper bound on $\omega(M)$ for some Bell functional M; it shows a limitation of *classical* strategies.[6] The *Bell inequality violation* demonstrated by a Bell functional M is defined as the ratio between the entangled and the classical value

$$\frac{\omega^*(M)}{\omega(M)}.$$

This provides a convenient quantitative way to measure the extra power provided by entanglement. This definition of Bell violation enjoys a rich mathematical structure, as witnessed by the numerous connections found to Banach space and operator space theory [JPP$^+$10, JP11, Duk10], and also has a beautiful geometrical interpretation as the “distance” between the set of all classical strategies and the set of all entangled strategies. (See Section 6.1 in [JP11].)

Clearly, any game G for which $\omega^*(G) \geq K\omega(G)$ gives a Bell violation of K by just taking the functional corresponding to G. But recall that in the introduction we said that one is also allowed to look at the ratio of biases around some center probability p. We now explain why this still agrees with the above definition of Bell violation. We claim that if G is a game for which the winning probability of any classical strategy cannot deviate from p by more than δ_1 and, moreover, there is an entangled strategy obtaining winning probability at least $p + \delta_2$ (or at most $p - \delta_2$), then we obtain a Bell violation of δ_2/δ_1. To see why, let M be the functional corresponding to the game, and let M' be the functional obtained by

[6] An upper bound on $\omega^*(M)$ is known as a *Tsirelson inequality*, and shows a limitation of entangled strategies.

subtracting from each M_{xy}^{ab} the probability of input pair (x, y) times p. Then, one can see that for any strategy P, $\langle M', P\rangle = \langle M, P\rangle - p$. Hence, $\omega(M')$ and $\omega^*(M')$ are exactly the bias around p of classical and entangled strategies, respectively, and the claim follows. The converse to this statement is also true: any Bell functional M can be converted to a game in such a way that the ratio between the entangled bias and the classical bias of the game (both around $1/2$, say) is exactly the Bell violation demonstrated by the functional. To prove this, consider the game in which an input pair (x, y) is chosen uniformly, and outputs a, b are accepted with probability $1/2 + \delta M_{xy}^{ab}$ for some sufficiently small $\delta > 0$ so that all these probabilities are in $[0, 1]$.

2.3 Hidden Matching game

In this section we define and analyze the Hidden Matching game. Here and below, unless stated otherwise, all probabilities and expectations are taken over the distributions on x and M specified by the game.

2.3.1 The Hidden Matching problem in communication complexity

While our focus is non-locality, it will actually be useful to first study the original version of the Hidden Matching problem in the context of protocols where communication from Alice to Bob is allowed. Both the problem and the efficient quantum protocol below come from [BYJK08].

Definition 2.3.1 (Hidden Matching (HM_n)). *Let n be a power of 2 and $\mathcal{M}_n$ be the set of all perfect matchings on the set $[n]$. Alice is given $x \in \{0, 1\}^n$ and Bob is given $M \in \mathcal{M}_n$, both distributed according to the uniform distribution. We allow one-way communication from Alice to Bob, and Bob outputs an $\{i, j\} \in M$ and $v \in \{0, 1\}$. They win if $v = x_i \oplus x_j$.*

Theorem 2.3.2. *For every n that is a power of 2, there is a protocol for* HM_n *with* $\log n$ *qubits of one-way communication that wins with probability 1 (*i.e., *$v = x_i \oplus x_j$ always holds).*

Proof. The protocol is the following:

1. Alice sends Bob the state $|\psi\rangle = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (-1)^{x_i} |i\rangle$.

2. Bob measures $|\psi\rangle$ in the n-element basis $B = \{\frac{1}{\sqrt{2}}(|i\rangle \pm |j\rangle) \mid \{i, j\} \in M\}$. If the outcome of the measurement is a state $\frac{1}{\sqrt{2}}(|i\rangle + |j\rangle)$ then Bob outputs $\{i, j\}$ and $v = 0$. If the outcome of the measurement is a state $\frac{1}{\sqrt{2}}(|i\rangle - |j\rangle)$, Bob outputs $\{i, j\}$ and $v = 1$.

For each $\{i,j\} \in M$, the probability to get $\frac{1}{\sqrt{2}}(|i\rangle + |j\rangle)$ equals $2/n$ if $x_i \oplus x_j = 0$, and equals 0 otherwise, and similarly for $\frac{1}{\sqrt{2}}(|i\rangle - |j\rangle)$ with the parity flipped. Hence Bob's output is always correct. □

2.3.2 Limits of classical protocols for HM_n

Here we show that classical protocols with little communication cannot have good success probability. To start, note that a protocol that uses shared randomness is just a probability distribution over deterministic protocols, hence the maximal winning probability is achieved by a deterministic protocol.

Theorem 2.3.3. *Every classical deterministic protocol for* HM_n *with* c *bits of one-way communication, where Bob outputs* $(\{i,j\}, v)$*, has*

$$\Pr[v = x_i \oplus x_j] \leq \frac{1}{2} + \frac{c+1}{\sqrt{n-1}}.$$

The intuition behind the proof is the following. If the communication c is small, the set X_m of inputs x for which Alice sends message m will typically be large (of size about 2^{n-c}), meaning Bob has little knowledge of most of the bits of x. The KKL inequality implies that for most of the $\binom{n}{2}$ pairs $\{i,j\}$, Bob cannot guess the parity $x_i \oplus x_j$ well. Of course, Bob has some freedom in which $\{i,j\}$ he outputs, but that freedom is limited to the $n/2$ pairs $\{i,j\}$ in his matching M, and it turns out that on average he will not be able to guess any of those parities well.

Proof. Fix a classical deterministic protocol. For each $m \in \{0,1\}^c$, let $X_m \subseteq \{0,1\}^n$ be the set of Alice's inputs for which she sends message m. These sets X_m together partition Alice's input space $\{0,1\}^n$. Define $p_m = |X_m|/2^n$. Note that $\sum_m p_m = 1$, so $\{p_m\}$ is a probability distribution over the 2^c messages m.

For each m, define the following probability distribution over all possible pairs $\{i,j\}$:

$$q_m(\{i,j\}) = \Pr[\text{Bob outputs } \{i,j\} \mid \text{Bob received } m].$$

We have

$$q_m(\{i,j\}) \leq \frac{1}{n-1}, \tag{2.4}$$

because we assume Bob always outputs an element from M and for fixed $i \neq j$ we have $\Pr[\{i,j\} \in M] = 1/(n-1)$, since each j is equally likely to be paired up with i under the uniform distribution on M. This implies

$$\sum_{\{i,j\}} q_m(\{i,j\})^2 \leq \max_{\{i,j\}} q_m(\{i,j\}) \cdot \sum_{\{i,j\}} q_m(\{i,j\}) = \max_{\{i,j\}} q_m(\{i,j\}) \leq \frac{1}{n-1}. \tag{2.5}$$

Define ε such that $\Pr[v = x_i \oplus x_j] = \frac{1}{2} + \varepsilon$, and ε_m such that $\Pr[v = x_i \oplus x_j \mid \text{Bob received } m] = \frac{1}{2} + \varepsilon_m$. Then $\varepsilon = \sum_m p_m \varepsilon_m$. The best Bob can do when guessing $x_i \oplus x_j$ given message m, is to output the value of $x_i \oplus x_j$ that occurs most often among the $x \in X_m$. Define

$$\beta_{mij} = \mathop{\mathbb{E}}_{x \in X_m}[(-1)^{x_i \oplus x_j}].$$

Intuitively, if X_m (and hence p_m) is large, then most of these β_{mij} should be small. We now use the KKL inequality (Theorem 2.2.2) to make this intuition precise.

Claim 2.3.4. $\sum_{\{i,j\}} \beta_{mij}^2 \leq \begin{cases} 4\log_2(1/p_m)^2 & \textit{if } p_m \leq 1/2 \\ 2 & \textit{if } p_m > 1/2 \end{cases}.$

Proof. Let $f : \{0,1\}^n \to \{0,1\}$ be the characteristic function of the set X_m, so that $\Pr[f(x) \neq 0] = |X_m|/2^n = p_m$. Observe that β_{mij} is proportional to the Fourier coefficient $\widehat{f}(\{i,j\})$:

$$\beta_{mij} = \mathop{\mathbb{E}}_{x \in X_m}[(-1)^{x_i \oplus x_j}] = \frac{2^n}{|X_m|} \mathop{\mathbb{E}}_{x \in \{0,1\}^n}[f(x)(-1)^{x_i \oplus x_j}] = \frac{1}{p_m}\widehat{f}(\{i,j\}).$$

Using the KKL inequality with a $\delta \in [0,1]$ to be specified later, we get

$$\delta^2 \sum_{\{i,j\}} \widehat{f}(\{i,j\})^2 \leq \sum_{S \subseteq [n]} \delta^{|S|} \widehat{f}(S)^2 \leq \Pr[f(x) \neq 0]^{2/(1+\delta)} = p_m^{2/(1+\delta)},$$

and therefore,

$$\sum_{\{i,j\}} \beta_{mij}^2 = \frac{1}{p_m^2} \sum_{\{i,j\}} \widehat{f}(\{i,j\})^2 \leq \frac{1}{\delta^2}(1/p_m)^{2-2/(1+\delta)}.$$

For $p_m > 1/2$ simply choose $\delta = 1$. For $p_m \leq 1/2$, choose $\delta = 1/\log_2(1/p_m)$, which is in $[0,1]$, so that the above is

$$\log_2(1/p_m)^2(1/p_m)^{2\delta/(1+\delta)} \leq \log_2(1/p_m)^2(1/p_m)^{2\delta} = 4\log_2(1/p_m)^2. \qquad \square$$

The fraction of $x \in X_m$ where $x_i \oplus x_j = 0$ is $1/2 + \beta_{mij}/2$, hence Bob's optimal success probability when guessing $x_i \oplus x_j$ is $1/2 + |\beta_{mij}|/2$. This implies, for fixed m,

$$\mathop{\mathbb{E}}_{\{i,j\} \sim q_m}\left[\frac{1}{2} + \frac{|\beta_{mij}|}{2}\right] \geq \Pr[v = x_i \oplus x_j] = \frac{1}{2} + \varepsilon_m,$$

where "$\{i,j\} \sim q_m$" means that $\{i,j\}$ is distributed according to distribution q_m. This allows us to upper bound ε_m for m where $p_m \leq 1/2$:

$$\begin{aligned} 2\varepsilon_m &\leq \mathop{\mathbb{E}}_{\{i,j\}\sim q_m}[|\beta_{mij}|] \\ &= \sum_{\{i,j\}} q_m(\{i,j\})|\beta_{mij}| \\ &\leq \sqrt{\sum_{\{i,j\}} q_m(\{i,j\})^2} \cdot \sqrt{\sum_{\{i,j\}} \beta_{mij}^2} \\ &\leq \frac{2\log_2(1/p_m)}{\sqrt{n-1}}, \end{aligned}$$

where the second inequality is by Cauchy-Schwarz and the third uses both Eq. (2.5) and the first part of Claim 2.3.4. Since the p_m sum to 1, there can be at most one m for which $p_m > 1/2$. For that m we have $\varepsilon_m \leq 1/\sqrt{n-1}$ by an analogous argument combined with the second part of Claim 2.3.4.

Finally we can bound ε, treating the (at most one) m with $p_m > 1/2$ separately:

$$\begin{aligned} \varepsilon &= \sum_{m\in\{0,1\}^c} p_m\varepsilon_m \\ &\leq \sum_m p_m \frac{\log_2(1/p_m)}{\sqrt{n-1}} + \frac{1}{\sqrt{n-1}} \\ &= \frac{1}{\sqrt{n-1}}(H(p)+1) \\ &\leq \frac{c+1}{\sqrt{n-1}}, \end{aligned}$$

where $H(p) = \sum_m p_m \log_2(1/p_m)$ denotes the binary entropy function, which is at most c since the distribution $\{p_m\}$ is on 2^c elements. □

2.3.3 Classical protocol for HM_n

Here we design a classical protocol that achieves the above upper bound on the success probability. This protocol has no bearing on the large Bell inequality violations that are our main goal in this chapter, but it is nice to know the previous upper bound on the maximal success probability is essentially tight.

Theorem 2.3.5. *For every n that is a power of 2, and every positive integer $c \leq \sqrt{n}$, there exists a classical protocol for HM_n with c bits of one-way communication, such that for all inputs x, M,*

$$\Pr[v = x_i \oplus x_j] = \frac{1}{2} + \Omega\left(\frac{c}{\sqrt{n}}\right).$$

Proof. Assume for simplicity that $\sqrt{n}$ is integer, and that c is even and sufficiently large. Alice and Bob use shared randomness to choose two disjoint subsets S_1, S_2 of $[n]$ of size $\sqrt{n}$ each. Let y denote the bits of x located in the indices given by the first subset, and z the bits located in the indices given by the second subset. Alice and Bob use shared randomness to produce $2^{c/2}$ random $\sqrt{n}$-bit strings $y^{(1)}, \ldots, y^{(2^{c/2})}$. For each ℓ, the distance $d(y, y^{(\ell)})$ is distributed binomially, as the sum of $\sqrt{n}$ fair coin flips.

The following well-known fact about the tail of binomial distribution can be seen for instance by estimating $\binom{k}{k/2-\beta\sqrt{k}}$ using Stirling's approximation.

Fact: There exists a universal constant $\gamma > 0$ such that if X is the sum of k fair coin flips, then for all $0 < \beta < \sqrt{k}/2$ we have $\Pr[X \leq k/2 - \beta\sqrt{k}] \geq 2^{-\gamma(1+\beta^2)}$.

Thus we have $\Pr[d(y, y^{(\ell)}) \leq \sqrt{n}/2 - \beta n^{1/4}] \geq 2^{-\gamma(1+\beta^2)}$. Hence by choosing $\beta = \Theta(\sqrt{c})$, with probability close to 1, there will be an ℓ such that y and $y^{(\ell)}$ are at relative distance $\leq 1/2 - \Omega(c^{1/2}/n^{1/4})$. If so, Alice sends Bob the first such ℓ, and otherwise she tells him there is no such ℓ. This costs $c/2$ bits of communication. Similarly, at the expense of another $c/2$ bits of communication, Bob obtains an approximation of z with relative distance at most $\leq 1/2 - \Omega(c^{1/2}/n^{1/4})$.

One can see that with probability at least 1/2, Bob's matching M contains an $\{i, j\}$ with $i \in S_1$ and $j \in S_2$. Bob can predict x_i with success probability $1/2 + \Omega(c^{1/2}/n^{1/4})$ from his approximation of y, and can predict x_j with success probability $1/2 + \Omega(c^{1/2}/n^{1/4})$ from his approximation of z. These success probabilities are independent, hence he can predict $x_i \oplus x_j$ with success probability $1/2 + \Omega(c/\sqrt{n})$. If there is no such $\{i, j\} \in M$, or if he did not get good approximations to y or z, then Bob just outputs any $\{i, j\} \in M$ and a random bit for v, giving success probability 1/2. Putting everything together, we have a protocol that wins with probability $1/2 + \Omega(c/\sqrt{n})$. □

2.3.4 Entangled value for $\mathrm{HM}_n^{\mathbf{NL}}$

We now port our results to the non-local setting, referring to Definition 2.1.1 for the game associated with the Hidden Matching problem.

Theorem 2.3.6. *For every n that is a power of 2, there exists an entangled strategy for $\mathrm{HM}_n^{\mathrm{NL}}$ using a maximally entangled state with local dimension n, such that condition* (2.1) *is always satisfied.*

Proof. The strategy is as follows. Alice and Bob share $|\psi\rangle = \frac{1}{\sqrt{n}} \sum_{i \in \{0,1\}^{\log n}} |i\rangle|i\rangle$.

1. Alice performs a phase-flip according to her input x. The state becomes

$$\frac{1}{\sqrt{n}} \sum_{i \in \{0,1\}^{\log n}} (-1)^{x_i} |i\rangle|i\rangle.$$

2. Bob performs a projective measurement with projectors $P_{ij} = |i\rangle\langle i| + |j\rangle\langle j|$, with $\{i,j\} \in M$. The state collapses to $\frac{1}{\sqrt{2}}[(-1)^{x_i}|i\rangle|i\rangle + (-1)^{x_j}|j\rangle|j\rangle]$ for some $\{i,j\} \in M$ known to Bob.

3. Both players apply Hadamard transforms $H^{\otimes \log n}$, and the state becomes

$$\frac{1}{\sqrt{2}n} \sum_{a,b \in \{0,1\}^{\log n}} \left((-1)^{x_i + a\cdot i + b\cdot i} + (-1)^{x_j + a\cdot j + b\cdot j}\right) |a\rangle|b\rangle.$$

Notice that in the latter state, any pair a, b with nonzero amplitude must satisfy that

$$(a \cdot (i \oplus j)) \oplus (b \cdot (i \oplus j)) = x_i \oplus x_j.$$

Hence, if the players measure the state, Alice outputs a, and Bob outputs $\{i,j\}$ and the bit $d = b \cdot (i \oplus j)$, then they win the game with certainty. □

2.3.5 Classical value for $\mathrm{HM}_n^{\mathbf{NL}}$

In contrast to the entangled value, the optimal classical winning probability is not much better than 1/2:

Theorem 2.3.7. *The winning probability of any classical strategy for* $\mathrm{HM}_n^{\mathrm{NL}}$ *differs from* $\frac{1}{2}$ *by at most* $O\left(\log(n)/\sqrt{n}\right)$.

Proof. A strategy that wins $\mathrm{HM}_n^{\mathrm{NL}}$ with success probability $1/2+\varepsilon$ can be turned into a protocol for HM_n with $\log n$ bits of communication and the same winning probability: the players play $\mathrm{HM}_n^{\mathrm{NL}}$, with Alice producing a and Bob producing $\{i,j\}, d$; Alice then sends a to Bob, who outputs $\{i,j\}, (a \cdot (i \oplus j)) \oplus d$. The latter bit equals $x_i \oplus x_j$ with probability $1/2 + \varepsilon$. This requires $c = \log n$ bits of communication, so Theorem 2.3.3 implies the upper bound $1/2 + O\left(\log(n)/\sqrt{n}\right)$ on the winning probability. The lower bound of $1/2 - O\left(\log(n)/\sqrt{n}\right)$ on the winning probability follows similarly. □

Next we show that our upper bound on the success probability of classical strategies for $\mathrm{HM}_n^{\mathrm{NL}}$ is nearly optimal: we can achieve advantage at least $\Omega(\sqrt{\log(n)/n})$. Later we also give an alternative strategy with a slightly weaker advantage of $\Omega(1/\sqrt{n})$. The correctness of that more elementary strategy can be proven from first principles, unlike the one presented here. Theorem 2.3.8 and the alternative strategy have no bearing on our separation, but are included here mostly for completeness.

Theorem 2.3.8. *For every n that is a power of 2, there exists a classical deterministic strategy for* $\mathrm{HM}_n^{\mathrm{NL}}$ *with winning probability* $\frac{1}{2} + \Omega\left(\sqrt{\log(n)/n}\right)$ *(under the uniform input distribution).*

Proof. The strategy is as follows. Bob finds the $j \in \{2, \ldots, n\}$ that is matched to 1 in M, and outputs $\{1, j\}$ and $d = 0$. Since the number $i = 1$ corresponds to the string $0^{\log n}$, the winning condition $(a \cdot (i \oplus j)) \oplus d = x_i \oplus x_j$ is now equivalent to $a \cdot j = x_1 \oplus x_j$. Alice, given her input $x \in \{0,1\}^n$, outputs the value $a \in \{0,1\}^{\log n}$ that maximizes the winning probability subject to j being uniformly distributed over $\{2, \ldots, n\}$. That is, she selects an a that maximizes the number $J_{ax} := |\{j \in \{2, \ldots, n\} : a \cdot j = x_1 \oplus x_j\}|$. The winning probability of this strategy, for fixed x and uniformly random M, is $\max_a J_{ax}/(n-1)$. In the remainder of this proof we show that

$$\mathop{\mathbb{E}}_x[\max_a J_{ax}] \geq n/2 + \Omega(\sqrt{n \log n}), \tag{2.6}$$

which implies the theorem.

We use the following result due to Talagrand [LT91, Proposition 4.13]. For a finite set $T \subseteq \mathbb{R}^n$, define

$$r(T) := \mathop{\mathbb{E}}_{z \in \{\pm 1\}^n} \left[\sup_{t \in T} \left| \sum_{i=1}^n z_i t_i \right| \right].$$

This is the expected maximal overlap between z and the elements of T, expectation taken over uniformly random $z \in \{\pm 1\}^n$. Let $N(T, d_2; \varepsilon)$ denote the minimal number of (open) balls of radius $\varepsilon > 0$ (in Euclidean distance d_2) needed to cover T.

Proposition 2.3.9 (Talagrand)**.** *There exists a constant $K > 0$ such that for any $\varepsilon > 0$ and $T \subseteq \mathbb{R}^n$, if $\max_{t \in T, i \in [n]} |t_i| \leq \varepsilon^2/(Kr(T))$, then*

$$\varepsilon \sqrt{\log N(T, d_2; \varepsilon)} \leq Kr(T).$$

We apply this proposition as follows. For $a \in \{0,1\}^{\log n}$ consider the Hadamard codeword $h(a) := ((-1)^{a \cdot j})_{j \in \{0,1\}^{\log n}}$, with bits represented as ± 1 instead of 0/1. Let $T = \{h(a) : a \in \{0,1\}^{\log n}\}$ be the set of n Hadamard codewords. Any two distinct elements of T differ in exactly $n/2$ positions, and hence are at Euclidean distance $\sqrt{2n}$. Therefore a ball of radius $\varepsilon := \sqrt{n}/2$ can contain at most one element of T, so we need exactly n balls of radius ε to cover T (i.e., $N(T, d_2; \varepsilon) = n$). Proposition 2.3.9 now implies

$$r(T) = \Omega(\sqrt{n \log n}).$$

The definition of $r(T)$ takes the absolute value of $\sum_{i=1}^n z_i t_i$, but by symmetry, with probability 1/2 this quantity is positive for the maximizing t, and moreover the sum can never be smaller than $-\sqrt{n}$ since T forms an orthogonal basis. Hence we also have

$$\mathop{\mathbb{E}}_{z \in \{\pm 1\}^n} \left[\sup_{t \in T} \sum_{i=1}^n z_i t_i \right] = \Omega(\sqrt{n \log n}).$$

This means that we expect z to be relatively close in Hamming distance to some Hadamard codeword: the expected number of positions where z agrees with the closest $t \in T$ is $n/2 + \Omega(\sqrt{n \log n})$. In our application, since x itself is uniformly random, the string $y := ((-1)^{x_1 \oplus x_j})_{j \in \{0,1\}^{\log n}}$ is uniformly random except for its first bit, which is fixed to 1. Our quantity $\max_a J_{ax}$ measures the number of positions where y agrees with the closest $t \in T$, ignoring the first position. We conclude that

$$\mathop{\mathbb{E}}_{x \in \{0,1\}^n} \left[\max_{a \in \{0,1\}^{\log n}} J_{ax} \right] \geq n/2 + \Omega(\sqrt{n \log n}) - 1.$$

This implies Eq. (2.6). □

An alternative strategy for $\mathrm{HM}_n^{\mathbf{NL}}$

Here we give an alternative and slightly weaker version of Theorem 2.3.8, with advantage $\Omega(1/\sqrt{n})$ instead of $\Omega(\sqrt{\log(n)/n})$.

Proof. Fix arbitrary inputs x, M. Bob always outputs $i = 1$ and j is whatever is matched to i by M. Consider the following two unit vectors in $\mathbb{R}^n$,

$$u = \left((-1)^{x_1 \oplus x_k}/\sqrt{n}\right)_{k \in [n]} \qquad\qquad v = e_j$$

where e_j is the vector with 1 in the jth coordinate and zero elsewhere. Notice that Alice knows u, Bob knows v, and that $\langle u, v \rangle = (-1)^{x_1 \oplus x_j}/\sqrt{n}$. The players use shared randomness to choose a random unit vector $w \in \mathbb{R}^n$. Bob outputs $d = 0$ if $\langle w, v \rangle > 0$, and $d = 1$ otherwise. Alice outputs $a = 0^{\log n}$ if $\langle w, u \rangle > 0$, and a uniform $a \in \{0,1\}^{\log n}$ otherwise.

We now analyze the success probability. Assume that $x_1 \oplus x_j = 0$ (the other case being similar). One can see that the probability of both $\langle w, u \rangle$ and $\langle w, v \rangle$ being positive is $\frac{1}{2} - \frac{1}{2\pi} \arccos\langle u, v \rangle$, as this is essentially a two-dimensional question. They have the same probability of both being negative, and probability $\frac{1}{2\pi} \arccos\langle u, v \rangle$ to be in each of the two remaining cases. In the two cases that $\langle w, u \rangle \leq 0$ (an event that happens with probability 1/2), $a \cdot (i \oplus j)$ is a uniform bit (since $i \neq j$) and the players win with probability exactly 1/2. Otherwise (*i.e.*, if $\langle w, u \rangle > 0$), the players win if and only if $d = 0$ (*i.e.*, if also $\langle w, v \rangle > 0$). Hence, using that $\arccos(z) = \pi/2 - \Theta(z)$ for small z, the overall winning probability is

$$\frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} - \frac{1}{2\pi} \arccos\langle u, v \rangle = \frac{1}{2} + \Theta\left(\frac{1}{\sqrt{n}}\right). \qquad \square$$

2.4 The Khot-Vishnoi game

2.4.1 The classical value

In this section we analyze the classical value of the Khot-Vishnoi game (Definition 2.1.2). Our main result is an upper bound on the classical value of $1/n^{\eta/(1-\eta)}$, based on the analysis from [KV05].

Before we give that upper bound, let us first argue that it is essentially tight, *i.e.*, there exists a strategy whose winning probability is approximately $1/n^{\eta/(1-\eta)}$. To get some intuition for this game, first think of η as some small constant (even though we will eventually choose it close to 1/2), and consider the following natural classical strategy:

> Alice and Bob each output the element of their coset that has highest Hamming weight.

The idea is that if a is the element of highest Hamming weight in Alice's coset x, we expect $a \oplus z$ to also be of high Hamming weight (because it is close to a in Hamming distance), and since $a \oplus z$ is in his coset y, Bob is somewhat likely to pick it as his output. We now give a brief back-of-the-envelope calculation suggesting that the winning probability of this strategy is approximately $1/n^{\eta/(1-\eta)}$; since it is not required for our main result, we will not attempt to make this argument rigorous.

Let $t \geq 0$ be such that the probability that a binomial $B(n, 1/2)$ variable is greater than $(n+t)/2$ is $1/n$. Recalling that the cumulative distribution function of the binomial distribution $B(n,p)$ can be approximated by that of the normal distribution $N(np, np(1-p))$, and that the probability that a normal variable is greater than its mean by s standard deviations is approximately $e^{-s^2/2}$, we can essentially choose t to be the solution to $e^{-t^2/(2n)} = 1/n$ (so $t = \sqrt{2n \ln n}$). Then we expect Alice's n-element coset to contain exactly one element of Hamming weight greater than $(n+t)/2$. Since the element a that Alice picks is the one of highest Hamming weight, we assume for simplicity that its Hamming weight is $(n+t)/2$. The players win the game if and only if $a \oplus z$ has the highest weight among Bob's n elements, which we heuristically approximate by the event that $a \oplus z$ has Hamming weight at least $(n+t)/2$. The Hamming weight of $a \oplus z$ is distributed as the sum of two independent binomial distributions $B((n+t)/2, 1-\eta)$ and $B((n-t)/2, \eta)$, which can be approximated as above by the normal distribution $N((n+t)/2 - \eta t, n\eta(1-\eta))$. Hence, for the Hamming weight of $a \oplus z$ to be at least $(n+t)/2$, the normal variable needs to be greater than its mean by $\eta t / \sqrt{n\eta(1-\eta)}$ standard deviations, and the probability of this happening is approximately $e^{-\eta^2 t^2/(2n\eta(1-\eta))} = 1/n^{\eta/(1-\eta)}$, as claimed.

Now we show that no classical strategy can be substantially better. The main technical tool used in the proof is the hypercontractive inequality (Theorem 2.2.1), which is applicable to our setting because we choose u uniformly and $u \oplus z$ may be viewed as a "noisy version" of u.

Theorem 2.4.1. *For every n that is a power of 2, and every $\eta \in [0, 1/2]$, every classical strategy for the Khot-Vishnoi game KV_n (Definition 2.1.2) has winning probability at most $1/n^{\eta/(1-\eta)}$.*

Proof. Recall that the inputs are generated as follows: we choose a uniformly random $u \in \{0,1\}^n$ and an η-biased $z \in \{0,1\}^n$, and define the respective inputs to be the cosets $u \oplus H$ and $u \oplus z \oplus H$. We can assume without loss of generality that Alice and Bob's behavior is deterministic. Define functions $A, B : \{0,1\}^n \to \{0,1\}$ by $A(u) = 1$ if and only if Alice's output on $u \oplus H$ is u, and similarly for Bob. Notice that by definition, these functions attain the value 1 on exactly one element of each coset, and hence $\mathbb{E}_u[A(u)] = \mathbb{E}_u[B(u)] = 1/n$.

Recall that the players win if and only if the sum of Alice's output and Bob's output equals z. Hence for all u, z, $\sum_{h \in H} A(u \oplus h)B(u \oplus z \oplus h)$ equals 1 if the players win on input pair $u \oplus H, u \oplus z \oplus H$, and equals 0 otherwise. Therefore, the winning probability is given by

$$\begin{aligned}\mathop{\mathbb{E}}_{u,z}\Big[\sum_{h \in H} A(u \oplus h)B(u \oplus z \oplus h)\Big] &= \sum_{h \in H} \mathop{\mathbb{E}}_{u,z}[A(u \oplus h)B(u \oplus z \oplus h)] \\ &= n \mathop{\mathbb{E}}_{u,z}[A(u)B(u \oplus z)],\end{aligned}$$

where the second equality uses the fact that for all h, $u \oplus h$ is uniformly distributed.

We use the Fourier analysis framework introduced in Section 2.2.1. We have

$$\begin{aligned}\mathop{\mathbb{E}}_{u,z}[A(u)B(u \oplus z)] &= \mathop{\mathbb{E}}_{u}[A(u)(T_{1-2\eta}B)(u)] \\ &= \langle A, T_{1-2\eta}B \rangle \\ &= \langle T_{\sqrt{1-2\eta}}A, T_{\sqrt{1-2\eta}}B \rangle \\ &\le \big\|T_{\sqrt{1-2\eta}}A\big\|_2 \cdot \big\|T_{\sqrt{1-2\eta}}B\big\|_2 \\ &\le \|A\|_{2-2\eta} \cdot \|B\|_{2-2\eta} \\ &= \Big(\mathop{\mathbb{E}}_{u}[A(u)]\Big)^{1/(2-2\eta)} \cdot \Big(\mathop{\mathbb{E}}_{u}[B(u)]\Big)^{1/(2-2\eta)} \\ &= \frac{1}{n^{1/(1-\eta)}}.\end{aligned}$$

Here the third equality follows from Eqs. (2.2) and (2.3), the first inequality is by Cauchy-Schwarz, the second is the hypercontractive inequality (Theorem 2.2.1 with $q = 2$, $p = 2 - 2\eta$ and $\rho = \sqrt{1 - 2\eta}$) applied to each of A and B, and the second to last equality uses the fact that A, B have range $\{0,1\}$. We complete the proof by noting that $n/n^{1/(1-\eta)} = 1/n^{\eta/(1-\eta)}$. □

2.4.2 Lower bound on the entangled value

In this section we describe a good entangled strategy for the Khot-Vishnoi game, following the ideas of Kempe, Regev, and Toner [KRT10] and the SDP-solution of [KV05].

Theorem 2.4.2. *For every n that is a power of 2, and every $\eta \in [0, 1/2]$, there exists an entangled strategy that wins KV_n with probability at least $(1-2\eta)^2$, using a maximally entangled state of local dimension n.*

Proof. For $a \in \{0,1\}^n$, let $v^a \in \mathbb{R}^n$ denote the unit vector $((-1)^{a_i}/\sqrt{n})_{i\in[n]}$. Notice that for all a, b, we have $\langle v^a, v^b\rangle = 1 - 2d(a,b)/n$, where $d(a,b)$ denotes the Hamming distance between a and b. In particular, the n vectors v^a, as a ranges over a coset of H, form an orthonormal basis of $\mathbb{R}^n$.

The entangled strategy is as follows. Alice and Bob start with the n-dimensional maximally entangled state. Alice, given coset $x = u \oplus H$ as input, performs a projective measurement in the orthonormal basis given by $\{v^a \mid a \in x\}$ and outputs the value a given by the measurement. Bob proceeds similarly with the basis $\{v^b \mid b \in y\}$ induced by his coset $y = x \oplus z \oplus H$. A standard calculation now shows that the probability to obtain the pair of outputs a, b is $\langle v^a, v^b\rangle^2/n$. Since the players win if and only if $b = a \oplus z$, the winning probability on inputs x, y is given by

$$\frac{1}{n}\sum_{a\in x}\langle v^a, v^{a\oplus z}\rangle^2 = \frac{1}{n}\sum_{a\in x}(1 - 2d(a, a\oplus z)/n)^2 = (1 - 2|z|/n)^2,$$

where $|z|$ denotes the Hamming weight of the η-biased string z. Taking expectation and using convexity, the overall winning probability is

$$\mathop{\mathbb{E}}_{z}[(1-2|z|/n)^2] \geq \left(\mathop{\mathbb{E}}_{z}[1-2|z|/n]\right)^2 = (1-2\eta)^2. \qquad \square$$

2.5 Concluding remarks and open problems

Although Bell violations provide an elegant way to quantify the non-locality exhibited in a game, in experimental realizations of such games it is often important to take into account the actual classical and entangled values, and not just their ratio, especially when one tries to take into account possible imperfections in the experimental set-up. The large Bell violation and the tiny success probability achievable by classical players seem to make the KV game attractive to an experimental realization. One should keep in mind, though, that the success probability achievable by entangled players is $1/(\log n)^2$, which is somewhat low in absolute terms, and might not be visible if the experiment has too many false positives. It also means that an experiment must be repeated about $(\log n)^2$ times before we expect to see the first win. In the HM game, on the other hand, entangled players can win with certainty, which seems beneficial in case there are few false negatives. Another advantage of the HM game over KV is its somewhat simpler description.

One natural open question is to improve the Bell violation of $n/(\log n)^2$ achieved by the KV game either by tweaking the game or defining another game, possibly even matching the $O(n)$ upper bound up to a constant factor.[7] Throughout this chapter we considered the Bell violation as a function of the number of outputs of the players and/or of the dimension of entanglement. One can also analyze the violation in terms of the number of possible *inputs*. We recall that in the KV game both players have inputs taken from an exponentially large set, and that in the HM game (when modified as in Footnote 2) Bob has only $n/2$ possible inputs, but Alice still has an exponentially large set of inputs. The Bell inequality violation of $\sqrt{n}/\log n$ presented by Junge and Palazuelos [JP11] has the advantage that the number of inputs is only $O(n)$. Accordingly, another open question presents itself: can we find a game with a (near-)linear Bell inequality violation, and linear number of inputs and outputs for both Alice and Bob?

Finally, while this chapter focuses on the two-party setting, obtaining stronger Bell inequality violations for settings with three or more parties is also a worthwhile goal. Pérez-García *et al.* [PWP$^+$08] (see also [BV11]) gave a randomized construction of a three-party *XOR game* (in such a game each party outputs a bit, and winning or losing depends only on the XOR of those three bits) that gives a Bell inequality violation of $\Omega(\sqrt{d})$ using an entangled state in dimensions $d \times D \times D$ with $D \gg d$.[8] In contrast, it is a known consequence of Grothendieck's inequality that such non-constant separations do not exist for *two*-party XOR games. We do not know how large Bell inequality violations can be for arbitrary three-party games. Note that it is possible to make a three-party version of the Hidden Matching game: Alice gets input $x \in \{0,1\}^n$, Bob gets input $y \in \{0,1\}^n$, and Charlie gets a matching M as input, all uniformly distributed. The goal is that Alice outputs $a \in \{0,1\}^{\log n}$, Bob outputs $b \in \{0,1\}^{\log n}$, Charlie outputs $d \in \{0,1\}$ and $\{i,j\} \in M$, such that $((a \oplus b) \cdot (i \oplus j)) \oplus d = x_i \oplus x_j \oplus y_i \oplus y_j$. By modifying the two-party proofs in this chapter, one can show that the winning probability using an n-dimensional GHZ state is 1, while the best classical winning probability deviates from $1/2$ by at most $O((\log n)^2/n)$. So going from two to three parties squares the Bell inequality violation for Hidden Matching. This improvement unfortunately does not scale up with more than three parties, as one can show the classical winning probability is always at least $1/2 + \Omega(1/n)$.

[7] Interestingly, very recently Palazuelos [Pal12a] showed that this cannot be done using the maximally entangled state (which is the state used in all our entangled strategies): he proved that the ratio between the optimal value of entangled strategies using the maximally entangled state of local dimension n, and the classical value, is at most $O(n/\sqrt{\log n})$.

[8] They also showed that using GHZ states, there is no superconstant Bell inequality violation for XOR games (see also [BBLV09].)

Chapter acknowledgments. I thank the coauthors of the work on which this chapter is based. Together, we thank Jop Briët, Dejan Dukaric, Carlos Palazuelos, and Thomas Vidick for useful discussions, and Carlos also for sending us his manuscripts [Pal12b, Pal12a]. We thank the anonymous ToC referees and John Watrous for comments and suggestions, which improved the presentation.

Chapter 3

Quantum Graph Parameters

This chapter is mainly based on three papers. (Section 3.5 is based on unpublished work.) The first paper is *"Kochen-Specker Sets and the Rank-1 Quantum Chromatic Number"*, by the author and S. Severini. The paper was presented as a poster at the Quantum Information Processing conference in December 2011, and published in the IEEE Transactions on Information Theory in April 2012.

The second paper is *"New Separations in Zero-error Channel Capacity through Projective Kochen-Specker Sets and Quantum Coloring"*, by L. Mančinska, the author and S. Severini. The paper was presented at the Asian Quantum Information Science conference and as an invited talk at the Workshop on Quantum Physics of Information in August 2012. It was then presented as a poster at the Quantum Information Processing conference in January 2013. The paper was published in the IEEE Transactions on Information Theory in June 2013.

The third paper is *"Exclusivity structures and graph representatives of local complementation orbits"*, by A. Cabello, M. G. Parker, the author and S. Severini. The paper was published in the Journal of Mathematical Physics in July 2013.

3.1 Introduction

The chromatic number and the independence number are important and well-studied graph parameters.

The notion of *quantum chromatic number* was described in its generality by Cameron *et al.* [CMN$^+$07] in 2007, but it has been studied in the context of quantum non-locality since the late '90s (see the seminal work by Brassard, Cleve, and Tapp [BCT99]; see also the recent survey by Galliard, Wolf and Tapp [GTW10] and the references therein). It also appears implicitly in works on communication complexity (see, for example, [BCW98] and [dW01, pages 148–150]).

The notion of *quantum independence number* was implicitly present in many works about zero-error information theory, because of its link with channel capacities (see, for example, [CLMW10, LMM$^+$12, BCGSM]). We worked on the quantum independence number in the same framework in [MSS13]. In the meanwhile, a new definition related to graph homomorphisms appeared in [RM12]. We will use such definition because it allows us to better express the quantum independence number as a graph parameter.

The value of these notions is at least twofold: first, they have a natural use as tools for isolating the difference between quantum and classical behavior, second, they are a new approach for studying many combinatorial parameters between well-known NP-hard quantities like the clique and the chromatic number (*e.g.*, the Lovász ϑ-function, the orthogonal rank, *etc.*).

The rest of the chapter is structured as follows. In Section 3.2 we give some definitions that will be useful later.

In Section 3.3 we focus on the quantum chromatic number. We give a formal definition in terms of a non-local game. We reprove a result from [CMN$^+$07] concerning the form of the strategies for the above non-local game. Then, we prove some of the basic properties of the quantum chromatic number, including the relation with other common graph-theoretic parameters.

In Section 3.3.1, we focus on the rank-1 quantum chromatic number. This quantity is obtained by using only rank-1 measurement operators in the quantum strategies for the above-mentioned non-local game. We prove that the rank-1 quantum chromatic number is equal to orthogonal rank of a particular Cartesian product graph. Then, we exhibit graphs where the rank-1 quantum chromatic number is strictly greater than the orthogonal rank thereby solving an open problem stated in [CMN$^+$07]. The proof technique is not based on a specific example, but on a general result which connects rank-1 quantum chromatic number and Kochen-Specker sets. These are collections of vectors originally used to prove the inadequacy of local hidden variable theories to model quantum mechanical behavior deterministically [KS67, PMMM05].

In Section 3.3.2, we use our newly-defined notion of *projective Kochen-Specker set*, to show that there is a separation between quantum and classical chromatic number. The characterization settles the graph-theoretic discussion started in [CMN$^+$07]. A characterization for the rank-1 case was already given in [SS12] but it is subsumed by this result. Interestingly, [FIG11] observed a separation between rank-1 and general rank quantum chromatic number. From their result it follows that the use of projective KS sets is necessary. This full characterization is valuable because until now the only examples of the separation were some orthogonality graphs, specifically the Hadamard graphs considered by Avis *et al.* [AHKS06] and introduced in [FR87], and an isolated example with 18 vertices [CMN$^+$07].

In Section 3.4 we focus on the quantum independence number. We give the definition as a graph parameter related to non-local games. We will see in Chapter 4 how this definition relates to zero-error channel capacity. Our main contribution about the quantum independence number is to exhibit three different constructions of graphs with a separation between quantum and classical independence number. In Section 3.4.1, we use projective KS sets to find graphs with such a separation. This generalizes a similar construction for KS sets given in [CLMW10]. Then, in Section 3.4.2, we show how to use graphs for which the chromatic and quantum chromatic number are different to construct graphs with a separation between quantum and classical independence number. The third construction, in Section 3.4.3, is a construction based on graph states, used in [CPSS12] to study orbits of graphs under local complementation.

Finally we dedicate Section 3.5 to some unpublished notes about graphs representing non-local games. These graphs are defined for a particular case in [CSW10]. It is known that the independence number of such graphs corresponds to the classical value of the game. Here we give a precise definition of game graphs and show that quantum independence number and the Lovász theta number are bounds on the quantum value of the game.

3.2 Preliminaries

3.2.1 Notions of graph theory

A *simple graph* $G = (V, E)$ consists of a finite vertex set V and its edge set $E \subsetneq V \times V$ (the inclusion here is strict because there are no edges of the form (v, v)). Two vertices $(v, w) \in E$ are "adjacent" or equivalently "form an edge".[1] All graphs considered in this chapter, unless otherwise specified, are simple graphs. For a graph $G = (V, E)$, we also denote its vertex set with $V(G)$ and its edge set with $E(G)$ whenever confusion has to be avoided.

A *proper c-coloring* of a graph is an assignment of c colors to the vertices of the graph such that every two adjacent vertices have different colors. The *chromatic number* of a graph G, denoted by $\chi(G)$, is the minimum number of colors c such that there exists a proper c-coloring of G.

An *independent set* of a graph is a subset I of $V(G)$ such that no two elements of I are adjacent. The *independence number* of a graph G, denoted by $\alpha(G)$, is the maximum size of an independent set of G. A little thought shows that $\alpha(G) \cdot \chi(G) \geq |V(G)|$.

The complement of G is $\overline{G}$, the graph with vertex set $V(G)$ where distinct vertices are adjacent if and only if they are not adjacent in G. A clique is a subset of vertices in which each pair is adjacent and the *clique number* $\omega(G)$ is the maximum cardinality of a clique in G. Clearly $\alpha(G) = \omega(\overline{G})$.

A *homomorphism* from a graph G to a graph H is a map $\phi : V(G) \to V(H)$ such that every edge $\{u, v\} \in E(G)$ in G is mapped to an edge $\{\phi(u), \phi(v)\} \in E(H)$ in H. If such a map exists, we write $G \longrightarrow H$.

[1] Since simple graphs are non-directed, it is equivalent to address the edges as pairs (u, v) or sets $\{u, v\}$. We will use both notations depending on the context.

A d-dimensional orthogonal representation of $G = (V, E)$ is a map $\phi : V \to \mathbb{C}^d$ such that for all $(v, w) \in E$, $\langle\phi(v)|\phi(w)\rangle = 0$. (If all the vectors have unit norm, this is called orthonormal representation.) The *orthogonal rank* of a graph G, denoted by $\xi(G)$, is defined as the minimum d such that there exists an orthogonal representation of G in $\mathbb{C}^d$.

We relate every multiset of projectors T to a graph. The *orthogonality graph* of T is the graph with vertex set T and edge set $\{(P, P') : \mathrm{Tr}(PP') = 0\}$. Let us denote by $\uplus$ the multiset union. Some authors construct orthogonality graphs from a set of vectors. This is just a special case of the above definition: associate to each vector v the rank-1 projector $|v\rangle\langle v|$.

For all pairs of graphs G and H, define their *Cartesian product* $G\Box H$ as follows. The vertex set $V(G\Box H) = V(G) \times V(H)$ is the Cartesian product of the vertex sets of G and H. We can therefore identify each vertex in $V(G\Box H)$ with a pair of vertices, one from each of the two original graphs. There is an edge in $E(G\Box H)$ between vertices (v, i) and (w, j) if either $v = w$ and $(i, j) \in E(H)$ or $(v, w) \in E(G)$ and $i = j$.

Lemma 3.2.1. *[Viz63] For all graphs G, H, the independence number of their Cartesian product satisfies*

$$\alpha(G\Box H) \leq \min\{\alpha(G) \cdot |V(H)|, \alpha(H) \cdot |V(G)|\}.$$

Proof. Assume w.l.o.g. that $\alpha(G) \cdot |V(H)| \leq \alpha(H) \cdot |V(G)|$. Suppose, towards a contradiction, that $\alpha(G\Box H) > \alpha(G) \cdot |V(H)|$. Then, there is a $i \in V(H)$ such that there are more than $\alpha(G)$ non-adjacent vertices of $G\Box H$ of the form $(*, i)$. But this implies the existence of an independent set of G of size larger than $\alpha(G)$, because there is an edge between (v, i) and (w, i) whenever $(v, w) \in E(G)$. □

For all pairs of graphs G and H, the *strong product* $G \boxtimes H$ is the graph whose vertex set is the cartesian product $V(G) \times V(H)$ and where two distinct vertices (v, i) and (w, j) are adjacent if and only if $(v, w) \in E(G)$ or $v = w$ and $(i, j) \in E(H)$ or $i = j$.

We finally introduce an important graph parameter:the *theta number* (a.k.a. Lovász number or theta function). It was originally defined by Lovász [Lov79] to solve a long-standing problem posed by Shannon [Sha56]: computing the Shannon capacity of the five-cycle. There are many equivalent formulations of the theta number (see [KD93] for a detailed survey). In this thesis we use two of them. The one that we use in this chapter is the following:

$$\vartheta(G) = \max \sum_{v\in V(G)} |\langle\psi|\psi_v\rangle|^2, \tag{3.1}$$

where the max is over unit vectors ψ and orthonormal representations $\{\psi_v\}_{v\in V(G)}$.

The second one is an SDP formulation that we use in Chapter 4.

$$\begin{aligned}\vartheta(G) = \min\Big\{ \lambda : \ \ & \exists\, Z \in \mathbb{R}^{V(G)\times V(G)},\ Z \succeq 0, \\ & Z(u,u) = \lambda - 1 \ \text{ for } u \in V(G), \\ & Z(u,v) = -1 \ \text{ for } \{u,v\} \notin E(G) \Big\}. \end{aligned} \tag{3.2}$$

Lovász [Lov79] proved that $\alpha(G) \leq \vartheta(G) \leq \chi(\overline{G})$ holds (this inequality is often referred to as the sandwich theorem [KD93]). The theta number can be approximated to within arbitrary precision in polynomial time, hence it gives a tractable and in many cases useful bound for both α and χ. Lovász proved that ϑ is *multiplicative* under the strong graph product, that is,

$$\vartheta(G \boxtimes H) = \vartheta(G)\vartheta(H). \tag{3.3}$$

3.2.2 Kochen-Specker sets

Consider a set of n-dimensional (complex) vectors $S \subseteq \mathbb{C}^n$.

Definition 3.2.2. *A function $f : S \to \{0,1\}$ is a* marking function *for S if for all orthonormal bases $b \subseteq S$ we have $\sum_{u\in b} f(u) = 1$.*

Gleason's theorem [Gle57] implies that for any $n \geq 3$ there does not exist a marking function for $\mathbb{C}^n$. Bell [Bel66] and independently Kochen and Specker [KS67] interpreted this statement in the framework of contextuality of physical theories. For this reason, this statement is also known as the (Bell-)Kochen-Specker theorem. Since then, finite sets of vectors in some given dimension giving rise to a proof of this theorem are known as Kochen-Specker (KS) sets. Note that although in principle there are KS sets of infinite size, in this thesis we are only interested in finite sets, since we will use them as a tool to work on finite graphs. In general, much importance is given to finding the smallest possible KS set (see [AOW11]).

Definition 3.2.3 (KS set)**.** *A set of unit vectors $S \subseteq \mathbb{C}^n$ is a* Kochen-Specker set *if there is no marking function for S.*

Renner and Wolf [RW04] considered a generalization of KS sets called *weak* KS sets. Intuitively, for a weak KS set there can be marking functions, but every such function evaluates to 1 for two orthogonal vectors in the set.

Definition 3.2.4 (weak KS set)**.** *A set of unit vectors $S \subseteq \mathbb{C}^n$ is a* weak Kochen-Specker set *if for all marking functions f for S there exist orthogonal vectors $u, v \in S$ such that $f(u) = f(v) = 1$.*

As explained in [RW04], every KS set is clearly a weak KS set but the converse does not always hold. However, every weak KS set can be completed to a KS set by adding $O(|S|^2)$ elements. Hence, a weak KS set also gives a proof of the Kochen-Specker theorem in some specific dimension. In fact it is more convenient to deal with weak KS sets, since they capture the essence of KS sets and can contain fewer vectors.

Generalizations of KS sets We introduce a natural generalization of weak KS sets by considering subsets of $\mathcal{Q}_n$, the set of all $n \times n$ projectors, instead of subsets of $\mathbb{C}^n$. Recall that an orthogonal projector is a Hermitian matrix P such that $P^2 = P$. For brevity, from now on we omit the word "orthogonal" when we talk about projectors.

Let $\mathcal{M}_n$ be the set of $n \times n$ matrices.

Definition 3.2.5. *A* marking function f for $S \subsetneq \mathcal{M}_n$ *is a function* $f : S \to \{0, 1\}$ *such that for all* $M \subseteq S$ *with* $\sum_{P \in M} P = I$*, we have* $\sum_{P \in M} f(P) = 1$*.*

Definition 3.2.6 (Projective KS set)**.** *A set* $S \subsetneq \mathcal{Q}_n$ *is a* projective Kochen-Specker set *if for all* marking functions f *for* S*, there exist* $P, P' \in S$ *for which* $\mathrm{Tr}(PP') = 0$ *and* $f(P) = f(P') = 1$*.*

Note that each set $M \subseteq \mathcal{Q}_n$ with $\sum_{P \in M} P = I$ is a projective measurement. Also, note that weak KS sets are a special case of projective KS sets, if we identify a vector with the corresponding rank-1 projector. Conversely, starting from any projective KS set one can construct (usually infinitely many) underlying weak KS sets. For example, one can take for each projector an arbitrary orthonormal basis of its span. It can be verified that union of all the bases is a weak KS set (see [MSS13, Appendix A]).

Although in the rest of the thesis we will only deal with projective KS sets, we can further generalize weak KS sets by considering subsets of $\mathcal{S}_+^n$, the set of all $n \times n$ positive semidefinite matrices.

Definition 3.2.7 (Generalized KS set)**.** *A set* $S \subsetneq \mathcal{S}_+^n$ *is a* generalized Kochen-Specker set *if for all* marking functions f *for* S *there exist* $E, E' \in S$ *for which* $E + E' \leq I$ *and* $f(E) = f(E') = 1$*.*

Note that each set $M \subseteq \mathcal{S}_+^n$ with $\sum_{E \in M} E = I$ is a POVM (where, as usual, "POVM" stands for positive operator-valued measure). Projective KS sets are a special case of generalized KS sets, because when S is a set of projectors the condition $E + E' \leq I$ is equivalent to $\mathrm{Tr}(EE') = 0$.

KS-like sets consisting of positive semidefinite matrices have already been considered by Cabello [Cab03]. Motivated by a recent analogue of Gleason's theorem for positive semidefinite operators in two dimensions [Bus03, CFMR04], Cabello exhibits what we here call a generalized KS set in $\mathcal{S}_+^2$. Hence, generalized KS sets exist even in two dimensions and have turned out to be useful for scenarios where regular KS sets do not apply (recall that there are no KS sets in $\mathbb{C}^2$).

KS sets and pseudo-telepathy games Here we generalize the results of [RW04] concerning the relationship between *weak* KS sets and a class of pseudo-telepathy games. We show that there is a relationship between *projective* KS sets and a class of pseudo-telepathy games that is larger than the one above.

Informally, a non-local game is called a *pseudo-telepathy game* if players sharing the entangled state win with certainty, while classical players have non-zero probability to lose. More formally:

Definition 3.2.8 (Pseudo-telepathy game)**.** *A non-local game with input sets X, Y, output sets A, B, input distribution π and verification function V is called a* pseudo-telepathy game *if:*

1. *There exists a quantum strategy consisting of a shared bipartite entangled state $|\psi\rangle$ and POVMs $\{E_a^x\}_{a\in A}$ for every $x \in X$ for Alice, and $\{F_b^y\}_{b\in B}$ for every $y \in Y$ for Bob, with the following property. For any (a, b, x, y) with $\pi(x, y) > 0$ it holds that $\langle\psi|E_a^x \otimes F_b^y|\psi\rangle \neq 0$ implies $V(a, b, x, y) = 1$.*

2. *For all deterministic classical strategies s_A, s_B, there exists a tuple (a, b, x, y) with $\pi(x, y) > 0$ such that $V(a, b, x, y) = 0$ but $s_A(x) = a$ and $s_B(y) = b$.*

The following theorem relates projective KS sets and a special kind of pseudo-telepathy games. We will only consider KS sets for which each projector is part of some projective measurement from S. Note that we can delete elements from any projective KS set until the resulting set is still a projective KS set and satisfies the above property. Therefore, projectors not contained in any measurement from S are inessential.

For all natural numbers n, let $|\psi_n\rangle = \frac{1}{\sqrt{n}}\sum_{i\in[n]}|i\rangle|i\rangle$, where $\{|i\rangle\}_{i\in[n]}$ is the standard basis of $\mathbb{C}^n$. Let $\overline{P}$ be the entry-wise complex conjugate of P. A useful property of $|\psi_n\rangle$ is the following. For all $A, B \in \mathbb{C}^{n\times n}$,

$$\langle\psi_n|A \otimes B|\psi_n\rangle = \sum_{ij}\langle ii|A \otimes B|jj\rangle = \sum_{ij}\langle i|A|j\rangle\langle i|B|j\rangle = \sum_{ij} A_{ij}B_{ij} = \mathrm{Tr}(A\overline{B}) \tag{3.4}$$

Theorem 3.2.9. *Let n be any natural number and let X, Y, A, B be sets. Consider projective measurements $\{P_a^x\}_{a\in A}$ for every $x \in X$, and $\{Q_b^y\}_{b\in B}$ for every $y \in Y$, acting on $\mathbb{C}^n$. The following statements are equivalent:*

1. *$S = \{P_a^x\}_{(x,a)} \cup \{\overline{Q_b^y}\}_{(y,b)}$ is a projective KS set.*

2. *There exists a pseudo-telepathy game with input sets X, Y and output set A, B for which there is a winning quantum strategy consisting of a shared state $|\psi_n\rangle$ and projective measurements $\{P_a^x\}_{a\in A}$ for every $x \in X$ for Alice, $\{Q_b^y\}_{b\in B}$ for every $y \in Y$ for Bob.*

Proof. ($1 \Rightarrow 2$) Assume S is a projective KS set. Let $\{S_1, \ldots, S_k\}$ be the set of all projective measurements contained in S (*i.e.* each S_i is a set of projectors that sum to identity). Consider a non-local game $\mathcal{G}$ where $X, Y = [k]$, $A, B = [\max_i |S_i|]$, π is the uniform distribution, and the verification function is defined by

$$\begin{aligned} V(a,b,x,y) = 1 &\Leftrightarrow \langle \psi_n | P_a^x \otimes \overline{P_b^y} | \psi_n \rangle \neq 0 \\ &\Leftrightarrow \mathrm{Tr}(P_a^x P_b^y) \neq 0. \end{aligned}$$

By definition, this game has an optimal quantum strategy in which Alice and Bob share $|\psi_n\rangle$ and they measure their part of the state using projective measurements S_x and $\overline{S_y}$ respectively upon receiving inputs (x, y).

Towards a contradiction, suppose there is an optimal classical strategy (s_A, s_B). Note that $s_A = s_B$ by definition of V, because if $x = y$ then the only winning outputs satisfy $a = b$. Let us now construct a marking function f for S from s_A:

$$\begin{aligned} \forall P \in S,\ f(P) = 1 \Leftrightarrow\ &\exists a \in A, x \in X \text{ such that} \\ &P = P_a^x \text{ and } s_A(x) = a. \end{aligned}$$

It is clear that f marks one projector per measurement and since it is based on a winning strategy, it never marks two orthogonal projectors. However, to show that f is indeed a function defined on S, we need to address a last issue. It can happen that the same projector appears in more than one measurement. We need to show that if f marks a projector, it marks the same projector in all the measurements that contain it. Suppose, towards a contradiction, that there exists x, x' such that $s_A(x) = a$, $P_a^x \in S'_x$, $s_A(x') = a'$ but $P_{a'}^{x'} \neq P_a^x$. Then, since $s_A = s_B$ and all elements of $S_{x'}$ different from P_a^x are orthogonal to it, the players would lose the game on input (x, x'). This contradicts the fact that s_A, s_B are winning strategies.

We have proved that f is a marking function for S constructed from s_A, s_B. This contradicts that S is a projective KS set, therefore classical players cannot have an optimal strategy and the desired statement follows.

$\Leftarrow$) Assume $\mathcal{G}$ is a pseudo-telepathy game and $S = \{P_a^x\}_{(x,a)} \cup \{\overline{Q_b^y}\}_{(y,b)}$ together with $|\psi_n\rangle$ is a winning quantum strategy. Every marking function f for S can be mapped to a classical strategy in the following way:

$$s_A(x) = a \Leftrightarrow f(P_a^x) = 1 \text{ and } s_B(y) = b \Leftrightarrow f(\overline{Q_b^y}) = 1.$$

Since $\mathcal{G}$ is a pseudo-telepathy game, for every f there exists a tuple (a, b, x, y) such that $s_A(x) = a$ and $s_B(y) = b$ (and therefore $f(P_a^x) = f(Q_b^y) = 1$), but $V(a, b, x, y) = 0$. Since the quantum players always win the game, we have that $\langle \Psi | P_a^x \otimes \overline{Q_b^y} | \Psi \rangle = 0$ and this implies $\mathrm{Tr}(P_a^x Q_a^y) = 0$ by (3.4). Therefore, for any marking function f for S we can find orthogonal projectors $P_a^x, Q_b^y \in S$ such that $f(P_a^x) = f(P_b^y) = 1$. Hence, S is projective KS set. □

3.3 Quantum Chromatic Number

In this section we define the quantum chromatic number of a graph. For the sake of completeness and to fix some useful facts, we present a comprehensive statement about its basic properties. This is done by extending and completing some observations contained in [CMN$^+$07].

Informally, the *c-coloring game* for a graph $G = (V, E)$ is as follows. Two players, Alice and Bob, claim that they have a proper c-coloring for G. A referee wants to test this claim with a one-round game, so he forbids communication between the players and separately asks Alice the color α for the vertex v and Bob the color β for the vertex w. The players are required to give the same color as output if $v = w$, and to give a different color if $(v, w) \in E$. A formal definition follows.

Definition 3.3.1. *The* c-coloring game *on the graph $G = (V, E)$ is a non-local game with input sets $X = Y = V$, output sets $A = B = [c]$ and uniform distribution on the inputs.*[2] *Alice gets input v and outputs α, Bob gets input w and outputs β. The players* lose the game *in the following two cases:*

1. *$v = w$ and $\alpha \neq \beta$*

2. *$(v, w) \in E$ and $\alpha = \beta$*

A classical strategy consists w.l.o.g. of two deterministic functions $c_A : V \to [c]$ for Alice and $c_B : V \to [c]$ for Bob. We can introduce shared randomness, but since this results in a probability distribution over deterministic strategies, it is not beneficial. A little thought will show that to win with probability 1, we must have $c_A = c_B$ (to avoid the first losing condition) and that c_A must be a valid c-coloring of the graph (to avoid the second losing condition). It follows that the classical players cannot win the c-coloring game with probability 1 when $c < \chi(G)$.

A quantum strategy for the c-coloring game uses an entangled state $|\psi\rangle$ of local dimension d and two families of POVMs: for all $v \in V$, Alice has $\{E^v_\alpha\}_{\alpha=1,\ldots,c}$ and Bob has $\{F^v_\beta\}_{\beta=1,\ldots,c}$. According to her input v, Alice applies the corresponding POVM $\{E^v_\alpha\}_{\alpha=1,\ldots,c}$ to her part of the entangled state and outputs the outcome α. Bob acts similarly and outputs β. The requirements for the game translate into the following *consistency conditions*. Alice and Bob win the coloring game with certainty, using a quantum strategy as described above, if and only if

$$\forall v \in V, \forall \alpha \neq \beta,\ \langle\psi|E^v_\alpha \otimes F^v_\beta|\psi\rangle = 0 \tag{3.5}$$

$$\forall (v, w) \in E, \forall \alpha,\ \langle\psi|E^v_\alpha \otimes F^w_\alpha|\psi\rangle = 0. \tag{3.6}$$

[2] Another variant of the game uses an uniform distribution over pairs of inputs that equal or form an edge (see for example [AHKS06]).

In this case, we call the strategy a *winning strategy* or a *quantum c-coloring of G*. Note that we do not bound the dimension of the entangled state or the rank of the measurement operators, we only care about the *number* of measurement operators needed to win the game with certainty (*i.e.*, the number of colors). We are now ready to give the central definition of this section.

Definition 3.3.2. *For all graphs G, the* quantum chromatic number $\chi_q(G)$ *is the minimum number c such that there exists a quantum c-coloring of G.*

We will see that without loss of generality a quantum c-coloring has a *normal form*, a clean and simple structure defined as follows.

Definition 3.3.3. *A quantum c-coloring of G is in* normal form *if there exists an integer r such that:*

1. *All POVMs are projective measurements with c real-valued projectors of rank r.*
2. *The shared state is $|\psi\rangle = \frac{1}{\sqrt{rc}} \sum_{i\in[rc]} |i\rangle|i\rangle$.*
3. *Alice's projectors are related to Bob's as follows: for all v, α, $E^v_\alpha = F^v_\alpha$.*
4. *The consistency conditions (3.5) and (3.6) can be expressed as the single condition:*
$$\forall (v,w) \in E, \forall \alpha \in [c],\ \mathrm{Tr}(E^v_\alpha E^w_\alpha) = 0. \tag{3.7}$$

Notice that we do not know whether the rank r of the projectors in the best normal form is equal to the minimum rank of a general strategy. It might be necessary to increase the rank to obtain the normal form.

The next proposition is a collection of statements from [CMN$^+$07], expanded and rearranged, that are useful to direct our discussion.

Proposition 3.3.4 ([CMN$^+$07, SS12])**.** *If G has a quantum c-coloring, then it has a quantum c-coloring in normal form.*

Proof. We start with a generic winning strategy consisting of entangled state $|\psi'''\rangle$, and POVMs $\{(E^v_\alpha)'''\}_{v\in V,\alpha\in[c]}$ for Alice and $\{(F^w_\beta)'''\}_{w\in V,\beta\in[c]}$ for Bob. Then, we will gradually construct an equivalent strategy with the desired properties. We will prove the statements in a few steps. Each bullet in the following list is a small statement that is proved right after. At the end of the steps, we will have the final strategy in normal form. The number of prime symbols of the notation for the entangled state and the POVM elements will reduce as soon as we get close to the final strategy, consisting of $|\psi\rangle$ and $\{E^v_\alpha\}_{v\in V,\alpha\in[c]}$, $\{F^w_\beta\}_{w\in V,\beta\in[c]}$.

- *The entangled state has full Schmidt rank.*

  Start with the entangled state $|\psi'''\rangle$, with local dimension d'. Consider the Schmidt decomposition $|\psi'''\rangle = \sum_{i=0}^{d'-1} \lambda_i|i\rangle|i\rangle$, where without loss of generality $\{|i\rangle\}_{i\in\{0,\dots,d'-1\}}$ is the computational basis. Say there are d non-zero λ_i. Then we define the new entangled state as $|\psi''\rangle = \sum_{i:\lambda_i\neq 0} \lambda_i|i\rangle|i\rangle$, and restrict the measurement operators to the respective supports of the reduced states as follows. Consider the projector $P = \sum_{i:\lambda_i\neq 0} |i\rangle\langle i|$. Then for all the POVM elements of Alice define $(E^v_\alpha)'' = P(E^v_\alpha)'''P$, and do the same for Bob's POVM elements. These restricted POVMs are valid measurements on $|\psi''\rangle$, and they still form a valid quantum coloring: $\sum_\alpha (E^v_\alpha)'' = I$ (on the d-dimensional subspace on which P projects) and

  $$\begin{aligned}\langle\psi''|(E^v_\alpha)'' \otimes (F^v_\beta)''|\psi''\rangle =&\langle\psi'''|P(E^v_\alpha)'''P \otimes P(F^v_\beta)'''P|\psi'''\rangle\\ =&\langle\psi'''|(E^v_\alpha)''' \otimes (F^v_\beta)'''|\psi'''\rangle.\end{aligned}$$

  We have that $|\psi''\rangle$ has full Schmidt rank d and together with the new POVMs is a winning strategy.

- *All POVM elements are projectors.*
  With some abuse of notation, we identify a projector with the support of the space on which it projects. We denote the support of an operator A by $\mathrm{supp}(A)$.

  It follows from consistency condition (3.5) that

  $$\forall v, \alpha, \sum_{\beta\neq\alpha} \langle (E^v_\alpha)'', \mathrm{Tr}_B(I \otimes (F^v_\beta)''|\psi''\rangle\langle\psi''|)\rangle = 0, \tag{3.8}$$

  Since both $(E^v_\alpha)''$ and $\mathrm{Tr}_B(I \otimes (F^v_\beta)''|\psi''\rangle\langle\psi''|$ are positive semidefinite operators, it follows from (3.8) that $\mathrm{supp}((E^v_\alpha)'')$ is a subspace of $\mathrm{supp}(\mathrm{Tr}_B(I \otimes (F^v_\beta)''|\psi''\rangle\langle\psi''|)$. By a symmetric argument, it also follows that $\mathrm{supp}(\mathrm{Tr}_B(I \otimes (F^v_\beta)''|\psi''\rangle\langle\psi''|)$ is a subspace of $\mathrm{supp}((E^v_\alpha)'')$.

  Therefore, without loss of generality for all v and α we can replace $(E^v_\alpha)''$ with $(E^v_\alpha)' = \mathrm{supp}(\mathrm{Tr}_B(I \otimes (F^v_\alpha)''|\psi''\rangle\langle\psi''|))$. A similar replacement can be done for Bob's POVM elements.

- *The state $|\psi''\rangle$ can be replaced by the maximally entangled state.*
  The winning strategies do not depend on the values of the Schmidt coefficients $\{\lambda_i\}$ of $|\psi''\rangle$, as long as they are non-zero. Thus we can set for all $i, \lambda_i = 1/\sqrt{d}$ and define $|\psi'\rangle = |\psi_d\rangle = \frac{1}{\sqrt{d}} \sum_{i\in[d]} |i\rangle|i\rangle$.

- *All projectors can be real-valued, of the same rank r and the maximally entangled state can have local dimension rc.*

We first map every complex-valued $d \times d$ projector into a real-valued $2d \times 2d$ projector using the following map.

$$R(A) = \begin{pmatrix} \mathcal{R}(A) & \mathcal{I}(A) \\ -\mathcal{I}(A) & \mathcal{R}(A) \end{pmatrix},$$

where $\mathcal{R}(A)$ and $\mathcal{I}(A)$ are respectively the real and imaginary part of A. One can verify that this map preserves matrix sum and matrix product, and that for all v we have $\sum_{\alpha \in [c]} R(E_\alpha^v) = I$.

Second, we extend the entangled state to $|\psi\rangle = |\psi_{2d}\rangle \otimes |\psi_c\rangle$ and then define new projectors for Alice $E_\alpha^v = \sum_{i=0}^{c-1} R((E^v_{\alpha+i (\operatorname{mod} c)})') \otimes |i\rangle\langle i|$ (and similarly for Bob). All have rank $r = \sum_\alpha \operatorname{rank}(R((E_\alpha^v)'))$ and act on the new state of local dimension rc. One can see that the new projectors still satisfy the consistency conditions.

- *We have for all v, α, $E_\alpha^v = F_\alpha^v$.*
  Call $\rho = \operatorname{Tr}_B(|\psi\rangle\langle\psi|) = \sum_i \frac{1}{\sqrt{rc}}|i\rangle\langle i| = \operatorname{Tr}_A(|\psi\rangle\langle\psi|)$. Then

$$\begin{aligned}
E_\alpha^v &= \operatorname{supp}(\operatorname{Tr}_B(I \otimes F_\alpha^v |\psi\rangle\langle\psi|)) \\
&= \operatorname{supp}\left(\operatorname{Tr}_B(I \otimes F_\alpha^v \sum_i \frac{1}{\sqrt{rc}}|i\rangle|i\rangle \sum_j \frac{1}{\sqrt{rc}}\langle j|\langle j|)\right) \\
&= \operatorname{supp}\left(\sum_{i,j} \frac{1}{rc}\operatorname{Tr}_B(|i\rangle\langle j| \otimes F_\alpha^v |i\rangle\langle j|)\right) \\
&= \operatorname{supp}\left(\sum_{i,j} \frac{1}{rc}|i\rangle\langle j|(F_\alpha^v)_{j,i}\right) \\
&= \operatorname{supp}\left(\sum_{i,j} \frac{1}{rc}|i\rangle\langle i|(F_\alpha^v)|j\rangle\langle j|\right) \\
&= \operatorname{supp}(\sqrt{\rho}(F_\alpha^v)\sqrt{\rho}).
\end{aligned}$$

Since all measurements are projective measurements, we have that $\sum_\alpha E_\alpha^v = I$ and that for $\alpha \neq \beta$

$$\begin{aligned}
F_\alpha^v F_\beta^v = 0 \Rightarrow & \sqrt{\rho} E_\alpha^v \rho E_\beta^v \sqrt{\rho} = 0 \\
\Rightarrow & \sum_{i,j} \lambda_i \lambda_j |i\rangle\langle i| \left(E_\alpha^v \rho E_\beta^v\right) |j\rangle\langle j| = 0 \\
\Rightarrow & E_\alpha^v \rho E_\beta^v = 0.
\end{aligned}$$

Hence, we have:

$$\begin{aligned}\rho =& I\rho I = \sum_\alpha E^v_\alpha \rho \sum_\beta E^v_\beta \\ =& \sum_{\alpha,\beta} E^v_\alpha \rho E^v_\beta = \sum_\alpha E^v_\alpha \rho E^v_\alpha .\end{aligned}$$

This last fact implies that ρ commutes with all operators (to see this, use the fact that $(E^v_\alpha)^2 = E^v_\alpha$). Hence we have, using the fact that $|\psi\rangle$ has full Schmidt rank,

$$E^v_\alpha = \mathrm{supp}(\sqrt{\rho} F^v_\alpha \sqrt{\rho}) = \mathrm{supp}(F^v_\alpha \rho) = F^v_\alpha . \tag{3.9}$$

- *We can express the consistency conditions* (3.5) *and* (3.6) *just in terms of Alice's projectors as:* $\forall (v, w) \in E$ *and* $\forall \alpha$, $\langle E^v_\alpha , E^w_\alpha \rangle = 0$.

  We have that $|\psi\rangle$ is the maximally entangled state with local dimension rc. It follows that for all v, α and β, $\mathrm{Tr}(E^v_\alpha \otimes F^w_\beta |\psi\rangle\langle\psi|) = \frac{1}{rc}\mathrm{Tr}(E^v_\alpha F^w_\beta)$. We also have that for all v and α, $E^v_\alpha = F^v_\alpha$. Then $\frac{1}{rc}\mathrm{Tr}(E^v_\alpha F^w_\alpha) = 0$ if and only if $\langle E^v_\alpha , E^w_\alpha \rangle = 0$, and we can write the consistency conditions as wanted.

Starting from any quantum c-coloring of G, we have constructed a quantum c-coloring in normal form, consisting of $|\psi\rangle, \{E^v_\alpha\}_{v\in V, \alpha\in[c]}$. □

It is natural to distinguish between different types of quantum chromatic number according to the rank of the POVM elements used in the strategies of Alice and Bob.

Definition 3.3.5. *The* rank-r quantum chromatic number $\chi_q^{(r)}(G)$ *of G is the minimum number of colors c such that G has a quantum c-coloring consisting of projectors of rank at most r and a maximally entangled state of local dimension rc.*

It follows that for all $r \geq s$ we have $\chi_q^{(r)}(G) \leq \chi_q^{(s)}(G)$ and therefore

$$\chi_q(G) = \min_r \{\chi_q^{(r)}(G)\}. \tag{3.10}$$

We remark that in this last definition the c-coloring need not be in normal form.

3.3.1 Rank-1 Quantum Chromatic Number

We now restrict our attention to the rank-1 quantum chromatic number $\chi_q^{(1)}(G)$. It follows from (3.10) that the rank-1 quantum chromatic number is an upper bound on the quantum chromatic number. We also know an example where the rank-1 quantum chromatic number is strictly greater than the quantum chromatic number [FIG11]. In a rank-1 quantum c-coloring, we have that the maximally entangled state has local dimension c and that the rank-1 projectors for each vertex v can be seen as outer products $|a_{v\alpha}\rangle\langle a_{v\alpha}|_{\alpha\in[c]}$ of an orthonormal basis $\{|a_{v\alpha}\rangle\}_{\alpha\in[c]}$. Then the consistency condition (3.7) becomes

$$\forall(v,w)\in E, \forall\alpha\in[c],\ \langle a_{v\alpha}|a_{w\alpha}\rangle = 0. \tag{3.11}$$

As explained in [CMN$^+$07], a rank-1 quantum c-coloring of G induces a *matrix representation* of G, which is a map $\Phi : V \to \mathbb{C}^{c\times c}$ such that for all $(v,w)\in E$, the diagonal of $\Phi(v)^\dagger\Phi(w)$ is 0. This is obtained as follows. For all vertices $v\in V$ consider the unitary matrix U_v mapping the computational basis $\{|i\rangle\}_{i\in[c]}$ to $\{|a_{v\alpha}\rangle\}_{\alpha\in[c]}$. This is a $c\times c$ matrix and because of condition (3.11), if (v,w) is an edge then the diagonal entries of $U_v^\dagger U_w$ are zero.

The results in [CMN$^+$07, Proposition 7], give the following:

Proposition 3.3.6. *For all graphs G,*

$$\xi(G) \le \chi_q^{(1)}(G) \le \chi(G).$$

Our results in this section answer some questions about the relation between these three quantities that were left open in [CMN$^+$07]. For all graphs G, we give a necessary and sufficient condition for $\xi(G) = \chi_q^{(1)}(G)$, using a relation between the rank-1 quantum chromatic number and the orthogonal representation of a particular Cartesian product. Then, using the properties of Kochen-Specker sets in two different ways, we first give a class of graphs for which the rank-1 quantum chromatic number is strictly greater than the orthogonal rank. Later, for all graphs G, we give a necessary and sufficient condition for $\chi_q^{(1)}(G) < \chi(G)$.

Equality between rank-1 quantum chromatic number and orthogonal rank of product graphs

The following proposition will help us to characterize the graphs for which there is equality between orthogonal rank and the rank-1 quantum chromatic number. Let K_c be the complete graph on c vertices and let $\square$ denote the graph Cartesian product, as defined in Section 3.2.

Proposition 3.3.7. *For all graphs G,*

$$\chi_q^{(1)}(G) = \min\{c : \xi(G\square K_c) = c\}.$$

Proof. We first prove that we can map any orthogonal representation in $\mathbb{C}^c$ of $G' = G\square K_c$ to a matrix representation in $\mathbb{C}^{c\times c}$ of G, and vice versa.

Let $\{1,\ldots,c\}$ be the vertex set of K_c. The vertex set of G' is $V(G') = V(G)\times\{1,\ldots,c\}$. There is an edge in $E(G')$ between vertices (v,i) and (w,j) if either $v = w$ and $i \neq j$ or $(v,w)\in E(G)$ and $i = j$. Thus an orthogonal representation $\{a_{(v,i)}\}_{(v,i)\in V(G')}$ of G' can be mapped to a matrix representation of G as follows: for all $v \in V(G)$, let the i-th column of U_v be $a_{(v,i)}/\left\|a_{(v,i)}\right\|$. One can check that this is a valid matrix representation in $\mathbb{C}^{c\times c}$ for G. Similarly we can map matrix representations of G to orthogonal representations of G'.

We now prove the main statement. Let G be a graph with $\chi_q^{(1)}(G) = d$, then by the discussion above we can find an orthogonal representation of $G\square K_d$ in d dimensions starting from the matrix representation.

We also know that $\xi(G\square K_d) \geq d$, because there exist subgraphs of $G\square K_d$ isomorphic to K_d and $\xi(K_d) = d$. Hence we have $\xi(G\square K_d) = \chi_q^{(1)}(G) = d$. Now suppose that $\min\{c : \xi(G\square K_c) = c\} = d' < d$. Then there exists an orthogonal representation of $G\square K_{d'}$ in $\mathbb{C}^{d'}$. We can map such orthogonal representation to a matrix representation for G in $\mathbb{C}^{d'\times d'}$. But then $\chi_q^{(1)}(G) = d'$, contradicting the assumption. Therefore we have $\chi_q^{(1)}(G) = d = \min\{c : \xi(G\square K_c) = c\}$. □

We are now able to prove the following theorem.

Theorem 3.3.8. *For all graphs G,*

$$\chi_q^{(1)}(G) = \xi(G) \iff \xi(G\square K_{\xi(G)}) = \xi(G).$$

Proof. One can see that for all pairs of graphs G, H we have $\xi(G\square H) \geq \max\{\xi(G), \xi(H)\}$ because there exist subgraphs of $G\square H$ isomorphic to G and subgraphs of $G\square H$ isomorphic to H. We also have that for all c, $\xi(K_c) = c$. Hence, we have that $\xi(G\square K_c) \geq \max\{\xi(G), c\}$. Using this and Proposition 3.3.7, we observe that

$$\chi_q^{(1)}(G) = \min\{c : \xi(G\square K_c) = c\} \geq \xi(G),$$

with equality if and only if $\xi(G\square K_{\xi(G)}) = \xi(G)$. □

On the basis of Proposition 3.3.7 and following [Hog07] we can upper bound the rank-1 quantum chromatic number, in terms of a positive-semidefinite rank. Let S_n denote the set of $n \times n$ real symmetric matrices. Then for $A \in S_n$, the graph $\mathcal{G}(A) = (V, E)$ is the graph with vertex set $V = \{1, \ldots, n\}$ and edge set $E = \{(i, j) : A_{ij} \neq 0\}$. The set of positive-semidefinite matrices of the graph G is

$$\mathcal{S}_+(G) = \{A \in S_n : A \succeq 0,\ \mathcal{G}(A) = G\},$$

and the *positive-semidefinite minimum rank* of G is

$$\mathrm{mr}_+(G) = \min\{\mathrm{rank}(A) : A \in \mathcal{S}_+(G)\}.$$

From [Hog07, Observation 1.2] we have $\xi(G) \leq \mathrm{mr}_+(\overline{G})$, and from Proposition 3.3.7 we have:

$$\chi_q^{(1)}(G) \leq \min\{c : \mathrm{mr}_+(\overline{G\square K_c}) = c\}.$$

This observation may be useful for future work about the complexity of computing the quantum chromatic number.

Separation between rank-1 quantum chromatic number and orthogonal rank using KS sets

We know that for all graphs G, we have $\xi(G) \leq \chi_q^{(1)}(G)$. One can see that, given a matrix representation of G in $\mathbb{C}^{n\times n}$, one can obtain an orthogonal representation of G in $\mathbb{C}^n$ (take the first row of each matrix). We now exhibit graphs with rank-1 quantum chromatic number *strictly larger* than the orthogonal rank. These graphs are the orthogonality graphs of Kochen-Specker sets (see Section 3.2.2).

Theorem 3.3.9. *There are graphs G such that $\xi(G) < \chi_q^{(1)}(G)$.*

Proof. Let $S \subseteq \mathbb{C}^3$ be a KS set. We exhibit a graph G_S such that $\xi(G_S) = 3$ but $\chi_q^{(1)}(G_S) > 3$. The vertices of G_S are all the vectors in S, and there is an edge between orthogonal vectors. The graph G_S has obviously an orthogonal representation of dimension 3. We show now that it is not 3-colorable. Suppose we are able to 3-color the graph, and let f be a function that maps a vector in S to 1 if it has color 1, and to zero otherwise. Every orthonormal basis $b \subseteq S$ is a clique in G_S, so we have $\sum_{u\in b} f(u) = 1$, contradicting the assumption that S is a KS set. In [CMN$^+$07, Proposition 11] it is proven that for all graphs G, $\chi_q^{(1)}(G) = 3$ if and only if $\chi(G) = 3$, so we conclude that $\chi_q^{(1)}(G_S) > 3$. □

Orthogonality graphs of KS sets are not the only ones that exhibit a separation between rank-1 quantum chromatic number and chromatic number. We know about the existence of a small graph with rank-1 quantum chromatic number and chromatic number equal to 4, but orthogonal rank 3, based on [CMN$^+$07, Proposition 11]. It is the orthogonality graph of a set of 13 vectors of dimension 3 [MO]. This set is not a Kochen-Specker set, as there are no KS sets in $\mathbb{C}^3$ smaller than 18 vectors [AOW11].

3.3.2 Quantum chromatic number and KS sets

We now prove that *projective* KS sets characterize all the graphs with a separation between the chromatic number and the quantum chromatic number.

Theorem 3.3.4 allows us to identify a quantum c-coloring with Alice's multiset of projectors, denoted as $\{P_\alpha^v\}_{v\in V,\alpha\in[c]}$.

Theorem 3.3.10. *For all graphs G, we have that $\chi(G) > \chi_q(G) =: c$ if and only if for all quantum c-colorings in normal form, $S = \bigcup_{v\in V,\alpha\in[c]}\{P_\alpha^v\}$ is a projective KS set.*

Proof. ⇒) Let $\chi(G) > \chi_q(G) =: c$ and let S be the union of Alice's projectors in a quantum c-coloring in normal form. We now show that if S *is not* a projective KS set, then we can properly c-color the graph, contradicting the assumption that $\chi(G) > c$. If S is not a projective KS set then there exists a marking function

$f : S \to \{0,1\}$ such that for all orthogonal $P, P' \in S$ we have $f(P) = 0$ or $f(P') = 0$. We can use the function f to c-color the graph as follows:

$$\text{color}(v) = \alpha \text{ if } f(P^v_\alpha) = 1.$$

This is a proper c-coloring for the following two reasons. First, the quantum coloring associates each vertex to a projective measurement, and since f is a marking function, exactly one projector per measurement is mapped to 1. Second, this property of f and the consistency condition (3.7) ensure that we never color adjacent vertices with the same color.

$\Leftarrow$) Let $\chi_q(G) = c$ and assume that for all quantum c-colorings in normal form, the union of Alice's projectors is a projective KS set. Now suppose, towards a contradiction, that it is possible to classically c-color the graph. Then for each $v \in V$ with classical color α, define the projective measurement $\{P^v_i = |i+\alpha\rangle\langle i+\alpha|\}_{i\in\{0,\ldots,c-1\}}$ (where the addition is modulo c). One can see that this is a valid quantum c-coloring, and the union of its vectors consists of one projective measurement only. Thus it is not a projective KS set, because you can define a function that maps $|1\rangle\langle 1|$ to 1 and all other projectors to 0. This is a contradiction with the assumption that the union of Alice's projectors is a projective KS set. □

We remark that Theorem 3.3.10 can also be proven starting from Theorem 3.2.9. However, we prefer the direct approach to underline the structural relationship between graphs with $\chi(G) > \chi_q(G)$ and orthogonality graphs of projective KS sets. Also notice that, because of the bijection between vectors and rank-1 projectors, *weak* KS sets characterize all the graphs with a separation between the chromatic number and the *rank-1* quantum chromatic number.

3.4 Quantum Independence Number

In this section we define the quantum independence number of a graph and study its properties.

In [MSS13] we presented the results of this chapter in terms of zero-error information theory. However, a new definition came in [RM12], after our paper. We decided to adopt this new definition, since it better captures the nature of the problem. The work by Roberson and Mančinska discusses a framework that defines in the quantum regime all the graph parameters that can be expressed as graph homomorphisms. For simplicity, here we leave homomorphisms apart and give a direct definition of the quantum independence number.

As with the chromatic number, the quantum independence number can be defined in terms of a non-local game. Informally, the *independent set game* with parameter t for a graph $G = (V, E)$ is as follows. Two players, Alice and Bob, claim that they know an independent set I of G consisting of t vertices. A referee wants to test this claim with a one-round game. He forbids communication between the players, generates two random numbers $x, y \in [t]$ and separately asks Alice to provide the x-th vertex of I and Bob to provide the y-th vertex of I. The players are required to output the same vertex if $x = y$, and to output non-adjacent vertices if $x \neq y$. A formal definition follows.

Definition 3.4.1. *The independent set game with parameter t on the graph $G = (V, E)$ is a non-local game with input sets $X = Y = [t]$, output sets $A = B = V$. Alice gets input x and outputs v, Bob gets input y and outputs w. The players* lose the game *in the following two cases:*

1. $x = y$ *and* $v \neq w$
2. $x \neq y$ *and* $(v, w) \in E$ *or* $v = w$

A classical strategy consists w.l.o.g. of two deterministic functions $f_A : [t] \to V$ for Alice and $f_B : [t] \to V$ for Bob. Shared randomness, as seen for the coloring game, is not beneficial. A little thought will show that to win with probability 1, we must have $f_A = f_B$ (to avoid the first losing condition) and that $\{f_A(1), \ldots, f_A(t)\}$ must be a valid independent set of the graph of size t (to avoid the second losing condition). It follows that the classical players cannot win the game with probability 1 when $t > \alpha(G)$.

It is proven in [RM12] (with a proof almost identical to the proof of Proposition 3.3.4) that w.l.o.g. quantum strategies for the independent set game consist of real-valued projective measurements on a maximally entangled state and that the projective measurements of Alice and Bob are the same. Therefore we can define a *quantum independent set* of size t as a collection of t real-valued projective measurements $\{P_v^x\}_{v \in V}$ for all $x \in [t]$ that have the whole vertex set as outputs, with the following consistency condition:

$$\text{for all } (u, v) \in E \text{ or } u = v \text{ and for all } x \neq x', \quad P_u^x P_v^{x'} = 0. \tag{3.12}$$

Definition 3.4.2. *For all graphs G, the* quantum independence number $\alpha_q(G)$ *is the maximum number t such that there exists a quantum independent set of G of size t.*

In the following sections we show three different ways to construct graphs with a separation between quantum and classical independence number. In Chapter 4 we will see that the quantum independence number is a lower bound on the one-shot entanglement-assisted channel capacity in zero-error information theory. Since the classical independence number is related to the classical one-shot channel capacity, all the separations in this chapter imply separations in the information theory framework.

3.4.1 Separation using projective KS sets

It follows from a result in [CLMW10] in the zero-error information theory context that every weak Kochen-Specker set can be used to construct a graph for which $\alpha(G) < \alpha_q(G)$. By the same line of argument, we now prove that also projective KS sets can be used for this purpose (Theorem 3.4.4) as well as a weak converse of this statement (Theorem 3.4.5).

Let $|\Psi\rangle = \frac{1}{\sqrt{n}} \sum_{i\in[n]} |i\rangle|i\rangle$, where $\{|i\rangle\}_{i\in[n]}$ is the standard basis of $\mathbb{C}^n$. We start by proving the following technical lemma.

Lemma 3.4.3. *Let $G = (V, E)$ be a graph whose vertex set V can be partitioned into k cliques $S_1, \dots, S_k$, not necessarily of the same size. Assume that there is an assignment of a projector P_v to each vertex v such that:*

1. *for all $i \in [k]$, we have $\sum_{v\in S_i} P_v = I$.*
2. *for all edges (v, w), we have $\mathrm{Tr}(P_v P_w) = 0$,*

Then, $\alpha_q(G) \geq k$.

Proof. We define projective measurements $\{P_v^x\}_{v\in V}$ for all $x \in [k]$ as follows: $P_v^x = P_v$ if $v \in S_x$ and $P_v^x = 0$ otherwise.

We now check that this is a winning strategy for the independent set game on G. By the first property, each projector is a valid projective measurement. By the second property and by the disjointness of the cliques, such projectors satisfy the consistency condition (3.12). Hence, the above strategy is a winning strategy and we conclude that $\alpha_q(G) \geq k$. □

We are now ready to prove the following.

Theorem 3.4.4 (Projective KS set ⇒ separation)**.** *Let S be a projective KS set. Let $S_1, \dots, S_k \subsetneq \mathcal{Q}_n$ be all the subsets of S such that $\sum_{P\in S_i} P = I$. Then the orthogonality graph G of the multiset $S_1 \uplus \cdots \uplus S_k$ satisfies $\alpha(G) < \alpha_q(G)$.*

Proof. Observe that every S_i is a projective measurement, so the vertices of G can be partitioned in k cliques $S_1, \dots, S_k$. Let T be a maximal independent set in G. Suppose towards a contradiction that $|T| = k$, i.e., T is a multiset of projectors containing exactly one element per clique. Since G contains one clique per measurement and orthogonal elements are joined by edges, we have that if $P \in T$ is part of ℓ measurements, then T contains ℓ copies of P. Define a marking function for S as:

$$f(P) = 1 \iff P \in T.$$

It is a marking function for S because by assumption $S_1, \dots, S_k \subseteq \mathcal{Q}_n$ are all the projective measurements in S and f selects exactly one element from each S_i. Moreover, f does not mark any pair of orthogonal elements because T is an independent set and G is an orthogonality graph. The existence of f contradicts the assumption that S is a projective KS set. Therefore, $\alpha(G) < k$. To see that $\alpha_q(G) \geq k$, partition the vertices of G into k cliques $S_1, \ldots, S_k$ and use Lemma 3.4.3. □

The following theorem provides a converse of Theorem 3.4.4.

Theorem 3.4.5 (Separation ⇒ projective KS set)**.** *Let $G = (V, E)$ be a graph with $\alpha(G) < k$. If there exists a quantum independent set $\{P_v^x\}_{v \in V}$ for all $x \in [k]$, then*

$$S = \{P_v^x : v \in V, x \in [k]\}$$

is a projective KS set.

Proof. We prove that if S is *not* a projective KS set, then we can construct an independent set of size k and thus $\alpha(G) \geq k$. If S is not a projective KS set, then there exists a marking function $f : S \to \{0, 1\}$ such that for any $P, P' \in S$ for which $\mathrm{Tr}(PP') = 0$, $f(P) = 0$ or $f(P') = 0$. Consider the set

$$J = \{v \in V : f(P_v^x) = 1, \text{ for some } x \in [k]\}.$$

We now show that J is an independent set of size k. From the fact that f selects one projector from each of the k measurements, we obtain that $|J| = k$. From the fact that f cannot mark two orthogonal projectors and from Condition 3.12 we get that J is an independent set. □

We remark that Theorems 3.4.4 and 3.4.5 also follow from Theorem 3.2.9, by using a reduction from one-shot zero-error channel capacity to pseudo-telepathy games given in [CLMW10]. However, the direct approach taken in the proofs above more clearly shows the relationship between orthogonality graphs of projective KS sets and graphs having separations between quantum and classical independence number.

3.4.2 Separation using properties of chromatic numbers

Here we use the relationships described in Section 3.3.2 to show that every graph with a separation between quantum and classical chromatic number can be used to construct a graph with separation between quantum and classical independence number. Using this fact we find a new class of graphs with large separation.

The main result of this section needs the following lemmas. Here, the symbol □ denotes the Cartesian graph product (see 3.2.1) and K_k is the complete graph on k vertices.

Lemma 3.4.6. *Let G be a graph on n vertices with $\chi(G) > k$. Then we have $\alpha(G \square K_k) < n$.*

Proof. The vertex set of $G\Box K_k$ can be partitioned into n disjoint cliques of size k. Towards a contradiction, suppose $\alpha(G\Box K_k) \geq n$. Then an independent set of size n must contain exactly one vertex from each clique in the partition. We can get a k-coloring for G, as follows: if (v,i) belongs to the independent set, color $v \in E(G)$ with the i-th color. This is a proper coloring because, by definition of the Cartesian product of graphs, for all $(u,v) \in E(G)$ we have $((u,i),(v,i)) \in E(G\Box K_k)$, and hence u and v will not both get color i. This contradicts the assumption that $\chi(G) > k$. □

Lemma 3.4.7. *Let G be a graph on n vertices and $\chi_q(G) \leq k$. Then we have $\alpha_q(G\Box K_k) = n$.*

Proof. Let $G' = G\Box K_k$. We first show that $\alpha_q(G') \geq n$. Note that the vertex set of G' can be partitioned into n disjoint cliques of size k. Now consider an quantum k-coloring of G in normal form, $\{P_i^v\}_{v\in V, i\in[k]}$. Assign each projector P_i^v to the vertex (v,i) of G'. By the properties of a quantum coloring in normal form, this assignment satisfies the requirements of Lemma 3.4.3. Therefore, $\alpha_q(G') \geq n$.

Now note that $\overline{K}_n \boxtimes K_k$ is a subgraph of $G\Box K_k$, where $\boxtimes$ denotes the strong product of graphs and $\overline{K}_n$ is the complement of K_n. Since for all graphs H and subgraphs H' we have $\vartheta(H) \leq \vartheta(H')$ (see [Lov79]) and since the theta number is multiplicative under the strong product (see (3.3)), we obtain

$$\vartheta(G\Box K_k) \leq \vartheta(\overline{K}_n \boxtimes K_k) = \vartheta(\overline{K}_n)\vartheta(K_k) = n.$$

Lovász ϑ is an upper bound for the quantum independence number [DSW13, Bei10], therefore we conclude

$$n \leq \alpha_q(G\Box K_k) \leq \vartheta(G\Box K_k) \leq n.$$

□

Combining the results from the above lemmas we obtain the following.

Theorem 3.4.8. *Let G be a graph on n vertices with $\chi(G) > \chi_q(G) =: k$. Then for $G' = G\Box K_k$:*

1. *$\alpha(G') < \alpha_q(G') = n$*
2. *$\alpha(G') \leq \alpha(G) \cdot k$.*

Proof. Since $\chi(G) > k$, $\chi_q(G) = k$ and G has n vertices, we have from Lemma 3.4.6 and Lemma 3.4.7 that $\alpha(G') < \alpha_q(G') = n$, as desired. The second bound follows directly from Lemma 3.2.1. □

Note that the second upper bound on Theorem 3.4.8 is very interesting in the case $\alpha(G) \cdot \chi_q(G) \ll n$. This happens only when there is a separation between quantum and classical chromatic number, because for the chromatic number we have $\alpha(G) \cdot \chi(G) \geq n$. Therefore, as we show in the next section, some graphs with a large separation between quantum and classical chromatic number induce graphs with large separation between quantum and classical independence number.

Orthogonality graphs

In this section we apply the observations made above. Specifically, we isolate a new family of graphs for which the quantum independence number is exponentially larger than its classical counterpart. Each member of this family is a Cartesian product of an orthogonality graph with a complete graph. Chromatic number and quantum chromatic number are known to be different for orthogonality graphs with a sufficiently large number of vertices [dKP07, AHKS06, GN08].

For each integer n multiple of 2, the orthogonality graph[3] Ω_n is a graph with vertex set $\{\pm 1\}^n$ and edge set $\{(u,v) : \langle u, v\rangle = 0\}$. Some of these graphs (for certain values of n) are also known in the literature as Hadamard graphs and Deutsch-Jozsa graphs.

Theorem 3.4.9. *For all $n > 8$ that are divisible by 4, there exists $\varepsilon > 0$ such that*

$$\frac{\alpha_q(\Omega_n \Box K_n)}{\alpha(\Omega_n \Box K_n)} \geq \frac{1}{n}\left(\frac{2}{2-\epsilon}\right)^n.$$

Proof. It is shown in [AHKS06] that $\chi_q(\Omega_n) \leq n$ for all $n \in \mathbb{Z}_+$. Since $|V(\Omega_n)| = 2^n$, using Lemma 3.4.7 we conclude that $\alpha_q(\Omega_n \Box K_n) \geq 2^n$.

On the other hand, from Theorem 1.11 in [FR87] it follows that for all n divisible by 4 and greater than 8, there exists $\epsilon > 0$ such that $\alpha(\Omega_n) \leq (2-\epsilon)^n$. Hence, by Lemma 3.2.1, we have that $\alpha(\Omega_n \Box K_n) \leq (2-\epsilon)^n \cdot n$. By putting the two observations together we obtain the desired statement. □

To conclude, we give an example that also for small n we can find a large ratio $\frac{\alpha_q(\Omega_n \Box K_n)}{\alpha(\Omega_n \Box K_n)}$. The following properties are proven in [dKP07, AHKS06]:

1. $\alpha(\Omega_{16}) = 2304$
2. $\chi(\Omega_{16}) \geq 29$
3. $\chi_q(\Omega_{16}) = 16$.

Take a graph $\Omega_{16} \Box K_{16}$. It follows from Theorem 3.4.8 that $\alpha_q(G) = 2^{16}$ while $\alpha(\Omega_{16} \Box K_{16}) \leq \alpha(\Omega_{16}) \cdot 16 = 36864$.

[3] We have defined orthogonality graphs in general in Section 3.2.1. The ones discussed here are a special case, but we keep the naming consistent with literature.

3.4.3 Separation using graph states

In this section we show how *starting from any graph G* we can construct a graph $H(G)$ with separation between quantum and classical independence number. The drawback of this method is that we do not know how to quantify such separation. The construction is used in [CPSS12] to study orbits of graphs under local complementation, but this is out of the scope of this thesis.

The useful tools for this construction are graph states, which we define below. Given a graph $G = (V, E)$, a generator for $i \in V(G)$ is defined as

$$g_i = X^{(i)} \bigotimes\nolimits_{j \in \mathcal{N}(i)} Z^{(j)}, \tag{3.13}$$

where $X^{(i)}$, $Y^{(i)}$, and $Z^{(i)}$ denote the Pauli matrices acting on the i-th qubit and $\mathcal{N}(i)$ is the neighborhood of i. Identity matrices act on the qubits corresponding to non-adjacent vertices, but we omit them to simplify the notation. Therefore, g_i can be obtained directly from G. The *graph state* $|G\rangle$ (see, for example, [HEB04, SW01]) associated to G is the unique (up to a phase) n-qubit state such that

$$g_i|G\rangle = |G\rangle \text{ for } i = 1, \dots, n. \tag{3.14}$$

The *stabilizer group* of the state $|G\rangle$ is the set S of the stabilizing operators s_j of $|G\rangle$ defined by the product of any number of generators g_i. For convenience, we remove the identity element from S. Therefore, the set S contains $2^n - 1$ elements.

We will now explain how to construct a graph $H(G)$ from any graph G. Let G be a graph on n vertices and consider the n-qubit graph state $|G\rangle$. Let S be the stabilizer group of G. For each $s_j \in S$ of the form $s_j = \bigotimes_{k=1}^{n} O^{(k)}$, where each $O^{(k)}$ is a Pauli matrix, let $w_j = |\{O^{(k)} : O^{(k)} \neq I\}|$ be the *weight of s_j*. Let $S_j = \{S_{(i,j)} : i = 1, 2, \dots, 2^{w_j - 1}\}$ be the set of the *events* of s_j, *i.e.* the measurement outcomes that can occur with non-zero probability when the system is in state $|G\rangle$ and the stabilizing operators s_j are measured with single-qubit measurements. The set of all events is $\mathcal{S} = \bigcup_{j=1,2,\dots,2^n-1} S_j$. Two events are *exclusive* if there exists a $k \in \{1, \dots, n\}$ for which the same single-qubit measurement gives a different outcome.

We give an example for events and exclusiveness. Let $n = 3$ and $s_2 = ZXZ$ (we omit the superscripts for simplicity). This means that $ZXZ|G\rangle = |G\rangle$, *i.e.*, if the system is prepared in $|G\rangle$ and s_2 is measured by measuring Z on the first qubit (with possible results -1 or 1), X on the second qubit, and Z on the third qubit, then the product of the three results must be 1. Therefore, $S_2 = \{zxz, z\underline{xz}, \underline{z}x\underline{z}, \underline{zx}z\}$, where hereafter $z\underline{xz}$ denotes the event "the result 1 is obtained when Z is measured on qubit 1, the result -1 is obtained when X is measured on qubit 2, and the result -1 is obtained when Z is measured on qubit 3". As another example: if $n = 2$ and $s_1 = XZ$, then $S_1 = \{zx, \underline{zx}\}$.

We now define the graph $H(G)$ starting from the exclusivity structure of the events explained above.

Definition 3.4.10. *Let G be a graph. Let S be the stabilizer group of the graph state of G. We denote by $H(G)$ the graph whose vertices are the events in S and the edges are all the pairs of exclusive events.*

We give an example of the construction. Let us consider $G = P_3$, the path on three vertices, with $V = \{1, 2, 3\}$ and $E = \{\{1, 2\}, \{2, 3\}\}$. We construct $H(P_3)$. The stabilizer group S (minus the identity) has the following elements: $s_1 = g_1 = XZI$, $s_2 = g_2 = ZXZ$, $s_3 = g_3 = IZX$, $s_4 = g_1g_2 = YYZ$, $s_5 = g_1g_3 = XIX$, $s_6 = g_2g_3 = ZYY$, and $s_7 = g_1g_2g_3 = -YXY$. For all $j = 1, \ldots, 2^3 - 1$, obtain all possible events (*i.e.*, those which can happen with non-zero probability) when three qubits are prepared in the state $|G\rangle$ and three parties measure the observables corresponding to s_j. For instance, when $j = 1$, Alice measures $X^{(1)}$, Bob measures $Z^{(2)}$, and Charlie does not perform any measurement. Since the three qubits are in state $|G\rangle$, there are only two possible outcomes: Alice obtains $X^{(1)} = +1$ and Bob obtains $Z^{(2)} = +1$, denoted as xzI; or Alice obtains $X^{(1)} = -1$ and Bob obtains $Z^{(2)} = -1$, denoted as $\underline{x}\underline{z}I$. For $j = 2$, the only events that can occur are zxz, $z\underline{x}\underline{z}$, $\underline{z}x\underline{z}$, and $\underline{z}\underline{x}z$. The other events for the remaining j's are obtained in a similar way. Now, let us construct the graph $H(P_3)$: the vertices represent possible events; two vertices are adjacent if and only events are exclusive (*e.g.*, xzI and $\underline{x}I\underline{x}$). Notice that each s_j of weight w_j generates $2^{w_j - 1}$ vertices. A drawing of $H(P_3)$ is in Fig. 3.1.

We now prove that this graph has separation between quantum and classical independence number. We now prove the classical upper bound.

Theorem 3.4.11. *Let G be a graph on $n > 2$ vertices and let $H(G)$ be as in Definition 3.4.10. Then,*

$$\alpha(H) < 2^n - 1. \tag{3.15}$$

Proof. For simplicity let $H = H(G)$. We use an argument very similar to [GTHB05, Lemma 1 and Theorem 1]. Each connected graph with more than two vertices has a subgraph with three vertices. For each of those we can see that $\alpha(H) < 7$ (by direct calculation, see [CPSS12, Table 1]). Therefore, we just need to show that if G' is a subgraph of G with n' vertices and $\alpha(H') < 2^{n'} - 1$, where $H' = H(G')$, then $\alpha(H) < 2^n - 1$ for $n > 2$. Notice that S', the stabilizer group of G', is a subset of S. Therefore, in the graph H we find cliques associated with S', but containing slightly different events. For each $s'_i \in S'$, the corresponding $s_i \in S$ has the same structure, with eventually some additional Z operators. Let $\tilde{H}$ be the subgraph of H induced by the vertices in cliques associated with the elements of S'. We need to show that if in H' there is no vertex per clique to form a maximal independent set then neither are there in $\tilde{H}$. Therefore, $\alpha(H) < 2^n - 1$. Towards a contradiction, suppose there is an independent set L of $\tilde{H}$ such that $|L| = 2^{n'} - 1$. We distinguish two cases:

- If the events at the vertices in L do not have any $\underline{z}$ element then we can map

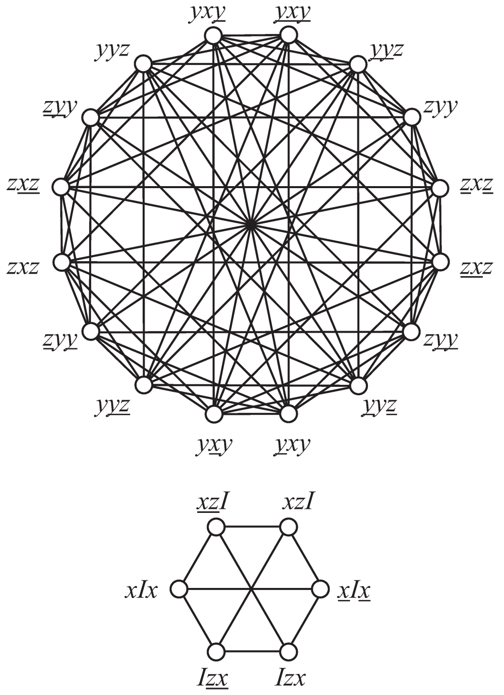


Figure 3.1: The graph $H(P_3)$ associated to the path on three vertices, P_3, consisting of two connected components.

them to an independent set in H' of size $2^{n'} - 1$, just by ignoring the additional Z operators. This contradicts the hypothesis that $\alpha(H') < 2^{n'} - 1$.

- If the events at the vertices in L do have $\underline{z}$ elements then we can find another independent set J with the same cardinality such that the events at its vertices do not have any $\underline{z}$ element. We can find J as follows. One can check that an operator s_i has the form $O^{(1)} \cdots Z^{(\ell)} \cdots O^{(n)}$ if and only if it has an odd number of $X^{(k)}$ and $Y^{(k)}$, with $\{\ell, k\} \in E(H)$. Therefore, complementing $\underline{z^{(\ell)}}$ and all occurrences of $X^{(k)}$ and $Y^{(k)}$ in the events at the vertices of the independent set L, we obtain the events in J with the desired properties, and so we are back to the previous case.

□

For each $H(G)$, we define an orthogonal representation and use it to calculate the Lovász number.

Definition 3.4.12 (Canonical orthogonal representation)**.** *Let $H = (V, E)$ be a graph as in Definition 3.4.10. Consider the vertex of H associated to an event $S_{(i,j)} = \left(s^{(1)}_{(i,j)}, s^{(2)}_{(i,j)}, \ldots, s^{(n)}_{(i,j)}\right)$, where $i = 1 \ldots 2^n - 1$, $j = 1 \ldots 2^{w_i - 1}$, and $s^{(k)}_{(i,j)} \in$*

$\{I, x, \underline{x}, y, \underline{y}, z, \underline{z}\}$, for each $k = 1, 2, \ldots, n$. Let $|s_{(i,j)}^{(k)}\rangle$ be defined as follows:

$$\begin{array}{c|ccccccc} s_{(i,j)}^{(k)} & x & \underline{x} & y & \underline{y} & z & \underline{z} & I \\ \hline |s_{(i,j)}^{(k)}\rangle & |+\rangle & |-\rangle & |y_+\rangle & |y_-\rangle & |0\rangle & |1\rangle & |\psi\rangle \end{array}. \tag{3.16}$$

Here, $|\psi\rangle$ is an arbitrary unit vector in $\mathbb{C}^2$ and $|y_+\rangle, |y_-\rangle$ are the eigenvectors of the Pauli matrix Y with eigenvalue $+1$ and -1, respectively. The canonical orthogonal representation *of H is the set of vectors $\{|s_{(i,j)}\rangle = |s_{(i,j)}^{(1)}\rangle \otimes |s_{(i,j)}^{(2)}\rangle \otimes \cdots \otimes |s_{(i,j)}^{(n)}\rangle : S_{(i,j)} \in V(H)\}$.*

For example, in $H(P_3)$ (see Fig. 3.1), the element of the canonical orthogonal representation of the vertex labeled by $\underline{x}I\underline{x}$ is $|-\rangle \otimes |\psi\rangle \otimes |-\rangle$.

Lemma 3.4.13. *Let G be a graph on n vertices and $H(G)$ as in Definition 3.4.10. Then,*

$$\vartheta(H(G)) = 2^n - 1.$$

Proof. Let $H = H(G)$. We first prove that $\vartheta(H) \geq 2^n - 1$. It follows directly from Eq. (3.14) that $\sum_{i=1}^{2^n-1}\langle G|s_i|G\rangle = 2^n - 1$. We know that the eigenvectors with eigenvalue $+1$ of each operator s_i are in one-to-one correspondence with the vertices of a clique in H: $|s_{(i,1)}\rangle, |s_{(i,2)}\rangle, \ldots, |s_{(i,2^{w_i-1})}\rangle$. These are elements of the canonical orthogonal representation of H. From the definition of the stabilizer group, for all $s_i \in S$ and for all eigenvectors $|\underline{s}^{(i,j)}\rangle$ ($j = 1, 2, \ldots, 2^{w_i-1}$) with eigenvalue -1, we have $\langle \underline{s}^{(i,j)}|G\rangle = 0$, because $|G\rangle$ is in the $+1$ eigenspace. Now, let $s_i = \sum_j \lambda_{ij}|s_{(i,j)}\rangle\langle s_{(i,j)}|$ be an Hermitian eigendecomposition of s_i. Thus,

$$\begin{aligned} 2^n - 1 &= \sum_{i=1}^{2^n-1}\langle G|s_i|G\rangle \\ &= \sum_{i=1}^{2^n-1}\sum_j \lambda_{ij}\langle G|s_{(i,j)}\rangle\langle s_{(i,j)}|G\rangle \\ &= \sum_{i=1}^{2^n-1}\sum_{j:\lambda_{ij}=1}\langle G|s_{(i,j)}\rangle\langle s_{(i,j)}|G\rangle \\ &= \sum_{i=1}^{2^n-1}\sum_{j:\lambda_{ij}=1}|\langle G|s_{(i,j)}\rangle|^2 \\ &\leq \vartheta(H), \end{aligned}$$

where the inequality in the last line follows because a canonical orthogonal representation of H together with the state $|G\rangle$ represents a feasible solution for the formulation (3.1) of the Lovász number.

We now prove the upper bound $\vartheta(H) \leq 2^n - 1$. We can partition H into $2^n - 1$ disjoint cliques by considering the events associated with each s_i. Since each clique of H is an independent set in the complement $\bar{H}$, we can associate each independent set to a color and obtain $\chi(\bar{H}) \leq 2^n - 1$. It follows from the sandwich theorem [KD93] that

$$\vartheta(H) \leq \chi(\bar{H}) \leq 2^n - 1.$$

Combining the two directions, we have the desired result. □

We are now ready to show that for every graph G on n vertices, the graph $H(G)$ has quantum independence number equal to $2^n - 1$, therefore strictly larger than the classical independence number by Theorem 3.4.11. This result opens directions for future studies, for example identifying subclasses or hierarchies where the separation is large or is easy to quantify.

Theorem 3.4.14. *Let G be a graph on n vertices and let $H(G)$ be as in Definition 3.4.10. Then,*

$$\alpha_q(H(G)) = 2^n - 1.$$

Proof. Let $H = H(G)$. We have the upper bound

$$\alpha_q(H) \leq \vartheta(H) = 2^n - 1,$$

where the inequality is [DSW13, Corollary 14] and the equality is Lemma 3.4.13. We need to show a matching lower bound on $\alpha_q(H)$. We do this by exhibiting a strategy for quantum players to win the independent set game on H with $t = 2^n - 1$. Observe that H can be partitioned into $2^n - 1$ cliques, one for each element of the stabilizer group. We denote by S_i the set of events related to the element of the stabilizer group s_i. The clique corresponding to $s_i \in S$ consists of the vertices associated with the mutually exclusive events in the set S_i. The strategy is as follows. Alice and Bob share a maximally entangled state of local dimension $2^n(n-1)$. Since w.l.o.g. Alice and Bob use the same strategy, we can describe only Alice's side. For each $i \in \{1, 2, \ldots, 2^n - 1\}$, Alice performs a projective measurement on her part of the shared state. The outcomes of the measurement are the elements of S_i. The strategy has to satisfy two properties to be correct:

1. For each $i \in \{1, 2, \ldots, 2^n - 1\}$, the projectors associated to elements of S_i form a projective measurement.

2. For each edge $\{u, v\} \in E(H)$, projectors associated with u and v must be orthogonal (to satisfy the consistency condition (3.12)).

The next step is to exhibit the projectors and show that both properties are satisfied. In what follows we use the notation in Definition 3.4.12.

We begin by examining the case where s_i does not contain any identity operator. In this case, each projective measurement will consist of projectors of rank

1 acting on $\mathbb{C}^{2^n(n-1)}$. Order the elements of S_i arbitrarily. Let s_i be of the form $O^{(1)} \cdots O^{(n)}$, where $O^{(k)} \in \{X, Y, Z\}$. Define for each $s^{(k)}_{(i,j)}$ the *occurrence number* $\nu(i, j, k)$ based on a chosen ordering: if the same eigenvector of $O^{(k)}$ occurs in $s^{(k)}_{(i,j)}$ for the ℓ-th time in the chosen ordering then $\nu(i, j, k) = \ell$. Construct projectors starting from the canonical orthogonal representation and an ancillary space of dimension $n - 1$. For $s^{(k)}_{(i,j)}$, let

$$P_{(i,j)} = \bigotimes_{k=1}^{n} |s^{(k)}_{(i,j)}\rangle\langle s^{(k)}_{(i,j)}| \otimes |\nu(i, j, k)\rangle\langle\nu(i, j, k)|. \tag{3.17}$$

We show that Property 1 is satisfied. These projectors are mutually orthogonal for all vertices (i, j). We need to prove that their sum is the identity. From the structure of the events in S_i we observe that, for each $O^{(k)}$, the eigenvectors with eigenvalue $+1$ (and -1) occur in half of the elements of S_i. Therefore, in the construction of the projectors, a pair of ± 1 eigenvectors for each $O^{(k)}$ is summed for each ancillary subspace. The sum of each subspace is the identity. Hence, the total sum is the identity for the whole space. We show now that also Property 2 is satisfied. If two projectors are in the same clique, orthogonality follows from the discussion above. Consider now two projectors of adjacent vertices from two different cliques that project to the same ancillary subspace. Since we started from an orthogonal representation, those projectors are orthogonal.

Now, consider the more general case where s_i can contain identity operators. Let s_i be of the form $O^{(1)} \cdots O^{(n)}$, where $O^{(k)} \in \{I, X, Y, Z\}$. We assume that s_i has weight w. First consider the case where the first w operators are different from Identity, $O^{(1)}, O^{(2)}, \dots, O^{(w)} \neq I$. To construct the projective measurement for S_i, we initially construct the projectors for the first w operators as in the previous case. We obtain rank-1 projectors acting on $\mathbb{C}^{2^w(w-1)}$. Choose a basis for $\mathbb{C}^{2^n(n-1)-2^w(w-1)}$ and let the projectors be

$$Q_{(i,j)} = \sum_{\ell=1}^{2^n(n-1)-2^w(w-1)} P_{(i,j)} \otimes |\ell\rangle\langle\ell|. \tag{3.18}$$

This ensures that the dimensions match and that Properties 1 and 2 hold. To finish the proof, we need to prove the general case where identity operators are in arbitrary positions and not all at the end. In this case, split the construction into subspaces so that each subspace has all the identities at the end. Obtain the projectors for the subspaces as described above and then obtain the final projectors by making tensor products of the projectors for the subspaces. □

3.5 Bounds on the value of non-local games through game graphs

This section contains previously unpublished results. It is based on a collaboration with A. Chailloux, L. Mančinska and S. Severini. We consider a construction of graphs associated with non-local games from [CSW10]. It is known that the independence number of such graphs corresponds to the classical value of the game. Here we show that quantum independence number and the Lovász theta number are bounds on the quantum value of the game.

Consider a two-prover game $\mathcal{G}$ with input sets X, Y, output sets A, B, predicate $\lambda : X \times Y \times A \times B \to \{0,1\}$ and uniform distribution on the inputs.[4]

Definition 3.5.1. *A graph $G = (V, E)$* associated *to the game $\mathcal{G}$ has:*

1. $V = \{xyab \mid x \in X, y \in Y, a \in A, b \in B \text{ and } \lambda(x,y,a,b) = 1\}$*,*
2. $E = \{\{xyab, x'y'a'b'\} \mid (x = x' \wedge a \neq a') \vee (y = y' \wedge b \neq b')\}$*.*

This definition is inspired by a construction in [CSW10] in the framework of contextuality of physical theories. The authors used something similar to Definition 3.5.1 for the special case of the CHSH game. Here we generalize to all games.

For simplicity, we prove the results in this section for the case where the game has the uniform distribution on the inputs and λ is a boolean function. It is straightforward to generalize to games with real-valued predicate and any probability distribution π of the inputs, as follows. Consider the (vertex) weighted graph with all the quadruples $xyab$ in the vertex set, labelled with weight$(xyab) = \lambda(x,y,a,b) \cdot \pi(x,y)$, and the same edge set as before. The classical bound and the Lovász theta bound that we will prove later can be adapted by considering the weighted versions of these parameters. However, we do not know how to generalize our last result because we do not define the quantum independence number for a weighted graph.

Now we prove that that the classical value of a game can be expressed in terms of the independence number of its game graph.

Theorem 3.5.2. *Let $\mathcal{G}$ be a non-local game with input sets X and Y, uniform input distribution and associated graph G. Then*

$$\omega(\mathcal{G}) = \frac{\alpha(G)}{|X \times Y|}.$$

[4] Note the change of notation: in this section we denote the game as $\mathcal{G}$ and its predicate as λ to avoid confusion with the standard notation for a graph G and for its vertex set V.

Proof. Let $k = |X \times Y|$. We begin by proving that $\omega(\mathcal{G}) \geq \alpha(G)/k$. Namely, we show that given a maximal independent set $I \subseteq V$ of size ℓ, we can exhibit a strategy for $\mathcal{G}$ that answers correctly to at least ℓ of the k questions. By the structure of G, the independent set I cannot contain vertices $xyab$ and $xy'a'b'$ such that $a \neq a'$. Similarly, I cannot contain vertices $xyab$ and $x'ya'b'$ such that $b \neq b'$. Hence, we have the following strategy: on input x, Alice outputs the unique a determined by the vertices in the independent set I. Bob behaves similarly. Since V contains only winning quadruples $xyab$, the size ℓ of the independent set means Alice and Bob answer correctly to at least ℓ input pairs. Hence, $\omega(\mathcal{G}) \geq \ell/k$.

Now we show that $\omega(\mathcal{G}) \leq \alpha(G)/k$, *i.e.*, if there exists a strategy that wins on ℓ of the k input pairs, then there exists an independent set with weight ℓ. We have that w.l.o.g. classical strategies consist of a pair of deterministic functions. Fix Alice and Bob's deterministic functions f_A and f_B that win on ℓ input pairs. Now take the set of quadruples $S = \{(x, y, f_A(x), f_B(y))\}_{x \in X, y \in Y}$. We have that $I = S \cap V$ is a set of ℓ vertices of G. Since f_A and f_B are deterministic, I cannot contain vertices $xyab$ and $xy'a'b'$ such that $a \neq a'$ nor vertices $xyab$ and $x'ya'b'$ such that $b \neq b'$. Therefore, there cannot be an edge between any pair of the elements of I and we have that I is an independent set of G of size ℓ. Hence, $\alpha(G) \geq \ell$. Combining the two directions proves the theorem. □

Bounds on the quantum value of a game

Cabello *et al.* [CSW10] observe that the quantum value of the CHSH game is equal to the theta number of its associated graph divided by the number of questions. We have found by direct calculation that this is not always true for general games, for example in the case of the 2-fold parallel repetition of CHSH. Instead, we have the following upper bound.

Theorem 3.5.3. *Let $\mathcal{G}$ be a non-local game with input sets X and Y, uniform input distribution and associated graph $G = (V, E)$. Then*

$$\omega^*(\mathcal{G}) \leq \frac{\vartheta(G)}{|X \times Y|}.$$

Proof. Let $k = |X \times Y|$. Consider a quantum strategy for $\mathcal{G}$ that achieves the value $\omega^*(\mathcal{G})$. It consists of a shared entangled state $|\psi\rangle$ and a collection of projective measurements $\{P_a^x\}, \{Q_b^y\}$, such that

$$\sum_{xyab} \frac{1}{k} \lambda(x, y, a, b) \langle\psi| P_a^x \otimes Q_b^y |\psi\rangle = \frac{1}{k} \sum_{xyab \in V} \langle\psi| P_a^x \otimes Q_b^y |\psi\rangle = \omega^*(\mathcal{G}).$$

For each quadruple $xyab$ let $|\psi_{xyab}\rangle = P_a^x \otimes Q_b^y |\psi\rangle$. This is an orthogonal representation of G, since for every edge $(xyab, x'y'a'b')$ either $P_a^x P_{a'}^{x'} = 0$ or

$Q^y_b Q^{y'}_{b'} = 0$. Now for each $xyab$ consider the normalized vector

$$|\psi'_{xyab}\rangle = \frac{|\psi_{xyab}\rangle}{||\psi_{xyab}||} = \frac{|\psi_{xyab}\rangle}{\sqrt{\langle\psi|P^x_a \otimes Q^y_b|\psi\rangle}}.$$

We have that $\{\psi'_{xyab}\}_{xyab\in V}$ and ψ are a feasible solution for the formulation (3.1) of $\vartheta(G)$.

We conclude

$$\begin{aligned}\vartheta(G) &\geq \sum_{xyab\in V} |\langle\psi|\psi_{xyab}\rangle|^2 \\ &= \sum_{xyab\in V} \left|\frac{\langle\psi|\psi_{xyab}\rangle}{||\psi_{xyab}||}\right|^2 \\ &= \sum_{xyab\in V} \frac{\langle\psi|P^x_a \otimes Q^y_b|\psi\rangle^2}{\langle\psi|P^x_a \otimes Q^y_b|\psi\rangle} \\ &= \sum_{xyab\in V} \langle\psi|P^x_a \otimes Q^y_b|\psi\rangle \\ &= k \cdot \omega^*(\mathcal{G}).\end{aligned}$$

□

We now have the following lower bound in terms of the quantum independence number.

Theorem 3.5.4. *Let $\mathcal{G}$ be a non-local game with input sets X and Y, uniform input distribution and associated graph $G = (V, E)$. Then*

$$\omega^*(\mathcal{G}) \geq \frac{\alpha_q(G)}{|X \times Y|}$$

To prove the theorem, we will use the following lemma.

Lemma 3.5.5. *Let M, N be positive semidefinite matrices. Then for any vector $|v\rangle$, we have that*

$$\langle v|\operatorname{supp}(M+N)|v\rangle \geq \langle v|\operatorname{supp}(M)|v\rangle,$$

where supp(M) *denotes the projector onto the support (*i.e., *the column space) of M.*

Proof. If P is a projector onto a subspace Π then $\langle v|P|v\rangle$ is the squared length of the projection of $|v\rangle$ into Π. Hence, to prove the lemma it suffices to show that $\operatorname{supp}(M) \subseteq \operatorname{supp}(M+N)$, where by abusing the notation we use supp to denote the support itself (rather than the projection onto it).

For contradiction, suppose that $\mathrm{supp}(M) \not\subseteq \mathrm{supp}(M+N)$. Then the orthogonal complement of $\mathrm{supp}(M)$ (i.e. the nullspace $\mathrm{Null}(M)$) does not contain $\mathrm{Null}(M+N)$. Hence we can pick a vector $|w\rangle$ such that $(M+N)|w\rangle = 0$ but $M|w\rangle \neq 0$. This further implies that

$$\langle w|N|w\rangle = \langle w|(M+N)|w\rangle - \langle w|M|w\rangle = -\langle w|M|w\rangle < 0,$$

since M is positive semidefinite and $M|w\rangle \neq 0$. This completes the proof as we have reached a contradiction with the initial assumption that N is positive semidefinite. □

Proof of Theorem 3.5.4. Given a quantum strategy $\{P^i_{xyab}\}$ for the independent set game on G with parameter t, we construct a strategy to win the game $\mathcal{G}$ with probability at least $t/|X \times Y|$, as follows.

Players share a maximally entangled state with local dimension d (which is the dimension of the projectors above). On input x, Alice measures her half of the state using the projective measurement$\{P^x_a\}_{a\in A} \bigcup \{I - \sum_a P^x_a\}$, where the individual elements are defined as follows:

$$P^x_a = \mathrm{supp}\left(\sum_{\substack{yb \\ xayb\in V}} \sum_i P^i_{xayb} \right),$$

where we use $\mathrm{supp}(M)$ to denote the projector onto the image of M. We show that this is a valid projective measurement. For all y, b, y', b' there is an edge $(xyab, xy'a'b') \in E$. Therefore in the strategy for the independent set game we have that for all i, j each projector P^i_{xyab} is orthogonal to $P^j_{xy'a'b'}$. Hence, for all $a \neq a'$ we have $P^x_a \cdot P^x_{a'} = 0$. Bob constructs projectors P^y_b similarly.

Now we lower bound the quantum value of $\mathcal{G}$ as follows:

$$\begin{aligned} |X \times Y| \cdot \omega^*(G) &\geq \sum_{xyab\in V} \langle\psi| P^x_a \otimes P^y_b |\psi\rangle \\ &= \sum_{xyab\in V} \langle\psi| \,\mathrm{supp}\Big(\sum_{i,j} \sum_{\substack{y'b' \\ xay'b'\in V}} \sum_{\substack{x'a' \\ x'a'yb\in V}} P^i_{xay'b'} \otimes P^j_{x'a'yb} \Big) |\psi\rangle, \end{aligned}$$

where we have used the fact that $\mathrm{supp}(M \otimes N) = \mathrm{supp}(M) \otimes \mathrm{supp}(N)$ for all matrices M, N to obtain the last equality. Now by applying Lemma 3.5.5, we drop all the terms except the ones with $i = j, a = a', b = b', x = x'$ and $y = y'$, and we have that

$$|X \times Y| \cdot \omega^*(G) \geq \sum_{xyab \in V} \langle\psi| \operatorname{supp}\Big(\sum_i P^i_{xayb} \otimes P^i_{xayb}\Big)|\psi\rangle \tag{3.19}$$

$$= \sum_{xyab \in V} \langle\psi|\Big(\sum_i P^i_{xayb} \otimes P^i_{xayb}\Big)|\psi\rangle \tag{3.20}$$

$$= \sum_{xyab \in V} \sum_i \frac{1}{d}\mathrm{Tr}(P^i_{xayb}) \tag{3.21}$$

$$= \sum_i \frac{1}{d}\mathrm{Tr}(I_d) \tag{3.22}$$

$$= \alpha_q(G). \tag{3.23}$$

In the above we have observed that $\operatorname{supp}(P+Q) = P+Q$ for mutually orthogonal projectors P and Q to get Expression (3.20). We have used properties of $|\psi\rangle = \frac{1}{\sqrt{d}}\sum_i |i,i\rangle$ to obtain Expression (3.21). We have used the fact that, for all i, $\{P^i_{xayb} : \lambda(x,a,y,b) = 1\}$ forms a measurement to obtain Expression (3.22). □

Pseudo-telepathy games

Here we show that from a class of pseudo-telepathy games (*i.e.*, games with quantum value 1 and classical value strictly less than 1), one can obtain graphs with separation between independence number and quantum independence number.

Theorem 3.5.6. Let $\mathcal{G}$ be a pseudo-telepathy game with a 0-1 valued verification function λ, such that the best quantum strategy uses a maximally entangled state $|\psi\rangle$. Let G be the corresponding game graph. Then,

$$\omega^*(\mathcal{G}) = \frac{\alpha_q(G)}{|X \times Y|}.$$

Proof. From Theorem 3.5.4 we have $\alpha_q(G) \leq |X \times Y| \cdot \omega^*(\mathcal{G})$. We need to prove the other direction.

Let $\{P^x_a\}, \{Q^y_b\}$ be the strategies that win the game $\mathcal{G}$ on $|\psi\rangle$. We have:

$$\sum_{xy} \pi(x,y) \sum_{ab:\lambda(xyab)=1} \langle\psi|P^x_a \otimes Q^y_b|\psi\rangle = 1,$$

so for all (x,y) we must have

$$\sum_{ab:\lambda(xyab)=1} \langle\psi|P^x_a \otimes Q^y_b|\psi\rangle = 1$$

and for all quadruples (x,y,a,b) such that $\lambda(xyab) = 0$ we have $P^x_a Q^y_b = 0$.

Let $\Pi_{xyab} = P^x_a Q^y_b$. We observe:

1. For all (x, y) we have

$$\sum_{ab:\lambda(xyab)=1} P_a^x Q_b^y = \sum_{ab} P_a^x Q_b^y = \sum_a P_a^x \sum_b Q_b^y = I,$$

   where the second equality follows from $Q_b^y Q_{b'}^y = \delta_{bb'}$.

2. For each edge $(x, y, a, b), (x', y', a', b')$ we have

$$\Pi_{xyab}\Pi_{x'y'a'b'} = 0,$$

   because if $x = x'$ and $a \neq a'$ then $P_a^x P_{a'}^x = 0$, and if $y = y'$ and $b \neq b'$ then $Q_b^y Q_{b'}^y = 0$.

Therefore, we can construct $|X \times Y|$ projective measurements that are a winning strategy for the independent set game with $t = |X \times Y|$ as follows. For each pair (x, y) consider the projective measurement $\{\Pi_{xyab}\}_{a,b:\lambda(xyab)=1}$ (and zero matrices for the other vertices of the graph). The first observation above proves that those are valid projective measurements; the second observation shows that they respect the consistency condition (3.12). □

3.6 Concluding remarks and open problems

The main contribution of Sections 3.3 and 3.4 is the introduction and use of a formal generalization of Kochen-Specker sets. In particular, we showed that projective KS sets lead to graphs for which the quantum independence number is strictly larger than the independence number. We have also shown that projective KS sets completely characterize the graphs for which the quantum chromatic number is strictly smaller than the chromatic number. Furthermore, we used projective KS sets to relate quantum chromatic number to quantum independence number. For all graphs obtained with our construction the Lovász theta function is equal to the quantum independence number. Hence, although our construction contributes to shed light on the link between the Lovász theta function and quantum graph parameters, it cannot directly be used to resolve whether or not the quantum independence number equals the Lovász theta function [DSW13]. An open question is: can we use a graph G with $\alpha(G) < \alpha_q(G)$ to construct a G' with $\chi_q(G') < \chi(G')$? Such a construction would be complementary to Theorem 3.4.8.

We showed in Section 3.5, with the use of a specific graph construction, that the quantum independence number is a bound to the value of non-local games. Moreover, we have shown that for a class of pseudo-telepathy games that quantum players can win using projective measurements on maximally entangled state, this bound is tight. The same class of games is shown in Section 3.2.2 to be in one-to-one correspondence with projective KS sets. It is not clear to us if those two results together could be used to prove something stronger. Perhaps the whole class could be interpreted as pseudo-telepathy games based on some graph parameter (maybe the homomorphism games in [RM12]) and the relationship to the quantum independence number would be a consequence of this. Another open question is weather generalized KS sets of Definition 3.2.7 that are not projective KS sets can be used to construct pseudo-telepathy games. Furthermore, can a wider class of pseudo-telepathy games be characterized using generalized KS sets than projective ones?

Finally, a fundamental open question. Determining the computational complexity of $\chi_q(G)$ and $\alpha_q(G)$ as a function of the number of vertices in G is now a long standing open question. Can we use the relationship with KS sets to answer this question?

Chapter acknowledgments I thank the coauthors of the works on which this chapter is based. Together, we are grateful to Jop Briët, André Chailloux, Will Matthews, Fernando de Melo, Teresa Piovesan, David Roberson, Dan Stahlke, Antonis Varvitsiotis, and Ronald de Wolf for helpful discussions. We also thank the referees of Asian Quantum Information Science 2012 for useful comments that improved the readability of some sections. We thank the Associate Editors of IEEE Transactions on Information theory and Journal of Mathematical Physics for valuable comments.

Chapter 4

Zero-error information theory

This chapter is based on the paper *"Zero-error source-channel coding with entanglement"*, by J. Briët, H. Buhrman, M. Laurent, T. Piovesan and the author.

The paper was presented at the Eurocomb conference in September 2013. An extended abstract was published in the conference proceedings.

4.1 Introduction

In this chapter, we study a problem from classical zero-error information theory: the *zero-error source-channel coding problem*, in the non-classical setting where a sender and receiver may use quantum entanglement. Viewed separately, the (dual) source coding problem asks a sender, Alice, to efficiently communicate data about which a receiver, Bob, already has some information, while the channel coding problem asks Alice to transmit data reliably in the presence of noise. In the combination of these two problems, Alice and Bob are each given an input from a random source and get access to a noisy channel through which Alice can send messages to Bob. Their goal is to minimize the average number of channel uses per source input such that Bob can learn Alice's inputs with zero probability of error.

Shannon's seminal paper [Sha56] on zero-error channel capacity kindled a large research area which involves not only information theorists but also researchers from combinatorics, computer science and mathematical programming (see for example Körner and Orlitsky [KO98] for an extensive survey and Lubetzky's PhD thesis [Lub07] for more recent results). In the zero-error regime, the optimal rates of source codes and channel codes are given by graph parameters known as the Witsenhausen rate and Shannon capacity, respectively. The Lovász theta number, which gives the best known efficiently-computable upper bound on the Shannon capacity, also upper bounds its entanglement-assisted counterpart. The line of research involving entanglement was started only recently by Cubitt *et* *al.* [CLMW10]. They show that the Shannon capacity can be increased if Alice and Bob may use entanglement.

Here we extend these results to source-coding problem and the more general source-channel coding problem. We prove a lower bound on the rate of entanglement-assisted source-codes in terms Szegedy's number (a strengthening of the theta number). This result implies that the theta number lower bounds the entangled variant of the Witsenhausen rate. We show that entanglement can allow for an unbounded improvement of the asymptotic rate of both classical source codes and classical source-channel codes. Our semidefinite programming bounds rely on a characterization of positive semidefinite matrices with a block form due to Gvozdenovic and Laurent. Our results use low-degree polynomials due to Barrington, Beigel and Rudich, Hadamard matrices due to Xia and Liu and a new application of the quantum teleportation scheme of Bennett *et al.*

The rest of the chapter is organized as follows. In Section 4.2 we explain the classical zero-error source-channel setting and we describe the two main tools we use later on: the class of "quarter orthogonality graphs" and the quantum teleportation scheme. In Section 4.3 we define the entanglement-assisted zero-error source-channel setting and prove some basic properties of the quantities involved. Then, in Section 4.4, we proceed with the first technical result: a lower bound on the entangled chromatic number in terms of the Szegedy's number. Sections 4.5-4.7 are dedicated to the entangled-classical separation results. We give some final comments and open questions in Section 4.8.

4.2 Preliminaries

4.2.1 Classical source-channel coding

In this section we describe the classical zero-error source, channel and source-channel coding problems. These are problems on graphs. Therefore, we refer back to Section 3.2.1 for an introduction on graph parameters, graph products and graph homomorphisms.

A *dual source* $\mathcal{M} = (\mathsf{X}, \mathsf{U}, P)$ consists of a finite set X, a (possibly infinite) set U and a probability distribution P over $\mathsf{X} \times \mathsf{U}$. In a dual-source instance, Alice is given an input $x \in \mathsf{X}$ and Bob an input $u \in \mathsf{U}$ with probability $P(x, u)$. Bob's input may already give him some information about Alice's. But if his input does not uniquely identify hers, she has to supply additional information for him to learn it exactly. For this they get access to a noiseless one-way binary channel which they aim to use as little as possible.[1] Here we consider only *memoryless* sources, which means that the probability distribution $P(x, u)$ of the source is unchanged after every instance.

[1] From now on we will assume that all binary channels are noiseless.

The source-coding problem can sometimes be solved more efficiently by jointly encoding sequences of inputs into single codewords. If the parties use *block codes* of length-n to deal with length-m input sequences, then after receiving an input sequence $\mathbf{x} = (x_1, \ldots, x_m)$, Alice applies encoding function $\mathsf{C} : \mathsf{X}^m \to \{0,1\}^n$ and sends $\mathsf{C}(\mathbf{x})$ through the binary channel by using it n times in a row. Bob, who received an input $\mathbf{u} = (u_1, \ldots, u_m) \in \mathsf{U}^m$, then applies a decoding function $\mathsf{D} : \mathsf{U}^m \times \{0,1\}^n \to \mathsf{X}^m$ to the pair $(\mathbf{u}, \mathsf{C}(\mathbf{x}))$ to get a string in X^m. The scheme works if Bob always gets the string $\mathbf{x}$. The *cost rate* of the scheme (C, D) is then n/m, which counts the average number of channel uses per source-input symbol.

Witsenhausen [Wit76] and Ferguson and Bailey [FB75] showed that the zero-error source coding problem can be studied in graph-theoretic terms. Associated with a dual source $\mathcal{M} = (\mathsf{X}, \mathsf{U}, P)$ is its *characteristic graph* $G = (\mathsf{X}, E)$, where $\{x, y\} \in E$ if there exists a $u \in \mathsf{U}$ such that $P(x, u) > 0$ and $P(y, u) > 0$. As such, the edge set identifies the pairs of inputs for Alice which Bob may not be able to distinguish based on his input. It is not difficult to see that every graph is the characteristic graph of a (non-unique) source. Solving a single instance of the zero-error source coding problem for $\mathcal{M}$ is equivalent to finding a proper coloring of G. Indeed, Bob's input u reduces the list of Alice's possible inputs to the set $\{x \in \mathsf{X} : P(u|x) > 0\}$ and this set forms a clique in G. So Bob can learn Alice's input if she sends him its color. Conversely, a length-1 block-code for $\mathcal{M}$ defines a proper coloring of G. To deal with length-m input sequences we consider the graph $G^{\boxtimes m}$ (the strong product of m copies of G), whose vertex set is X^m and where two distinct vertices $\mathbf{x} = (x_1, \ldots, x_m)$ and $\mathbf{y} = (y_1, \ldots, y_m)$ are adjacent in $G^{\boxtimes m}$ if, for every $i \in [m]$, either $x_i = y_i$ or $\{x_i, y_i\} \in E(G)$. The edges in $G^{\boxtimes m}$ are precisely the pairs of input sequences on Alice's side which Bob cannot distinguish. The *Witsenhausen rate*

$$R(G) = \lim_{m \to \infty} \frac{1}{m} \log \chi(G^{\boxtimes m}) \tag{4.1}$$

is the minimum asymptotic cost rate of a zero-error code for a source. The chromatic number is sub-multiplicative, *i.e.*, $\chi(G^{\boxtimes(m+m')}) \leq \chi(G^{\boxtimes m})\chi(G^{\boxtimes m'})$. Therefore Fekete's lemma implies that [2] the above limit exists and is equal to the infimum, $R(G) = \inf_m \log \chi(G^{\boxtimes m})/m$.

A *discrete channel* $\mathcal{N} = (\mathsf{S}, \mathsf{V}, Q)$ consists of a finite input set S, a (possibly infinite) output set V and a probability distribution $Q(\cdot|s)$ over V for each $s \in \mathsf{S}$. Throughout the paper we consider only memoryless channels. If Alice sends an input $s \in \mathsf{S}$ through the channel, then Bob receives the output $v \in \mathsf{V}$ with probability $Q(v|s)$. Their goal is to transmit a binary string $\mathbf{y}$ of, say, m bits from Alice to Bob while using the channel as little as possible. If the parties use a block

[2] Consider a sequence $(a_m)_{m\in\mathbb{N}}$ which is sub-additive: $a_{m+m'} \leq a_m + a_{m'}$ for all $m, m' \in \mathbb{N}$. Fekete's lemma says that the limit of the sequence $(a_m/m)_{m\in\mathbb{N}}$ exists and $\lim_{m\to\infty} a_m/m = \inf_{m\in\mathbb{N}} a_m/m$.

code of length n, then Alice has an encoding function $\mathsf{C} : \{0,1\}^m \to \mathsf{S}^n$ and sends $\mathsf{C}(\mathbf{y})$ through the channel by using it n times in sequence. Bob then receives an output sequence $\mathbf{v} = (v_1, \dots, v_n)$ on his side of the channel and applies a decoding function $\mathsf{D} : \mathsf{V}^n \to \{0,1\}^m$. The coding scheme (C, D) works if $\mathsf{D}(\mathbf{v}) = \mathbf{y}$. The *communication rate* of the scheme is m/n, the number of bits transmitted per channel use.

Shannon [Sha56] showed that the zero-error channel coding problem can be studied in graph-theoretic terms. Associated to a channel $\mathcal{N} = (\mathsf{S}, \mathsf{V}, Q)$ is its *confusability graph* $H = (\mathsf{S}, F)$ where $\{s,t\} \in F$ if there exists a $v \in \mathsf{V}$ such that both $Q(v|s) > 0$ and $Q(v|t) > 0$. The edge set identifies pairs of inputs which can lead to identical channel outputs on Bob's side. Sets of non-confusable inputs thus correspond to independent sets in H. Therefore, by identifying a maximal independent set in the confusability graph, the parties can send one out of $\alpha(H)$ messages with a single use of the channel. Conversely, any strategy that allows parties to perfectly communicate one out of t messages with a single use of the channel can be used to find and independent set of H. Shannon proved that the graph $H^{\boxtimes n}$ represents n uses of the channel. Then, codes of block-length n allow the zero-error transmission of $\alpha(H^{\boxtimes n})$ distinct messages. The *Shannon capacity*

$$c(H) = \lim_{n \to \infty} \frac{1}{n} \log \alpha(H^{\boxtimes n}) \tag{4.2}$$

is the maximum communication rate of a zero-error coding scheme. Similarly to the Witsenhausen rate, we can replace the above limit with the supremum: $c(H) = \sup_n \log \alpha(H^{\boxtimes n})/n$.

Now we combine the two settings above. In the *source-channel coding problem* the parties receive inputs from a dual source $\mathcal{M} = (\mathsf{X}, \mathsf{U}, P)$ and get access to a channel $\mathcal{N} = (\mathsf{S}, \mathsf{V}, Q)$. Their goal is to solve the source coding problem, but now using the channel $\mathcal{N}$ instead of a binary channel. An (m,n)*-coding scheme* for this problem consists of an encoding function $\mathsf{C} : \mathsf{X}^m \to \mathsf{S}^n$ and a decoding function $\mathsf{D} : \mathsf{U}^m \times \mathsf{V}^n \to \mathsf{X}^m$ (see Figure 4.1). The *cost rate* is n/m.

Nayak, Tuncel and Rose [NTR06] showed that if $\mathcal{M}$ has characteristic graph G and $\mathcal{N}$ has confusability graph H, then a zero-error (m,n)-coding scheme is equivalent to a homomorphism from $G^{\boxtimes m}$ to $\overline{H^{\boxtimes n}}$.[3] Then, for G and $\overline{H}$ containing at least one edge, the parameter

$$\eta(G,H) := \lim_{m \to \infty} \frac{1}{m} \min\left\{ n \in \mathbb{N} : G^{\boxtimes m} \longrightarrow \overline{H^{\boxtimes n}} \right\} \tag{4.3}$$

gives the minimum asymptotic cost rate of a zero-error code. To see that the limit exists, observe that the parameter

$$\eta_m(G,H) := \min\left\{ n \in \mathbb{N} : G^{\boxtimes m} \longrightarrow \overline{H^{\boxtimes n}} \right\}$$

[3] An intuition for this is that in the source-channel problem, parties need to map *confusable* pairs of source inputs to *non-confusable* pairs of channel inputs, and that is what an homomorphism from G to the complement of H does.

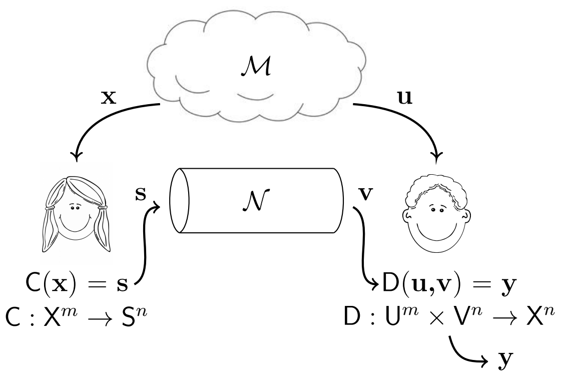


Figure 4.1: The figure illustrates a classical source-coding instance where Alice and Bob use an (m, n)-coding scheme (C, D). The parties receive length-m input strings $\mathbf{x}$ and $\mathbf{u}$, respectively, from a dual source $\mathcal{M} = (\mathsf{X}, \mathsf{U}, P)$ and have a one-way channel $\mathcal{N} = (\mathsf{S}, \mathsf{V}, Q)$. Using C, Alice encodes her input into a string $\mathbf{s} \in \mathsf{S}^n$ which she sends through the channel. After receiving a channel output $\mathbf{v}$, Bob applies D to the pair $(\mathbf{u}, \mathbf{v})$ to get a string $\mathbf{y}$. The scheme works if $\mathbf{y} = \mathbf{x}$.

is sub-additive and apply Fekete's lemma, which shows that $\eta(G, H) = \lim_{m\to\infty} \eta_m(G, H)/m$ is also equal to the infimum $\inf_m \eta_m(G, H)/m$.

If the channel $\mathcal{N}$ is replaced by a binary channel we regain the source coding problem. Conversely, if Alice receives binary inputs from the source and Bob's source inputs give him no information about Alice's at all, then we regain the channel coding problem. More formally, we can reformulate $R(G)$ and $c(H)$ in the following way.

Lemma 4.2.1. *Let G and H be graphs such that both G and $\overline{H}$ have at least one edge. Then,*

$$R(G) = \eta(G, \overline{K}_2) \quad \textit{and} \quad 1/c(H) = \eta(K_2, H).$$

Proof. For the proof of the identity $R(G) = \eta(G, \overline{K}_2)$ we use the following fact: for a graph G' and $t \in \mathcal{N}$, there exists a homomorphism from G' to K_t if and only if $\chi(G') \leq t$, which implies

$$\log \chi(G') \leq \min\{n : \, G' \longrightarrow K_{2^n}\} < \log \chi(G') + 1.$$

Combining these inequalities applied to $G' = G^{\boxtimes m}$ with the identity $\overline{\overline{K}_2^{\boxtimes n}} = K_{2^n}$, we obtain

$$\begin{aligned} \eta(G, \overline{K}_2) &= \lim_{m\to\infty} \frac{1}{m} \min\{n : \, G^{\boxtimes m} \longrightarrow \overline{\overline{K}_2^{\boxtimes n}} = K_{2^n}\} \\ &= \lim_{m\to\infty} \frac{1}{m} \log \chi(G^{\boxtimes m}) \\ &= R(G). \end{aligned}$$

The proof of the identity $1/c(H) = \eta(K_2, H)$ uses the fact that, for a graph H' and $t \in \mathcal{N}$, there exists a homomorphism from K_t to $\overline{H'}$ if and only if $\alpha(H') \geq t$. Since $K_2^{\boxtimes m} = K_{2^m}$, we get

$$\begin{aligned} \eta_m(K_2, H) &= \min\left\{n : \, K_2^{\boxtimes m} = K_{2^m} \longrightarrow \overline{H^{\boxtimes n}}\right\} \\ &= \min\left\{n : \, \alpha(H^{\boxtimes n}) \geq 2^m\right\} \\ &= \min\left\{n : \, \log \alpha(H^{\boxtimes n}) \geq m\right\}. \end{aligned}$$

Setting $n(m) := \eta_m(K_2, H)$, this implies

$$\log \alpha(H^{\boxtimes (n(m)-1)}) < m \leq \log \alpha(H^{\boxtimes n(m)})$$

and thus

$$\frac{n(m)}{\log \alpha(H^{\boxtimes n(m)})} \leq \frac{n(m)}{m} < \frac{n(m)}{\log \alpha(H^{\boxtimes (n(m)-1)})}. \tag{4.4}$$

As $c(H) = \sup_n \log \alpha(H^{\boxtimes n})/n$, using the left most inequality in (4.4) we deduce

$$\frac{1}{c(H)} \leq \frac{n(m)}{\log \alpha(H^{\boxtimes n(m)})} \leq \frac{n(m)}{m}$$

for all m. Taking the limit, we obtain the inequality $1/c(H) \leq \lim_{m\to\infty} n(m)/m = \eta_m(K_2, H)$. Next, as $\eta_m(K_2, H) = \inf_m n(m)/m$, using the right most inequality in (4.4) we deduce that

$$\eta_m(K_2, H) \leq \frac{n(m)}{m} < \frac{n(m)}{\log \alpha(H^{\boxtimes (n(m)-1)})} = \frac{n(m)-1}{\log \alpha(H^{\boxtimes (n(m)-1)})} \frac{n(m)}{n(m)-1}.$$

It is clear that $\lim_{m\to\infty} n(m) = \infty$. Therefore we can conclude that the limit of the right most term in the above inequalities is equal to $1/c(H)$. This shows the reverse inequality $\eta(K_2, H) \leq 1/c(H)$ and thus $\eta(K_2, H) = 1/c(H)$. □

Source and channel coding are often treated separately (as such, they motivate the two main branches of Shannon theory). The main reason for this are so-called *separation theorems*, which roughly say that source and channel code design can be separated without asymptotic loss in the code rate in the limit of large block lengths. Such results typically hold in a setting of asymptotically vanishing error probability [VVS95]. But when no errors can be tolerated at all, Nayak, Tuncel and Rose [NTR06] showed that separated codes can be highly suboptimal. In terms of the above graph parameters, this says that in general the inequality $\eta(G, H) \leq R(G)/c(H)$ holds (see Proposition 4.3.8), but that for some families of graphs there can be a large separation: $\eta(G, H) \ll R(G)/c(H)$.

4.2.2 Quarter-orthogonality graphs

To show separations between the classical and entangled variants of the above-mentioned parameters in Sections 4.5-4.7, we will use the following family of graphs (also considered in [BBG12] for similar reasons).

Definition 4.2.2 (Quarter-orthogonality graph H_k)**.** *For an odd positive integer k, the* quarter-orthogonality graph *H_k has as vertex set all vectors in $\{-1,1\}^k$ that have an even number of "-1" entries, and as edge set the pairs with inner product -1. Equivalently, the vertices of H_k are the k-bit binary strings with even Hamming weight and its edges are the pairs with Hamming distance $(k+1)/2$.*

We first give some intuition about the structure of these graphs, explain why we call them quarter-orthogonality graph and state some useful properties. The usual *orthogonality graph* has vertex set $\{-1,1\}^k$ and two vertices are adjacent if they are orthogonal. (This is the class of graphs we used in Section 3.4.2.) The quarter-orthogonality graph is a subgraph of the orthogonality graph. To see this, consider the map $\phi : \{-1,1\}^k \to \{-1,1\}^{k+1}$ that sends every vector u to $\phi(u) = (u^{\mathsf{T}}, 1)^{\mathsf{T}}$ (*i.e.*, the vector u with a "1" appended to it). This map embeds the graph H_k in the usual orthogonality graph (on 2^{k+1} vertices) since $\phi(u)^T\phi(v) = -1+1 = 0$ for every $\{u,v\}$ edge in H_k. Since H_k has 2^{k-1} vertices it is a subgraph of size a quarter of $\{-1,1\}^k$. We later use the following map, which sends vertices of H_k to the unit sphere in $\mathbb{R}^{k+1}$ and adjacent vertices to orthogonal vectors:

$$\begin{array}{rcccc} f & : & V(H_k) & \longrightarrow & \mathbb{R}^{k+1} \\ & & u & \longmapsto & \phi(u)/\sqrt{k+1}, \end{array} \tag{4.5}$$

Lemma 4.2.3. *For every k odd positive integer, we have $\alpha(H_k) \geq 2^{(k-3)/2}$.*

Proof. The lemma follows by considering the subset W of all the vectors in $V(H_k)$ (in the $\{0,1\}^k$ setting) that have zeros in their last $(k+1)/2$ coordinates. One can see that $|W| = 2^{(k-3)/2}$ and that W is an independent set since it does not contain pairs of strings at Hamming distance $(k+1)/2$. □

Some of our results rely on the existence of certain Hadamard matrices. A *Hadamard matrix* is a square matrix $A \in \{-1,1\}^{\ell\times\ell}$ that satisfies $AA^{\mathsf{T}} = \ell I$. The size ℓ of a Hadamard matrix must necessarily be 2 or a multiple of 4 and the famous Hadamard conjecture states that for every ℓ that is a multiple of 4 there exists an $\ell \times \ell$ Hadamard matrix. This conjecture is usually attributed to Paley [Pal33], who wrote:

> *"It seems probable that, whenever m is divisible by 4, it is possible to construct an orthogonal matrix of order m composed of ± 1, but the general theorem has every appearance of difficulty."*

Indeed, the conjecture has remained unproved despite sustained efforts. However, many infinite families of Hadamard matrices are known. We will use a family constructed by Xia and Liu [XL91] (see for example [XL96, WX97, Che97, Xia98, XSX06] for closely related constructions).

Theorem 4.2.4 (Xia and Liu [XL91]). *Let q be a prime power such that $q \equiv 1 \bmod 4$. Then, there exists a Hadamard matrix of size $4q^2$.*

We also use the following result regarding the graph H_k.

Proposition 4.2.5 (Briët, Buhrman and Gijswijt [BBG12]). *Let $k > 0$ be an integer such that there exists a Hadamard matrix of size $k+1$. Then, $\omega(H_k) \geq k+1$.*

Proof. Let A be a Hadamard matrix of size $k+1$. Without loss of generality the first row and column of A contain only "1" entries. Consider the submatrix A' of A where we remove the first column. Then by the orthogonality of the rows of A, each row of A' (but the first one) has $(k+1)/2$ entries with value "-1" and each pair of rows have inner product equal to -1. Since $k+1$ is a multiple of 4, the rows of A' form a clique in H_k. □

4.2.3 Quantum teleportation

Next we briefly explain the quantum teleportation scheme of Bennett *et al.* [BBC$^+$93]. The scheme allows Alice and Bob to transport a d-dimensional state from Alice to Bob by using only one-way classical communication and local operations on a pre-shared entangled state. It will be the crucial tool for obtaining the lower bound in Section 4.6.

The essential features of this scheme are as follows (we refer to [BBC$^+$93] and [NC00, pp. 26–28] for the details). Suppose that Alice has a local d-dimensional quantum system $\mathcal{A}$ in state ρ. Suppose in addition that Alice and Bob have local d-dimensional systems $\mathcal{X}$ and $\mathcal{Y}$, respectively. For this set-up, it follows from the basic quantum teleportation scheme of [BBC$^+$93] that there exist:

(QT1) a state σ of the pair $(\mathcal{X}, \mathcal{Y})$ (known as the *maximally entangled state*),

(QT2) a measurement $\mathsf{M} = \{M_i \in \mathbb{C}^{d\times d} \otimes \mathbb{C}^{d\times d} : i \in [d^2]\}$ (which is independent of ρ) and

(QT3) for every $i \in [d^2]$, a unitary operation $U_i \in \mathbb{C}^{d\times d}$

with which Alice and Bob can transfer ("teleport") the state ρ of Alice's system $\mathcal{A}$ to Bob's system $\mathcal{Y}$. To achieve this, the parties may follow the following protocol:

1. Alice performs the measurement M on the system $(\mathcal{A}, \mathcal{X})$ and gets some measurement outcome $i \in [d^2]$ with probability $\mathrm{Tr}[(M_i \otimes I)(\rho \otimes \sigma)]$;
2. Alice communicates her measurement outcome i to Bob;

3. Bob applies the unitary operation U_i to his system $\mathcal{Y}$.

That is, at the end of the protocol

(QT4) Bob's system $\mathcal{Y}$ is in state proportional to $U_i \mathrm{Tr}_{\mathcal{A},\mathcal{X}}\big((M_i \otimes I)(\rho \otimes \sigma)\big)U_i^\dagger$, which is guaranteed to be equal to ρ.

4.3 Source-channel coding with entanglement

We now explain the model of entanglement-assisted source-channel coding, also pictured by Figure 4.2. In the next sections, we derive algebraic definitions of entangled graph parameters and prove some of their basic properties.

In the entanglement-assisted source-channel coding, Alice and Bob receive inputs from a dual source $\mathcal{M} = (\mathsf{X}, \mathsf{U}, P)$ and Alice can send messages through a classical channel $\mathcal{N} = (\mathsf{S}, \mathsf{V}, Q)$. Their goal is for Bob to learn Alice's input, minimizing the number of channel uses per input sequence of given length. In addition Alice and Bob each have a local quantum system $\mathcal{A}$ and $\mathcal{B}$, respectively, and share an entangled state σ in $(\mathcal{A}, \mathcal{B})$ on which they can perform measurements. The entanglement-assisted source-channel coding protocol goes as follows:

1. Alice and Bob receive inputs $\mathbf{x} \in \mathsf{X}^m$ and $\mathbf{u} \in \mathsf{U}^m$, respectively, from the dual source $\mathcal{M}$;

2. Alice performs a measurement $\{A^{\mathbf{x}}_{\mathbf{s}}\}_{\mathbf{s}\in\mathsf{S}^n}$ (which can depend on $\mathbf{x}$) on $\mathcal{A}$ and gets $\mathbf{s}$ as outcome;

3. Alice sends $\mathbf{s}$ through the channel $\mathcal{N}$ and Bob receives $\mathbf{v} \in \mathsf{V}^n$;

4. Bob performs a measurement $\{B^{\mathbf{u},\mathbf{v}}_{\mathbf{y}}\}_{\mathbf{y}\in\mathsf{X}^m}$ (which can depend on $\mathbf{u}$ and $\mathbf{v}$) on $\mathcal{B}$ and gets $\mathbf{y} \in \mathsf{X}^m$ as outcome.

Recall that if the two parties share no entanglement, then a zero-error (m, n)-coding scheme is equivalent to a homomorphism from $V(G^{\boxtimes m})$ to $V(\overline{H^{\boxtimes n}})$, *i.e.*, a map that sends edges of $G^{\boxtimes m}$ to non-edges of $H^{\boxtimes n}$, where G is the characteristic graph of $\mathcal{M}$ and H is the confusability graph of $\mathcal{N}$. Analogously, the entanglement-assisted protocol is successful if and only if, for every edge $\{\mathbf{x}, \mathbf{y}\}$ in $G^{\boxtimes m}$ and every non-edge $\{\mathbf{s}, \mathbf{t}\}$ in $H^{\boxtimes n}$, we have that $\mathrm{Tr}_{\mathcal{A}}\big((A^{\mathbf{x}}_{\mathbf{s}} \otimes I)\sigma\big)$ is orthogonal to $\mathrm{Tr}_{\mathcal{A}}\big((A^{\mathbf{y}}_{\mathbf{t}} \otimes I)\sigma\big)$. The intuition is that indistinguishable pairs of Alice's inputs must be related to channel inputs that will not create confusion in Bob's measurement, thus allowing him to output correctly. We will see in the next section how this requirement gives rise to the algebraic definition of the entangled source-channel rate.

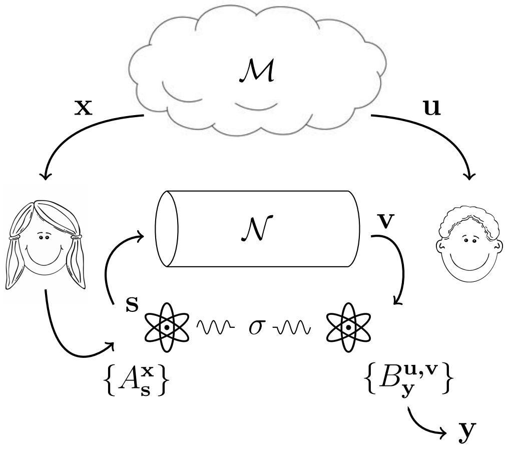


Figure 4.2: The figure illustrates the entanglement-assisted source-channel coding protocol. After receiving a source-input $\mathbf{x} \in \mathsf{X}^m$, Alice performs a measurement $\{A_{\mathbf{s}}^{\mathbf{x}} : \mathbf{s} \in \mathsf{S}^n\}$ on her part of an entangled state σ which she shares with Bob. She sends her measurement outcome $\mathbf{s}$ through the channel, upon which Bob—who previously already received a source-input $\mathbf{u}$—receives a channel-output $\mathbf{v} \in \mathsf{V}^n$. Bob performs a measurement $\{B_{\mathbf{y}}^{\mathbf{u},\mathbf{v}} : \mathbf{y} \in \mathsf{X}^m\}$ on his part of σ and obtains outcome $\mathbf{y} \in \mathsf{X}^m$.

4.3.1 Entangled source-channel rate, Witsenhausen rate and Shannon capacity

In graph-theoretic terms this model gives the following algebraic definition of the entangled variant of $\eta(G,H)$. It can be derived by considering the protocol described in the previous section and putting $\rho_{\mathbf{s}}^{\mathbf{x}} = \mathrm{Tr}_{\mathcal{A}}\big((A_{\mathbf{s}}^{\mathbf{x}} \otimes I)\sigma\big)$ and $\rho = \mathrm{Tr}_{\mathcal{A}}(\sigma)$. Conversely, given a solution to the algebraic definition, it is possible to recover the entangled state and Alice's measurements required in the protocol in a standard way.

Definition 4.3.1 (Entangled cost rate)**.** *For graphs G, H and $m \in \mathbb{N}$, define $\eta_m^\star(G,H)$ as the minimum integer $n \in \mathbb{N}$ for which there exist a $d \in \mathbb{N}$ and $d \times d$ positive semidefinite matrices ρ and $\{\rho_{\mathbf{s}}^{\mathbf{x}} : \mathbf{x} \in V(G^{\boxtimes m}), \mathbf{s} \in V(H^{\boxtimes n})\}$ such that* $\mathrm{Tr}(\rho) = 1$ *and*

$$\begin{aligned} \rho_{\mathbf{s}}^{\mathbf{x}} \rho_{\mathbf{t}}^{\mathbf{y}} &= 0 \quad \forall \mathbf{x},\mathbf{y},\mathbf{s},\mathbf{t} : \{\mathbf{x},\mathbf{y}\} \in E(G^{\boxtimes m}),\ \mathbf{s} = \mathbf{t} \textit{ or } \{\mathbf{s},\mathbf{t}\} \in E(H^{\boxtimes n}) \\ \sum_{\mathbf{s} \in V(H^{\boxtimes n})} \rho_{\mathbf{s}}^{\mathbf{x}} &= \rho \quad \forall \mathbf{x} \in V(G^{\boxtimes m}). \end{aligned}$$

The entangled cost rate is defined by

$$\eta^\star(G,H) = \lim_{m\to\infty} \frac{1}{m} \eta_m^\star(G,H).$$

As for the classical counterpart, we assume throughout that both graphs G and $\bar{H}$ contain at least one edge. We regain the parameter $\eta(G,H)$ if we restrict the above matrices ρ and $\rho_{\mathbf{s}}^{\mathbf{x}}$ to be chosen among $|0\rangle\langle 0|$ and $|1\rangle\langle 1|$. Thus sharing an entangled quantum system cannot make the coding scheme worse and $\eta^\star(G,H) \leq \eta(G,H)$. As in the classical case, the parameter $\eta_m^\star(G,H)$ is subadditive (see Lemma 4.3.5), hence the parameter $\eta^\star(G,H)$ is well defined and can be equivalently written as the infimum of $\eta_m^\star(G,H)/m$.

Similarly we also define an entangled variant of the chromatic and independence number.

Definition 4.3.2 (Entangled chromatic number). *For a graph G define $\chi^\star(G)$ as the minimum integer $t \in \mathbb{N}$ for which there exist a $d \in \mathbb{N}$ and $d \times d$ positive semidefinite matrices ρ and $\{\rho_i^u : u \in V(G),\, i \in [t]\}$ such that* $\mathrm{Tr}(\rho) = 1$ *and*

$$\begin{aligned} \rho_i^u \rho_i^v &= 0 \quad \forall i,u,v : i \in [t],\ \{u,v\} \in E(G) \\ \sum_{i\in[t]} \rho_i^u &= \rho \quad \forall u \in V(G). \end{aligned}$$

The entangled Witsenhausen rate is defined by

$$R^\star(G) = \lim_{m\to\infty} \frac{1}{m} \log \chi^\star(G^{\boxtimes m}).$$

In Lemma 4.3.6 we show that $\chi^\star$ is sub-multiplicative and thus the entangled Witsenhausen rate can be equivalently defined as the infimum: $R^\star(G) = \inf_m \log \chi^\star(G^{\boxtimes m})/m$.

Definition 4.3.3 (Entangled independence number). *For a graph H define $\alpha^\star(H)$ as the maximum integer $M \in \mathbb{N}$ for which there exist $d \in \mathbb{N}$ and $d \times d$ positive semidefinite matrices ρ and $\{\rho_u^i : i \in [M],\, u \in V(H)\}$ such that* $\mathrm{Tr}(\rho) = 1$ *and*

$$\begin{aligned} \rho_u^i \rho_v^j &= 0 \quad \forall i,j,u,v : i \neq j,\ u = v \text{ or } \{u,v\} \in E(H) \\ \sum_{u\in V(H)} \rho_u^i &= \rho \quad \forall i \in [M]. \end{aligned}$$

The entangled Shannon capacity is defined by

$$c^\star(H) = \lim_{n\to\infty} \frac{1}{n} \log \alpha^\star(H^{\boxtimes n}).$$

The parameter $\alpha^\star(H)$ was introduced by Cubitt *et al.* [CLMW10] and it is known to be super-multiplicative. Hence, in the definition of $c^\star(H)$ the limit can be replaced with the supremum.

Analogous to the classical setting, we can reformulate the entangled variants of the Witsenhausen rate and Shannon capacity as follows.

Lemma 4.3.4. *Let G and H be graphs such that both G and $\overline{H}$ have at least one edge. Then,*

$$R^\star(G) = \eta^\star(G, \overline{K}_2) \quad \textit{and} \quad 1/c^\star(H) = \eta^\star(K_2, H).$$

Proof. Since the graph $\overline{K}_2^{\boxtimes n}$ has 2^n vertices and no edges, it follows from the definitions that $\eta^\star_m(G, \overline{K}_2) = \lceil \log \chi^\star(G^{\boxtimes m}) \rceil$. The identity $R^\star(G) = \eta^\star(G, \overline{K}_2)$ follows by dividing by m and letting m go to infinity.

Since $K_2^{\boxtimes m} = K_{2^m}$, it follows from the definitions that $\eta^\star_m(K_2, H)$ is the minimum $n \in \mathcal{N}$ such that $\alpha^\star(H^{\boxtimes n}) \geq 2^m$ or, equivalently, $\log \alpha^\star(H^{\boxtimes n}) \geq m$. Using the same techniques as in Lemma 4.2.1, we have that $1/c^\star(H) = \eta^\star(K_2, H)$. □

In [CLMW10] it is shown that $\alpha^\star(H)$ can be strictly larger than $\alpha(H)$, meaning that the number of messages that can be sent with a single use of a channel can be increased with the use of entanglement (see also Mančinska, Severini and the author [MSS13]). This result was subsequently strengthened by Leung, Mančinska, Matthews, Ozols and Roy [LMM+12] and Briët, Buhrman and Gijswijt [BBG12], who found families of graphs for which $c^\star(H) > c(H)$.

To the best of our knowledge, neither source nor source-channel coding were considered in the context of shared entanglement before. However, in the context of Bell inequalities, Cameron *et al.* [CMN+07] studied the *quantum chromatic number* $\chi_q(G)$, and Roberson and Mančinska [RM12] considered a variant of the quantum independence number $\alpha_q(H)$. These are the parameters that we studied in Chapter 3. They can be obtained from the respective definitions of $\chi^\star$ and $\alpha^\star$ given above, if we set ρ to be proportional to the identity matrix and if we further restrict the other positive semidefinite matrices to be scalar multiples of orthogonal projections (matrices that satisfy $P^2 = P$). Furthermore, we regain χ and α if we restrict these matrices further still by requiring that they all be chosen among $|0\rangle\langle 0|$ and $|1\rangle\langle 1|$. It thus follows immediately that

$$\chi^\star(G) \leq \chi_q(G) \leq \chi(G) \quad \text{and} \quad \alpha(H) \leq \alpha_q(H) \leq \alpha^\star(H).$$

It is well-known that determining the classical chromatic and independence numbers of a graph are an NP-hard problems. The problem of determining the Shannon capacity and the Witsenhausen rate is not known to be computable.

Despite substantial efforts, the properties of these parameters are still only partially understood (see [Alo02, AL06] and references therein). For example, the largest odd cycle for which the Shannon capacity has been determined is C_5 and the computability of the Shannon capacity and the Witsenhausen rate are still unknown. Clearly the parameter η is at least as hard to compute as R and c since it contains them as special cases. Even less is known about the quantum variants of these parameters and determining the computational complexity of the quantities $\chi^\star, \alpha^\star, \chi_q, \alpha_q, R^\star$ and $c^\star$ is an open problem.

4.3.2 Basic properties of the entangled parameters

We have already mentioned that the parameter $\eta_m^\star$ is sub-additive, $\chi^\star$ is sub-multiplicative and that a coding scheme for the source-channel problem can be solved by concatenating a coding scheme for a source with one for a channel. Here we prove these facts.

Sub-additivity of $\eta_m^\star$ and sub-multiplicativity of $\chi^\star$

Lemma 4.3.5. *Let G and H be graphs and assume that both G and $\overline{H}$ have at least one edge. For every $m, m' \in \mathcal{N}$, we have*

$$\eta^\star_{m+m'}(G,H) \leq \eta^\star_m(G,H) + \eta^\star_{m'}(G,H).$$

Proof. Let $\varphi, \{\varphi^{\mathbf{s}}_{\mathbf{x}} : \mathbf{x} \in V(G^{\boxtimes m}), \mathbf{s} \in V(H^{\boxtimes n})\}$ be a set of positive semidefinite matrices that witness $\eta^\star_m(G,H) = n$ (Definition 4.3.1) and let $\psi, \{\psi^{\mathbf{t}}_{\mathbf{y}} : \mathbf{y} \in V(G^{\boxtimes m'}), \mathbf{t} \in V(H^{\boxtimes n'})\}$ be a collection of matrices which are a solution for $\eta^\star_{m'}(G,H) = n'$. Notice that every vertex $\mathbf{w}$ of $G^{\boxtimes(m+m')}$ can be written as $\mathbf{w} = (\mathbf{x}, \mathbf{y})$ where $\mathbf{x} \in V(G^{\boxtimes m})$ and $\mathbf{y} \in V(G^{\boxtimes m'})$ and similarly any $\mathbf{r} \in V(H^{\boxtimes(n+n')})$, $\mathbf{r} = (\mathbf{s}, \mathbf{t})$ where $\mathbf{s} \in V(H^{\boxtimes n})$ and $\mathbf{t} \in V(H^{\boxtimes n'})$. We create a solution for $\eta^\star_{m+m'}(G,H)$ as follows. Let $\rho = \varphi \otimes \psi$ and for every vertex $(\mathbf{x}, \mathbf{y}) \in V(G^{\boxtimes(m+m')})$ and $(\mathbf{s}, \mathbf{t}) \in V(H^{\boxtimes(n+n')})$ define

$$\rho^{(\mathbf{s},\mathbf{t})}_{(\mathbf{x},\mathbf{y})} = \varphi^{\mathbf{s}}_{\mathbf{x}} \otimes \psi^{\mathbf{t}}_{\mathbf{y}}.$$

Then, for every $(\mathbf{x}, \mathbf{y}) \in V(G^{\boxtimes(m+m')})$, we have

$$\begin{aligned}\sum_{(\mathbf{s},\mathbf{t}) \in V(H^{\boxtimes(n+n')})} \rho^{(\mathbf{s},\mathbf{t})}_{(\mathbf{x},\mathbf{y})} &= \sum_{\mathbf{s} \in V(H^{\boxtimes n})} \sum_{\mathbf{t} \in V(H^{\boxtimes n'})} \varphi^{\mathbf{s}}_{\mathbf{x}} \otimes \psi^{\mathbf{t}}_{\mathbf{y}} \\ &= \left(\sum_{\mathbf{s} \in V(H^{\boxtimes n})} \varphi^{\mathbf{s}}_{\mathbf{x}} \right) \otimes \left(\sum_{\mathbf{t} \in V(H^{\boxtimes n'})} \psi^{\mathbf{t}}_{\mathbf{y}} \right) = \varphi \otimes \psi = \rho.\end{aligned}$$

Suppose $(\mathbf{x}, \mathbf{y})$ and $(\mathbf{x}', \mathbf{y}')$ are adjacent in $G^{\boxtimes(m+m')}$ and $(\mathbf{s}, \mathbf{t})$ and $(\mathbf{s}', \mathbf{t}')$ are either equal or adjacent in $H^{\boxtimes(n+n')}$. We have that

$$\rho^{(\mathbf{s},\mathbf{t})}_{(\mathbf{x},\mathbf{y})} \rho^{(\mathbf{s}',\mathbf{t}')}_{(\mathbf{x}',\mathbf{y}')} = \left(\varphi^{\mathbf{s}}_{\mathbf{x}} \otimes \psi^{\mathbf{t}}_{\mathbf{y}} \right) \left(\varphi^{\mathbf{s}'}_{\mathbf{x}'} \otimes \psi^{\mathbf{t}'}_{\mathbf{y}'} \right) = \left(\varphi^{\mathbf{s}}_{\mathbf{x}} \varphi^{\mathbf{s}'}_{\mathbf{x}'} \right) \otimes \left(\psi^{\mathbf{t}}_{\mathbf{y}} \psi^{\mathbf{t}'}_{\mathbf{y}'} \right) = 0.$$

Now since $\mathrm{Tr}(\rho) = \mathrm{Tr}(\varphi \otimes \psi) = 1$, it follows that the collection of positive semidefinite matrices ρ, $\{\rho^{(\mathbf{s},\mathbf{t})}_{(\mathbf{x},\mathbf{y})} : (\mathbf{x}, \mathbf{y}) \in V(G^{\boxtimes(m+m')}), (\mathbf{s}, \mathbf{t}) \in V(H^{\boxtimes(n+n')})\}$ is a solution for $\eta^\star_{m+m'}(G,H) \leq n + n' = \eta^\star_m(G,H) + \eta^\star_{m'}(G,H)$. □

Lemma 4.3.6. *For two graphs G and H, $\chi^\star(G \boxtimes H) \leq \chi^\star(G)\chi^\star(H)$.*

Proof. Let $\varphi, \{\varphi_i^u : u \in V(G), i \in [s]\}$ be a collection of positive semidefinite matrices that witness $\chi^\star(G) = s$ with $s \in \mathbb{N}$ and let $\psi, \{\psi_j^v : v \in V(H), j \in [t]\}$ be a set of matrices which are a solution for $\chi^\star(H) = t$, $t \in \mathbb{N}$. Let $\rho = \varphi \otimes \psi$ and for every vertex (u, v) in $G \boxtimes H$ and $\mathbf{k} = (i, j) \in [s] \times [t]$ define

$$\rho_{\mathbf{k}}^{(u,v)} = \varphi_u^i \otimes \psi_v^j.$$

Using similar techniques as in the previous proof, one can see that the set of matrices ρ, $\{\rho_{\mathbf{k}}^{(u,v)} : (u, v) \in V(G) \times V(H), \mathbf{k} \in [s] \times [t]\}$ is a feasible solution for $\chi^\star(G \boxtimes H) \leq |[s] \times [t]| = \chi^\star(G)\chi^\star(H)$. □

Separate coding schemes

By concatenating an entanglement-assisted coding scheme for a source with one for a channel, one obtains a coding scheme for the combined source-channel problem. For this to work, the number of bits one can send perfectly with n uses of the channel must be at least as large as the number of bits required to solve m instances of the source problem. In other words, for a source with characteristic graph G and a channel with confusability graph H, we need the condition $\chi^\star(G^{\boxtimes m}) \leq \alpha^\star(H^{\boxtimes n})$ in order to send length-m source-input sequences with n uses of the channel and shared entanglement. If this condition holds, then it follows that $\eta_m^\star(G, H) \leq n$. We now give a formal proof of this statement which we also prove for the classical case.

Lemma 4.3.7. *Given graphs G, H and positive integers n, m, we have*

$$\chi(G^{\boxtimes m}) \leq \alpha(H^{\boxtimes n}) \Longrightarrow \eta_m(G, H) \leq n, \tag{4.6}$$

$$\chi^\star(G^{\boxtimes m}) \leq \alpha^\star(H^{\boxtimes n}) \Longrightarrow \eta_m^\star(G, H) \leq n. \tag{4.7}$$

Proof. If $\chi(G^{\boxtimes m}) \leq \alpha(H^{\boxtimes n})$, then there is a homomorphism from $G^{\boxtimes m}$ to $\overline{H^{\boxtimes n}}$ and thus $\eta_m(G, H) \leq n$, which shows (4.6). We now show (4.7). For this set $t = \chi^\star(G^{\boxtimes m})$ and $M = \alpha^\star(H^{\boxtimes n})$, with $t \leq M$ by assumption. Let $\varphi, \{\varphi_i^{\mathbf{x}} : \mathbf{x} \in V(G^{\boxtimes m}), i \in [t]\}$ be a collection of positive semidefinite matrices forming a solution for $\chi^\star(G^{\boxtimes m})$ and let the set of positive semidefinite matrices $\psi, \{\psi_{\mathbf{s}}^i : \mathbf{s} \in V(H^{\boxtimes n}), i \in [M]\}$ be feasible for $\alpha^\star(H^{\boxtimes n})$. We construct a solution for $\eta_m^\star(G, H)$ as follows. For $\mathbf{x} \in V(G^{\boxtimes m})$ and $\mathbf{s} \in V(H^{\boxtimes n})$ set

$$\rho_{\mathbf{s}}^{\mathbf{x}} = \sum_{i \in [t]} \varphi_i^{\mathbf{x}} \otimes \psi_{\mathbf{s}}^i \quad \text{and} \quad \rho = \varphi \otimes \psi.$$

Then, we have that $\mathrm{Tr}(\rho) = \mathrm{Tr}(\varphi \otimes \psi) = 1$ and, for every $\mathbf{x} \in V(G^{\boxtimes m})$, we get that $\sum_{\mathbf{s} \in V(H^{\boxtimes n})} \rho_{\mathbf{s}}^{\mathbf{x}} = \sum_{\mathbf{s} \in V(H^{\boxtimes n})} \sum_{i \in [t]} \varphi_i^{\mathbf{x}} \otimes \psi_{\mathbf{s}}^i = \sum_{i \in [t]} \varphi_i^{\mathbf{x}} \otimes (\sum_{\mathbf{s} \in V(H^{\boxtimes n})} \psi_{\mathbf{s}}^i)$ is equal to

$\varphi \otimes \psi = \rho$. Moreover, for every $\{\mathbf{x}, \mathbf{y}\} \in E(G^{\boxtimes m})$ and $\{\mathbf{s}, \mathbf{t}\} \in V(H^{\boxtimes n}) \cup E(H^{\boxtimes n})$,

$$\begin{aligned} \rho^{\mathbf{x}}_{\mathbf{s}} \rho^{\mathbf{y}}_{\mathbf{t}} &= \Big(\sum_{i \in [t]} \varphi^{\mathbf{x}}_i \otimes \psi^i_{\mathbf{s}}\Big)\Big(\sum_{j \in [t]} \varphi^{\mathbf{y}}_j \otimes \psi^j_{\mathbf{t}}\Big) = \sum_{i \in [t]} \sum_{j \in [t]} \varphi^{\mathbf{x}}_i \varphi^{\mathbf{y}}_j \otimes \psi^i_{\mathbf{s}} \psi^j_{\mathbf{t}} \\ &= \sum_{i \in [t]} \varphi^{\mathbf{x}}_i \varphi^{\mathbf{y}}_i \otimes \psi^i_{\mathbf{s}} \psi^i_{\mathbf{t}} + \sum_{i,j \in [t], i \neq j} \varphi^{\mathbf{x}}_i \varphi^{\mathbf{y}}_j \otimes \psi^i_{\mathbf{s}} \psi^j_{\mathbf{t}} = 0, \end{aligned}$$

where the last identity uses the orthogonality conditions of the matrices $\varphi^{\mathbf{x}}_i$ and $\psi^i_{\mathbf{s}}$. Hence $\rho, \{\rho^{\mathbf{x}}_{\mathbf{s}} : \mathbf{x} \in V(G^{\boxtimes m}), \mathbf{s} \in V(H^{\boxtimes n})\}$ is a feasible solution for $\eta^\star_m(G, H) \leq n$. □

We now relate the minimum cost rate to the ratio of the Witsenhausen rate and the Shannon capacity in both classical and entangled assisted cases.

Proposition 4.3.8. *Let G and H be graphs and assume that both G and $\overline{H}$ have at least one edge. Then,*

$$\eta(G, H) \leq \frac{R(G)}{c(H)} = \lim_{m \to \infty} \frac{1}{m} \min\{n : \chi(G^{\boxtimes m}) \leq \alpha(H^{\boxtimes n})\}, \tag{4.8}$$

$$\eta^\star(G, H) \leq \frac{R^\star(G)}{c^\star(H)} = \lim_{m \to \infty} \frac{1}{m} \min\{n : \chi^\star(G^{\boxtimes m}) \leq \alpha^\star(H^{\boxtimes n})\}. \tag{4.9}$$

Proof. We show (4.8); we omit the proof of (4.9) which is analogous (and uses (4.7)). From (4.6) we have the inequality:

$$\eta_m(G, H) \leq \epsilon_m(G, H) := \min\{n : \chi(G^{\boxtimes m}) \leq \alpha(H^{\boxtimes n})\},$$

which implies $\eta(G, H) \leq \lim_{m \to \infty} \epsilon_m(G, H)/m$. Next we show that this limit is equal to $R(G)/c(H)$, which concludes the proof of (4.8). Setting $n = \epsilon_m(G, H)$, we have that $\alpha(H^{\boxtimes n-1}) < \chi(G^{\boxtimes m}) \leq \alpha(H^{\boxtimes n})$, implying

$$\frac{R(G)}{c(H)} \leq \frac{\log \chi(G^{\boxtimes m})}{m} \frac{n}{\log \alpha(H^{\boxtimes n})} \leq \frac{n}{m} \leq \frac{n}{n-1} \frac{\log \chi(G^{\boxtimes m})}{m} \frac{n-1}{\log \alpha(H^{\boxtimes n-1})}.$$

Taking limits as $m \to \infty$ in the right-most terms we obtain that

$$\frac{R(G)}{c(H)} = \lim_{m \to \infty} \frac{\epsilon_m(G, H)}{m}.$$

□

We also record the following bound, which we use later.

Proposition 4.3.9. *For graphs G and H and positive integer m, we have*

$$\eta^\star_m(G, H) \leq \left\lceil \frac{\log \chi^\star(G^{\boxtimes m})}{\log \alpha^\star(H)} \right\rceil.$$

Proof. Set $n = \lceil \log \chi^\star(G^{\boxtimes m}) / \log \alpha^\star(H) \rceil$. Using the super-multiplicativity of $\alpha^\star(H)$ we get

$$\log \alpha^\star(H^{\boxtimes n}) \quad \geq \quad n \log \alpha^\star(H) = \left\lceil \frac{\log \chi^\star(G^{\boxtimes m})}{\log \alpha^\star(H)} \right\rceil \log \alpha^\star(H) \geq \log \chi^\star(G^{\boxtimes m}).$$

From Lemma 4.3.7 it then follows that $\eta^\star_m(G, H) \leq n$. □

4.4 Szegedy's number lower bound on the entangled chromatic number

Here we explain our lower bound on the entangled chromatic number. We show that $\chi^\star(G)$ is lower bounded by an efficiently computable graph parameter, namely a variant of the famous *theta number* introduced by Szegedy [Sze94].

Szegedy [Sze94] introduced the following strengthening of the theta number, which includes an extra linear constraint to the formulation (3.2).

$$\begin{aligned} \vartheta^+(G) = \min \Big\{ \lambda : \quad & \exists\, Z \in \mathbb{R}^{V(G) \times V(G)},\ Z \succeq 0, \\ & Z(u,u) = \lambda - 1 \ \text{ for } u \in V(G), \\ & Z(u,v) = -1 \ \text{ for } \{u,v\} \notin E(G), \\ & Z(u,v) \geq -1 \ \text{ for } \{u,v\} \in E(G) \Big\}. \end{aligned} \tag{4.10}$$

Szegedy's number satisfies $\vartheta(G) \leq \vartheta^+(G)$ and $\alpha(G) \leq \vartheta^+(G) \leq \chi(\overline{G})$. Recall that Lovász proved that ϑ is *multiplicative* under the strong graph product, that is, $\vartheta(G \boxtimes H) = \vartheta(G)\vartheta(H)$. Moreover Knuth [KD93] showed that $\vartheta(\overline{G \boxtimes H}) = \vartheta(\overline{G})\vartheta(\overline{H})$. It is unknown if the latter identity holds for ϑ^+ [Meu05]. The identities of Lovász and Knuth give for any graph G and $m \in \mathbb{N}$:

$$\vartheta(\overline{G}^{\boxtimes m}) = \vartheta(\overline{G^{\boxtimes m}}) = \vartheta(\overline{G})^m. \tag{4.11}$$

Combining these properties of ϑ with the Sandwich Theorem shows that

$$c(G) \leq \log \vartheta(G) \leq R(\overline{G}).$$

These inequalities capture the best known efficiently computable bounds for the Shannon capacity and the Witsenhausen rate.

Now we prove that the parameter ϑ^+ (and thus ϑ as well) lower bounds the entangled chromatic number and hence $\log \vartheta$ lower bounds the entangled Witsenhausen rate.

Theorem 4.4.1. *For any graph G, we have*

$$\vartheta^+(G) \leq \chi^\star(\overline{G}), \tag{4.12}$$

$$\log \vartheta(G) \leq R^\star(\overline{G}). \tag{4.13}$$

In [RM12] it is observed that $\vartheta(G) \leq \chi_q(\overline{G})$ holds. Theorem 4.4.1 thus strengthens this bound as it gives $\vartheta(G) \leq \vartheta^+(G) \leq \chi^\star(\overline{G}) \leq \chi_q(\overline{G})$.

Beigi [Bei10] and Duan, Severini and Winter [DSW13] proved that $\vartheta(G)$ upper bounds $\alpha^\star(G)$. The above-mentioned relations therefore imply the following sequence of inequalities.

$$c(G) \leq c^\star(G) \leq \log \vartheta(G) \leq R^\star(\overline{G}) \leq R(\overline{G}).$$

We will use the following result about positive semidefinite matrices with a special block form (which can be found, e.g., in [GL08]).

Lemma 4.4.2. *Let X be a $t \times t$ block matrix, with a matrix A as diagonal blocks and a matrix B as non-diagonal blocks, of the form*

$$X = \underbrace{\begin{pmatrix} A & B & \dots & B \\ B & A & \dots & B \\ \vdots & \vdots & \ddots & \vdots \\ B & B & \dots & A \end{pmatrix}}_{t \text{ blocks}}.$$

Then, $X \succeq 0$ if and only if $A - B \succeq 0$ and $A + (t-1)B \succeq 0$.

Proof of Theorem 4.4.1. We show that relations (4.12) and (4.13) hold for the graph $\overline{G}$. First we observe that (4.13) follows from (4.12). Indeed, relation (4.12) combined with the identity (4.11) implies $\vartheta(\overline{G})^m = \vartheta(\overline{G^{\boxtimes m}}) \leq \chi^\star(G^{\boxtimes m})$ and thus $\log \vartheta(\overline{G}) \leq R^\star(G)$ follows after taking limits.

We now prove (4.12) for the graph $\overline{G}$, i.e., we show the inequality $\vartheta^+(\overline{G}) \leq \chi^\star(G)$. For this let $\rho, \{\rho_u^i : u \in V(G), i \in [t]\}$ be a set of positive semidefinite matrices which form a solution for $\chi^\star(G) = t$. We may assume that $\langle \rho, \rho \rangle = 1$. Here, $\langle \cdot, \cdot \rangle$ is the trace inner product, defined by $\langle A, B \rangle = \mathrm{Tr}(A^* B)$ for matrices A, B of the same size. Define the matrix X, indexed by all pairs $\{u, i\} \in V(G) \times [t]$, with entries $X_{ui,vj} := \langle \rho_u^i, \rho_v^j \rangle$. By construction, X is a non-negative positive semidefinite matrix which satisfies $X_{ui,vi} = 0$ for every $\{u, v\} \in E(G)$ and $i \in [t]$.

For any element σ of $\mathrm{Sym}(t)$, the group of permutations of $[t]$, we define the new (permuted) matrix $\sigma(X) = (X_{u\sigma(i),v\sigma(j)})$. Then we average the matrix X over the group $\mathrm{Sym}(t)$, obtaining the new matrix

$$Y = \frac{1}{|\mathrm{Sym}(t)|} \sum_{\sigma \in \mathrm{Sym}(t)} \sigma(X).$$

By construction, the matrix Y is invariant under any permutation of $[t]$, *i.e.*, $\sigma(Y) = Y$ for any $\sigma \in \mathrm{Sym}(t)$. Therefore, Y has the block form of Lemma 4.4.2 with, moreover,

$$A_{uv} = 0 \text{ for all } \{u, v\} \in E(G). \tag{4.14}$$

As each matrix $\sigma(X)$ is positive semidefinite, the matrix Y is positive semidefinite as well. From Lemma 4.4.2, this implies that $A-B$ and $A+(t-1)B$ are positive semidefinite matrices. Using the definition of the matrix X combined with the properties of the matrices ρ_u^i and the invariance of Y, we obtain the following relation for any $u, v \in V(G)$:

$$1 = \langle \rho, \rho \rangle = \langle \sum_{i\in[t]} \rho_u^i, \sum_{j\in[t]} \rho_v^j \rangle = \sum_{i\in[t]} \sum_{j\in[t]} \langle \rho_u^i, \rho_v^j \rangle = \sum_{i\in[t]} \sum_{j\in[t]} X_{ui,vj} = \sum_{i\in[t]} \sum_{j\in[t]} Y_{ui,vj},$$

implying

$$1 = \sum_{i\in[t]} \sum_{j\in[t]} Y_{ui,vj} = t \sum_{j\in[t]} Y_{ui,vj} = t(A_{uv} + (t-1)B_{uv}). \tag{4.15}$$

We are now ready to define a matrix Z which is a feasible solution for the program (4.10) defining $\vartheta^+(G)$. Namely, set $Z = t(t-1)(A-B)$. Then, Z is a positive semidefinite matrix. For any edge $\{u, v\} \in E(G)$, the relations (4.14) and (4.15) give $A_{uv} = 0$ and $t(t-1)B_{uv} = 1$ and thus $Z_{uv} = -1$. For a non-edge $\{u, v\}$, relation (4.15) combined with the fact that $A_{uv} \geq 0$ implies that $Z_{uv} \geq -1$. Finally, for any $u \in V(G)$, relation (4.15) combined with the fact that $B_{uu} \geq 0$ implies that $Z_{uu} \leq t-1$. Define the vector c with entries $c_u = t-1-Z_{uu} \geq 0$ for $u \in V(G)$, the diagonal matrix $D(c)$ with c as diagonal, and the matrix $Z' = Z + D(c)$. Then, Z' is positive semidefinite and satisfies all the conditions of the program (4.10) defining $\vartheta^+(\overline{G})$. This shows that $\vartheta^+(\overline{G}) \leq \chi^\star(G)$, which concludes the proof. □

4.5 Separation between classical and entangled Witsenhausen rate

Our first separation result shows an exponential gap between the entangled and classical Witsenhausen rates of quarter-orthogonality graphs (Definition 4.2.2).

Theorem 4.5.1. *For every odd k integer, we have*

$$R^\star(H_k) \leq \log(k+1). \tag{4.16}$$

Moreover, if $k = 4p^\ell - 1$ where p is an odd prime and $\ell \in \mathbb{N}$, then

$$R(H_k) \geq 0.154k - 1. \tag{4.17}$$

In the following sections we show separately the upper bound and lower bound that together prove Theorem 4.5.1.

4.5.1 Upper bound on the entangled Witsenhausen rate

Here we prove the upper bound (4.16) stated in Theorem 4.5.1 on $R^\star(H_k)$. Recall that a *d-dimensional orthonormal representation* of a graph G is a map f from $V(G)$ to the unit sphere in $\mathbb{C}^d$, having the property that adjacent vertices are mapped to orthogonal vectors.[4] Also, recall that the *orthogonal rank* $\xi(G)$ of G is the minimum d such that there exists a d-dimensional orthonormal representation of G. Following [CMN+07] we define $\xi'(G)$ to be the minimum dimension d such that there exists a d-dimensional orthonormal representation f of G such that, for every vertex $u \in V(G)$, the d entries of the vector $f(u)$ all have absolute value $1/\sqrt{d}$.

The following bound on $\chi^\star(G)$ follows from the fact that $\chi^\star(G) \leq \chi_q(G)$ and a result proved in [CMN+07] stating that $\chi_q(G) \leq \xi'(G)$. We give a self-contained proof of the implied bound on $\chi^\star(G)$ for completeness.

Lemma 4.5.2. *For every graph G, we have $\chi^\star(G) \leq \xi'(G)$.*

Proof. Set $d = \xi'(G)$, $\omega_d = e^{2i\pi/d}$ and let $h_j = [\omega_d^j, \omega_d^{j+1}, \ldots, \omega_d^{j+d-1}]^\mathsf{T} \in \mathbb{C}^d$ for every $j \in [d]$. One can see that $\{h_1, h_2, \ldots, h_d\}$ is a complete orthogonal basis for $\mathbb{C}^d$. Set $\rho = I/d$. Then $\mathrm{Tr}(\rho) = 1$.

Let $f : V(G) \to \mathbb{C}^d$ be an orthonormal representation of G where each vector $f(u)$ is such that $(f(u)_i)^* f(u)_i = 1/d$ for every $i \in [d]$, as guaranteed to exist by the fact that $\xi'(G) = d$. For every $u \in V(G)$ and $i \in [d]$ define $\rho_i^u = |f(u) \circ h_i\rangle\langle f(u) \circ h_i|$, where $\circ$ denotes the entrywise product. We have

$$\langle f(u) \circ h_i, f(v) \circ h_j \rangle = \begin{cases} \langle h_i, h_j \rangle / d & \text{if } u = v \\ \langle f(u), f(v) \rangle & \text{if } i = j. \end{cases}$$

It follows that for every $u \in V(G)$ we have $\rho_1^u + \rho_2^u + \cdots \rho_d^u = I/d = \rho$. Moreover, for each $\{u, v\} \in E(G)$ and $i \in [d]$, we have $\rho_i^u \rho_i^v = 0$. As the matrices ρ, ρ_i^u are also positive semidefinite, they satisfy all the requirements of Definition 4.3.2 and so $\chi^\star(G) \leq d$. □

The above lemma gives a bound on the entangled chromatic number of powers of H_k from which it will be easy to get the upper bound on $R(H_k)$ given in (4.16).

Lemma 4.5.3. *Let k be an odd positive integer and $m \in \mathbb{N}$. Then,*

$$\chi^\star(H_k^{\boxtimes m}) \leq (k+1)^m.$$

Moreover, if there exists a Hadamard matrix of size $k + 1$, then equality holds.

[4] We stress that in our definition *orthogonality corresponds to adjacency.* Some authors prefer to demand orthogonality for non-adjacent vertices instead.

Proof. We first prove that $\chi^\star(H_k) \leq k+1$ by using Lemma 4.5.2. To this end we use the map f defined in (4.5), which is an orthonormal representation from $V(H_k)$ to $\mathbb{R}^{k+1}$ where the representing vectors have entries with equal moduli. We conclude that $\xi'(H_k) \leq k+1$ and so by Lemma 4.5.2 we get $\chi^\star(H_k) \leq k+1$. Using the sub-multiplicativity of $\chi^\star$ (Lemma 4.3.6) we get $\chi^\star(H_k^{\boxtimes m}) \leq (k+1)^m$.

We now prove that if there exists a Hadamard matrix of size $k+1$ then also the other direction of the inequality holds. Recall from Proposition 4.2.5 the existence of a Hadamard matrix of size $k+1$ implies $\omega(H_k) \geq k+1$. Combining this with Theorem 4.4.1 and the Sandwich Theorem [KD93] gives that for every positive integer m, we have

$$\chi^\star(H_k^{\boxtimes m}) \geq \vartheta(\overline{H_k^{\boxtimes m}}) \geq \omega(H_k^{\boxtimes m}) \geq \omega(H_k)^m \geq (k+1)^m, \tag{4.18}$$

where the second-last inequality uses the fact that if a subset $W \subseteq V(G)$ forms a clique in a graph G, then the set W^m of m-tuples forms a clique in $G^{\boxtimes m}$. □

The desired result now follows as a corollary.

Proof of (4.16). Combining Lemmas 4.3.6 and 4.5.3 gives

$$R^\star(H_k) \leq \log \chi^\star(H_k) \leq \log(k+1).$$

□

We also record the following additional corollary, which we use in Section 4.7.

Corollary 4.5.4. *For every odd integer k such that there is a Hadamard matrix of size $k+1$, we have $\omega(H_k^{\boxtimes m}) = (k+1)^m$.*

Proof. Combining Proposition 4.2.5, Lemma 4.5.3 and (4.18) gives the result. □

4.5.2 Lower bound on the classical Witsenhausen rate

To prove the lower bound (4.17) on $R(H_k)$ stated in Theorem 4.5.1 we use the following upper bound on the classical independence number of the graphs $H_k^{\boxtimes m}$ for certain values of k.

Lemma 4.5.5. *Let p be an odd prime number, $\ell \in \mathbb{N}$ and set $k = 4p^\ell - 1$. Then, for every $m \in \mathbb{N}$, we have*

$$\alpha(H_k^{\boxtimes m}) \leq \left(\binom{k}{0} + \binom{k}{1} + \cdots + \binom{k}{p^\ell - 1}\right)^m \leq 2^{k\,m\,H(3/11)} < 2^{0.846\,k\,m}, \tag{4.19}$$

where $H(t) = -t\log t - (1-t)\log(1-t)$ is the binary entropy function.

The proof of this lemma is an instance of the linear algebra method due to Alon [Alo98] (see also Gopalan [Gop06]), which we recall below for completeness. Let G be a graph and $\mathbb{F}$ be a field. Let $\mathcal{F} \subseteq \mathbb{F}[x_1, \dots, x_k]$ be a subspace of the space of k-variate polynomials over $\mathbb{F}$. A *representation* of G over $\mathcal{F}$ is an assignment $\big((f_u, c_u)\big)_{u \in V(G)} \subseteq \mathcal{F} \times \mathbb{F}^k$ of polynomial-point pairs to the vertices of G such that

$$f_u(c_u) \neq 0 \;\forall u \in V(G), \; f_u(c_v) = 0 \;\forall u \neq v \in V(G) \text{ with } \{u, v\} \notin E(G).$$

Lemma 4.5.6 (Alon [Alo98])**.** *Let G be a graph, $\mathbb{F}$ be a field, $k \in \mathbb{N}$ and $\mathcal{F}$ be a subspace of $\mathbb{F}[x_1, \dots, x_k]$. If $\big((f_u, c_u)\big)_{u \in V} \subseteq \mathcal{F} \times \mathbb{F}^k$ represents G, then $\alpha(G^{\boxtimes n}) \leq \dim(\mathcal{F})^n$ for all $n \in \mathbb{N}$.*

Proof. Let $I \subseteq V^n$ be an independent set in $G^{\boxtimes n}$. Each element $\mathbf{u} = (u_1, \dots, u_n)$ of I is a n-tuple of vertices of G and for every distinct pair $\mathbf{u}, \mathbf{v} \in I$ there is at least one index $i \in [n]$ such that u_i and v_i are neither equal nor adjacent in G.

For each $\mathbf{u} \in I$, define the polynomial $f_{\mathbf{u}} = \otimes_{i=1}^n f_{u_i} \in \mathcal{F}^{\otimes n}$, which takes as input n-tuples of vectors $\mathbf{c} = (c_1, \dots, c_n) \in (\mathbb{F}^k)^n$ and assumes the value $f_{\mathbf{u}}(\mathbf{c}) = f_{u_1}(c_1) \cdots f_{u_n}(c_n)$. Now define the n-tuple of vectors $c_{\mathbf{u}} = (c_{u_1}, \dots, c_{u_n})$. For all $\mathbf{u}$ we have $f_{u_1}(c_{u_1}) \cdots f_{u_n}(c_{u_n}) \neq 0$ and for all distinct $\mathbf{u}, \mathbf{v} \in I$ we have $f_{u_i}(c_{v_i}) = 0$ for at least one $i \in [n]$. It follows that the pairs $\big((f_{\mathbf{u}}, c_{\mathbf{u}})\big)_{\mathbf{u} \in V^n}$ represent $G^{\boxtimes n}$.

Now let $(a_{\mathbf{u}})_{\mathbf{u} \in I} \in \mathbb{F}^I$ be a sequence of scalars and consider the polynomial

$$f = \sum_{\mathbf{u} \in I} a_{\mathbf{u}} f_{\mathbf{u}}.$$

Then, by definition of a representation, for every $\mathbf{v} \in I$ such that $a_{\mathbf{v}} \neq 0$, we have

$$f(c_{\mathbf{v}}) = \sum_{\mathbf{u} \in I} a_{\mathbf{u}} f_{\mathbf{u}}(c_{\mathbf{v}}) = a_{\mathbf{v}} f_{\mathbf{v}}(c_{\mathbf{v}}) \neq 0.$$

It follows that f can only be the zero polynomial if $a_{\mathbf{u}}$ are zero for all $\mathbf{u}$ and hence the polynomials $f_{\mathbf{u}}$, for $\mathbf{u} \in I$, are linearly independent. This implies that $\alpha(G^{\boxtimes n}) \leq \dim(\mathcal{F}^{\otimes n}) = \dim(\mathcal{F})^n$. □

We will get a representation for the graph H_k, for $k = 4p^\ell - 1$, from the following result of Barrington, Beigel and Rudich [BBR94]. The proof we give here closely follows Yekhanin's [Yek12, Lemma 5.6] but is slightly more explicit. Below, a *multilinear* polynomial is a polynomial in which the degree of each variable is at most 1.

Lemma 4.5.7 (Barrington, Beigel and Rudich [BBR94])**.** *Let p be a prime number and let k, ℓ and w be integers such that $k > p^\ell$. There exists a multilinear* *polynomial* $f \in \mathbb{Z}_p[x_1, \ldots, x_k]$ *of degree* $\deg(f) \leq p^\ell - 1$ *such that for every* $c \in \{0,1\}^k$, *we have*

$$f(c) \equiv \begin{cases} 1 & \textit{if } c_1 + c_2 + \cdots c_k \equiv w \bmod p^\ell \\ 0 & \textit{otherwise.} \end{cases}$$

The proof of this lemma relies on Lucas' Theorem from number theory.

Theorem 4.5.8 (Lucas' Theorem). *Let* p *be a prime and* $a, b \in \mathbb{N}$ *with* p*-ary expansions* $a = \sum_i a_i p^i$ *and* $b = \sum_i b_i p^i$, *where* $0 \leq a_i, b_i < p$. *Then,*

$$\binom{a}{b} \equiv \prod_i \binom{a_i}{b_i} \bmod p.$$

Proof of Lemma 4.5.7. For $c \in \{0,1\}^k$, note that the value modulo p^ℓ of the Hamming weight $|c|$ depends only on the first ℓ coefficients $|c|_0, |c|_1, \ldots, |c|_{\ell-1}$ of the p-ary expansion of $|c|$. The k-variate symmetric polynomial of degree d is defined by

$$P_d(x_1, \ldots, x_k) = \sum_{S \in \binom{[k]}{d}} \prod_{i \in S} x_i.$$

For every $c \in \{0,1\}^k$, we have

$$P_{p^i}(c) = \binom{|c|}{p^i} \equiv \binom{|c|_i}{1} \bmod p \equiv |c|_i \bmod p,$$

where the second identity follows from Lucas' theorem and the p-ary expansion of p^i, in which the coefficient of value 1 multiplying p^i is the only nonzero coefficient. Now, define the polynomial $\hat{f} \in \mathbb{Z}_p[x_1, \ldots, x_k]$ by

$$\hat{f}(x_1, \ldots, x_k) = \prod_{i=0}^{\ell-1} \Big(1 - \big(P_{p^i}(x) - w_i\big)^{p-1}\Big),$$

where w_i are the coefficients in p-ary expansion of w. For $c \in \{0,1\}^k$, we have $\hat{f}(c) \equiv 1 \bmod p$ if $|c|_i \equiv w_i$ for every $i = 0, 1, \ldots, \ell - 1$ (*i.e.*, if $|c| \equiv w \bmod p^\ell$) and $\hat{f}(c) \equiv 0 \bmod p$ otherwise. Here, we have used Fermat's Little Theorem, which states that, for p prime and $a \in \mathbb{N}$, $a^{p-1} \equiv 1 \bmod p$. Clearly the polynomial $\hat{f}$ has only integer coefficients. Now let f be the multilinear polynomial obtained from $\hat{f}$ by replacing each monomial $x_1^{d_1} \cdots x_k^{d_k}$ by $x_1^{i_1} \cdots x_k^{i_k}$ where $i_h = \min\{d_h, 1\}$ for every $h \in [k]$. Then, the degree of the polynomial f is bounded by $\deg(f) \leq \deg(\hat{f}) \leq (p-1)(1 + p + p^2 + \cdots + p^{\ell-1}) = p^\ell - 1$. Moreover, f agrees with $\hat{f}$ on $\{0,1\}^k$ and satisfies the conditions of the lemma. □

With this we can now prove Lemma 4.5.5.

Proof of Lemma 4.5.5. Let $c \in \{0,1\}^k$ be a string such that its Hamming weight $|c|$ is even and satisfies $|c| \equiv 0 \bmod p^\ell$. Then, since p is odd and $k < 4p^\ell$, we have $|c| \in \{0, 2p^\ell\}$. Hence, if $|c| \notin \{0, 2p^\ell\}$, then $|c| \not\equiv 0 \bmod p^\ell$.

Recall from Definition 4.2.2 that H_k can be defined as the graph whose vertices are the strings of $\{0,1\}^k$ with an even Hamming weight and where two distinct vertices u, v are adjacent if their Hamming distance $|u \oplus v|$ is equal to $(k+1)/2 = 2p^\ell$. Here $u \oplus v$ is the sum modulo 2. For $u, v \in V(H_k)$, their Hamming distance $|u \oplus v|$ is an even number. Hence if $u \neq v$ are not adjacent in H_k, then $|u \oplus v| \notin \{0, 2p^\ell\}$ and thus $|u \oplus v| \not\equiv 0 \mod p^\ell$.

Let $f \in \mathbb{Z}_p[x_1, \dots, x_k]$ be a multilinear polynomial of degree at most $p^\ell - 1$ such that for every $c \in \{0,1\}^k$, we have

$$f(c) \equiv \begin{cases} 1 & \text{if } |c| \equiv 0 \bmod p^\ell \\ 0 & \text{otherwise,} \end{cases}$$

as is promised to exist by Lemma 4.5.7 (applied to $w = 0$).

We use f to define a representation for H_k. To this end define for each $u \in \{0,1\}^k$ vertex in $V(H_k)$ the polynomial $f_u \in \mathbb{Z}_p[x_1, \dots, x_k]$ obtained by replacing in the polynomial f the variable x_i by $1 - x_i$ if $u_i = 1$ and leaving it unchanged otherwise. For example, if $u = (1, 1, 0, \dots, 0)$, then $f_u(x_1, \dots, x_k) = f(1 - x_1, 1 - x_2, x_3, \dots, x_k)$. Moreover, associate to the vertex u the point $c_u = u$ seen as a 0/1 vector in $\mathbb{Z}_p^k$. We claim that $\big((f_u, c_u)\big)_{u \in V(H_k)}$ is a representation of H_k. To see this, observe that $f_u(c_v) = f(u \oplus v)$ for any $u, v \in V(H_k)$, so that $f_u(c_u) = f(0) = 1$, and $f_u(c_v) = 0$ if u, v are distinct and non-adjacent.

Since the polynomials f_u are multilinear and have degree at most $p^\ell - 1$, they span a space of dimension at most $\binom{k}{0} + \binom{k}{1} + \cdots + \binom{k}{p^\ell - 1}$, which is the number of multilinear monomials of degree at most $p^\ell - 1$. Applying Lemma 4.5.6 we obtain that

$$\alpha(H_k^{\boxtimes m}) \leq \left(\binom{k}{0} + \binom{k}{1} + \cdots + \binom{k}{p^\ell - 1} \right)^m. \tag{4.20}$$

We now use the well known fact that for $q, k \in \mathbb{N}$ with $1 < q < k/2$, $\binom{k}{0} + \ldots + \binom{k}{q-1} \leq 2^{kH(q/k)}$. From this, since $p^\ell/(4p^\ell - 1) \leq 3/11$, we deduce that the right hand side in (4.20) can be upper bounded by $2^{k\,m\,H(3/11)} < 2^{0.846\,k\,m}$. □

The bound (4.17) stated in Theorem 4.5.1 is a corollary of Lemma 4.5.5.

Proof of (4.17). By Lemma 4.5.5, for every integer m we have

$$\chi(H_k^{\boxtimes m}) \geq \frac{|V(H_k^{\boxtimes m})|}{\alpha(H_k^{\boxtimes m})} \geq \frac{2^{(k-1)m}}{2^{0.846km}} = 2^{(0.154k-1)m}.$$

Taking the logarithm, dividing by m and taking the limit $m \to \infty$ gives that for $k = 4p^\ell - 1$, we have

$$R(H_k) \geq 0.154k - 1.$$

□

4.6 Separation between classical and entangled Shannon capacity

Our second separation result is a strengthening of the following result of [BBG12], which shows that for some values of k, the entangled Shannon capacity of H_k can be strictly larger than its (classical) Shannon capacity.

Theorem 4.6.1 (Briët, Buhrman and Gijswijt [BBG12]). *Let p be an odd prime such that there exists a Hadamard matrix of size $4p$. Set $k = 4p - 1$. Then,*

$$\begin{aligned} c^\star(H_k) &\geq k - 1 - 2\log(k+1) \\ c(H_k) &\leq 0.846k. \end{aligned}$$

Note that here we consider the exact bounds on $c^\star(H_k)$ and $c(H_k)$ rather than the asymptotic ones as originally written in [BBG12]. It is not known if Hadamard matrices of size $4p$ exist for infinitely many primes p. Theorem 4.6.1 requires the existence of Hadamard matrices due to the technique used to lower bound $c^\star(H_k)$, which originates from [LMM$^+$12]. It also requires that k is of the form $rp - 1$ for some odd prime p and positive integer $r \geq 4$ due to the technique used to upper-bound $c(H_k)$, which is based on a result of Frankl and Wilson [FW81].

Here we relax the conditions in Theorem 4.6.1 and our result does not rely anymore on the existence of a Hadamard matrix. We show the existence of an infinite family of quarter-orthogonality graphs whose entangled capacity exceeds their Shannon capacity.

Theorem 4.6.2. *For every odd integer $k \geq 5$, we have*

$$c^\star(H_k) \geq (k-1)\left(1 - \frac{4\log(k+1)}{k-3}\right). \tag{4.21}$$

Moreover, if $k = 4p^\ell - 1$ where p is an odd prime and $\ell \in \mathbb{N}$, then

$$c(H_k) \leq 0.846\,k. \tag{4.22}$$

In the next sections we prove Theorem 4.6.2.

4.6.1 Lower bound on the entangled Shannon capacity

The proof of the bound (4.21) on the entangled Shannon capacity is based on quantum teleportation (see Section 4.2.3). In operational terms the proof can be interpreted as showing that with $t + 1$ sequential uses of a channel with confusability graph H_k, Alice can send Bob $|V|^t$ distinct messages with zero probability of error provided that $t \leq \log\alpha(H_k)/(2\log(k+1))$. To give some intuition we explain this operational interpretation before moving on to the proof.

Let f be the map defined in (4.5) and define $\rho^x = f(x)f(x)^\mathsf{T}$ for $x \in V(H_k)$. To transmit a sequence $\mathbf{x} = (x_1, \dots, x_t) \in V(H_k)^t$ Alice and Bob may follow the following four-step procedure. First, Alice prepares $(k+1)$-dimensional quantum systems $\mathcal{A}_1, \dots, \mathcal{A}_t$ to be in the states $\rho^{x_1}, \dots, \rho^{x_t}$, respectively. Second, Alice sends the sequence $\mathbf{x}$ through the channel by using it t times in a row. This will result in t channel-outputs on Bob's end of the channel from which he can infer that each x_i belongs to a particular clique in H_k. Third, Alice and Bob execute a quantum teleportation scheme after which Bob ends up with quantum systems $\mathcal{Y}_1, \dots, \mathcal{Y}_t$ in states $\rho^{x_1}, \dots, \rho^{x_t}$, respectively. The teleportation step requires that Alice communicates a total of $2t\lceil\log(k+1)\rceil$ bits to Bob. We are now ready to prove the lower bound.

Proof of (4.21). Set $V = V(H_k)$ and let $t \in \mathbb{N}$ be such that $(k+1)^{2t} \leq \alpha(H_k)$. (Note that this choice of t follows from the fact that Alice needs to use the channel classically to send the measurement outcomes of the teleportation to Bob.) In what follows we construct a trace-1 positive semidefinite matrix ρ and, for every $\mathbf{x} \in V^t$, positive semidefinite matrices $\{\rho^{\mathbf{x}}_{\mathbf{u}} : \mathbf{u} \in V^{t+1}\}$ satisfying the conditions of Definition 4.3.3, i.e.,

$$\sum_{\mathbf{u} \in V^{t+1}} \rho^{\mathbf{x}}_{\mathbf{u}} = \rho \quad \forall \mathbf{x}, \tag{4.23}$$

$$\rho^{\mathbf{x}}_{\mathbf{u}} \rho^{\mathbf{y}}_{\mathbf{v}} = 0 \quad \forall \mathbf{x} \neq \mathbf{y},\ \{\mathbf{u}, \mathbf{v}\} \in V^{t+1} \cup E(H_k^{\boxtimes(t+1)}). \tag{4.24}$$

This implies that $\alpha^\star(H_k^{\boxtimes(t+1)}) \geq |V|^t$.

Let $f : V \to \mathbb{R}^{k+1}$ be the orthonormal representation of H_k defined in (4.5). For $x \in V$ define $\rho^x = f(x)f(x)^\mathsf{T}$ and, for $\mathbf{x} = (x_1, \dots, x_t) \in V^t$, define $\rho^{\mathbf{x}} = \rho^{x_1} \otimes \rho^{x_2} \otimes \cdots \otimes \rho^{x_t}$. Notice that $\mathrm{Tr}(\rho^{\mathbf{x}}) = 1$ and that $\rho^{\mathbf{x}}\rho^{\mathbf{y}} = 0$ for every $\{\mathbf{x}, \mathbf{y}\} \in E(H_k^{\boxtimes t})$. We now consider the quantum teleportation scheme from Section 4.2.3, for the setting where Alice would want to transmit the state $\rho^{\mathbf{x}}$ of a $(k+1)^t$-dimensional quantum system $\mathcal{A}$ to Bob. According to (QT1), let σ be the maximally entangled state defined over a pair of $(k+1)^t$-dimensional quantum systems $(\mathcal{X}, \mathcal{Y})$, where $\mathcal{X}$ belongs to Alice and $\mathcal{Y}$ to Bob. With $T = (k+1)^{2t}$, let $\{M_i : i \in [T]\}$ be Alice's measurement on the system $(\mathcal{A}, \mathcal{X})$ provided by (QT2), and let $U_1, \dots, U_T$ be Bob's unitary operators on $\mathcal{Y}$ given by (QT3). Define

$$\begin{aligned} \rho &= \mathrm{Tr}_{\mathcal{X}}(\sigma), \\ \rho^{\mathbf{x}}_i &= \mathrm{Tr}_{(\mathcal{A},\mathcal{X})}\big((M_i \otimes I)(\rho^{\mathbf{x}} \otimes \sigma)\big) \quad \forall x \in V^t, i \in [T]. \end{aligned}$$

Since the M_i sum to the identity, for every $\mathbf{x}$, we have

$$\sum_{i=1}^{T} \rho^{\mathbf{x}}_i = \mathrm{Tr}_{(\mathcal{A},\mathcal{X})}(\rho^{\mathbf{x}} \otimes \sigma) = \mathrm{Tr}_{\mathcal{A}}(\rho^{\mathbf{x}})\,\mathrm{Tr}_{\mathcal{X}}(\sigma) = \rho. \tag{4.25}$$

By (QT4), we know that that the identity $U_i \rho_i^{\mathbf{x}} U_i^\dagger = \beta_i^{\mathbf{x}} \rho^{\mathbf{x}}$ holds, where $\beta_i^{\mathbf{x}} = \mathrm{Tr}(\rho_i^{\mathbf{x}})$. Hence, since f is an orthonormal representation, for every edge $\{\mathbf{x}, \mathbf{y}\} \in E(H_k^{\boxtimes t})$, we have

$$\rho_i^{\mathbf{x}} \rho_i^{\mathbf{y}} = (U_i^\dagger \beta_i^{\mathbf{x}} \rho^{\mathbf{x}} U_i)\,(U_i^\dagger \beta_i^{\mathbf{y}} \rho^{\mathbf{y}} U_i) = \beta_i^{\mathbf{x}} \beta_i^{\mathbf{y}} U_i^\dagger \rho^{\mathbf{x}} \rho^{\mathbf{y}} U_i = 0. \tag{4.26}$$

Let $W \subseteq V$ be an independent set in H_k with cardinality $|W| = T$ and let $\phi : W \to [T]$ be some bijection. For every $\mathbf{u} \in V^{t+1}$ and $\mathbf{x} \in V^t$ define

$$\rho_{\mathbf{u}}^{\mathbf{x}} = \begin{cases} \rho_{\phi(u_{t+1})}^{\mathbf{x}} & \text{if } (u_1, \ldots, u_t) = \mathbf{x} \text{ and } u_{t+1} \in W \\ 0 & \text{otherwise.} \end{cases}$$

Then, $\sum_{\mathbf{u} \in V^{t+1}} \rho_{\mathbf{u}}^{\mathbf{x}} = \sum_{u_{t+1} \in W} \rho_{\phi(u_{t+1})}^{\mathbf{x}} = \sum_{i=1}^{T} \rho_i^{\mathbf{x}} = \rho$ by (4.25). Next, let $\mathbf{x} \neq \mathbf{y} \in V^t$ and $\{\mathbf{u}, \mathbf{v}\} \in V^{t+1} \cup E(H_k^{\boxtimes(t+1)})$; we show that $\rho_{\mathbf{u}}^{\mathbf{x}} \rho_{\mathbf{v}}^{\mathbf{y}} = 0$. This is clear if $\mathbf{x} \neq (u_1, \ldots, u_t)$, or $\mathbf{y} \neq (v_1, \ldots, v_t)$, or $\{u_{t+1}, v_{t+1}\} \not\subseteq W$. So we may assume $\mathbf{u} = (\mathbf{x}, u_{t+1})$, $\mathbf{v} = (\mathbf{y}, v_{t+1})$ and $\{u_{t+1}, v_{t+1}\} \subseteq W$ and thus $\{\mathbf{u}, \mathbf{v}\} \in E(H_k^{\boxtimes(t+1)})$, $\{\mathbf{x}, \mathbf{y}\} \in E(H_k^{\boxtimes t})$ and $u_{t+1} = v_{t+1}$. Then we have that $\rho_{\mathbf{u}}^{\mathbf{x}} \rho_{\mathbf{v}}^{\mathbf{y}} = \rho_{\phi(u_{t+1})}^{\mathbf{x}} \rho_{\phi(u_{t+1})}^{\mathbf{y}} = 0$ by (4.26).

Hence, for t such that $(k+1)^{2t} \leq \alpha(H_k)$, we have $\alpha^\star(H_k^{\boxtimes(t+1)}) \geq |V|^t = 2^{(k-1)t}$. This implies

$$c^\star(H_k) \geq \frac{1}{t+1} \log \alpha^\star(H_k^{\boxtimes(t+1)}) \geq \frac{1}{t+1} t(k-1). \tag{4.27}$$

By Lemma 4.2.3 we have $\alpha(H_k) \geq 2^{(k-3)/2}$. Hence, for $k \geq 5$ we can choose the integer t to be equal to $t = \lfloor (k-3)/4\log(k+1) \rfloor$. From (4.27) we then get

$$c^\star(H_k) \;\geq\; \frac{4\log(k+1)}{k-3}\Big(\frac{k-3}{4\log(k+1)} - 1\Big)(k-1) \geq (k-1)\left(1 - \frac{4\log(k+1)}{k-3}\right)$$

which gives the claimed result. □

4.6.2 Upper bound on the Shannon capacity

The upper bound (4.22) on the Shannon capacity of H_k (for certain values of k) stated in Theorem 4.6.2 is a corollary of Lemma 4.5.5.

Proof of (4.22). By taking the logarithm, dividing by m and taking the limit $m \to \infty$ on both sides of (4.19) we get that for p odd prime, $\ell \in \mathbb{N}$ and $k = 4p^\ell - 1$,

$$c(H_k) \leq 0.846k.$$

□

4.7 Separation between classical and entangled source-channel cost rate

Our last contribution concerns the combined source-channel problem for a source that has H_k as characteristic and a channel that has H_k as confusability graph. The result is the following.

Theorem 4.7.1. *Let p be an odd prime and $\ell \in \mathbb{N}$ such that there exists a Hadamard matrix of size $4p^\ell$. Set $k = 4p^\ell - 1$. Then,*

$$\eta^\star(H_k, H_k) \leq \frac{\log(k+1)}{(k-1)\left(1 - \frac{4\log(k+1)}{k-3}\right)}, \tag{4.28}$$

$$\eta(H_k, H_k) > \frac{0.154\,k - 1}{k - 1 - \log(k+1)}. \tag{4.29}$$

The proof of Theorem 4.7.1 is given below. The bound on the entangled source-channel cost rate is obtained by concatenating an entanglement-assisted coding scheme for a source with one for a channel. In this way, one obtains a "separated" coding scheme for the source-channel problem, see Section 4.3.2 for details. There we show that the asymptotic cost rate of a separate coding scheme is $R^\star(H_k)/c^\star(H_k)$ and thus $\eta^\star(H_k, H_k) \leq R^\star(H_k)/c^\star(H_k)$. The bound for the classical parameter $\eta(H_k, H_k)$ relies on properties of the fractional chromatic number and the fact that H_k is vertex-transitive. Let us point out that Theorem 4.7.1 holds for an infinite family of graphs. This follows from the result of Xia and Lu [XL91] in Theorem 4.2.4, since there exist infinitely many (p, ℓ)-pairs such that $p^{\ell/2} \equiv 1 \bmod 4$. (For instance, for $p = 5$ and $\ell = 2i$ with $i \in \mathbb{N}$, $5^i = (4+1)^i \equiv 1 \bmod 4$.)

Hence, for any k satisfying the condition of the theorem, we have an exponential separation between the entangled and the classical source-channel cost rate as

$$\eta^\star(H_k, H_k) \leq \frac{R^\star(H_k)}{c^\star(H_k)} \leq O\Big(\frac{\log k}{k}\Big) \ll \Omega(1) \leq \eta(H_k, H_k).$$

As shown in [NTR06], a large separation $\eta(G, H) \ll R(G)/c(H)$ exists for some graphs. But this is not the case for our source-channel combination using $G = H = H_k$. Indeed,

$$\Omega(1) \leq \eta(H_k, H_k) \leq \frac{R(H_k)}{c(H_k)} \leq \frac{\log \chi(H_k)}{\log \alpha(H_k)} \leq \frac{2(k-1)}{k-3} \leq O(1),$$

where in the second last inequality we use that $\log \chi(H_k) \leq \log |V(H_k)| = k - 1$ and that $\log \alpha(H_k) \geq (k-3)/2$ (Lemma 4.2.3).

Now we prove Theorem 4.7.1, separately showing the two bounds (4.28) for $\eta^\star$ and (4.29) for η. The bound (4.28) is obtained by combining (4.16), (4.21) with Proposition 4.3.8. The proof of (4.29) relies on some basic properties of the *fractional chromatic number* of vertex-transitive graphs, which we now recall.

Let G be a graph and let $\mathcal{I}_G$ denote the collection of its independent sets. For $I \subseteq V(G)$, $\chi^I \in \{0,1\}^{V(G)}$ denotes its characteristic vector, with $\chi^I_u = 1$ if and only if $u \in I$. We let $\mathbf{1}$ denotes the all-ones vector. The *fractional chromatic number* $\chi_{\mathsf{f}}(G)$ is a lower bound to the chromatic number, defined by

$$\chi_{\mathsf{f}}(G) = \min\left\{ \sum_{I\in\mathcal{I}_G} \lambda_I : \sum_{I\in\mathcal{I}_G} \lambda_I \chi^I \geq \mathbf{1},\ \lambda_I \geq 0\ \forall I \in \mathcal{I}_G \right\}.$$

An *automorphism* of G is a permutation π of $V(G)$ preserving edges, i.e., $\{\pi(u),\pi(v)\} \in E(G)$ if and only if $\{u,v\} \in E(G)$. The graph G is *vertex-transitive* if, for any $u, v \in V(G)$, there exists an automorphism π of G such that $v = \pi(u)$. We use the following well known facts.

Lemma 4.7.2. *(see e.g. [GR01, Corollaries 7.4.2, 7.5.2])*

(i) For graphs G and H, if $G \longrightarrow H$ then $\chi_{\mathsf{f}}(G) \leq \chi_{\mathsf{f}}(H)$.

(ii) For a graph G, $\chi_{\mathsf{f}}(G) \geq |V(G)|/\alpha(G)$, with equality if G is vertex-transitive.

Corollary 4.7.3. *Let G and H be vertex-transitive graphs. If there is a homomorphism from G to H, then*

$$\frac{|V(G)|}{\alpha(G)} \leq \frac{|V(H)|}{\alpha(H)}.$$

As observed in [BBG12], the graph H_k is vertex-transitive; indeed, for any $u \in V(H_k)$, consider the map $v \mapsto u \oplus v$. One can see that taking the strong product and complement of graphs preserves vertex-transitivity. Hence, $\overline{H_k^{\boxtimes n}}$ is vertex-transitive for any $n \in \mathbb{N}$.

We are now ready to prove the bound (4.29).

Proof of (4.29). Recall the definition of $\eta(H_k, H_k)$ from (4.3). Consider integers $m, n \in \mathbb{N}$ for which $H_k^{\boxtimes m} \longrightarrow \overline{H_k^{\boxtimes n}}$. Applying Corollary 4.7.3, we deduce that

$$\frac{|V(H_k^{\boxtimes m})|}{\alpha(H_k^{\boxtimes m})} \leq \frac{|V(\overline{H_k^{\boxtimes n}})|}{\alpha(\overline{H_k^{\boxtimes n}})} = \frac{|V(\overline{H_k^{\boxtimes n}})|}{\omega(H_k^{\boxtimes n})}. \tag{4.30}$$

From Corollary 4.5.4 we have $\omega(H_k^{\boxtimes n}) = (k+1)^n$. As $|V(H_k)| = 2^{k-1}$ and applying Lemma 4.5.5, we get

$$\frac{2^{(k-1)\,m}}{2^{k\,m\,H(3/11)}} \overset{\text{Lemma 4.5.5}}{\leq} \frac{|V(H_k^{\boxtimes m})|}{\alpha(H_k^{\boxtimes m})} \overset{(4.30)}{\leq} \frac{|V(\overline{H_k^{\boxtimes n}})|}{\omega(H_k^{\boxtimes n})} = \frac{2^{(k-1)\,n}}{(k+1)^n}.$$

After a few algebraic manipulations and taking logarithms the above inequality reduces to

$$\frac{n}{m} \geq \frac{k(1 - H(3/11)) - 1}{k - 1 - \log(k+1)} > \frac{0.154\,k - 1}{k - 1 - \log(k+1)}.$$

This gives directly the lower bound (4.29). □

4.7.1 Stronger bounds based on Hadamard matrices

The reader may have noticed that for the purpose of proving Theorem 4.7.1, we may assume that the integer k appearing in (4.28) is such that there exists a Hadamard matrix of size $k+1$. The reason for this is that the lower bound on $\eta(H_k, H_k)$ given in (4.29) is conditional on the existence of such a matrix. With this additional assumption a stronger upper bound on $\eta^\star(H_k, H_k)$ can be proved without the use of quantum teleportation.

To prove this, we bound $\eta_1^\star(H_k, H_k)$ by the rate achievable with separate entangled coding schemes for the source-coding and channel-coding problem, respectively (see Section 4.3.2). To do so, we need a lower bound on the entangled independence number that was obtained previously in [BBG12].

Lemma 4.7.4 ([BBG12])**.** *Let k be a positive integer such that there is a Hadamard matrix of size $k+1$. Then,*

$$\log \alpha^\star(H_k) \geq k - 1 - 2\log(k+1).$$

Lemma 4.7.5. *Let k be a positive integer such that there is a Hadamard matrix of size $k+1$. Then,*

$$\eta_1^\star(H_k, H_k) \leq \left\lceil \frac{\log(k+1)}{k - 1 - 2\log(k+1)} \right\rceil.$$

Proof. Putting together Proposition 4.3.9, Lemma 4.5.3 and Lemma 4.7.4 we have that, for every k such that there exists a Hadamard matrix of size $(k+1)$,

$$\eta_1^\star(H_k, H_k) \leq \left\lceil \frac{\log \chi^\star(H_k)}{\log \alpha^\star(H_k)} \right\rceil \leq \left\lceil \frac{\log(k+1)}{k - 1 - 2\log(k+1)} \right\rceil$$

which proves the claim. □

Since we have $\eta^\star(H_k, H_k) \leq \eta_1^\star(H_k, H_k)$, the above result also implies an upper bound of the cost rate attainable by encoding infinitely long sequences of source inputs into single codewords.

4.8 Concluding remarks and open problems

We have shown a separation between classical and entanglement-assisted coding for the zero-error source-channel, source and channel problems. Note that these separations do not hold if asymptotically vanishing error is allowed. We have presented an infinite family of instances for which there is an exponential saving in the minimum asymptotic cost rate of communication for the source-channel and the source coding problems. Moreover, for the channel coding problem we have shown an infinite family of channels for which the entangled Shannon capacity exceeds the classical Shannon capacity by a constant factor. It would be interesting to find a family of channels with a larger separation.

The main result in [NTR06] is that, for the classical source-channel coding problem, there exist situations for which separate encoding is highly suboptimal. Does this happen also in the entanglement-assisted case? This question has a positive answer if there exists a graph G with $R^\star(G) > c^\star(\overline{G})$. In [NTR06] a sufficient condition for a separate encoding to be optimal is also proven, namely that the characteristic or the confusability graph is a perfect graph. It is straightforward to see that this is also a sufficient condition for a separate entanglement-assisted encoding to be optimal. Are there stronger conditions that hold for the entangled case?

One of the most interesting open questions in zero-error classical information theory is the computational complexity of the Witsenhausen rate and of the Shannon capacity. The same question is also open for the entangled counterparts, as well as for the parameters $\chi^\star$ and $\alpha^\star$.

In Section 4.1, we have seen that the entangled chromatic and independence number generalize the parameters χ_q and α_q which arise in the context of Bell inequalities and non-local games. In [RM12] it is conjectured that $\alpha^\star(G) = \alpha_q(G)$ for every graph G. A possible approach to prove that entangled chromatic number and quantum chromatic number are distinct quantities is to prove that the relationship between Kochen-Specker sets and χ_q found in Chapter 3 does not hold for $\chi^\star$.

Additionally, we mention that the existence of a graph G for which $\chi^\star(G) < \chi_q(G)$ or $\alpha_q(G) < \alpha^\star(G)$ would prove the existence of a non-local game such that every quantum strategy that wins with probability one does not use a maximally entangled state. This is because the source-channel settings studied in this chapter can be seen as non-local games as follows. Alice receives input $\mathsf{x} \in \mathsf{X}^m$ and Bob receives inputs $\mathsf{u}, \mathsf{v} \in \mathsf{U}^m \times \mathsf{V}^n$. Alice produces as output $\mathsf{s} \in \mathsf{S}^n$, and Bob produces output $\mathsf{y} \in \mathsf{X}^m$. The winning condition is the following: if s, v is an invalid channel input-output pair, then players win, otherwise players win if $\mathsf{x} = \mathsf{y}$. By setting the parameters n, m according to the entangled source-channel rate, it is possible to obtain a pseudo-telepathy game (although the classical value can be very close to 1, due to the many winning instances).

Chapter acknowledgments I thank the coauthors of the paper on which this chapter is based. Together, we thank Ronald de Wolf for useful discussions and helpful comments on a preliminary version of the paper, and Dion Gijswijt for helpful discussions.

Bibliography

[AGR81] A. Aspect, P. Grangier, and G. Roger. Experimental tests of realistic local theories via Bell's theorem. *Phys. Rev. Lett.*, 47(7):460–463, 1981.

[AHKS06] D. Avis, J. Hasegawa, Y. Kikuchi, and Y. Sasaki. A quantum protocol to win the graph colouring game on all hadamard graphs. *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.*, E89-A(5):1378–1381, 2006.

[AL06] N. Alon and E. Lubetzky. The Shannon capacity of a graph and the independence numbers of its powers. *Information Theory, IEEE Transactions on*, 52(5):2172–2176, 2006.

[Alo98] N. Alon. The shannon capacity of a union. *Combinatorica*, 18:301–310, 1998.

[Alo02] N. Alon. Graph powers. *Contemporary Combinatorics,(B. Bollobás, ed.), Bolyai Society Mathematical Studies, Springer*, pages 11–28, 2002.

[AOW11] F. Arends, J. Ouaknine, and C. W. Wampler. On searching for small Kochen-Specker vector systems. In *Proceedings of WG 11*, 2011.

[BB84] C. H. Bennett and G. Brassard. Quantum cryptography: Public key distribution and coin tossing. In *Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing*, pages 175–179. IEEE, 1984.

[BBC⁺93] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Phys. Rev. Lett.*, 70:1895–1899, 1993.

[BBG12] J. Briët, H. Buhrman, and D. Gijswijt. Violating the Shannon capacity of metric graphs with entanglement. *Proceedings of the National Academy of Sciences*, 2012.

[BBLV09] J. Briët, H. Buhrman, T. Lee, and T. Vidick. Multiplayer XOR games and quantum communication complexity with clique-wise entanglement. arXiv:0911.4007, 2009.

[BBR94] D. A. M. Barrington, R. Beigel, and S. Rudich. Representing Boolean functions as polynomials modulo composite numbers. *Computational Complexity*, 4(4):367–382, 1994.

[BCGSM] J. Briët, S. Chakraborty, D. García-Soriano, and A. Matsliah. Monotonicity testing and shortest-path routing on the cube. *Combinatorica*, 32(1):1–19, 2012.

[BCT99] G. Brassard, R. Cleve, and A. Tapp. The cost of exactly simulating quantum entanglement with classical communication. *Phys. Rev. Lett.*, 83(9):1874–1877, 1999. quant-ph/9901035.

[BCW98] H. Buhrman, R. Cleve, and A.Wigderson. Quantum vs. classical communication and computation. *Proceedings of the 30th ACM STOC*, pages 63–68, 1998.

[Bei10] S. Beigi. Entanglement-assisted zero-error capacity is upper-bounded by the Lovász ϑ function. *Phys. Rev. A*, 82(1):010303, 2010.

[Bel64] J. S. Bell. On the Einstein-Podolsky-Rosen paradox. *Physics*, 1:195–200, 1964.

[Bel66] J. S. Bell. On the problem of hidden variables in quantum mechanics. *Rev. Mod. Phys.*, 38:447–452, 1966.

[BSST02] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal. Entanglement-assisted capacity of a quantum channel and the reverse shannon theorem. *IEEE Trans. Inf. Theory*, 48(10):2637–2655, 2002.

[Bus03] P. Busch. Quantum states and generalized observables: A simple proof of Gleason's theorem. *Phys. Rev. Lett.*, 91:120403–120406, 2003.

[BV11] J. Briët and T. Vidick. Explicit lower and upper bounds on the entangled value of multiplayer XOR games. arXiv:1108.5647, 2011.

[BYJK08] Z. Bar-Yossef, T. S. Jayram, and I. Kerenidis. Exponential separation of quantum and classical one-way communication complexity. *SIAM J. Comp.*, 38(1):366–384, 2008. Earlier version in STOC'04.

[Cab03] A. Cabello. Kochen-Specker theorem for a single qubit using positive operator-valued measures. *Phys. Rev. Lett.*, 90:190401–190404, 2003.

[CFMR04] C. Caves, C. Fuchs, K. Manne, and J. Renes. Gleason-type derivations of the quantum probability rule for generalized measurements. *Found. Phys.*, 34:193–209, 2004.

[Che97] Y. Q. Chen. On the existence of abelian Hadamard difference sets and a new family of difference sets. *Finite fields Appl.*, 3(3):234–256, 1997.

[CHSH69] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.*, 23(15):880–884, 1969.

[CLMW10] T. S. Cubitt, D. Leung, W. Matthews, and A. Winter. Improving zero-error classical communication with entanglement. *Phys. Rev. Lett.*, 104(23):230503, 2010.

[CMN$^+$07] P. J. Cameron, A. Montanaro, M. W. Newman, S. Severini, and A. Winter. On the quantum chromatic number of a graph. *Electr. J. Comb.*, 14(1), 2007.

[CPSS12] A. Cabello, M. G. Parker, G. Scarpa, and S. Severini. Exclusivity structures and graph representatives of local complementation orbits, 2012. *J. Math. Phys.*, 54(7), 2013.

[CSUU07] R. Cleve, W. Slofstra, F. Unger, and S. Upadhyay. Strong parallel repetition theorem for quantum XOR proof systems. In *Proceedings of 22nd IEEE CCC*, pages 282–299, 2007. quant-ph/0608146.

[CSW10] A. Cabello, S. Severini, and A. Winter. (Non-)contextuality of physical theories as an axiom, 2010. arXiv:1010.2163.

[DKLR09] J. Degorre, M. Kaplan, S. Laplante, and J. Roland. The communication complexity of non-signaling distributions. In *Proceedings of 34th MFCS*, pages 270–281, 2009. arXiv:0804.4859.

[dKP07] E. de Klerk and D. V. Pasechnik. A note on the stability number of an orthogonality graph. *Europ. J. Combinat.*, 28:1971–1979, 2007.

[DSW13] R. Duan, S. Severini, and A. Winter. Zero-error communication via quantum channels, noncommutative graphs, and a quantum Lovász number. *IEEE Trans. Inf. Theory*, 59(2):1164 –1174, 2013.

[Duk10] D. D. Dukaric. The Hilbertian tensor norm and its connection to quantum information theory. arXiv:1008.1948v2, 2010.

[dW01] R. de Wolf. *Quantum computing and communication complexity.* PhD thesis, Universiteit van Amsterdam, 2001.

[EBB71] A. Einstein, M. Born, and H. Born. *The Born-Einstein letters: correspondence between Albert Einstein and Max and Hedwig Born from 1916-1955, with commentaries by Max Born.* Macmillan, 1971.

[EPR35] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47:777–780, 1935.

[FB75] M. J. Ferguson and D. W. Bailey. Zero-error coding for correlated sources, Unpublished, 1975.

[FIG11] J. Fukawa, H. Imai, and F. Le Gall. Quantum coloring games via symmetric SAT games. In *Proceedings of (AQIS'11)*, 2011.

[FR87] P. Frankl and V. Rödl. Forbidden intersections. *Trans. Amer. Math. Soc.*, 300(1):259–286, 1987.

[FW81] P. Frankl and R.M. Wilson. Intersection theorems with geometric consequences. *Combinatorica*, 1(4):357–368, 1981.

[Gav09] D. Gavinsky. Classical interaction cannot replace quantum nonlocality. arXiv:0901.0956, 2009.

[GKK$^+$08] D. Gavinsky, J. Kempe, I. Kerenidis, R. Raz, and R. de Wolf. Exponential separation for one-way quantum communication complexity, with applications to cryptography. *SIAM J. Comp.*, 38(5):1695–1708, 2008. Earlier version in STOC'07. quant-ph/0611209.

[GKRdW09] D. Gavinsky, J. Kempe, O. Regev, and R. de Wolf. Bounded-error quantum state identification and exponential separations in communication complexity. *SIAM J. Comp.*, 39(1):1–24, 2009. Special issue on STOC'06.

[GL08] N. Gvozdenovic and M. Laurent. The operator Ψ for the chromatic number of a graph. *SIAM J. Opt.*, 19(2):572–591, July 2008.

[Gle57] A. M. Gleason. Measures on the closed subspaces of a Hilbert space. *J. Math. Mech.*, 6:885–893, 1957.

[GN08] C. D. Godsil and M. W. Newman. Coloring an orthogonality graph. *SIAM J. Discret. Math.*, 22(2):683–692, 2008.

[Gop06] P. Gopalan. Constructing Ramsey graphs from boolean function representations. In *Proceedings of the 21st IEEE CCC*, 2006.

[GR01] C. Godsil and G. Royle. *Algebraic graph theory*, volume 207 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2001.

[GTHB05] O. Gühne, G. Tóth, P. Hyllus, and H. J. Briegel. Bell inequalities for graph states. *Phys. Rev. Lett.*, 95:120405, 2005.

[GTW10] V. Galliard, A. Tapp and S. Wolf. Deterministic quantum nonlocality and graph colorings. *Theor. Comp. Sci.*, 486:20–26, 2013.

[HEB04] M. Hein, J. Eisert, and H. J. Briegel. Multiparty entanglement in graph states. *Phys. Rev. A*, 69(6):062311–1–062311–20, June 2004.

[Hei71] W. Heisenberg. *Physics and Beyond: Encounters and Conversations*. Harper torchbooks. The Academy library. Allen and Unwin, 1971.

[Hog07] L. Hogben. Orthogonal representations, minimum rank, and graph complements, *Linear Algebra Appl.*, 428, 2008.

[JP11] M. Junge and C. Palazuelos. Large violation of Bell inequalities with low entanglement. *Comm. Math. Phys.*, 306(3):695–746, 2011. arXiv:1007.3043v2.

[JPP$^+$10] M. Junge, C. Palazuelos, D. Pérez-García, I. Villanueva, and M. Wolf. Unbounded violations of bipartite Bell inequalities via Operator Space theory. *Comm. Math. Phys.*, 300(3):715–739, 2010. Shorter version appeared in PRL 104:170405.

[KD93] D.E. Knuth and Stanford University. Computer Science Dept. *The sandwich theorem*. Stanford University, Dept. of Computer Science, 1993.

[Kho02] Subhash Khot. On the power of unique 2-prover 1-round games. In *Proceedings of 34th ACM STOC*, pages 767–775, 2002.

[KKL88] J. Kahn, G. Kalai, and N. Linial. The influence of variables on Boolean functions. In *Proceedings of 29th IEEE FOCS*, pages 68–80, 1988.

[KKM$^+$11] J. Kempe, H. Kobayashi, K. Matsumoto, B. Toner, and T. Vidick. Entangled games are hard to approximate. *SIAM J. Comp.*, 40(3):848–877, 2011. Preliminary version in FOCS'08.

[KKMV09] J. Kempe, H. Kobayashi, K. Matsumoto, and T. Vidick. Using entanglement in quantum multi-prover interactive proofs. *Computational Complexity*, 18(2):273–307, 2009. Preliminary version in Complexity 2008.

[KO98] J. Körner and A. Orlitsky. Zero-error information theory. *IEEE Trans. Inf. Theory*, 44(6):2207–2229, 1998.

[KRT10] J. Kempe, O. Regev, and B. Toner. Unique games with entangled provers are easy. *SIAM J. Comp*, 39(7):3207–3229, 2010. Preliminary version in FOCS'08. arXiv:0710.0655.

[KS67] S. Kochen and E. P. Specker. The problem of hidden variables in quantum mechanics. *J. of Mathematics and Mechanics*, 17:59–87, 1967.

[KV05] S. A. Khot and N. K. Vishnoi. The unique games conjecture, integrality gap for cut problems and embeddability of negative type metrics into ℓ_1. In *Proceedings of 46th IEEE FOCS*, pages 53–62, 2005.

[LMM$^+$12] D. Leung, L. Mancinska, W. Matthews, M. Ozols, and A. Roy. Entanglement can increase asymptotic rates of zero-error classical communication over classical channels. *Comm. Math. Phys.*, 311:97–111, 2012.

[Lov79] L. Lovász. On the Shannon capacity of a graph. *IEEE Trans. Inf. Theory*, 25(1):1–7, 1979.

[LT91] M. Ledoux and M. Talagrand. *Probability in Banach Spaces.* Springer, 1991.

[Lub07] E. Lubetzky. *Graph Powers and Related Extremal Problems.* PhD thesis, Tel Aviv University, 2007.

[Meu05] P. Meurdesoif. Strengthening the Lovász Theta(G) bound for graph coloring. *Math. Program.*, 102(3):577–588, 2005.

[MO] L. Mančinska and M. Ozols. private communications.

[MSS13] L. Mančinska, G. Scarpa, and S. Severini. New separations in zero-error channel capacity through projective Kochen-Specker sets and quantum coloring. *IEEE Trans. Inf. Theory*, 59(6):4025–4032, 2013.

[NC00] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information.* Cambridge University Press, New York, 2000.

[NTR06] J. Nayak, E. Tuncel, and K. Rose. Zero-error source-channel coding with side information. *IEEE Trans. Inf. Theory*, 52(10):4626–4629, 2006.

[O'D08] R. O'Donnell. Some topics in analysis of boolean functions. Technical report, ECCC Report TR08–055, 2008. Paper for an invited talk at STOC'08.

[Pai79] A. Pais. Einstein and the quantum theory. *Reviews of Modern Physics*, 51(4):863–914, October 1979.

[Pal33] R. E. A. C. Paley. On orthogonal matrices. *Journal of Mathematics and Physics (now called Studies in Applied Mathematics)*, 12:311–320, 1933.

[Pal12a] C. Palazuelos. Bounding the largest Bell violation of the maximally entangled state. Preprint, 2012.

[Pal12b] C. Palazuelos. Super-activation of quantum non-locality. *Phys. Rev. Lett.*, 109, 190401 (2012)

[Per91] A. Peres. Two simple proofs of the Kochen-Specker theorem. *J. Phys. A*, 24:L175–L178, 1991.

[PMMM05] M. Pavicic, J. Merlet, B. McKay, and N. D. Megill. Kochen-Specker vectors. *J. Phys. A*, 38:1577, 2005.

[PWP$^+$08] D. Pérez-García, M. M. Wolf, C. Palazuelos, I. Villanueva, and M. Junge. Unbounded violations of tripartite Bell inequalities. *Comm. Math. Phys.*, 279:455, 2008. quant-ph/0702189.

[Reg12] O. Regev. Bell violations through independent bases games. *Quant. Inf. Comp.*, 12(1):9–20, 2012. arXiv:1101.0576.

[RM12] D. E. Roberson and L. Mancinska. Graph Homomorphisms for Quantum Players. 2012. arXiv:1212.1724.

[RW04] R. Renner and S. Wolf. Quantum pseudo-telepathy and the Kochen-Specker theorem. In *Proc. Int. Symp. Inf. Theory*, pages 322–329, 2004.

[Sha56] C. E. Shannon. The zero error capacity of a noisy channel. IT-2(3):8–19, September 1956.

[SS12] G. Scarpa and S. Severini. Kochen-Specker sets and the rank-1 quantum chromatic number. *IEEE Trans. Inf. Theory*, 58(4):2524–2529, 2012.

[SW01] D. Schlingemann and R. F. Werner. Quantum error-correcting codes associated with graphs. *Phys. Rev. A*, 65:012308, Dec 2001.

[Sze94] M. Szegedy. A note on the ϑ number of Lovász and the generalized Delsarte bound. *Proceedings of the 35th IEEE FOCS*, pages 36–39, 1994.

[Tsi87] B. S. Tsirelson. Quantum analogues of the Bell inequalities. the case of two spatially separated domains. *Journal of Soviet Mathematics*, 36:557–570, 1987.

[Viz63] V. G. Vizing. The Cartesian product of graphs. *Vycisl. Sistemy*, 9:30–43, 1963.

[VVS95] S. Vembu, S. Verdu, and Y. Steinberg. The source-channel separation theorem revisited. *IEEE Trans. Inf. Theory*, 41(1):44–54, 1995.

[Wit76] H. S. Witsenhausen. The zero-error side information problem and chromatic numbers. *IEEE Trans. Inf. Theory*, 22(5):592–593, 1976.

[Wol08] R. de Wolf. A brief introduction to Fourier analysis on the Boolean cube. *Theory of Computing*, ToC Library, Graduate Surveys 1, 2008.

[WX97] R.M. Wilson and Q. Xiang. Constructions of Hadamard difference sets. *Journal of combinatorial theory. Series A*, 77(1):148–160, 1997.

[Xia98] Q. Xiang. Difference families from lines and half lines. *European J. of Combinatorics*, 19(3):395–400, 1998.

[XL91] M. Xia and G. Liu. An infinite class of supplementary difference sets and Williamson matrices. *J. of Combinatorial Theory, Series A*, 58(2):310–317, 1991.

[XL96] M. Xia and G. Liu. A new family of supplementary difference sets and Hadamard matrices. *Journal of Statistical Planning and Inference*, 51(3):283–291, 1996.

[XSX06] T. Xia, J. Seberry, and M. Xia. New constructing of regular Hadamard matrices. In *Proceedings of the 10th ICCOMP* pages 1294–1299, 2006.

[Yek12] S. Yekhanin. Locally decodable codes. *Foundations and Trends in Theoretical Computer Science*, 6(3):139–255, 2012.

Index

Samenvatting

In dit proefschrift bestuderen we kwantumverstrengeling, en enkele toepassingen in grafentheorie en *zero-error* informatietheorie.

In Hoofdstuk 1 introduceren we kwantumverstrengeling en andere fundamentele concepten uit de kwantumtheorie. Kwantumverstrengeling werd voor het eerst beschreven in 1935 door Einstein, Podolsky en Rosen, die inzagen dat verstrengeling correlaties tussen deeltjes toestaat die sterker zijn dan mogelijk is in het klassieke geval. Zij geloofden aanvankelijk niet dat zulke correlaties in de werkelijkheid konden bestaan, en interpreteerden dit als bewijs voor de onvolledigheid of incorrectheid van de kwantummechanica. De volgende stap werd genomen in 1964 door Bell, die een experiment voorstelde om te testen of zulk niet-klassiek gedrag kan voorkomen in de realiteit. Hij toonde aan dat klassieke invoer-uitvoercorrelaties een bepaalde ongelijkheid vervullen, die sindsdien de ongelijkheid van Bell wordt genoemd. Vervolgens liet hij zien dat de kwantummechanica correlaties toestaat die deze ongelijkheid overtreden. Aspect *et al.* waren in de vroege jaren tachtig voor het eerst in staat een dergelijk experiment uit te voeren, en demonstreerden een geval waarin de ongelijkheid van Bell overtreden wordt. Dus de wereld volgt niet de wetten van de klassieke natuurkunde!

We maken in dit proefschrift gebruik van kwantumverstrengeling als een hulpmiddel voor verscheidene informatieverwerkingstaken. In Hoofdstuk 2 behandelen we de vraag *in welke mate* kwantumcorrelaties kunnen afwijken van klassieke voorspellingen. We richten ons specifiek op *nonlokale spelen*: experimenten waarin twee spelers, die van elkaar gescheiden zijn en niet mogen communiceren, samen moeten werken om een taak te volbrengen. In dit geval geeft de ongelijkheid van Bell een bovengrens aan de maximale kans van slagen van klassieke spelers. We zien een overtreding van de ongelijkheid van Bell als kwantumspelers een verstrengelde toestand gebruiken om hun kans op succes te vergroten. In 2009 gaven Junge *et al.* bovengrenzen aan overtredingen van de ongelijkheid van Bell. Zij bewezen dat de maximale overtreding afhangt van het aantal mogelijke in- en uitvoervariabelen van elke speler, en van de dimensie van de verstrengelde toestand. Dit proefschrift geeft *ondergrenzen* in de vorm van twee spelen die overtredingen laten zien die zeer dicht bij de ondergrenzen van Junge *et al.* liggen. Het is interessant om op te merken dat onze resultaten in de theo-retische natuurkunde zijn geïnspireerd door de theoretische informatica. Het eerste nonlokale spel dat we beschrijven is gebaseerd op een communicatiecomplexiteitsprobleem dat bekend staat als *hidden matching*, terwijl het tweede spel voortkomt uit het werk van Khot en Vishnoi omtrent de bekende *Unique Games Conjecture* uit de computationele complexiteitstheorie.

Hoofdstuk 3 is gewijd aan de studie van *kwantumgraafparameters*. Bekende grootheden zoals het chromatische getal en het onafhankelijksgetal van een graaf kunnen worden geïnterpreteerd als parameters voor nonlokale spelen. Een voorbeeld is een "scheidsrechter" die kan testen of twee spelers, Alice en Bob, een k-kleuring van een graaf G hebben (d.w.z., een toekenning van één uit k kleuren aan elke knoop van de graaf zodanig dat incidente knopen verschillende kleuren krijgen). Hij haalt de twee spelers uit elkaar en vraagt beide een kleur (van de k mogelijke kleuren) te noemen van een knoop van G. Als hij Alice en Bob dezelfde knoop zou geven, dan verwacht hij dezelfde kleur terug te krijgen; als hij ze twee incidente knopen geeft, dan verwacht hij twee verschillende kleuren. Het chromatische getal van G is het kleinste aantal kleuren k zodanig dat Alice en Bob een klassieke strategie hebben die altijd slaagt. Analoog daaraan is het *kwantum*chromatische getal van graaf G de kleinste k zodanig dat Alice en Bob altijd winnen als ze kwantumverstrengeling mogen gebruiken. Voor sommige grafen is het laatste getal strikt kleiner dan het klassieke getal. De studie van kwantumgraafparameters, hoewel al impliciet in eerder werk, nam een serieuze aanvang met twee artikelen van Cameron *et al.* in 2010 en Cubitt *et al.* in 2009. Dit proefschrift vormt een bijdrage aan dit onderzoek in een aantal opzichten. We bestuderen de relatie tussen de kwantumgraafparameters en andere parameters zoals het Lovász ϑ-getal en de orthogonale rang (*orthogonal rank*). We beschrijven een verrassende karakterisering van de grafen met afwijkende kwantum- en klassieke chromatische getallen. deze is gerelateerd aan de stelling van Kochen-Specker, een resultaat uit 1967 binnen de grondslagen van de kwantummechanica. We geven verscheidene constructies van grafen met afwijkende kwantum- en klassieke chromatische getallen, bijvoorbeeld gebaseerd op graafproducten en graaftoestanden. Ten slotte gebruiken we kwantumgraafparameters om grenzen voor de waardes van algemene nonlokale spelen te vinden.

In Hoofdstuk 4 verleggen we onze aandacht naar de *zero-error informatietheorie*. We bestuderen de *zero-error*-capaciteit van een klassiek kanaal met ruis als de zender en ontvanger kwantumverstrengeling mogen gebruiken. Daarnaast bestuderen we het *broncoderingsprobleem* (het *source problem*), waar de ontvanger de beschikking heeft over partiële informatie over het te ontvangen bericht, en het gecombineerde *bron-kanaalprobleem* (het *source-channel problem*). De belangrijkste gereedschappen in dit domein zijn grafen en hun parameters, waardoor zero-error informatietheorie met verstrengeling veel baat heeft van de concepten die het onderwerp waren van de voorgaande hoofdstukken. Zo kan de capaciteit van een kanaal berekend worden met het onafhankelijkheidsgetal van een graaf, en is het broncoderingsprobleem verbonden met het chromatische getal. We definiëren in dit hoofdstuk het *verstrengelde* chromatische getal en het *verstrengelde* onafhankelijksheidsgetal, naast andere, verwante grootheden. De precieze relatie met de parameters van Hoofdstuk 3 vormt nog een open probleem. We initiëren de studie van het broncoderingsprobleem en het bron-kanaalprobleem met verstrengeling en we vinden kanalen en bronnen die een sterke divergentie in kwantum- en klassiek gedrag laten zien. Daartoe gebruiken we resultaten uit combinatoriek, lineaire algebra, optimalisatie en getaltheorie.

Dit proefschrift is gebaseerd op de volgende artikelen:

1. *"Near-optimal and explicit Bell inequality violations"*, by H. Buhrman, O. Regev, the author and R. de Wolf. Theory of Computing, december 2012.

2. *"Kochen-Specker Sets and the Rank-1 Quantum Chromatic Number"*, by the author and S. Severini. IEEE Transactions on Information Theory, april 2012.

3. *"New Separations in Zero-error Channel Capacity through Projective Kochen-Specker Sets and Quantum Coloring"*, by L. Mančinska, the author and S. Severini. IEEE Transactions on Information Theory, juni 2013.

4. *"Exclusivity structures and graph representatives of local complementation orbits"*, by A. Cabello, M. G. Parker, the author and S. Severini. Journal of Mathematical Physics, juli 2013.

5. *"Zero-error source-channel coding with entanglement"*, by J. Briët, H. Buhrman, M. Laurent, T. Piovesan and the author. Proceedings of Eurocomb'13, september 2013.

Abstract

We study quantum entanglement and some of its applications in graph theory and zero-error information theory.

In Chapter 1 we introduce entanglement and other fundamental concepts of quantum theory. Entanglement was first described in 1935 by Einstein, Podolsky and Rosen, who observed that it allowed for stronger-than-classical correlations between distant particles. They did not believe such correlations existed in nature, and interpreted this as evidence for incompleteness or incorrectness of quantum mechanics. Bell took the next step in 1964, proposing an experiment to test whether such non-classical behavior occurs in nature. He showed that classical input-output correlations satisfy a certain inequality, called a Bell inequality. Then, he showed that quantum mechanics allows for correlations that violate it. Aspect *et al.*, in the early 1980s, realized such an experiment for the first time, showing that Nature violated a Bell inequality. Therefore, Nature does not follow the rules of classical physics!

In this thesis we make use of quantum entanglement as a resource for various information-processing tasks. In Chapter 2 we address the question of *how much* quantum correlations can deviate from classical predictions. We focus on *non-local games*: experiments in which two players are separated and forbidden to communicate, and have to collaborate to accomplish a task. In this case, a Bell inequality is an upper bound on the maximum success probability of classical players. We have a Bell inequality violation when quantum players use an entangled state to achieve a larger success probability. Upper bounds on Bell inequality violation of Gs were given by Junge *et al.* in 2009. They proved that the maximum violation depends on the number of possible inputs and outputs of each player, and on the dimension of the entangled state. We give *lower bounds* in the form of two games exhibiting violations that are very close to the upper bounds of Junge *et al.* Remarkably, our results in theoretical physics are inspired by theoretical computer science. The first non-local game is based on a communication complexity problem called "hidden matching", while the second one derives from the work of Khot and Vishnoi on the famous Unique Games Conjecture from computational complexity theory.

Chapter 3 is dedicated to the study of *quantum graph parameters*. Interestingly, well-known quantities such as the chromatic number and the independence number of a graph can be interpreted as parameters for non-local games. For example, a "referee" can test if two players, Alice and Bob, have a k-coloring of a graph G, *i.e.*, an assignment of one out of k colors for each vertex of the graph such that adjacent vertices get different colors. He separates the two players, gives each of them a vertex of G, and asks them for the color (out of the k possible) of that vertex. If he gave Alice and Bob the same vertex, he expects to receive back the same color, if he gave them two adjacent vertices, he expects to get two different colors. The minimum number of colors k such that Alice and Bob have a classical strategy that always succeeds is the chromatic number of G. If Alice and Bob can use quantum entanglement, then the minimum k such that they always win is called the *quantum* chromatic number of the graph. For some graphs, it can be strictly smaller than the classical one. The study of quantum graph parameters officially started with two papers by Cameron *et al.* in 2010 and Cubitt *et al.* in 2009, but was implicit in earlier work. We contribute to the field in a number of ways. We study the relationship between these quantum graph parameters and other parameters such as the Lovász ϑ number and the orthogonal rank. We find a surprising characterization of the graphs having a separation between the quantum and classical chromatic numbers. This is related to the Kochen-Specker theorem, a result in the foundations of quantum mechanics from 1967. We also find various constructions of graphs that feature separations between the quantum and classical independence numbers, based for example on graph products and graph states. Additionally, we use quantum graph parameters to give bounds on the value of general non-local games.

In Chapter 4, we move to *zero-error information theory*. We study the *zero-error capacity* of a classical noisy channel when the sender and the receiver can use quantum entanglement. We also study the *source problem*, where the receiver has partial information about the message that is going to be delivered, as well as the combined *source-channel problem*. The main tools in this field are graphs and their parameters, so zero-error information theory with entanglement is a fertile field for the concepts studied in previous chapters. For example, the capacity of a channel can be calculated using the independence number of a graph and the source-problem is related to the chromatic number of a graph. We define the *entangled* chromatic number, the *entangled* independence number and other related quantities. The exact relation with the parameters of Chapter 3 is still an open problem. We initiate the study of the source problem and source-channel problem with entanglement and we find channels and sources that exhibit a strong divergence in quantum and classical behaviors. To do that, we use results in combinatorics, linear algebra, optimization and number theory.

This thesis is based on the following articles:

- *"Near-optimal and explicit Bell inequality violations"*, by H. Buhrman, O. Regev, the author and R. de Wolf. Theory of Computing, December 2012.
- *"Kochen-Specker Sets and the Rank-1 Quantum Chromatic Number"*, by the author and S. Severini. IEEE Transactions on Information Theory, April 2012.
- *"New Separations in Zero-error Channel Capacity through Projective Kochen-Specker Sets and Quantum Coloring"*, by L. Mančinska, the author and S. Severini. IEEE Transactions on Information Theory, June 2013.
- *"Exclusivity structures and graph representatives of local complementation orbits"*, by A. Cabello, M. G. Parker, the author and S. Severini. Journal of Mathematical Physics, July 2013.
- *"Zero-error source-channel coding with entanglement"*, by J. Briët, H. Buhrman, M. Laurent, T. Piovesan and the author. Proceedings of Eurocomb'13, September 2013.

Titles in the ILLC Dissertation Series:

ILLC DS-2009-01: **Jakub Szymanik**
Quantifiers in TIME and SPACE. Computational Complexity of Generalized Quantifiers in Natural Language

ILLC DS-2009-02: **Hartmut Fitz**
Neural Syntax

ILLC DS-2009-03: **Brian Thomas Semmes**
A Game for the Borel Functions

ILLC DS-2009-04: **Sara L. Uckelman**
Modalities in Medieval Logic

ILLC DS-2009-05: **Andreas Witzel**
Knowledge and Games: Theory and Implementation

ILLC DS-2009-06: **Chantal Bax**
Subjectivity after Wittgenstein. Wittgenstein's embodied and embedded subject and the debate about the death of man.

ILLC DS-2009-07: **Kata Balogh**
Theme with Variations. A Context-based Analysis of Focus

ILLC DS-2009-08: **Tomohiro Hoshi**
Epistemic Dynamics and Protocol Information

ILLC DS-2009-09: **Olivia Ladinig**
Temporal expectations and their violations

ILLC DS-2009-10: **Tikitu de Jager**
"Now that you mention it, I wonder...": Awareness, Attention, Assumption

ILLC DS-2009-11: **Michael Franke**
Signal to Act: Game Theory in Pragmatics

ILLC DS-2009-12: **Joel Uckelman**
More Than the Sum of Its Parts: Compact Preference Representation Over Combinatorial Domains

ILLC DS-2009-13: **Stefan Bold**
Cardinals as Ultrapowers. A Canonical Measure Analysis under the Axiom of Determinacy.

ILLC DS-2010-01: **Reut Tsarfaty**
Relational-Realizational Parsing

ILLC DS-2010-02: **Jonathan Zvesper**
Playing with Information

ILLC DS-2010-03: **Cédric Dégremont**
The Temporal Mind. Observations on the logic of belief change in interactive systems

ILLC DS-2010-04: **Daisuke Ikegami**
Games in Set Theory and Logic

ILLC DS-2010-05: **Jarmo Kontinen**
Coherence and Complexity in Fragments of Dependence Logic

ILLC DS-2010-06: **Yanjing Wang**
Epistemic Modelling and Protocol Dynamics

ILLC DS-2010-07: **Marc Staudacher**
Use theories of meaning between conventions and social norms

ILLC DS-2010-08: **Amélie Gheerbrant**
Fixed-Point Logics on Trees

ILLC DS-2010-09: **Gaëlle Fontaine**
Modal Fixpoint Logic: Some Model Theoretic Questions

ILLC DS-2010-10: **Jacob Vosmaer**
Logic, Algebra and Topology. Investigations into canonical extensions, duality theory and point-free topology.

ILLC DS-2010-11: **Nina Gierasimczuk**
Knowing One's Limits. Logical Analysis of Inductive Inference

ILLC DS-2010-12: **Martin Mose Bentzen**
Stit, Iit, and Deontic Logic for Action Types

ILLC DS-2011-01: **Wouter M. Koolen**
Combining Strategies Efficiently: High-Quality Decisions from Conflicting Advice

ILLC DS-2011-02: **Fernando Raymundo Velazquez-Quesada**
Small steps in dynamics of information

ILLC DS-2011-03: **Marijn Koolen**
The Meaning of Structure: the Value of Link Evidence for Information Retrieval

ILLC DS-2011-04: **Junte Zhang**
System Evaluation of Archival Description and Access

ILLC DS-2011-05: **Lauri Keskinen**
Characterizing All Models in Infinite Cardinalities

ILLC DS-2011-06: **Rianne Kaptein**
Effective Focused Retrieval by Exploiting Query Context and Document Structure

ILLC DS-2011-07: **Jop Briët**
Grothendieck Inequalities, Nonlocal Games and Optimization

ILLC DS-2011-08: **Stefan Minica**
Dynamic Logic of Questions

ILLC DS-2011-09: **Raul Andres Leal**
Modalities Through the Looking Glass: A study on coalgebraic modal logic and their applications

ILLC DS-2011-10: **Lena Kurzen**
Complexity in Interaction

ILLC DS-2011-11: **Gideon Borensztajn**
The neural basis of structure in language

ILLC DS-2012-01: **Federico Sangati**
Decomposing and Regenerating Syntactic Trees

ILLC DS-2012-02: **Markos Mylonakis**
Learning the Latent Structure of Translation

ILLC DS-2012-03: **Edgar José Andrade Lotero**
Models of Language: Towards a practice-based account of information in natural language

ILLC DS-2012-04: **Yurii Khomskii**
Regularity Properties and Definability in the Real Number Continuum: idealized forcing, polarized partitions, Hausdorff gaps and mad families in the projective hierarchy.

ILLC DS-2012-05: **David García Soriano**
Query-Efficient Computation in Property Testing and Learning Theory

ILLC DS-2012-06: **Dimitris Gakis**
Contextual Metaphilosophy - The Case of Wittgenstein

ILLC DS-2012-07: **Pietro Galliani**
The Dynamics of Imperfect Information

ILLC DS-2012-08: **Umberto Grandi**
Binary Aggregation with Integrity Constraints

ILLC DS-2012-09: **Wesley Halcrow Holliday**
Knowing What Follows: Epistemic Closure and Epistemic Logic

ILLC DS-2012-10: **Jeremy Meyers**
Locations, Bodies, and Sets: A model theoretic investigation into nominalistic mereologies

ILLC DS-2012-11: **Floor Sietsma**
Logics of Communication and Knowledge

ILLC DS-2012-12: **Joris Dormans**
Engineering emergence: applied theory for game design

ILLC DS-2013-01: **Simon Pauw**
Size Matters: Grounding Quantifiers in Spatial Perception

ILLC DS-2013-02: **Virginie Fiutek**
Playing with Knowledge and Belief

ILLC DS-2013-03: **Giannicola Scarpa**
Quantum entanglement in non-local games, graph parameters and zero-error information theory