Quantum entanglement: insights via graph parameters and conic optimization

Teresa Piovesan

Quantum entanglement: insights via graph parameters and conic optimization

ILLC Dissertation Series DS-2016-09

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/

The investigations were performed at the Centrum Wiskunde & Informatica (CWI). They were partially supported by SIQS European project and QuSoft.

Printed and bound by Ipskamp Drukkers.

ISBN: 978–94–028–0348–8

Quantum entanglement: insights via graph parameters and conic optimization

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. K.I.J. Maex
ten overstaan van een door het College voor
Promoties ingestelde commissie, in het openbaar
te verdedigen in de Agnietenkapel
op donderdag 27 oktober 2016, te 10.00 uur

door

Teresa Piovesan

geboren te San Doná di Piave, Italië.

Promotiecommissie:

Promotores:

Prof. dr. H.M. Buhrman	CWI & Universiteit van Amsterdam
Prof. dr. M. Laurent	CWI & Tilburg University

Overige leden:

Dr. J. Briët	CWI
Dr. D.C. Gijswijt	TU Delft
Prof. dr. J.D.M. Maassen	Universiteit van Amsterdam
Prof. dr. E.M. Opdam	Universiteit van Amsterdam
Dr. C. Schaffner	CWI & Universiteit van Amsterdam
Prof. dr. A. Winter	Universitat Autónoma de Barcelona
Prof. dr. R.M. de Wolf	CWI & Universiteit van Amsterdam

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

Contents

Acknowledgments

First and foremost, I would like to thank my PhD supervisors, Harry Buhrman and Monique Laurent, for their guidance and support throughout the last four years. It has been a great pleasure working with the two of you. Harry, your passion for science is truly admirable, doing research with you has been rewarding and great fun. Monique, thank you so much for your patient and thorough supervision, it was extremely insightful. But most of all, thanks for your trust and encouragement. I want to thank Michele Conforti for various life advices throughout the years, the one about pursuing a PhD with Monique was especially a good one! For willing to be on my PhD committee and for helpful comments on this thesis, I thank Jop Briët, Harry Buhrman, Dion Gijswijt, Monique Laurent, Hans Maassen, Eric Opdam, Christian Schaffner, Andreas Winter and Ronald de Wolf.

I greatly enjoyed working and interacting with my colleagues at CWI and for this I want to thank the various members of the Algorithms & Complexity and Networks & Optimization groups. A special thank goes to Jop Briët for being so enthusiastic and helpful from the very beginning of my PhD and to Ronald de Wolf for his sharp advices about giving presentations and scientific writing. Most of all, I am indebted with my coauthors: Emma Bauxis-Aussalet, Jop Briët, Harry Buhrman, Sabine Burgdorf, Monique Laurent, Debbie Leung, Erik Quaeghebeur, Giannicola Scarpa, Christian Schaffner, Florian Speelman, Tom Sterkenburg and Chris Wesseling.

During my time in Amsterdam I met some great people. Among others I want to thank Gianni, Fernando, Gabriele, Giulia, Matteo, Deba, Emma, Tom, Michael, Eleni and the entire Prague team. A very special thank goes to Pablo and Valerio for being great friends. I will cherish all the amazing memories.

My staying in Amsterdam would have been very different if it wasn't for Carlo, Candida and Leo. Thank you for all the dinners, coffees, breakfasts, random conversations and the inevitable cakes. Mi mancherete piú di quanto non voglia ammettere.

Huge thanks to all my friends and in particular to Giorgia and the Pearson gang for keeping me grounded.

Finally, I want to thank my family: Lucia, Lucio, Carlo, Davide, Leo, Alessandra, Candida, Ivana, Franca for the never ending supply of life lessons.

Amsterdam Teresa Piovesan
September, 2016.

Chapter 1

Introduction

Our daily experiences, whether we are riding a bike or lazily floating on the water, can be perfectly explained using the laws of classical mechanics. However, if we want to predict the behavior of small particles, such as atoms and photons, a different mathematical framework is needed. The development of this new model of nature known as *quantum mechanics*, has been one of the most important scientific paradigm shifts of the last century. A very peculiar and unique feature of this theory is that a quantum state can be in a *superposition* of various different states at the same time. Moreover, an ensemble of spatially separated quantum systems can be *entangled* and operations on a single quantum system can influence the state of the other ones.

Large parts of this thesis are devoted to the study of the effects of *quantum entanglement* in *nonlocal games* and communication problems in *zero-error information theory*, using *graph parameters* and tools from *conic optimization*.

Nonlocal games. A *nonlocal game* is a thought physical experiment in which two or more cooperating players, who can agree on a strategy but not exchange information, interact with an extra party, usually called the referee. At the beginning of the game, the referee sends to each player a question to which they have to reply with an answer. Based on the questions asked and the answers received, the referee decides if the players win or lose. The predicate that decides whether the game is won or lost is known to all the parties involved beforehand. However, the players only know the question that was asked directly to them, not the ones aimed to the other players. As the players cannot communicate during the game, they can only use prearranged strategies to coordinate their answers with the goal of maximizing the chances of winning. If the players can only use the laws of classical mechanics, the optimal course of action is to agree on how to answer to each question. More sophisticated strategies can be implemented if the players have access to entangled physical systems. In this case each of the players bases its answer on the outcome

of an experiment performed on their private system. As it was first noticed by Bell [Bel64] using a slightly different language, this type of strategies can produce answers that are correlated in a way that cannot be obtained in a classical world. The existence of these non-classical correlations is also supported by increasingly convincing experimental evidences [ADR82, HBD+15]. Moreover there exist games, such as the CHSH one (named after Clauser, Horne, Shimony, and Holt [CHSH69]), for which the maximal probability of winning using a quantum strategy is strictly larger than the one using a classical strategy [CHTW04]. A *perfect* classical, or quantum, strategy is one that guarantees the players to win on any possible set of questions. For a fixed game, we will focus on the problem: *Does a perfect classical, or quantum, strategy exist?*

Zero-error information theory. Information theory is a mathematical field that studies the way information can be communicated and stored. The foundations of this subject were laid by Shannon in the paper "A Mathematical Theory of Communication" [Sha48]. One of the main tasks Shannon considered was the channel coding problem, where a sender wants to communicate messages over a noisy channel in a way that allows the receiver to reconstruct the messages with low probability of error. In a follow-up paper, Shannon [Sha56] studied the same problem but now without tolerating any error: the transmission of the message must be error free. This paper started the field of *zero-error information theory*, which studies various communication problems where no error is allowed and which has developed into a large research area involving information theory, combinatorics, computer science, and mathematical programming (see for example the survey of Körner and Orlitsky [KO98] and Lubetzky's PhD thesis [Lub07]). In this thesis we will approach various zero-error classical communication problems with two main questions in mind: *Given a classical communication task, does entanglement allow for communication schemes that are better than the classical ones?* and *How much more efficient can the communication be when quantum states are transmitted rather than classical ones?*

Graph parameters. The unifying link among the various problems that we study in this thesis is their combinatorial nature. Indeed, the majority of them will have a graph theoretical formulation, mainly concerning the chromatic and stability numbers and some quantum generalizations thereof. For instance, we will consider a nonlocal game in which two players want to convince a referee that they can color a given graph G using at most t colors. (A coloring of a graph is an assignment of colors to the vertex set such that adjacent vertices receive different colors.) At the start of the game, the referee sends to each player a vertex of the graph as question, to which they have to answer with a color from $\{1, \ldots, t\}$. The players win the game if they answer the same color upon receiving the same vertex and different colors upon receiving adja-

cent vertices. One can easily show that the chromatic number of the graph G is the minimum $t \in \mathbb{N}$ for which there exists a perfect classical strategy for this nonlocal game (see Section 3.3.1 for details). Analogously, the quantum chromatic number of the graph G is the minimum $t \in \mathbb{N}$ for which there exists a perfect quantum strategy. This parameter can be equivalently reformulated as minimum $t \in \mathbb{N}$ for which there exists a collection of positive semidefinite matrices satisfying certain linear and orthogonality constraints (which only depend on the graph G).

An extensive overview of the various graph parameters considered in this thesis is given in Chapter 3.

Conic optimization. Many hard combinatorial problems, as for example the chromatic and stability numbers, can be reformulated as linear optimization programs over an appropriate convex cone $\mathcal{K}$. In this way the complexity of the problem is pushed to the cone $\mathcal{K}$ and one can exploit the properties of the cone to study the original problem. For instance, approximations can be built by replacing $\mathcal{K}$ with a hierarchy of linear or semidefinite subcones (or supercones). In this thesis, we define a new matrix cone, the completely positive semidefinite cone, to be able to reformulate some quantum generalizations of the classical graph parameters as conic optimization programs.

1.1 Overview

Whenever one wants to study the effect of quantum entanglement on a classical problem, whether this is a nonlocal game or a communication problem, there are two main questions that naturally arise. What are the intrinsic mathematical properties and differences between the classical and entanglement-assisted scenarios? Can entanglement give an advantage?

Our contribution to the first question is a novel approach to the study of quantum strategies using the paradigm of conic optimization. We introduce the completely positive semidefinite cone $\mathcal{CS}^n_+$, a new matrix cone consisting of all $n \times n$ symmetric matrices that admit a Gram representation by positive semidefinite matrices (i.e., $A \in \mathcal{CS}^n_+$ if there exists a family of positive semidefinite matrices $X_1, \ldots, X_n \in \mathcal{S}^d_+$, for some $d \in \mathbb{N}$, such that $A = (\langle X_i, X_j \rangle)^n_{i,j=1}$), and use it to model quantum variants of classical parameters. Chapter 4 will be entirely dedicated to the study of the completely positive semidefinite cone and some of its structural properties. Moreover, we will investigate its strong ties with a well-studied cone: the completely positive cone $\mathcal{CP}^n$, which consists of the set of $n \times n$ matrices admitting a Gram representation by nonnegative scalars (i.e., restricting to the case where $d = 1$ in the above definition). Using ideas that have been developed to approximate the completely positive

cone, we will construct two hierarchies, a semidefinite and a linear one, that approach the dual of the completely positive semidefinite cone.

In Chapter 5 we will draw the connection between the completely positive semidefinite cone and the quantum graph parameters. In a nutshell, a quantum graph parameter, such as the quantum chromatic number, can be defined as the minimum (or maximum) integer of which there exists a collection of positive semidefinite matrices satisfying some linear and orthogonality constraints. We can then reformulate it as a conic linear program over the completely positive semidefinite cone. Moreover, approximations can be obtained applying the above mentioned hierarchies.

Chapters 6-9 will mainly focus on the second question by finding separations between classical and quantum strategies in some standard problems from information theory.

In Chapter 6 we study the *channel coding problem*, which asks a sender to transmit data reliably to a receiver in the presence of noise. If we want the receiver to recover the message with a probability of error that asymptotically goes to zero (as the number of channel uses tends to infinity), then Bennett, Shor, Smolin, and Thapliyal [BSST02] proved that entanglement cannot provide any advantage. We thus focus on the zero-error case, where the transmission has to happen error free. Building on the work of Briët, Buhrman and Gijswijt [BBG12] we will exhibit an infinite family of channel coding problems for which the entanglement-assisted strategies are strictly better than the classical ones. The main contribution is a novel entanglement-assisted channel coding protocol that uses remote state preparation.

In Chapter 7 we study two generalizations of the channel coding problem to multiparty settings. In the first scenario we consider there is one sender who wants to transmit a common message to multiple receivers; in the second one we have multiple collaborating senders that want to communicate a message to a single receiver. We will prove some separation results as well as show limitations of the entangled strategies. Moreover, we will show that entanglement allows for a peculiar amplification of information which cannot happen classically.

In Chapter 8 we study the *source coding problem*, where a sender has to efficiently communicate data about which a receiver has already some information, and the *source-channel coding problem* which is a combination of the source and the channel coding problem. Here the sender can only use a noisy channel to communicate the data to the receiver. For both these problems, using families of problems that are related to the one studied in Chapter 6, we will show that entanglement allows for strategies that are exponentially better than the classical ones.

Finally, in Chapter 9 we consider two communication complexity problems: the *promise equality* and the *list* problems. We are interested in the minimum

number of classical, or quantum, messages that have to be exchanged between two parties (Alice and Bob) to be able to solve the problem without error, especially when making the distinction between one-round communication complexity, where the communication flows from Alice to Bob, and multi-round communication complexity, where the parties take turns in the transmission of the messages.

In the promise equality problem, Alice and Bob must decide whether their inputs are equal or not. An instance of this problem was used by Buhrman, Cleve, and Wigderson [BCW98] to show the first large gap between classical and quantum communication complexity. Here, for a different promise equality problem, we prove that there exists an exponential gap between the one-round and the two-round quantum communication complexity.

In the list problem Bob gets a list, Alice gets an element from Bob's list and their goal is for Bob to learn Alice's element. We will show various results, regarding both the classical and the quantum communication complexity.

Chapter 2

Preliminaries

We start by introducing some basic notions that are used in this thesis. In this chapter we will give a brief introduction to the necessary concepts from linear algebra, conic optimization, and quantum information theory. In Chapter 3, we will give an overview of the needed notions from graph theory. Throughout, we will use standard notation. The reader can find a List of Symbols at the end of the thesis where the notation is defined.

2.1 Matrices

We begin with some standard definitions and properties of matrices. The interested reader can find the omitted proofs in the book of Horn and Johnson [HJ12].

We denote by $\mathcal{S}^n$ the set of $n \times n$ real symmetric matrices, which is equipped with the standard trace inner product: $\langle A, B \rangle = \mathsf{Tr}(AB) = \sum_{i,j=1}^{n} A_{ij}B_{ij}$ and the corresponding Frobenius norm: $\|A\|_{\mathrm{F}} = \sqrt{\langle A, A \rangle}$. The trace of matrix A is defined as the sum of the elements on the main diagonal; i.e., $\mathsf{Tr}(A) = \sum_i A_{ii}$. The trace is a linear mapping and is invariant under cyclic permutations; i.e., for any $A, B \in \mathcal{S}^n$ and scalar $\lambda \in \mathbb{R}$, we have $\mathsf{Tr}(A + B) = \mathsf{Tr}(A) + \mathsf{Tr}(B)$, $\mathsf{Tr}(\lambda A) = \lambda \mathsf{Tr}(A)$ and $\mathsf{Tr}(AB) = \mathsf{Tr}(BA)$. The *rank* of a matrix A, denoted by $\mathrm{rank}(A)$, is the largest number of linearly independent columns. The *image* of a matrix $A \in \mathcal{S}^n$ is the set of all vectors $Ax \in \mathbb{R}^n$ for $x \in \mathbb{R}^n$.

A matrix $Q \in \mathbb{R}^{n \times n}$ is said to be *orthogonal* if $QQ^T = I$ (or equivalently $Q^TQ = I$), which means that the rows (respectively, the columns) of Q form an orthonormal basis of $\mathbb{R}^n$. The *real spectral decomposition theorem* says that any real symmetric matrix $A \in \mathcal{S}^n$ can be decomposed as $A = \sum_{i \in [n]} \lambda_i v_i v_i^T$ where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A and $v_1, \ldots, v_n \in \mathbb{R}^n$ are the corresponding eigenvectors. Equivalently, $A = QDQ^T$ where Q is an orthogonal matrix and D is a diagonal matrix whose entries are the eigenvalues of A.

Let $A \in \mathcal{S}^n$ be a matrix with strictly positive entries. The Perron-Frobenius theorem says that the largest eigenvalue in absolute value of A has multiplicity one and the corresponding eigenvector can be chosen to have strictly positive entries.

The vector space of the complex matrices $\mathbb{C}^{n\times n}$ is equipped with the trace inner product defined as $\langle A, B\rangle = \mathsf{Tr}(A^*B) = \sum_{ij} \overline{A_{ij}} B_{ij}$, where $\overline{A_{ij}}$ is the complex conjugate of A_{ij}. A complex matrix A is called *Hermitian* if $A^* = A$; i.e., if A is equal to its conjugate transpose. All the eigenvalues of a Hermitian matrix are real. A complex matrix $U \in \mathbb{C}^{n\times n}$ is called *unitary* if $U^*U = I$. Unitary matrices preserve inner products between vectors; i.e., for any pair $x, y \in \mathbb{C}^n$ we have $\langle Ux, Uy\rangle = \langle x, y\rangle$.

A *permutation matrix* is a square matrix where each row and each column has exactly one entry equal to 1 and all others are equal to 0.

For a pair of matrices A, B we let $A \oplus B = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ denote their direct sum and $A \circ B$ denote the entrywise product, where the ij-th entry of $A \circ B$ is equal to $A_{ij}B_{ij}$. The tensor product (also known as Kronecker product) of A and B is denoted by $A \otimes B$. If A is an $m \times n$ matrix, it is defined as the block matrix $A \otimes B = \begin{pmatrix} A_{11}B & \dots & A_{1n}B \\ \vdots & \ddots & \vdots \\ A_{m1}B & \dots & A_{mn}B \end{pmatrix}$. The tensor product is both associative (i.e., $(A \otimes B) \otimes C = A \otimes (B \otimes C)$) and distributive (i.e., for matrices B, C of the same size we have $A \otimes (B + C) = A \otimes B + A \otimes C$).

2.1.1 Positive semidefinite matrices

A matrix $A \in \mathcal{S}^n$ is called *positive semidefinite* if the associated quadratic form x^TAx is nonnegative; i.e., $x^TAx \geq 0$ for any vector $x \in \mathbb{R}^n$. We write $A \succeq 0$ to denote the fact that A is positive semidefinite and we let $\mathcal{S}^n_+$ denote the set of $n \times n$ positive semidefinite matrices. The *Gram matrix* of a set of vectors $x_1, \dots, x_n \in \mathbb{R}^d$ is the $n \times n$ matrix A where $A_{ij} = \langle x_i, x_j\rangle$ for all $i, j \in [n]$ and we say that the vectors $x_1, \dots, x_n$ form a *Gram representation* of A. There are several equivalent characterizations of a positive semidefinite matrix. In the theorem below we summarize the ones that are more relevant for this thesis.

2.1.1. THEOREM. *Consider a matrix $A \in \mathcal{S}^n$. The following statements are equivalent:*

- *A is positive semidefinite; i.e., $x^TAx \geq 0$ for any vector $x \in \mathbb{R}^n$.*
- *All the eigenvalues of A are nonnegative.*
- *A is the Gram matrix of a family of real vectors.*

- $A = X^T X$ *for some real matrix* X.

A complex Hermitian matrix $A \in \mathbb{C}^{n\times n}$ is *positive semidefinite* if $x^* Ax \geq 0$ for all $x \in \mathbb{C}^n$. This is identical to requiring that A has only real nonnegative eigenvalues, or that A is the Gram matrix of a family of complex vectors, or that $A = X^* X$ for some complex matrix X.

A matrix A is *positive definite* if the associated quadratic form $x^T Ax$ is positive (i.e., $x^T Ax > 0$ for any nonzero vector $x \in \mathbb{R}^n$) and if this is the case we write $A \succ 0$. In analogy to Theorem 2.1.1, a matrix is positive definite if and only if all its eigenvalues are strictly positive.

A symmetric matrix A that satisfies $A^2 = A$ is called a *projector*.

We now collect some useful, basic properties of positive semidefinite matrices.

2.1.2. LEMMA. *A symmetric matrix A is positive semidefinite if and only if $A = X^2$ for some $X \in \mathcal{S}^n$.*

PROOF: Let $A \in \mathcal{S}^n$ be a positive semidefinite matrix. By the spectral decomposition theorem $A = QDQ^T$ for some orthogonal matrix Q and nonnegative diagonal matrix D. Let $\sqrt{D}$ be the square root of the matrix D (i.e., $D = \sqrt{D}\sqrt{D}$), then the matrix $X = Q\sqrt{D}Q^T$ is symmetric and $X^2 = A$.

Conversely, let X be a symmetric matrix. Then the matrix $X^2 = X^T X$ must be positive semidefinite by Theorem 2.1.1. □

Let $A \in \mathcal{S}^n$ be of the form $A = \begin{pmatrix} \alpha & b^T \\ b & M \end{pmatrix}$, where $\alpha > 0$, $b \in \mathbb{R}^{n-1}$ and $M \in \mathcal{S}^{n-1}$. Then,

$$A \succeq 0 \iff M - bb^T/\alpha \succeq 0. \tag{2.1}$$

The matrix $M - bb^T/\alpha$ is called the *Schur complement* of M in A with respect to the entry α.

2.1.3. PROPOSITION. *Let A and B be $n \times n$ positive semidefinite matrices. Then the following holds:*

(i) $\langle A, B\rangle = 0$ *if and only if* $AB = 0$.

(ii) The matrices $A \oplus B$, $A \circ B$ and $A \otimes B$ are all positive semidefinite.

2.1.4. LEMMA. *Consider the matrices $A, B \in \mathcal{S}^n_+$. If $\mathsf{Tr}(AB) = \mathsf{Tr}(A^2) = \mathsf{Tr}(B^2)$, then we have that $A = B$.*

PROOF: The matrix $A - B$ is symmetric and therefore, by Lemma 2.1.2, we have $(A - B)^2 \in \mathcal{S}^n_+$. Hence, $0 \leq \mathsf{Tr}((A - B)^2) = \mathsf{Tr}(A^2) + \mathsf{Tr}(B^2) - 2\,\mathsf{Tr}(AB)$ which, by the assumptions, is equal to zero. Since the trace of a positive semidefinite matrix X is zero if and only if X is the zero matrix, we conclude that $A = B$. □

Consider the matrices $I, J \in \mathcal{S}^n$, which are, respectively, the identity and the all-one matrix. Then the matrix $nI - J$ is positive semidefinite. Indeed, one can easily check that the only eigenvalues of $nI - J$ are 0 and n.

2.2 Cones of matrices

A *convex cone* $\mathcal{K}$ is a set satisfying the two following properties: (*i*) for all $x \in \mathcal{K}$ and $\lambda > 0$ we have $\lambda x \in \mathcal{K}$; (*ii*) for all $x, x' \in \mathcal{K}$ we have $x + x' \in \mathcal{K}$. Given a cone $\mathcal{K} \subseteq \mathcal{S}^n$, its *dual cone* is $\mathcal{K}^* = \{M \in \mathcal{S}^n : \langle A, M\rangle \geq 0 \quad \forall\, A \in \mathcal{K}\}$. The set $\mathcal{K}^*$ is a closed convex cone and $\mathcal{K}^{**} = \mathcal{K}$ if and only if $\mathcal{K}$ is a closed convex cone. A cone $\mathcal{K} \subseteq \mathcal{S}^n$ is *full-dimensional* if it contains a basis of $\mathcal{S}^n$ and *pointed* if the only linear subspace contained in it is the trivial subspace consisting only of the zero matrix. Consider a full-dimensional cone $\mathcal{K}$. A matrix A lies in the *interior* of $\mathcal{K}$, denoted by $\text{int}(\mathcal{K})$, if and only if $\langle A, M\rangle > 0$ for all nonzero matrices $M \in \mathcal{K}^*$. A cone is called *proper* if it is convex, closed, full-dimensional, pointed, and with non-empty interior. The dual set of a proper cone is also proper.

2.2.1 The positive semidefinite cone

The cone of positive semidefinite matrices has been widely studied. We summarize some of its useful properties.

2.2.1. THEOREM. *The cone of positive semidefinite matrices $\mathcal{S}^n_+$ has the following properties:*

(i) *$\mathcal{S}^n_+$ is a proper cone (i.e., it is convex, closed, full-dimensional, pointed, and with non-empty interior).*

(ii) *The interior of the positive semidefinite cone is the set of positive definite matrices.*

(iii) *$\mathcal{S}^n_+$ is a self-dual cone; i.e., $M \in \mathcal{S}^n_+$ if and only if $\langle M, A\rangle \geq 0$ for all $A \in \mathcal{S}^n_+$.*

(iv) *The extreme rays of $\mathcal{S}^n_+$ are the rank 1 matrices yy^T where $y \in \mathbb{R}^n$.*

The *doubly nonnegative cone*, denoted by $\mathcal{DNN}^n$, is the set of positive semidefinite matrices in $\mathcal{S}^n$ with nonnegative entries. The cone $\mathcal{DNN}^n$ is proper.

2.2.2 The completely positive and copositive cone

A matrix $A \in \mathcal{S}^n$ is called *completely positive* if A is the Gram matrix of a set of *nonnegative* vectors $x_1, \ldots, x_n \in \mathbb{R}^d_+$ for some $d \geq 1$. We let $\mathcal{CP}^n$ denote the

set of completely positive matrices. Clearly, any completely positive matrix is positive semidefinite and its entries are nonnegative. Thus, $\mathcal{CP}^n \subseteq \mathcal{DNN}^n$. Moreover, we observe that any rank 1 doubly nonnegative matrix is completely positive. Indeed, if $A = yy^T$ with $y \in \mathbb{R}^n$ and A is entrywise nonnegative then, without loss of generality, $y \in \mathbb{R}^n_+$ and the scalars $y_1, \ldots, y_n$ form a Gram representation of A. It is well-known that the set $\mathcal{CP}^n$ is closed. This can be proven using the fact that its extreme rays are the rank 1 matrices yy^T where $y \in \mathbb{R}^n_+$. Therefore, any matrix in $\mathcal{CP}^n$ can be written as $\sum_{i=1}^N y_i y_i^T$, where $y_1, \ldots, y_N \in \mathbb{R}^n_+$ and $N \leq \binom{n+1}{2}$ (using Carathéodory's theorem), and thus closedness follows using a compactness argument (see e.g. [BSM03, Theorem 2.2] for the full proof). Having an explicit description of the extreme rays of the $\mathcal{CP}$ cone is a key ingredient in many proofs concerning $\mathcal{CP}$. We refer the reader to the book of Berman and Shaked-Monderer [BSM03] for a detailed account on the properties of the $\mathcal{CP}$ cone.

The dual of the completely positive cone $\mathcal{CP}^n$ is the copositive cone $\mathcal{COP}^n$, which consists of the matrices $M \in \mathcal{S}^n$ for which the n-variate polynomial $p_M = \sum_{i,j=1}^n M_{ij} x_i^2 x_j^2$ is nonnegative over $\mathbb{R}^n$; i.e., $\sum_{i,j=1}^n M_{ij} x_i^2 x_j^2 \geq 0$ for all $x_1, \ldots, x_n \in \mathbb{R}$. The following simple lemma shows that it is only necessary to check that the polynomial p_M is nonnegative over the ball or, equivalently, over the sphere.

2.2.2. LEMMA. *A matrix $M \in \mathcal{COP}^n$ if and only if $p_M = \sum_{i,j=1}^n M_{ij} x_i^2 x_j^2$ is nonnegative over the ball (i.e., $x \in \mathbb{R}^n : \sum_{i=1}^n x_i^2 \leq 1$) or, equivalently, p_M is nonnegative over the sphere (i.e., $x \in \mathbb{R}^n : \sum_{i=1}^n x_i^2 = 1$).*

PROOF: Observe that any nonzero vector $x \in \mathbb{R}^n$ can be rescaled such that $\sum_{i=1}^n x_i^2 = 1$. Therefore, as p_M is a homogeneous polynomial, p_M is nonnegative over $\mathbb{R}^n$ if and only if it is nonnegative over the ball or, equivalently, over the sphere. □

2.3 Conic optimization

A *conic optimization* problem consists in finding the supremum (or infimum) of a convex function over an affine slice of a convex cone. Linear and semidefine programming are two of the most well-known classes of conic optimization programs. We now briefly introduce the needed concepts and we refer the reader to the book of Boyd and Vandenberghe [BV04] for further details.

Let $\mathcal{K}$ be a proper cone. Given $C, A_j \in \mathcal{S}^n$ and $b_j \in \mathbb{R}$ for $j \in [m]$, a standard conic program over a convex cone $\mathcal{K} \subseteq \mathcal{S}^n$ has the form:

$$p^* = \sup \langle C, X \rangle \text{ s.t. } \langle A_j, X \rangle = b_j \quad \forall j \in [m],\ X \in \mathcal{K}. \tag{2.2}$$

The matrix X is the *variable* of the program. The conditions $X \in \mathcal{K}$, $\langle A_j, X \rangle = b_j$ for all $j \in [m]$ are the *constraints* and the quantity $\langle C, X \rangle$ is the *objective value* of the program. A matrix that satisfies all the constraints is called *feasible* and a feasible matrix which lies in the interior of $\mathcal{K}$ is called *strictly feasible*. A feasible matrix that maximizes the objective value is an *optimal solution* for the program. The corresponding dual program has the form:

$$d^* = \inf \sum_{j=1}^{m} b_j y_j \text{ s.t. } Z = \sum_{j=1}^{m} y_j A_j - C \in \mathcal{K}^*. \tag{2.3}$$

By weak duality we have that $p^* \leq d^*$. Moreover, assume that $d^* > -\infty$ and (2.3) is strictly feasible, then strong duality holds: $p^* = d^*$ and (2.2) attains its supremum.

If $\mathcal{K}$ is the set of nonnegative diagonal matrices, (2.2) and (2.3) are linear programs and they can be solved in polynomial time. If $\mathcal{K} = \mathcal{S}_+^n$ then (2.2) and (2.3) are called positive semidefinite programs. The optimal value of such programs can be approximated to within fixed arbitrary precision in polynomial time (see e.g. [GLS88, BTN01]).

2.4 Quantum information theory

We now give some basic mathematical background information on quantum information theory. For more on quantum information theory we refer to the book of Nielsen and Chuang [NC00] and the lecture notes of Watrous [Wat11].

2.4.1 Quantum states and operations

Quantum states. A *quantum register* is an idealized physical system with which experimenters (commonly called Alice and Bob) may interact and it is represented by a finite-dimensional complex vector space. The set of possible *states* of a d-dimensional quantum register is formed by the $d \times d$ complex positive semidefinite matrices whose trace equals 1. When such a state is ρ, the quantum register $\mathcal{A}$ is said to be *in* state ρ. A state with rank 1 is called a *pure state*; i.e., $\rho = vv^*$ for some unit vector $v \in \mathbb{C}^d$. By the complex spectral decomposition theorem, we have that any state $\rho \in \mathbb{C}^{d \times d}$ is the convex combination of pure states; i.e., $\rho = \sum_{i \in [d]} \lambda_i v_i v_i *$ where $\lambda_1, \ldots, \lambda_d \geq 0$ are the eigenvalues and $v_1, \ldots, v_d \in \mathbb{C}^d$ are the corresponding eigenvectors. A state which is not pure is called a *mixed state*. At times, a complex unit vector $v \in \mathbb{C}^d$ is also referred to as a state. In that case we are identifying the vector v to the pure state $\rho = vv^*$.

Quantum operations. An experimenter can alter a state $\rho \in \mathbb{C}^{d \times d}$ in two possible ways: applying a unitary transformation or performing a measurement.

A *unitary transformation* is simply a mapping $\rho \mapsto U\rho U^*$, where $U \in \mathbb{C}^{d\times d}$ satisfies $UU^* = I$. A *t-outcome measurement* is a collection $\{F_i \in \mathbb{C}^{d\times d} : i \in [t]\}$ of positive semidefinite matrices F_i that satisfy $\sum_{i=1}^{t} F_i = I$, where I is the identity matrix. Such a collection of matrices is also called *positive operator valued measure*, or *POVM* for short. If Alice performs a t-outcome measurement $\{F_i\}_{i\in[t]}$, where $F_i = M_i^* M_i$, on a register $\mathcal{A}$ which is in a state ρ, then she will observe a random variable λ over the set $[t]$ whose probability distribution is given by $\Pr[\lambda = i] = \mathsf{Tr}(F_i\rho)$. In the event that $\lambda = i$, we say that Alice gets measurement outcome i and that the state collapses to $M_i\rho M_i^* / \mathsf{Tr}(F_i\rho)$.

Dirac notation. In quantum information theory it is common to write vectors using the Dirac notation. The canonical unit vectors in $\mathbb{C}^d$ are denoted by $|1\rangle, \ldots, |d\rangle$; that is, $|i\rangle$ is the vector with a 1 in position i and 0's elsewhere. Greek letters are used to denote unit vectors; e.g., $|\phi\rangle = \sum_{i\in[d]} \alpha_i|i\rangle \in \mathbb{C}^d$ where $\sum_{i\in[d]} |\alpha_i|^2 = 1$. The conjugate transpose of a vector $|\phi\rangle \in \mathbb{C}^d$ is denoted by $\langle\phi|$. When we take the tensor product between two vectors, we often omit the tensor product symbol: we abbreviate $|\phi\rangle \otimes |\psi\rangle$ with $|\phi\rangle|\psi\rangle$.

Superposition. A fundamental feature of quantum mechanics is the fact that quantum states can be in superposition. The pure state $|\phi\rangle = \sum_{i\in[d]} \alpha_i|i\rangle$ is said to be in *superposition* of the states $|1\rangle, \ldots, |d\rangle$ and the complex number α_i is called the *amplitude* of $|i\rangle$ in $|\phi\rangle$. The rough idea is that a quantum state can be in various states at the same time, but this phenomenon cannot be directly observed. Indeed, if the state $|\phi\rangle$ is measured in the computational basis (i.e., the d-outcome measurement $\{F_i\}_{i\in[d]}$ where $F_i = |i\rangle\langle i|$), then the experimenter will observe the state $|j\rangle$ with probability $\mathsf{Tr}(F_j|\phi\rangle\langle\phi|) = |\alpha_j|^2$. In other words, once the state $|\phi\rangle$ is measured it collapses to one of the states forming its superposition and the probability of this happening depends on the corresponding amplitude.

Qubits. The basic unit of classical computing is the *bit*, which is either 0 or 1. Its quantum counterpart is the *qubit* (quantum bit), which is a superposition of the basis states $|0\rangle = \begin{pmatrix}1\\0\end{pmatrix}$ and $|1\rangle = \begin{pmatrix}0\\1\end{pmatrix}$ in $\mathbb{C}^2$. Indeed, a qubit is a state $\alpha_0|0\rangle + \alpha_1|1\rangle \in \mathbb{C}^2$, where $|\alpha_0|^2 + |\alpha_1|^2 = 1$. More generally, an n-qubit is a superposition of 2^n basis states, each of the form $|b_1\rangle|b_2\rangle \ldots |b_n\rangle$ where $b_i \in \{0, 1\}$. As n-bit strings can be viewed as numbers between 0 and $2^n - 1$, the basis states can also written as $|0\rangle, \ldots, |2^n - 1\rangle$. An n-qubit is then any state of the form $\sum_{i=0}^{2^n-1} \alpha_i|i\rangle$ with $\sum_{i=0}^{2^n-1} |\alpha_i|^2 = 1$.

Entangled states. The possible states of a *pair* of quantum registers $(\mathcal{A}, \mathcal{B})$ are the trace-1 positive semidefinite matrices in $\mathbb{C}^{d_A \times d_A} \otimes \mathbb{C}^{d_B \times d_B}$. Here, d_A and d_B are the dimensions of $\mathcal{A}$ and $\mathcal{B}$, respectively. The pair of systems $(\mathcal{A}, \mathcal{B})$ is said to be *entangled* if it is in a state ρ which is not a convex combination of states of the form $\rho_A \otimes \rho_B$. In that case, we also say that the state ρ is entangled. If a state is not entangled, it is called *separable*. The most famous entangled state is the so-called *EPR pair*: $(|0\rangle|0\rangle + |1\rangle|1\rangle)/\sqrt{2} \in \mathbb{C}^2 \otimes \mathbb{C}^2$ named after Einstein, Podolsky, and Rosen [EPR35], who first observed that quantum mechanics predicts the existence of entangled states. In matrix notation, this is the state $\rho = vv^*$ where $v = (e_1 \otimes e_1 + e_2 \otimes e_2)/\sqrt{2}$. More generally, the d-dimensional *maximally entangled state* $\sigma = vv^*$ is defined by the vector $v = (\sum_{\ell \in [d]} e_\ell \otimes e_\ell)/\sqrt{d}$, where e_ℓ denotes the ℓ-th canonical basis vector.

The *partial trace* is a linear operator defined as follows: for $A \in \mathbb{C}^{d_A \times d_A}$ and $B \in \mathbb{C}^{d_B \times d_B}$ define $\mathsf{Tr}_{\mathcal{A}}(A \otimes B) = \mathsf{Tr}(A)B$ and $\mathsf{Tr}_{\mathcal{B}}(A \otimes B) = A\,\mathsf{Tr}(B)$, and extend these definitions in a linear fashion to all matrices of $\mathbb{C}^{d_A \times d_A} \otimes \mathbb{C}^{d_B \times d_B}$. The partial trace of a state $\rho \in \mathbb{C}^{d_A \times d_A} \otimes \mathbb{C}^{d_B \times d_B}$ with respect to the system $\mathcal{A}$ is called the *reduced state* of ρ on system $\mathcal{B}$. The concept of partial trace can be generalized to the case where there are more than two quantum registers. Suppose that the ℓ quantum registers $(\mathcal{B}_1, \ldots, \mathcal{B}_\ell)$ are in state ρ. We denote with $\mathsf{Tr}_{\mathcal{B}_{-k}}(\rho)$ the partial trace of ρ over all the subspaces but the k-th one; i.e., $\mathsf{Tr}_{\mathcal{B}_{-k}}(\rho) = \mathsf{Tr}_{\mathcal{B}_1,\ldots,\mathcal{B}_{k-1},\mathcal{B}_{k+1},\ldots,\mathcal{B}_\ell}(\rho)$.

Suppose that the pair of quantum registers $(\mathcal{A}, \mathcal{B})$ is in the state ρ. If Alice performs a unitary $U \in \mathbb{C}^{d_A \times d_A}$ on her register then the state ρ of the system $(\mathcal{A}, \mathcal{B})$ is mapped to $(U \otimes I)\rho(U \otimes I)^*$, and similarly if Bob performs a unitary on his register. Moreover, if Alice performs a t-outcome measurement $\{F_i\}_{i \in [t]}$ on $\mathcal{A}$ then the probability that Alice gets measurement outcome i equals $p_i = \mathsf{Tr}\left((F_i \otimes I)\rho\right)$ and if this happens Bob's register $\mathcal{B}$ is left in the state $\rho^i = \mathsf{Tr}_{\mathcal{A}}\left((F_i \otimes I)\rho\right)/p_i$. If Bob now performs an r-outcome measurement $\{F'_j\}_{j \in [r]}$ on $\mathcal{B}$, then the probability that he gets outcome $j \in [r]$ equals $\mathsf{Tr}_{\mathcal{B}}(F'_j \rho^i)$.

The following lemma says that there exists a measurement that allows to perfectly distinguish among a collection of states if and only if these are pairwise orthogonal.

2.4.1. LEMMA (ORTHOGONALITY LEMMA). *Let $\rho_1, \ldots, \rho_\ell \in \mathbb{C}^{d \times d}$ be a collection of Hermitian positive semidefinite matrices. Then the following are equivalent:*

(1) We have $\rho_i \rho_j = 0$ for every $i \neq j \in [\ell]$.

(2) There exists a set of projectors $P^1, \ldots, P^\ell, P^\perp \in \mathbb{C}^{d \times d}$ forming an $(\ell + 1)$-outcome measurement (i.e., $\sum_{i \in [\ell]} P^i + P^\perp = I$) and such that $\mathsf{Tr}(P^i \rho_j) = \delta_{ij}\,\mathsf{Tr}(\rho_j)$ and $\mathsf{Tr}(P^\perp \rho_j) = 0$ for every $i, j \in [\ell]$.

In particular, a collection of pure states $|\phi_1\rangle, \ldots, |\phi_\ell\rangle \in \mathbb{C}^d$ can be perfectly distinguished with a measurement if and only if they are pairwise orthogonal.

PROOF: (1) $\Rightarrow$ (2): Let $V_i \subseteq \mathbb{C}^d$ be the image of the matrix ρ_i. Condition (1) implies that the spaces $V_1, \ldots, V_\ell$ are pairwise orthogonal. To see this, observe that for any vectors $u \in V_i$ and $v \in V_j$ there exist $x, y \in \mathbb{C}^d$ such that $u = \rho_i x$ and $v = \rho_j y$. By Hermiticity, we have $u^* v = x^* \rho_i \rho_j y = 0$. Let P^i be the orthogonal projection onto V_i and let $P^\perp = I - \sum_{i=1}^{\ell} P^i$. It is now trivial to verify that these projectors satisfy the desired properties.

(2) $\Rightarrow$ (1): Let V_i be the image of ρ_i and let W_i be the image of P^i. We start by proving that $V_i \subseteq W_i$. To this end, we expand ρ_i in its spectral decomposition: $\rho_i = \sum_{\ell \in [d]} \lambda_\ell v_\ell v_\ell^*$ where $\lambda_1, \ldots, \lambda_d \in \mathbb{R}_+$ are the eigenvalues and $v_1, \ldots, v_d \in \mathbb{C}^d$ are the corresponding eigenvectors. Then we have that

$$\sum_{\ell \in [d]} \lambda_\ell = \mathsf{Tr}(\rho_i) = \mathsf{Tr}(P^i \rho_i) = \sum_{\ell \in [d]} \lambda_\ell \, \mathsf{Tr}\left(P^i v_\ell v_\ell^*\right) = \sum_{\ell \in [d]} \lambda_\ell v_\ell^* P^i v_\ell. \tag{2.4}$$

As P^i is a projector, we have that (2.4) holds if and only if $v_\ell^* P^i v_\ell = 1$ for each $\ell \in [d]$ such that $\lambda_\ell \neq 0$, which in turns implies that for such $\ell \in [d]$ each v_ℓ is an eigenvector of P^i with eigenvalue 1. Therefore, $V_i \subseteq W_i$ holds. Similarly, the condition $\mathsf{Tr}(P^i \rho_j) = 0$ if $i \neq j$ implies that W_i is orthogonal to V_j and thus $V_i \subseteq W_i \subseteq V_j^\perp$ if $i \neq j$. Considering again the spectral decomposition of ρ_i, for $i \neq j$ we have: $\rho_j \rho_i = \sum_{\ell \in [d]} \lambda_\ell \rho_j v_\ell v_\ell^* = 0$, since for every $\lambda_\ell \neq 0$ the vector v_ℓ lies in $V_i \subseteq V_j^\perp$. □

Hilbert spaces. A *Hilbert space* $\mathcal{H}$ is a vector space endowed with an inner product $\langle \cdot, \cdot \rangle$ such that the induced norm $\|x\| = \sqrt{\langle x, x \rangle}$ turns $\mathcal{H}$ into a complete metric space; i.e., a metric space in which every Cauchy sequence converges. (A sequence $(x_i)_{i \in \mathbb{N}} \subseteq \mathcal{H}$ is a Cauchy sequence if for every $\varepsilon > 0$ there exists a $N \in \mathbb{N}$ such that $d(x_i, x_j) < \varepsilon$ for all $i, j > N$, where $d(\cdot, \cdot)$ is the distance of the metric space.) A bounded operator $T : \mathcal{H} \to \mathcal{H}$ is said to be *positive* if $\langle Tx, x \rangle \geq 0$ for all $x \in \mathcal{H}$.

The Euclidean spaces $\mathbb{R}^n$ and $\mathbb{C}^n$ endowed with the standard (Euclidean) inner product are Hilbert spaces.

2.4.2 Quantum correlations

The most interesting difference between separable and entangled states is that the latter type can lead to measurement outcomes which are correlated in a non-classical fashion. We only consider the case of two parties (aka the bipartite setting). The sets X, Y (respectively, A, B) model the possible inputs

(respectively, outputs) of the two parties. We assume throughout that these sets are finite.

A classical correlation is a bipartite probability distribution that can be obtained using local and shared randomness. More formally, we have the following mathematical definition.

2.4.2. DEFINITION. [Classical correlations] A bipartite probability distribution $P = (P(a,b|x,y))_{a\in A,b\in B,x\in X,y\in Y}$ is called *classical* if $P(a,b|x,y)$ admits a local hidden variable model. Formally, there exists a distribution $Q(\lambda)$ over the hidden variable λ as well as probabilities $P(a|x,\lambda)$ and $P(b|y,\lambda)$ such that $P(a,b|x,y) = \sum_\lambda Q(\lambda)P(a|x,\lambda)P(b|y,\lambda)$.

We denote by $\mathcal{L}$ the set of bipartite classical correlations.

Any probability distribution that is not in $\mathcal{L}$ is called *nonlocal*.

We now define the set of bipartite quantum correlations $\mathcal{Q}$, consisting of the conditional probabilities that two physically separated parties can generate by performing measurements on a shared quantum state.

2.4.3. DEFINITION. [Quantum correlations] A bipartite probability distribution $P = (P(a,b|x,y))_{a,b,x,y}$ is called *quantum* if

$$P(a,b|x,y) = \langle \psi, (E_x^a \otimes F_y^b)\psi\rangle,$$

where $\psi \in \mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$ is a unit vector (for some $d_A, d_B \in \mathbb{N}$) and for some sets of positive semidefinite matrices (aka POVM) $\{E_x^a\}_{a\in A}$ and $\{F_y^b\}_{b\in B}$ satisfying $\sum_{a\in A} E_x^a = I$ and $\sum_{b\in B} F_y^b = I$ for all $x \in X, y \in Y$.

The set of bipartite quantum correlations $\mathcal{Q}$ consists of all bipartite quantum probabilities.

In the above definition we have only considered pure states because any mixed state can be viewed as the reduced state of a pure state. This property is known as the purification of a quantum state.

Note that we can without loss of genality assume that the unit vector ψ is real valued and that E_x^a, F_y^b are real valued positive symmetric matrices. This is due to the fact that the map that sends a Hermitian matrix $A \in \mathbb{C}^{d\times d}$ to the symmetric matrix $\frac{1}{\sqrt{2}}\begin{pmatrix} \mathrm{Re}(A) & \mathrm{Im}(A) \\ -\mathrm{Im}(A) & \mathrm{Re}(A)\end{pmatrix} \in \mathcal{S}^{2d}$ is an isometry that preserves positive semidefiniteness.

While the set of classical correlations (those obtained using only local and shared randomness) forms a polytope so that membership can be decided using linear programming, the set $\mathcal{Q}$ of quantum correlations is convex but with

infinitely many extreme points and its structure is much harder to characterize. In particular, it is not known whether the set of quantum correlations $\mathcal{Q}$ is closed.

In the above definition we required the composite Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ to be finite dimensional. One can generalize such definition by allowing an infinite amount of entanglement (the space $\mathcal{H}_A \otimes \mathcal{H}_B$ is now infinite dimensional). We let $\mathcal{Q}_\infty$ define the set of correlations arising this way.

One can also consider probability distributions arising from the relativistic point of view. Roughly, instead of assuming that the measurement operators act on different Hilbert spaces so that joint measurements have a tensor product structure, in the relativistic model the measurement operators act on a common (possibly infinite dimensional) Hilbert space and the operators of the two parties mutually commute. In this case, joint measurement operators have a product structure.

2.4.4. DEFINITION. [Relativistic quantum field theory correlations] A correlation $P = (P(a,b|x,y))_{a,b,x,y}$ is obtained from relativistic quantum field theory if $P(a,b|x,y) = \langle \psi, E_x^a F_y^b \psi \rangle$, where ψ is a unit vector in a (possibly infinite dimensional) Hilbert space $\mathcal{H}$; E_x^a and F_y^b are positive operators on $\mathcal{H}$ satisfying $\sum_{a\in A} E_x^a = I$ and $\sum_{b\in B} F_y^b = I$ for all $x \in X, y \in Y$; and $E_x^a F_y^b = F_y^b E_x^a$ for all $a \in A, b \in B, x \in X, y \in Y$.

We denote by $\mathcal{Q}_c$ the set of bipartite quantum correlations arising from the relativistic point of view.

The set $\mathcal{Q}_c$ is closed (see e.g. [Fri12, Proposition 3.4]) and the following inclusions hold:

$$\mathcal{Q} \subseteq \mathcal{Q}_\infty \subseteq \mathsf{cl}(\mathcal{Q}) \subseteq \mathcal{Q}_c. \tag{2.5}$$

In a very recent breakthrough Slofstra [Slo16] showed that $\mathcal{Q}_\infty \subsetneq \mathcal{Q}_c$, while it is still an open problem to determine whether $\mathcal{Q} = \mathcal{Q}_\infty$ holds [WCD08]. Moreover, deciding whether the identity $\mathsf{cl}(\mathcal{Q}) = \mathcal{Q}_c$ holds is known to be equivalent to *Connes' embedding conjecture* (see [Oza13, Fri12, JNP$^+$11]).

Any quantum correlation respects the no-signaling principle, which says that information cannot propagate faster than the speed of light. The non-signaling correlations are the set of correlations that obey the no-signaling principle.

2.4.5. DEFINITION. [Non-signaling correlations] A bipartite probability distribution $P = (P(a,b|x,y))_{a,b,x,y}$ is called *non-signaling* if the marginal distribution of each party only depends on its corresponding input, which means that $\sum_{a\in A} P(a,b|x,y) = \sum_{a\in A} P(a,b|x',y)$ and $\sum_{b\in B} P(a,b|x,y) = \sum_{b\in B} P(a,b|x,y')$ for all a, b, x, x', y, y'.

We denote by $\mathcal{NS}$ the set of bipartite non-signaling correlations.

It is well-known that the following relationships hold among the various bipartite correlations:

$$\mathcal{L} \subset \mathcal{Q} \subseteq \mathcal{Q}_c \subset \mathcal{NS}.$$

We briefly mention that recent works from Mančinska and Roberson [MR14], and independently Sikora and Varvitsiotis [SV15], showed that the set $\mathcal{Q}$ can be described using the matrix cone which we introduce in this thesis, the completely positive semidefinite cone. (We will give the details in Section 5.5.) In particular, they showed that

$$\mathcal{Q} = \pi(\mathcal{CS}_+^N \cap \mathcal{B}^N), \tag{2.6}$$

where $\mathcal{B}^N$ is an affine space and π is the projection onto a subspace (see Theorem 5.5.3 for the specifics).

2.4.3 Nonlocal games

In a nonlocal game, two (or more) cooperating players determine a common strategy to answer questions posed by a referee. A question pair (x, y) is drawn from a finite set $X \times Y$ and the referee sends a question to each of the players. Without communicating, the players must each respond to their question and send the answer to the referee. Upon collecting the answer pair $(a, b) \in A \times B$, using the rules of the game the referee determines whether the players have won or lost. A strategy is said to be successful with probability p if it wins any instance of the game with probability at least p. A perfect strategy is one that always succeeds with probability 1.

A *deterministic classical strategy* is determined by the maps $f_A : X \to A$, $f_B : Y \to B$ that each player respectively uses to determine the answer given the question. Classical strategies might involve shared and private randomness where the players also use coin flips to determine their answers and any probabilistic strategy can be seen as a probability distribution over deterministic classical strategies. In other words, the probability distribution $P = (P(a, b|x, y))_{a,b,x,y}$ that arises from a probabilistic strategy must lie in the set $\mathcal{L}$.

In a *quantum strategy*, the players share a quantum state on which they perform local measurements to obtain their answers (see also Figure 2.1). Suppose that the two players, Alice and Bob, have quantum registers $\mathcal{A}$ and $\mathcal{B}$, respectively, that are initialized to be in some entangled state ρ. Upon receiving question $x \in X$, Alice performs a measurement $\{E_x^a\}_{a \in A}$; that is, the set $\{E_x^a\}_{a \in A}$ is a collection of positive semidefinite matrices that sums up to the identity. Simultaneously, upon receiving question $y \in Y$, Bob performs a measurement $\{F_y^b\}_{b \in B}$. The measurement outcomes determine the answers. Therefore, if (x, y) is the question pair then the probability of answering (a, b) is equal to

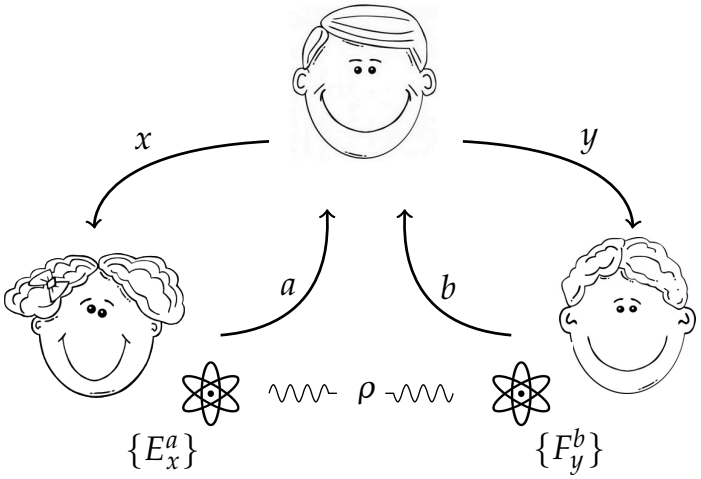

Figure 2.1: A quantum strategy for a nonlocal game.

$\mathrm{Tr}((E_s^a \otimes F_t^b)\rho)$. Equivalently, the corresponding probability $P = (P(a,b|x,y))$ lies in $\mathcal{Q}$.

We will study, in particular, the problem of whether for a fixed game a perfect classical (or quantum) strategy exists. Note that since any perfect classical strategy is given by the convex combination of perfect deterministic classical strategies, we can without loss of generality restrict our attention to deterministic classical strategies.

2.4.4 Remote state preparation

We end this chapter with the description of a quantum protocol that will be useful in Chapters 6-9.

Suppose that Alice has in mind a pure state $\rho = uu^*$ where $u \in \mathbb{C}^d$ is some unit vector that is unknown to Bob. *Remote state preparation* [BDVS$^+$01] is a protocol that enables the parties to prepare a quantum register belonging to Bob in the state ρ using only local measurements on a pair of entangled quantum registers and classical communication from Alice to Bob. This task can be achieved using the teleportation scheme of Bennett et al. [BBC$^+$93], which allows to remotely prepare a d-dimensional state ρ with the communication of one among d^2 distinct messages. (We refer interested readers to [BBC$^+$93] and [NC00, pp. 26–28] for the details of the teleportation scheme.) However, for certain states, remote state preparation can be performed with less communication. In particular, for the case where the vector u has only entries of absolute value $d^{-1/2}$ there exists a remote state preparation protocol that requires only sending one among d distinct messages, or equivalently $\lceil \log d \rceil$

bits. The protocol is due to Zeng and Zhang [ZZ02] and we will describe it here for completeness.

To remotely prepare the state $\rho = uu^*$, where $u \in \mathbb{C}^d$ is a vector with entries of modulus $d^{-1/2}$ that is unknown to Bob, in a d-dimensional register $\mathcal{B}$ that he possesses the two parties can use the following protocol: Assume Alice has a d-dimensional quantum register $\mathcal{A}$, such that $(\mathcal{A}, \mathcal{B})$ is in the maximally entangled state $\sigma = vv^*$ where $v = (\sum_{\ell \in [d]} e_\ell \otimes e_\ell)/\sqrt{d}$.

1. Alice performs on her register $\mathcal{A}$ the unitary transformation given by $U = \sqrt{d}\,\mathrm{Diag}(u)$, where $\mathrm{Diag}(u)$ is the diagonal matrix whose main diagonal is vector u.

2. Next, Alice performs on her register the d-dimensional discrete Fourier transform, given by the unitary $F \in \mathbb{C}^{d \times d}$ whose (ℓ, m)-th entry is equal to $F_{\ell,m} = e^{2\pi i(\ell-1)(m-1)/d}/\sqrt{d}$.

3. She then measures in the canonical basis (i.e., she performs the measurement $\{e_\ell e_\ell^* : \ell \in [d]\}$ on her register) and gets an outcome $\ell \in [d]$ which she communicates to Bob.

4. Last, Bob performs on his register $\mathcal{B}$ the unitary given by the diagonal matrix whose main diagonal is the vector $(e^{-2\pi i(\ell-1)(m-1)/d})_{m=1}^d$.

The correctness of the protocol follows easily from the following observations. After step (2) the register pair $(\mathcal{A}, \mathcal{B})$ is in state $(FU \otimes I)vv^*(FU \otimes I)^*$. For a matrix $A \in \mathbb{C}^{d \times d}$ with columns $a_1, \ldots, a_d$ and rows $b_1, \ldots, b_d$, we have

$$\sum_{\ell=1}^{d} a_\ell \otimes e_\ell = \sum_{\ell=1}^{d} e_\ell \otimes b_\ell.$$

Since the ℓth row of the matrix FU is given by

$$\sum_{m=1}^{d} c_m\, e_m \text{ where } c_m = u_m\, e^{2\pi i(\ell-1)(m-1)/d},$$

this is exactly the state of Bob's register $\mathcal{B}$ after step (3). Thus, after step (4) Bob's register is in the state $\rho = uu^*$ as desired.

Chapter 3

Graph parameters

In this thesis we will discuss various graph parameters. Here we introduce them and present some of their properties.

3.1 Basic notations

We start by defining some basic graph theory notation.

Throughout, all graphs are assumed to be finite, undirected and without loops. For a graph $G = (V, E)$, the sets V and E denote its vertex and edge set, respectively (equivalently denoted by $V(G)$ and $E(G)$). Given two vertices $u, v \in V(G)$, we write $u \simeq v$ if u, v are adjacent or equal and we write $u \sim v$ when u and v are adjacent, in which case the corresponding edge is denoted as $\{u, v\}$ or simply as uv. 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 non-adjacent in G. A subgraph of a graph G is a graph formed from a subset of the vertices and edges of G. An induced subgraph of G is a subgraph that contains all the edges whose endpoints belong to the vertex set.

We denote with K_t the *complete graph* on t vertices, where each pair of distinct vertices is adjacent, and with C_n the *n-cycle*, where the n vertices are connected through a single cycle. The *disjoint union* between the two graphs G and H is denoted by $G + H$, where its vertex set is the disjoint union between $V(G)$ and $V(H)$ and the edge set is equal to $E(G) \cup E(H)$. We denote with G^{+t} the disjoint union of t copies of G; i.e., $V(G^{+t}) = V(G) \times [t]$ and the vertices (u, i) and (v, j) are adjacent if $u \sim v$ in G and $i = j \in [t]$. The *Cartesian product graph* $G \square K_t$ has $V(G) \times [t]$ as vertex set and two vertices (u, i) and (v, j) are adjacent if ($u = v$ and $i \neq j$) or if ($u \sim v$ and $i = j$). The *strong graph product* $G \boxtimes H$ of G and H is the graph whose vertex set is the cartesian product $V(G) \times V(H)$ and where two distinct vertices (u_1, u_2), (v_1, v_2) are adjacent if and only if it holds that $u_1 = v_1$ or $\{u_1, v_1\} \in E(G)$ and that $u_2 = v_2$ or $\{u_2, v_2\} \in E(H)$. The *m-th strong graph power* of G, denoted by $G^{\boxtimes m}$, is the

strong product graph of m copies of G. Its vertex set is the cartesian product of m copies of $V(G)$ and the pair of distinct vertices $(u_1 \dots, u_m), (v_1, \dots, v_m)$ forms an edge in $G^{\boxtimes m}$ if $u_i \simeq v_i$ in G for all $i \in [m]$. Similarly, the *disjunctive product* (or *coproduct*) of G and H is denoted by the graph $G * H$. Its vertex set is $V(G) \times V(H)$ and two vertices (u_1, u_2) and (v_1, v_2) are adjacent if and only if either $\{u_1, v_1\} \in E(G)$ or $\{u_2, v_2\} \in E(H)$. We denote by G^{*m} the m-th disjunctive power of G, where the vertex set is the cartesian product of m copies of $V(G)$ and the vertices $(u_1 \dots, u_m), (v_1, \dots, v_m)$ form an edge in G^{*m} if there exists a $j \in [m]$ such that $\{u_j, v_j\} \in E(G)$. One can easily check that

$$G^{\boxtimes m} = \overline{\overline{G}^{*m}}. \tag{3.1}$$

An *automorphism* of a graph G is a permutation π of $V(G)$ that preserves the 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 every pair of vertices $u, v \in V(G)$ there is an automorphism $\pi : V(G) \to V(G)$ such that $\pi(u) = v$. Moreover, the graph G is *edge-transitive* if for every pair of edges $\{u_1, v_1\}, \{u_2, v_2\} \in E(G)$, there exists an automorphism $\pi : V(G) \to V(G)$ where $\pi(u_1) = u_2$ and $\pi(v_1) = v_2$. A *homomorphism* from a graph H to a graph G is a map $\phi : V(H) \to V(G)$ such that every edge $\{u, v\}$ in H is mapped to an edge $\{\phi(u), \phi(v)\}$ in G. If such a map exists, we write $H \longrightarrow G$.

3.2 Classical graph parameters

3.2.1 Stability number and chromatic number

Given a graph G, a stable set of G is a subset of pairwise non-adjacent vertices. The *stability number* $\alpha(G)$ is the cardinality of the largest stable set in G. A proper coloring of a graph is an assignment of colors to the vertex set such that adjacent vertices receive different colors. The *chromatic number* $\chi(G)$ is the minimum number of colors needed for a proper coloring of G. As each color class must define a subset of pairwise non-adjacent vertices, $\chi(G)$ is also the minimum number of stable sets one needs to vertex-cover the graph G. Thus, the inequality $\chi(G)\alpha(G) \geq |V(G)|$ holds trivially. The stability and the chromatic numbers are NP-hard [Kar72] and also hard to approximate [Hås99, FK98].

Another interesting relationship between these two parameters was shown by Chvátal [Chv73] who related the chromatic number of a graph G to the stability number of an appropriate graph product.

3.2.1. THEOREM (CHVÁTAL [CHV73]). *For any graph G and any integer $t \geq 1$, we have $\chi(G) \leq t$ if and only if $\alpha(G \square K_t) = |V(G)|$. Hence, $\chi(G)$ is the minimum $t \in \mathbb{N}$ for which $\alpha(G \square K_t) = |V(G)|$ holds.*

A clique is a set of pairwise adjacent vertices and the *clique number* $\omega(G)$ is the maximum cardinality of a clique in G. Clearly, the clique number of a graph G is the stability number of its complement; that is, $\omega(G) = \alpha(\overline{G})$. As in any proper coloring all elements of a clique must receive different colors, we have $\omega(G) \leq \chi(G)$.

Given two integers $a \geq b \geq 1$, an (a,b)-coloring is an assignment of b colors, out of a available ones, to each vertex of the graph such that adjacent vertices have no colors in common. The *fractional chromatic number* $\chi_f(G)$ is the minimum ratio a/b such that there exists an (a,b)-coloring. Equivalently, $\chi_f(G)$ is the smallest $\sum_{h=1}^{k} \lambda_h$ for which there exist stable sets $S_1, \ldots, S_k$ of G and nonnegative scalars $\lambda_1, \ldots, \lambda_k$ such that $\sum_{h:v \in S_h} \lambda_h = 1$ for all $v \in V(G)$. By the latter definition, one can see that the fractional chromatic number can be written as a linear program. Nevertheless, computing $\chi_f(G)$ is an NP-hard problem [LY94]. Clearly, we have

$$\omega(G) \leq \chi_f(G) \leq \chi(G) \text{ and } \alpha(G) \leq \chi_f(\overline{G}) \leq \chi(\overline{G}).$$

We remark that the separation between $\omega(G)$ and $\chi_f(G)$ can be arbitrarily large. Indeed, there exists a family of graphs M_n, called Mycielski graphs, such that $\omega(M_n) = 2$ for every $n \in \mathbb{N}$, while for every $k \in \mathbb{R}$ there exists a number $n_k \in \mathbb{N}$ such that $\chi_f(M_{n_k}) \geq k$ [LPU95]. Such a large separation however cannot exist between $\chi_f(G)$ and $\chi(G)$, for which Lovász [Lov75] proved that $\chi(G)/(1 + \ln \alpha(G)) \leq \chi_f(G)$. Nonetheless, the fractional chromatic and the chromatic numbers can differ significantly. To see this, we define the Kneser graph $K_{a:b}$, for $a, b \in \mathbb{N}$ where $a \geq 2b$, to be the graph whose vertices are all the subsets of size b of $[a]$ and where two vertices are adjacent if the sets are disjoint. Lovász [Lov79] showed that $\chi_f(K_{a:b}) = a/b$, while he showed in [Lov78] that $\chi(K_{a:b}) = a - 2b + 2$.

Note that we can reformulate all the parameters we have introduced so far using graph homomorphism: the stability number $\alpha(G)$ is the maximum integer t for which there exists a homomorphism from the complete graph K_t to $\overline{G}$; $\chi(G)$ is the minimum $t \in \mathbb{N}$ such that $G \longrightarrow K_t$ and $\chi_f(G)$ is the minimum a/b such that there is a graph homomorphism from G to the Kneser graph $K_{a:b}$.

We have already seen that the inequality $\chi(G)\alpha(G) \geq |V(G)|$ holds for any graph G. There exists a similar, stronger relationship between the fractional chromatic number and the stability number: $\chi_f(G)\alpha(G) \geq |V(G)|$ with equality if the graph is vertex-transitive (see Corollary 3.2.10 below).

There is yet another way to reformulate the parameters $\alpha(G), \chi(G)$ and $\chi_f(G)$: as linear optimization problems over the completely positive cone $\mathcal{CP}$. These characterizations will be very useful in Section 5.1. Using a result of Motzkin and Straus [MS65], de Klerk and Pasechnik [dKP02] showed the following reformulation of the stability number.

3.2.2. THEOREM (DE KLERK–PASECHNIK [DKP02]). *For any graph G, its stability number $\alpha(G)$ is equal to the optimum value of the following program:*

$$\max \langle J, X \rangle \ \ s.t. \ \ X \in \mathcal{CP}^{|V(G)|}, \quad \mathsf{Tr}(X) = 1, \quad X_{uv} = 0 \quad \forall \{u, v\} \in E(G).$$

Combining Theorems 3.2.1 and 3.2.2, Gvozdenović and Laurent [GL08] obtained a reformulation of the chromatic number as a completely positive optimization program. Furthermore, Dukanovic and Rendl [DR10] gave the following reformulation for the fractional chromatic number $\chi_{\mathsf{f}}(G)$.

3.2.3. THEOREM (DUKANOVIC–RENDL [DR10]). *For any graph G, its fractional chromatic number $\chi_{\mathsf{f}}(G)$ is equal to the optimum value of the following program:*

$$\begin{gathered}\min \ \ t \ \ s.t. \ \ X \in \mathcal{CP}^{|V(G)|}, \quad X - J \succeq 0, \quad X_{uu} = t \quad \forall u \in V(G), \\ X_{uv} = 0 \quad \forall \{u, v\} \in E(G).\end{gathered}$$

3.2.2 Shannon capacity

To study the zero-error channel coding problem (see Section 6.1 for details), Shannon [Sha56] introduced a graph parameter, known as the Shannon capacity of a graph, which is defined from the stability number of strong graph products.

The *Shannon capacity of a graph* is

$$\Theta(G) = \lim_{n \to \infty} \sqrt[n]{\alpha(G^{\boxtimes n})}.$$

Combining a result due to Fekete (Lemma 6.1.1) with the observation that the sequence $\left(\sqrt[n]{\alpha(G^{\boxtimes n})} \right)_{n \in \mathbb{N}}$ is monotone nondecreasing, one can derive that the Shannon capacity of graph is also equivalent to $\Theta(G) = \sup_n \sqrt[n]{\alpha(G^{\boxtimes n})}$ (see Section 6.1 for details). Furthermore, the following chain of inequalities holds:

$$\alpha(G) \leq \Theta(G) \leq \chi_{\mathsf{f}}(\overline{G}) \leq \chi(\overline{G}). \tag{3.2}$$

The left most inequality follows directly from the supremum formulation of the Shannon capacity and the right most one is trivial. Moreover, the inequality $\Theta(G) \leq \chi_{\mathsf{f}}(\overline{G})$ will follow from (3.4) below.

All the inequalities in (3.2) can be strict. For this, consider the 5-cycle graph C_5. One can easily compute that $\alpha(C_5) = 2$, $\chi_{\mathsf{f}}(C_5) = 5/2$ and that $\chi(C_5) = 3$. Moreover, Shannon [Sha56] showed that $\alpha(C_5^{\boxtimes 2}) = 5$ and, thus, $\alpha(C_5) < \sqrt{\alpha(C_5^{\boxtimes 2})} \leq \Theta(C_5)$. Only after a couple of decades Lovász [Lov79] was able to prove that $\Theta(C_5) = \sqrt{5}$.

The fact that the Shannon capacity is hard to determine even for small graphs should not be surprising since it is defined as the limit of a sequence of NP-hard parameters. Interestingly, we do not even know whether the Shannon capacity is decidable.

3.2.3 Lovász theta number

We briefly mentioned that Lovász was able to prove that the Shannon capacity of the 5-cycle is equal to $\sqrt{5}$. He managed to do this by introducing a parameter $\vartheta(G)$, known as the *Lovász theta number*, which is an upper bound on the Shannon capacity.

3.2.4. DEFINITION. [Lovász theta number] Let G be a graph with $|V(G)| = n$, the Lovász theta number $\vartheta(G)$ of G is defined as follows:

$$\begin{aligned}
\vartheta(G) = \max \ & \langle J, X\rangle && = \min \ && t \\
\text{s.t.} \ & X \in \mathcal{S}_+^n && \text{s.t.} \ && Z \in \mathcal{S}_+^n,\ Z - J \in \mathcal{S}_+^n \\
& \mathsf{Tr}(X) = 1 && && Z_{uu} = t \ \ \forall\, u \in V(G) \\
& X_{uv} = 0 \ \ \forall\, \{u,v\} \in E(G); && && Z_{uv} = 0 \ \ \forall\, \{u,v\} \in E(\overline{G}).
\end{aligned}$$

Since the Lovász theta number is the optimum value of a positive semidefinite program, it can be computed up to any precision in polynomial time in the number of vertices. Furthermore, this parameter is a well-known bound for both the stability and the chromatic numbers. Indeed, Lovász [Lov79] showed the following 'sandwich' inequalities:

$$\alpha(G) \leq \vartheta(G) \leq \chi_{\mathsf{f}}(\overline{G}) \leq \chi(\overline{G}). \tag{3.3}$$

By definition, the parameter $\vartheta(G)$ is monotone non-decreasing under taking subgraphs and from (3.3) we get that $\vartheta(K_t) = 1$, $\vartheta(\overline{K}_t) = t$ where $t \in \mathbb{N}$ and K_t is the complete graph. Moreover, in [Lov79] it is shown that for any graph G we have $\vartheta(G)\vartheta(\overline{G}) \geq |V(G)|$, with equality if the graph is vertex-transitive (see also Corollary 3.2.10 below).

As can be seen in the following lemma, the parameter $\vartheta(G)$ behaves well under various graph products. (For the proofs and further properties of the Lovász theta number, we refer the reader to the survey of Knuth [Knu94].)

3.2.5. LEMMA. *Consider two graphs G and H. The following identities hold for the Lovász theta number: (i) $\vartheta(G + H) = \vartheta(G) + \vartheta(H)$; (ii) $\vartheta(G \boxtimes H) = \vartheta(G)\vartheta(H)$; (iii) $\vartheta(G * H) = \vartheta(G)\vartheta(H)$.*

Using Lemma 3.2.5 (*ii*) and by (3.3), we get that $\vartheta(G)$ is an upper bound for the Shannon capacity. Therefore, we have that

$$\alpha(G) \leq \Theta(G) \leq \vartheta(G) \leq \chi_{\mathsf{f}}(\overline{G}) \leq \chi(\overline{G}). \tag{3.4}$$

Lovász was able to conclude that the Shannon capacity of the 5-cycle is equal to $\sqrt{5}$ since $\vartheta(C_5) = \sqrt{5}$ and therefore $\sqrt{5} \leq \Theta(C_5) \leq \vartheta(C_5) = \sqrt{5}$. In general, however, $\vartheta(G)$ is not a tight bound on $\Theta(G)$. We will see examples of such graphs in Chapter 6.

Inspired by the chain of inequalities (3.4), Berge introduced the notion of a perfect graph. This is a graph such that for every induced subgraph $H \subseteq G$ we have $\alpha(H) = \chi(\overline{H})$. This in particular implies that for a perfect graph both the stability number and the Shannon capacity can be computed in polynomial time. As the complement of a perfect graph is still perfect [Lov72], also the chromatic number of a perfect graph can be efficiently computed.

Generalizations of the Lovász theta number. Several strengthenings of the Lovász theta number toward $\alpha(G)$ and $\chi(G)$ have been proposed, in particular, the parameters $\vartheta'(G)$, introduced independently by Schrijver [Sch79] and McEliece et al. [MRR78], and $\vartheta^+(G)$, introduced by Szegedy [Sze94].

3.2.6. DEFINITION. Let G be a graph with $|V(G)| = n$, the parameters $\vartheta'(G)$ and $\vartheta^+(G)$ are defined as follows:

$$\begin{aligned} \vartheta'(G) = \max\ & \langle J, X\rangle & \vartheta^+(G) = \min\ & t \\ \text{s.t. } & X \in \mathcal{DNN}^n & \text{s.t. } & Z \in \mathcal{DNN}^n,\ Z - J \in \mathcal{S}^n_+ \\ & \mathsf{Tr}(X) = 1 & & Z_{uu} = t\ \forall\, u \in V(G) \\ & X_{uv} = 0\ \forall \{u,v\} \in E(G); & & Z_{uv} = 0\ \forall\, \{u,v\} \in E(\overline{G}). \end{aligned}$$

Dukanovic and Rendl [DR10] introduced a further generalization of the Lovász theta number.

3.2.7. DEFINITION. Let G be a graph with $|V(G)| = n$ and $\mathcal{K}$ be a convex cone such that $\mathcal{CP} \subseteq \mathcal{K} \subseteq \mathcal{S}_+$. The parameters $\vartheta^{\mathcal{K}}(G)$ and $\Theta^{\mathcal{K}}(G)$ are defined as follows:

$$\begin{aligned} \vartheta^{\mathcal{K}}(G) = \sup\ & \langle J, X\rangle & \Theta^{\mathcal{K}}(G) = \inf\ & t \\ \text{s.t. } & X \in \mathcal{K}^n & \text{s.t. } & Z \in \mathcal{K}^n,\ Z - J \in \mathcal{S}^n_+ \\ & \mathsf{Tr}(X) = 1 & & Z_{uu} = t\ \forall\, u \in V(G) \\ & X_{uv} = 0\ \forall\, \{u,v\} \in E(G); & & Z_{uv} = 0\ \forall\, \{u,v\} \in E(G). \end{aligned}$$

3.2.8. REMARK. We observe a 'monotonicity' property for the program above characterizing $\vartheta^{\mathcal{K}}(G)$, that will be useful later in Section 5.1.2. Set $n = |V(G)|$ and consider scalars $1 \leq t < T$. Assume that a matrix X is feasible for the program defining $\vartheta^{\mathcal{K}}(G)$ with value $\langle J, X\rangle = T$. Then we have that the matrix $X' = \frac{t-1}{T-1}X + \frac{T-t}{n(T-1)}I$ is again feasible for $\vartheta^{\mathcal{K}}(G)$ and it has value $\langle J, X'\rangle = t$.

3.2.9. PROPOSITION (DUKANOVIC–RENDL [DR10]). *Let $\mathcal{K}$ be a cone such that $\mathcal{CP} \subseteq \mathcal{K} \subseteq \mathcal{S}_+$. Then, we have $\vartheta^{\mathcal{K}}(G)\Theta^{\mathcal{K}}(G) \geq |V(G)|$. Moreover, it holds with equality if the graph G is vertex-transitive and the cone $\mathcal{K}$ is invariant under simultaneous permutation of the rows and the columns.*

Clearly, if in Definition 3.2.7 we replace the cone $\mathcal{K}$ with the positive semidefinite cone, we get back the original Lovász theta number. Moreover, by Definition 3.2.6 we have that $\vartheta^{\mathcal{DNN}}(G) = \vartheta'(G)$ and $\Theta^{\mathcal{DNN}}(G) = \vartheta^+(\overline{G})$. At last, if in Definition 3.2.7 we set $\mathcal{K} = \mathcal{CP}$, we find $\alpha(G)$ and $\chi_{\mathrm{f}}(G)$ (respectively by Theorems 3.2.2 and 3.2.3). Summarizing, we get:

$$\vartheta^{\mathcal{DNN}}(G) = \vartheta'(G),\ \vartheta^{\mathcal{CP}}(G) = \alpha(G), \tag{3.5}$$

$$\Theta^{\mathcal{DNN}}(G) = \vartheta^+(\overline{G}),\ \Theta^{\mathcal{CP}}(G) = \chi_{\mathrm{f}}(G). \tag{3.6}$$

Combining Proposition 3.2.9 with the above considerations, we obtain the following relationships.

3.2.10. COROLLARY. *For a graph G, the following inequalities hold:*

$$\alpha(G)\chi_{\mathrm{f}}(G) \geq |V(G)|, \vartheta(G)\vartheta(\overline{G}) \geq |V(G)| \text{ and } \vartheta'(G)\vartheta^+(\overline{G}) \geq |V(G)|.$$

Moreover, if G is vertex-transitive graph all the above inequalities hold with equality.

As $\mathcal{CP}^n \subseteq \mathcal{DNN}^n \subseteq \mathcal{S}^n_+$ and by (3.5) and (3.6), we get the following inequalities, which refine (3.3):

$$\alpha(G) \leq \vartheta'(G) \leq \vartheta(G) \leq \vartheta^+(G) \leq \chi_{\mathrm{f}}(\overline{G}) \leq \chi(\overline{G}).$$

All the above inequalities can be strict and the separation between $\vartheta'(G)$ and $\vartheta(G)$, as well as the one between $\vartheta(G)$ and $\vartheta^+(G)$, can be exponentially large. Indeed, let n be an even integer and consider the graph $G_n = (\{0,1\}^n, E)$ where E is given by all pairs of strings with Hamming distance in $\{n/2, \ldots, n\}$. Samorodnitsky [Sam98] showed that $\vartheta'(\overline{G_n}) \leq O(n)$ while $2^{\Omega(n)} \leq \vartheta(\overline{G_n})$. Moreover, one can easily check that the graph G_n is vertex-transitive and using Corollary 3.2.10 we obtain that there is an exponential separation also between the parameters $\vartheta(G_n)$ and $\vartheta^+(G_n)$.

Using Lemma 3.2.5 (*ii*), we have derived that the Lovász theta number is an upper bound on the Shannon capacity. Such reasoning cannot be applied to the parameter $\vartheta'(G)$ because Cubitt et al. [CMR+14] exhibited a graph G for which $\vartheta'(G \boxtimes G) > \vartheta'(G)^2$ holds. In the same paper and for the same graph, it was also proven that $\vartheta^+(G * G) < \vartheta^+(G)^2$ and therefore that ϑ^+ is not multiplicative under the disjunctive product.

3.2.4 Orthogonal rank

A d-coloring of a graph G can be thought of as a map that assigns to each vertex one of the canonical basis vectors $\{e_1, \ldots, e_d\}$ of $\mathbb{C}^d$ such that adjacent vertices receive distinct vectors. As a straightforward generalization, a *d-dimensional orthogonal representation* of a graph G is a map f from the vertex set to nonzero

vectors in $\mathbb{C}^d$ such that adjacent vertices are mapped to orthogonal vectors. (We stress that we consider the representations over complex vectors and not, as more usual in the combinatorial literature, over real ones and that orthogonalities are required for adjacent vertices.)

3.2.11. DEFINITION. [Orthogonal rank] The orthogonal rank $\xi(G)$ of a graph G is the minimum integer d such that there exists a d-dimensional orthogonal representation of G.

Clearly, we have that $\xi(G) \leq \chi(G)$ and Peeters [Pee96, Theorem 3.1] proved that the orthogonal rank is an NP-hard parameter. Following [CMN$^+$07], we introduce a slight variation of the orthogonal rank.

3.2.12. DEFINITION. For a graph G, $\xi'(G)$ is the minimum $d \in \mathbb{N}$ for which there exists a d-dimensional orthogonal representation f of G such that for each $u \in V(G)$ the entries of the vector $f(u)$ all have absolute value one.

In the paper where Lovaśz introduced the parameter $\vartheta(G)$, he proved that this is a lower bound on the minimum dimension of an orthogonal representation where the vectors are real valued. We show that the Lovász theta number, and in particular $\vartheta^+(\overline{G})$, is also a lower bound for the orthogonal rank $\xi(G)$, where the vectors can have complex entries. The proof is an adaptation to the complex case of a known proof [Lau14].

3.2.13. LEMMA. *For any graph G, we have that $\vartheta^+(\overline{G}) \leq \xi(G)$.*

PROOF: Let $n = |V(G)|$ and label the vertices of the graph G by $\{1, 2, \ldots, n\}$. Suppose that the orthogonal rank of G is equal to d and that $u_1, \ldots, u_n \in \mathbb{C}^d$ are the nonzero vectors forming an orthogonal representation of G. By scaling, we can without loss of generality assume that $u_1, \ldots, u_n$ are unit vectors. For every vertex of the graph $i \in [n]$, define a matrix $U_i = u_i u_i^*$ and $U_0 = I \in \mathcal{S}^d$. Let Z be a $(n+1) \times (n+1)$ matrix where the i, j-th entry $Z_{ij} = \langle U_i, U_j \rangle = \mathsf{Tr}(U_j^* U_i)$ for every $i, j \in \{0\} \cup [n]$. The matrix Z is positive semidefinite, as it is the Gram matrix of a set of complex vectors, and is also real valued. Moreover, we have that $Z_{00} = d$, $Z_{0i} = \langle I, u_i u_i^* \rangle = 1$ and $Z_{ii} = \langle u_i u_i^*, u_i u_i^* \rangle = (u_i^* u_i)^2 = 1$ for all $i \in V(G)$ and that $Z_{ij} = (u_i^* u_j)(u_j^* u_i) \geq 0$ for all $i, j \in V(G)$ with equality if $ij \in E(G)$. By taking the Schur complement in Z with respect to the entry Z_{00} (see (2.1)), we obtain a new symmetric positive semidefinite matrix X with $X_{ii} = 1 - 1/d$ for all $i \in V(G)$, $X_{ij} \geq -1/d$ for all $i, j \in V(G)$ and with equality if $ij \in E(G)$. The matrix $Y = dX + J$ is then a feasible solution for the minimization program in Definition 3.2.6 of $\vartheta^+(\overline{G})$ with value d. We conclude that $d = \xi(G) \geq \vartheta^+(\overline{G})$. □

Therefore, the following chain of inequalities hold:

$$\vartheta(\overline{G}) \leq \vartheta^+(\overline{G}) \leq \xi(G) \leq \xi'(G) \leq \chi(G), \tag{3.7}$$

where for the latter one we use the following observation. Let $t = \chi(G)$ and fix an optimal coloring of G. To any vertex with color ℓ, we assign the vector $f(\ell) \in \mathbb{C}^t$ whose k-th entry is $f(\ell)_k = e^{2\pi i \ell(k-1)/t}$. As the vectors $\{f(\ell)\}_{\ell \in [t]}$ form an orthogonal basis, this is a t-orthogonal representation where all the entries have absolute value one and thus $\xi'(G) \leq \chi(G)$.

3.3 Quantum graph parameters

We now introduce quantum variants of the chromatic and stability numbers.

3.3.1 Quantum chromatic number

Consider the nonlocal game where two players want to convince a referee that they can color a graph using at most a fixed amount of colors. Fix a graph G and an integer $t \in \mathbb{N}$. As questions each player receives a vertex of the graph to which they have to answer a color from $\{1, \ldots, t\}$. To win the game, they have to answer the same color upon receiving the same vertex and different colors if the vertices are adjacent. We say that the players can classically t-color the graph G if the above nonlocal game admits a perfect classical strategy. We will now show that such a strategy exists if and only if t is least the chromatic number of G. Take any perfect classical strategy, which we can assume to be deterministic. Since the players have to answer the same color upon receiving the same vertex, both players use the same map $f : V(G) \to [t]$ as strategy. This map assigns different color to adjacent vertices and therefore induces a coloring of the graph. For the other direction, any coloring of the graph can be used as a strategy: each player upon receiving a vertex answer the color of that vertex. Therefore, the chromatic number $\chi(G)$ is the minimum number for which the players can classically color the graph. Similarly, we define the *quantum chromatic number* $\chi_q(G)$ as the minimum number of colors for which the game admits a perfect quantum strategy. Reformulating, this means that $\chi_q(G)$ is the minimum integer $t \in \mathbb{N}$ for which there exists an entangled state σ and measurements $\{E_u^i\}_{i \in [t]}$, $\{F_u^i\}_{i \in [t]}$ for each $u \in V(G)$ such that:

$$\begin{aligned} \mathsf{Tr}((E_u^i \otimes F_v^i)\sigma) &= 0 \quad \forall i \in [t],\, \forall \{u, v\} \in E(G); \\ \mathsf{Tr}((E_u^i \otimes F_u^j)\sigma) &= 0 \quad \forall i \neq j \in [t],\, \forall u \in V(G). \end{aligned} \tag{3.8}$$

This parameter was first introduced in [AHKS06] and then more formally studied in [CMN$^+$07]. In particular, in the latter paper (see also [Rob13, Section 6.5])

it was proven that if a perfect quantum strategy exists then there is one with the following special form: E_u^i, F_u^i are $d \times d$ projectors all of the same rank (for some $d \in \mathbb{N}$), $E_u^i = F_u^{iT}$ for all $i \in [t]$ and $u \in V(G)$, and the state σ is the maximally entangled one (i.e., $\sigma = vv^*$ where $v = \frac{1}{\sqrt{d}} \sum_{\ell=1}^{d} e_\ell \otimes e_\ell$). Denote $D = I/\sqrt{d}$ and notice that $\mathrm{vec}(D) = v$. We get that:

$$\mathsf{Tr}((E_u^i \otimes F_v^j)vv^*) = \mathsf{Tr}(v^*(E_u^i \otimes F_v^j)v) = \mathsf{Tr}(D^*(E_u^i D F_v^j)) = \mathsf{Tr}(E_u^i F_v^j D^2),$$

where the first equivalence holds by the cyclicity of the trace, the second one due to the identity $(A \otimes B)\mathrm{vec}(X) = \mathrm{vec}(AXB^T)$, and the latter one due to the observations that D commutes with F_v^j and that $D = D^*$. Now, we observe that $\mathsf{Tr}(E_u^i F_v^{jT} D^2) = \mathsf{Tr}(E_u^i F_v^{jT} I/d) = \mathsf{Tr}(E_u^i F_v^{jT})/d = 0$ if and only if $E_u^i F_v^{jT} = 0$ (applying Proposition 2.1.3 (*i*)) and the latter is equivalent to $E_u^i E_v^j = 0$. Therefore, we can reformulate conditions (3.8) and get the following definition of the quantum chromatic number.

3.3.1. DEFINITION. [Quantum chromatic number [CMN$^+$07]] For a graph G, $\chi_q(G)$ is the minimum $t \in \mathbb{N}$ for which there exist $d \times d$ projectors E_u^i for $i \in [t]$, $u \in V(G)$ (for some $d \geq 1$) satisfying the conditions:

$$\begin{aligned} \textstyle\sum_{i\in[t]} E_u^i &= I \quad \forall u \in V(G), \\ E_u^i E_v^i &= 0 \quad \forall i \in [t], \forall \{u,v\} \in E(G), \\ E_u^i E_u^j &= 0 \quad \forall i \neq j \in [t], \forall u \in V(G). \end{aligned}$$

This parameter has recently received a notable amount of attention (see among others [AHKS06, CMN$^+$07, FILG11, SS12, MSS13, MR16, Ji13, PSS$^+$16, PT15]). Ji [Ji13] proved that $\chi_q(G)$ is an NP-hard parameter and the following inclusions are known to hold:

$$\vartheta(\overline{G}) \leq \vartheta^+(\overline{G}) \leq \chi_q(G) \leq \xi'(G) \leq \chi(G), \tag{3.9}$$

using Corollary 5.1.8, [CMN$^+$07, Proposition 7] and (3.7), respectively. One of the most interesting questions is to find and characterize graphs for which there is a separation between the quantum chromatic number and its classical counterpart. Clearly there is no such separation when G is a perfect graph because then the identities $\alpha(\overline{G}) = \vartheta(\overline{G}) = \chi(G)$ hold. Moreover, the only graphs for which $\chi_q(G) = 1$ are the empty ones and the only graphs for which $\chi_q(G) = 2$ are the bipartite ones. Therefore, the smallest separation possible is to have a graph with $\chi_q(G) = 3$ and $\chi(G) = 4$ and such a graph was found in [FILG11]. In [AHKS06] they show that there exists a family of graphs for which the quantum chromatic number is exponentially smaller than the classical one. These graphs are known as the *orthogonality graphs* and they are defined as follows.

3.3.2. DEFINITION. [Orthogonality graph] For $k \in \mathbb{N}$ even, the *orthogonality graph* Ω_k has vertex set $\{-1,1\}^k$ and two vertices are adjacent if they are orthogonal.

By a well-known result of Frankl and Rödl (see Theorem 9.2.4), for k multiple of 4 large enough, we have that $\chi(\Omega_k) \geq \frac{|V(\Omega_k)|}{\alpha(\Omega_k)} \geq (\frac{2}{2-\varepsilon})^n$ where ε is a small positive constant. At the same time, by Definitions 3.2.12 and 3.3.2, we know that $\xi'(\Omega_k) \leq k$ holds and, using (3.9), we derive that $\chi_q(\Omega_k) \leq k$.

Beside the inequality $\chi_q(G) \leq \xi'(G)$, the only other non-trivial method to upper bound the quantum chromatic number is given by the following proposition, which holds for graphs having small dimensional real valued orthogonal representations.

3.3.3. PROPOSITION (CAMERON ET AL. [CMN$^+$07]). *Let G be a graph with an orthogonal representation in $\mathbb{R}^c$. If $c \in \{3,4\}$ then $\chi_q(G) \leq 4$; if $5 \leq c \leq 8$ then $\chi_q(G) \leq 8$.*

In [PSS$^+$16, PT15], Paulsen and coauthors have introduced many variants of the quantum chromatic number motivated by the study of quantum correlations. We briefly recall two of them, the parameters $\chi_{qa}(G)$ and $\chi_{qc}(G)$, which we will later use in Section 5.3.

Consider a conditional bipartite probability distribution $(P(i,j|u,v))$ with input sets $X = Y = V(G)$ and output sets $A = B = [t]$. Recall that $\mathcal{Q}$ is the set of quantum correlations (Definition 2.4.3) and that $\mathcal{Q}_c$ is the set of probability distributions arising from the relativistic field theory point of view (Definition 2.4.4). We can rewrite constraints (3.8) by defining the linear map $\mathcal{L}_{G,t} : \mathbb{R}^{(nt)^2} \to \mathbb{R}$

$$\mathcal{L}_{G,t}(P) = \sum_{i \neq j \in [t], u \in V(G)} P(i,j|u,u) + \sum_{i \in [t], uv \in E(G)} P(i,i|u,v).$$

Then the players have a perfect quantum winning strategy if and only if the probability distribution P lies in $\mathcal{Q}$ and satisfies $\mathcal{L}_{G,t}(P) = 0$. Therefore, $\chi_q(G)$ is equal to

$$\chi_q(G) = \min t \in \mathbb{N} \text{ s.t. } \exists P \in \mathcal{Q} \text{ with } \mathcal{L}_{G,t}(P) = 0. \tag{3.10}$$

The parameter $\chi_{qa}(G)$ defined in [PSS$^+$16] asks the probability distribution P to be in the closure of $\mathcal{Q}$:

$$\chi_{qa}(G) = \min t \in \mathbb{N} \text{ s.t. } \exists P \in \mathsf{cl}(\mathcal{Q}) \text{ with } \mathcal{L}_{G,t}(P) = 0. \tag{3.11}$$

In [PSS$^+$16] (see also [PT15]) the parameter $\chi_{qc}(G)$ is defined as

$$\chi_{qc}(G) = \min t \in \mathbb{N} \text{ s.t. } \exists P \in \mathcal{Q}_c \text{ with } \mathcal{L}_{G,t}(P) = 0. \tag{3.12}$$

In [PSS+16] it is shown that $\chi_{qc}(G)$ can be computed by a positive semidefinite program (after rounding). This result is existential in the sense that the semidefinite program is not explicitly known. For this the authors of [PSS+16] use the semidefinite programming hierarchy developed by Navascués, Pironio and Acín [NPA08] for noncommutative polynomial optimization. As pointed out in [PSS+16], in view of the inclusions in (2.5), the following relationships hold between the parameters:

$$\chi_{qc}(G) \leq \chi_{qa}(G) \leq \chi_q(G).$$

Furthermore, if Connes' embedding conjecture has a positive answer then the identity $\chi_{qc}(G) = \chi_{qa}(G)$ holds for every graph.

3.3.2 Quantum stability number

Mančinska and Roberson [MR16] introduced a quantum version of the stability number. For a fixed graph G and an integer t, suppose that two players want to convince a referee that there exists a stable set of cardinality at least t. The nonlocal game is as follows: each player receives a number from $\{1, \dots t\}$ as question and has to answer with a vertex of the graph. They win if upon receiving the same number they reply with the same vertex as answer and non-adjacent vertices if the questions were different. We are interested in determining for which $t \in \mathbb{N}$ there exists a perfect classical (or quantum) strategy for this nonlocal game. A possible classical strategy is to label the vertices of a (maximum) stable set of the graph with $\{1, \dots, t\}$ and upon receiving a number each player answers with the corresponding vertex. Actually, one can easily see that any perfect classical strategy is symmetric (meaning that both players apply the same map) and identifies a stable set. Therefore, the stability number $\alpha(G)$ is the maximum integer $t \in \mathbb{N}$ such that the players have a perfect classical strategy. Analogously, the *quantum stability number* $\alpha_q(G)$ is the maximum $t \in \mathbb{N}$ for which there exists a perfect quantum strategy. Using a similar reasoning as the one used for the quantum chromatic number, in [MR16] it is proven that $\alpha_q(G)$ can be equivalently reformulated as follows.

3.3.4. DEFINITION. [Quantum stability number [MR16]] For a graph G, $\alpha_q(G)$ is the maximum $t \in \mathbb{N}$ for which there exist $d \times d$ projectors E_i^u for $i \in [t]$, $u \in V(G)$ (for some $d \geq 1$) satisfying the conditions:

$$\begin{aligned} \textstyle\sum_{u \in V(G)} E_i^u &= I \quad \forall i \in [t], \\ E_i^u E_j^v &= 0 \quad \forall i \neq j \in [t], \forall u \simeq v \in V(G), \\ E_i^u E_i^v &= 0 \quad \forall i \in [t], \forall u \neq v \in V(G). \end{aligned}$$

In [MR16] it is shown that $\alpha(G) \leq \alpha_q(G) \leq \vartheta'(G) \leq \vartheta(G)$ holds and that there exists a quantum equivalent to Theorem 3.2.1:

$$\chi_q(G) \leq t \text{ if and only if } \alpha_q(G \square K_t) = |V(G)| \text{ holds.} \tag{3.13}$$

As noticed in [MR16] (using an idea from [MSS13]), the above relation can be used to show separations between the quantum stability number and the classical one starting from graphs G where $\chi_q(G) < \chi(G)$. Indeed, let $k \geq 8$ be a multiple of 4 and consider the orthogonality graph Ω_k (Definition 3.3.2) for which we know that $\chi_q(\Omega_k) \leq k < \chi(\Omega_k)$. Vizing [Viz63] proved that, for any pair of graphs G and H, $\alpha(G\square H) \leq \min\{\alpha(G)|V(H)|, \alpha(H)|V(G)|\}$ and therefore $\alpha(\Omega_k\square K_k) \leq k(2-\varepsilon)^k < 2^k = |V(\Omega_k)| = \alpha_q(\Omega_k\square K_k)$ where in the last identity we used (3.13). Furthermore, since $\chi_q(G)$ is an NP-hard parameter, relation (3.13) implies that also $\alpha_q(G)$ is NP-hard.

In Section 3.2.1, we have seen that the inequality $\chi(G)\alpha(G) \geq |V(G)|$ always holds. Interestingly, in [MR16] it is shown that if k is a multiple of 4 but not a power of 2 then $\chi_q(\Omega_k)\alpha_q(\Omega_k) < |V(\Omega_k)|$.

3.3.3 Zero-error information theory graph parameters

In Sections 6.2 and 8.1.1, we will introduce two further quantum variants of the stability and chromatic numbers: the entangled stability number $\alpha^\star(G)$ (Definition 6.2.1) and the entangled chromatic number $\chi^\star(G)$ (Definition 5.1.2), which arise in the context of zero-error communication scenarios. To avoid unnecessary repetitions, we refer the reader to the respective sections for the definitions and properties of these parameters. As we will see in Corollaries 5.1.8 and 5.1.17, the following chain of inequality holds:

$$\alpha(G) \leq \alpha_q(G) \leq \alpha^\star(G) \leq \vartheta(G) \leq \chi^\star(\overline{G}) \leq \chi_q(\overline{G}) \leq \chi(\overline{G}). \tag{3.14}$$

Interestingly, it is not known whether the inequalities $\alpha_q(G) \leq \alpha^\star(G)$ and $\chi^\star(\overline{G}) \leq \chi_q(\overline{G})$ can be strict.

We finish this chapter by exhibiting an example of a graph for which both the inequalities $\alpha_q(G) > \alpha(G)$ and $\chi_q(G) < \chi(G)$ hold. This graph is depicted in Figure 3.1 and it was used by Cubitt, Leung, Matthews, and Winter [CLMW10] to show a separation between $\alpha^\star(G)$ and $\alpha(G)$. It has 24 vertices which are defined by the vectors:

$$\begin{array}{llll}
1:(1,\ 0,\ 0,0) & 2:(0,\ 1,\ 0,\ 0) & 3:(0,\ 0,\ 1,\ 0) & 4:(0,0,\ 0,\ 1)\\
5:(0,\ 1,\ 1,0) & 6:(1,\ 0,\ 0,-1) & 7:(1,\ 0,\ 0,\ 1) & 8:(0,1,-1,\ 0)\\
9:(1,\ 1,\ 1,1) & 10:(1,-1,1,-1) & 11:(1,-1,-1,1) & 12:(1,1,-1,-1)\\
13:(1,\ -1,0,0) & 14:(1,\ 1,\ 0,\ 0) & 15:(0,\ 0,\ 1,\ 1) & 16:(0,0,\ 1,\ -1)\\
17:(-1,\ 1,\ 1,1) & 18:(1,\ 1,\ 1,-1) & 19:(1,-1,\ 1,\ 1) & 20:(1,1,-1,\ 1)\\
21:(1,\ 0,\ 1,0) & 22:(0,\ 1,\ 0,\ 1) & 23:(1,\ 0,\ -1,0) & 24:(0,1,\ 0,\ -1)
\end{array}$$

and two vertices are adjacent if the corresponding vectors are orthogonal. We observe that the sets $\mathcal{B}_1 = \{1,2,3,4\}$, $\mathcal{B}_2 = \{5,6,7,8\}$, $\mathcal{B}_3 = \{9,10,11,12\}$,

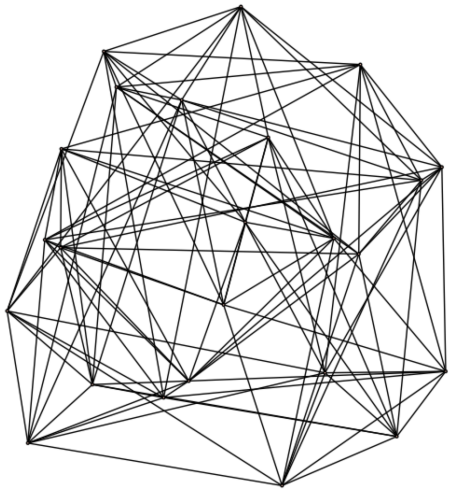

Figure 3.1: Graph for which $\alpha_q(G) > \alpha(G)$ and $\chi_q(G) < \chi(G)$.

$\mathcal{B}_4 = \{13, 14, 15, 16\}$, $\mathcal{B}_5 = \{17, 18, 19, 20\}$ and $\mathcal{B}_6 = \{21, 22, 23, 24\}$ are 6 disjoint cliques. One can check that $\alpha(G) = 5$ and $\chi(G) = 5$ hold. Moreover, by construction G has a 4-dimensional real valued orthogonal representation and, using Proposition 3.3.3, this implies that $\chi_q(G) \leq 4$. In particular, equality holds since the clique number is equal to 4. Cubitt et al. [CLMW10] showed that $\alpha^\star(G) \geq 6$ by proving that if a graph has orthogonal rank d and M disjoint d-cliques then $\alpha^\star(G) \geq M$ (see Theorem 6.2.4). We will use a very similar reasoning to show that $\alpha_q(G) \geq 6$. For every $k \in [24]$, let f_k be the normalized vector associated to vertex k. For $i \in [6]$ and $k \in [24]$, we define $E_i^k = f_k f_k^T$ if $k \in \mathcal{B}_i$ and $E_i^k = 0$ otherwise. One can easily check that the projectors E_i^k satisfy all the conditions of the program of Definition 3.3.4, which in turn implies that $\alpha_q(G) \geq 6$.

Chapter 4

The completely positive semidefinite cone

In this chapter we introduce the completely positive semidefinite cone $\mathcal{CS}_+$ and establish some of its basic properties. We investigate the relationship among the completely positive semidefinite cone, the completely positive cone and the doubly nonnegative cone, and their dual counterparts. Moreover, we present two different constructions that aim to approximate the $\mathcal{CS}_+$ cone. The first one is based on noncommutative trace polynomial optimization (Section 4.4) and the second one gives a hierarchy of polyhedral cones covering the interior of the completely positive semidefinite cone (Section 4.5). At last in Section 4.6, we give an explicit description of the closure of the completely positive semidefinite cone using the tracial ultraproduct of matrix algebras $\mathbb{R}^{k\times k}$.

The content of this chapter is based on the results of two papers: one is joint work with Monique Laurent [LP15] and the other is joint work with Sabine Burgdorf and Monique Laurent [BLP15].

4.1 Basic properties

Associated to any positive semidefinite matrix there is a set of vectors which forms its Gram representation. That is, for any positive semidefinite matrix $A \in \mathcal{S}^n$ there exists a set of vectors $x_1, \dots, x_n \in \mathbb{R}^d$, for some $d \in \mathbb{N}$, such that $A = (\langle x_i, x_j \rangle)_{i,j=1}^n$. Similarly, a matrix is completely positive if it has a Gram representation by nonnegative vectors. We now consider Gram representations by positive semidefinite matrices.

4.1.1. DEFINITION. A matrix $A \in \mathcal{S}^n$ is called *completely positive semidefinite* if there exists a set of positive semidefinite matrices $X_1, \dots, X_n \in \mathcal{S}_+^d$, for some $d \in \mathbb{N}$, such that $A = (\langle X_i, X_j \rangle)_{i,j=1}^n$. We then say that the set $\{X_i\}_{i\in[n]}$ forms a Gram representation of A. We denote by $\mathcal{CS}_+^n$ the set of all $n \times n$ completely positive semidefinite matrices.

4.1.2. LEMMA. *$\mathcal{CS}_+^n$ is a convex cone.*

PROOF: Given a matrix $A \in \mathcal{CS}_+^n$, let $X_1, \ldots, X_n \in \mathcal{S}_+^d$ be its Gram representation. Fix a $\lambda \geq 0$ and consider the set of positive semidefinite matrices $\sqrt{\lambda}X_1, \ldots, \sqrt{\lambda}X_n$. These form a Gram representation of λA and thus $\lambda A \in \mathcal{CS}_+^n$.

Now, let $B \in \mathcal{CS}_+^n$ and $Y_1, \ldots, Y_n \in \mathcal{S}_+^k$ be its Gram representation. We show that the matrix $A + B$ lies in $\mathcal{CS}_+^n$. To this end, consider the set of matrices $X_1 \oplus Y_1, \ldots, X_n \oplus Y_n \in \mathcal{S}_+^{d+k}$, where $X_i \oplus Y_i$ is the direct sum between X_i and Y_i. These matrices are positive semidefinite and form a Gram representation of $A + B$. We conclude that $\mathcal{CS}_+^n$ is a convex cone. □

The completely positive semidefinite cone was implicitly introduced by Frenkel and Weiner [FW14]. The following lemmas contain simple but useful results about completely positive semidefinite matrices.

4.1.3. LEMMA. *Any principal submatrix of a completely positive semidefinite matrix is itself completely positive semidefinite.*

PROOF: Let A be a completely positive semidefinite matrix and $X_1, \ldots, X_n$ be its Gram representation. Consider a principal submatrix $A[I]$ of A, obtained from A by keeping only the rows and columns which are indexed by $I \subseteq [n]$. The set $\{X_i\}_{i \in I}$ forms a Gram representation for $A[I]$ and thus $A[I] \in \mathcal{CS}_+^{|I|}$. □

4.1.4. LEMMA. *The matrix $A \oplus B$ is completely positive semidefinite if and only if A and B are both completely positive semidefinite.*

PROOF: One direction follows directly from Lemma 4.1.3. Indeed, if $A \oplus B$ is a completely positive semidefinite matrix then also its two principal submatrices A and B must be completely positive semidefinite.

Suppose now that $A \in \mathcal{CS}_+^n$, $B \in \mathcal{CS}_+^m$ and that $X_1, \ldots, X_n \in \mathcal{S}_+^d$ and $Y_1, \ldots, Y_m \in \mathcal{S}_+^k$ form a Gram representation for A and B, respectively. The set of positive semidefinite matrices $X_1 \oplus 0_k, \ldots, X_n \oplus 0_k, 0_d \oplus Y_1, \ldots, 0_d \oplus Y_m$ (where 0_n is the $n \times n$ zero matrix) forms a Gram representation for $A \oplus B$ and thus $A \oplus B \in \mathcal{CS}_+^{n+m}$. □

4.1.5. LEMMA. *Let $A \in \mathcal{S}^n$ and P be an $n \times n$ permutation matrix. Then $A \in \mathcal{CS}_+^n$ if and only if $P^T AP \in \mathcal{CS}_+^n$.*

PROOF: The claim follows by an appropriate reordering of the matrices forming the Gram representation of A. □

4.1.6. LEMMA. *The following chain of inclusions holds: $\mathcal{CP}^n \subseteq \mathcal{CS}_+^n \subseteq \mathcal{DNN}^n$.*

PROOF: Let $A \in \mathcal{CP}^n$ and the nonnegative vectors $x_1, \ldots, x_n \in \mathbb{R}^d_+$ be its Gram representation. For each $i \in [n]$, define $X_i = \mathrm{Diag}(x_i)$; i.e., X_i is a diagonal matrix having vector x_i as diagonal. Clearly, $X_1, \ldots, X_n \in \mathcal{S}^d_+$ form a Gram representation of A and thus the inclusion $\mathcal{CP}^n \subseteq \mathcal{CS}^n_+$ holds.

Furthermore, any completely positive semidefinite matrix is also positive semidefinite (any matrix can be thought of as a vector by stacking on top of each others its columns). As the inner product of positive semidefinite matrices is always nonnegative, any matrix in $\mathcal{CS}^n_+$ must have nonnegative entries. Therefore, we deduce that $\mathcal{CS}^n_+ \subseteq \mathcal{DNN}^n$ holds. □

Since the cone $\mathcal{CP}^n$ is full-dimensional, the same holds for the cone $\mathcal{CS}^n_+$. All the cones $\mathcal{CP}^n, \mathcal{CS}^n_+, \mathcal{DNN}^n, \mathcal{S}^n_+$ are pointed. Moreover, the sets $\mathcal{CP}^n$, $\mathcal{S}^n_+$ and $\mathcal{DNN}^n$ are closed, while we do not know whether the cone $\mathcal{CS}^n_+$ is closed. One of the difficulties in proving this, as well as other properties of the completely positive semidefinite cone, lies in the fact that we do not have an alternative description for it. In particular, we do not know its extreme rays. This topic will be further discussed in Section 4.6.

We have that the following relations hold:

$$\mathcal{CP}^n \subseteq \mathcal{CS}^n_+ \subseteq \mathrm{cl}(\mathcal{CS}^n_+) \subseteq \mathcal{DNN}^n$$

and, taking their dual, we get the corresponding inclusions:

$$\mathcal{DNN}^{n*} = \mathcal{S}^n_+ + (\mathcal{S}^n \cap \mathbb{R}^{n\times n}_+) \subseteq \mathcal{CS}^{n*}_+ \subseteq \mathcal{CP}^{n*} = \mathcal{COP}^n.$$

Consider a convex, pointed, full-dimensional cone $\mathcal{K}^n \subseteq \mathcal{S}^n$. Recall that its *dual cone* is the set of symmetric matrices having nonnegative inner product with any matrix in the original cone: $\mathcal{K}^{n*} = \{M \in S^n : \langle M, A\rangle \geq 0 \quad \forall\, A \in \mathcal{K}^n\}$. Moreover, a matrix A lies in the *interior* of $\mathcal{K}^n$ if and only if $\langle A, M\rangle > 0$ for all nonzero matrices $M \in \mathcal{K}^{n*}$. Equivalently, A lies on the *boundary* of $\mathcal{K}^n$ if and only if there exists a nonzero matrix $M \in \mathcal{K}^{n*}$ such that $\langle A, M\rangle = 0$. We give two sufficient conditions for a matrix to lie on the boundary of $\mathcal{CS}_+$.

4.1.7. LEMMA. *Let $A \in \mathcal{CS}^n_+$ be a matrix having a zero entry, then A lies on the boundary of $\mathcal{CS}^n_+$.*

PROOF: Let $A \in \mathcal{CS}^n_+$ and suppose $A_{ij} = 0$. Consider the elementary matrix E_{ij} (with entries equal to 1 at positions (i, j) and (j, i) and zero elsewhere). Clearly, $E_{ij} \in \mathcal{S}^n \cap \mathbb{R}^{n\times n}_+$ and thus $E_{ij} \in \mathcal{CS}^{n*}_+$. Moreover, $\langle A, E_{ij}\rangle = 0$ and we conclude that A must lie on the boundary of $\mathcal{CS}^n_+$. □

Let $\mathcal{A}^t$ be the affine space in $\mathcal{S}^{nt}$ defined by the equations

$$\sum_{i,j\in[t]} A_{ui,vj} = 1 \text{ for } u, v \in [n].$$

We show that any completely positive semidefinite matrix that lies on the affine space $\mathcal{A}^t$ is on the boundary of the completely positive semidefinite cone.

4.1.8. LEMMA. *Let $A \in \mathcal{S}^{nt}$ be a completely positive semidefinite matrix that lies on $\mathcal{A}^t$, then A is on the boundary of $\mathcal{CS}^{nt}_+$.*

PROOF: Take a matrix $A \in \mathcal{CS}^{nt}_+ \cap \mathcal{A}^t$. Pick two distinct elements $u, v \in [n]$ and define F to be $n \times n$ matrix with $F_{uu} = F_{vv} = 1$, $F_{uv} = F_{vu} = -1$ and zero elsewhere. Let J be the $t \times t$ all-one matrix, then $M = J \otimes F$ is a positive semidefinite matrix as $J, F \succeq 0$. Therefore $M \in \mathcal{CS}^{nt*}_+$ and, since $\langle A, M \rangle = 0$, this shows that A lies on the boundary of $\mathcal{CS}^{nt}_+$. □

We end this section by mentioning that linear optimization over affine sections of the completely positive semidefinite cone is an NP-hard problem (see Remark 5.2.4).

4.2 Links with completely positive and doubly nonnegative matrices

In this section we study the links among completely positive, completely positive semidefinite and doubly nonnegative matrices. In particular, we give criteria that reduce the question of determining whether a matrix lies in $\mathcal{CS}^n_+$ to the one of determining if it lies in $\mathcal{CP}^n$. These are then use to show that both strict inclusions: $\mathcal{CP}^n \subsetneq \mathcal{CS}^n_+$ and $\mathcal{CS}^n_+ \subsetneq \mathcal{DNN}^n$ hold for any $n \geq 5$.

We have already seen that the following inclusions hold:

$$\mathcal{CP}^n \subseteq \mathcal{CS}^n_+ \subseteq \mathsf{cl}(\mathcal{CS}^n_+) \subseteq \mathcal{DNN}^n \tag{4.1}$$

and, by taking their dual, we get:

$$\mathcal{S}^n_+ + (\mathcal{S}^n \cap \mathbb{R}^{n\times n}_+) \subseteq \mathcal{CS}^{n*}_+ \subseteq \mathcal{COP}^n. \tag{4.2}$$

For any $n \leq 4$, Maxfield and Minc [MM62] and Diananda [Dia62] showed, respectively, that $\mathcal{CP}^n = \mathcal{DNN}^n$ and that $\mathcal{S}^n_+ + (\mathcal{S}^n \cap \mathbb{R}^{n\times n}_+) = \mathcal{COP}^n$. Hence equality holds throughout in (4.1) and (4.2). Moreover, as we will see, for any $n \geq 5$ the inclusions $\mathcal{CP}^n \subseteq \mathcal{DNN}^n$ and $\mathcal{S}^n_+ + (\mathcal{S}^n \cap \mathbb{R}^{n\times n}_+) \subseteq \mathcal{COP}^n$ are known to be strict.

One may wonder about why there is a change in behavior between 4×4 and 5×5 matrices. This can be explained by the following result of Kogan and Berman [KB93] that uses graph theoretical arguments. Given a matrix $A \in \mathcal{S}^n$, its *support graph* is the graph $G(A) = ([n], E)$ where there is an edge $\{i, j\}$ whenever $A_{ij} \neq 0$ with $i \neq j$. A graph G is said to be *completely positive* if every doubly nonnegative matrix with support G is completely positive.

4.2.1. THEOREM (KOGAN–BERMAN [KB93]). *A graph G is completely positive if and only if it does not contain an odd cycle of length at least 5 as a subgraph.*

This in particular implies that any matrix $A \in \mathcal{S}^n$ with $n \leq 4$ has completely positive support graph $G(A)$ and, therefore, $A \in \mathcal{CP}^n$ if and only if $A \in \mathcal{DNN}^n$.

On the other hand, to find a doubly nonnegative matrix which is not completely positive (see Example 4.2.3 below), we can use the following result characterizing completely positive matrices whose support graph is triangle-free. For a matrix $A \in \mathcal{S}^n$, define its *comparison matrix* $C(A) \in \mathcal{S}^n$ to be the matrix with entries $C(A)_{ii} = A_{ii}$ for all $i \in [n]$ and $C(A)_{ij} = -A_{ij}$ for all $i \neq j \in [n]$.

4.2.2. THEOREM (DREW–JOHNSON–LOEWY [DJL94]). *Let $A \in \mathcal{S}^n$ and assume that its support graph is triangle-free. Then, A is completely positive if and only if its comparison matrix $C(A)$ is positive semidefinite.*

We notice that to prove strict inclusions in (4.1) and consequentially in (4.2), it suffices to show them for $n = 5$. Indeed, in view of Lemma 4.1.4, a matrix A belongs to $\mathcal{CS}^5_+$ if and only if the extended matrix $\tilde{A}$, obtained by adding a border of zero entries to A, belongs to $\mathcal{CS}^n_+$. One can easily observe that an equivalent statement holds true for both $\mathcal{CP}$ and $\mathcal{DNN}$. We thus focus on 5×5 matrices and, in particular, on circulant matrices with the following form:

$$M(b,c) = \begin{pmatrix} 1 & b & c & c & b \\ b & 1 & b & c & c \\ c & b & 1 & b & c \\ c & c & b & 1 & b \\ b & c & c & b & 1 \end{pmatrix} \quad \text{where } b, c \in \mathbb{R}.$$

4.2.3. EXAMPLE. The matrix $W = M((\sqrt{5}-1)/2, 0)$ is doubly nonnegative but not completely positive. Indeed, one can easily check that W is positive semidefinite while its comparison matrix $C(W) = M((1-\sqrt{5})/2, 0) \notin \mathcal{S}^5_+$. Applying Theorem 4.2.2 we then conclude that W is not completely positive and hence $\mathcal{CP}^n \subsetneq \mathcal{DNN}^n$ for any $n \geq 5$.

4.2.4. EXAMPLE. The matrix $H = M(-1, 1)$, known as the *Horn matrix*, is copositive but cannot be decomposed as the sum of a positive semidefinite matrix and a nonnegative matrix (see e.g. [Hal86, Section 16.2]).

4.2.5. EXAMPLE. The matrix $L = M(\cos^2(4\pi/5), \cos^2(2\pi/5))$ (equivalently, $L = M((3+\sqrt{5})/8, (3-\sqrt{5})/8)$) was given in [FGP+15] as an example of a completely positive semidefinite matrix which is not completely positive. To see this, consider the matrix $\widehat{L} = M(\cos(4\pi/5), \cos(2\pi/5))$, so that L is the entrywise square of $\widehat{L}$. Then, the vectors $x_i = (\cos(4i\pi/5), \sin(4i\pi/5)) \in \mathbb{R}^2$ (for $i \in [5]$) form a Gram representation of $\widehat{L}$ and thus the positive semidefinite

matrices $x_i x_i^T \in \mathcal{S}_+^2$ ($i \in [5]$) form a Gram representation of L. This shows that $L \in \mathcal{CS}_+^5$. On the other hand, $L \notin \mathcal{CP}^5$ as its inner product with the Horn matrix is negative: $\langle L, H\rangle = 5(1-\sqrt{5}/2) < 0$ and $H \in \mathcal{CP}^{5*}$. Therefore,

$$L \in \mathcal{CS}_+^5 \setminus \mathcal{CP}^5 \text{ and } H \in \mathcal{COP}^5 \setminus \mathcal{CS}_+^{5*}.$$

Example of a matrix in $\mathcal{DNN}^5 \setminus \mathcal{CS}_+^5$. To find a matrix which is doubly nonnegative but not completely positive semidefinite (or even better is not in the closure of the $\mathcal{CS}_+$ cone), we need a sufficient criterion for a matrix to be not completely positive semidefinite. In Theorem 4.2.9 below, we prove that, for matrices whose support graph is a cycle, being completely positive is equivalent to being completely positive semidefinite. From this we deduce that the matrix $W = M((\sqrt{5}-1)/2, 0)$ (from Example 4.2.3) is doubly nonnegative but not completely positive semidefinite. Frenkel and Weiner [FW14] were the first to prove that $W \in \mathcal{DNN}^5 \setminus \mathcal{CS}_+^5$. At the end of this section, we will present their original proof and show how we can use their ideas to prove that, in fact, W does not even belong to the closure of $\mathcal{CS}_+^5$ (see Theorem 4.2.14 and the preceding discussion).

We start with an observation about matrices supported by bipartite graphs.

4.2.6. LEMMA. *Consider a matrix $A \in \mathcal{S}^n$ and assume that $G(A)$ is a bipartite graph. Then, $A \in \mathcal{CS}_+^n$ if and only if $A \in \mathcal{CP}^n$.*

PROOF: We only have to prove that $A \in \mathcal{CS}_+^n$ implies $A \in \mathcal{CP}^n$, as the reverse implication holds trivially. Assume that $A \in \mathcal{CS}_+^n$ and, say, $X_1, \ldots, X_n \in \mathcal{S}_+^d$ form its Gram representation. As the support graph $G(A)$ is bipartite, consider a bipartition of its vertex set as $U \cup W$ so that all edges of $G(A)$ are of the form $\{i, j\}$ with $i \in U$ and $j \in W$. Observe that the matrices X_i for $i \in U$, and $-X_j$ for $j \in W$ form a Gram representation of the comparison matrix $C(A)$. Thus $C(A)$ is positive semidefinite and, by Theorem 4.2.2, $A \in \mathcal{CP}^n$. □

4.2.7. REMARK. We could have derived the statement of Lemma 4.2.6 directly from Theorem 4.2.1. Indeed, a graph G is bipartite if and only if it does not contain any odd-length cycles. Hence, in particular, the support graph $G(A)$ does not contain an odd cycle of length at least 5 as a subgraph and by Theorem 4.2.1 we get that: $A \in \mathcal{CP} \iff A \in \mathcal{CS}_+ \iff A \in \mathcal{DNN}$.

Next we state an useful elementary result about positive semidefinite matrices.

4.2.8. LEMMA. *Let A and B be positive semidefinite matrices with block-form:*

$$A = \begin{pmatrix} A_1 & A_2 \\ A_2^T & A_3 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} B_1 & B_2 \\ B_2^T & B_3 \end{pmatrix},$$

where A_i and B_i have the same dimension. If $\langle A, B\rangle = 0$ holds, then we have $\langle A_1, B_1\rangle = \langle A_3, B_3\rangle = -\langle A_2, B_2\rangle$.

PROOF: As both A, B are positive semidefinite matrices, $\langle A, B\rangle = 0$ implies $AB = 0$ and thus $A_1B_1 + A_2B_2^T = 0$ and $A_2^TB_2 + A_3B_3 = 0$. Taking the trace and noticing that A_i, B_i for $i \in \{1, 3\}$ are symmetric matrices, we obtain the desired identities. □

We can now characterize the completely positive semidefinite matrices supported by a cycle.

4.2.9. THEOREM. *Consider a matrix $A \in \mathcal{S}^n$ and assume that its support graph $G(A)$ is a cycle. Then, $A \in \mathcal{CS}_+^n$ if and only if $A \in \mathcal{CP}^n$.*

PROOF: As $\mathcal{CP}^n \subseteq \mathcal{CS}_+^n$ is always true, one direction is obvious.

Assume that $A \in \mathcal{CS}_+^n$ with $n \geq 5$ (otherwise there is nothing to prove) and that its support graph $G(A)$ is a cycle. In view of Lemma 4.1.4, we can without loss of generality restrict our attention to the case where the graph $G(A)$ is a n-cycle. Moreover, due to Lemma 4.1.5, we can assume that the nonzero entries of A are $A_{i,i+1}$ for $i \in [n]$ (taking indices modulo n). Let $X^1, \ldots, X^n \in \mathcal{S}_+^d$ be a positive semidefinite Gram representation of A. Due to Theorem 4.2.2, to prove that $A \in \mathcal{CP}^n$ it suffices to show that the comparison matrix $C(A)$ is positive semidefinite.

Consider first the easy case when n is even. Then, the matrices $Y^1 = -X^1$, $Y^2 = X^2, Y^3 = -X^3, Y^4 = X^4, \ldots, Y^{n-1} = -X^{n-1}, Y^n = X^n$ form a Gram representation of $C(A)$, thus showing that $C(A) \in \mathcal{S}_+^n$ and concluding the proof. Notice that, since even cycles are bipartite graphs, this case also follows from Lemma 4.2.6.

Now suppose that n is odd. As we will see, in order to construct a Gram representation of $C(A)$, we can choose the same matrices $Y^1, \ldots, Y^{n-1}$ as above but we need to look in more detail into the structure of the X^i's in order to be able to tell how to define the last matrix Y^n.

For this, we now show that the matrices $X^1, \ldots, X^n$ can be assumed to be $(n-2) \times (n-2)$ block-matrices, where we denote the blocks of X^k as X_{rs}^k for $r, s \in [n-2]$ (with $X_{sr}^k = (X_{rs}^k)^T$) and the index sets of the blocks as $I_1, \ldots, I_{n-2}$. Indeed, without loss of generality we can assume that $X^1 = \begin{pmatrix} X_{11}^1 & 0 \\ 0 & 0 \end{pmatrix}$ where X_{11}^1 is positive definite and the index set of X_{11}^1 defines the index set I_1 of the first block. Next, X^2 has the form $\begin{pmatrix} X_{11}^2 & X_{12}^2 & 0 \\ X_{21}^2 & X_{22}^2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, where $X_{22}^2 \succ 0$ and its index set defines the index set I_2 of the second block. Then, we can write $X^3 = \begin{pmatrix} X_{11}^3 & X_{12}^3 & X_{13}^3 & 0 \\ X_{21}^3 & X_{22}^3 & X_{23}^3 & 0 \\ X_{31}^3 & X_{32}^3 & X_{33}^3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, where $X_{33}^3 \succ 0$ and I_3 is the index set of X_{33}^3.

Hence X^3 has its blocks indexed by I_1, I_2, I_3 and $[d]\backslash(I_1 \cup I_2 \cup I_3)$. Iteratively, for each $k \in \{2, 3, \ldots, n-3\}$, the matrix X^k has blocks $X^k_{r,s}$ for $r, s \in [k]$ with $X^k_{k,k} \succ 0$ and it has zero entries outside of these blocks. The index sets of the blocks X^k_{kk} for $1 \leq k \leq n-3$ define the sets $I_1, I_2, \ldots I_{n-3}$ and the set $I_{n-2} = [d]\backslash(I_1 \cup I_2 \cdots \cup I_{n-3})$ collects all the remaining indices.

By looking at the zero-pattern of the matrix A, we show some structural properties of the X^k matrices and that each set I_k is nonempty. Since $A_{12} \neq 0$, we know that $I_1 \neq \emptyset$. As $A_{13} = 0$ we can conclude that $X^3_{11} = 0$ (and thus $X^3_{12} = X^3_{13} = 0$) and that the only nonzero blocks of X^3 are $X^3_{22}, X^3_{23}, X^3_{32}$ and X^3_{33}. Moreover, as $A_{23} \neq 0$ we get that $I_2 \neq \emptyset$. With the same reasoning, for each $k \in \{2, 3, \ldots, n-3\}$, as $A_{k',k} = 0$ for all $k' \leq k-2$ we have that all blocks of the matrix X^k are equal to zero except its blocks $X^k_{k-1,k-1}$, $X^k_{k-1,k}$, $X^k_{k,k-1}$ and X^k_{kk}. For all $k \in [n-2]$, the fact that $A_{k,k+1} \neq 0$ implies that the index set I_k is nonempty. Additionally, using the fact that $A_{n-2,k} = 0$ for each $k \in \{1, \ldots, n-4\}$ we obtain that each block X^{n-2}_{kk} is equal to zero. Similarly, X^{n-1}_{kk} is the zero matrix for every $k \in \{1, \ldots, n-3\}$ as $A_{n-1,k} = 0$. For the matrix X^n we cannot make any consideration on the presence of zero blocks.

Next we indicate how to construct the (non-symmetric) matrix Y^n from X^n: we just change signs to its two blocks $X^n_{n-3,n-2}$ and $X^n_{n-2,n-2}$. In other words, we let Y^n be the $(n-2) \times (n-2)$ block matrix where $Y^n_{n-3,n-2} = -X^n_{n-3,n-2}$, $Y^n_{n-2,n-2} = -X^n_{n-2,n-2}$ and $Y^n_{rs} = X^n_{rs}$ for all other blocks. Let us stress that we do not change the sign of the block $X^n_{n-2,n-3}$. As in the case when n is even, for any $1 \leq i \leq n-1$, we set $Y^i = -X^i$ for odd i and $Y^i = X^i$ for even i.

We claim that $Y^1, \ldots, Y^n$ form a Gram representation of the comparison matrix $C(A)$. It is clear that $\langle Y^i, Y^j \rangle = C(A)_{ij}$ for all $i, j \in [n-1]$ and that $\langle Y^1, Y^n \rangle = -A_{1n} = C(A)_{1n}$ and $\langle Y^i, Y^n \rangle = 0$ for $2 \leq i \leq n-3$ (since the blocks indexed by $[n-3]$ in Y^n are the same as in X^n and for any $r \in [n]$ each block $Y^i_{r,n-2}$ is equal to the zero matrix). Moreover, $\langle Y^n, Y^n \rangle = \langle X^n, X^n \rangle = C(A)_{nn}$ and $\langle Y^{n-1}, Y^n \rangle = -A_{n-1,n} = C(A)_{n-1,n}$. Finally, we use Lemma 4.2.8 to verify that $\langle Y^{n-2}, Y^n \rangle = 0$. Indeed, we have that

$$0 = A_{n-2,n} = \langle X^{n-2}, X^n \rangle = \langle \begin{pmatrix} X^{n-2}_{n-3,n-3} & X^{n-2}_{n-3,n-2} \\ X^{n-2}_{n-2,n-3} & X^{n-2}_{n-2,n-2} \end{pmatrix}, \begin{pmatrix} X^n_{n-3,n-3} & X^n_{n-3,n-2} \\ X^n_{n-2,n-3} & X^n_{n-2,n-2} \end{pmatrix} \rangle.$$

By Lemma 4.2.8, this then implies $\langle X^{n-2}_{n-3,n-3}, X^n_{n-3,n-3} \rangle = \langle X^{n-2}_{n-2,n-2}, X^n_{n-2,n-2} \rangle$. Therefore,

$$\langle Y^{n-2}, Y^n \rangle = \langle \begin{pmatrix} -X^{n-2}_{n-3,n-3} & -X^{n-2}_{n-3,n-2} \\ -X^{n-2}_{n-2,n-3} & -X^{n-2}_{n-2,n-2} \end{pmatrix}, \begin{pmatrix} X^n_{n-3,n-3} & -X^n_{n-3,n-2} \\ X^n_{n-2,n-3} & -X^n_{n-2,n-2} \end{pmatrix} \rangle$$

is equal to zero. □

4.2.10. EXAMPLE. The matrix $W = M((\sqrt{5}-1)/2, 0)$ is doubly nonnegative but not completely positive semidefinite. Recall from Example 4.2.3 that W is doubly nonnegative but not completely positive. Noticing that the support graph of W is the 5-cycle, by Theorem 4.2.9 we conclude that $W \notin \mathcal{CS}_+^5$ and, hence, that $\mathcal{CS}_+^n \subsetneq \mathcal{DNN}^n$ for any $n \geq 5$.

Using a result from Hamilton-Jester and Li [HJL96] we can construct a class of matrices in $\mathcal{DNN}^5 \setminus \mathcal{CS}_+^5$, see Lemma 4.2.12 below.

4.2.11. THEOREM (HAMILTON-JESTER–LI [HJL96]). *(i) Assume $n \geq 5$ is odd number and consider a matrix $A \in \mathcal{DNN}^n$ with rank $n-2$. Then, A lies on an extreme ray of $\mathcal{DNN}^n$ if and only if $G(A)$ is the n-cycle.*

(ii) If A lies on an extreme ray of $\mathcal{DNN}^5$, then A has rank 1 or 3.

4.2.12. LEMMA. *For odd $n \geq 5$, any matrix $A \in \mathcal{S}^n$ with rank $n-2$ and whose support graph $G(A)$ is the n-cycle is not completely positive semidefinite. Moreover, for $n = 5$, if $A \in \mathcal{S}^5$ lies on an extreme ray of $\mathcal{DNN}^5$, then A is completely positive semidefinite if and only if A is completely positive.*

PROOF: Let $A \in \mathcal{CS}_+^n$ be a matrix having the n-cycle as support graph and with rank $n-2$, for some odd number $n \geq 5$. By Theorem 4.2.9 we know that $A \in \mathcal{CP}^n$. Moreover, from Theorem 4.2.11(i), A lies on an extreme ray of $\mathcal{DNN}^n$ and thus also of $\mathcal{CP}^n$. Since the extreme rays of $\mathcal{CP}$ are rank 1 matrices, we get a contradiction: $1 = \operatorname{rank} A = n-2$. Hence, $A \notin \mathcal{CS}_+^n$.

For the second claim, assume that A lies on an extreme ray of $\mathcal{DNN}^5$. We show that if $A \notin \mathcal{CP}^5$ then $A \notin \mathcal{CS}_+^5$ (the reverse implication is clear). By Theorem 4.2.11(ii), any matrix on an extreme ray of $\mathcal{DNN}^5$ has rank 1 or 3. Recall that any doubly nonnegative matrix with rank 1 is completely positive (see Section 2.2.2). Hence, if $A \notin \mathcal{CP}^5$ then A has rank 3 and its support graph is the 5-cycle (by Theorem 4.2.11(i)). Using the first part of the lemma we can conclude that $A \notin \mathcal{CS}_+^5$. □

Furthermore, as a simple application of Lemma 4.2.6 and of Theorem 4.2.9, we get the following result. Recall that the parameters $\vartheta^{\mathcal{K}}$ and $\Theta^{\mathcal{K}}$ were introduced in Definition 3.2.7.

4.2.13. LEMMA. *If a graph G is bipartite or an odd cycle, then $\vartheta^{\mathcal{CS}_+}(\overline{G}) = \alpha(\overline{G})$ and $\Theta^{\mathcal{CS}_+}(\overline{G}) = \chi_f(\overline{G})$.*

PROOF: It suffices to show $\vartheta^{\mathcal{CS}_+}(\overline{G}) \leq \alpha(\overline{G})$ (as the reverse inequality is clear). For this pick a matrix $A \in \mathcal{CS}_+$ feasible for the program defining $\vartheta^{\mathcal{CS}_+}(\overline{G})$. Then the support of A is contained in G and thus is bipartite or an odd cycle. By Lemma 4.2.6 and Theorem 4.2.9, A is completely positive. Using Theorem 3.2.2, this implies $\alpha(\overline{G}) \geq \langle J, A \rangle$ and thus $\alpha(\overline{G}) \geq \vartheta^{\mathcal{CS}_+}(\overline{G})$. The identity $\Theta^{\mathcal{CS}_+}(\overline{G}) = \chi_f(\overline{G})$ follows analogously using Theorem 3.2.3. □

Example of a matrix in $\mathcal{DNN}^5 \setminus \mathsf{cl}(\mathcal{CS}^5_+)$. At last we show that the matrix $W = M((\sqrt{5}-1)/2, 0)$ does not belong to the closure of $\mathcal{CS}^5_+$. For this we use results that will be proven in Section 4.6 and findings from [FW14] regarding Gram representations by positive elements in a general *finite von Neumann algebra* (an infinite dimensional analog of Gram representations by positive semidefinite matrices). At this stage, we only need to know that a finite von Neumann algebra $\mathcal{N}$ is equipped with a trace τ (an analog of the usual matrix trace) which satisfies the following properties: for any $A, B \in \mathcal{N}$, $\tau(AB) = \tau(BA)$; if A is positive then $\tau(A) \geq 0$ with equality if and only if $A = 0$; and

$$\text{if } A, B \text{ are positive and } \tau(AB) = 0 \text{ then } AB = 0. \tag{4.3}$$

On the one hand, in Theorem 4.6.9 we show that there exists a finite von Neumann algebra $\mathcal{N}$ (with trace τ) with the property that any matrix A lying in the closure of $\mathcal{CS}^n_+$ admits a Gram representation by positive elements of $\mathcal{N}$; i.e., $A = (\tau(X_i X_j))_{i,j\in[n]}$ for some positive $X_1, \ldots, X_n \in \mathcal{N}$. On the other hand, it is shown in [FW14] that the matrix W does not admit a Gram representation by positive elements in any finite von Neumann algebra, see Theorem 4.2.14 below. Hence, by combining these two results, we deduce that the matrix W does not belong to the closure of $\mathcal{CS}^5_+$.

4.2.14. THEOREM (FRENKEL–WEINER [FW14]). *Let $\mathcal{N}$ be a finite von Neumann algebra with trace τ. For the matrix $W = M((\sqrt{5}-1)/2, 0)$, there do not exist positive elements $X_1, \ldots, X_5 \in \mathcal{N}$ such that $W = (\tau(X_i X_j))^5_{i,j=1}$.*

PROOF: We give a proof for completeness. Assume that $W = (\tau(X_i X_j))^5_{i,j=1}$ for some positive $X_1, \ldots, X_5 \in \mathcal{N}$. Using (4.3), $W_{i,i+2} = 0$ implies $X_i X_{i+2} = 0$ for all $i \in [5]$ (taking indices modulo 5). As W is a rank 3 positive semidefinite matrix, there exist vectors $u_1, \ldots, u_5 \in \mathbb{R}^3$ forming a Gram representation of W and the set $\{u_1, \ldots, u_5\}$ has rank 3. Moreover, one can check that the set $\{u_1, u_3, u_4\}$ is a base of $\mathbb{R}^3$. Hence, $u_2 = \alpha u_1 + \beta u_3 + \gamma u_4$ for some $\alpha, \beta, \gamma \in \mathbb{R}$. Since $W = (u_i^T u_j)^5_{i,j=1} = (\tau(X_i X_j))^5_{i,j=1}$, we obtain the analogous relation: $X_2 = \alpha X_1 + \beta X_3 + \gamma X_4$. Multiplying both sides by X_1 gives $X_1 X_2 = \alpha X_1^2$. Analogously, expressing u_1 in the base $\{u_2, u_4, u_5\}$ implies that $X_1 X_2 = \alpha' X_2^2$ for some $\alpha' \in \mathbb{R}$. Thus, $\alpha X_1^2 = \alpha' X_2^2$ which, as $W_{11} = W_{22} = 1$, implies that $\alpha = \alpha'$ and thus $X_1 = X_2$. Since $W_{12} \neq 1$, we reached a contradiction. □

4.2.15. COROLLARY. *For any $n \geq 5$, the following inclusions hold*

$$\mathcal{CP}^n \subsetneq \mathcal{CS}^n_+ \subseteq \mathsf{cl}(\mathcal{CS}^n_+) \subsetneq \mathcal{DNN}^n \text{ and } \mathcal{S}^n_+ + (\mathcal{S}^n \cap \mathbb{R}^{n\times n}_+) \subsetneq \mathcal{CS}^{n*}_+ \subsetneq \mathcal{COP}^n.$$

PROOF: Combine relationships (4.1), (4.2) together with Examples 4.2.5, 4.2.10 and the discussion preceding Theorem 4.2.14. □

4.3 The dual of the completely positive semidefinite cone

We now investigate the dual of the cone $\mathcal{CS}_+^n$.

4.3.1. LEMMA. *Given a matrix $M \in \mathcal{S}^n$, the following assertions are equivalent:*

(i) $M \in \mathcal{CS}_+^{n*}$, *i.e.,* $\sum_{i,j=1}^n M_{ij}\langle X_i, X_j\rangle \geq 0$ *for all* $X_1, \ldots, X_n \in \mathcal{S}_+^d$ *and* $d \in \mathbb{N}$.

(ii) $\mathsf{Tr}(\sum_{i,j=1}^n M_{ij} X_i^2 X_j^2) \geq 0$ *for all* $X_1, \ldots, X_n \in \mathcal{S}^d$ *and* $d \in \mathbb{N}$.

PROOF: Statement (i) is simply conic duality. By linearity of the trace we have: $\sum_{i,j=1}^n M_{ij}\langle X_i, X_j\rangle = \sum_{i,j=1}^n M_{ij}\,\mathsf{Tr}(X_iX_j) = \mathsf{Tr}(\sum_{i,j=1}^n M_{ij}X_iX_j)$. Statement (ii) follows since any matrix $X \in \mathcal{S}_+^d$ can be written as $X = Y^2$ for some $Y \in \mathcal{S}^d$. □

4.3.2. COROLLARY. *A matrix $M \in \mathcal{CS}_+^{n*}$ if and only if the associated polynomial $p_M = \sum_{i,j=1}^n M_{ij}X_i^2X_j^2$ in the noncommutative variables $X_1, \ldots, X_n$ is* trace positive*, meaning that the evaluation of p_M at any symmetric matrices $X_1, \ldots, X_n$ (of the same arbitrary size $d \geq 1$) produces a matrix with nonnegative trace.*

Recalling that a matrix M is copositive if and only if the n-variate polynomial $p_M = \sum_{i,j=1}^n M_{ij}x_i^2x_j^2$ is nonnegative over $\mathbb{R}^n$, in the above corollary we recover copositivity by restricting to symmetric matrices X_i of size $d = 1$; i.e., to real numbers.

Interestingly, understanding which matrices lie in $\mathcal{CS}_+^{n*}$ is deeply connected with Connes' embedding conjecture [Con76], one of the most important conjectures in von Neumann algebra (see Section 4.6.3). A reformulation of the conjecture that shows this connection is given by Klep and Schweighofer [KS08] (Conjecture 4.3.5 below). Connes' embedding conjecture plays an important role also in the description of the matrices lying in the closure of $\mathcal{CS}_+^n$. We investigate this link in Section 4.6.

We introduce some useful notation. Let $\mathbb{R}[\underline{x}]$ denote the set of real polynomials in the commutative variables $x_1, \ldots, x_n$. Similarly, we denote by $\mathbb{R}\langle\underline{X}\rangle$ the set of real polynomials in the noncommutative variables $X_1, \ldots, X_n$. The set $\mathbb{R}\langle\underline{X}\rangle$ is endowed with the involution $* : \mathbb{R}\langle\underline{X}\rangle \to \mathbb{R}\langle\underline{X}\rangle$ that sends each variable to itself, each monomial $X_{i_1}X_{i_2}\cdots X_{i_t}$ to its reverse $X_{i_t}\cdots X_{i_2}X_{i_1}$ and extends linearly to arbitrary polynomials; e.g., $(X_1^2X_2 + X_2X_3)^* = X_2X_1^2 + X_3X_2$. Let $f \in \mathbb{R}\langle\underline{X}\rangle$ be polynomial, then f^* evaluated at $(X_1, \ldots, X_n) \in (\mathcal{S}^d)^n$ is equal to $f(X_1, \ldots, X_n)^T$. Given a constant $\varepsilon \in \mathbb{R}$ and $f \in \mathbb{R}\langle\underline{X}\rangle$, the polynomial $f + \varepsilon$ evaluated at $(X_1, \ldots, X_n) \in (\mathcal{S}^d)^n$ is equal to $f(X_1, \ldots, X_n) + \varepsilon/dI_d$. We say that a polynomial $f \in \mathbb{R}\langle\underline{X}\rangle$ is *symmetric* if $f^* = f$ and $S\mathbb{R}\langle\underline{X}\rangle$ denotes the set of symmetric polynomials in $\mathbb{R}\langle\underline{X}\rangle$. A polynomial $f \in \mathbb{R}\langle\underline{X}\rangle$ is said to be

trace positive if $\mathsf{Tr}(f(X_1,\dots,X_n)) \geq 0$ for all $(X_1,\dots,X_n) \in \cup_{d\geq 1}(\mathcal{S}^d)^n$. A polynomial of the form ff^* is called a *Hermitian square*. Note that any Hermitian square when evaluated by symmetric matrices gives a positive semidefinite matrix and it is, therefore, trace positive. Moreover, any polynomial of the form $[f,g] = fg - gf$ is called a *commutator*. By the cyclic and linear properties of the trace, it is clear that when evaluated at any n-tuple of matrices the trace of any commutator vanishes.

Membership of a matrix M in $\mathcal{CS}_+^{n*}$ requires that the polynomial p_M is trace positive on *all* symmetric matrices. However, since p_M is a homogeneous polynomial, it suffices to check trace positivity over the (noncommutative version of the) hypercube Q_{nc} or over the (noncommutative) ball B_{nc}, where we set

$$Q_{\text{nc}} = \bigcup_{d\geq 1} \left\{ (X_1,\dots,X_n) \in (\mathcal{S}^d)^n : I - X_i^2 \succeq 0 \quad \forall i \in [n] \right\},$$

$$B_{\text{nc}} = \bigcup_{d\geq 1} \left\{ (X_1,\dots,X_n) \in (\mathcal{S}^d)^n : I - \sum_{i=1}^n X_i^2 \succeq 0 \right\}.$$

A similar observation was made by Burgdorf [Bur11], we include a proof for completeness.

4.3.3. LEMMA. *A matrix $M \in \mathcal{S}^n$ belongs to $\mathcal{CS}_+^{n*}$ if and only if the associated polynomial p_M is trace positive on the cube Q_{nc} or, equivalently, on the ball B_{nc}.*

PROOF: One direction is clear as being trace positive on all symmetric matrices implies being trace positive on the cube Q_{nc} or on the ball B_{nc}.

Suppose that p_M is trace positive on the ball B_{nc}. Consider any n-tuple of matrices $\underline{X} = (X_1,\dots,X_n) \in (\mathcal{S}^d)^n$, we show that $\mathsf{Tr}(f(\underline{X})) \geq 0$. Let λ be the largest eigenvalue of $\sum_{i=1}^n X_i^2 \in \mathcal{S}_+^d$. If $\lambda = 0$, then each X_i is the zero matrix and $f(\underline{X}) = 0$. Otherwise, if $\lambda > 0$ the matrix $\lambda I - \sum_{i=1}^n X_i^2$ is positive semidefinite and $\widetilde{\underline{X}} = (\widetilde{X}_1,\dots,\widetilde{X}_n)$, where $\widetilde{X}_i = X_i/\sqrt{\lambda}$ (for $i \in [n]$), is in B_{nc}. Hence, we have $0 \leq \mathsf{Tr}(f(\widetilde{\underline{X}})) = (1/\sqrt{\lambda})^n \,\mathsf{Tr}(f(\underline{X}))$ and thus $\mathsf{Tr}(f(\underline{X})) \geq 0$. We can conclude that $M \in \mathcal{CS}_+^{n*}$.

As $B_{\text{nc}} \subseteq Q_{\text{nc}}$, if p_M is trace positive on the hypercube Q_{nc} it is also so on B_{nc} and, by the above reasoning, $M \in \mathcal{CS}_+^{n*}$. □

In view of Lemma 4.3.3, to check whether $M \in \mathcal{CS}_+^{n*}$ we need to determine if p_M is trace positive on Q_{nc} (or equivalently on B_{nc}). We define two sets of polynomials $\text{tr}\mathcal{M}_{\text{nc}}^{\text{cube}}$ and $\text{tr}\mathcal{M}_{\text{nc}}^{\text{ball}}$ which by construction are trace positive on Q_{nc} and on B_{nc}, respectively.

4.3.4. DEFINITION. The *tracial quadratic module* $\text{tr}\mathcal{M}$ generated by a set of polynomials $p_1,\dots,p_m \in S\mathbb{R}\langle\underline{X}\rangle$ is defined as the set of all polynomials of the

form $h + \sum_{j=1}^{m_0} f_j f_j^* + \sum_{i=1}^{m} \sum_{j_i=1}^{m_i} g_{j_i} p_i g_{j_i}^*$, where $h \in \mathbb{R}\langle \underline{X} \rangle$ is a sum of commutators, $f_j, g_{j_i} \in \mathbb{R}\langle \underline{X} \rangle$ and $m_0, m_i \in \mathbb{N}$.

In particular, the tracial quadratic module $\mathrm{tr}\mathcal{M}_{\mathrm{nc}}^{\mathrm{cube}}$ is the set of all polynomials of the form $h + \sum_{j=1}^{m_0} f_j f_j^* + \sum_{i=1}^{m} \sum_{j_i=1}^{m_i} g_{j_i}(I - X_i^2) g_{j_i}^*$, where $h \in \mathbb{R}\langle \underline{X} \rangle$ is a sum of commutators, $f_j, g_{j_i} \in \mathbb{R}\langle \underline{X} \rangle$ and $m_0, m_i \in \mathbb{N}$.

The tracial quadratic module $\mathrm{tr}\mathcal{M}_{\mathrm{nc}}^{\mathrm{ball}}$ consists of all polynomials of the form $h + \sum_{j=1}^{m_0} f_j f_j^* + \sum_{j=1}^{m_1} g_j (I - \sum_{i=1}^{n} X_i^2) g_j^*$, where $h \in \mathbb{R}\langle \underline{X} \rangle$ is a sum of commutators, $f_j, g_j \in \mathbb{R}\langle \underline{X} \rangle$ and $m_0, m_1 \in \mathbb{N}$.

Klep and Schweighofer [KS08] (see also [BDKS14]) showed that Connes' embedding conjecture is equivalent to the following conjecture which characterizes the trace positive polynomials on Q_{nc}.

4.3.5. CONJECTURE (KLEP–SCHWEIGHOFER [KS08]). *Let $f \in S\mathbb{R}\langle \underline{X} \rangle$. The following assertions are equivalent:*

(i) f is trace positive on Q_{nc}, i.e., $\mathsf{Tr}(f(X_1, \ldots, X_n)) \geq 0$ for all n-tuples of matrices $(X_1, \ldots, X_n) \in Q_{\mathrm{nc}}$.

(ii) For any $\varepsilon > 0$, $f + \varepsilon \in \mathrm{tr}\mathcal{M}_{\mathrm{nc}}^{\mathrm{cube}}$, i.e., $f + \varepsilon = g + h$, where h is a sum of commutators and $g = \sum_{j=1}^{m_0} f_j f_j^* + \sum_{i=1}^{n} \sum_{j_i=1}^{m_i} g_{j_i}(1 - X_i^2) g_{j_i}^*$ for some polynomials $f_j, g_{j_i} \in \mathbb{R}\langle \underline{X} \rangle$ and $m_0, m_i \in \mathbb{N}$.

The implication $(ii) \implies (i)$ is true. Indeed, suppose $f + \varepsilon \in \mathrm{tr}\mathcal{M}_{\mathrm{nc}}^{\mathrm{cube}}$ for any $\varepsilon > 0$. Then, for all tuples $\underline{X} \in Q_{\mathrm{nc}}$ and any $\varepsilon > 0$, we have that $\mathsf{Tr}(f(\underline{X})) \geq -\varepsilon$. Thus, f is trace positive on Q_{nc}.

As a matter of fact, Connes' embedding conjecture is also equivalent to Conjecture 4.3.5 when we restrict the polynomial f to have degree at most 4 (see [Bur11, Proposition 2.14]). Note that the polynomials p_M, associated to matrix M, involve only monomials of the form $X_i^2 X_j^2$. Interestingly, in the proof that Conjecture 4.3.5 is equivalent to Connes' embedding conjecture, the monomials $X_i^2 X_j^2$ play a fundamental role (due to a result of Rădulescu [Răd99]). Finally, let us point out that, as observed by Burgdorf [Bur11, Remark 2.8], Connes' conjecture is also equivalent to Conjecture 4.3.5 where the ball is used instead of the hypercube, i.e., replacing the tracial quadratic module $\mathrm{tr}\mathcal{M}_{\mathrm{nc}}^{\mathrm{cube}}$ by the tracial quadratic module $\mathrm{tr}\mathcal{M}_{\mathrm{nc}}^{\mathrm{ball}}$.

4.4 Approximations to the dual of the completely positive semidefinite cone

A matrix $M \in \mathcal{CS}_+^{n*}$ if and only if the associated polynomial p_M belongs to the tracial quadratic module $\mathrm{tr}\mathcal{M}_{\mathrm{nc}}^{\mathrm{ball}}$ (Lemma 4.3.3). We define the set $\mathcal{K}_{\mathrm{nc},\varepsilon}$ consisting of all matrices M for which the perturbed polynomial $p_M + \varepsilon$ belongs

to $\mathrm{tr}\mathcal{M}_{\mathrm{nc}}^{\mathrm{ball}}$. To simplify the notation, in $\mathcal{K}_{\mathrm{nc},\varepsilon}$ we omit the dependence on the size n of the matrices.

4.4.1. DEFINITION. For $\varepsilon \geq 0$, let $\mathcal{K}_{\mathrm{nc},\varepsilon}$ denote the set of matrices $M \in \mathcal{S}^n$ for which the polynomial $p_M + \varepsilon$ belongs to the tracial quadratic module $\mathrm{tr}\mathcal{M}_{\mathrm{nc}}^{\mathrm{ball}}$.

4.4.2. LEMMA. *For any $\varepsilon \geq 0$, $\mathcal{K}_{\mathrm{nc},\varepsilon}$ is a convex set. Moreover, we have inclusion $\bigcap_{\varepsilon>0} \mathcal{K}_{\mathrm{nc},\varepsilon} \subseteq \mathcal{CS}_+^{n*}$, with equality if Connes' embedding conjecture holds.*

PROOF: Assume that $M, M' \in \mathcal{K}_{\mathrm{nc},\varepsilon}$ and that $\lambda \in [0,1]$, then the polynomial $p_{\lambda M+(1-\lambda)M'} + \varepsilon = \lambda(p_M + \varepsilon) + (1-\lambda)(p_{M'} + \varepsilon) \in \mathrm{tr}\mathcal{M}_{\mathrm{nc}}^{\mathrm{ball}}$ and, therefore, $\lambda M + (1-\lambda)M' \in \mathcal{K}_{\mathrm{nc},\varepsilon}$. Hence, the set $\mathcal{K}_{\mathrm{nc},\varepsilon}$ is convex.

Consider a matrix $M \in \bigcap_{\varepsilon>0} \mathcal{K}_{\mathrm{nc},\varepsilon}$. Then, for any $\varepsilon > 0$, the polynomial $p_M + \varepsilon$ is trace positive on B_{nc}. By letting ε tend to 0, we obtain that p_M is trace positive on B_{nc}. Thus, by Lemma 4.3.3 we get $M \in \mathcal{CS}_+^{n*}$. Finally, equality $\bigcap_{\varepsilon>0} \mathcal{K}_{\mathrm{nc},\varepsilon} = \mathcal{CS}_+^{n*}$ holds under Connes' embedding conjecture since, as mentioned above, by results of [KS08, BDKS14] Connes' embedding conjecture is equivalent to Conjecture 4.3.5, also when the ball is used instead of the hypercube. □

We point out a connection between the set $\mathcal{K}_{\mathrm{nc},\varepsilon}$ and the following set $\mathcal{K}_{\mathrm{c}}$, used in the commutative setting. Let Σ denote the set of sums of squares of (commutative) polynomials. Following [Par00] define the cone

$$\begin{aligned}\mathcal{K}_{\mathrm{c}} &= \left\{ M \in \mathcal{S}^n : p_M(\sum_{i=1}^n x_i^2)^r \in \Sigma \text{ for some } r \in \mathbb{N} \right\} \\ &= \left\{ M \in \mathcal{S}^n : p_M \in \Sigma + (1 - \sum_{i=1}^n x_i^2)\mathbb{R}[\underline{x}] \right\}\end{aligned}$$

(see [dKLP05, Proposition 2] for the equivalence between both definitions). Clearly, we have the inclusion $\mathcal{K}_{\mathrm{c}} \subseteq \mathcal{COP}$. Moreover, Parrilo [Par00, Section 5.3] showed that $\mathcal{K}_{\mathrm{c}}$ covers the interior of $\mathcal{COP}$. By adding degree constraints on the terms entering the decomposition of p_M, he defined a hierarchy of subcones of $\mathcal{COP}$, whose first level is equal to the dual of the doubly nonnegative cone: $\mathcal{K}_{\mathrm{c}}^{(0)} = \{M \in \mathcal{S}^n : p_M \in \Sigma\} = \mathcal{S}_+^n + (\mathcal{S}^n \cap \mathbb{R}_+^{n\times n}) = \mathcal{DNN}^{n*}$. It turns out that, for $\varepsilon = 0$, the set $\mathcal{K}_{\mathrm{nc},0}$ is equal to $\mathcal{K}_{\mathrm{c}}^{(0)}$.

4.4.3. LEMMA. *We have: $\mathcal{DNN}^{n*} = \mathcal{K}_{\mathrm{c}}^{(0)} = \mathcal{K}_{\mathrm{nc},0} \subseteq \mathcal{K}_{\mathrm{nc},\varepsilon}$ for any $\varepsilon > 0$.*

PROOF: The inclusion $\mathcal{K}_{\mathrm{nc},0} \subseteq \mathcal{K}_{\mathrm{nc},\varepsilon}$ is obvious.

Firstly, we show the inclusion $\mathcal{K}_{\mathrm{nc},0} \subseteq \mathcal{K}_{\mathrm{c}}^{(0)}$. Assume that $M \in \mathcal{K}_{\mathrm{nc},0}$; i.e., the associated polynomial $p_M = h + g$, where h is a sum of commutators and

$g = \sum_{j=1}^{m_0} f_j f_j^* + \sum_{j=1}^{m_1} g_j(1 - \sum_{i=1}^n X_i^2)g_j^*$ with $f_j, g_j \in \mathbb{R}\langle \underline{X}\rangle$. If we evaluate p_M at commutative variables $\underline{x} = (x_1, \ldots, x_n)$, we see that $h(\underline{x})$ vanishes and thus we obtain $p_M(\underline{x}) = g(\underline{x}) \in \Sigma + (1 - \sum_{i=1}^n x_i^2)\Sigma$. As p_M is a homogeneous polynomial, we can derive that $p_M \in \Sigma$ and thus $M \in \mathcal{K}_{\mathrm{c}}^{(0)}$. This follows from [dKLP05, Proposition 4] which shows that, for a homogeneous polynomial of even degree, membership in $\Sigma + (1 - \sum_{i=1}^n x_i^2)\Sigma$ implies membership in Σ.

Secondly, we prove that $\mathcal{K}_{\mathrm{c}}^{(0)} \subseteq \mathcal{K}_{\mathrm{nc},0}$. As $\mathcal{K}_{\mathrm{c}}^{(0)} = \mathcal{S}_+^n + (\mathcal{S}^n \cap \mathbb{R}_+^{n\times n})$, it suffices to show that if $M \succeq 0$ or if $M \geq 0$ then p_M is a sum of commutators and of Hermitian squares, which implies $M \in \mathcal{K}_{\mathrm{nc},0}$. Assume that $M \succeq 0$ and let $u_1, \ldots, u_n \in \mathbb{R}^d$ be a set of vectors forming a Gram representation of M. Then, $p_M(\underline{X}) = \sum_{i,j=1}^n \sum_{h=1}^d u_i(h)u_j(h)X_i^2X_j^2 = \sum_{h=1}^d (\sum_{i=1}^n u_i(h)X_i^2)^2$ is a sum of Hermitian squares. Assume now that M is entrywise nonnegative. Then each term $M_{ij}X_i^2X_j^2 = M_{ij}([X_i^2X_j, X_j] + X_jX_i^2X_j)$ is sum of a commutator and a Hermitian square and, therefore, p_M is sum of commutators and Hermitian squares. □

We conclude with some remarks concerning how well $\mathcal{K}_{\mathrm{c}}$ and $\mathcal{K}_{\mathrm{nc},\varepsilon}$ approximate the cones $\mathcal{COP}$ and $\mathcal{CS}_+^*$, respectively. As mentioned above, Parrilo [Par00] showed that $\mathsf{int}(\mathcal{COP}) \subseteq \mathcal{K}_{\mathrm{c}} \subseteq \mathcal{COP}$. This can also be derived using the following result of Schmüdgen [Sch91].

4.4.4. THEOREM (SCHMÜDGEN [SCH91]). *If $f \in \mathbb{R}[\underline{x}]$ is positive on the sphere, i.e., $f(\underline{x}) > 0$ for all $\underline{x} \in \mathbb{R}^n$ with $\sum_{i=1}^n x_i^2 = 1$, then $f \in \Sigma + (1 - \sum_{i=1}^n x_i^2)\mathbb{R}[\underline{x}]$.*

In the noncommutative case, membership of a matrix M in $\mathcal{K}_{\mathrm{nc},\varepsilon}$ means that the polynomial $p_M + \varepsilon$ belongs to the tracial quadratic module $\mathrm{tr}\mathcal{M}_{\mathrm{nc}}^{\mathrm{ball}}$, but there is no clear link between this and membership in the interior of the cone $\mathcal{CS}_+^*$. To explain this difference of behavior between $\mathcal{K}_{\mathrm{c}}$ and $\mathcal{K}_{\mathrm{nc},\varepsilon}$ let us point out that, in the commutative (scalar) case, working with the ball is, in some sense, equivalent to working with the sphere (recall Lemma 2.2.2). However, when working with matrices $X_1, \ldots, X_n$, one can rescale them to ensure that $I - \sum_{i=1}^n X_i^2 \succeq 0$ but one cannot ensure equality: $\sum_{i=1}^n X_i^2 = I$. Hence, in the noncommutative case one cannot equivalently switch between the ball and the sphere.

At last, we observe that, for a fix $\varepsilon \geq 0$, checking whether M in $\mathcal{K}_{\mathrm{nc},\varepsilon}$ can be done using a sequence of semidefinite programs. The matrix $M \in \mathcal{K}_{\mathrm{nc},\varepsilon}$ if the polynomial $p_M + \varepsilon$ admits a decomposition of the form $p_M + \varepsilon = g + h$, where $g = \sum_{j=1}^{m_0} f_j f_j^* + \sum_{j=1}^{m_1} g_j(1 - \sum_{i=1}^n X_i^2)g_j^*$ with $f_j, g_j \in \mathbb{R}\langle \underline{X}\rangle$ and $m_0, m_1 \in \mathbb{N}$, and h is a sum of commutators. Fixing an integer k, we can determine via a semidefinite program whether $p_M + \varepsilon$ has a decomposition $p_M + \varepsilon$ where the terms $f_j f_j^*$ and $g_j(1 - \sum_{i=1}^n X_i^2)g_j^*$ have degree at most $2k$ (see e.g. [Bur11] for details).

4.5 Polyhedral approximations of the completely positive semidefinite cone and of its dual cone

In this section we construct hierarchies of polyhedral cones converging asymptotically to the completely positive cone and its dual. The construction of our polyhedral hierarchy for $\mathcal{CS}^n_+$ is directly inspired from the classical case where analogous hierarchies of polyhedral cones exist for approximating the completely positive cone $\mathcal{CP}^n$ and the copositive cone $\mathcal{COP}^n$. In Section 4.5.1 we recall this construction and we introduce the new hierarchy in Section 4.5.2.

4.5.1 Polyhedral approximations of the completely positive cone and of its dual cone

A matrix $M \in \mathcal{S}^n$ is copositive if and only if the polynomial p_M is nonnegative over the standard simplex $\Delta_n = \{x \in \mathbb{R}^n_+ : \sum_{i=1}^n x_i = 1\}$ (Lemma 2.2.2). The idea for constructing outer approximations of the copositive cone is simple and relies on requiring nonnegativity of the polynomial p_M over all rational points in the standard simplex with given denominator r and letting r grow. This is made explicit in [Yil12] and goes back to earlier work on how to design tractable approximations for quadratic optimization problems over the standard simplex [BdK02, dKP02] and more general polynomial optimization problems [dKLP06]. More precisely, for an integer $r \geq 1$, define the sets

$$\Delta(n,r) = \{x \in \Delta_n : rx \in \mathbb{Z}^n\}, \quad \widetilde{\Delta}(n,r) = \bigcup_{s=1}^{r} \Delta(n,s)$$

where we restrict to rational points in Δ_n with given denominators. The sets $\widetilde{\Delta}(n,r)$ are nested within the standard simplex: $\widetilde{\Delta}(n,r) \subseteq \widetilde{\Delta}(n,r+1) \subseteq \Delta_n$. Following Yildirim [Yil12], we define the cone:

$$\mathcal{O}^n_r = \{M \in \mathcal{S}^n : x^\mathsf{T} M x \geq 0 \quad \forall x \in \widetilde{\Delta}(n,r)\}$$

and its dual cone $\mathcal{O}^{n*}_r$, which is the conic hull of all matrices of the form vv^T for some $v \in \widetilde{\Delta}(n,r)$. By construction, the cones $\mathcal{O}^n_r$ form a hierarchy of outer approximations for $\mathcal{COP}^n$ and their dual cones form a hierarchy of inner approximations for $\mathcal{CP}^n$:

$$\mathcal{COP}^n \subseteq \mathcal{O}^n_{r+1} \subseteq \mathcal{O}^n_r \text{ and } \mathcal{O}^{n*}_r \subseteq \mathcal{O}^{n*}_{r+1} \subseteq \mathcal{CP}^n.$$

4.5.1. THEOREM (YILDIRIM [YIL12]). *We have:* $\mathcal{COP}^n = \bigcap_{r\geq 1} \mathcal{O}^n_r$. *Moreover, we have the inclusions* $\mathsf{int}(\mathcal{CP}^n) \subseteq \bigcup_{r\geq 1} \mathcal{O}^{n*}_r \subseteq \mathcal{CP}^n$ *and* $\mathsf{cl}(\bigcup_{r\geq 1} \mathcal{O}^{n*}_r) = \mathcal{CP}^n$.

4.5.2 The new polyhedral approximations

We introduce the cones $\mathcal{C}_r^n$, which will form a hierarchy of inner approximations for the cone $\mathcal{CS}_+^n$, and the cones $\mathcal{D}_r^n$, which will form a hierarchy of outer approximations for the dual cone $\mathcal{CS}_+^{n*}$. These cones are in fact dual to each other, so it suffices to define the cones $\mathcal{D}_r^n$. The idea is simple and analogous to the one used in the classical (scalar) case: instead of requiring trace nonnegativity of the polynomial p_M over the full set $\bigcup_{d\geq 1}(\mathcal{S}_+^d)^n$, we only ask trace nonnegativity over specific finite subsets. We start with defining the set

$$\boldsymbol{\Delta}_n = \{\underline{X} = (X_1, \ldots, X_n) \in \bigcup_{d\geq 1} (\mathcal{S}_+^d)^n : \sum_{i=1}^n \mathsf{Tr}(X_i) = 1\}, \tag{4.4}$$

which can be seen as a dimension-free matrix analog of the standard simplex Δ_n in $\mathbb{R}^n$. Next we observe that a matrix M belongs to $\mathcal{CS}_+^{n*}$ if and only if its associated polynomial p_M is trace nonnegative over all n-tuples of *rational* matrices in $\boldsymbol{\Delta}_n$.

4.5.2. LEMMA. *For $M \in \mathcal{S}^n$ the following assertions are equivalent:*

(i) $M \in \mathcal{CS}_+^{n*}$, *i.e.,* $\mathsf{Tr}(p_M(\underline{X})) \geq 0$ *for all* $\underline{X} \in \bigcup_{d\geq 1}(\mathcal{S}_+^d)^n$.

(ii) $\mathsf{Tr}(p_M(\underline{X})) \geq 0$ *for all* $\underline{X} \in \boldsymbol{\Delta}_n$.

(iii) $\mathsf{Tr}(p_M(\underline{X})) \geq 0$ *for all* $\underline{X} = (X_1, \ldots, X_n) \in \boldsymbol{\Delta}_n$ *with* $X_1 \succ 0, \ldots, X_n \succ 0$.

(iv) $\mathsf{Tr}(p_M(\underline{X})) \geq 0$ *for all* $\underline{X} = (X_1, \ldots, X_n) \in \boldsymbol{\Delta}_n$ *with* $X_1 \succ 0, \ldots, X_n \succ 0$ *and with rational entries.*

(v) $\mathsf{Tr}(p_M(\underline{X})) \geq 0$ *for all* $\underline{X} \in \boldsymbol{\Delta}_n$ *with rational entries.*

PROOF: The implications $(i) \implies (ii) \implies (iii) \implies (iv)$, $(i) \implies (v)$ and $(v) \implies (iv)$ are clear. We show that $(iv) \implies (iii) \implies (ii) \implies (i)$.

Implication $(ii) \implies (i)$ follows by scaling. Indeed, take $\underline{X} \in (\mathcal{S}_+^d)^n$ with $\lambda = \sum_{i=1}^n \mathsf{Tr}(X_i) > 0$ (else $\underline{X}$ is identically zero and $\mathsf{Tr}(p_M(\underline{X})) = 0$). Then we have $\underline{X}/\lambda \in \boldsymbol{\Delta}_n$ and thus $\mathsf{Tr}(p_M(\underline{X}/\lambda)) \geq 0$, which implies $\mathsf{Tr}(p_M(\underline{X})) \geq 0$.

The remaining implications follow using continuity arguments. Namely, for $(iv) \implies (iii)$, use the fact that the set of rational positive definite matrices is dense within the set of positive definite matrices. For $(iii) \implies (ii)$, use that the set of positive definite matrices is dense within the set of positive semidefinite matrices (Theorem 2.2.1 (*ii*)). □

This motivates introducing the subset $\boldsymbol{\Delta}(n, r)$ of the set $\boldsymbol{\Delta}_n$ obtained by considering only n-tuples of rational positive semidefinite matrices with denominator at most r. This set can be seen as a matrix analog of the rational grid point subsets of the standard simplex Δ_n and it permits to define the new cones $\mathcal{D}_r^n$.

4.5.3. DEFINITION. Given an integer $r \in \mathbb{N}$, define the set

$$\boldsymbol{\Delta}(n,r) = \{\underline{X} \in \boldsymbol{\Delta}_n : \text{ each } X_i \text{ has rational entries with denominator} \leq r\}$$

and define the cone

$$\mathcal{D}_r^n = \{M \in \mathcal{S}^n : \mathsf{Tr}(p_M(\underline{X})) \geq 0 \quad \forall \underline{X} \in \boldsymbol{\Delta}(n,r)\}.$$

Next we show that $\mathcal{D}_r^n$ is a polyhedral cone. Notice that the set $\boldsymbol{\Delta}(n,r)$ is not finite as there is no bound on the dimension of the matrices in $\boldsymbol{\Delta}_n$. However, in the next lemma we observe that, without loss of generality, we can replace in the definition of $\mathcal{D}_r^n$ the set $\boldsymbol{\Delta}(n,r)$ by its subset $\underline{\boldsymbol{\Delta}}(n,r)$, obtained by restricting to $r \times r$ matrices $X_1, \ldots, X_n$.

4.5.4. LEMMA. *Define the set*

$$\underline{\boldsymbol{\Delta}}(n,r) = \{\underline{X} \in (\mathcal{S}_+^r)^n \cap \boldsymbol{\Delta}_n : \textit{each } X_i \textit{ has rational entries with denominator} \leq r\}.$$

Then, the following identity holds:

$$\mathcal{D}_r^n = \{M \in \mathcal{S}^n : \mathsf{Tr}(p_M(\underline{X})) \geq 0 \quad \forall \underline{X} \in \underline{\boldsymbol{\Delta}}(n,r)\}.$$

PROOF: The inclusion "$\supseteq$" is clear since $\underline{\boldsymbol{\Delta}}(n,r) \subseteq \boldsymbol{\Delta}(n,r)$.

To prove the reverse inclusion, take a matrix M such that $\mathsf{Tr}(p_M(\underline{X})) \geq 0$ for all $\underline{X} \in \underline{\boldsymbol{\Delta}}(n,r)$. Consider a n-tuple $\underline{X} = (X_1, \ldots, X_n) \in \boldsymbol{\Delta}(n,r)$, we show that $\mathsf{Tr}(p_M(\underline{X})) \geq 0$. By assumption, the matrices $X_1, \ldots, X_n$ are rational with denominator at most r, $\sum_{i=1}^n \mathsf{Tr}(X_i) = 1$ and $X_1, \ldots, X_n \in \mathcal{S}_+^d$ with $d > r$ (else there is nothing to prove). For each $i \in [n]$, set $I_i = \{k \in [d] : X_i(k,k) \neq 0\}$ and notice that $\mathsf{Tr}(X_i) \geq |I_i|/r$ (since each diagonal entry $X_i(k,k)$ indexed by $k \in I_i$ is at least $1/r$). Hence we have $1 = \sum_{i=1}^n \mathsf{Tr}(X_i) \geq \sum_{i=1}^n |I_i|/r$, implying $\sum_{i=1}^n |I_i| \leq r$. Then we can find a set I that contains $\bigcup_{i\in[n]} I_i$ and with cardinality $|I| = r$. As each matrix X_i has zero entries outside of its principal submatrix $X_i[I]$ indexed by I, then $\mathsf{Tr}(p_M(X_1, \ldots, X_n)) = \mathsf{Tr}(p_M(X_1[I], \ldots, X_n[I])) \geq 0$, where the last inequality follows from the fact that $(X_1[I], \ldots, X_n[I])$ belongs to the set $\underline{\boldsymbol{\Delta}}(n,r)$. □

The cardinality of the set $\underline{\boldsymbol{\Delta}}(n,r)$ is clearly finite. Moreover, in the following lemma we provide a simple upper bound on the cardinality of $\underline{\boldsymbol{\Delta}}(n,r)$.

4.5.5. LEMMA. *For any fixed r, the cardinality of the set $\underline{\boldsymbol{\Delta}}(n,r)$ is polynomial in terms of n. More precisely, let γ_r denote the number of $r \times r$ positive semidefinite matrices whose entries are rational with denominator at most r and whose trace is at most one. Then, $|\underline{\boldsymbol{\Delta}}(n,r)| \leq (\gamma_r)^r$ if $n \leq r$, and $|\underline{\boldsymbol{\Delta}}(n,r)| \leq \binom{n}{r}(\gamma_r)^r$ if $n > r$.*

PROOF: Consider an n-tuple of matrices whose sum of the traces is equal to one and whose entries are rational with denominator at most r. Clearly only at most r of them can be a nonzero matrix. Using this simple observation the statement of the lemma becomes straightforward. □

Notice that the simple identity $\mathsf{Tr}(p_M(\underline{X})) = \sum_{i,j} M_{ij}\langle X_i, X_j\rangle$ holds for any $\underline{X} = (X_1, \ldots, X_n)$. Therefore, the cone $\mathcal{D}_r^n$ can be equivalently defined as the set of matrices $M \in \mathcal{S}^n$ satisfying the finitely many linear inequalities: $\sum_{i,j=1}^n M_{ij}\langle X_i, X_j\rangle \geq 0$ for all $(X_1, \ldots, X_n) \in \underline{\boldsymbol{\Delta}}(n, r)$. This implies the following corollary.

4.5.6. COROLLARY. *The cone $\mathcal{D}_r^n$ is polyhedral.*

As $\underline{\boldsymbol{\Delta}}(n, r) \subseteq \underline{\boldsymbol{\Delta}}(n, r+1) \subseteq \boldsymbol{\Delta}_n$, the sets $\mathcal{D}_r^n$ form a hierarchy of outer approximations for $\mathcal{CS}_+^{n*}$:

$$\mathcal{CS}_+^{n*} \subseteq \mathcal{D}_{r+1}^n \subseteq \mathcal{D}_r^n \subseteq \cdots \subseteq \mathcal{D}_1^n.$$

Hence, $\mathcal{CS}_+^{n*} \subseteq \bigcap_{r\geq 1} \mathcal{D}_r^n$. In fact, as a direct application of the equivalence between (i) and (v) in Lemma 4.5.2, equality holds. The proof of the next theorem is thus omitted.

4.5.7. THEOREM. *The identity $\mathcal{CS}_+^{n*} = \bigcap_{r\geq 1} \mathcal{D}_r^n$ holds.*

The following property of the cones $\mathcal{D}_r^n$ will be useful.

4.5.8. LEMMA. *Consider a sequence of matrices $(M_r)_{r\geq 1}$ in $\mathcal{S}^n$ converging to a matrix $M \in \mathcal{S}^n$. If $M_r \in \mathcal{D}_r^n$ for all $r \in \mathbb{N}$, then $M \in \mathcal{CS}_+^{n*}$.*

PROOF: By Lemma 4.5.2, it suffices to show that $\mathsf{Tr}(p_M(\underline{X})) \geq 0$ whenever $\underline{X} \in \boldsymbol{\Delta}_n$ is rational valued. Fix a rational valued $\underline{X} \in \boldsymbol{\Delta}_n$ with, say, $\underline{X} \in (\mathcal{S}_+^d)^n$ and all entries have denominator at most t. Then, for all $r \geq r_0 = \max\{d, t\}$, we have $\underline{X} \in \underline{\boldsymbol{\Delta}}(n, r)$. Hence $\mathsf{Tr}(p_{M_r}(\underline{X})) \geq 0$ for all M_r with $r \geq r_0$. When r tends to infinity, $\mathsf{Tr}(p_{M_r}(\underline{X}))$ tends to $\mathsf{Tr}(p_M(\underline{X}))$ and thus $\mathsf{Tr}(p_M(\underline{X})) \geq 0$. □

We turn to the description of the dual cone $\mathcal{C}_r^n = \mathcal{D}_r^{n*}$. As a direct application of Lemma 4.5.4, we derive that $\mathcal{C}_r^n$ is the set of conic combinations of matrices which have a Gram representation by matrices in $\underline{\boldsymbol{\Delta}}(n, r)$; i.e.,

$$\mathcal{C}_r^n = \mathrm{cone}\{A \in \mathcal{S}^n : A = (\langle X_i, X_j\rangle)_{i,j=1}^n \text{ for some } \underline{X} \in \underline{\boldsymbol{\Delta}}(n, r)\}.$$

By construction, the cones $\mathcal{C}_r^n$ are polyhedral and they form a hierarchy of inner approximations of $\mathcal{CS}_+^n$: $\mathcal{C}_1^n \subseteq \cdots \subseteq \mathcal{C}_r^n \subseteq \mathcal{C}_{r+1}^n \subseteq \mathcal{CS}_+^n$. Moreover, strict inclusion holds.

4.5.9. LEMMA. *We have: $\mathcal{C}_r^n \subsetneq \mathcal{C}_{r+1}^n \subsetneq \mathcal{CS}_+^n$ for any $n \geq 2$ and $r \geq 1$.*

PROOF: We only need to prove that each inclusion is strict. It suffices to show this for $n = 2$ since one can extend a matrix A in $\mathcal{C}_r^2$ to a matrix in $\mathcal{C}_r^n$ by adding a border of zeros, and similarly for $\mathcal{CS}_+$. For this, we consider a rank 1 matrix $A = vv^T$, where $v = (1\ a)^T$ and a is a nonnegative scalar. Then $A \in \mathcal{CS}_+^2$. If we choose a to be an irrational number, A cannot belong to any cone $\mathcal{C}_r^2$, and if we choose $a = 1/(r+1)$, A belongs to $\mathcal{C}_{r+1}^2$ but not to $\mathcal{C}_r^2$. □

We show that the union of the cones $\mathcal{C}_r^n$ covers the interior of $\mathcal{CS}_+^n$.

4.5.10. THEOREM. *We have the inclusions:* $\text{int}(\mathcal{CS}_+^n) \subseteq \bigcup_{r\geq 1} \mathcal{C}_r^n \subseteq \mathcal{CS}_+^n$.

PROOF: We only have to show: $\text{int}(\mathcal{CS}_+^n) \subseteq \bigcup_{r\geq 1} \mathcal{C}_r^n$. For a contradiction, let A be a matrix in the interior of the cone $\mathcal{CS}_+^n$ and assume that A does not belong to $\bigcup_{r\geq 1} \mathcal{C}_r^n$. Then, for each $r \geq 1$, there exists a hyperplane strictly separating A from the (closed convex) cone $\mathcal{C}_r^n$. That is, there exists a matrix $M_r \in \mathcal{D}_r^n$ such that $\langle M_r, A\rangle < 0$ and $\|M_r\|_{\text{F}} = 1$. Since all matrices M_r lie in a compact set, the sequence $(M_r)_r$ admits a converging subsequence $(M_{r_i})_{i\geq 1}$ which converges to a matrix $M \in \mathcal{S}^n$. By Lemma 4.5.8 we know that the matrix M belongs to the cone $\mathcal{CS}_+^{n*}$ and thus $\langle A, M\rangle \geq 0$. On the other hand, as $\langle A, M_{r_i}\rangle < 0$ for all the indexes i, by taking the limit as i tends to infinity, we get that $\langle A, M\rangle \leq 0$. Hence we obtain $\langle A, M\rangle = 0$, which contradicts the assumption that A lies in the interior of $\mathcal{CS}_+^n$. □

It is easy to give an explicit description of the cones $\mathcal{C}_r^n$ for small r. For example, $\mathcal{C}_1^n$ is the set of $n \times n$ diagonal nonnegative matrices and $\mathcal{C}_2^n$ is the convex hull of the matrices E_{ii} and $E_{ii} + E_{ij} + E_{jj}$ (for $i, j \in [n]$), where E_{ij} denote the elementary matrices in $\mathcal{S}^n$.

4.6 The closure of the completely positive semidefinite cone

One of the most interesting, but hard, open questions regarding the completely positive semidefinite cone is whether it is closed. We make a small progress by giving an alternative description of the closure of $\mathcal{CS}_+$ using the tracial ultraproduct of matrix algebras $\mathbb{R}^{k\times k}$. More precisely, $\text{cl}(\mathcal{CS}_+)$ consists of the symmetric matrices having a Gram representation by positive operators which belong to the mentioned tracial ultraproduct. This ultraproduct is an algebra of bounded operators on an infinite dimensional Hilbert space.

Before introducing tracial ultraproducts, we observe an easier connection between the closure of $\mathcal{CS}_+$ and Gram matrices of operators on infinite dimensional Hilbert spaces. Let $\mathcal{S}^{\mathbb{N}}$ denote the set of all infinite symmetric matrices $X = (X_{ij})_{i,j\geq 1}$ with finite norm: $\sum_{i,j\geq 1} X_{ij}^2 < \infty$. Thus $\mathcal{S}^{\mathbb{N}}$ is a Hilbert space,

equipped with the inner product $\langle X, Y\rangle = \sum_{i,j\geq 1} X_{ij}Y_{ij}$. A matrix $X \in \mathcal{S}^{\mathbb{N}}$ is called *positive semidefinite* if all its finite principal submatrices are positive semidefinite, i.e., $X[I] \in \mathcal{S}_+^{|I|}$ for all finite subsets $I \subseteq \mathbb{N}$, and let $\mathcal{S}_+^{\mathbb{N}}$ denote the set of all positive semidefinite matrices in $\mathcal{S}^{\mathbb{N}}$. Finally, let $\mathcal{CS}_{\infty+}^n$ denote the set of matrices $A \in \mathcal{S}^n$ having a Gram representation by elements of $\mathcal{S}_+^{\mathbb{N}}$. As for $\mathcal{CS}_+^n$, one can verify that $\mathcal{CS}_{\infty+}^n$ is a convex cone. Moreover, we can show the following relationships between these two cones.

4.6.1. THEOREM. *We have:* $\mathcal{CS}_+^n \subseteq \mathcal{CS}_{\infty+}^n \subseteq \mathsf{cl}(\mathcal{CS}_{\infty+}^n) = \mathsf{cl}(\mathcal{CS}_+^n)$.

PROOF: The inclusion $\mathcal{CS}_+^n \subseteq \mathcal{CS}_{\infty+}^n$ is clear. Indeed, any matrix $X \in \mathcal{S}_+^d$ can be viewed as an element of $\mathcal{S}_+^{\mathbb{N}}$ by adding zero entries.

Next we prove the inclusion: $\mathcal{CS}_{\infty+}^n \subseteq \mathsf{cl}(\mathcal{CS}_+^n)$. For this, let $A \in \mathcal{CS}_{\infty+}^n$ and $X_1, \ldots, X_n \in \mathcal{S}_+^{\mathbb{N}}$ be its Gram representation; i.e., $A_{ij} = \langle X_i, X_j\rangle$ for $i, j \in [n]$. For any $\ell \in \mathbb{N}$ and $i \in [n]$, let $X_i^\ell = X_i[\{1, \ldots, \ell\}]$ be the $\ell \times \ell$ upper left principal submatrix of X_i and let $\widetilde{X}_i^\ell \in \mathcal{S}^{\mathbb{N}}$ be the infinite matrix obtained by adding zero entries to X_i^ℓ. Thus, $X_i^\ell \in \mathcal{S}_+^\ell$ and $\widetilde{X}_i^\ell \in \mathcal{S}_+^{\mathbb{N}}$. Now, let A^ℓ denote the Gram matrix of $X_1^\ell, \ldots, X_n^\ell$, so that $A^\ell \in \mathcal{CS}_+^n$. We claim that the sequence $(A^\ell)_{\ell\geq 1}$ converges to A as ℓ tends to infinity, which shows that $A \in \mathsf{cl}(\mathcal{CS}_+^n)$. Indeed, for any $i, j \in [n]$ and $\ell \in \mathbb{N}$, we have:

$$\begin{aligned} |A_{ij} - A_{ij}^\ell| &= |\langle X_i, X_j\rangle - \langle X_i^\ell, X_j^\ell\rangle| \\ &\leq |\langle X_i - \widetilde{X}_i^\ell, X_j\rangle| + |\langle \widetilde{X}_i^\ell, X_j - \widetilde{X}_j^\ell\rangle| \\ &\leq \|X_i - \widetilde{X}_i^\ell\|_{\mathrm{F}}\|X_j\|_{\mathrm{F}} + \|\widetilde{X}_i^\ell\|_{\mathrm{F}}\|X_j - \widetilde{X}_j^\ell\|_{\mathrm{F}} \end{aligned}$$

using the Cauchy-Schwarz inequality in the last step. Clearly, we have that $\|\widetilde{X}_i^\ell\|_{\mathrm{F}} \leq \|X_i\|_{\mathrm{F}} = \sqrt{A_{ii}}$ for all $\ell \in \mathbb{N}$ and $i \in [n]$. Hence $\lim_{\ell\to\infty} |A_{ij} - A_{ij}^\ell| = 0$ for all $i, j \in [n]$, concluding the proof.

Taking the closure in the inclusions: $\mathcal{CS}_+^n \subseteq \mathcal{CS}_{\infty+}^n \subseteq \mathsf{cl}(\mathcal{CS}_+^n)$, we conclude that $\mathsf{cl}(\mathcal{CS}_{\infty+}^n) = \mathsf{cl}(\mathcal{CS}_+^n)$ holds. □

4.6.1 Preliminaries

Let $\mathcal{A}$ be a (Banach) algebra. A subalgebra $\mathcal{A}'$ of $\mathcal{A}$ is a subset $\mathcal{A}' \subseteq \mathcal{A}$ closed under the algebra's operations. The subalgebra $\mathcal{A}'$ is said to be unital if it contains a unit I (i.e., an identity element) and I is also the unit of the original algebra $\mathcal{A}$. The center of an algebra $\mathcal{A}$ is the set of all elements of $\mathcal{A}$ commuting with every element in the algebra. A subset $\mathcal{I} \subset \mathcal{A}$ is called a two-sided ideal if $\mathcal{I}$ is a subspace and $ai, ia \in \mathcal{A}$ whenever $a \in \mathcal{A}$ and $i \in \mathcal{I}$. An ideal $\mathcal{I}$ is said to be maximal if the existence of an ideal $\mathcal{J}$ containing $\mathcal{I}$ (i.e., $\mathcal{J} \supseteq \mathcal{I}$) implies that either $\mathcal{J} = \mathcal{I}$ or $\mathcal{J} = \mathcal{A}$. Every maximal ideal is closed.

Let $\mathcal{M}_k = \mathbb{R}^{k\times k}$ denote the matrix algebra of all $k \times k$ real matrices. We assume that each $\mathcal{M}_k$ is endowed with the normalized trace $\mathsf{tr}_k = \frac{1}{k}\,\mathsf{Tr}$ (if clear from the context, we may omit the dimension and simply write tr) and the corresponding inner product, so that $\|I\|_2^2 = \mathsf{tr}(I) = 1$ where I is the identity matrix. For $T \in \mathcal{M}_k$, $\|T\|_{\mathrm{op}}$ denotes its operator norm and $\|T\|_2$ its L_2-norm. They satisfy the inequality: $\|ST\|_2 \leq \|S\|_{\mathrm{op}} \|T\|_2$ for $S, T \in \mathcal{M}_k$. This, in particular, implies that $\|S\|_2 \leq \|S\|_{\mathrm{op}}$ holds for any $S \in \mathcal{M}_k$.

We denote by $\mathcal{B}(\mathcal{H})$ the (Banach) algebra of bounded linear operators on a Hilbert space $\mathcal{H}$. This is endowed with an involution $*$, which maps an operator T to its adjoint T^*. An operator $T \in \mathcal{B}(\mathcal{H})$ is self-adjoint if $T^* = T$. The operator norm of an element $T \in \mathcal{B}(\mathcal{H})$ is $\|T\|_{\mathrm{op}} = \sup\{\|Tx\| : x \in \mathcal{H}, \|x\| \leq 1\}$. The algebra $\mathcal{B}(\mathcal{H})$ satisfies the identity $\|T^*T\|_{\mathrm{op}} = \|T\|_{\mathrm{op}}^2$ for all $T \in \mathcal{B}(\mathcal{H})$, i.e., $\mathcal{B}(\mathcal{H})$ is a C^*-algebra. The positive operators of $\mathcal{B}(\mathcal{H})$ are exactly the squares of (symmetric) operators.

A *von Neumann algebra* $\mathcal{N}$ is a unital $*$-subalgebra (i.e., a subalgebra closed under the involution $*$) of the algebra $\mathcal{B}(\mathcal{H})$ that is closed in the weak operator topology. The weak operator topology on $\mathcal{B}(\mathcal{H})$ is the weakest topology for which the map $\mathcal{B}(\mathcal{H}) \to \mathbb{C}$ that sends $T \mapsto \langle Tx, y\rangle$ is continuous for any $x, y \in \mathcal{H}$. In other words, a sequence $(T_k)_{k\in\mathbb{N}} \in \mathcal{B}(\mathcal{H})$ converges to $T \in \mathcal{B}(\mathcal{H})$ if, for any $x, y \in \mathcal{H}$, the sequence $(\langle T_k x, y\rangle)_{k\in\mathbb{N}}$ converges to $\langle Tx, y\rangle$ as k tends to infinity.

A *tracial state* (or trace) τ on a von Neumann algebra $\mathcal{N}$ is a linear map $\tau : \mathcal{N} \to \mathbb{C}$ satisfying: (i) $\tau(I) = 1$; (ii) $\tau(T) \geq 0$ for all positive $T \in \mathcal{N}$; and (iii) $\tau(TU) = \tau(UT)$ for any $T, U \in \mathcal{N}$. The tracial state τ is said to be normal if $\tau(T^*T) = 0$ implies $T = 0$, and to be faithful if τ is continuous on the unit ball of $\mathcal{N}$ with respect to the weak operator topology.

4.6.2 Tracial ultraproducts

Tracial ultraproducts of matrix algebras, or more generally of finite von Neumann algebras, are an adapted version of classical ultraproducts from model theory. Their construction is a standard technique in von Neumann algebras (see e.g. the appendix of the book of Brown and Ozawa [BO08]). One usually considers complex Hilbert spaces but the construction works similarly over real Hilbert spaces. Alternatively, one can use the complex construction and 'realify' the resulting algebra afterwards, see for instance [ARU97, Li03]. Ultraproducts are constructions with respect to an ultrafilter. Here we only consider ultrafilters on the natural numbers $\mathbb{N}$. Recall that $\mathcal{P}(\mathbb{N})$ is the collection of all subsets of $\mathbb{N}$.

4.6.2. DEFINITION. An ultrafilter on the set $\mathbb{N}$ is a subset $\mathcal{U} \subseteq \mathcal{P}(\mathbb{N})$ satisfying the conditions: (i) $\emptyset \notin \mathcal{U}$; (ii) if $A \subseteq B \subseteq \mathbb{N}$ and $A \in \mathcal{U}$ then $B \in \mathcal{U}$; (iii) if

$A, B \in \mathcal{U}$ then $A \cap B \in \mathcal{U}$; (iv) for every $A \in \mathcal{P}(\mathbb{N})$ either $A \in \mathcal{U}$ or $\mathbb{N} \setminus A \in \mathcal{U}$.

Combining (i) and (iii) in Definition 4.6.2, any two elements in $\mathcal{U}$ must have nonempty intersection. This allows only two kinds of ultrafilters: either all elements of $\mathcal{U}$ contains a common element $n_0 \in \mathbb{N}$ or $\mathcal{U}$ contains the cofinite sets of $\mathbb{N}$. We are only interested in the second kind of ultrafilters, which are called *free ultrafilters*. For a given free ultrafilter $\mathcal{U}$ on $\mathbb{N}$, we define the *ultralimit* $\lim_{\mathcal{U}} a_k$ of a bounded sequence $(a_k)_{k\in\mathbb{N}}$ of real numbers as follows:

$$\lim_{\mathcal{U}} a_k = a \text{ if } I_\varepsilon \in \mathcal{U} \text{ for all } \varepsilon > 0, \text{ where } I_\varepsilon = \{k \in \mathbb{N} : \ |a_k - a| < \varepsilon\}.$$

4.6.3. REMARK. For any fixed ultrafilter, the ultralimit of any bounded sequence of real numbers is unique.

4.6.4. EXAMPLE. Let $\mathcal{U}$ be a non-free ultrafilter; i.e., $\mathcal{U} = \{A \in \mathcal{P}(\mathbb{N}) : \ k_0 \in A\}$ for some $k_0 \in \mathbb{N}$. Then, $\lim_{\mathcal{U}} a_k = a_{k_0}$ for any sequence $(a_k)_{k\in\mathbb{N}} \subseteq \mathbb{R}$.

4.6.5. EXAMPLE. Let $\mathcal{U}$ be a free ultrafilter, then the ultralimit of a bounded sequence $(a_k)_{k\in\mathbb{N}} \subseteq \mathbb{R}$ is one of its accumulation points.

Consider the sequence $a_k = (-1)^k$ for all $k \in \mathbb{N}$. This has two accumulation points and both can be attained as an ultralimit depending on the choice of the ultrafilter $\mathcal{U}$. By conditions (iii) and (iv) in Definition 4.6.2, we know that any ultrafilter contains either the set $2\mathbb{N}$ of even numbers or its complement, but not both. Hence, there is an ultrafilter $\mathcal{U}$, which contains $2\mathbb{N}$, with $\lim_{\mathcal{U}} a_k = 1$ and an ultrafilter $\mathcal{U}'$, containing the odd numbers $2\mathbb{N} + 1$, with $\lim_{\mathcal{U}'} a_k = -1$.

Next we use ultralimits to construct the tracial ultraproduct of a sequence $(\mathcal{M}_{d_k})_{k\in\mathbb{N}}$ of matrix algebras for some $d_k \in \mathbb{N}$. Here we consider the full sequence $(\mathcal{M}_k)_{k\in\mathbb{N}}$, but the same construction would work for the sequence $(\mathcal{M}_{d_k})_{k\in\mathbb{N}}$. We define the C^*-algebra

$$\ell^\infty(\mathbb{N}, (\mathcal{M}_k)_k) = \{(T_k)_{k\in\mathbb{N}} \in \prod_{k\in\mathbb{N}} \mathcal{M}_k : \ \sup_{k\in\mathbb{N}} \|T_k\|_{\mathrm{op}} < \infty\}.$$

Every free ultrafilter $\mathcal{U}$ on $\mathbb{N}$ defines a two-sided ideal of $\ell^\infty(\mathbb{N}, (\mathcal{M}_k)_k)$

$$\mathcal{I}_{\mathcal{U}} = \{(T_k)_{k\in\mathbb{N}} \in \ell^\infty(\mathbb{N}, (\mathcal{M}_k)_k) : \ \lim_{\mathcal{U}} \|T_k\|_2 = 0\},$$

which is well-defined as sequences in $\ell^\infty(\mathbb{N}, (\mathcal{M}_k)_k)$ are also bounded in the Hilbert-Schmidt norm. The ideal $\mathcal{I}_{\mathcal{U}}$ is maximal and thus closed. The quotient algebra

$$\mathcal{M}_{\mathcal{U}} = \ell^\infty(\mathbb{N}, (\mathcal{M}_k)_k)/\mathcal{I}_{\mathcal{U}}$$

is called the *tracial ultraproduct* of $(\mathcal{M}_k)_k$ along $\mathcal{U}$. One can check that the map

$$\tau_{\mathcal{U}} : \mathcal{M}_{\mathcal{U}} \to \mathbb{R}, \quad (T_k)_{k\in\mathbb{N}} + \mathcal{I}_{\mathcal{U}} \mapsto \lim_{\mathcal{U}} \mathrm{tr}_k(T_k)$$

defines a tracial state on $\mathcal{M}_{\mathcal{U}}$. In fact, $\mathcal{M}_{\mathcal{U}}$ is a finite von Neumann algebra of type II_1 (see definition below). In particular, $\mathcal{M}_{\mathcal{U}}$ is a subalgebra of bounded operators on an infinite dimensional Hilbert space.

4.6.3 Von Neumann algebras and Connes' embedding problem

We give a short overview of the needed concepts; we refer the reader to the book of Takesaki [Tak03] for details. A von Neumann algebra $\mathcal{N}$ is a unital $*$- subalgebra of the algebra $\mathcal{B}(\mathcal{H})$ of bounded operators on a Hilbert space $\mathcal{H}$ that is closed in the weak operator topology. A *factor* is a von Neumann algebra with trivial center (i.e., the center contains only the scalar multiples of the identity I). Every von Neumann algebra on a separable Hilbert space is isomorphic to a direct integral of factors (the appropriate analog of matrix block decomposition).

A factor $\mathcal{F}$ is *finite* if it possesses a normal faithful tracial state $\tau : \mathcal{F} \to \mathbb{C}$. This tracial state τ is unique and gives rise to the Hilbert-Schmidt norm on $\mathcal{F}$ given by $\|T\|_2^2 = \tau(T^*T)$ for $T \in \mathcal{F}$. A von Neumann algebra is finite if it decomposes into finite factors. Every finite von Neumann algebra comes with a trace, which might not be unique.

Von Neumann algebras can be classified into two types depending on the behavior of their projections (i.e., the elements $P \in \mathcal{N}$ satisfying $P = P^* = P^2$). If for a given finite factor $\mathcal{F}$ with trace τ the range of τ over all projections $P \in \mathcal{F}$ is discrete, then $\mathcal{F}$ is of type I. A von Neumann algebra is of type I if it consists only of type I factors. Any finite type I von Neumann algebra is isomorphic to a matrix algebra over $\mathbb{C}$. The only other possibility for a finite factor is that τ maps projections (surjectively) onto $[0, 1]$. Those are II_1 factors, and a von Neumann algebra is of type II_1 if it is finite and contains at least one II_1 factor.

Connes' embedding problem asks to what extent II_1 factors are close to matrix algebras. Murray and von Neumann [MvN36] showed that there is a unique II_1 factor $\mathcal{R}$ which contains an ascending sequence of finite-dimensional von Neumann subalgebras (i.e., matrix algebras) with dense union. This factor $\mathcal{R}$ is called the *hyperfinite* II_1 *factor*. There are several constructions of $\mathcal{R}$, one is as infinite tensor product $\overline{\bigotimes}_{n\in\mathbb{N}} M_2(\mathbb{C})$ of the von Neumann algebras $M_2(\mathbb{C})$, which is the weak closure of the algebraic tensor product $\bigotimes_{n\in\mathbb{N}} M_2(\mathbb{C})$. In fact, any infinite countable sequence of matrix algebras will do.

Connes [Con76] conjectured that all separable II_1 factors embed (in a trace-preserving way) into an ultrapower $\mathcal{R}^{\mathcal{U}}$ of the hyperfinite II_1 factor $\mathcal{R}$, where the ultrapower $\mathcal{R}^{\mathcal{U}}$ is the ultraproduct $\ell^\infty(\mathbb{N}, (\mathcal{R})_k)/\mathcal{I}_{\mathcal{U}}$. As $\mathcal{R}$ contains ascending sequences of matrix algebras with dense union, any matrix algebra $\mathcal{M}_k$ embeds into $\mathcal{R}$. One can extend these embeddings of $\mathcal{M}_k$ into $\mathcal{R}$ to an embedding of the tracial ultraproduct $\mathcal{M}_{\mathcal{U}}$ into $\mathcal{R}^{\mathcal{U}}$ (using a more general construction of ultralimits). Therefore, the finite von Neumann algebra $M_{\mathcal{U}}$ satisfies Connes' embedding conjecture.

This conjecture is equivalent to a huge variety of other important conjec-

tures in e.g. operator theory, noncommutative real algebraic geometry and quantum information theory. In particular, we have already mentioned that it is equivalent to Conjecture 4.3.5 and to deciding whether $\mathsf{cl}(\mathcal{Q}) = \mathcal{Q}_c$ holds.

For an alternative description of $\mathsf{cl}(\mathcal{CS}_+)$ in the case that Connes' embedding conjecture holds true, we will use the following result on finite von Neumann algebras which embed into $\mathcal{R}^{\mathcal{U}}$ (for a proof see e.g. [CD08]). The claim is that tracial moments of an embeddable finite factor can be approximated up to arbitrary precision by matricial tracial moments.

4.6.6. PROPOSITION (COLLINS AND DYKEMA [CD08]). *Let $(\mathcal{F}, \tau)$ be a II_1 factor which embeds into $\mathcal{R}^{\mathcal{U}}$ for some free ultrafilter $\mathcal{U}$. Then $\mathcal{F}$ has matricial microstates, i.e., for any $n \in \mathbb{N}$ and given self-adjoint $T_1, \ldots, T_n \in \mathcal{F}$ the following holds: for every $k \in \mathbb{N}$ and $\varepsilon > 0$ there exist $d \in \mathbb{N}$ and $B_1, \ldots, B_n \in \mathcal{S}^d$ such that*

$$|\tau(T_{i_1} \cdots T_{i_t}) - \mathsf{tr}_d(B_{i_1} \ldots B_{i_t}))| < \varepsilon \text{ for all } i_1, \ldots, i_t \in [n],\ t \leq k.$$

4.6.4 Ultraproduct description of the closure of $\mathcal{CS}_+$

We define a new cone $\mathcal{CS}_{\mathcal{U}+}$ which turns out to be equal to $\mathsf{cl}(\mathcal{CS}_+)$. Fix a free ultrafilter $\mathcal{U}$ on $\mathbb{N}$. Consider the tracial ultraproduct $\mathcal{M}_{\mathcal{U}} = \ell^\infty(\mathbb{N}, (\mathcal{M}_k)_k)/\mathcal{I}_{\mathcal{U}}$, we define

$$\mathcal{CS}^n_{\mathcal{U}+} = \{A \in \mathcal{S}^n_+ : A = (\tau_{\mathcal{U}}(X_i X_j))^n_{i,j=1} \text{ for positive } X_1, \ldots, X_n \in \mathcal{M}_{\mathcal{U}}\}.$$

Note that the trace $\tau_{\mathcal{U}}$ is normalized, i.e., $\tau_{\mathcal{U}}(I) = 1$, whereas we used the (not normalized) trace Tr in the definition of $\mathcal{CS}_+$. Nonetheless, both descriptions agree up to a rescaling of the X_i's.

We first show the inclusion: $\mathsf{cl}(\mathcal{CS}^n_+) \subseteq \mathcal{CS}^n_{\mathcal{U}+}$. To this end, let $A^{(k)}$ be a sequence of matrices in $\mathcal{CS}^n_+$ converging to some $A \in \mathcal{S}^n$; i.e., $\lim_{k\to\infty} A^{(k)}_{ij} = A_{ij}$ for all $i, j \in [n]$. A priori, for each k, there exist an integer d_k and matrices $X^{(k)}_1, \ldots, X^{(k)}_n \in \mathcal{S}^{d_k}_+$ such that $A^{(k)} = (\mathsf{tr}(X^{(k)}_i X^{(k)}_j))^n_{i,j=1}$. The following technical lemma says that without loss of generality we can assume $d_k = k$ for all $k \in \mathbb{N}$.

4.6.7. LEMMA. *Suppose that $(X_k)_k, (Y_k)_k \in \prod_{k\in\mathbb{N}} \mathcal{S}^{d_k}_+$ are such that the sequence $(\mathsf{tr}_{d_k}(X_k Y_k))_{k\in\mathbb{N}}$ converges to some $\gamma \in \mathbb{R}$, then there exist $(X'_k)_k, (Y'_k)_k \in \prod_{k\in\mathbb{N}} \mathcal{S}^k_+$ such that $\mathsf{tr}_k(X'_k Y'_k)$ goes to γ as k tends to infinity.*

PROOF: By possibly reordering the indices, we can assume that the sequence $(d_k)_{k\in\mathbb{N}}$ is monotonically nondecreasing. First, we modify the sequence $(X_k)_k$ in such a way that $d_k \leq k$ holds for all $k \in \mathbb{N}$. For this, if there is some $k \in \mathbb{N}$ with $d_k > k$ we repeat the preceding element X_{k-1} exactly $d_k - k$ times before

the element X_k. For instance, suppose $X_1 \in \mathbb{R}_+$ and $X_2 \in \mathcal{S}_+^3$ (i.e., $d_1 = 1$ and $d_2 = 3$), then we replace the sequence $(X_1, X_2, X_3, \dots)$ by $(X_1, X_1, X_2, X_3, \dots)$. The position of X_k is shifted by $d_k - k$ to $k + d_k - k = d_k$. If $k = 1$ we simply add $d_1 - 1$ zero matrices before X_1. We apply the same procedure to the sequence $(Y_k)_k$. The new sequence of inner products is obtained from the original sequence $(\mathrm{tr}_{d_k}(X_k Y_k))_{k\in\mathbb{N}}$ by replacing each $\mathrm{tr}_{d_k}(X_k Y_k)$ by $d_k - k + 1$ copies of it if $d_k > k$, and thus still converges to the limit γ.

Thus, we have that $d_k \leq k$ for all $k \in \mathbb{N}$. We set $X_k' = \sqrt{\frac{k}{d_k}}(X_k \oplus 0_{k-d_k}) \in \mathcal{S}_+^k$ and $Y_k' = \sqrt{\frac{k}{d_k}}(Y_k \oplus 0_{k-d_k}) \in \mathcal{S}_+^k$ for every $k \in \mathbb{N}$. Therefore we have

$$\mathrm{tr}_k(X_k' Y_k') = \frac{1}{k}\,\mathrm{Tr}(X_k' Y_k') = \frac{1}{k}\frac{k}{d_k}\,\mathrm{Tr}(X_k Y_k) = \mathrm{tr}_{d_k}(X_k Y_k)$$

for every $k \in \mathbb{N}$ and the final sequence $(\mathrm{tr}_k(X_k' Y_k'))_{k\in\mathbb{N}}$ still converges to γ. □

We proceed by showing that the closure of $\mathcal{CS}_+$ is a subset of $\mathcal{CS}_{\mathcal{U}+}$, using Remark 4.6.3 together with Lemma 4.6.7.

4.6.8. LEMMA. *For any free ultrafilter $\mathcal{U}$ on $\mathbb{N}$, we have $\mathsf{cl}(\mathcal{CS}_+^n) \subseteq \mathcal{CS}_{\mathcal{U}+}^n$.*

PROOF: Take a matrix $A \in \mathsf{cl}(\mathcal{CS}_+^n)$, then there exists a sequence of matrices $A^{(k)} \in \mathcal{CS}_+^n$ converging to A: $\lim_{k\to\infty} A_{ij}^{(k)} = A_{ij}$ for all $i, j \in [n]$. For each $k \in \mathbb{N}$, we have $A^{(k)} = (\mathrm{tr}(X_i^{(k)} X_j^{(k)}))_{i,j=1}^n$ for some positive semidefinite matrices $X_1^{(k)}, \dots, X_n^{(k)}$. By Lemma 4.6.7, we can assume that $X_1^{(k)}, \dots, X_n^{(k)} \in \mathcal{S}_+^k$. As the matrices $A^{(k)}$ are bounded, the matrices $X_i^{(k)}$ are bounded as well. Hence, the sequence $(X_i^{(k)})_{k\in\mathbb{N}}$ belongs to $\ell^\infty(\mathbb{N}, (\mathcal{M}_k)_k)$ and we can consider its image X_i in the tracial ultrapower $\mathcal{M}_\mathcal{U}$. By the theorem of Łos (see e.g. [FHS14, Proposition 4.3] and references therein), the operators X_i are positive since all the matrices $X_i^{(k)}$ are positive semidefinite. To conclude that $A \in \mathcal{CS}_{\mathcal{U}+}^n$, we need to show that $A = (\tau_\mathcal{U}(X_i X_j))_{i,j=1}^n$. Observe that, by definition of $\tau_\mathcal{U}$, we have $\tau_\mathcal{U}(X_i X_j) = \lim_\mathcal{U} \mathrm{tr}_k(X_i^{(k)} X_j^{(k)}) = \lim_\mathcal{U} A_{ij}^{(k)}$. On the other hand, since the sequence $(A_{ij}^{(k)})_{k\in\mathbb{N}}$ converges to A_{ij}, by Remark 4.6.3 we have that $\lim_\mathcal{U} A_{ij}^{(k)} = A_{ij}$. This concludes the proof. □

4.6.9. THEOREM. *For any free ultrafilter $\mathcal{U}$ on $\mathbb{N}$, we have $\mathsf{cl}(\mathcal{CS}_+^n) = \mathcal{CS}_{\mathcal{U}+}^n$.*

PROOF: In view of Lemma 4.6.8 we only have to show that $\mathcal{CS}_{\mathcal{U}+}^n \subseteq \mathsf{cl}(\mathcal{CS}_+^n)$. Consider a matrix $A \in \mathcal{CS}_{\mathcal{U}+}^n$, then $A = (\tau_\mathcal{U}(X_i X_j))_{i,j=1}^n$ for some positive operators $X_1, \dots, X_n \in \mathcal{M}_\mathcal{U}$. As the operators X_i are positive, for any $i \in [n]$ there

exists a operator $Y_i \in \mathcal{M}_{\mathcal{U}}$ such that $X_i = Y_i^2$ and each Y_i is given by a sequence of symmetric matrices $(Y_i^{(k)})_{k\in\mathbb{N}} \in \prod_{k\in\mathbb{N}} \mathcal{M}_k$. For $s \in \mathbb{N}$, define the index set $I_s = \{k \in \mathbb{N} : |\tau_{\mathcal{U}}(Y_i^2 Y_j^2) - \mathrm{tr}_k((Y_i^{(k)})^2 (Y_j^{(k)})^2)| \leq 1/s \text{ for all } i, j \in [n]\}$. By definition of $\tau_{\mathcal{U}}$, I_s belongs to $\mathcal{U}$ and, therefore, is nonempty. For any $s \in \mathbb{N}$, we can thus find an index $k_s \in I_s$. Hence the operators $X_i^{(s)} = (Y_i^{(k_s)})^2$ belong to $\mathcal{S}_+^{k_s}$ and satisfy

$$\left|\tau_{\mathcal{U}}(X_i X_j) - \mathrm{tr}_{k_s}(X_i^{(s)} X_j^{(s)})\right| < \frac{1}{s} \text{ for all } i, j \in [n] \text{ and all } s \geq 1. \tag{4.5}$$

For each $s \in \mathbb{N}$, the matrix $A^{(s)} = (\mathrm{tr}_{k_s}(X_i^{(s)} X_j^{(s)}))_{i,j=1}^n$ belongs to $\mathcal{CS}_+^n$ and, by (4.5), the sequence $(A^{(s)})_{s\in\mathbb{N}}$ converges to the matrix A as s tends to infinity. We deduce that A belongs to the closure of $\mathcal{CS}_+^n$. □

We finish this section by giving a possibly alternative description of the closure of $\mathcal{CS}_+$ in the case that Connes' embedding conjecture holds true.

As shown in Theorem 4.6.1, $\mathrm{cl}(\mathcal{CS}_+)$ contains matrices which have a Gram representation by some class of positive semidefinite infinite dimensional matrices. The given description of $\mathrm{cl}(\mathcal{CS}_+)$ as $\mathcal{CS}_{\mathcal{U}+}$ also involves Gram representations by operators on an infinite dimensional Hilbert space. One might ask for the most general infinite dimensional version of $\mathcal{CS}_+$. As we are restricted to operators for which one can define an inner product (or trace), a decent candidate is the following.

4.6.10. DEFINITION. We define the set

$$\mathcal{CS}_{\mathrm{vN}+}^n = \{A \in \mathcal{S}_+^n : A = (\tau_{\mathcal{N}}(X_i X_j))_{i,j=1}^n \text{ for a finite von Neumann algebra } \mathcal{N} \text{ with trace } \tau_{\mathcal{N}} \text{ and some positive } X_1, \ldots, X_n \in \mathcal{N}\},$$

where we allow *any* finite von Neumann algebra $\mathcal{N}$ (with trace $\tau_{\mathcal{N}}$).

Obviously, we have that $\mathcal{CS}_+^n \subseteq \mathcal{CS}_{\mathcal{U}+}^n \subseteq \mathcal{CS}_{\mathrm{vN}+}^n$. Moreover, using the general theory of tracial ultraproducts of von Neumann algebras (instead of just matrix algebras), one can show with a similar line of reasoning as in Lemma 4.6.8 that $\mathcal{CS}_{\mathrm{vN}+}$ is a closed cone. Indeed, take a sequence of matrices $A^{(k)} \in \mathcal{CS}_{\mathrm{vN}+}^n$ converging to some $A \in \mathcal{S}^n$. For each k there exist a finite von Neumann algebra $\mathcal{N}_k$ with trace τ_k and bounded positive operators $X_1^{(k)}, \ldots, X_n^{(k)} \in \mathcal{N}_k$ such that $A^{(k)} = (\tau_k(X_i^{(k)} X_j^{(k)}))_{i,j=1}^n$. Fixing a free ultrafilter $\mathcal{U}$ one can conclude that the images X_i of the sequences $(X_i^{(k)})_{k\in\mathbb{N}}$ in the tracial ultraproduct $\mathcal{N}_{\mathcal{U}} = \ell^\infty(\mathbb{N}, (\mathcal{N}_k)_k)/\mathcal{I}_{\mathcal{U}}$ of the corresponding finite von Neumann algebras provide a Gram representation for A in the von Neumann algebra $\mathcal{N}_{\mathcal{U}}$. This implies the following statement.

4.6.11. THEOREM. *$\mathcal{CS}^n_{\mathrm{vN}+}$ is a closed cone.*

Theorem 4.2.14, due to Frenkel and Weiner [FW14], implies the strict inclusion $\mathcal{CS}^n_{\mathrm{vN}+} \subsetneq \mathcal{S}^n_+ \cap \mathbb{R}^{n\times n}_+$ for any $n \geq 5$. Summarizing we have that:

$$\mathsf{cl}(\mathcal{CS}^n_+) = \mathcal{CS}^n_{\mathcal{U}+} \subseteq \mathcal{CS}^n_{\mathrm{vN}+} \subseteq \mathcal{S}^n_+ \cap \mathbb{R}^{n\times n}_+.$$

Finally, if Connes' embedding conjecture is a true statement, the cone $\mathcal{CS}_{\mathrm{vN}+}$ coincides with the closure of $\mathcal{CS}_+$.

4.6.12. THEOREM. *If Connes' embedding conjecture is true, then* $\mathsf{cl}(\mathcal{CS}^n_+) = \mathcal{CS}^n_{\mathrm{vN}+}$.

PROOF: We only need to show the inclusion $\mathcal{CS}_{\mathrm{vN}+} \subseteq \mathsf{cl}(\mathcal{CS}_+)$. As the line of reasoning is similar to the one in the proof of Theorem 4.6.9, we only give a sketch of the proof. Fix a matrix $A \in \mathcal{CS}_{\mathrm{vN}+}$. Suppose first that $Y_1^2, \ldots, Y_n^2 \in \mathcal{F}$ is a Gram representation of A, where $\mathcal{F}$ is a finite II_1 factor. Since by assumption Connes' embedding conjecture holds, $\mathcal{F}$ embeds into an ultrapower $\mathcal{R}^{\mathcal{U}}$ of the hyperfinite II_1 factor $\mathcal{R}$ for some free ultrafilter $\mathcal{U}$. By Proposition 4.6.6, we can find, for every $k \in \mathbb{N}$, finite dimensional matrices $(Y_1^{(k)})^2, \ldots, (Y_n^{(k)})^2$ approximating the tracial moments $A_{ij} = \tau(Y_i^2 Y_j^2)$ for $i, j \in [n]$ within a distance $1/k$. The corresponding Gram matrices $A^{(k)}$ of $(Y_1^{(k)})^2, \ldots, (Y_n^{(k)})^2$ then belong to $\mathcal{CS}_+$ and therefore the limit point $\lim_{k\to\infty} A^{(k)} = A$ lies in $\mathsf{cl}(\mathcal{CS}_+)$.

Consider now the more general case where $A \in \mathcal{CS}_{\mathrm{vN}+}$ is a Gram matrix of operators Y_i^2 in a finite von Neumann algebra $\mathcal{N}$. Then we can use the same reasoning as in the previous case since any finite von Neumann algebra can be decomposed into finite factors. □

We conclude with mentioning that a hierarchy of semidefinite outer approximations of the cone $\mathcal{CS}_+$ was recently formulated in [BFS15]. These in fact also form outer approximations for the larger cone $\mathcal{CS}_{\mathrm{vN}+}$.

Chapter 5

Applications of the completely positive semidefinite cone

While in the previous chapter we presented some theoretical properties of the completely positive semidefinite cone, here we will see some of its applications. In Section 5.1, we study the quantum graph parameters as conic feasibility programs over the cone $\mathcal{CS}_+$ and, in Sections 5.2 and 5.3, apply to these parameters the hierarchies constructed in Sections 4.4 and 4.5, respectively. Moreover, we will see how to apply the polyhedral hierarchy of Section 4.5 to general optimization problems over the (closure of the) $\mathcal{CS}_+$ cone (Section 5.4) and to construct an hierarchy of polytopes that form an inner approximation to the set of bipartite quantum correlations and cover its relative interior (Section 5.5).

The content of this chapter is based on the results of two papers: one is joint work with Monique Laurent [LP15] and the other is joint work with Sabine Burgdorf and Monique Laurent [BLP15].

5.1 Conic reformulation of quantum graph parameters

We show how to use the completely positive semidefinite cone to study the quantum graph parameters $\chi_q(G), \chi^\star(G)$, $\alpha_q(G)$ and $\alpha^\star(G)$. The idea is to give an alternative definition of each quantum parameter as a conic feasibility program over the completely positive semidefinite cone. This will allow us to make a neat comparison between these quantum parameters, their classical counterparts and the (appropriate variant of the) Lovász theta number.

For simplicity, we will assume throughout that the graph G has a vertex set of cardinality $|V(G)| = n$.

5.1.1 Conic reformulation of the quantum chromatic numbers

We start by reformulating the two quantum variants of the chromatic number as conic feasibility programs over the completely positive semidefinite cone. This prospective allows to easily derive lower and upper bounds for both parameters (Corollary 5.1.8 and Proposition 5.1.10).

The quantum chromatic number $\chi_q(G)$ was introduced in Section 3.3.1. It roughly corresponds to the minimum number of colors needed for the existence of a perfect quantum strategy in the graph coloring game.

For the purposes of this chapter, it will be useful to reformulate Definition 3.3.1 in the following way.

5.1.1. PROPOSITION. *For a graph G, $\chi_q(G)$ is the minimum $t \in \mathbb{N}$ for which there exist positive semidefinite matrices $\rho, \rho_u^i \in \mathcal{S}_+^d$ for $i \in [t]$, $u \in V(G)$ (for some $d \geq 1$) satisfying the conditions:*

$$\begin{aligned} \langle \rho, \rho \rangle &= 1, && (5.1)\\ \textstyle\sum_{i\in[t]} \rho_u^i &= \rho \quad \forall u \in V(G), && (5.2)\\ \langle \rho_u^i, \rho_v^i \rangle &= 0 \quad \forall i \in [t], \forall \{u, v\} \in E(G), && (5.3)\\ \langle \rho_u^i, \rho_u^j \rangle &= 0 \quad \forall i \neq j \in [t], \forall u \in V(G). && (5.4) \end{aligned}$$

PROOF: Suppose there exists $d \times d$ projectors E_u^i that satisfy the conditions of Definition 3.3.1, then the positive semidefinite matrices $\rho_u^i = E_u^i/\sqrt{d}$, $\rho = I/\sqrt{d}$ form a feasible solution for Proposition 5.1.1.

Conversely, suppose that there exist matrices $\rho, \rho_u^i \in \mathcal{S}_+^d$ satisfying conditions (5.1)-(5.4) of Proposition 5.1.1. Let $W \subseteq \mathbb{R}^d$ be the image of ρ and, similarly, let W_u^i be the image of ρ_u^i for all $i \in [t]$, $u \in V(G)$. Then W is the orthogonal sum of the subspaces $\{W_u^i\}_{i\in[t]}$, i.e., $W = \oplus_{i\in[t]} W_u^i$ for all $u \in V(G)$, and therefore the matrices ρ_u^i are projectors. Suppose for the moment that $W = \mathbb{R}^d$, then we let E_u^i be the projection from $\mathbb{R}^d$ onto W_u^i (for all $i \in [t]$, $u \in V(G)$) and these clearly form a feasible solution for Definition 3.3.1. For the case where $W \subset \mathbb{R}^d$, we let E_u^i be the projection from $\mathbb{R}^d$ onto W_u^i for all $i \in [t-1]$, $u \in V(G)$, and E_u^t be the projection from $\mathbb{R}^d$ onto $(\oplus_{i\in[t-1]} W_u^i)^T$ for all $u \in V(G)$. One can easily check that the set of projectors $\{E_u^i\}$ form a feasible solution for Definition 3.3.1. □

The parameter $\chi^\star(G)$ arises in the context of an entanglement-assisted communication problem. We refer the reader to Section 8.1 for further details. For the time being, we will only need to know the following definition.

5.1.2. DEFINITION. [Entangled chromatic number] For a graph G, $\chi^\star(G)$ is the minimum $t \in \mathbb{N}$ for which there exist positive semidefinite matrices $\rho, \rho_u^i \in \mathcal{S}_+^d$ for $i \in [t]$, $u \in V(G)$ (for some $d \geq 1$) satisfying the conditions (5.1), (5.2) and (5.3).

We can now reformulate both quantum variants of the chromatic number as conic feasibility problem over the completely positive semidefinite cone.

5.1.3. PROPOSITION. *For a graph G, $\chi_q(G)$ is equal to the minimum integer t for which there exists a matrix $A \in \mathcal{CS}_+^{|V(G)|t}$ satisfying the following conditions:*

$$\sum_{i,j\in[t]} A_{ui,vj} = 1 \quad \forall u, v \in V(G). \tag{C1}$$

$$A_{ui,vi} = 0 \quad \forall i \in [t],\ \forall \{u, v\} \in E(G), \tag{O1}$$

$$A_{ui,uj} = 0 \quad \forall i \neq j \in [t],\ \forall u \in V(G). \tag{O2}$$

Moreover, the parameter $\chi^\star(G)$ is equal to the minimum integer t for which there exists a matrix $A \in \mathcal{CS}_+^{|V(G)|t}$ satisfying (C1) and (O1).

PROOF: By Proposition 5.1.1, there exist positive semidefinite matrices ρ, ρ_u^i for $i \in [t]$, $u \in V(G)$ satisfying conditions (5.1)-(5.4). Let A be the Gram matrix of the set $\{\rho_u^i\}$; i.e., $A_{ui,vj} = \langle \rho_u^i, \rho_v^j \rangle$ for all $i, j \in [t]$ and $u, v \in V(G)$. By construction $A \in \mathcal{CS}_+^{|V(G)|t}$ and it satisfies (O1) and (O2). Using (5.1) and (5.2), we have $1 = \langle \rho, \rho \rangle = \langle \sum_{i\in[t]} \rho_u^i, \sum_{j\in[t]} \rho_v^j \rangle = \sum_{i,j\in[t]} A_{ui,vj}$ for any $u, v \in V(G)$, which shows (C1).

Conversely, suppose that $A \in \mathcal{CS}_+^{|V(G)|t}$ satisfies conditions (C1), (O1) and (O2) and let ρ_u^i be the positive semidefinite matrices forming a Gram representation of A. It is clear that both (5.3) and (5.4) hold. Let $\rho_u = \sum_{i\in[t]} \rho_u^i$ for any $u \in V(G)$. Then, from (C1) we have that $1 = \sum_{i,j\in[t]} \langle \rho_u^i, \rho_v^j \rangle = \langle \rho_u, \rho_v \rangle$ for all $u, v \in V(G)$, which implies that the positive semidefinite matrices ρ_u are actually all equal (see Lemma 2.1.4 for a proof of this simple claim). Thus, also (5.1) and (5.2) are satisfied and this concludes the proof for the parameter $\chi_q(G)$.

The proof is analogous for $\chi^\star(G)$ and therefore omitted. □

Next we observe that, in Proposition 5.1.3, we can restrict without loss of generality to solutions that are invariant under the action of the permutation group $\mathrm{Sym}(t)$ (consisting of all permutations of $[t] = \{1, \ldots, t\}$). We sketch this well-known symmetry reduction, which has been used in particular for the study of the chromatic number in [GL08].

Given a matrix $A \in \mathcal{S}^{|V(G)|t}$ and a permutation $\pi \in \mathrm{Sym}(t)$, define the new matrix $\pi(A)$ with entries $\pi(A)_{ui,vj} = A_{u\pi(i),v\pi(j)}$ for $i, j \in [t]$, $u, v \in V(G)$, and the matrix $A' = \frac{1}{|\mathrm{Sym}(t)|} \sum_{\pi\in\mathrm{Sym}(t)} \pi(A)$, called the *symmetrization* of A under the action of $\mathrm{Sym}(t)$. Then, A' is invariant under the action of $\mathrm{Sym}(t)$, i.e., $\pi(A') = A'$ for all $\pi \in \mathrm{Sym}(t)$, and thus A' has the following block-form:

$$\begin{pmatrix} X & Y & \ldots & Y \\ Y & X & \ldots & Y \\ \vdots & \vdots & \ddots & \vdots \\ Y & Y & \ldots & X \end{pmatrix} \quad \text{for some } X, Y \in \mathcal{S}^{|V(G)|}. \tag{5.5}$$

Notice that the programs described in Proposition 5.1.3 are invariant under the action of $\mathrm{Sym}(t)$; that is, if A is feasible for one of them then any permutation $\pi(A)$ is feasible too and thus its symmetrization A' as well. Therefore both programs have a feasible solution in block-form (5.5) (assuming one exists).

To prove Proposition 5.1.6 below, we will need the following lemma whose proof can be found for example in [GL08].

5.1.4. LEMMA. *Let A be a $t \times t$ block-matrix with the block structure (5.5), having X as diagonal blocks and Y as off-diagonal blocks, where $X, Y \in \mathcal{S}^k$ (for some $k \geq 1$). Then, $A \succeq 0$ if and only if $X - Y \succeq 0$ and $X + (t-1)Y \succeq 0$.*

Next we consider again the programs introduced in Proposition 5.1.3 to reformulate the parameters $\chi_q(G)$ and $\chi^\star(G)$, and we investigate what is their optimum value when the cone $\mathcal{CS}_+$ is replaced by any of the two cones $\mathcal{CP}$ or $\mathcal{DNN}$. For this we will use the following theorem regarding a property of completely positive matrices. We will apply it in the proof of Proposition 5.1.6 for the choice of B having $\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$ as its 2×2 nonzero principal submatrix.

5.1.5. THEOREM (BARIOLI [BAR01]). *Let $A, B \in \mathcal{S}^n$. Assume that A is a completely positive matrix, B is positive semidefinite with all its entries equal to zero except for a 2×2 principal submatrix, and that $A + B$ is a nonnegative matrix. Then the matrix $A + B$ is completely positive.*

5.1.6. PROPOSITION. *Let G be a graph, $t \geq 1$ be an integer, and let $\mathcal{K}$ denote the cone $\mathcal{DNN}$ or $\mathcal{CP}$. Consider the following three assertions.*

(i) There exists a matrix $A \in \mathcal{K}^{|V(G)|}$ such that $\lceil A_{uu} \rceil = t$ for every $u \in V(G)$, $A_{uv} = 0$ for all $\{u, v\} \in E(G)$ and $A - J \succeq 0$.

(ii) There exists a matrix $A \in \mathcal{K}^{|V(G)|t}$ satisfying the conditions (C1), (O1) and (O2).

(iii) There exists a matrix $A \in \mathcal{K}^{|V(G)|t}$ satisfying the conditions (C1) and (O1).

Then, $(i) \iff (ii) \iff (iii)$ if $\mathcal{K} = \mathcal{DNN}$, and $(iii) \iff (ii) \implies (i)$ if $\mathcal{K} = \mathcal{CP}$.

PROOF: Note that statement (i) is equivalent to saying that $\Theta^{\mathcal{K}}(G) \leq t$ holds.

Assume first $\mathcal{K} = \mathcal{DNN}$. We show: $(i) \Rightarrow (iii) \Rightarrow (ii) \Rightarrow (i)$.

$(i) \Rightarrow (iii)$: Let A be a matrix that satisfies the conditions of (i). By adding an appopriate nonnegative diagonal matrix to A we can assume that $A_{uu} = t$ for all $u \in V(G)$. Set $A' = A - J \in \mathcal{S}^{|V(G)|}$. Then $A' \succeq 0$, $A'_{uu} = t - 1$ for all $u \in V(G)$ and, for all $u \neq v$, $A'_{uv} = A_{uv} - 1 \geq -1$ with equality when $\{u, v\} \in E(G)$. Moreover, $A'_{uv} \geq -(t-1)$ since $A' \succeq 0$ and has diagonal entries equal to $t - 1$.

In the case $t = 1$, we have $A' = 0$, hence G is the empty graph and the all-ones matrix satisfies (iii). We now assume $t \geq 2$. We define the matrices $\widetilde{X} = \frac{1}{t^2}A'$, $\widetilde{Y} = -\frac{1}{t^2(t-1)}A'$, $X = \widetilde{X} + \frac{1}{t^2}J$ and $Y = \widetilde{Y} + \frac{1}{t^2}J \in \mathcal{S}^{|V(G)|}$. We let $B \in \mathcal{S}^{|V(G)|t}$ be the block-matrix as in (5.5) with X as diagonal blocks and Y as off-diagonal blocks and show that B satisfies (iii).

The constrain (O1) holds by construction and (C1) follows from the simple observation that $tX + t(t-1)Y = J$ and thus, for every $u, v \in V(G)$, we have $\sum_{i,j\in[t]} B_{ui,vj} = tX_{uv} + t(t-1)Y_{uv} = 1$. At last, we argue that $B \in \mathcal{DNN}$. Notice that $X, Y \geq 0$ and thus $B \geq 0$. Moreover, since $X + (t-1)Y = J/t \succeq 0$ and $X - Y = A'/t \succeq 0$, using Lemma 5.1.4 we deduce that $B \succeq 0$.

$(iii) \Rightarrow (ii)$: Let B' be a feasible matrix for (iii), we construct a new matrix B satisfying (ii). For this, it suffices to modify each (u,u)-th diagonal block of B' in such a way that all its off-diagonal entries become zero. The idea is simple: move the value of each off-diagonal entry $B'_{ui,uj}$ to the diagonal entry $B'_{ui,ui}$. Formally, for $i \neq j \in [t]$, define $F^{ij} \in \mathcal{S}^t$ to be the matrix with entries $F^{ij}(ij) = F^{ij}(ji) = -1$, $F^{ij}(ii) = F^{ij}(jj) = 1$, and all remaining entries equal to 0. Clearly, F^{ij} is positive semidefinite. Moreover, for $u \in V(G)$, define the matrix $F^{ij}_u \in \mathcal{S}^{|V(G)|t}$ with F^{ij} as its (u,u)-th diagonal block and all remaining entries equal to 0, so that $F^{ij}_u \succeq 0$. The new matrix

$$B = B' + \sum_{u\in V(G)} \sum_{1\leq i<j\leq t} B'_{ui,uj} F^{ij}_u, \tag{5.6}$$

is entrywise nonnegative, positive semidefinite and satisfies (ii).

$(ii) \Rightarrow (i)$: Let $B \in \mathcal{DNN}$ satisfy (ii). Without loss of generality, we can assume that B has the block-form (5.5). Then, $X_{uu} = 1/t$ for all $u \in V(G)$ by (C1) together with (O2), $X_{uv} = 0$ for $\{u,v\} \in E(G)$ by (O1), and $Y_{uu} = 0$ for $u \in V(G)$ by (O2). Moreover, we can rewrite (C1) as $tX_{uv} + t(t-1)Y_{uv} = 1$ for all $u, v \in V(G)$. The matrix $X + (t-1)Y$ is nonnegative, all its diagonal entries are equal to $1/t$ and, by Lemma 5.1.4, is positive semidefinite. Therefore, for any $u, v \in V(G)$ the (u,v)-th entry of $X + (t-1)Y$ lies in the interval $[0, 1/t]$. However, since $\sum_{u,v\in V(G)} X_{uv} + (t-1)Y_{uv} = |V(G)|^2/t$, we have that the matrix $X + (t-1)Y$ must be equal to J/t.

We now construct a matrix $A \in \mathcal{S}^{|V(G)|}$ satisfying (i). Namely, set $A = t^2X$. Thus, $A \in \mathcal{DNN}$, $A_{uu} = t$ for $u \in V(G)$, and $A_{uv} = 0$ for $\{u,v\} \in E(G)$. We are left to show that $A - J \succeq 0$. For this, using the identity $X + (t-1)Y = J/t$ we have that $A - J = t^2X - J = t(t-1)(X - Y)$. Positive semidefiniteness now follows because, due to Lemma 5.1.4, we have $X - Y \succeq 0$. This concludes the proof in the case $\mathcal{K} = \mathcal{DNN}$.

We now consider the case $\mathcal{K} = \mathcal{CP}$. The implication $(ii) \Rightarrow (iii)$ is clear.

$(iii) \Rightarrow (ii)$: We can mimic the above proof of this implication for the $\mathcal{DNN}$ cone. The only thing to notice is that the new matrix B in (5.6) is completely positive, which can proved by applying Theorem 5.1.5. Indeed, $B' \in \mathcal{CP}$, each term $B'_{ui,uj}F_u^{ij}$ is a positive semidefinite matrix whose entries are all zero except for a 2×2 principal submatrix, and one gets a nonnegative matrix at each intermediate step of the summation. Hence, Theorem 5.1.5 can be applied at every step and one can conclude that $B \in \mathcal{CP}$.

$(ii) \Rightarrow (i)$: Again we can mimic the above proof of this implication in the case of $\mathcal{DNN}$. Indeed, we can assume that there exists a matrix $B \in \mathcal{CP}^{|V(G)|t}$ satisfying (ii) and with block-form (5.5), where the block matrices X, Y satisfy the identity: $X + (t-1)T = J/t$. Then, the matrix $X = t^2 A$ belongs to $\mathcal{CP}^{|V(G)|}$ and satisfies (i). □

5.1.7. COROLLARY. *For any graph G, the minimum integer t for which there exists a matrix $A \in \mathcal{K}^{|V(G)|t}$ satisfying the conditions (C1), (O1) and (O2) (or, equivalently, the conditions (C1) and (O1)) is equal to the parameter $\lceil \vartheta^+(\overline{G}) \rceil$ when $\mathcal{K} = \mathcal{DNN}$ and it is equal to the chromatic number $\chi(G)$ when $\mathcal{K} = \mathcal{CP}$.*

PROOF: In the case $\mathcal{K} = \mathcal{DNN}$, the result follows using Proposition 5.1.6 combined with the program of Definition 3.2.6 defining $\vartheta^+(G)$.

Consider the case $\mathcal{K} = \mathcal{CP}$. In view of Proposition 5.1.6, we know that the two conditions (ii) and (iii) are equivalent. Let t denote the minimum integer for which the condition (ii) of Proposition 5.1.6 holds; we show that $\chi(G) = t$. First, we show that $\chi(G) \leq t$. Consider a matrix $A \in \mathcal{CP}^{|V(G)|t}$ satisfying (ii) which has block-form (5.5) and X, Y denote its diagonal and off-diagonal blocks, respectively. As in the proof of implication $(ii) \Rightarrow (i)$ in Proposition 5.1.6, we can deduce that $X - Y \succeq 0$, $X + (t-1)Y = J/t$ and that $\mathsf{Tr}(X) = |V(G)|/t$. This then implies that $\mathsf{Tr}(A) = |V(G)|$ and $\langle J, A \rangle = |V(G)|^2$. Now we use the result of Theorem 3.2.2 for computing the value of $\alpha(G \square K_t)$. For this, set $A' = \frac{1}{|V(G)|} A \in \mathcal{CP}^{|V(G)|t}$. We see that A' satisfies the conditions of the program in Definition 3.2.2 applied to the graph $G \square K_t$. Indeed the orthogonality conditions (O1) and (O2) correspond exactly to the edges of the cartesian graph $G \square K_t$. Therefore, we can deduce that $\alpha(G \square K_t) \geq |V(G)|$. As the reverse inequality also holds (since $G \square K_t$ can be covered by $|V(G)|$ cliques K_t), we have $\alpha(G \square K_t) = |V(G)|$. Using the reduction of Chvátal in Theorem 3.2.1, we can conclude that $\chi(G) \leq t$.

For the reverse inequality, let $\chi(G) = s$. It is easy to see that $G \square K_s$ can be properly colored with s colors and therefore $\chi(G \square K_s) = s$. We construct a matrix $A \in \mathcal{CP}^{|V(G)|s}$ satisfying the conditions of (ii), which will imply $t \leq s$ and thus conclude the proof. For this, select s subsets $S_1, \ldots, S_s \subseteq V(G \square K_s)$ which are stable sets in $G \square K_s$ and partition the vertex set of $G \square K_s$. Since for each $u \in V(G)$ the collection of vertices $\{ui\}_{i \in [s]}$ forms a complete graph, each

set S_k must contain exactly one element from $\{ui\}_{i\in[s]}$. In particular, this implies that for any fixed $u, v \in V(G)$ and set S_k there exists a unique pair (i, j) such that $ui, vj \in S_k$. For $k \in [s]$, let $x_k \in \mathbb{R}^{|V(G)|s}$ denote the incidence vector of S_k and define the matrix $A = \frac{1}{s}\sum_{k=1}^{s} x_k x_k^T$. By construction, $A \in \mathcal{CP}^{|V(G)|s}$, $A_{ui,ui} = 1/s$ for all $u \in V(G), i \in [s]$ and A satisfies conditions (O1) and (O2). Moreover, $\sum_{i,j\in[s]} A_{ui,vj} = \sum_{k=1}^{s}\sum_{i,j\in[s]} \frac{1}{s} x_k(ui)x_k(vj) = 1$ for all $u, v \in V(G)$ by using the above mentioned property of the sets S_k, that is A also satisfies (C1). Hence A is feasible for (ii). This concludes the proof. □

As an application we obtain the following 'sandwich' inequalities for the quantum variants of the chromatic number.

5.1.8. COROLLARY. *For any graph G, $\lceil \vartheta^+(\overline{G}) \rceil \leq \chi^\star(G) \leq \chi_q(G) \leq \chi(G)$.*

We further observe that, in Proposition 5.1.6, the implication $(i) \Rightarrow (ii)$ does not hold when selecting the cone $\mathcal{K} = \mathcal{CP}$.

5.1.9. REMARK. By Corollary 5.1.7, the smallest integer t for which there exists a matrix $A \in \mathcal{CP}^{|V(G)|t}$ satisfying Proposition 5.1.6 (ii) is equal to the chromatic number $\chi(G)$. On the other hand, as a direct application of Theorem 3.2.3, we have that the smallest integer t for which there exists a matrix $X \in \mathcal{CP}^{|V(G)|}$ satisfying Proposition 5.1.6 (i) is equal to $\lceil \chi_f(G) \rceil$, where $\chi_f(G)$ is the fractional chromatic number of G. The inequality $\lceil \chi_f(G) \rceil \leq \chi(G)$ is consistent with the inequality $t \leq s$ corresponding to implication $(ii) \Rightarrow (i)$ in Proposition 5.1.6.

Moreover, as we have mentioned in Section 3.2, the parameters $\lceil \chi_f(G) \rceil$ and $\chi(G)$ can differ significantly. Indeed, for the Kneser graph $K_{a:b}$ where $a \geq 2b$, $\chi_f(K_{a:b}) = a/b$ and $\chi(K_{a:b}) = a - 2b + 2$. This shows that the implication $(i) \Rightarrow (ii)$ does not hold in Proposition 5.1.6 for $\mathcal{K} = \mathcal{CP}$.

We conclude with a comparison between the quantum chromatic numbers and the generalized theta number $\Theta^{\mathcal{CS}_+}(G)$, obtained by selecting the cone $\mathcal{CS}_+$ in Definition 3.2.7.

5.1.10. PROPOSITION. *For any graph G, the following chain of inequalities holds: $\lceil \vartheta^+(\overline{G}) \rceil \leq \lceil \Theta^{\mathsf{cl}(\mathcal{CS}_+)}(G) \rceil \leq \lceil \Theta^{\mathcal{CS}_+}(G) \rceil \leq \chi^\star(G) \leq \chi_q(G)$.*

PROOF: Combining the identity $\Theta^{\mathcal{DNN}}(G) = \vartheta^+(\overline{G})$ (from (3.6)) with the inclusions $\mathcal{CS}_+ \subseteq \mathsf{cl}(\mathcal{CS}_+) \subseteq \mathcal{DNN}$, we obtain the two left most inequalities. Moreover, the right most inequality derives from Corollary 5.1.8.

For the inequality $\lceil \Theta^{\mathcal{CS}_+}(G) \rceil \leq \chi^\star(G)$, we use the fact that $\lceil \Theta^{\mathcal{CS}_+}(G) \rceil$ is the minimum integer t for which Proposition 5.1.6 (i) holds when selecting $\mathcal{K} = \mathcal{CS}_+$, and that $\chi^\star(G)$ is by definition the minimum integer t for which Proposition 5.1.6 (iii) holds with $\mathcal{K} = \mathcal{CS}_+$. Therefore, in order to prove that

$\lceil \Theta^{\mathcal{CS}_+}(G) \rceil \leq \chi^\star(G)$ holds, it suffices to show that Proposition 5.1.6 (iii) implies Proposition 5.1.6 (i) also in the case $\mathcal{K} = \mathcal{CS}_+$. This is what we do next.

Let $B \in \mathcal{CS}_+$ satisfy Proposition 5.1.6 (iii) with $\mathcal{K} = \mathcal{CS}_+$. Again we may assume without loss of generality that B has the block-form (5.5) with blocks X, Y. We can use condition (C1) to show that $X + (t-1)Y = J/t$, following the same steps as in the proof $(ii) \Rightarrow (i)$. Next we consider the matrix $A = t^2 X$. Then $A \in \mathcal{CS}_+$, $A_{uv} = 0$ for every $\{u, v\} \in E(G)$ and $A - J \succeq 0$. Since we started with a solution B of (iii) (instead of a solution for (ii)), we can only derive that $A_{uu} \leq t$ for any $u \in V(G)$. We build a solution A' by adding to A a diagonal matrix D with entries $D_{uu} = t - A_{uu} \geq 0$ for any $u \in V(G)$. Hence $A' \in \mathcal{CS}_+$ and it satisfies all the conditions of Proposition 5.1.6(i). □

5.1.2 Conic reformulation for quantum stability numbers

Analogously to what was done in Section 5.1.1, we reformulate the two quantum stability numbers $\alpha_q(G)$ and $\alpha^\star(G)$ as conic feasibility programs over the completely positive semidefinite cone and use this to retrieve lower and upper bounds for those parameters.

In Section 3.3.2 we have introduced the quantum stability number $\alpha_q(G)$. This is the maximum number t such that there exists a perfect quantum strategy that persuades a referee of the existence of a stable set of cardinality t. Similarly as what we did at the beginning of Section 5.1.1, we reformulate Definition 3.3.4 as follows. We omit the proof since it goes along the same lines as the one of Proposition 5.1.1.

5.1.11. PROPOSITION (QUANTUM STABILITY NUMBER [MR16]). *For a graph G, $\alpha_q(G)$ is the maximum integer $t \in \mathbb{N}$ for which there exist positive semidefinite matrices $\rho, \rho_i^u \in \mathcal{S}_+^d$ for $i \in [t]$, $u \in V(G)$ (for some $d \geq 1$) satisfying the conditions:*

$$\begin{aligned}
\langle \rho, \rho \rangle &= 1, && (5.7)\\
\textstyle\sum_{u \in V(G)} \rho_i^u &= \rho \quad \forall i \in [t], && (5.8)\\
\langle \rho_i^u, \rho_j^v \rangle &= 0 \quad \forall i \neq j \in [t],\ \forall u \simeq v \in V(G), && (5.9)\\
\langle \rho_i^u, \rho_i^v \rangle &= 0 \quad \forall i \in [t],\ \forall u \neq v \in V(G). && (5.10)
\end{aligned}$$

The parameter $\alpha^\star(G)$ is useful to study a zero-error communication problem, which will be explained in Section 6.2. For the purpose of this section, we only need to know the following definition.

5.1.12. DEFINITION. [Entangled stability number [CLMW10]] For a graph G, $\alpha^\star(G)$ is the maximum $t \in \mathbb{N}$ for which there exist positive semidefinite matrices $\rho, \rho_i^u \in \mathcal{S}_+^d$ for $i \in [t]$, $u \in V(G)$ (for some $d \geq 1$) satisfying the conditions (5.7), (5.8) and (5.9).

We can reformulate the two quantum variants $\alpha_q(G)$ and $\alpha^\star(G)$ of the stability number as conic feasibility programs over the cone $\mathcal{CS}_+$. The proof is omitted since it is easy and along the same lines as the one of Proposition 5.1.3.

5.1.13. PROPOSITION. *For a graph G, the parameter $\alpha_q(G)$ is equal to the maximum $t \in \mathbb{N}$ for which there exists a matrix $A \in \mathcal{CS}_+^{|V(G)|t}$ satisfying the conditions:*

$$\sum_{u,v \in V(G)} A_{ui,vj} = 1 \quad \forall i,j \in [t], \tag{C2}$$

$$A_{ui,vj} = 0 \quad \forall i \neq j \in [t], \forall u \simeq v \in V(G), \tag{O3}$$

$$A_{ui,vi} = 0 \quad \forall i \in [t], \forall u \neq v \in V(G). \tag{O4}$$

Moreover, the parameter $\alpha^\star(G)$ is equal to the maximum integer t for which there exists a matrix $A \in \mathcal{CS}_+^{|V(G)|t}$ satisfying (C2) and (O3).

Next we show an analog of Proposition 5.1.6 for the stability numbers and prove that when choosing the cone $\mathcal{K} = \mathcal{CP}$ we find the classical stability number $\alpha(G)$ while, when using the cone $\mathcal{DNN}$, we find the parameter $\lfloor \vartheta'(G) \rfloor$ (Corollary 5.1.16). For this, we will use Lemma 5.1.14 below.

Given a graph G and an integer $t \geq 1$, we introduce the graph G_t which models the orthogonality conditions (O3), (O4); i.e., its vertex set is $V(G) \times [t]$ and two distinct vertices are adjacent in G_t if $i \neq j \in [t]$ and $u \simeq v \in V(G)$, or if $i = j \in [t]$ and $u \neq v \in V(G)$.

5.1.14. LEMMA. *Let G be a graph and let $t \geq 1$ be an integer such that $\vartheta'(G) \geq t$. Then, we have $\vartheta'(G_t) \geq t$.*

PROOF: Let X be a matrix which is an optimal solution for the program of Definition 3.2.6 defining $\vartheta'(G)$, that is $\langle J, X \rangle = \vartheta'(G)$, $\mathsf{Tr}(X) = 1$, $X_{uv} = 0$ for all pairs $\{u,v\} \in E(G)$ and $X \in \mathcal{DNN}$. Set $n = |V(G)|$ and $T = \vartheta'(G)$. Define the diagonal matrix $D \in \mathcal{S}^n$ with $D_{uu} = X_{uu}$ for all $u \in V$ and the matrix $M = (T-1)D \otimes I_t - (D - X) \otimes (J_t - I_t)$ in $\mathcal{S}^{nt}$. Then, M is entrywise nonnegative, its entries are zero at all the positions corresponding to edges of G_t, $\mathsf{Tr}(M) = (T-1)t$, and $\langle J, M \rangle = (T-1)t^2$. Hence, if we can show that $M \succeq 0$, then the matrix $\tilde{M} = \frac{M}{t(T-1)}$ is feasible for the program defining $\vartheta'(G_t)$ with $\langle J, \tilde{M} \rangle = t$, thus showing the desired inequality $\vartheta'(G_t) \geq t$.

We now show that $M \succeq 0$. We may assume that all diagonal entries of X are positive (else replace X by its principal submatrix having only positive diagonal entries). Then, $D \succ 0$ and define $M' = (D^{-1/2} \otimes I_t) M (D^{-1/2} \otimes I_t) = (T-1)I_{nt} - (I_n - D^{-1/2} X D^{-1/2}) \otimes (J_t - I_t)$. It is clear that $M \succeq 0$ if and only if $M' \succeq 0$, which in turn is equivalent to checking whether all the eigenvalues of the matrix $Y = (I_n - D^{-1/2} X D^{-1/2}) \otimes (J_t - I_t)$ are at most $T-1$. Let $0 \leq \lambda_1 \leq \ldots \leq \lambda_n$ denote the eigenvalues of the positive semidefinite matrix

$D^{-1/2}XD^{-1/2}$. Then, the eigenvalues of Y are $(1-\lambda_i)(t-1)$ and $(1-\lambda_i)(-1)$ for $i \in [n]$. Clearly, $(1-\lambda_i)(t-1) \leq t-1 \leq T-1$ for all $i \in [n]$ and thus it suffices to show that $(1-\lambda_i)(-1) = \lambda_i - 1 \leq T-1$ for all $i \in [n]$ or, equivalently, that $\lambda_n \leq T$. To this end, notice that since the matrix $D^{-1/2}XD^{-1/2}$ is nonnegative, by Perron-Frobenius it admits a nonnegative (unit) eigenvector u for its largest eigenvalue λ_n. Define the matrix $X' = D^{-1/2}XD^{-1/2} \circ uu^T \in \mathcal{S}^n$ (taking the entrywise product). Then, $X' \in \mathcal{DNN}^n$, $X'_{uv} = 0$ if $\{u,v\} \in E(G)$, $\mathsf{Tr}(X') = \|x\|_{\mathrm{F}}^2 = 1$, and $\langle J, X' \rangle = u^T D^{-1/2}XD^{-1/2}u = \lambda_n$. As X' is feasible for the program defining $\vartheta'(G)$, it follows that $\lambda_n \leq \vartheta'(G) = T$. □

5.1.15. PROPOSITION. *Let G be a graph, let $t \geq 1$ be an integer, and let $\mathcal{K}$ denote the cone $\mathcal{DNN}$ or $\mathcal{CP}$. The following statements are equivalent.*

(i) There exists a matrix $A \in \mathcal{K}^{|V(G)|}$ satisfying $\lfloor \langle J, X \rangle \rfloor = t$, $\mathsf{Tr}(X) = 1$ and $X_{uv} = 0$ for all $\{u,v\} \in E(G)$.

(ii) There exists a matrix $A \in \mathcal{K}^{|V(G)|t}$ satisfying the conditions (C2), (O3) and (O4).

(iii) There exists a matrix $A \in \mathcal{K}^{|V(G)|t}$ satisfying the conditions (C2) and (O3).

PROOF: Notice that statement (i) is equivalent to $\vartheta^{\mathcal{K}}(G) \geq t$. We will show the implications $(i) \Rightarrow (iii) \Rightarrow (ii) \Rightarrow (i)$, starting with the case $\mathcal{K} = \mathcal{DNN}$.

$(i) \Rightarrow (iii)$: Assume first $t = 1$. If (i) holds with $t = 1$, then by Remark 3.2.8 there exists a matrix $A \in \mathcal{DNN}^{|V(G)|}$ with $\mathsf{Tr}(A) = \langle J, A \rangle = 1$ and therefore $A_{uv} = 0$ for all $u \neq v \in V(G)$. Thus A satisfies (iii).

Assume now that $t \geq 2$. If (i) holds, then $\vartheta'(G) \geq t$ and from Lemma 5.1.14 we can conclude that $\vartheta'(G_t) \geq t$. Using Remark 3.2.8, we know there exists a matrix $A \in \mathcal{S}^{|V(G)|t}$ feasible for the program of Definition 3.2.6 defining $\vartheta'(G_t)$ with value $\langle J, A \rangle = t$. Hence the matrix $B = tA \in \mathcal{DNN}^{|V(G)|t}$ satisfies $\langle J, B \rangle = t^2$, $\mathsf{Tr}(B) = t$ and $B_{ui,vj} = 0$ for all edges $\{(u,i),(v,j)\}$ of G_t. Moreover, after symmetrization by $\mathrm{Sym}(t)$, we can assume that B has the block-form (5.5), where X is a diagonal matrix and $Y_{uv} = 0$ for all edges $\{u,v\}$ of G. Then, $t = \mathsf{Tr}(B) = t\,\mathsf{Tr}(X) = t\langle J, X \rangle$ and $t^2 = \langle J, B \rangle = t\langle J, X \rangle + t(t-1)\langle J, Y \rangle$. This implies that $\mathsf{Tr}(X) = \langle J, X \rangle = \langle J, Y \rangle = 1$ (since $t \geq 2$). Then B satisfies (C2) and therefore (iii).

$(iii) \Rightarrow (ii)$: Assume that B' satisfies (iii); we construct a new matrix B satisfying (ii). Similarly to the proof of implication $(iii) \Rightarrow (ii)$ in Proposition 5.1.6, it suffices to modify each (i,i)-th diagonal block $B'[ii] = (B'_{ui,vi})_{u,v \in V(G)}$ of B' in such a way that its off-diagonal entries become zero. For any $u \neq v \in V(G)$, define the matrix $F^{uv} \in \mathcal{S}^{|V(G)|}$ where the entries $F^{uv}(uv) = F^{uv}(vu) = -1$, $F^{uv}(uu) = F^{uv}(vv) = 1$ and all remaining entries are zero. Clearly, $F^{uv} \succeq 0$. Moreover, for $i \in [t]$, define the matrix $F_i^{uv} \in \mathcal{S}^{|V(G)|t}$ with F^{uv} as its (i,i)-th

diagonal block and all remaining entries equal to 0, so that $F_i^{uv} \succeq 0$. Fix an arbitrary ordering of the vertices of G. Define the new matrix

$$B = B' + \sum_{i\in[t]} \sum_{u<v\in V(G)} B'_{ui,vi} F_i^{uv}. \tag{5.11}$$

By construction, the sum of entries of the (i,i)-th diagonal block of B is equal to the sum of entries of the (i,i)-th diagonal block of B' and thus equal to 1. The matrix B is entrywise nonnegative and it is a sum of positive semidefinite matrices. It then follows that B satisfies (ii).

$(ii) \Rightarrow (i)$: Let B be a matrix satisfying (ii). As $B \succeq 0$, there exist vectors y_i^u (for $u \in V(G), i \in [t]$) forming a Gram representation of B. For any $i \in [t]$, let $y_i = \sum_{u\in V(G)} y_i^u$. Using condition (C2), $1 = \sum_{u,v\in V(G)} \langle y_i^u, y_j^v\rangle = \langle y_i, y_j\rangle$ holds for all $i, j \in [t]$, which implies that the vectors y_i are all equal. Define the vectors $x_u = \sum_{i\in[t]} y_i^u$ for all $u \in V(G)$ and let $A \in \mathcal{S}^{|V(G)|}$ denote their Gram matrix. Then, $A \succeq 0$, $\langle J, A\rangle = \|\sum_{u\in V(G)} \sum_{i=1}^t y_i^u\|^2 = \|ty_i\|^2 = t^2$, and has trace $\mathrm{Tr}(A) = \sum_{u\in V(G)} \|x_u\|^2 = \sum_{i,j\in[t]} \sum_{u\in V(G)} \langle y_i^u, y_j^u\rangle = \sum_{i\in[t]} \sum_{u\in V(G)} B_{ui,ui} = t$. Moreover, the entry $A_{uv} = \langle x_u, x_v\rangle = \sum_{i,j\in[t]} \langle y_i^u, y_j^v\rangle = \sum_{i,j\in[t]} B_{ui,vj} \geq 0$ for any $u, v \in V(G)$, with equality for $\{u, v\} \in E(G)$. Rescaling the matrix A by $1/t$, we obtain a feasible solution for (i).

We now consider the case $\mathcal{K} = \mathcal{CP}$.

$(i) \Rightarrow (iii)$: Let A be a matrix that satisfies (i). Applying Theorem 3.2.2, we obtain that $\alpha(G) \geq t$. Let $S \subseteq V(G)$ be a stable set of cardinality t. Say, $V(G) = [n]$ and $S = \{1, \ldots, t\}$. Define the vector $x \in \mathbb{R}_+^{nt}$ with block-form $x = (e_1, \ldots, e_t)$, where $e_1, \ldots, e_t$ are the first t standard unit vectors in $\mathbb{R}^n$. Define the matrix $B' = xx^T$ which, by construction, belongs to $\mathcal{CP}^{nt}$. One can easily verify that B' satisfies (iii).

$(iii) \Rightarrow (ii)$: We can mimic the above proof of this implication in the case of the cone $\mathcal{DNN}$. We only need to observe that the new matrix B in (5.11) is completely positive. This is the case because Theorem 5.1.5 can be applied at every step of the summation, since one gets a nonnegative matrix at each step.

$(ii) \Rightarrow (i)$: The reasoning is analogous to the above proof of this implication for $\mathcal{DNN}$. □

As an application, if in Proposition 5.1.13 we replace the cone $\mathcal{CS}_+$ by the cone $\mathcal{DNN}$ in the definition of $\alpha_q(G)$ or of $\alpha^\star(G)$, then we obtain the parameter $\lfloor\vartheta'(G)\rfloor$; analogously, if we replace the cone $\mathcal{CS}_+$ by the cone $\mathcal{CP}$ then we obtain $\alpha(G)$.

5.1.16. COROLLARY. *For any graph G, the maximum integer t for which there exists a matrix $X \in \mathcal{K}^{|V(G)|t}$ satisfying the conditions (C2), (O3) and (O4) (or, equivalently, the conditions (C2) and (O3)) is equal to the parameter $\lfloor\vartheta'(G)\rfloor$ when $\mathcal{K} = \mathcal{DNN}$ and it is equal to the stability number $\alpha(G)$ when $\mathcal{K} = \mathcal{CP}$.*

PROOF: We simply apply Proposition 5.1.15 combined with the program of Definition 3.2.6 defining ϑ' when $\mathcal{K} = \mathcal{DNN}$ and with Theorem 3.2.2 when $\mathcal{K} = \mathcal{CP}$. □

In turn this permits to derive the following 'sandwich inequalities' for the quantum analogs of the stability number.

5.1.17. COROLLARY. *For any graph G, $\alpha(G) \leq \alpha_q(G) \leq \alpha^\star(G) \leq \lfloor \vartheta'(G) \rfloor$.*

The bound $\alpha^\star(G) \leq \lfloor \vartheta'(G) \rfloor$ was recently shown, with a different method, by Cubitt et al. [CMR+14]. The inequality $\alpha(G) \leq \alpha_q(G)$ can be strict (see [MR16]), but it is not known whether the other two inequalities can be strict.

Observe that, if one could prove that the two conditions (ii) and (iii) in Proposition 5.1.15 are equivalent also when setting $\mathcal{K} = \mathcal{CS}_+$, this would imply the identity $\alpha_q(G) = \alpha^\star(G)$. This would work if we could show an analog of Theorem 5.1.5 when replacing the condition of being 'completely positive' by the condition of being 'completely positive semidefinite', since then the reasoning used in the proof of Proposition 5.1.15 for the implication $(iii) \Rightarrow (ii)$ would extend to the case of $\mathcal{CS}_+$. However, the following example shows that Theorem 5.1.5 does not extend to the cone $\mathcal{CS}_+$.

5.1.18. EXAMPLE. Consider the matrix $L = M(\cos^2(\frac{4\pi}{5}), \cos^2(\frac{2\pi}{5}))$, which was presented in Example 4.2.5 as an example of a matrix which is completely positive semidefinite but not completely positive. For $i \neq j \in [5]$, let $F^{ij} \in \mathcal{S}^5_+$ be the matrix with all zero entries except $F^{ij}_{ii} = F^{ij}_{jj} = 1$ and $F^{ij}_{ij} = F^{ij}_{ji} = -1$. Define $L' = L + \cos^2(\frac{2\pi}{5})(F^{13} + F^{24} + F^{35} + F^{14} + F^{25})$. This matrix is not completely positive since its inner product with the Horn matrix is negative: $\langle H, L' \rangle = 5(1 + 2\cos^2(\frac{2\pi}{5})) - 10\cos^2(\frac{4\pi}{5}) = 5(2 - \sqrt{5})/2 < 0$. As the support of L' is equal to the 5-cycle, we can conclude using Theorem 4.2.9 that L' is not completely positive semidefinite.

Therefore, although one starts from a completely positive semidefinite matrix and at each step of the summation nonnegativity is preserved, the final matrix L' does not belong to the completely positive semidefinite cone. We deduce that Theorem 5.1.5 does not extend to the cone $\mathcal{CS}_+$.

Finally, we relate the quantum stability number $\alpha_q(G)$ with the generalized theta number $\vartheta^{\mathcal{CS}_+}(G)$, obtained when selecting the cone $\mathcal{K} = \mathcal{CS}_+$ in Definition 3.2.7.

5.1.19. PROPOSITION. *For any graph G, we have the following chain of inequalities: $\alpha_q(G) \leq \lfloor \vartheta^{\mathcal{CS}_+}(G) \rfloor \leq \lfloor \vartheta^{\mathrm{cl}(\mathcal{CS}_+)}(G) \rfloor \leq \lfloor \vartheta'(G) \rfloor$.*

PROOF: From (3.5), we have the identity $\vartheta^{\mathcal{DNN}}(G) = \vartheta'(G)$ which together with $\mathcal{CS}_+ \subseteq \mathsf{cl}(\mathcal{CS}_+) \subseteq \mathcal{DNN}$ gives $\vartheta^{\mathcal{CS}_+}(G) \leq \vartheta^{\mathsf{cl}(\mathcal{CS}_+)}(G) \leq \vartheta'(G)$ and thus the two right most inequalities.

For the inequality $\alpha_q(G) \leq \lfloor \vartheta^{\mathcal{CS}_+}(G) \rfloor$, we revisit the proof of Proposition 5.1.15. First we observe that the implication $(ii) \Rightarrow (i)$ remains true in Proposition 5.1.15 if we select the cone $\mathcal{K} = \mathcal{CS}_+$. (Indeed, the same proof applies as in the case $\mathcal{K} = \mathcal{DNN}$, except that y_i^u are now positive semidefinite matrices and we need to use Lemma 2.1.4 to be able to claim that all the y_i are equal.) By definition, $\alpha_q(G)$ is the largest integer t for which Proposition 5.1.15 (ii) holds with $\mathcal{K} = \mathcal{CS}_+$. In turn, by the above, this largest number is at most the largest integer t for which Proposition 5.1.15 (i) holds with $\mathcal{K} = \mathcal{CS}_+$, the latter being equal to $\lfloor \vartheta^{\mathcal{CS}_+}(G) \rfloor$. Thus $\alpha_q(G) \leq \lfloor \vartheta^{\mathcal{CS}_+}(G) \rfloor$ holds. □

We do not know whether $\vartheta^{\mathcal{CS}_+}(G)$ also provides an upper bound for $\alpha^\star(G)$, since we cannot show that Proposition 5.1.15 (iii) implies Proposition 5.1.15 (i) when $\mathcal{K} = \mathcal{CS}_+$. The proof used when $\mathcal{K} \in \{\mathcal{DNN}, \mathcal{CP}\}$ does not extend to the case $\mathcal{K} = \mathcal{CS}_+$ since Theorem 5.1.5 does not hold if we consider matrices in $\mathcal{CS}_+$ (as shown in Example 5.1.18).

5.2 Approximations using the set $\mathcal{K}_{\mathrm{nc},\varepsilon}$

We show how one can use the convex sets $\mathcal{K}_{\mathrm{nc},\varepsilon}$ introduced earlier in Section 4.4 to define parameters that approximate the quantum graph parameters. We give the details only for the quantum chromatic number $\chi_q(G)$, but the same reasoning can be extended to the other parameters $\chi^\star(G)$, $\alpha_q(G)$ and $\alpha^\star(G)$.

The construction will go as follows. In a first step we reformulate $\chi_q(G)$ as a single 'aggregated' minimization program over an affine section of the cone $\mathcal{CS}_+$. When replacing the cone $\mathcal{CS}_+$ by its closure $\mathsf{cl}(\mathcal{CS}_+)$ we get the parameter $\widetilde{\chi}_q(G)$, satisfying $\chi_q(G) \geq \widetilde{\chi}_q(G)$. The second step will consist of writing the dual of this aggregated conic program over the cone $\mathsf{cl}(\mathcal{CS}_+)$, which is thus a maximization program over the dual cone $\mathcal{CS}_+^*$, and showing that strong duality holds. Finally, we define new parameters $\Psi_\varepsilon(G)$ by replacing in this dual conic program the cone $\mathcal{CS}_+^*$ by the convex sets $\mathcal{K}_{\mathrm{nc},\varepsilon}$.

We start with a slightly different conic formulation of the quantum chromatic number $\chi_q(G)$ than the one in Proposition 5.1.3. Given a set of positive semidefinite matrices that satisfies the conditions in Proposition 5.1.1, consider the matrix $A \in \mathcal{CS}_+^{|V(G)|t+1}$ defined as the Gram matrix of the set $\rho, \{\rho_u^i\}$ for $u \in V(G)$ and $i \in [t]$. Constructing the matrix in this way allows for the following conic reformulation of $\chi_q(G)$, the proof is omitted as it is equivalent to the one in Proposition 5.1.3.

5.2.1. PROPOSITION. *For a graph G, $\chi_q(G)$ is equal to the minimum integer t for which there exists a matrix $A \in \mathcal{CS}_+^{|V(G)|t+1}$ (indexed by $\{0\} \cup V(G) \times [t]$) satisfying the following conditions: (i) $A_{0,0} = 1$; (ii) $\sum_{i\in[t]} A_{0,ui} = 1$ for all $u \in V(G)$; (iii) $\sum_{i,j\in[t]} A_{ui,uj} =$ for all $u \in V(G)$; (iv) $A_{ui,vi} = 0$ for all $i \in [t]$, $\{u,v\} \in E(G)$; and (v) $A_{ui,uj} = 0$ for all $i \neq j \in [t], u \in V(G)$.*

5.2.2. REMARK. The minimum natural number t for which there exists a matrix $A \in \mathcal{DNN}^{|V(G)|t+1}$ satisfying conditions (*i*), (*ii*), (*iii*), (*iv*) and (*v*) of Proposition 5.2.1 is equal to $\lceil \vartheta^+(\overline{G}) \rceil$.

Indeed, consider a matrix A satisfying the above conditions and let $x, \{x_u^i\}$ (for $u \in V(G)$, $i \in [t]$) be its Gram representation. Then, $x = \sum_{i\in[t]} x_u^i$ for any $u \in V(G)$, as $\|x - \sum_{i\in[t]} x_u^i\|^2 = A_{0,0} - 2\sum_{i\in[t]} A_{0,ui} - \sum_{i,j\in[t]} A_{ui,uj} = 0$. Therefore, $\sum_{i,j\in[t]} A_{ui,vj} = \langle x, x\rangle = 1$ for any $u, v \in V(G)$ and the matrix A' which has Gram representation $\{x_u^i\}$ (for $u \in V(G)$, $i \in [t]$) is a feasible solution for Proposition 5.1.6 (*ii*) when $\mathcal{K} = \mathcal{DNN}$.

Conversely, take a matrix $A \in \mathcal{DNN}^{|V(G)|t}$ feasible for Proposition 5.1.6 (*ii*) and let $\{x_u^i\}$ (for $u \in V(G)$, $i \in [t]$) be its Gram representation. We define $x_u = \sum_{i\in[t]} x_u^i$ for any $u \in V(G)$. Since $\sum_{i,j\in[t]} A_{ui,vj} = 1$ for any $u, v \in V(G)$, we derive that the vectors $\{x_u\}_{u\in V(G)}$ are all equal to, say, x. Consider the matrix A' which is the Gram of the vectors $x, \{x_u^i\}$ (for $u \in V(G)$, $i \in [t]$). One can easily check that A' satisfies conditions (*i*), (*ii*), (*iii*), (*iv*) and (*v*) of Proposition 5.2.1. Using Corollary 5.1.7, we then derive the claim.

We define new matrices that are useful to formulate the constraints of Proposition 5.2.1. Let $D_u^t \in \mathcal{S}^{|V(G)|t+1}$ (for $u \in V(G)$, $t \in [|V(G)|]$) be the matrix with entries $D_u^t(0,0) = D_u^t(ui,uj) = 1$ for all $i, j \in [t]$, $D_u^t(0,ui) = D_u^t(ui,0) = -1$ for all $i \in [t]$ and zero elsewhere, and set $D^t = \sum_{u\in V(G)} D_u^t$. Observe that each matrix D_u^t is positive semidefinite (with rank 1). Notice that using Proposition 5.2.1, $\chi_q(G)$ is equal to the smallest $t \in \mathbb{N}$ for which there exists a matrix $A \in \mathcal{CS}_+^{|V(G)|t+1}$ satisfying the conditions (*i*), (*iv*) and (*v*) of Proposition 5.2.1 together with $\langle D^t, A\rangle = 0$. We can now reformulate $\chi_q(G)$ as the optimal value of a single conic optimization program over the cone $\mathcal{CS}_+$.

5.2.3. PROPOSITION. *Let G be a graph and set $n = |V(G)|$. The quantum chromatic number $\chi_q(G)$ is equal to the optimal value of the program:*

$$\begin{aligned} \min \ \sum_{t\in[n]} t A_{0,0}^t \quad \text{s.t. } & A^t \in \mathcal{CS}_+^{nt+1} \quad \forall t \in [n], \\ & \textstyle\sum_{t\in[n]} A_{0,0}^t = 1, \ \sum_{t\in[n]} \langle D^t, A^t\rangle = 0, \\ & A_{ui,vi}^t = 0 \quad \forall i \in [t], \forall \{u,v\} \in E(G), \ \forall t \in [n], \\ & A_{ui,uj}^t = 0 \quad \forall i \neq j \in [t], \forall u \in V(G), \ \forall t \in [n]. \end{aligned} \tag{5.12}$$

PROOF: Set $t = \chi_q(G)$ and let μ denote the optimal value of the program (5.12).

Let $A \in \mathcal{CS}_+^{|V(G)|t+1}$ be a solution for the program from Proposition 5.2.1 defining $\chi_q(G)$. We obtain a solution $A^1, \ldots, A^n$ to the program (5.12) by setting $A^t = A$ and $A^i = 0$ if $i \in [n] \setminus \{t\}$. This shows that $\mu \leq t$.

Conversely, let $A^1, \ldots, A^n$ be a solution for the program (5.12) with value $\widetilde{\mu}$ and let s be the minimum $i \in [n]$ such that $A^i_{0,0} \neq 0$. Then, consider the matrix $A = A^s/A^s_{0,0}$ which is feasible for the program in Proposition 5.2.1. We have: $t \leq s = s\sum_{i\in[n]} A^i_{0,0} = s\sum_{i\geq s} A^i_{0,0} \leq \sum_{i\geq s} iA^i_{0,0} = \sum_{i\in[n]} iA^i_{0,0} = \widetilde{\mu}$. This shows that $t \leq \mu$ and thus the identity $\chi_q(G) = \mu$ holds. Moreover, this also gives that program (5.12) has indeed an optimal solution, thus justifying writing 'min' rather than 'inf' in (5.12). □

5.2.4. REMARK. Ji [Ji13] proved that deciding whether $\chi_q(G) \leq 3$ is an NP-hard problem. Combining this with Proposition 5.2.3 one gets that linear optimization over affine sections of the completely positive semidefinite cone is also an NP-hard problem.

It is convenient to rewrite program (5.12) in a more compact way. For this set $N = \sum_{t=1}^n (nt+1)$, where $n = |V(G)|$, and define $D = \oplus_{t=1}^n D^t \in \mathcal{S}^N$. Let $E^t_{0,ui}$, $E^t_{ui,vj}$ denote the elementary matrices in $\mathcal{S}^{nt+1}$ and let $\widetilde{E}^t_{0,ui}$, $\widetilde{E}^t_{ui,vj}$ denote their extensions to $\mathcal{S}^N$ obtained by adding a border of zero entries. Moreover, set $F = \oplus_{t=1}^n tE^t_{0,0}$ and $\hat{F} = \oplus_{t=1}^n E^t_{0,0} \in \mathcal{S}^N$. We rewrite the program (5.12) as

$$\begin{aligned} \chi_q(G) = \min \langle F, A\rangle \text{ s.t. } & A \in \mathcal{CS}_+^N,\ \langle \hat{F}, A\rangle = 1,\ \langle D, A\rangle = 0, \\ & \langle \widetilde{E}^t_{ui,vi}, A\rangle = 0\ \forall i \in [t], \forall \{u,v\} \in E(G), \forall t \in [n], \\ & \langle \widetilde{E}^t_{ui,uj}, A\rangle = 0\ \forall i \neq j \in [t], \forall u \in V(G), \forall t \in [n]. \end{aligned} \tag{5.13}$$

5.2.5. REMARK. Any feasible solution A of program (5.13) defining $\chi_q(G)$ lies on the border of the $\mathcal{CS}_+$ cone (due to Lemma 4.1.7).

If in the program (5.13) we replace the cone $\mathcal{CS}_+$ by its closure $\mathsf{cl}(\mathcal{CS}_+)$, then its optimal value is equal to $\widetilde{\chi}_q(G)$ and we have: $\widetilde{\chi}_q(G) \leq \chi_q(G)$. By Remark 5.2.5 it is not clear whether these two parameters coincide. On the other hand, one can verify that the result of Proposition 5.2.3 (and its proof) extend to the case when the cone $\mathcal{CS}_+$ is replaced by its closure $\mathsf{cl}(\mathcal{CS}_+)$. Hence, $\widetilde{\chi}_q(G)$ can be equivalently defined by using the program from Proposition 5.2.1 after replacing the cone $\mathcal{CS}_+$ by its closure $\mathsf{cl}(\mathcal{CS}_+)$. (Another equivalent formulation of $\widetilde{\chi}_{q(G)}$ is given in Definition 5.3.1 in the next section.) Using this, Corollary 5.1.7 and the fact that $\mathcal{CS}_+ \subseteq \mathsf{cl}(\mathcal{CS}_+) \subseteq \mathcal{DNN}$, we get:

$$\lceil \vartheta^+(\overline{G}) \rceil \leq \widetilde{\chi}_q(G) \leq \chi_q(G).$$

The dual program of (5.13) reads:

$$\sup\ \lambda \text{ s.t. } M = F - \lambda \hat{F} - \mu D - \sum y^t_{u,v,i} \widetilde{E}^t_{ui,vi} - \sum z^t_{u,i,j} \widetilde{E}^t_{ui,uj} \in \mathcal{CS}^{N*}_+, \quad (5.14)$$

where the variables are λ, μ, $y^t_{u,v,i}$ and $z^t_{u,i,j}$, the first summation is over $t \in [n]$, $i \in [t]$ and $\{u,v\} \in E(G)$, and the second summation is over $t \in [n]$, $i \neq j \in [t]$ and $u \in V(G)$. Let $\lambda_q(G)$ be the optimal value of the program (5.14). By weak duality, we have: $\lambda_q(G) \leq \widetilde{\chi}_q(G) \leq \chi_q(G)$.

Moreover, the program (5.14) is strictly feasible, hence there is no duality gap and the optimal value of (5.14) is equal to $\widetilde{\chi}_q(G)$. That is, we have that $\lambda_q(G) = \widetilde{\chi}_q(G) \leq \chi_q(G)$. To see that (5.14) is strictly feasible, define the matrix $M^t = (t+n^2)E^t_{0,0} + D^t - \sum_{u \in V(G)} \sum_{i \neq j \in [t]} E^t_{ui,uj}$ and set $M = \oplus_{t=1}^n M^t$. Then, M is feasible for the program (5.14). Moreover, M lies in the interior of $\mathcal{CS}^{N*}_+$ since $M \succ 0$, as $M^t \succ 0$ for all t. (Indeed, the entries of M^t are equal to $M^t_{0,0} = n + t + n^2$, $M^t_{0,ui} = -1$, $M^t_{ui,ui} = 1$ and zero otherwise, and take a Schur complement, see (2.1), to derive that $M^t \succ 0$).

We now introduce the new parameter $\Psi_\varepsilon(G)$, which is obtained by replacing in the program (5.14) the cone $\mathcal{CS}^*_+$ by the convex set $\mathcal{K}_{\mathrm{nc},\varepsilon}$.

5.2.6. DEFINITION. For $\varepsilon \geq 0$, let $\Psi_\varepsilon(G)$ be the optimal value of the program:

$$\sup\ \lambda \text{ s.t. } M = F - \lambda \hat{F} - \mu D - \sum y^t_{u,v,i} \widetilde{E}^t_{ui,vi} - \sum z^t_{u,i,j} \widetilde{E}^t_{ui,uj} \in \mathcal{K}_{\mathrm{nc},\varepsilon}. \quad (5.15)$$

5.2.7. LEMMA. *For $\varepsilon \geq 0$, we have: $\lceil \vartheta^+(\overline{G}) \rceil \leq \Psi_\varepsilon(G)$, with equality if $\varepsilon = 0$.*

PROOF: By Lemma 4.4.3, we have the inclusion $\mathcal{DNN}^* \subseteq \mathcal{K}_{\mathrm{nc},\varepsilon}$, with equality if $\varepsilon = 0$. Hence the claim will follow if we can show that the optimal value of the program (5.15), when we replace the set $\mathcal{K}_{\mathrm{nc},\varepsilon}$ by its subset $\mathcal{DNN}^*$, is equal to $\lceil \vartheta^+(\overline{G}) \rceil$.

In other words, let us consider the program (5.14) where we replace the cone $\mathcal{CS}^*_+$ by the cone $\mathcal{DNN}^*$. Using the same argument as above, we can conclude that its optimal value is equal to the optimal value of the program (5.13) where we replace the cone $\mathcal{CS}_+$ by the cone $\mathcal{DNN}$ (strong duality holds and use the fact that the cone $\mathcal{DNN}$ is closed).

Next, observe that this latter value (which is equal to the optimal value of the program (5.12) when we replace $\mathcal{CS}_+$ by $\mathcal{DNN}$) is equal to $\lceil \vartheta^+(\overline{G}) \rceil$. This can be seen by combining Remark 5.2.2 together with the fact that the result of Proposition 5.2.3 (and its proof) extends to the case when we replace the cone $\mathcal{CS}_+$ by the cone $\mathcal{DNN}$. □

As the sets $\mathcal{K}_{\mathrm{nc},\varepsilon}$ aim to approximate the dual cone $\mathcal{CS}^*_+$, the parameters $\Psi_\varepsilon(G)$ aim to approximate the quantum coloring number $\chi_q(G)$. However,

as there is no apparent inclusion relationship between $\mathcal{CS}_+^*$ and $\mathcal{K}_{\mathrm{nc},\varepsilon}$, we do not know the exact relationship between $\Psi_\varepsilon(G)$ and $\chi_q(G)$. Moreover, as the cone $\mathcal{CS}_+$ is not known to be closed, there is a possible gap between the two parameters $\chi_q(G)$ and $\widetilde{\chi}_q(G)$. Nevertheless, what we can claim is the following relationship under Connes' embedding conjecture.

5.2.8. LEMMA. *If Connes' embedding conjecture is true, then* $\widetilde{\chi}_q(G) \leq \inf_{\varepsilon>0} \Psi_\varepsilon(G)$.

PROOF: If Connes' conjecture holds then $\mathcal{CS}_+^* \subseteq \mathcal{K}_{\mathrm{nc},\varepsilon}$ for any $\varepsilon > 0$ (from Lemma 4.4.2). The result now follows using the definition of $\Psi_\varepsilon(G)$ and the definition of $\widetilde{\chi}_q(G)$ as the optimal value of (5.14). □

Finally, we observe that the parameter $\Psi_\varepsilon(G)$ can be obtained as the limit of a sequence of semidefinite programs. For this, recall that M lies in $\mathcal{K}_{\mathrm{nc},\varepsilon}$ if the polynomial $p_M + \varepsilon$ admits a decomposition of the form $p_M + \varepsilon = g + h$, where $g = \sum_{j=1}^{m_0} f_j f_j^* + \sum_{i=1}^{n} \sum_{j_i=1}^{m_i} g_{j_i}(1 - X_i^2) g_{j_i}^*$ for some $f_j, g_{j_i} \in \mathbb{R}\langle \underline{X} \rangle$ and $m_0, m_i \in \mathbb{N}$, and h is a sum of commutators. Fixing an integer k and restricting to those decompositions of $p_M + \varepsilon$ where all terms $f_j f_j^*$ and $g_{j_i}(1 - X_i^2) g_{j_i}^*$ have degree at most $2k$, we get a parameter $\Psi_\varepsilon^k(G)$ which can be computed via a semidefinite program (see e.g. [Bur11] for details). Moreover, $\Psi_\varepsilon^k(G)$ tends to $\Psi_\varepsilon(G)$ as k goes to infinity.

5.3 Linear programming lower bounds to the quantum graph parameters

In this section we apply the polyhedral hierarchy $\mathcal{C}_r^n$ defined in Section 4.5.2 to get linear programming bounds for the quantum graph parameters. We will show the construction in details for the quantum chromatic number $\chi_q(G)$, but the same ideas extend also to the parameters $\chi^\star(G)$, $\alpha_q(G)$ and $\alpha^\star(G)$.

In Theorem 4.5.10, we showed that the hierarchy $\mathcal{C}_r^n$ asymptotically covers the full interior of $\mathcal{CS}_+^n$. However, from Remark 5.2.5 (or also Lemma 4.1.8) we know that any feasible solution for $\chi_q(G)$ lies in the border of the $\mathcal{CS}_+$ cone. To ensure the existence of a feasible solution in the interior of the cone $\mathcal{CS}_+$ and thus to be able to use the hierarchy, we will relax the affine constraints defining $\chi_q(G)$ (using a small perturbation). In this way we will be able to get a hierarchy of parameters that can be computed through linear programming and give the exact value of $\widetilde{\chi}_q(G)$ (see Definition 5.3.1 below). We remark that this result is existential. We can prove the existence of a linear program permitting to compute the quantum parameter, but we do not know at which stage this happens. This result should be seen in the light of a recent result of the same flavor proved by Paulsen et al. [PSS$^+$16]. The authors of [PSS$^+$16] consider yet another variant $\chi_{qc}(G)$ of the quantum chromatic number (see (3.12)

in Section 3.3.1 for a definition), satisfying $\chi_{qc}(G) \leq \chi_q(G)$, and they show that $\chi_{qc}(G)$ can be computed with a positive semidefinite program (also not explicitly known).

In the same paper, Paulsen et al. [PSS+16] introduced the parameter $\chi_{qa}(G)$. Using the same approach we will also show that a variant $\widetilde{\chi}_{qa}(G)$ of the parameter $\chi_{qa}(G)$ can be written as a linear program.

We start by recalling some definitions and by underlining the link among all the various parameters.

Let $\mathcal{A}^t$ represent the affine space in $\mathcal{S}^{|V(G)|t}$ defined by the equations

$$\sum_{i,j\in[t]} A_{ui,vj} = 1 \text{ for } u, v \in V(G), \tag{5.16}$$

and $L_{G,t} : \mathcal{S}^{|V(G)|t} \to \mathbb{R}$ denote the linear map defined by

$$L_{G,t}(A) = \sum_{u\in V(G), i\neq j\in[t]} A_{ui,uj} + \sum_{uv\in E(G), i\in[t]} A_{ui,vi}. \tag{5.17}$$

We can then reformulate the definition of $\chi_q(G)$ in Proposition 5.1.3 as

$$\chi_q(G) = \min t \in \mathbb{N} \text{ s.t. } \exists A \in \mathcal{CS}_+^{|V(G)|t},\ A \in \mathcal{A}^t \text{ and } L_{G,t}(A) = 0.$$

We introduced the variant $\widetilde{\chi}_q(G)$ by replacing the cone $\mathcal{CS}_+$ by its closure in the above definition.

5.3.1. DEFINITION. For a graph G, the parameter $\widetilde{\chi}_q(G)$ is defined as follows:

$$\widetilde{\chi}_q(G) = \min t \in \mathbb{N} \text{ s.t. } \exists A \in \mathsf{cl}(\mathcal{CS}_+^{|V(G)|t}),\ A \in \mathcal{A}^t \text{ and } L_{G,t}(A) = 0,$$

where $\mathcal{A}^t$ and $L_{G,t}$ are defined by (5.16) and (5.17), respectively.

The parameter $\chi_{qa}(G)$, which was defined in (3.11), can also be written as a feasibility program over the affine section of the cone $\mathsf{cl}(\mathcal{CS}_+)$. Indeed, using (2.6) (see Theorem 5.5.3 for details), we can rewrite the identities (3.10) and (3.11), which define, respectively, $\chi_q(G)$ and $\chi_{qa}(G)$, as follows:

$$\begin{aligned}\chi_q(G) &= \min t \text{ s.t. } \exists P \in \pi(\mathcal{CS}_+^{2nt} \cap \mathcal{B}^{2nt}) \text{ with } \mathcal{L}_{G,t}(P) = 0, \text{ and}\\ \chi_{qa}(G) &= \min t \text{ s.t. } \exists P \in \mathsf{cl}(\pi(\mathcal{CS}_+^{2nt} \cap \mathcal{B}^{2nt})) \text{ with } \mathcal{L}_{G,t}(P) = 0.\end{aligned}$$

Analogously to the way we have defined the variant $\widetilde{\chi}_q(G)$ of the parameter $\chi_q(G)$, we introduce the variant $\widetilde{\chi}_{qa}(G)$ by replacing $\mathcal{CS}_+$ by its closure in the above definition of $\chi_{qa}(G)$. Namely,

$$\widetilde{\chi}_{qa}(G) = \min t \text{ s.t. } \exists P \in \pi(\mathsf{cl}(\mathcal{CS}_+^{2nt}) \cap \mathcal{B}^{2nt}) \text{ with } \mathcal{L}_{G,t}(P) = 0. \tag{5.18}$$

Note that the set $\mathsf{cl}(\mathcal{CS}_+) \cap \mathcal{B}^{2nt}$ is bounded and thus compact, so that its projection $\pi(\mathsf{cl}(\mathcal{CS}_+) \cap \mathcal{B}^{2nt})$ is compact too.This is the reason why in (5.18) we have written $P \in \pi(\mathsf{cl}(\mathcal{CS}_+^{2nt}) \cap \mathcal{B}^{2nt})$ instead of $P \in \mathsf{cl}(\pi(\mathsf{cl}(\mathcal{CS}_+^{2nt}) \cap \mathcal{B}^{2nt}))$. The inclusion $\mathcal{CS}_+ \cap \mathcal{B}^{2nt} \subseteq \mathsf{cl}(\mathcal{CS}_+) \cap \mathcal{B}^{2nt}$ implies:

$$\mathsf{cl}(\pi(\mathcal{CS}_+ \cap \mathcal{B}^{2nt})) \subseteq \pi(\mathsf{cl}(\mathcal{CS}_+) \cap \mathcal{B}^{2nt})$$

and thus the following relationship: $\widetilde{\chi}_{qa}(G) \leq \chi_{qa}(G)$.

Note that if a matrix A is feasible for the program of Definition 5.3.1, then the matrix $R = \left(\begin{smallmatrix} A & A \\ A & A \end{smallmatrix}\right)$ is feasible for the program (5.18) defining $\widetilde{\chi}_{qa}(G)$. Hence, $\widetilde{\chi}_{qa}(G) \leq \widetilde{\chi}_q(G)$ holds.

The relationship among the parameters $\chi_q(G), \chi_{qc}(G), \chi_{qa}(G)$ and $\widetilde{\chi}_{qa}(G)$, $\widetilde{\chi}_q(G)$ can be summarized as follows:

$$\begin{array}{ccc} \chi_{qc}(G) & \leq & \chi_{qa}(G) \leq \chi_q(G) \\ & & \vee_I \qquad \vee_I \\ & & \widetilde{\chi}_{qa}(G) \leq \widetilde{\chi}_q(G). \end{array}$$

We are ready now to explain how to use the hierarchy $\mathcal{C}_r^n$ to build linear relaxations for the parameters $\widetilde{\chi}_q(G), \widetilde{\chi}_{qa}(G)$. We will illustrate the method for $\widetilde{\chi}_q(G)$ but, as it will be mention at the end of the section, a similar approach can be taken for $\widetilde{\chi}_{qa}(G)$.

A first natural approach is to replace the cone $\mathsf{cl}(\mathcal{CS}_+^{nt})$ in the definition of $\widetilde{\chi}_q(G)$ by the subcone $\mathcal{C}_r^{nt}$ leading to the parameter

$$\ell_r(G) = \min t \in \mathbb{N} \text{ s.t. } \exists A \in \mathcal{C}_r^{nt},\ A \in \mathcal{A}^t \text{ and } L_{G,t}(A) = 0.$$

(Recall that throughout we assume $|V(G)| = n$.) As $\mathcal{C}_r^{nt} \subseteq \mathcal{CS}_+^{nt}$, we have $\widetilde{\chi}_q(G) \leq \chi_q(G) \leq \ell_r(G)$. Moreover, the sequence $(\ell_r(G))_{r\in\mathbb{N}}$ of natural numbers is monotonically nonincreasing and thus has a limit (it even becomes stationary). However, it is not clear whether the limit is equal to $\chi_q(G)$. If one could claim that for $t = \chi_q(G)$ there is a feasible matrix A for Definition 5.3.1 which lies in the interior of $\mathcal{CS}_+^{nt}$ then, by Theorem 4.5.10, A would belong to some cone $\mathcal{C}_r^{nt}$ which in turn would imply the identity $\chi_q(G) = \ell_r(G)$. But this idea cannot work because, as observed in Lemma 4.1.8 (see also Lemma 4.1.7), any matrix feasible for Definition 5.3.1 lies on the boundary of $\mathcal{CS}_+^{nt}$. To go around this difficulty, our strategy is to relax the affine constraints in Definition 5.3.1 so as to allow feasible solutions in the interior of $\mathcal{CS}_+^{nt}$.

For any integer $k \geq 1$, let $\mathcal{A}_k^t$ be the affine space defined by the equations:

$$|\sum_{i,j\in[t]} A_{ui,vj} - 1| \leq \frac{1}{k} \text{ for } u, v \in V(G).$$

We define the parameter:

$$\lambda_k(G) = \min t \text{ s.t. } \exists A \in \mathsf{cl}(\mathcal{CS}_+^{nt}),\ A \in \mathcal{A}_k^t \text{ and } L_{G,t}(A) \leq \frac{1}{k}. \tag{5.19}$$

In a first step we show that $\lambda_k(G) = \widetilde{\chi}_q(G)$ for k large enough.

5.3.2. LEMMA. *For any graph G, there exists $k_0 \in \mathbb{N}$ such that $\widetilde{\chi}_q(G) = \lambda_k(G)$ holds for all $k \geq k_0$.*

PROOF: Notice that $\lambda_k(G) \leq \widetilde{\chi}_q(G)$ holds for every $k \in \mathbb{N}$. Indeed, any matrix solution for $\widetilde{\chi}_q(G)$ is also a solution for $\lambda_k(G)$. Moreover, as the sequence $(\lambda_k(G))_{k\in\mathbb{N}}$ is a monotone nondecreasing sequence of natural numbers upper bounded by $\widetilde{\chi}_q(G)$, there exists a k_0 such that $\lambda_k(G) = \lambda_{k_0}(G)$ for all $k \geq k_0$. Let $t = \lambda_{k_0}(G)$. For all $k \geq k_0$, there exists a matrix $A_k \in \mathsf{cl}(\mathcal{CS}_+^{nt})$ with $A_k \in \mathcal{A}_k^t$ and $L_{G,t}(A_k) \leq 1/k$. Consider the sequence $(A_k)_{k\geq k_0}$, which is bounded as all A_k lie in $\mathcal{A}_{k_0}^t$. Therefore, the sequence has a converging subsequence to, say, A where $A \in \mathsf{cl}(\mathcal{CS}_+^{nt})$, $A \in \mathcal{A}^t$ and $L_{G,t}(A) = 0$. Hence, A is a feasible solution for $\widetilde{\chi}_q(G)$ and $\widetilde{\chi}_q(G) \leq t = \lambda_{k_0}(G) = \lambda_k(G)$ for all $k \geq k_0$. □

In a second step we show that the new parameter $\lambda_k(G)$ can be computed by a linear program. For this we replace in the definition of $\lambda_k(G)$ the cone $\mathsf{cl}(\mathcal{CS}_+^{nt})$ by the polyhedral cone $\mathcal{C}_r^{nt}$, leading to the following parameter:

$$\lambda_k^r(G) = \min t \text{ s.t. } \exists A \in \mathcal{C}_r^{nt},\ A \in \mathcal{A}_k^t \text{ and } L_{G,t}(A) \leq \frac{1}{k}. \tag{5.20}$$

Notice that this parameter $\lambda_k^r(G)$ can be computed through a linear program since $\mathcal{C}_r^{nt}$ is a polyhedral cone. We will show that for any graph G there exist integers k_0 and r_0 such that $\widetilde{\chi}_q(G) = \lambda_{k_0}^{r_0}(G)$. We emphasize that this is an existential result: we do not know for which integers k_0 and r_0 such a convergence happens. One of the ingredients to prove the result is to show the existence of a matrix in the interior of $\mathcal{CS}_+$ satisfying certain constraints. To this end, we will use the matrix $Z = I + J \in \mathcal{S}^{nt}$ where I and J are, respectively, the identity and the all-ones matrix.

5.3.3. LEMMA. *The matrix $Z = I + J \in \mathcal{S}^{nt}$ lies in the interior of $\mathcal{CS}_+^{nt}$. Moreover, we have that $\sum_{i,j\in[t]} Z_{ui,uj} = t^2 + t$ for all $u \in V(G)$, $\sum_{i,j\in[t]} Z_{ui,vj} = t^2$ for all $u \neq v \in V(G)$ and $L_{G,t}(Z) = nt^2 - nt + mt$, where m is the number of edges of the graph G.*

PROOF: We only show that $I + J$ lies in the interior of $\mathcal{CS}_+^{nt}$, the other claims follow from direct computations. Assume that there exists a matrix $M \in \mathcal{CS}_+^{nt*}$ such that $\langle M, I + J\rangle = 0$, we show that $M = 0$. Indeed, as both I and J lie in $\mathcal{CS}_+^{nt}$ we get that $\mathsf{Tr}(M) = 0$ and $\langle J, M\rangle = 0$. Observe that, since M is copositive with zero diagonal entries, all entries of M must be nonnegative. Combining this with $\langle J, X\rangle = 0$, we deduce that M is identically zero. □

5.3.4. THEOREM. *For any graph G, there exist integers k_0 and $r_0 \in \mathbb{N}$ such that $\widetilde{\chi}_q(G) = \lambda_k^r(G)$ for all $k \geq k_0$ and all $r \geq r_0$. Moreover $\lambda_{k_0}^{r_0}(G)$, and thus $\widetilde{\chi}_q(G)$, can be computed via a linear program.*

PROOF: From Lemma 5.3.2 we know that there exists an integer $k_0 \in \mathbb{N}$ such that $\lambda_k(G) = \widetilde{\chi}_q(G)$ for all $k \geq k_0$. In view of this, we just need to show that for this k_0 there exists an integer $r_0 \in \mathbb{N}$ for which $\lambda_{k_0}^{r_0}(G) = \lambda_{k_0}(G)$. Let $t = \lambda_{k_0}(G) = \widetilde{\chi}_q(G)$.

By the definitions (5.19) and (5.20) and the inclusion relationships between the cones $\mathcal{C}_r^{nt}$, we have that the sequence $(\lambda_{k_0}^r)_{r\in\mathbb{N}}$ of natural numbers is non-increasing and it is lower bounded by $\lambda_{k_0}(G)$. Hence there exists an $r_0 \in \mathbb{N}$ such that $\lambda_{k_0}^r(G) = \lambda_{k_0}^{r_0}(G) \geq \lambda_{k_0}(G)$ for all $r \geq r_0$. We are left to prove that $\lambda_{k_0}^{r_0}(G) \leq \lambda_{k_0}(G) = t$.

To this end, we show that there exists a matrix $Y_{k_0} \in \text{int}(\mathcal{CS}_+^{nt})$ such that $Y_{k_0} \in \mathcal{A}_{k_0}^t$ and $L_{G,t}(Y_{k_0}) \leq 1/k_0$. This will suffice since then, by Theorem 4.5.10, $Y_{k_0} \in \mathcal{C}_{r_0}^{nt}$ for some r_0. Therefore, Y_{k_0} satisfies the conditions in program (5.20) and thus $\lambda_{k_0}^{r_0}(G) \leq t = \lambda_{k_0}(G)$. To show the existence of such a matrix Y_{k_0}, let $A \in \text{cl}(\mathcal{CS}_+^{nt})$ be a feasible solution for the program of Definition 5.3.1 defining $\widetilde{\chi}_q(G) = t$ and consider the matrix $Z = I + J$ which belongs to $\text{int}(\mathcal{CS}_+^{nt})$ by Lemma 5.3.3. Any convex combination $Z_\varepsilon = (1-\varepsilon)A + \varepsilon Z$ (for $0 < \varepsilon < 1$) lies in the interior of $\mathcal{CS}_+^{nt}$. If we can tune ε so that the new matrix Z_ε satisfies the conditions in program (5.20), then we can choose $Y_{k_0} = Z_\varepsilon$ and we are done. We claim that selecting $\varepsilon = \min\{\frac{1}{k_0(t^2+t-1)}, \frac{1}{k_0(nt^2-nt+2mt)}\}$ will do the trick. Indeed, for such ε we have $Z_\varepsilon \in \text{int}(\mathcal{CS}_+^{nt})$ and $L_{G,t}(Z_\varepsilon) = \varepsilon L_{G,t}(Z) \leq 1/k_0$ (use Lemma 5.3.3). Moreover, Z_ε lies in $\mathcal{A}_{k_0}^t$ since for all $u, v \in V(G)$ the following holds

$$\begin{aligned} \Big|\sum_{i,j\in[t]} Z_\varepsilon(ui,vj) - 1\Big| &= \Big|(1-\varepsilon) + \varepsilon \sum_{i,j\in[t]} Z_{ui,vj} - 1\Big| \\ &\leq \Big| -\varepsilon + \varepsilon \sum_{i,j\in[t]} Z_{ui,uj}\Big| = \big|\varepsilon(t^2+t-1)\big| \leq \frac{1}{k_0}. \end{aligned}$$

Summarizing, from Lemma 5.3.2 we know that there exists an integer $k_0 \in \mathbb{N}$ such that $\lambda_{k_0}(G) = \widetilde{\chi}_q(G)$ and we just proved that for this k_0 there exists an integer $r_0 \in \mathbb{N}$ with the property that $\lambda_{k_0}^{r_0}(G) = \lambda_{k_0}(G) = \widetilde{\chi}_q(G)$. □

The same result holds for the parameter $\widetilde{\chi}_{qa}(G)$ introduced in (5.18). For clarity we rewrite its definition in the following form:

$$\widetilde{\chi}_{qa}(G) = \min t \in \mathbb{N} \text{ s.t. } \exists A \in \text{cl}(\mathcal{CS}_+^{2nt}), A \in \mathcal{B}^{2nt} \text{ with } \mathcal{L}_{G,t}(\pi(A)) = 0.$$

This parameter is quite similar to $\widetilde{\chi}_q(G)$, the only differences being that we now work with matrices A of size $2nt$ (instead of nt) lying in the affine space $\mathcal{B}^{2nt}$

(instead of $\mathcal{A}^t$) and satisfying $\mathcal{L}_{G,t}(\pi(A)) = 0$ (instead of $L_{G,t}(A) = 0$). In analogy to the parameter $\lambda_k(G)$ we can define the parameter $\Lambda_k(G)$ by doing these replacements and defining the relaxed affine space $\mathcal{B}_k^{2nt}$ in the same way as $\mathcal{A}_k^t$ was defined from $\mathcal{A}^t$. Then the analog of Lemma 5.3.2 holds: there exists an integer k_0 such that $\widetilde{\chi}_{qa}(G) = \Lambda_k(G)$ for all $k \geq k_0$. Next, replacing the cone $\mathsf{cl}(\mathcal{CS}_+^{2nt})$ by $\mathcal{C}_r^{2nt}$, we get the parameter $\Lambda_k^r(G)$ (the analog of $\lambda_r^k(G)$):

$$\Lambda_k^r(G) = \min t \in \mathbb{N} \text{ s.t. } A \in \mathcal{C}_r^{2nt}, A \in \mathcal{B}_k^{2nt} \text{ with } \mathcal{L}_{G,t}(\pi(A)) \leq \frac{1}{k}.$$

The analog of Theorem 5.3.4 holds, whose proof is along the same lines and thus omitted.

5.3.5. THEOREM. *For any graph G, there exist natural numbers k_0 and r_0 such that $\widetilde{\chi}_{qa}(G) = \Lambda_k^r(G)$ for all $k \geq k_0$ and $r \geq r_0$. Hence the parameter $\widetilde{\chi}_{qa}(G)$ can be computed by a linear program.*

5.4 Polyhedral approach for optimization over the completely positive semidefinite cone

Here we explain how to extend the polyhedral approach explained in the previous section to a more general class of optimization problems over the (closure of the) $\mathcal{CS}_+^n$ cone. Consider the following optimization program:

$$\min \ \langle C, A\rangle \text{ s.t. } A \in \mathsf{cl}(\mathcal{CS}_+^n), \ A \in \mathcal{A} \text{ and } \mathcal{L}(A) = 0, \tag{5.21}$$

where $C \in \mathcal{S}^n$, $\mathcal{L}$ a linear functional on $\mathcal{S}^n$, and $\mathcal{A} \subseteq \mathcal{S}^n$ is an affine subspace of $\mathcal{S}^n$ with the property that $\mathcal{A} \cap \mathcal{CS}_+^n$ is bounded. In particular, assume that $\mathcal{A}$ is defined by the affine equations $\langle B_j, A\rangle = b_j$ (for $j \in [m]$).

Analogously to what was done in Section 5.3, we can define a double hierarchy yielding a sequence of two-parameters LP-based bounds which converges *asymptotically* to the optimum value of the above optimization program. More concretely, for any integer $k \geq 1$ define the parameter

$$\lambda_k = \min \ \langle C, A\rangle \text{ s.t. } A \in \mathsf{cl}(\mathcal{CS}_+^n), A \in \mathcal{A}_k \text{ and } |\mathcal{L}(A)| \leq \frac{1}{k},$$

where the affine space $\mathcal{A}_k$ is a perturbed version of $\mathcal{A}$ defined by the constraints $|\langle B_j, A\rangle - b_j| \leq 1/k$ (for $j \in [m]$). Using similar arguments as for Lemma 5.3.2, one can show that the sequence $(\lambda_k)_{k\in\mathbb{N}}$ is monotone non-decreasing and converges to the optimum value of program (5.21) as k tends to infinity. However, in contrast to Lemma 5.3.2, we cannot guarantee finite convergence in general (the finite convergence in Lemma 5.3.2 followed from the fact that the

parameter $\lambda_k(G)$ is integer valued, which is generally the case). Next, for any integer $r \geq 1$ define the parameter

$$\lambda_k^r = \min \ \langle C, A \rangle \text{ s.t. } A \in \mathcal{C}_r^n, A \in \mathcal{A}_k \text{ and } |\mathcal{L}(A)| \leq \frac{1}{k}.$$

Using similar arguments as for the proof of Theorem 5.3.4, one can show that the sequence $(\lambda_k^r)_{r\in\mathbb{N}}$ is monotone non-increasing and converges to λ_k. Hence, we obtain the sequence of parameters $(\lambda_k^r)_{k,r\in\mathbb{N}}$ asymptotically convergences to the optimum value of program (5.21) as both the parameters k, r tend to infinity.

5.5 Polyhedral approximations for the set of quantum correlations

As yet another application of the polyhedral hierarchy $\mathcal{C}_r^n$ (defined in Section 4.5.2), we construct a hierarchy of polytopes that form inner approximations to the set of bipartite quantum correlations $\mathcal{Q}$ and cover its relative interior.

5.5.1 The set of bipartite quantum correlations

The set of bipartite quantum correlations, commonly denoted as $\mathcal{Q}$, consists of the conditional probabilities that two physically separated parties can generate by performing measurements on a shared entangled state. More formally, recall the following definition.

5.5.1. DEFINITION. [Quantum correlations] A conditional bipartite probability distribution $(P(a, b|x, y))_{a\in A, b\in B, x\in X, y\in Y}$ is called *quantum* if

$$P(a, b|x, y) = \langle \psi, (E_x^a \otimes F_y^b)\psi \rangle,$$

where $\psi \in \mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$ is a unit vector (for some $d_A, d_B \in \mathbb{N}$) and for some sets of positive semidefinite matrices (aka POVM) $\{E_x^a\}_{a\in A}$ and $\{F_y^b\}_{b\in B}$ satisfying $\sum_{a\in A} E_x^a = I$ and $\sum_{b\in B} F_y^b = I$ for all $x \in X, y \in Y$.

The set of bipartite quantum correlations $\mathcal{Q}$ consists of all bipartite quantum probabilities.

5.5.2. REMARK. Without loss of generality, in the above definition we can assume that all positive semidefinite matrices E_x^a, F_y^b are real valued and that, for some $d \in \mathbb{N}$, the matrices E_x^a, F_y^b lie in $\mathcal{S}_+^d$ and the vector ψ is in $\mathbb{R}^{d^2}$.

We introduce some notation. For $x \in X$, we let $\underline{E}_x$ denote the tuple $(E_x^a)_{a\in A}$ and then the tuple $\underline{E} = (\underline{E}_x)_{x\in X}$ contains all matrices E_x^a for $a \in A, x \in X$. Analogously, for $y \in Y$, let $\underline{F}_y$ denote the tuple $(F_y^b)_{b\in B}$ and $\underline{F} = (\underline{F}_y)_{y\in Y}$ contains all matrices F_y^b for $b \in B, y \in Y$.

Let Γ' denote the set of all triples $(\underline{E}, \underline{F}, \psi)$, where $\underline{E} = (\underline{E}_x)_{x\in X}$, $\underline{F} = (\underline{F}_y)_{y\in Y}$ and each $\underline{E}_x, \underline{F}_y$ is a POVM, and where ψ is a unit vector. By definition the elements of $\mathcal{Q}$ are characterized by triples in the set Γ'. Now consider the following dimension-dependent set of triples:

$$\begin{aligned}\Gamma_d = \{(\underline{E}, \underline{F}, \psi) \ : \ & \underline{E} = (\underline{E}_x)_{x\in X} \text{ where each } \underline{E}_x = (E_x^a)_{a\in A} \in (\mathcal{S}_+^d)^{|A|} \text{ is a POVM,} \\ & \underline{F} = (\underline{F}_y)_{y\in Y} \text{ where each } \underline{F}_y = (F_y^b)_{b\in B} \in (\mathcal{S}_+^d)^{|B|} \text{ is a POVM} \\ & \text{and } \psi \in \mathbb{R}^{d^2}, \psi \neq 0, \|\psi\|^2 \leq 1\}\end{aligned}$$

and the union

$$\Gamma = \bigcup_{d\geq 1} \Gamma_d.$$

By Remark 5.5.2, the elements of $\mathcal{Q}$ can be equivalently described as

$$\mathcal{Q} = \left\{ P = \left(\frac{1}{\|\psi\|^2} \langle \psi, (E_a^x \otimes F_b^y)\psi \rangle \right)_{a,b,x,y} \text{ for some } (\underline{E}, \underline{F}, \psi) \in \Gamma \right\}.$$

Mančinska and Roberson [MR14], and independently Sikora and Varvitsiotis [SV15], recently showed that the set of bipartite quantum correlations $\mathcal{Q}$ can be described in terms of the completely positive semidefinite cone. They show that $\mathcal{Q}$ can be obtained as the projection of an affine section of the cone $\mathcal{CS}_+$.

5.5.3. THEOREM ([MR14, SV15]). *A bipartite conditional probability distribution $P = (P(a, b|x, y))$ with input sets X, Y and output sets A, B is quantum (i.e., $P \in \mathcal{Q}$) if and only if there exists a matrix $R \in \mathcal{CS}_+$ indexed by $(X \times A) \cup (Y \times B)$ satisfying the conditions:*

$$\sum_{a,a'\in A} R_{xa,x'a'} = 1 \text{ for all } x, x' \in X, \tag{5.22}$$

$$\sum_{b,b'\in B} R_{yb,y'b'} = 1 \text{ for all } y, y' \in Y, \tag{5.23}$$

$$\sum_{a\in A, b\in B} R_{xa,yb} = 1 \text{ for all } x \in X, y \in Y, \tag{5.24}$$

$$R_{xa,yb} = P(a, b|x, y) \text{ for all } a \in A, b \in B, x \in X, y \in Y. \tag{5.25}$$

In other words,

$$\mathcal{Q} = \pi(\mathcal{CS}_+^N \cap \mathcal{B}^N),$$

where $N = |(X \times A) \cup (Y \times B)|$, $\mathcal{B}^N$ is the affine space defined by the constraints (5.22), (5.23) and (5.24), and where π is the projection onto the subspace indexed by $(X \times A) \times (Y \times B)$ (defined by (5.25)).

Any feasible matrix R to the above program has the form $\begin{pmatrix} R_1 & P \\ P^T & R_2 \end{pmatrix}$, where R_1 is indexed by $X \times A$, R_2 is indexed by $Y \times B$ and each entry of P is such that $P_{xa,yb} = P(a,b|x,y)$.

As shown in [MR14, SV15], if the completely positive semidefinite cone is closed then the set $\mathcal{Q}$ of bipartite quantum correlations is also closed. Indeed, the constraints (5.22)-(5.24) imply that the set $\mathcal{CS}_+ \cap \mathcal{B}^t$ is bounded. Hence, if $\mathcal{CS}_+$ is closed then $\mathcal{CS}_+ \cap \mathcal{B}^t$ is compact and therefore its projection $\mathcal{Q} = \pi(\mathcal{CS}_+ \cap \mathcal{B}^t)$ is also compact.

5.5.2 Inner polyhedral hierarchy for the set $\mathcal{Q}$

We now construct an inner polyhedral hierarchy that approximates the set $\mathcal{Q}$ and covers its relative interior.

We start by introducing a discretization of the set Γ which we then use to define the polyhedral inner approximations of the set $\mathcal{Q}$.

5.5.4. DEFINITION. Given an integer $r \in \mathbb{N}$, define the sets

$$\Gamma(r) = \{(\underline{E}, \underline{F}, \psi) \in \Gamma_d \quad : \quad d \leq r \text{ and each element has rational entries with denominator at most } r\}$$

and

$$\mathcal{Q}(r) = \text{Conv}\left\{ P = \left(\frac{1}{\|\psi\|^2} \langle \psi, (E_a^x \otimes F_b^y)\psi \rangle \right)_{a,b,x,y} \text{ for } (\underline{E}, \underline{F}, \psi) \in \Gamma(r) \right\}.$$

By construction, the set $\Gamma(r)$ is finite and thus the set $\mathcal{Q}(r)$ is a polytope. Clearly, $\mathcal{Q}(r) \subseteq \mathcal{Q}(r+1) \subseteq \mathcal{Q}$ holds for every $r \in \mathbb{N}$ and therefore the polytopes $\mathcal{Q}(r)$ form a hierarchy of inner approximations for $\mathcal{Q}$. Moreover, as we see below, the union of the sets $\mathcal{Q}(r)$ covers the relative interior of $\mathcal{Q}$.

5.5.5. THEOREM. *The relative interior of the set $\mathcal{Q}$ is contained in $\bigcup_{r \geq 1} \mathcal{Q}(r)$.*

The statement of the above theorem has a similar flavor to the one of Theorem 4.5.10. In Section 4.5 we considered the set $\boldsymbol{\Delta}_n$, consisting of the n-tuples of positive semidefinite matrices such that $\text{Tr}(\sum_{i=1}^n X_i) = 1$ (see (4.4)), as a dimension-free matrix analog of the standard simplex Δ_n and we used a discretization of $\boldsymbol{\Delta}_n$ to obtain the polyhedral hierarchy. Here, in order to prove Theorem 5.5.5, we will use a different normalization: we will study the set of n-tuples $\underline{X} = (X_1, \ldots, X_n)$ forming a POVM; i.e., a collection of positive semidefinite matrices such that $\sum_{i=1}^n X_i = I$. This is another possible way to define the dimension-free matrix analog of the standard simplex Δ_n.

The rest of the section will be devoted to the proof of Theorem 5.5.5. For this, we will first prove that for any triple $(\underline{E}, \underline{F}, \psi) \in \Gamma$ we can find a triple $(\widetilde{\underline{E}}, \widetilde{\underline{F}}, \widetilde{\psi}) \in \Gamma(r)$ (for some $r \in \mathbb{N}$) which is arbitrarly close to it and then we will prove some useful geometric properties of the set $\mathcal{Q}$.

In what follows, for a n-tuple of matrices $\underline{X} = (X_1, \ldots, X_n)$ we define the norm $\|\underline{X}\| = \sqrt{\sum_{i=1}^n \|X_i\|_{\text{op}}^2}$.

5.5.6. LEMMA. *Given an n-tuple $\underline{X} = (X_1, \ldots, X_n) \in (\mathcal{S}_+^d)^n$ such that $\sum_{i=1}^n X_i = I$ and a constant $\varepsilon > 0$, there exists a n-tuple $\underline{Y} = (Y_1, \ldots, Y_n) \in (\mathcal{S}_+^d)^n$ of rational valued matrices with $\sum_{i=1}^n Y_i = I$ and such that $\|\underline{X} - \underline{Y}\| < \varepsilon$.*

PROOF: Let $\underline{X} = (X_1, \ldots, X_n)$ be a POVM, i.e., $\sum_{i=1}^n X_i = I$ and $X_i \succeq 0$ for all $i \in [n]$, and fix a constant $\varepsilon > 0$. We will prove the statement in two steps: firstly we build a n-tuple $\underline{Z}$ of positive definite matrices such that $\sum_{i=1}^n Z_i = I$ and $\|\underline{X} - \underline{Z}\|_{\text{op}} < \varepsilon/2$ and secondly we construct a n-tuple of rational valued positive semidefinite matrices $\underline{Y}$ such that $\sum_{i=1}^n Y_i = I$ and $\|\underline{Z} - \underline{Y}\|_{\text{op}} < \varepsilon/2$. Combining these two results, we then get the statement of the lemma.

Let $0 < \lambda < 1$ be a constant and define $Z_i = (1-\lambda)X_i + \lambda/nI$ for all $i \in [n]$. Then $\sum_{i=1}^n Z_i = I$, each Z_i is a positive definite matrix, and we have $\|X_i - Z_i\|_{\text{op}} = \lambda \|X_i + I/n\|_{\text{op}}$. Hence we can choose λ to be small enough such that the n-tuples $\underline{X}$ and $\underline{Z}$ are arbitrarly close.

As the set of rational positive semidefinite matrices is dense within the set of positive definite matrices, for each $i \in [n-1]$ and $0 < \gamma < 1$, we can pick a rational valued positive semidefinite matrix Y_i such that $\|Z_i - Y_i\|_{\text{op}} < \gamma$. We show that also the matrix $Y_n = I - \sum_{i=1}^{n-1} Y_i$ is positive semidefinite if we choose γ small enough. Since $Z_n = I - \sum_{i=1}^{n-1} Z_i \succ 0$, we have $\|\sum_{i=1}^{n-1} Z_i\|_{\text{op}} < 1$. Thus, $\|\sum_{i=1}^{n-1} Y_i\|_{\text{op}} - \|\sum_{i=1}^{n-1} Z_i\|_{\text{op}} \leq \|\sum_{i=1}^{n-1}(Y_i - Z_i)\|_{\text{op}}$, which implies that $\|\sum_{i=1}^{n-1} Y_i\|_{\text{op}} \leq \gamma(n-1) + \|\sum_{i=1}^{n-1} Z_i\|_{\text{op}}$. Then, for any $\gamma > 0$ small enough, in particular $\gamma < (1 - \|\sum_{i=1}^{n-1} Z_i\|_{\text{op}})/(n-1)$, we have that $\|\sum_{i=1}^{n-1} Y_i\|_{\text{op}} < 1$ and equivalently $Y_n \succ 0$. Hence we have constructed a rational valued POVM n-tuple $\underline{Y}$ which is arbitrarily close to $\underline{Z}$. □

The above lemma says that we can approximate any POVM by a rational valued one of the same dimension. Moreover, as the set of rational numbers is dense in the set of real numbers, any nonzero vector can be approximated by a rational valued one. By noticing that any element of the set Γ is composed of a collection of POVM's and a nonzero vector, we get the following corollary.

5.5.7. COROLLARY. *Given a triple $(\underline{E}, \underline{F}, \psi) \in \Gamma_d$ (for some $d \in \mathbb{N}$) and a constant $\varepsilon > 0$, there exist an integer $r \in \mathbb{N}$ and a triple $(\widetilde{\underline{E}}, \widetilde{\underline{F}}, \widetilde{\psi}) \in \Gamma(r)$ satisfying the inequality $\|(\underline{E}, \underline{F}, \psi) - (\widetilde{\underline{E}}, \widetilde{\underline{F}}, \widetilde{\psi})\| < \varepsilon$.*

We now prove some useful geometrical properties of the set $\mathcal{Q}$ of bipartite quantum correlations. As is well-known, the set $\mathcal{Q}$ is a convex bounded subset of the space $\mathbb{R}^{A\times X\times B\times Y}$, which for convenience is denoted below as $\mathcal{V}$ and can be seen as the set of all $(X \times A) \times (Y \times B)$ matrices. For $x \in X, y \in Y$, let $H_{x,y}$ denote the hyperplane:

$$H_{x,y} = \{P \in \mathcal{V} : \sum_{a\in A, b\in B} P(a,b|x,y) = 1\} = \{P \in \mathcal{V} : \langle J_{xy}, P\rangle = 1\},$$

where $J_{xy} \in \mathcal{V}$ is the matrix whose entries are equal to 1 at the positions within the block $(\{x\} \times A) \times (\{y\} \times B)$ and zero otherwise. Since any $P \in \mathcal{Q}$ is a conditional probability distribution, we have that the inclusion $\mathcal{Q} \subseteq \bigcap_{x\in X, y\in Y} H_{x,y}$ holds and that any $P \in \mathcal{Q}$ is entrywise nonnegative. The combination of these two simple observations gives that the set $\mathcal{Q}$ is bounded. We show that the hyperplanes $H_{x,y}$ are (essentially) the only ones containing $\mathcal{Q}$.

5.5.8. LEMMA. *Assume that the hyperplane $\{P \in \mathcal{V} : \langle M, P\rangle = \alpha\}$ contains the set $\mathcal{Q}$. Then there exist scalars $\lambda_{x,y}$ such that $M = \sum_{x\in X, y\in Y} \lambda_{x,y} J_{xy}$ with $\sum_{x\in X, y\in Y} \lambda_{x,y} = \alpha$.*

PROOF: Notice that the set $\mathcal{Q}$ contains the set of deterministic conditional probability distributions; i.e., the elements $P \in \mathcal{V}$ having exactly one entry equal to 1 in each of its (x, y)-blocks and all other entries equal to zero. This implies that, for any $a \in A$, $b \in B$, the entries $M_{xa,yb}$ have to be equal to a common value, say $\lambda_{x,y}$, which in turn implies that $\alpha = \sum_{x\in X, y\in Y} \lambda_{x,y}$. This concludes the proof. □

As $\mathcal{Q}$ is not full-dimensional, any linear inequality $\langle M, P\rangle \leq \alpha$ that is valid for $\mathcal{Q}$ admits several possible forms obtained by adding a linear combination of the equations $\langle J_{xy}, P\rangle = 1$ to it. We say that the inequality $\langle M, P\rangle \leq \alpha$ is *non-trivial* if $\langle M, P\rangle < \alpha$ for some $P \in \mathcal{Q}$; i.e., if $\mathcal{Q}$ is not contained in the hyperplane $\langle M, P\rangle = \alpha$. In the following lemma, we observe that any non-trivial valid linear inequality for $\mathcal{Q}$ can be assumed to have a unique representation of a special form.

5.5.9. LEMMA. *Any linear inequality which is valid for $\mathcal{Q}$ and non-trivial has, without loss of generality, the form:*

$$\langle M, P\rangle \leq 1 \text{ where } M \geq 0 \text{ and } \min_{a\in A, b\in B} M_{xa,yb} = 0 \ \forall x \in X, y \in Y. \quad (5.26)$$

Moreover, the same holds for any valid non-trivial inequality for $\mathcal{Q}(r)$ with $r \in \mathbb{N}$.

PROOF: Let $\langle M, P\rangle \leq \alpha$ be a non-trivial valid inequality for $\mathcal{Q}$. Up to adding suitable scalar multiples of the matrices J_{xy} and modifying accordingly the right hand side α, we can assume M to be nonnegative and that $\alpha > 0$. Scaling

by α we thus can assume that $\alpha = 1$. Finally, let $\mu_{x,y}$ denote the smallest of the entries $M_{xa,yb}$ for $x \in X$, $y \in Y$ and suppose that $\mu_{x,y} > 0$ for some x, y. Now, if we replace M by $M' = (M - \sum_{x,y} \mu_{x,y} J_{xy})/(1 - \sum_{x,y} \mu_{x,y})$, then we obtain a reformulation of the form $\langle M', P\rangle \leq 1$ as desired. This can be done since the inequality $\langle M, P\rangle \leq 1$ being non-trivial implies that $1 - \sum_{x,y} \mu_{x,y} > 0$. Indeed, by definition of $\mu_{x,y}$ we have that $M - \sum_{x,y} \mu_{x,y} J_{xy} \geq 0$. So, $1 = \sum_{x,y} \mu_{x,y}$ implies that for all $P \in \mathcal{Q}$ we have $\langle M, P, \rangle \geq \sum_{x,y} \mu_{x,y} \langle J_{xy}, P\rangle = 1$ and thus that $\langle M, P\rangle \leq 1$ is a trivial inequality, which is a contradiction of the assumption.

The same reasoning proves that, for any $r \in \mathbb{N}$, one may assume that any non-trivial valid linear inequality for $\mathcal{Q}(r)$ has the form (5.26). □

The following corollary is a direct consequence of Lemma 5.5.9.

5.5.10. COROLLARY. *The set $\mathcal{Q}$ can be defined as the solution set of all its valid inequalities, which can be assumed to be of the form (5.26). Moreover, an element $P \in \mathcal{Q}$ lies in the relative interior of $\mathcal{Q}$ precisely when $\langle M, P\rangle < 1$ for all the non-trivial valid inequalities for $\mathcal{Q}$.*

For the proof of Theorem 5.5.5, we will also need the following lemma.

5.5.11. LEMMA. *Assume $\langle M_r, P\rangle \leq 1$ is valid for $\mathcal{Q}(r)$ for all $r \geq 1$ and assume that the sequence $(M_r)_{r\in\mathbb{N}}$ converges to M. Then the inequality $\langle M, P\rangle \leq 1$ is valid for $\mathcal{Q}$.*

PROOF: For any fixed $d \in \mathbb{N}$, consider the function $f_d : \Gamma_d \to \mathcal{Q}$ that maps $(\underline{E}, \underline{F}, \psi)$ to $P = (\langle \psi, (\underline{E}^a_x \otimes \underline{F}^b_y)\psi\rangle / \|\psi\|^2)_{a,b,x,y}$. Notice that each f_d is a continuous function.

Take a $P \in \mathcal{Q}$, then there exist a $d \in \mathbb{N}$ and a triple $(\underline{E}, \underline{F}, \psi) \in \Gamma_d$ such that $f_d(\underline{E}, \underline{F}, \psi) = P$. As f_d is continuos, for any fixed $\varepsilon > 0$ there exists a $\eta > 0$ with the property that for all $(\widetilde{\underline{E}}, \widetilde{\underline{F}}, \widetilde{\psi}) \in \Gamma_d$ such that $\|(\underline{E}, \underline{F}, \psi) - (\widetilde{\underline{E}}, \widetilde{\underline{F}}, \widetilde{\psi})\| < \eta$ then we have $\|P - \widetilde{P}\|_{\mathrm{op}} < \varepsilon$ where $\widetilde{P} = f_d(\widetilde{\underline{E}}, \widetilde{\underline{F}}, \widetilde{\psi})$. Moreover, from Corollary 5.5.7 we know that there exists a triple $(\widetilde{\underline{E}}, \widetilde{\underline{F}}, \widetilde{\psi})$ with these properties and rational valued. Suppose that the denominator of the entries of all the matrices in $\widetilde{\underline{E}}, \widetilde{\underline{F}}$ and in the vector $\widetilde{\psi}$ is at most ℓ and let $r_0 = \max\{\ell, d\}$. Then, for all $r \geq r_0$, we have $\widetilde{P} = f_d(\widetilde{\underline{E}}, \widetilde{\underline{F}}, \widetilde{\psi}) \in \mathcal{Q}(r)$ and thus $\langle M_r, \widetilde{P}\rangle \leq 1$ holds by assumption. We get the following chain of inequalities:

$$\begin{aligned}
\langle M, P\rangle &= \langle M, P - \widetilde{P}\rangle + \langle M_r, \widetilde{P}\rangle + \langle -M_r + M, \widetilde{P}\rangle \\
&\leq 1 + \|M\|_{\mathrm{F}} \|P - \widetilde{P}\|_{\mathrm{F}} + \|\widetilde{P}\|_{\mathrm{F}} \|M - M_r\|_{\mathrm{F}} \\
&< 1 + \varepsilon \|M\|_{\mathrm{F}} + \|\widetilde{P}\|_{\mathrm{F}} \|M - M_r\|_{\mathrm{F}},
\end{aligned}$$

using the Cauchy-Schwarz inequality. As M_r tends to M, for any r large enough also $\|M - M_r\|_{\mathrm{F}} \leq \varepsilon$ holds. Hence, for any fixed $\varepsilon > 0$ there exist a $r \in \mathbb{N}$

and a $\widetilde{P} \in \mathcal{Q}(r)$ such that $\langle M, P\rangle < 1 + \varepsilon(\|M\|_{\mathrm{F}} + \|\widetilde{P}\|_{\mathrm{F}})$. As $\mathcal{Q}$ is bounded, $\|M\|_{\mathrm{F}} + \|\widetilde{P}\|_{\mathrm{F}}$ is upper bounded by an absolute constant. Therefore by letting ε tend to zero, we deduce that the inequality $\langle M, P\rangle \leq 1$ is valid for $\mathcal{Q}$. □

We can finally prove the statement of Theorem 5.5.5, namely that the relative interior of the set $\mathcal{Q}$ is contained in $\bigcup_{r\geq 1} \mathcal{Q}(r)$

PROOF OF THEOREM 5.5.5: Consider an element P_0 lying in the relative interior of $\mathcal{Q}$ and, for a contradiction, assume that it does not belong to any of the sets $\mathcal{Q}(r)$. Then, for each $r \geq 1$, there exists a non-trivial inequality valid for $\mathcal{Q}(r)$ which separates P_0 from the closed convex set $\mathcal{Q}(r)$; i.e., there exist matrices M_r and $\alpha_r > 0$ such that $\langle M_r, P\rangle \leq \alpha_r$ for all $P \in \mathcal{Q}(r)$ while $\langle M_r, P_0\rangle \geq \alpha_r$. By Lemma 5.5.9, the inequalities can be chosen of the form $\langle M_r, P\rangle \leq 1$ and satisfying (5.26). Since all the entries of M_r lie in $[0, 1]$, the sequence $(M_r)_{r\in\mathbb{N}}$ admits a converging subsequence $(M_{r_i})_{i\geq 1}$ that converges to, say, M. Moreover, $\langle M_{r_i}, P\rangle \leq 1$ for all $P \in \mathcal{Q}(r_i)$ $(i \geq 1)$ and, from Lemma 5.5.11, we deduce that the inequality $\langle M, P\rangle \leq 1$ is valid for $\mathcal{Q}$. Hence, we have $\langle M, P_0\rangle \leq 1$. At the same time, $\langle M_r, P_0\rangle \geq 1$ holds for all r by construction. Taking the limit as i tends to infinity, we obtain that $\langle M, P_0\rangle \geq 1$. Therefore the equality $\langle M, P_0\rangle = 1$ holds. However, since P_0 lies in the relative interior of $\mathcal{Q}$, by Corollary 5.5.10 the inequality $\langle M, P\rangle \leq 1$ must be trivial for $\mathcal{Q}$ and it thus defines a hyperplane that contains the set $\mathcal{Q}$. Using Lemma 5.5.8 we know that $M = \sum_{x,y} \lambda_{x,y} J_{xy}$ for some scalars $\lambda_{x,y}$. We now show that for all x, y the scalar $\lambda_{x,y}$ is equal to zero. This means that $M = 0$ and gives a contradiction.

Fix some $x \in X, y \in Y$. As $\langle M_r, P\rangle \leq 1$ is a valid non-trivial inequality for $\mathcal{Q}(r)$, by Lemma 5.5.9 it follows that each M_r has at least one zero entry within the block indexed by $(\{x\} \times A) \times (\{y\} \times B)$. Hence, there must exist a pair $(a, b) \in A \times B$ and an infinite subsequence $(M_{r_j})_{j\geq 1}$ of the sequence $(M_r)_{r\in\mathbb{N}}$ such that all M_{r_j} have a zero entry at the same position (xa, yb). Taking the limit as j tends to infinity, we obtain that the (xa, yb)-entry of M must be equal to 0. However, this entry is equal to $\lambda_{x,y}$, which implies that $\lambda_{x,y} = 0$, as desired. □

Chapter 6

Channel coding

In this chapter we study a problem from information theory: the zero-error channel coding problem, in the setting where the sender and the receiver may use quantum entanglement. The task is to transmit data reliably using a noisy channel. In Section 6.1 we review the classical problem and its graph theoretical reformulation. In Section 6.2 we describe the entangled assisted version of the problem and present the known results. At last, in Section 6.3 we explain a new entanglement-assisted protocol and use it to present an infinite family of channels for which entanglement-assisted protocols are more efficient than classical ones.

The content of this chapter is based on joint work with Jop Briët, Harry Buhrman, Monique Laurent, and Giannicola Scarpa [BBL$^+$15a].

6.1 The channel coding problem

Imagine the following scenario: a sender, Alice, wants to transmit some information to a receiver, Bob, and in order to do that they can communicate through a one-way classical noisy channel. How much information can she send to him on average, such that Bob learns Alice's message with zero probability of error? This question was first posed by Shannon in his seminal paper [Sha56] and spurred a large research area which involves information theory, combinatorics, computer science and mathematical programming. We refer the interested reader to the survey of Körner and Orlitsky [KO98] and to Lubetzky's PhD thesis [Lub07] for more recent results.

A noisy discrete *channel* $\mathcal{N}$ is fully characterized by a finite input set V, a (possibly infinite) output set W and a probability distribution $\mathcal{N}(\cdot|v)$ over W for each $v \in \mathsf{V}$. Here we only consider *memoryless* channels, where the probability distribution of the outputs depends only on the current channel input. If Alice sends an input $v \in \mathsf{V}$ through the channel, Bob then receives output

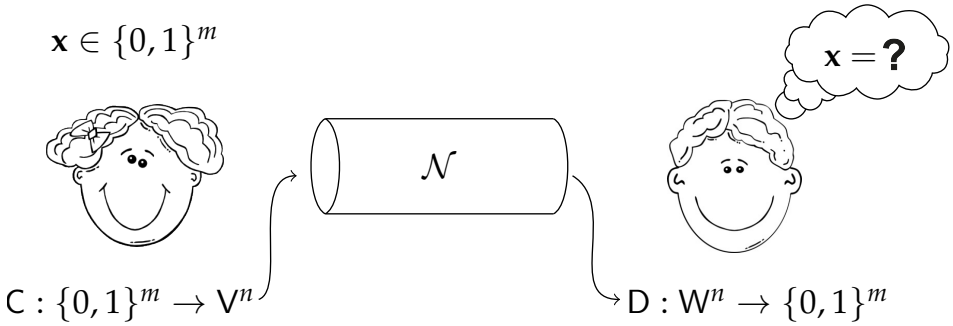

Figure 6.1: A classical channel coding protocol.

$w \in \mathsf{W}$ with probability $\mathcal{N}(w|v)$. Their goal is to transmit a binary string $\mathbf{x}$ of, say, m bits from Alice to Bob while using the channel as little as possible.

A communication protocol using a block code of length n is depicted in Figure 6.1 and works as follows. To communicate an m-bit string $\mathbf{x}$, Alice uses an encoding function $\mathsf{C} : \{0,1\}^m \to \mathsf{V}^n$ and sends $\mathsf{C}(\mathbf{x}) = (\mathsf{C}(\mathbf{x})_1, \ldots, \mathsf{C}(\mathbf{x})_n) \in \mathsf{V}^n$ through the channel by using it n times in a sequence. On the i-th instance Alice sends input $\mathsf{C}(\mathbf{x})_i$ and consequentially Bob receives output $w_i \in \mathsf{W}$ with probability $\mathcal{N}(w_i|\mathsf{C}(\mathbf{x})_i)$. He then applies a decoding function $\mathsf{D} : \mathsf{W}^n \to \{0,1\}^m$ to the entire sequence $\mathbf{w} = (w_1, \ldots, w_n)$. The coding scheme (C, D) works if $\mathsf{D}(\mathbf{w}) = \mathbf{x}$. The *communication rate* of the scheme is m/n, which is the number of bits transmitted per channel use.

Shannon [Sha48] proved that there is a computable quantity, called the *channel capacity*, which gives the maximum communication rate that can be achieved nearly error-free asymptotically in the number of uses of the channel. The great challenge in information theory has been to construct encoding and decoding schemes that achieve a communication rate close to the optimal one. Here we focus on a slightly different situation. Instead of allowing an arbitrary small probability of error, we do not allow any error in the communication. The key notion to study is then the *zero-error capacity* of a channel: the maximum number of bits that can be transmitted without error per channel use. This notion was first introduced by Shannon [Sha56].

Suppose $\mathbf{v}_1, \mathbf{v}_2 \in \mathsf{V}^n$ are two channel input sequences that can both lead to a channel output sequence $\mathbf{w}$ with positive probability. Then no decoding function D can decide with zero probability of error which one of the two sequences was used by the sender. Such sequences are called *indistinguishable* and, to be able to communicate with zero error, one has to select a subset $\mathcal{C} \subseteq \mathsf{V}^n$, called a codebook, having the property that all the elements in $\mathcal{C}$ are pairwise distinguishable. As was shown by Shannon [Sha56], this problem can be reformulated in graph-theoretic terms.

Associated to a channel $\mathcal{N}$ is its *confusability graph* $G = (\mathsf{V}, E)$ where the pair $\{u, v\}$ forms an edge if there exists a $w \in \mathsf{W}$ such that both $\mathcal{N}(w|u) > 0$ and $\mathcal{N}(w|v) > 0$. The edge set identifies indistinguishable inputs; i.e., pairs of inputs which can lead to identical channel outputs on Bob's side. Therefore, any stable set in G can be used as codebook for a single use of the channel. It is easy to see that any graph is the confusability graph of a (non-unique) channel.

To model n uses of the channel, we take the graph $G^{\boxtimes n}$ (the strong product of n copies of G), whose edges are the pairs of input sequences for Alice which Bob cannot distinguish. That is, $\mathbf{u} = (u_1, \ldots, u_n)$ and $\mathbf{v} = (v_1, \ldots, v_n)$ form an edge in $G^{\boxtimes n}$ if and only if for every $i \in [n]$ either $u_i v_i \in E(G)$ or $u_i = v_i$. Any stable set in $G^{\boxtimes n}$ is a feasible codebook for V^n and codes of block-length n allow the zero-error transmission of $\alpha(H^{\boxtimes n})$ distinct messages. The *Shannon capacity*

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

is the maximum communication rate of a zero-error coding scheme. In other words, $c(G)$ is the zero-error capacity of the channel $\mathcal{N}$.

As is well-known and easy to check, the stability number of a graph is super-multiplicative; i.e., $\alpha(G^{\boxtimes(m+n)}) \geq \alpha(G^{\boxtimes m})\alpha(G^{\boxtimes n})$. Combining this with the following lemma (commonly known as Fekete's Lemma), we obtain that in the above definition the limit exists and it coincides with the supremum: $c(G) = \sup_{n\in\mathbb{N}} \frac{1}{n} \log \alpha(G^{\boxtimes n})$.

6.1.1. LEMMA (FEKETE'S LEMMA (SEE E.G. [SCH03] THEOREM 2.2)). *Consider a sequence $\{a_m\}_{m\in\mathbb{N}}$ of real numbers with the property that $a_{n+m} \geq a_n + a_m$ for all $n, m \in \mathbb{N}$. Then the sequence $(a_n/n)_{n\in\mathbb{N}}$ has a limit which is equal to its supremum:* $\lim_{n\to\infty} a_n/n = \sup_{n\in\mathbb{N}} a_n/n$.

With a slight abuse of terminology, we call *Shannon capacity of a graph* also the parameter:

$$\Theta(G) = \lim_{n\to\infty} \sqrt[n]{\alpha(G^{\boxtimes n})} = \sup_n \sqrt[n]{\alpha(G^{\boxtimes n})},$$

which is linked to $c(G)$ by the simple identity: $c(G) = \log \Theta(G)$. By the supremum formulation of $\Theta(G)$, it is clear that $\alpha(G) \leq \Theta(G)$ holds.

The smallest graph (on the number of vertices) for which $\alpha(G) < \Theta(G)$ is the 5-cycle C_5 [Sha56]. Indeed, $\alpha(C_5) = 2$ while $\alpha(C_5^{\boxtimes 2}) = 5$, which implies $\alpha(C_5) = 2 < \sqrt{5} = \sqrt{\alpha(C_5^{\boxtimes 2})} \leq \Theta(C_5)$. It took more than 20 years to prove that Shannon's lower bound is actually tight: $\Theta(C_5) = \sqrt{5}$. Lovász [Lov79] introduced an upper bound of the Shannon capacity $\Theta(G)$, now known as the *Lovász theta number* $\vartheta(G)$, and showed that $\vartheta(C_5) = \sqrt{5}$. (The definition and some useful properties of the Lovász theta number can be found in Section 3.2.3.) In general however the Lovász theta number is not (and far from

being) a tight bound on $\Theta(G)$ and the problem of determining the Shannon capacity of a graph is wide open. Indeed, the value of $\Theta(G)$ is not known even for very small graphs, like the 7-cycle, and the Shannon capacity is not known to be decidable.

Curiously, every graph G for which the Shannon capacity is known has the property that $\Theta(G)$ is either attained with block codes of length one or of length two, or not attained at any finite length. An example of the latter is the graph $H = C_5 + K_1$, which is the disjoint union of C_5 and an isolated vertex. Then $\sqrt{5}+1 = \Theta(C_5) + \Theta(K_1) \leq \Theta(H) \leq \vartheta(H) = \vartheta(C_5) + \vartheta(K_1) = \sqrt{5}+1$, where for the last two identities we used Lemma 3.2.5 (i) and that $\vartheta(K_1) = 1$. Therefore, $\Theta(H) = \sqrt{5}+1$. One can easily observe that there is no natural number $n \geq 1$ such that $(\sqrt{5}+1)^n$ is an integer, hence the Shannon capacity of H can never be attained at any finite length. Moreover, Alon and Lubetzky [AL06] proved that if one knows an arbitrarily large, but fixed, sequence of values $\sqrt[n]{\alpha(G^{\boxtimes n})}$ this cannot be used to approximate the Shannon capacity, not even if the sequence stabilizes.

These are only some peculiar features of the zero-error capacity. The ordinary channel capacity (which allows asymptotically vanishing error) is both multiplicative and additive, while both of these plausible properties do not hold for the zero-error case. Shannon [Sha56] proved that the Shannon capacity is super-multiplicative and super-additive; that is, for any pair of graphs G and H we have that $\Theta(G \boxtimes H) \geq \Theta(G)\,\Theta(H)$ and $\Theta(G+H) \geq \Theta(G) + \Theta(H)$, where $G + H$ denotes the disjoint union of the graphs G and H. More interestingly, Haemers [Hae78] proved that the Shannon capacity is not multiplicative, i.e., there exist two distinct graphs G, H such that $\Theta(G \boxtimes H) > \Theta(G)\Theta(H)$, and Alon [Alo98] showed that it is not additive, i.e., there exist distinct graphs G, H such that $\Theta(G+H) > \Theta(G) + \Theta(H)$. A key ingredient in both of these results is an upper bound on the Shannon capacity due to Haemers [Hae78].

6.1.2. THEOREM (HAEMERS [HAE78]). *Let $G = ([n], E)$ be a graph and A be an $n \times n$ matrix over a field F having all diagonal entries equal to one and such that $A_{ij} = 0$ if and only if the pair of vertices $\{i, j\}$ is non-adjacent in G. Then, the inequality $\Theta(G) \leq \operatorname{rank}(A)$ holds.*

6.2 Entanglement-assisted channel coding

Consider again the zero-error channel coding problem, but now the parties are allowed to use entanglement. That is, Alice and Bob have quantum registers $\mathcal{A}$ and $\mathcal{B}$, respectively, that are initialized to be in some entangled state. Their most general course of action, for a single use of the channel, is as follows (see also Figure 6.2):

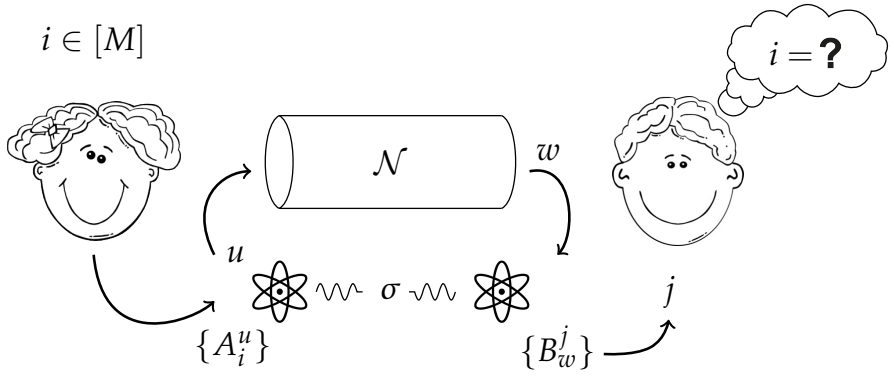

Figure 6.2: An entanglement-assisted channel coding protocol.

1. To communicate a message $i \in [M]$, Alice performs a measurement on her register $\mathcal{A}$ and uses the measurement outcome, say u, as input to the channel $\mathcal{N}$;

2. After receiving output w with probability $\mathcal{N}(w|u)$, Bob performs a measurement on his register $\mathcal{B}$ and obtains a measurement outcome $j \in [M]$.

This protocol is successful if Bob's outcome j is always equal to the original message i. Let σ denote the state in which the pair of registers $(\mathcal{A}, \mathcal{B})$ is initialized before the protocol starts. The measurement Alice performs in step (1) is given by a collection of positive semidefinite matrices $\{A_i^v\}_{v \in \mathsf{V}}$ that adds up to the identity. If Alice gets outcome $u \in \mathsf{V}$, then after step (1) Bob's register is left in a state proportional to $\rho_i^u = \mathrm{Tr}_{\mathcal{A}}((A_i^u \otimes I)\sigma)$. Note that for each $i \in [M]$ the matrices $\{\rho_i^v\}_{v \in \mathsf{V}}$ sum to Bob's reduced density matrix $\rho = \mathrm{Tr}_{\mathcal{A}}(\sigma)$. With probability $\mathcal{N}(w|u)$, Bob receives outcome $w \in \mathsf{W}$ from the channel and this allows him to reduce the list of Alice's possible inputs to a clique $\mathcal{C}$ in the confusability graph $G = (\mathsf{V}, E)$ that contains Alice's input u. For the protocol to be successful, Bob measurement $\{B_w^i\}_{i \in [M]}$ must be able to discriminate between state ρ_i^u and ρ_j^v for any $i \neq j \in [M]$ and $u, v \in \mathcal{C}$. Hence, by the Orthogonality Lemma (Lemma 2.4.1), the states ρ_i^u must be such that $\rho_i^u \rho_j^v = 0$ for any $i \neq j$ and $u \simeq v \in \mathsf{V}$. This justifies Definition 6.2.1 below, first introduced by Cubitt et al. [CLMW10], and shows that as in the classical case the protocol depends only on the confusability graph of the channel.

6.2.1. DEFINITION. [Entangled stability number and Shannon capacity] For a graph G, define $\alpha^\star(G)$ as the maximum integer $M \in \mathbb{N}$ for which there exist $d \in \mathbb{N}$ and positive semidefinite matrices ρ and $\{\rho_i^u : i \in [M], u \in V(G)\}$

in $\mathbb{C}^{d\times d}$ such that $\mathsf{Tr}(\rho) = 1$ and

$$\begin{aligned} \rho_i^u \rho_j^v &= 0 \quad \forall i \neq j \text{ and } \forall u, v \in V(G) \text{ such that } u = v \text{ or } \{u, v\} \in E(G), \\ \sum_{u \in V(G)} \rho_i^u &= \rho \quad \forall i \in [M]. \end{aligned}$$

The *entangled Shannon capacity* is defined by

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

We still have to show that if there exists a set of matrices satisfying Definition 6.2.1, then we can construct a channel coding protocol.

6.2.2. PROPOSITION. *Suppose there exists a collection of positive semidefinite matrices $\{\rho_i^u : i \in [M], u \in V(G)\}$ in $\mathbb{C}^{d\times d}$ such that (i) $\sum_{u\in V(G)} \rho_i^u = \rho$ for all $i \in [M]$, (ii) $\mathsf{Tr}(\rho) = 1$, and (iii) $\rho_i^u \rho_j^v = 0$ for all $i \neq j \in [M]$, $u \simeq v \in V(G)$. Then, there exists an M-message entanglement-assisted protocol for the channel coding problem.*

This proposition follows almost directly from the following well-known theorem (see e.g. [SR02, HJW Theorem, pp. 74], where it is attributed to Hughston, Jozsa and Wootters).

6.2.3. THEOREM (HUGHSTON–JOZSA–WOOTTERS). *Let d, n be positive integers, $p_1, \ldots, p_n \geq 0$ satisfying $p_1 + \cdots + p_n = 1$, and let $\rho_1, \ldots, \rho_n \in \mathbb{C}^{d\times d}$ be positive semidefinite matrices with trace 1. Then, there exists a state σ for a pair of registers $(\mathcal{A}, \mathcal{B})$ and a measurement on $\mathcal{A}$ consisting of a collection of positive semidefinite matrices $A^1, \ldots, A^n$ that add up to the identity, such that for each $i \in [n]$ we have $\mathsf{Tr}_{\mathcal{A}}((A^i \otimes I)\sigma) = p_i\rho_i$. Moreover, σ depends only on $p_1\rho_1 + \cdots + p_n\rho_n$.*

PROOF OF PROPOSITION 6.2.2: It suffices to find an entangled state σ and measurement $\{A_i^u\}_{u\in V(G)}$ for each message $i \in [M]$ such that $\rho_i^u = \mathsf{Tr}_{\mathcal{A}}((A_i^u \otimes I)\sigma)$. Indeed, suppose Alice gets outcome u which she uses as input to the channel and Bob receives output w. Then Bob knows that his register is in the state $\rho_j^v / \mathsf{Tr}(\rho_j^v)$ for some $j \in [M]$ and $v \in \mathcal{C}$, where $\mathcal{C}$ is a clique in G that contains Alice's outcome u. By (i) and the Orthogonality Lemma (Lemma 2.4.1), there exists a measurement $\{B^j : j \in [M]\} \cup \{B^\perp\}$ such that $\mathsf{Tr}(B^j \rho_i^u) = \delta_{ij} \mathsf{Tr}(\rho_i^u)$ for all $u \in \mathcal{C}$, which thus allows Bob to correctly identify Alice's message i.

For each $i \in [M]$, define the nonnegative numbers $p_i^u = \mathsf{Tr}(\rho_i^u)$, where $u \in V(G)$, and trace-1 matrices $\widetilde{\rho}_i^u$ given by ρ_i^u / p_i^u if p_i^u is nonzero and an arbitrary trace-1 matrix otherwise. Then, $\sum_{u\in V(G)} p_i^u = 1$ and $\sum_{u\in V(G)} p_i^u \widetilde{\rho}_i^u = \rho$. Hence, Theorem 6.2.3 gives the desired state and measurements. □

As we already mentioned, the parameter $\alpha^\star(G)$ was introduced by Cubitt et al. [CLMW10] and gives the maximum number of messages that can be sent without error using entanglement and a single use of a channel with confusability graph G. It follows that $c^\star(G)$ equals the maximum asymptotic communication rate of such a channel when entanglement can be used.

Using the operational interpretation of $\alpha^\star(G)$, one can easily see that this parameter is super-multiplicative with respect to strong graph products; i.e., $\alpha^\star(G \boxtimes G') \geq \alpha^\star(G)\,\alpha^\star(G')$. Indeed, let $\mathcal{N}$ and $\mathcal{N}'$ be channels with confusability graphs G and G' and suppose that each has an entanglement-assisted protocol that allow to transmit M and M' messages, respectively. If Alice and Bob are connected through the product channel $\mathcal{N} \otimes \mathcal{N}'$ (which has confusability graph $G \boxtimes G'$), then they can communicate at least $M \cdot M'$ different messages by running the two protocols separately. Alternatively, the super-multiplicativity can be derived from Definition 6.2.1 using simple matrix manipulations. Thus, using Fekete's Lemma (Lemma 6.1.1), in the definition of $c^\star(G)$ the limit can be replaced with the supremum.

Whenever one considers the entanglement-assisted version of a problem the first natural question to ask is whether there is something to gain: Are there channels for which the entangled Shannon capacity is strictly greater than the classical one? What about for a single use of a channel?

The second question was positively answered by the authors of [CLMW10], who found a graph G such that $\alpha^\star(G) > \alpha(G)$. The key ingredient is the following non-trivial lower bound on $\alpha^\star(G)$.

6.2.4. THEOREM (CUBITT–LEUNG–MATTHEWS–WINTER [CLMW10]). *If G is a graph with $\xi(G) \leq d$ and it has M disjoint d-cliques, then $\alpha^\star(G) \geq M$.*

While for sufficiently small, or structured, graphs the stability number can be computed, the Shannon capacity is a much harder quantity. Thus, to find a graph that exhibits a separation between entangled and classical Shannon capacity, one has to find an upper bound on $c(G)$ which potentially could be smaller than $c^\star(G)$. We have seen that the Lovász theta number is an upper bound on the classical Shannon capacity. This is however not good enough because $\vartheta(G)$ is also an upper bound for the entangled Shannon capacity and hence $c(G) \leq c^\star(G) \leq \log \vartheta(G)$ holds. Indeed, [Bei10, DSW13] proved that $\alpha^*(G) \leq \lfloor \vartheta(G) \rfloor$ and using the multiplicativity of $\vartheta(G)$ under strong graph products one can conclude that $c^*(G) \leq \log \vartheta(G)$. What turns out to be useful is the upper bound on $c(G)$ due to Haemers (Theorem 6.1.2), which is the only other known non-trivial upper bound on the Shannon capacity. Haemers' bound and the Lovász theta number are incomparable and in most cases $\vartheta(G)$ provides a better bound. Leung, Mančinska, Matthews, Ozols and Roy [LMM$^+$12] and subsequently Briët, Buhrman and Gijswijt [BBG12]

found families of graphs for which $c^\star(G) > c(G)$ by combining Haemers' bound together with the lower bound given by Theorem 6.2.4.

Such type of separation results holds only for the special case where we want the communication to succeed with zero error. As shown in [BSST02, Theorem 1], sharing entanglement does not provide any advantage in the case of vanishing error probability, where we ask the probability of error to asymptotically go to zero as the number of uses of the channel goes to infinity. However, if one restricts to a finite number of channels uses Prevedel et al. [PLM+11] experimentally showed that entanglement allows for a better error rate than the optimal classical code.

We briefly mention that Cubitt et al. [CLMW10] (see also [CLMW11]) studied also the case where the two parties can share non-signaling correlations, instead of sharing an entangled state. They showed that the non-signaling zero-error channel capacity has an elegant closed-form formula that can be computed from the description of the channel via a linear program. We will consider a generalization of this scenario in Section 7.1.2.

6.3 Separation between classical and entangled Shannon capacity

In this section we introduce a new method to lower bound the entangled Shannon capacity (Theorem 6.3.9) which allows us to strengthen the above mentioned result of Briët, Buhrman and Gijswijt [BBG12]. More specifically, we use the same family of graphs as in [BBG12], but using the new lower bound technique we can relax the conditions on which graphs of the family we can use and get an *infinite* family of graphs whose entangled capacity exceeds their Shannon capacity.

6.3.1 Quarter-orthogonal graphs

We use the following family of graphs which was also considered in [BBG12] for similar reasons.

6.3.1. DEFINITION. [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 $(k+1)/2$.

6.3.2. REMARK. The quarter-orthogonality graph is an induced subgraph of the orthogonality graph (Definition 3.3.2) containing a quarter of its vertices. This can be seen by appending a '1' to the vertices of H_k.

Some results will 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 (usually attributed to Paley [Pal33]) states that for every ℓ that is a multiple of 4 there exists an $\ell \times \ell$ Hadamard matrix. Although this conjecture is still open, many infinite families of Hadamard matrices are known. In Section 8.2.3, we will use a family constructed by Xia and Liu [XL91] (see for example [Xia96, Che97, XSX06] for closely related constructions).

6.3.3. THEOREM (XIA–LIU [XL91]). *Let q be a prime power with $q \equiv 1 \bmod 4$. Then, there exists a Hadamard matrix of size $4q^2$.*

In [BBG12], it is shown that for some values of k, the entangled Shannon capacity of H_k can be strictly larger than the classical one.

6.3.4. THEOREM (BRIËT–BUHRMAN–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] and that it is not known if Hadamard matrices of size $4p$ exist for infinitely many primes p. Theorem 6.3.4 requires the existence of Hadamard matrices because to lower bound $c^\star(H_k)$ Theorem 6.2.4 is used. Moreover, k has to be 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 6.3.4 and our result does not rely anymore on the existence of a Hadamard matrix. This is obtained by introducing a new lower bound technique for $c^\star(G)$. We show the existence of an infinite family of quarter-orthogonality graphs whose entangled capacity exceeds their Shannon capacity.

6.3.5. THEOREM. *For every odd integer $k \geq 11$, we have*

$$c^\star(H_k) \geq (k-1)\left(1 - \frac{2\log(k+1)}{k-3}\right). \tag{6.1}$$

Moreover, if $k = 4p^\ell - 1$ where p is an odd prime and $\ell \in \mathbb{N}$, then

$$c(H_k) \leq 0.846\, k. \tag{6.2}$$

We prove (6.1) in Section 6.3.2, using a new technique based on quantum remote state preparation [BDVS+01]. In Section 6.3.3 we show the bound (6.2) by combining an instance of the linear algebra method due to Alon [Alo98] with a construction of certain low-degree polynomials over a finite field for a low-degree representation of the OR-function due to Barrington, Beigel and Rudich [BBR94].

Before doing that we record some useful results regarding the graph H_k.

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

PROOF: The statement 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. It is easy to see that $|W| = 2^{(k-3)/2}$ and that W is a stable set since it does not contain pairs of strings at Hamming distance $(k+1)/2$. □

6.3.7. PROPOSITION (BRIËT–BUHRMAN–GIJSWIJT [BBG12]). *Let k be an odd integer such that there exists a Hadamard matrix of size $k+1$. Then, $\omega(H_k) \geq k+1$.*

PROOF: We include a proof for completeness. Let A be a size $k+1$ Hadamard matrix. Without loss of generality, we may assume that the first row and the first column of A contain only '+1' elements. Indeed, the property of being a Hadamard matrix is preserved under changing the sign all entries of a row (or column). The first row is orthogonal to each of the last k rows, these therefore contain $(k+1)/2$ entries equal to '-1'. Take now the $k+1$ vectors in $\{-1,1\}^k$ obtained from the rows of A by deleting the first element. One can easily check that these form a clique in H_k and thus we can conclude. □

Recall that $\xi'(G)$ is the minimum d for which there exists a d-dimensional orthogonal representation f of G such that all entries of each vector $f(u)$ have absolute value one.

6.3.8. LEMMA. *For every odd positive integer k, we have $\xi'(H_k) \leq k+1$.*

PROOF: This was already observed in Remark 6.3.2. Indeed, the map f from $V(H_k)$ to $\{-1,1\}^{k+1}$ defined by $f(u) = (u,1)^\mathsf{T}$ has all the needed properties. □

6.3.2 Lower bound on the entangled Shannon capacity

Here we prove the bound (6.1). The idea is to show that with $t+1$ sequential uses of a channel with confusability graph G, Alice can perfectly transmit one out of $|V(G)|^t$ messages to Bob, provided that $t \leq \log \alpha(G) / \log \xi'(G)$. This is achieved using the remote state preparation protocol (Section 2.4.4).

6.3.9. THEOREM. *For a graph G and integer $t \geq 1$ such that $t \leq \log \alpha(G) / \log \xi'(G)$, we have*

$$c^\star(G) \geq \frac{t}{t+1} \log |V(G)|.$$

PROOF: Let $\mathcal{N}$ be a channel with confusability graph G. Let $d = \xi'(G)$ and let f be a d-dimensional orthogonal representation of G such that its vectors have entries of modulus one. For each $v \in V(G)$ define $\rho_v = f(v)f(v)^*/d$. Let t be a positive integer such that

$$t \leq \frac{\log \alpha(G)}{\log d}. \tag{6.3}$$

It suffices to find an entanglement-assisted protocol for the noiseless transmission of $|V(G)|^t$ distinct messages based on at most $t+1$ uses of the channel $\mathcal{N}$. Indeed, this then implies that $\alpha^\star(G^{\boxtimes(t+1)}) \geq |V(G)|^t$ and therefore

$$c^\star(G) \geq \frac{\log \alpha^\star(G^{\boxtimes(t+1)})}{t+1} \geq \frac{t \log |V(G)|}{t+1}$$

as claimed. To this end, consider the following four-step protocol for transmitting a sequence $\mathbf{v} = (v_1, \ldots, v_t) \in V(G)^t$. First, Alice prepares d-dimensional quantum registers $\mathcal{A}_1, \ldots, \mathcal{A}_t$ to be in the states $\rho_{v_1}, \ldots, \rho_{v_t}$, respectively. Second, Alice sends the sequence $\mathbf{v}$ 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 v_i belongs to a particular clique in G. Third, Alice and Bob execute the remote state preparation scheme described in Section 2.4.4 t times in a row, once for each of the states $\rho_{v_1}, \ldots, \rho_{v_t}$ separately. (Recall that $\rho_{v_i} = f(v_i)f(v_i)^*/d$ where $f(v_i) \in \mathbb{C}^d$ has norm $\sqrt{d}$, so in the notation of Section 2.4.4 we are setting $u = f(v_i)/\sqrt{d}$.) This requires that Alice communicates a total of $t\lceil \log d \rceil$ bits to Bob. To do so, Alice uses the channel one more time to send, without error, the bits required to perform the remote state preparation. This can be done if $\log \alpha(G) \geq t\lceil \log d \rceil$, which holds by our assumed bound (6.3). At this point Bob's quantum registers $\mathcal{B}_1, \ldots, \mathcal{B}_t$ are in states $\rho_{v_1}, \ldots, \rho_{v_t}$. Moreover, for each v_i Bob knows a clique in the graph G that contains v_i and by construction elements of a clique have pairwise orthogonal states. In the last step, for every $i \in [t]$, Bob can perform a measurement on register $\mathcal{B}_i$ such that he gets outcome v_i with probability one (due to Lemma 2.4.1). Hence, Bob can recover any sequence $\mathbf{v} = (v_1, \ldots, v_t) \in V(G)^t$ with zero probability of error, completing the proof. □

The bound (6.1) can now be derived as a simple corollary.

6.3.10. COROLLARY. *For every odd integer $k \geq 11$, we have*

$$c^\star(H_k) \geq (k-1)\left(1 - \frac{2\log(k+1)}{k-3}\right).$$

PROOF: By Lemmas 6.3.6 and 6.3.8, we have that $\log \alpha(H_k) \geq (k-3)/2$ and that $\log \xi'(H_k) \leq \log(k+1)$. Therefore, for any $k \geq 11$ we can choose t to be equal to $\lfloor (k-3)/(2\log(k+1)) \rfloor \leq \lfloor \log \alpha(H_k)/\log \xi'(H_k) \rfloor$. (We require $k \geq 11$ to ensure that $t \geq 1$.) Applying Lemma 6.3.9 combined with the fact that $|V(H_k)| = 2^{(k-1)}$, we obtain

$$c^\star(H_k) \geq \frac{t(k-1)}{t+1} \geq (k-1)\left(1 - \frac{2\log(k+1)}{k-3}\right)$$

which gives the result. □

6.3.3 Upper bound on the Shannon capacity

We prove the bound (6.2) by using the following upper bound on the stability number of the graphs $H_k^{\boxtimes m}$ for certain values of k.

6.3.11. LEMMA. *Let p be an odd prime, $\ell \in \mathbb{N}$ and set $k = 4p^\ell - 1$. Then, for every $m \in \mathbb{N}$, we have*

$$\alpha(H_k^{\boxtimes m}) \leq \left(\binom{k}{0} + \binom{k}{1} + \cdots + \binom{k}{p^\ell - 1}\right)^m \leq 2^{H(3/11)km} < 2^{0.846km} \quad (6.4)$$

where $H(t) = -t\log t - (1-t)\log(1-t)$ for $t \in [0,1]$ 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 itself is inspired by Haemers' bound (Theorem 6.1.2). We recall this method below for completeness. Let G be a graph and $\mathbb{F}$ be a field. Let $\mathcal{F} \subseteq \mathbb{F}[x_1, \ldots, x_k]$ 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

$$f_u(c_u) \neq 0 \quad \forall u \in V(G) \text{ and } f_u(c_v) = 0 \quad \forall u \neq v \in V(G) \text{ with } \{u,v\} \notin E(G).$$

6.3.12. LEMMA (ALON [ALO98]). *Let G be a graph, $\mathbb{F}$ be a field, $k \in \mathbb{N}$ and $\mathcal{F}$ be a subspace of $\mathbb{F}[x_1, \ldots, x_k]$. If $\big((f_u, c_u)\big)_{u \in V(G)} \subseteq \mathcal{F} \times \mathbb{F}^k$ represents G, then we have $\alpha(G^{\boxtimes n}) \leq \dim(\mathcal{F})^n$ for all $n \in \mathbb{N}$.*

We get a representation for the graph H_k, for $k = 4p^\ell - 1$, from the following result of Barrington, Beigel and Rudich [BBR94] (see [Yek12, Lemma 5.6] for the statement as it appears below).

6.3.13. LEMMA (BARRINGTON–BEIGEL–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*

$$f(c) \equiv \begin{cases} 1 & \text{if } c_1 + c_2 + \cdots c_k \equiv w \bmod p^\ell \\ 0 & \text{otherwise.} \end{cases}$$

With this we can now prove Lemma 6.3.11.

PROOF OF LEMMA 6.3.11: 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, as 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 6.3.1 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 non-adjacent in H_k, then $|u \oplus v| \notin \{0, 2p^\ell\}$ and thus $|u \oplus v| \not\equiv 0 \mod p^\ell$.

Let $f \in \mathbb{Z}_p[x_1, \ldots, x_k]$ 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 6.3.13 (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 $f_u \in \mathbb{Z}_p[x_1, \ldots, x_k]$ 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 $u = (1, 1, 0, \ldots, 0)$, then $f_u(x_1, \ldots, x_k) = f(1 - x_1, 1 - x_2, x_3, \ldots, x_k)$. 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 $p^\ell - 1$. Using Lemma 6.3.12 we obtain that

$$\alpha(H_k^{\boxtimes m}) \leq \left(\binom{k}{0} + \binom{k}{1} + \cdots + \binom{k}{p^\ell - 1} \right)^m. \tag{6.5}$$

We now use the well-known fact that for $q, k \in \mathbb{N}$ with $1 < q < k/2$, the sum $\binom{k}{0} + \ldots + \binom{k}{q-1} \leq 2^{kH(q/k)}$. Since $p^\ell/(4p^\ell - 1) \leq 3/11$, we deduce that the right hand side in (6.5) can be upper bounded by $2^{H(3/11)km} < 2^{0.846km}$. □

The upper bound (6.2) on the Shannon capacity of H_k stated in Theorem 6.3.5 is an easy corollary of Lemma 6.3.11.

6.3.14. COROLLARY. *Let p be an odd prime, $\ell \in \mathbb{N}$ and set $k = 4p^\ell - 1$. Then, $c(H_k) \leq 0.846k$.*

PROOF: By taking the logarithm, dividing by m and taking the limit m goes to infinity on both sides of (6.4) we get the result. □

Chapter 7

Multiparty channel coding

We study which effects entanglement can have on the performance of two generalizations of the zero-error channel coding problem. In the first task one sender wants to communicate a common message to multiple receivers (Section 7.1). For this we show that entanglement-assisted strategies might provide an advantage only if the number of receivers is below a certain threshold. In the second task multiple collaborating senders want to transmit a message to one receiver (Section 7.2). In Theorem 7.2.6 we show that entanglement allows for a peculiar amplification of information which cannot happen classically.

The content of this chapter is based on joint work with Giannicola Scarpa and Christian Schaffner [PSS15].

7.1 Multiple receivers

Suppose there are ℓ receivers that want to decode a common message sent by a single sender, as for example in TV broadcasting. This is known as the *compound channel* model. We focus on the zero-error case, where each receiver perfectly learns the original message, and on the scenario where the sender is connected to each of the receivers through identical classical channels. Our results are twofold. If the block length of the code is fixed, entanglement may be helpful only up to a certain number of receivers (Theorem 7.1.2). On the other hand, for any constant number of receivers, we can build a compound channel (based on Section 6.3) for which there is an entanglement-assisted protocol that is more efficient than any classical one (Corollary 7.1.10).

7.1.1 The compound channel problem

Consider a family of channels $\boldsymbol{\mathcal{N}} = \{\mathcal{N}_1, \ldots, \mathcal{N}_\ell\}$ with the same input set V where $\mathcal{N}_k$ connects the sender with the k-th receiver. A common input $v \in \mathsf{V}$

is sent to all the receivers and the k-th receiver gets the output w_k according to the distribution $\mathcal{N}_k(w_k|v)$. The goal is for each receiver to retrieve the original input v with zero probability of error. As for the two parties case, this problem can be treated from a graph-theoretical perspective associating to each channel $\mathcal{N}_k$ its confusability graph $G_k = (\mathsf{V}, E_k)$. As the input set is in common, the family of graphs $\mathcal{G} = \{G_1, \dots G_\ell\}$ share the same vertex set V. Suppose $\mathcal{C} \subseteq \mathsf{V}$ is a stable set in each graph $G_k \in \mathcal{G}$ then, for each of the receivers, $\mathcal{C}$ forms a set of non-confusable inputs and it therefore can be used for zero-error communication. We define $\alpha(\mathcal{G}, n)$ to be the maximum cardinality of a set $\mathcal{C} \subseteq \mathsf{V}^n$ such that $\mathcal{C}$ is a stable set in $G_k^{\boxtimes n}$ for each $G_k \in \mathcal{G}$. Thus, $\alpha(\mathcal{G}, n)$ is the maximum number of messages that can be transmitted perfectly to each of the receivers using codes of block length n. Observing that $\alpha(\mathcal{G}, \cdot)$ is super-multiplicative (i.e., $\alpha(\mathcal{G}, m+n) \geq \alpha(\mathcal{G}, m)\alpha(\mathcal{G}, n)$ for every $m, n \in \mathbb{N}$), the Shannon capacity of a family of graphs $\mathcal{G}$ is well-defined as $c(\mathcal{G}) = \lim_{n\to\infty} \frac{1}{n} \log \alpha(\mathcal{G}, n)$.

This parameter was introduced by Cohen, Körner and Simonyi [CKS90] as a generalization of the zero-error Shannon capacity (see [Sin09] and references therein for recent results). Determining the Shannon capacity of a family of graphs seems a hopeless endeavor since already computing the Shannon capacity of a single graph is not known to be decidable. However, there are positive results for a slightly different task. Suppose that one knows the capacity of every graph in the family $\mathcal{G}$, then Gargano, Körner and Vaccaro [GKV94] proved that it is possible to determine the capacity of the whole family.

Here we focus on the particular instance where all the receivers are connected to the sender through the same channel $\mathcal{N}$ with confusability graph G. Note that this leads to the trivial situation: $\alpha(\mathcal{G}, 1) = \alpha(G)$ and, for any $n \in \mathbb{N}$, $\alpha(\mathcal{G}, n) = \alpha(G^{\boxtimes n})$. Therefore, we have that $c(\mathcal{G}) = c(G)$. As all the elements of the family $\mathcal{G}$ are equal, we introduce the following new notation: let ℓ be the cardinality of the family $\mathcal{G}$, then $\alpha_{1,\ell}(G) = \alpha(\mathcal{G}, 1)$ and $c_{1,\ell}(G) = c(\mathcal{G})$.

Consider now the scenario where the sender shares a single entangled state with all the receivers. That is, Alice has a quantum register $\mathcal{A}$, each Bob has a quantum register $\mathcal{B}_k$ and the tuple $(\mathcal{A}, \mathcal{B}_1, \dots, \mathcal{B}_\ell)$ is in some entangled state. The entanglement-assisted version of the compound channel coding scheme, with a single use of the channels, is as follows (see also Figure 7.1):

1. To communicate a message $i \in [M]$, Alice performs a measurement on her register $\mathcal{A}$ and uses the measurement outcome, say, u as input to each channel $\mathcal{N}_k$;

2. After the k-th Bob receives output w_k with probability $\mathcal{N}_k(w_k|u)$, he performs a measurement on his register $\mathcal{B}_k$ and obtains a measurement outcome $j_k \in [M]$.

The protocol works if j_k is equal to the original message i for every $k \in [\ell]$; i.e., if every Bob is able to perfectly learn Alice's message. As in the single-receiver

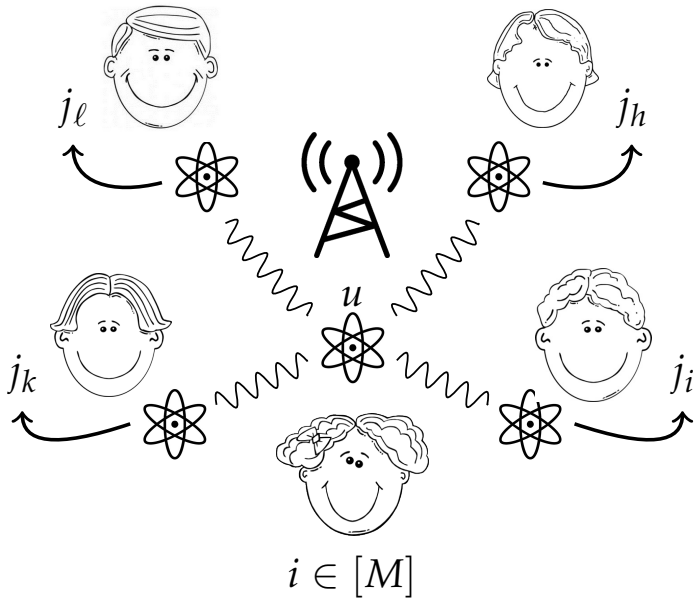

Figure 7.1: An entanglement-assisted compound channel coding protocol.

case, the protocol depends only on the confusability graph of the channels and, using the same line of reasoning as in Section 6.2, we can define the following quantities.

7.1.1. DEFINITION. [Entangled compound stability number and capacity] For a graph G, the *entangled compound stability number* with ℓ receivers $\alpha^\star_{1,\ell}(G)$ is defined as the maximum $M \in \mathbb{N}$ such that there exist $d \in \mathbb{N}$ and positive semidefinite matrices ρ and $\{\rho_i^u, i \in [M],\, u \in V(G)\}$ in $(\mathbb{C}^{d\times d})^{\otimes \ell}$, denoted as $\mathcal{B}_1 \otimes \mathcal{B}_2 \otimes \cdots \otimes \mathcal{B}_\ell$, such that $\mathsf{Tr}(\rho) = 1$ and

$$\begin{aligned} &\mathsf{Tr}_{\mathcal{B}_{-k}}(\rho_i^u)\ \mathsf{Tr}_{\mathcal{B}_{-k}}(\rho_j^v) = 0 \quad \forall k \in [\ell], \forall i \neq j,\, \forall u \simeq v \in V(G), \\ &\sum_{u\in V(G)} \rho_i^u = \rho \quad \forall i \in [M], \end{aligned}$$

where the operator $\mathsf{Tr}_{\mathcal{B}_{-k}}$ denotes the partial trace over all the subspaces but the k-th one; i.e., $\mathsf{Tr}_{\mathcal{B}_{-k}}(\rho_i^u) = \mathsf{Tr}_{\mathcal{B}_1,\ldots,\mathcal{B}_{k-1},\mathcal{B}_{k+1},\ldots,\mathcal{B}_\ell}(\rho_i^u)$.

The *entangled compound Shannon capacity* with one sender and ℓ receivers is

$$c^\star_{1,\ell}(G) = \lim_{n\to\infty} \frac{1}{n} \log \alpha^\star_{1,\ell}(G^{\boxtimes n}).$$

Operationally, $\alpha^\star_{1,\ell}(G)$ is the maximum number of messages that Alice can perfectly communicate to all the Bobs through identical channels with confusabil-

ity graph G and an entangled state. This interpretation allows us to immediately see that $\alpha^\star_{1,\ell}(G)$ is super-multiplicative with respect to strong graph product; i.e., $\alpha^\star_{1,\ell}(G \boxtimes G') \geq \alpha^\star_{1,\ell}(G)\,\alpha^\star_{1,\ell}(G')$. Indeed, suppose that two channels $\mathcal{N}$ and $\mathcal{N}'$ allow an entanglement-assisted compound channel protocol with ℓ receivers that communicates M and M' messages, respectively. Then for the combined channel $\mathcal{N} \otimes \mathcal{N}'$, with confusability graph $G \boxtimes G'$, the sender can transmit $M \cdot M'$ messages by running subsequently the two protocols. Combining this observation with Fekete's Lemma (Lemma 6.1.1) we obtain that $c^\star_{1,\ell}(G)$ is a well-defined quantity.

7.1.2 Entanglement for fixed block length

Here we prove a limitation of entanglement-assisted communication for the compound channel problem. For any fixed channel and number of channel uses, entanglement cannot be advantageous if the number of receivers is above a certain threshold. Let $\theta_e(G)$ denote the *edge-clique cover number* of G; i.e., the smallest number of cliques needed to cover all the edges of the graph, and $\theta'_e(G)$ denote the edge-clique cover number plus the number of isolated vertices of G. The goal of this section is to prove the following theorem.

7.1.2. THEOREM. *For any graph G, if $\ell \geq \theta'_e(G)$ then $\alpha^\star_{1,\ell}(G) = \alpha(G)$.*

This statement follows directly from Theorem 7.1.5 below, where we prove an analogous result for the situation when the players can use arbitrary non-signaling correlations. Recall that a probability distribution is non-signaling if the marginal distribution of the output of each subset of parties depends only on the corresponding inputs.

7.1.3. DEFINITION. An n-partite probability distribution $P(a_1, \ldots, a_n|x_1, \ldots, x_n)$ is called *non-signaling* if for all outputs $a_1, \ldots, a_n$ and all inputs $x_1, \ldots, x_n$ the marginal distribution for each subset of parties $I = \{i_1, i_2 \ldots, i_k\} \subseteq [n]$ only depends on the corresponding inputs

$$P(a_{i_1}, a_{i_2}, \ldots, a_{i_k}|x_1, x_2, \ldots, x_n) = P(a_{i_1}, a_{i_2}, \ldots, a_{i_k}|x_{i_1}, x_{i_2}, \ldots, x_{i_k}).$$

Since any entanglement-assisted strategy is also non-signaling, the amount of information that can be communicated using a non-signaling strategy is always at least as much as it can be done using entanglement.

We have seen that to study the zero-error channel coding problem we can restrict our attention to the properties of the confusability graph of the channel. However, many different channels have the same confusability graph and, unlike the classical and entanglement-assisted capacities, the non-signaling capacity depends on the particular channel. For our purposes we are interested

in the particular channel that minimizes the number of outputs while keeping the same confusability graph. Notice that every output of a channel defines a clique or an isolated vertex in the confusability graph. Therefore, we fix an edge-clique covering of the confusability graph G of minimum cardinality $\theta_e(G)$ (which might not be unique), we add the isolated vertices to obtain a clique covering of cardinality $\theta'_e(G)$, and we consider the channel that has $\theta'_e(G)$ outputs. In other words, we take a channel which has one output per element of the edge-clique covering plus one output per isolated vertex.

We mentioned that the non-signaling version of the two-party zero-error channel coding problem was studied in [CLMW10]. We consider a generalization of this and study the compound channel with ℓ receivers scenario, where Alice and the Bobs share an $(\ell+1)$-partite non-signaling probability distribution. Alice is connected to each Bob through a fixed channel, which has confusability graph G and is constructed as above. To communicate message $i \in [M]$, Alice inputs i to the non-signaling distribution $P(u, j_1, \ldots, j_\ell | i, c_1, \ldots, c_\ell)$ and uses her output $u \in V(G)$ as input of the channel. Each Bob gets a $c_k \subseteq V(G)$ as channel output, where $u \in c_k$ and c_k is either a clique or an isolated vertex of G. If c_k is used as input to the non-signaling distribution it gives j_k as output. The protocol works if every single Bob learns i with zero probability of error, that is $i = j_k$ for every $k \in [\ell]$.

7.1.4. DEFINITION. [Non-signaling compound stability number] For a graph G, the *non-signaling compound stability number* with ℓ receivers $\alpha^{\mathrm{ns}}_{1,\ell}(G)$ is the maximum $M \in \mathbb{N}$ such that there exists a non-signaling distribution

$$P(u, j_1, \ldots, j_\ell | i, c_1, \ldots, c_\ell)$$

between Alice and $\mathrm{Bob}_1, \ldots, \mathrm{Bob}_\ell$, where $i \in [M]$ and $c_k \subseteq V(G)$ are elements of a fixed clique covering of cardinality $\theta'_e(G)$. Additionally, for all $i \in [M]$ we require that: If vertex u is contained in c_k for all $k \in [\ell]$ and there exists a $k' \in [\ell]$ such that $i \neq j_{k'}$, then $P(u, j_1, \ldots, j_\ell | i, c_1, \ldots, c_\ell) = 0$.

The last condition imposes the perfect correctness of the protocol. Every Bob must output the correct message i upon receiving as channel output a c_k which is compatible with Alice's channel input u (i.e., if $u \in c_k$ for every $k \in [\ell]$).

Since every entanglement-assisted strategy is also non-signaling, for every graph G and $\ell \in \mathbb{N}$, we have: $\alpha^{\mathrm{ns}}_{1,\ell}(G) \geq \alpha^{\star}_{1,\ell}(G)$. Moreover, we now prove that for ℓ large enough equality holds.

7.1.5. THEOREM. *For any graph G, if $\ell \geq \theta'_e(G)$ then $\alpha^{\mathrm{ns}}_{1,\ell}(G) = \alpha^{\star}_{1,\ell}(G) = \alpha(G)$.*

To prove this, we use a property known as monogamy of non-signaling distributions as derived by Masanes, Acin and Gisin [MAG06]. For convenience, we reproduce the definition and result here.

7.1.6. DEFINITION. A non-signaling probability distribution $P(a,b|x,y)$ is called *ℓ-shareable with respect to Bob*, if there exists an $(\ell+1)$-partite non-signaling probability distribution $Q(a,b_1,\ldots,b_\ell|x,y_1,\ldots,y_\ell)$ such that:

1. For any permutation $\pi \in \Pi(\ell)$, we have that
$$Q(a,b_{\pi(1)},\ldots,b_{\pi(\ell)}|x,y_{\pi(1)},\ldots,y_{\pi(\ell)}) = Q(a,b_1,\ldots,b_\ell|x,y_1,\ldots,y_\ell)\,.$$
2. It holds that
$$\sum_{b_2,\ldots,b_\ell} Q(a,b_1,\ldots,b_\ell|x,y_1,\ldots,y_\ell) = P(a,b_1|x,y_1)\,.$$

Note that if both conditions hold, we have that for all $k \in [\ell]$
$$\sum_{b_1,\ldots,b_{k-1},b_{k+1},\ldots,b_\ell} Q(a,b_1,\ldots,b_\ell|x,y_1,\ldots,y_\ell) = P(a,b_k|x,y_k)\,. \tag{7.1}$$

7.1.7. THEOREM (MASANES–ACIN–GISIN [MAG06]). *Let Y be the set of different values for the input y and suppose $\ell \geq |Y|$. If $P(a,b|x,y)$ is a non-signaling distribution which is ℓ-shareable with respect to Bob, then $P(a,b|x,y)$ admits a local hidden variable model. Formally, there exists a distribution $Q(\lambda)$ over the hidden variable λ as well as local strategies $A(a|x,\lambda)$ for Alice and $B(b|y,\lambda)$ for Bob such that $P(a,b|x,y) = \sum_\lambda Q(\lambda)A(a|x,\lambda)B(b|y,\lambda)$.*

PROOF: Assume without loss of generality that $Y = \{1,2,\ldots,|Y|\}$. The idea of the proof is to ask all possible questions $y = 1,2,\ldots,|Y|$ to $|Y|$ different Bobs (which is possible because $\ell \geq |Y|$) and use their answers $b_1,\ldots,b_{|Y|}$ to these questions as hidden variable λ.

Assume for now that $\ell = |Y|$. Let us fix the questions to the ℓ Bobs as $y_1 = 1, y_2 = 2, \ldots, y_\ell = \ell$ and abbreviate this event with $\mathcal{E}$. We can then write
$$\begin{aligned} P(a,b|x,y) &\overset{(7.1)}{=} \sum_{\substack{b_1,\ldots,b_\ell \\ b_y = b}} Q(a,b_1,\ldots,b_\ell|x,\mathcal{E}) \\ &= \sum_{b_1,\ldots,b_\ell} Q(b_1,\ldots,b_\ell|x,\mathcal{E}) \cdot Q(a|b_1,\ldots,b_\ell,x,\mathcal{E}) \cdot \delta_{b,b_y}\,. \end{aligned}$$

Due to non-signaling, $Q(b_1,\ldots,b_\ell|x,\mathcal{E}) = Q(b_1,\ldots,b_\ell|\mathcal{E}) = Q(\lambda|\mathcal{E})$. The conditional distribution $Q(a|b_1,\ldots,b_\ell,x,\mathcal{E})$ defines Alice's strategy $A(a|\lambda,x,\mathcal{E})$. Bob's strategy $B(b|\lambda,y,\mathcal{E})$ is defined by giving the answer $b = b_y$ of the y-th Bob. In summary, we obtain a local-hidden-variable representation of P:
$$P(a,b|x,y) = \sum_\lambda Q(\lambda|\mathcal{E}) \cdot A(a|\lambda,x,\mathcal{E}) \cdot B(b|\lambda,y,\mathcal{E})\,.$$

In case that $\ell > |Y|$, we observe that ℓ-shareability of $P(a,b|x,y)$ implies $|Y|$-shareability. Hence, the above proof applies. □

PROOF OF THEOREM 7.1.5: Let $P(u, j_1, \ldots, j_\ell | i, c_1, \ldots, c_\ell)$ be the optimal non-signaling probability distribution achieving $\alpha^{\text{ns}}_{1,\ell(G)}$. We define the following distribution

$$Q(u, j_1, \ldots, j_\ell | i, c_1, \ldots, c_\ell) = \sum_{\pi \in \Pi(\ell)} \frac{1}{|\Pi(\ell)|} P(u, j_{\pi(1)}, \ldots, j_{\pi(\ell)} | i, c_{\pi(1)}, \ldots, c_{\pi(\ell)})$$

which clearly fulfills the first condition of Definition 7.1.6. By assumption, we have that for all $i \in [M]$, $P(u, j_1, \ldots, j_\ell | i, c_1, \ldots, c_\ell) = 0$ whenever $u \in c_k$ for all $k \in [\ell]$ and there is a k' such that $i \neq j_{k'}$. As this condition holds for each pair of Alice and Bob_k individually, it is invariant under permutations of Bobs. Therefore, the same condition also holds for Q. Since any convex combination of non-signaling distributions is also non-signaling, it follows that Q can also be used to achieve $\alpha^{\text{ns}}_{1,\ell}(G)$. We now focus on the marginal distribution $Q(u, j_1 | i, c_1)$ between Alice and the first Bob. This distribution is non-signaling and ℓ-shareable by construction where $\ell \geq \theta'_e(G)$; i.e., ℓ is greater or equal to the number of outputs of the specific channel we consider. By Theorem 7.1.7, Q admits a local-hidden-variable model. In other words, Alice and the first Bob can achieve the distribution by using classical shared randomness. However, as we are considering the zero-error scenario, shared randomness does not improve over the deterministic classical setting. Therefore Alice and the first Bob are unable to transmit more than $\alpha(G)$ messages over the channel, showing that $\alpha^{\text{ns}}_{1,\ell}(G) \leq \alpha(G)$. The claim of the theorem then follows by combining this inequality with $\alpha(G) = \alpha_{1,\ell}(G) \leq \alpha^{\star}_{1,\ell}(G) \leq \alpha^{\text{ns}}_{1,\ell}(G)$. □

7.1.3 Entanglement can improve the capacity for a fixed number of receivers

For any fixed number of receivers we construct a channel such that its entangled compound Shannon capacity is strictly bigger than the classical one. To this end, we will use the quarter-orthogonality graph H_k introduced in Definition 6.3.1 and prove a generalization to the compound channel setting of the lower bound on the entanglement-assisted capacity obtained in Theorem 6.3.9.

7.1.8. THEOREM. *For a graph G, a natural number ℓ and integer $t \geq 1$ such that $t \leq \log \alpha(G) / \log \xi'(G)$, we have*

$$c^{\star}_{1,\ell}(G) \geq \frac{t}{t+\ell} \log |V(G)|.$$

PROOF: In Theorem 6.3.9 we proved the special case when ℓ is equal to one. Let $\mathcal{N}$ be a channel with confusability graph G, $d = \xi'(G)$ and let f be a d-dimensional orthogonal representation of G such that its vectors have entries

of modulus one. For each $v \in V(G)$ define $\rho_v = f(v)f(v)^*/d$. Moreover, suppose $t \in \mathbb{N}$ is such that $t \leq \log \alpha(G)/\log d$.

Consider the following protocol for a compound channel with ℓ receivers. Let the entangled state be such that Alice shares with each individual Bob, say the k-th, an independent tuple of registers $(\mathcal{A}_1^k, \ldots, \mathcal{A}_t^k, \mathcal{B}_1^k, \ldots, \mathcal{B}_t^k)$, where $(\mathcal{A}_1^k, \ldots, \mathcal{A}_t^k)$ are Alice's registers and $(\mathcal{B}_1^k, \ldots, \mathcal{B}_t^k)$ are the ones of Bob_k. Suppose that Alice wants to transmit the sequence $\mathbf{v} = (v_1, \ldots, v_t) \in V(G)^t$. Alice prepares d-dimensional quantum registers $\mathcal{A}_1^k, \ldots, \mathcal{A}_t^k$ to be in the states $\rho_{v_1}, \ldots, \rho_{v_t}$, respectively, for each $k \in [\ell]$. Then, she sends the sequence $\mathbf{v}$ through the channels by using each of them t times in a row. For each Bob this will result in t channel outputs from which he can infer that each v_i belongs to a particular clique in G. Now Alice execute with each individual Bob the remote state preparation scheme (Section 2.4.4) t times in a row, once for each of the states $\rho_{v_1}, \ldots, \rho_{v_t}$ separately. This requires that Alice communicates a total of $t\lceil \log d \rceil$ bits to each Bob. To do this, Alice uses the channels ℓ additional times and the k-th Bob will consider the $(t+k)$-th use of the channel as his output and ignore the others. Due to the way we chose t, this communication suffices to perform the remote state preparation. At this step, each Bob has registers $(\mathcal{B}_1^k, \ldots, \mathcal{B}_t^k)$ in states $\rho_{v_1}, \ldots, \rho_{v_t}$. Using their channel outputs, they can each perform a measurement on their registers and all recover the sequence $\mathbf{v} = (v_1, \ldots, v_t) \in V(G)^t$ with zero probability of error.

This concludes the proof as we have provided an entanglement-assisted protocol for the noiseless transmission of $|V(G)|^t$ distinct messages based on at most $t + \ell$ uses of the compound channel $\mathcal{N}$ with ℓ receivers. Indeed, this implies that $\alpha_{1,\ell}^\star(G^{\boxtimes(t+\ell)}) \geq |V(G)|^t$ and therefore

$$c_{1,\ell}^\star(G) \geq \frac{\log \alpha_{1,\ell}^\star(G^{\boxtimes(t+\ell)})}{t+\ell} \geq \frac{\log |V(G)|^t}{t+\ell},$$

as claimed. □

We show that this lower bound technique allows to obtain a separation between $c_{1,\ell}^\star(H_k)$ and $c_{1,\ell}(H_k)$ for certain k and $\ell \geq 1$ (Corollary 7.1.10 below).

7.1.9. THEOREM. *For every odd integer $k \geq 5$ and integer $\ell \in \mathbb{N}$, we have*

$$c_{1,\ell}^\star(H_k) \geq \frac{t}{t+\ell}(k-1)$$

with $t = \lfloor \frac{k-3}{2\log(k+1)} \rfloor$.

PROOF: Note that if $t = \lfloor \frac{k-3}{2\log(k+1)} \rfloor$, then from Lemmas 6.3.6 and 6.3.8 we get that $\xi'(H_k)^t \leq (k+1)^t \leq 2^{(k-3)/2} \leq \alpha(H_k)$. Therefore we can apply Theorem 7.1.8 and obtain the desired bound, since $|V(H_k)| = 2^{k-1}$. □

7.1.10. COROLLARY. *Consider any $k = 4p^s - 1$ such that p is an odd prime and s is a natural number. If $\ell < \frac{0.144k-1}{0.856k}\lfloor\frac{k-3}{2\log(k+1)}\rfloor$ then $c^{\star}_{1,\ell}(H_k) > c_{1,\ell}(H_k)$.*

PROOF: Easy algebraic manipulations give that for every $\ell < \left(\frac{0.144k-1}{0.856k}\right) t$ with $t = \lfloor\frac{k-3}{2\log(k+1)}\rfloor$ we have:

$$c^{\star}_{1,\ell}(H_k) \geq \frac{t}{t+\ell}(k-1) > 0.846k \geq c_{1,\ell}(H_k),$$

where we used Theorem 7.1.9 and Corollary 6.3.14. □

This means that our lower bound on the entangled compound Shannon capacity for $k \approx 1000$ is strictly larger than the classical capacity up to $\ell = 8$, for $k \approx 2000$ up to $\ell = 15$. Moreover, the upper bound on ℓ tends to infinity as k goes to infinity.

7.2 Multiple senders

We now move to a different zero-error communication scenario which can be seen as having multiple senders and a single receiver. Classically, cooperation among the senders might allow them to communicate more messages than the sum of their individual possibilities. We show that whenever a channel allows single-sender entanglement-assisted advantage, then the gain extends also to the multi-sender case (Theorem 7.2.3). Furthermore, for a fixed number of channel uses entanglement allows for a peculiar amplification of information which cannot happen classically (Theorem 7.2.6).

7.2.1 Cooperating senders channel coding

Suppose there are ℓ senders, each of whom gets access to a classical channel which connects her to the single receiver. We are interested in the total amount of messages that the senders, as a group, can transmit perfectly. At every stage of the communication only one of the senders uses her channel to communicate a message. We assume that inputs of one sender cannot be confused with inputs from another sender. In other words, the receiver knows which one of the senders sent him the message. We want to find what is the maximum cardinality of a message set that the senders are able to perfectly communicate to the receiver when they are allowed to cooperate.

Equivalently, this communication scenario can be depicted as single-sender single-receiver where the sender can choose among ℓ channels $\{\mathcal{N}_1, \mathcal{N}_2, \ldots \mathcal{N}_\ell\}$ to use for the communication. At every round of communication, the receiver learns the output of the channel as well as which channel has been used.

Suppose that the k-th Alice is connected to Bob through a channel $\mathcal{N}_k$ with confusability graph G_k. As noticed in [AL07], the confusability graph related to ℓ cooperating Alices is given by the disjoint union $G_1 + G_2 + \cdots + G_\ell$. The intuition being that, as inputs from different senders cannot be confused, there is no edge between vertices of G_i and G_j if $i \neq j$. The maximum size of a message set that can be perfectly transmitted with one use of the channels is then $\alpha(G_1 + G_2 + \cdots + G_\ell) = \sum_{i\in[\ell]} \alpha(G_i)$. However, the capacity of such a graph is in general non additive. Indeed, as previously mentioned, Alon [Alo98] showed the existence of a pair of distinct graphs G and H having the property that $\Theta(G + H) > \Theta(G) + \Theta(H)$. (However, if G and H are equal then $\Theta(G + G) = 2\Theta(G)$ must hold. We will show this simple fact at the end of this section.) From an information-theoretical perspective, the example of Alon says that when two senders are allowed to cooperate, the average number of messages they can communicate is strictly more than the sum of their individual possibilities. This result was extended by Alon and Lubetzky [AL07] for a larger number of senders, where they showed that it is possible to assign a channel to each sender such that only privileged subsets of senders are allowed to communicate with high capacity.

Suppose now that the parties can use an entanglement-assisted protocol. We focus on the particular case where all ℓ senders have access to the same channel $\mathcal{N}$ with confusability graph G. We notice that, since the senders are cooperating and there is no restriction on the amount of shared entanglement, we can assume that only one of the senders performs quantum operations on the entangled state. Hence, without loss of generality, the quantum state is bipartite and the entanglement-assisted strategy is equal to the one in Section 6.2 for a channel with confusability graph $G^{+\ell}$, where $G^{+\ell}$ denotes the disjoint union of ℓ copies of the graph G. An instance of such an entanglement-assisted strategy is pictured and explained in Figure 7.2.

7.2.1. DEFINITION. [Entangled multi-sender stability number and Shannon capacity] For a graph G, the *entangled multi-sender stability number* with ℓ senders is $\alpha^\star_{\ell,1}(G) = \alpha^\star(G^{+\ell})$. The *entangled multi-sender Shannon capacity* with ℓ senders is $c^\star_{\ell,1}(G) = c^\star(G^{+\ell})$ which by definition is equal to $\lim_{n\to\infty} \frac{1}{n} \log \alpha^\star\left((G^{+\ell})^{\boxtimes n}\right)$.

A useful observation is that $\alpha^\star_{\ell,1}(G) \geq \ell \cdot \alpha^\star(G)$ holds for every G and $\ell \in \mathbb{N}$. Indeed, each Alice can individually communicate $\alpha^\star(G)$ messages using entanglement and in our model Bob learns for free which Alice performed the communication. Therefore, the ℓ cooperating Alices can communicate at least one among $\ell \cdot \alpha^\star(G)$ distinct messages with one use of the channels and entanglement. Somewhat surprisingly, we present in Section 7.2.3 an example of a graph for which ℓ senders have a better joined strategy. In other words, there is a graph G and $\ell \in \mathbb{N}$ for which $\alpha^\star_{\ell,1}(G) > \ell \cdot \alpha^\star(G)$. This does not happen in the

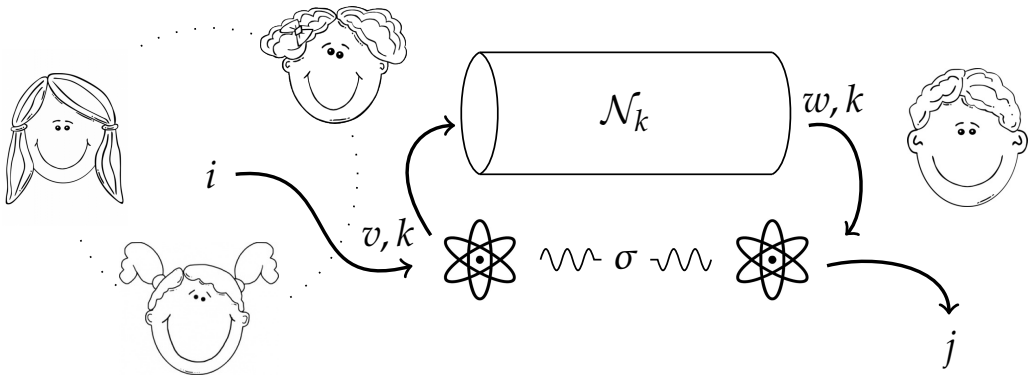

Figure 7.2: An instance of an entanglement-assisted multi-sender channel coding protocol: Suppose that the Alices want to communicate the message i to Bob, then one of them performs a measurement $\{A_i^{v,k}\}$ (that depends on i) on her part of the entangled state. The outcome (v, k) indicates that the k-th Alice should use her channel $\mathcal{N}_k$ to send input v. Bob receives an outcome w and, by assumption, he knows that channel $\mathcal{N}_k$ has been used for the communication. He can then perform measurement $\{B_{w,k}^j\}$ which depends on w and k, and outputs j. The protocol works if j is equal to i with zero probability of error.

classical case where, using analogous notation, for every G and $\ell \in \mathbb{N}$ we have $\alpha_{\ell,1}(G) = \alpha(G^{+\ell}) = \ell \cdot \alpha(G)$ and $c_{\ell,1}(G) = c(G^{+\ell}) = c(G) + \log \ell$ (or, analogously, $\Theta(G^{+\ell}) = \ell \cdot \Theta(G)$). We give a proof of this latter identity, which was also mentioned by Shannon [Sha56]. The key fact is that the strong graph product distributes over the disjoint union; i.e., $G \boxtimes (H_1 + H_2) = G \boxtimes H_1 + G \boxtimes H_2$ for every G, H_1, H_2 (see for example [HIK11] for a proof). This in particular implies that $(G^{+\ell})^{\boxtimes n} = (G^{\boxtimes n})^{+\ell^n}$. Using this last equality, we have

$$\begin{aligned} c(G^{+\ell}) &= \lim_{n\to\infty} \frac{1}{n} \log \alpha\big((G^{+\ell})^{\boxtimes n}\big) = \lim_{n\to\infty} \frac{1}{n} \log \alpha\big((G^{\boxtimes n})^{+\ell^n}\big) \\ &= \lim_{n\to\infty} \frac{1}{n} \log \left(\ell^n \cdot \alpha(G^{\boxtimes n})\right) = \lim_{n\to\infty} \frac{1}{n} \log \left(\alpha(G^{\boxtimes n})\right) + \log \ell \\ &= c(G) + \log \ell. \end{aligned}$$

7.2.2 Separation between classical and entangled multi-sender capacities

Here we show the following: for every graph with a separation between the classical and entangled capacity, there is also a separation in the multi-sender setting independently of the number of senders (Theorem 7.2.3). The same

type of result holds when we restrict to single use of the channels (Lemma 7.2.2). The latter is to be expected since we mentioned above that there is an easy quantum strategy that allows the ℓ Alices to communicate $\ell \cdot \alpha^\star(G)$ messages with a single communication round.

7.2.2. LEMMA. *Let G be a graph such that $\alpha^\star(G) > \alpha(G)$, then for every $\ell \in \mathbb{N}$ we have $\alpha^\star_{\ell,1}(G) > \alpha_{\ell,1}(G)$.*

PROOF: Each Alice can individually communicate $\alpha^\star(G)$ messages using an entanglement-assisted protocol. Since Bob also learns which Alice has sent him the message, the senders can transmit one among $\ell \cdot \alpha^\star(G)$ messages with entanglement. Thus, the claim follows from the chain of inequalities: $\alpha^\star_{\ell,1}(G) \geq \ell \cdot \alpha^\star(G) > \ell \cdot \alpha(G) = \alpha(G^{+\ell}) = \alpha_{\ell,1}(G)$. □

7.2.3. THEOREM. *Let G be a graph such that $c^\star(G) > c(G)$, then for every $\ell \in \mathbb{N}$ we have $c^\star_{\ell,1}(G) > c_{\ell,1}(G)$.*

PROOF: Recall that for any $\ell \in \mathbb{N}$ and graph G we have $(G^{+\ell})^{\boxtimes n} = (G^{\boxtimes n})^{+\ell^n}$ and $\alpha^\star_{\ell,1}(G) = \alpha^\star(G^{+\ell}) \geq \ell \cdot \alpha^\star(G)$. Therefore, we get

$$\begin{aligned}
c^\star_{\ell,1}(G) &= \lim_{n\to\infty} \left(\frac{1}{n} \log \left(\alpha^\star\left((G^{+\ell})^{\boxtimes n} \right) \right) \right) = \lim_{n\to\infty} \left(\frac{1}{n} \log \left(\alpha^\star\left((G^{\boxtimes n})^{+\ell^n} \right) \right) \right) \\
&\geq \lim_{n\to\infty} \left(\frac{1}{n} \log \left(\ell^n \cdot \alpha^\star(G^{\boxtimes n}) \right) \right) = \lim_{n\to\infty} \left(\frac{1}{n} \log(\alpha^\star(G^{\boxtimes n})) + \log \ell \right) \\
&= \lim_{n\to\infty} \left(\frac{1}{n} \log \left(\alpha^\star(G^{\boxtimes n}) \right) \right) + \log \ell = c^\star(G) + \log \ell \\
&> c(G) + \log \ell = c_{\ell,1}(G).
\end{aligned}$$

□

7.2.3 Improving communication by joint entanglement-assisted strategy

We exhibit a graph G and natural number ℓ for which $\alpha^\star_{\ell,1}(G) > \ell \cdot \alpha^\star(G)$. In other words, there is a joint entanglement-assisted strategy for ℓ senders which is strictly better than the sum of their optimal individual strategies. More generally, we are able to prove that there exist graphs for which cooperation among the senders allows for a better entanglement-assisted strategy for any finite number of channels uses. This is a peculiar property of the entanglement-assisted setting since in the classical case $\alpha_{\ell,1}(G) = \ell \cdot \alpha(G)$ and $c_{\ell,1}(G) = c(G) + \log \ell$ always hold. We do not know whether this improvement gained by cooperation in the entanglement-assisted setting extends also to the asymptotic regime.

In order to prove the result, we need to briefly describe a different two-party entanglement-assisted communication scenario that will be studied in detail in Section 8.1. Let G be a graph. Suppose that Alice receives a vertex $x \in V(G)$ and Bob receives (as side information) a clique $\mathcal{C} \subseteq V(G)$ under the promise that $x \in \mathcal{C}$. Moreover, Alice can send classical messages to Bob without error. What is the minimum cardinality of a message set that Alice has to use to communicate to Bob such that he can perfectly learn Alice's input x? In the classical scenario, the minimum cardinality is given by the chromatic number $\chi(G)$. Indeed, if the players agree on an optimal coloring of the graph, Alice can simply send the color corresponding to x to Bob. This suffices since elements of a clique all have different colors. Conversely, any deterministic strategy yields a coloring of the graph and, since we are in the zero-error regime, we can always assume the optimal strategy to be deterministic. Similarly, the entangled chromatic number $\chi^\star(G)$ is the minimum cardinality of a message set that Alice has to send to Bob such that he can perfectly learn x when Alice and Bob can share an arbitrary entangled state. This parameter has two useful properties: $\vartheta(\overline{G}) \leq \chi^\star(G)$ and $\chi^\star(G^{\boxtimes m}) \leq \chi^\star(G)^m$ for every graph G and $m \in \mathbb{N}$ (see Section 8.1.1).

We will use the following technical lemma, which can be seen as an entangled version of Theorem 3.2.1. Recall that $G\square K_t$ denotes the Cartesian product graph between the graph G and the complete graph K_t and that $\vartheta(G)$ is the Lovász theta number.

7.2.4. LEMMA. *For any graph G, if $t = \chi^\star(G)$ then $\alpha^\star(G\square K_t) = |V(G)|$.*

PROOF: First, we prove that for any $t \in \mathbb{N}$ the inequality $\alpha^\star(G\square K_t) \leq |V(G)|$ holds. For convenience let $|V(G)| = n$. For every $t \in \mathbb{N}$, we have that $\alpha^\star(G\square K_t) \leq \vartheta(G\square K_t) \leq \vartheta(\overline{K}_n \boxtimes K_t) = \vartheta(\overline{K}_n) \cdot \vartheta(K_t) = n = |V(G)|$. This chain of inequalities uses the fact that ϑ upper bounds $\alpha^\star$, $\overline{K}_n \boxtimes K_t$ is a subgraph of $G\square K_t$ and ϑ is monotone non-decreasing under taking subgraphs, ϑ is multiplicative under strong graph products, and that $\vartheta(\overline{K}_n) = n$ and $\vartheta(K_t) = 1$.

For the reverse inequality, let $t = \chi^\star(G)$ and suppose that Alice and Bob are connected through a classical channel $\mathcal{N}$ with confusability graph $G\square K_t$. We present a strategy that uses entanglement and allows to communicate $|V(G)|$ messages with a single use, thus implying $\alpha^\star(G\square K_t) \geq |V(G)|$. Suppose that Alice wants to send message $x \in V(G)$ to Bob. Using the strategy for the entangled chromatic number $\chi^\star(G)$, Alice makes a measurement on her part of the entangled state and gets an outcome $i \in [\chi^\star(G)]$. She sends message (x, i) through the channel. To any channel output $w \in W$ we can associate a clique in the confusability graph, describing the set of messages that are confusable to Bob given w. There are two types of cliques in $G\square K_t$: either $\{(z, h) : z \text{ is in a clique } \mathcal{C} \text{ of } G\}$ or $\{(z, h) : h \in H \subseteq [\chi^\star(G)]\}$. Suppose for the moment that from his channel output Bob infers that Alice's input (x, i) is an element

of the set $\{(y, i) : y \in \mathcal{C}_x$ where $\mathcal{C}_x$ is a clique of G containing $x\}$. Since Bob learns $\mathcal{C}_x$, he can use message i to finish the protocol of the entangled chromatic number. As mentioned above the protocol allows Bob to recover x with zero probability of error. For the other case, Bob from his output learns that Alice's input is an element in the set $\{(x, j) : j \in J \subseteq [\chi^\star(G)]$ with $i \in J\}$. Then he can directly deduce that x is the message that Alice wanted to send. Hence, we have shown an entanglement-assisted protocol that allows to perfectly communicate $|V(G)|$ classical messages through a channel with confusability graph $G\Box K_t$. Therefore, if $t = \chi^\star(G)$ then $\alpha^\star(G\Box K_t) \geq |V(G)|$ holds. Combining this inequality with the one derived at the beginning of the proof, we can conclude. □

The following lemma was proven in a more general context by Gvozdenović and Laurent [GL08, Lemma 2.4].

7.2.5. LEMMA. *Let G be a graph such that $\chi^\star(G)\,\alpha^\star(G) < |V(G)|$ and $t = \chi^\star(G)$. Then, $\alpha^\star(G^{+t}) > t \cdot \alpha^\star(G)$.*

PROOF: We have: $t \cdot \alpha^\star(G) = \chi^\star(G)\,\alpha^\star(G) < |V(G)| = \alpha^\star(G\Box K_t) \leq \alpha^\star(G^{+t})$, where we used Lemma 7.2.4 and, in the last inequality, that G^{+t} is a subgraph of $G\Box K_t$ and that the parameter $\alpha^\star$ is monotone non-decreasing under taking subgraphs. □

The above lemma is the key fact to obtain the desired result. We now simply have to exhibit a graph G where $\chi^\star(G)\,\alpha^\star(G) < |V(G)|$. We remark that classically $\chi(G)\,\alpha(G) \geq |V(G)|$ always holds and indeed $\alpha(G^{+\ell}) = \ell \cdot \alpha(G)$ for every $\ell \in \mathbb{N}$ and graph G.

To this end, we will use the orthogonality graph Ω_k (Definition 3.3.2); i.e., the graph with $\{\pm 1\}^k$ as vertex set and two vectors are adjacent if orthogonal. From [MR16], we know that $\vartheta(\Omega_k) = 2^k/k$ and $\vartheta(\overline{\Omega_k}) = k$ if k is a multiple of four. Consider the orthogonal representation $f : V(\Omega_k) \to \mathbb{R}^k$ with $f(v) = v/\sqrt{k}$ that maps vertices of Ω_k to the unit sphere and adjacent vertices to orthogonal vectors. Thus, we have $\xi'(\Omega_k) \leq k$. Since $\chi^\star$ is upper bounded by the minimum dimension of an orthogonal representation in which all the entries of the vectors have equal moduli (Lemma 8.1.3), we have that $k = \vartheta(\overline{\Omega_k}) \leq \chi^\star(\Omega_k) \leq \xi'(\Omega_k) \leq k$, and thus $\chi^\star(\Omega_k) = k$, for every k multiple of four.

We can now prove the main result of the section.

7.2.6. THEOREM. *Let Ω_k be the orthogonality graph with k a multiple of four but not a power of two. Then $\alpha^\star_{k,1}(\Omega_k) = \alpha^\star(\Omega_k^{+k}) > k \cdot \alpha^\star(\Omega_k)$.*

PROOF: From the discussion above we know that $\chi^\star(\Omega_k) = k$. Moreover, we have $\alpha^\star(\Omega_k) \leq \lfloor \vartheta(\Omega_k) \rfloor = \lfloor 2^k/k \rfloor < 2^k/k = \vartheta(\Omega_k)$. Using a similar argument as in [MR16], we get that $\chi^\star(\Omega_k)\,\alpha^\star(\Omega_k) < |V(\Omega_k)|$ since

$$\chi^\star(\Omega_k)\,\alpha^\star(\Omega_k) \leq k \cdot \lfloor 2^k/k \rfloor < k \cdot 2^k/k = 2^k = |V(\Omega_k)|.$$

Using Lemma 7.2.5 we conclude that $\alpha^\star_{k,1}(\Omega_k) = \alpha^\star(\Omega_k^{+k}) > k \cdot \alpha^\star(\Omega_k)$. □

With a similar reasoning, we can prove that for every finite number of uses of the channel, cooperation among the players improves the entanglement-assisted communication. Let $\alpha^\star_{\ell,1}(G,n) = \alpha^\star((G^{+\ell})^{\boxtimes n})$ be the maximum cardinality of a message set that ℓ Alices can communicate perfectly to Bob with n uses of the channel and entanglement.

In the next lemma, we show that there exist a graph G and $\ell \in \mathbb{N}$ such that $\alpha^\star_{\ell,1}(G,n) > \ell^n \cdot \alpha^\star(G^{\boxtimes n})$ for every $n \in \mathbb{N}$. This is equivalent to saying that there exists a channel and a certain number of senders for which cooperation among the senders strictly improves the communication of n channel uses for every $n \in \mathbb{N}$.

7.2.7. THEOREM. *Let Ω_k be the orthogonality graph with k a multiple of four but not a power of two. Then, $\alpha^\star_{k,1}(\Omega_k,n) > k^n \cdot \alpha^\star(\Omega_k,n)$ for every $n \in \mathbb{N}$.*

PROOF: Using Lemma 3.2.5 (*ii*) and the above discussion, we deduce that $\vartheta(\Omega_k^{\boxtimes n}) = \vartheta(\Omega_k)^n = \left(\frac{2^k}{k}\right)^n$ for every $n \in \mathbb{N}$. Moreover, from (3.1) together with Lemma 3.2.5 (*iii*), we have that $\vartheta(\overline{\Omega_k^{\boxtimes n}}) = \vartheta(\overline{\Omega}_k^{*n}) = \vartheta(\overline{\Omega}_k)^n = k^n$ for any $n \in \mathbb{N}$. Then by sub-multiplicativity of $\chi^\star(G)$ and since $\chi^\star(\Omega_k) = k$, we have $k^n = \vartheta(\overline{\Omega_k^{\boxtimes n}}) \leq \chi^\star(\Omega_k^{\boxtimes n}) \leq \chi^\star(\Omega_k)^n = k^n$. This implies that for any integer n,

$$\alpha^\star(\Omega_k^{\boxtimes n}) \leq \lfloor \vartheta(\Omega_k^{\boxtimes n}) \rfloor = \left\lfloor \left(\frac{2^k}{k}\right)^n \right\rfloor < \left(\frac{2^k}{k}\right)^n = \frac{|V(\Omega_k^{\boxtimes n})|}{\chi^\star(\Omega_k^{\boxtimes n})}.$$

Applying Lemma 7.2.5, we then have that $\alpha^\star\big((\Omega_k^{\boxtimes n})^{+k^n}\big) > k^n \cdot \alpha^\star(\Omega_k^{\boxtimes n})$ for every $n \in \mathbb{N}$. We can conclude that

$$\alpha^\star_{k,1}(\Omega_k,n) = \alpha^\star\big((\Omega_k^{\boxtimes n})^{+k^n}\big) > k^n \cdot \alpha^\star(\Omega_k^{\boxtimes n}) = k^n \cdot \alpha^\star(\Omega_k,n).$$

□

A particular graph. In this section, we exhibited examples of graphs whose entangled stability number is strictly greater than the sum of the entangled stability number of the disjoint components. A smaller graph that has this property was recently found by Mančinska and Roberson [MR15].

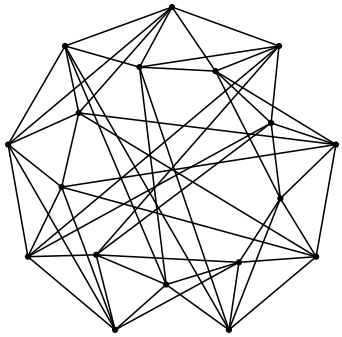

Figure 7.3: Line graph of $K_3 \square K_3$.

Consider the following 9 orthogonal bases of $\mathbb{R}^4$, say, $B_1, \ldots, B_9$:

$$\begin{bmatrix} (0,\ 0,\ 0, 1) \\ (0,\ 0,\ 1, 0) \\ (1,\ 1,\ 0, 0) \\ (1, -1, 0, 0) \end{bmatrix} \quad \begin{bmatrix} (0, 0,\ 0,\ 1) \\ (0, 1,\ 0,\ 0) \\ (1, 0,\ 1,\ 0) \\ (1, 0, -1, 0) \end{bmatrix} \quad \begin{bmatrix} (1, -1,\ 1,\ -1) \\ (1, -1, -1,\ 1) \\ (1,\ 1,\ \ 0,\ \ 0) \\ (0,\ 0,\ \ 1,\ \ 1) \end{bmatrix} \quad \begin{bmatrix} (1, -1,\ 1,\ -1) \\ (1,\ 1,\ \ 1,\ \ 1) \\ (1,\ 0,\ -1,\ 0) \\ (0,\ 1,\ \ 0,\ -1) \end{bmatrix} \quad \begin{bmatrix} (0, 0, 1,\ 0) \\ (0, 1, 0,\ 0) \\ (1, 0, 0,\ 1) \\ (1, 0, 0, -1) \end{bmatrix}$$

$$\begin{bmatrix} (1, -1, -1,\ 1) \\ (1,\ 1,\ \ 1,\ \ 1) \\ (1,\ 0,\ \ 0,\ -1) \\ (0,\ 1,\ -1,\ 0) \end{bmatrix} \quad \begin{bmatrix} (1,\ 1,\ -1,\ 1) \\ (1,\ 1,\ \ 1,\ -1) \\ (1, -1,\ 0,\ \ 0) \\ (0,\ 0,\ \ 1,\ \ 1) \end{bmatrix} \quad \begin{bmatrix} (1,\ 1, -1,\ 1) \\ (-1, 1,\ 1,\ \ 1) \\ (1,\ 0,\ 1,\ \ 0) \\ (0,\ 1,\ 0,\ -1) \end{bmatrix} \quad \begin{bmatrix} (1,\ 1,\ 1,\ -1) \\ (-1, 1,\ 1,\ \ 1) \\ (1,\ 0,\ 0,\ \ 1) \\ (0,\ 1, -1,\ 0) \end{bmatrix}$$

One can easily check that each vector is contained in exactly two bases. Construct the graph G as follows: to each of the above vectors we associate a vertex and they are adjacent if the corresponding vectors are elements of the same basis. (Note that orthogonal vectors do not have to be adjacent.) Equivalently, G is the line graph of the Cartesian product graph $K_3 \square K_3$; i.e., the graph whose vertices represents the edges of the original graph and where two vertices are adjacent if the corresponding edges share an endpoint. This graph has 18 vertices and 54 edges. It is depicted in Figure 7.3.

By construction $\xi(G) \leq 4$ and equality holds because $4 = \omega(G) \leq \xi(G)$. Using Proposition 3.3.3 and the fact that $\chi^\star(G) \leq \chi_q(G)$, we have $\chi^\star(G) \leq 4$. At the same time, $\alpha^\star(G) = 4$ since $4 = \alpha(G) \leq \alpha^\star(G) \leq \vartheta(G) = 9/2$. As $\alpha^\star(G)\,\chi^\star(G) = 16 < 18 = |V(G)|$, Lemma 7.2.5 applies and we have $\alpha^\star(G^{+4}) \geq 18 > 16 = 4\,\alpha^\star(G)$. Actually, $\alpha^\star(G^{+4}) = 18$ holds because $\vartheta(G^{+4}) = 4 \cdot 9/2 = 18$.

Even better, Mančinska and Roberson [MR15] gave an entanglement-assisted protocol with which one can transmit 9 classical messages with a channel whose confusability graph is $G + G$. Moreover, as $\vartheta(G + G) = 9$, we have the following identities $\alpha^\star(G + G) = \Theta^\star(G + G) = 9$.

7.2.8. LEMMA (MANČINSKA–ROBERSON [MR15]). *Let G be the line graph of the Cartesian product $K_3 \square K_3$, then $\alpha^\star(G+G) = 9$. Furthermore, this implies that $\alpha^\star(G+G) > 2\alpha^\star(G)$.*

PROOF: We only prove the non-trivial inequality: $\alpha^\star(G+G) \geq 9$. In what follows, we associate to each vertex $v \in V(G)$ its corresponding (normalized) vector f_v. The protocol goes as follows. Let the pair of quantum registers $(\mathcal{A}, \mathcal{B})$ be initialized to be in the 4-dimensional maximally entangled state $\sigma = zz^*$ where $z = (\sum_{\ell=1}^4 e_\ell \otimes e_\ell)/\sqrt{4}$. If Alice wants to communicate message $i \in [9]$, she performs the measurement $\{f_u f_u^T\}_{f_u \in B_i}$ and gets an outcome $f_u \in B_i$. Bob's register is then in a state proportional to $\rho_i^u = \mathsf{Tr}_{\mathcal{A}}((f_u f_u^T \otimes I)\sigma) = (f_u f_u^T)^T/4$. Let j be the index such that $f_u \in B_i \cap B_j$. If $i < j$, Alice uses the first copy of the graph G to send vertex u as input. Otherwise, to send u she uses the second copy of the graph G. As outcome of the channel, Bob gets a clique $\mathcal{C}$, where $u \in \mathcal{C}$, together with an index $k \in \{1,2\}$ that tells him which copy of the channel Alice has used. From $\mathcal{C}$, Bob constructs a measurement $\{P^v : v \in \mathcal{C}\} \cup \{P^T\}$ where $P^v = \{f_v f_v^T\}$ and $P^T = I - \sum_{v \in \mathcal{C}} P^v$, which he uses to measure the state ρ_i^u and gets u as outcome. Now, he knows that Alice either wanted to send index i or index j. If as output of the channel Bob received index $k = 1$ he outputs the minimum between i and j, otherwise he outputs $\max(i,j)$. Therefore, Alice can perfectly communicate to Bob 9 different messages and $\alpha^\star(G+G) \geq 9$. □

At last, we mention that this graph has another peculiar property. As observed in [LMM+12], for the very few graphs for which we know a separation between $\alpha^\star(G)$ and $\alpha(G)$ we have $\alpha^\star(G) = \Theta^\star(G)$. This might, or might not, be the case for the quarter-orthogonality graphs (Definition 6.3.1). However, we will now show that for this graph we have: $\alpha^\star(G) < \Theta^\star(G)$. As far as we know, this is the only graph with such a property.

We already know that $\alpha^\star(G) = 4$ and now prove that $\Theta^\star(G) \geq 3\sqrt{2}$. We claim that Alice can transmit $|V(G)|$ different messages to Bob with two uses of the channel and entanglement. The bound now follows as $|V(G)| = 18$. We use the same protocol as in the proof of Theorem 6.3.9.

For each $v \in V(G)$, define the state $\rho_v = f_v f_v^T$, where as before f_v is the normalized vector corresponding to vertex v. Suppose that Alice wants to send message $u \in V(G)$. She prepares her 4-dimensional quantum register $\mathcal{A}$ to be in state ρ_u and sends u through the channel. Using his output, Bob infers a clique $\mathcal{C}$ of the graph containing u. Since the state ρ_u lives in $\mathbb{R}^{4\times 4}$, Zeng and Zhang [ZZ02] have a scheme that allows to remote-state prepare ρ_u requiring only the transmission of 2 classical bits. As $\alpha(G) = 4$, this can be achieved by using the channel one additional time. Now Bob's register is in state ρ_u and, knowing $\mathcal{C}$, he can construct a measurement such that he is guaranteed to get u as outcome. Therefore, $\Theta^\star(G) \geq \sqrt{\alpha^\star(G^{\boxtimes 2})} \geq \sqrt{18} = 3\sqrt{2}$.

It would be of particular interest if this lower bound was tight. There reason begin that since $\Theta^\star(G+G) = 9$ the graph G is a candidate for having the property $\Theta^\star(G+G) > 2\Theta^\star(G)$. If this were to be true it would show that the parameter $\Theta^\star$ can be strictly smaller than the Lovász theta number.

Chapter 8

Source and source-channel coding

In this chapter we study two zero-error coding problems: the source and the source-channel coding problems, in the scenario where the parties may use entanglement. Here, Alice and Bob are given correlated inputs from a random source and they are connected by a one-way classical noiseless channel (source coding) or a noisy one (source-channel coding). Their goal is for Bob to learn Alice's input with zero probability of error, while using the channel as little as possible. In Section 8.1 we study the source coding problem and show that entanglement can allow for source coding schemes that are exponentially more efficient than classical ones. In Section 8.2 we present the source-channel coding problem and prove that, also in this case, entanglement allows for coding schemes which are exponentially better than the classical ones.

The content of this chapter is based on joint work with Jop Briët, Harry Buhrman, Monique Laurent, and Giannicola Scarpa [BBL$^+$15a].

8.1 The source coding problem

The *source coding problem* asks a sender to communicate data about which a receiver has already some information. The sender can supply additional information to the receiver by using a noiseless binary channel.

A *source* $\mathcal{M}$ consists of a finite set X, a (possibly infinite) set S and a probability distribution P over $\mathsf{X} \times \mathsf{S}$. In a source instance, Alice is given an input $x \in \mathsf{X}$ and Bob an input $s \in \mathsf{S}$ with probability $P(x, s)$. Bob's input may already give him some information about Alice's. But if his input does not uniquely identify hers, she can supply some additional information by getting access to a noiseless one-way binary channel. Their goal is for Bob to learn Alice's input while minimizing the use of the channel. Here we consider only *memoryless* sources, which means that the probability distribution $P(x, s)$ of the source is unchanged after every instance.

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 $\mathbf{s} = (s_1, \ldots, s_m) \in \mathsf{S}^m$, then applies a decoding function $\mathsf{D} : \mathsf{S}^m \times \{0,1\}^n \to \mathsf{X}^m$ to the pair $(\mathbf{s}, \mathsf{C}(\mathbf{x}))$ 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.

In the vanishing error regime, Slepian and Wolf [SW74] showed that the amount of communication that Alice needs to supply is equal to the information theoretical lower bound. That is, let (X, S) be the random variable pair generated according to the distribution $P(x, s)$. Then, asymptotically, Alice has to transmits to Bob at a cost rate which is equal to the conditional entropy function $H(X|S)$. Here we focus however on the zero-error scenario.

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 source $\mathcal{M}$ is its *characteristic graph* $H = (\mathsf{X}, E)$, where $\{x, y\} \in E$ if there exists a $s \in \mathsf{S}$ such that $P(x, s) > 0$ and $P(y, s) > 0$. As such, the edge set identifies the pairs of inputs for Alice which Bob cannot distinguish based on his input. Notice that every graph is the characteristic graph of a (non-unique) source. Solving one instance of the zero-error source coding problem for $\mathcal{M}$ is equivalent to finding a proper coloring of H. Indeed, Bob's input s reduces the list of Alice's possible inputs to the set $\{x \in \mathsf{X} : P(x, s) > 0\}$ and this set forms a clique in H. 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 H. To deal with length-m input sequences we take the strong product graph $H^{\boxtimes m}$, whose edges are the pairs of input sequences for Alice which Bob cannot distinguish. The *Witsenhausen rate*

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

is the minimum asymptotic cost rate of a zero-error code for a source. As is well-known and easy to check, the chromatic number is sub-multiplicative; i.e., $\chi(H^{\boxtimes(m+n)}) \leq \chi(H^{\boxtimes m})\chi(H^{\boxtimes n})$. Therefore, by Fekete's Lemma (Lemma 6.1.1) the above limit exists and is equal to the infimum: $R(H) = \inf_m \log \chi(H^{\boxtimes m})/m$.

8.1.1 Entanglement-assisted source coding

Consider the same setup as before, except now Alice and Bob have quantum registers $\mathcal{A}$ and $\mathcal{B}$, respectively, that are initialized to be in some entangled

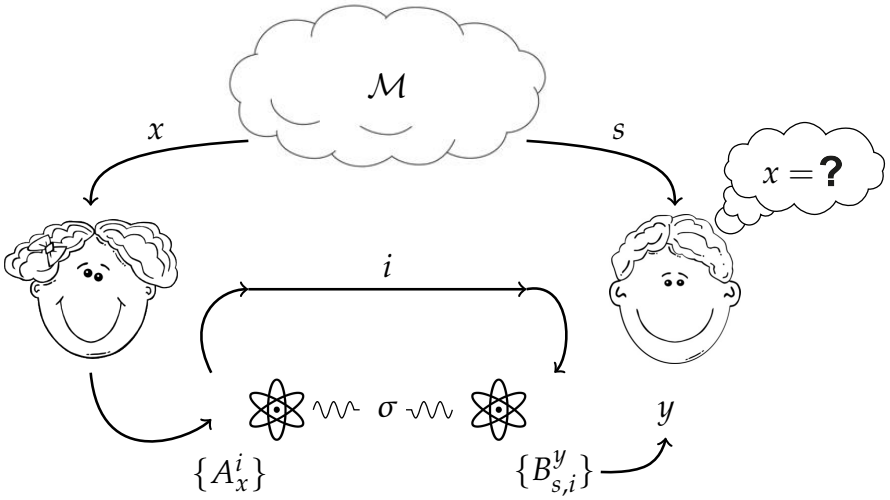

Figure 8.1: An entanglement-assisted source coding protocol.

state. If the source $\mathcal{M}$ gives the parties inputs $x \in \mathsf{X}$ and $s \in \mathsf{S}$, respectively, then their most general course of action is as follows (see also Figure 8.1):

1. After receiving her input x, Alice performs a measurement on her register $\mathcal{A}$ and communicates the measurement outcome, say i, to Bob;

2. After receiving both his input s and Alice's measurement outcome i, Bob performs a measurement on his register $\mathcal{B}$ and obtains a measurement outcome $y \in \mathsf{X}$.

The protocol is successful if Bob gets outcome $y = x$ for each possible input pair (x, s). In a fashion similar to Section 6.2, one arrives to the following entangled variants of the chromatic number and Witsenhausen rate.

8.1.1. DEFINITION. [Entangled chromatic number and Witsenhausen rate] For a graph H, define $\chi^\star(H)$ as the minimum integer $t \in \mathbb{N}$ for which there exist $d \in \mathbb{N}$ and positive semidefinite matrices ρ and $\{\rho_x^i : x \in V(H), i \in [t]\}$ in $\mathbb{C}^{d\times d}$ such that $\mathsf{Tr}(\rho) = 1$ and

$$\begin{aligned} \rho_x^i \rho_y^i &= 0 \quad \forall i \in [t] \text{ and } x, y \in V(H) \text{ such that } \{x, y\} \in E(H), \\ \sum_{i\in[t]} \rho_x^i &= \rho \quad \forall x \in V(H). \end{aligned}$$

The *entangled Witsenhausen rate* is defined by

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

The operational interpretation of the parameter $\chi^\star(H)$ makes it easy to see that it is sub-multiplicative with respect to strong graph products; that is, $\chi^\star(H \boxtimes H') \leq \chi^\star(H)\,\chi^\star(H')$. This implies that the entangled Witsenhausen rate is also given by the infimum: $R^\star(H) = \inf_m \log \chi^\star(H^{\boxtimes m})/m$.

Recall that in Section 5.1.1 we have shown that the Lovász theta number is a lower bound on the entangled chromatic number; i.e., $\vartheta(\overline{H}) \leq \chi^\star(H)$. By the multiplicativity of $\vartheta(H)$ under strong graph products, we can directly deduce the inequality: $\log \vartheta(\overline{H}) \leq R^\star(H)$.

8.1.2 Separation between classical and entangled Witsenhausen rate

We exhibit an exponential gap between the entangled and the classical Witsenhausen rate for the quarter-orthogonality graphs H_k (Definition 6.3.1).

We remark that, as for the channel coding problem, entanglement-assisted protocols may provide an advantage only in the zero-error scenario. Indeed, if one allows a vanishing probability of error, the above mentioned result of Slepian and Wolf [SW74] implies that, asymptotically, entanglement cannot improve the communication cost rate.

The remaining of the section will be use to prove the following result.

8.1.2. THEOREM. *For every odd integer k, we have*

$$R^\star(H_k) \leq \log(k+1). \tag{8.1}$$

Moreover, if $k = 4p^\ell - 1$ where p is an odd prime and $\ell \in \mathbb{N}$, then

$$R(H_k) \geq 0.154k - 1. \tag{8.2}$$

Upper bound on the entangled Witsenhausen rate. In order to prove the bound (8.1) on $R^\star(H_k)$, we first show that $\chi^\star(H) \leq \xi'(H)$. This inequality can be derived from the fact that $\chi^\star(H) \leq \chi_q(H)$ and a result of [CMN$^+$07] stating that $\chi_q(H) \leq \xi'(H)$. We give a self-contained proof of the implied bound on $\chi^\star(H)$. From our proof, it follows almost immediately that also $\chi^\star(H) \leq \xi(H)^2$ holds; the only thing to change in the proof is to replace the remote state preparation scheme described in Section 2.4.4 by quantum teleportation.

For any graph H, $\xi(H)$ is the minimum dimension of an orthogonal representation; while $\xi'(H)$ is the minimum dimension of an orthogonal representation where each vector has all the entries with absolute value one.

8.1.3. LEMMA. *For every graph H, we have $\chi^\star(H) \leq \xi'(H)$.*

PROOF: Consider a source $\mathcal{M}$ with characteristic graph H. Let $d = \xi'(H)$ and let f be a d-dimensional orthogonal representation of G such that all the entries of the vectors have modulus one. It suffices to find an entanglement-assisted protocol for $\mathcal{M}$ that involves only d-outcome measurements on Alice's part. To this end, let us recall the observation that Bob's input s allows him to reduce the list of Alice's possible inputs to a clique $\mathcal{C}$ in H that contains Alice's actual input x. Next, notice that the set of states $f(y)f(y)^*/d$ for $y \in \mathcal{C}$ are pairwise orthogonal. This suggests the following protocol. First the parties perform the remote state preparation protocol (Section 2.4.4) to put a quantum register belonging to Bob in the state $f(x)f(x)^*/d$. Now Bob performs the measurement with outcomes in $\mathcal{C} \cup \{\perp\}$ as promised to exist by the Orthogonality Lemma (Lemma 2.4.1) to learn which of the states $f(y)f(y)^*/d$ his register is in, thus learning x. The result now follows because the remote state preparation involves only d-outcome measurements. □

Combining Lemma 6.3.8 with the one above, we easily get the following upper bound on the entangled chromatic number.

8.1.4. LEMMA. *Let k be an odd positive integer and $m \in \mathbb{N}$. Then, the inequality $\chi^\star(H_k^{\boxtimes m}) \leq (k+1)^m$ holds. Moreover, we have equality if there exists a Hadamard matrix of size $k+1$.*

PROOF: By the sub-multiplicativity of $\chi^\star(H)$, Lemma 8.1.3 and Lemma 6.3.8, we have $\chi^\star(H_k^{\boxtimes m}) \leq \chi^\star(H_k)^m \leq \xi'(H_k)^m \leq (k+1)^m$.

Suppose now that there exists a Hadamard matrix of size $k+1$. Recall from Proposition 6.3.7 that the existence of a Hadamard matrix of size $k+1$ implies $\omega(H_k) \geq k+1$. Combining this with the fact that for every graph G the inequalities $\chi^\star(G) \geq \vartheta(\overline{G}) \geq \omega(G)$ hold (see (3.14)), then for every $m \in \mathbb{N}$ we have

$$\chi^\star(H_k^{\boxtimes m}) \geq \vartheta(\overline{H_k^{\boxtimes m}}) \geq \omega(H_k^{\boxtimes m}) \geq \omega(H_k)^m \geq (k+1)^m,$$

where the third inequality uses the simple fact that if a subset $W \subseteq V(G)$ forms a clique in a graph G, then the set W^m of m-tuples of elements from W forms a clique in $G^{\boxtimes m}$. □

The bound (8.1) can now be derived as a simple corollary.

8.1.5. COROLLARY. *Let k be an odd positive integer. Then $R^\star(H_k) \leq \log(k+1)$.*

PROOF: We have $R^\star(H_k) = \inf_m \log \chi^\star(H_k^{\boxtimes m})/m \leq \log(k+1)$, where in the last inequality we used Lemma 8.1.4. □

Lower bound on the Witsenhausen rate. The bound (8.2) on $R(H_k)$ follows from the upper bound on the classical stability number of the graphs $H_k^{\boxtimes m}$ for certain values of k given in Lemma 6.3.11.

8.1.6. COROLLARY. *Let p be an odd prime and $\ell \in \mathbb{N}$. Then, for $k = 4p^\ell - 1$, we have $R(H_k) \geq 0.154k - 1$.*

PROOF: By Lemma 6.3.11, for every integer m, we have

$$\chi(H_k^{\boxtimes m}) \geq \frac{|V(H_k^{\boxtimes m})|}{\alpha(H_k^{\boxtimes m})} > \frac{2^{(k-1)m}}{2^{0.846km}} = 2^{(0.154k-1)m}.$$

Taking the logarithm, dividing by m and taking the limit as m tends to infinity gives the result. □

8.2 The source-channel coding problem

In the *source-channel coding problem* the parties receive inputs from a source $\mathcal{M}$ and get access to a channel $\mathcal{N}$. Their goal is to solve the source coding problem, but now using the (noisy) channel $\mathcal{N}$ instead of a (noiseless) binary channel. An (m, n)*-coding scheme* for this problem consists of an encoding function $\mathsf{C} : \mathsf{X}^m \to \mathsf{V}^n$ and a decoding function $\mathsf{D} : \mathsf{S}^m \times \mathsf{W}^n \to \mathsf{X}^m$. The *cost rate* is n/m and gives the average number of channel uses per source-input symbol.

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

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

gives the minimum asymptotic cost rate of a zero-error code. We will assume throughout that both H and $\overline{G}$ contain at least one edge. (Indeed, if H has no edge then $\eta(H, G) = 0$ for any G and, if H has at least one edge, then $\eta(H, G)$ is well-defined only if $\overline{G}$ has at least one edge.) To see that the limit exists, observe that the parameter

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

is sub-additive and apply Fekete's Lemma (Lemma 6.1.1), which shows that $\eta(H, G) = \lim_{m\to\infty} \eta_m(H, G)/m$ is also equal to the infimum $\inf_m \eta_m(H, G)/m$.

If the channel $\mathcal{N}$ is replaced by a noiseless 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 at all about Alice's one, then we regain the channel coding problem. More formally, we can reformulate $R(H)$ and $c(G)$ in the following way.

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

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

PROOF: For the proof of the identity $R(H) = \eta(H, \overline{K}_2)$ we use the following simple fact: for a graph H' and $t \in \mathbb{N}$, there exists a homomorphism from H' to K_t if and only if $\chi(H') \leq t$. That is, $\chi(H') = \min\{t : H' \longrightarrow K_t\}$ and taking the logarithms we have

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

Combining these inequalities applied to $H' = H^{\boxtimes m}$ with the simple identity $\overline{K_2^{\boxtimes n}} = K_{2^n}$, we obtain

$$\begin{aligned} \eta(H, \overline{K}_2) &= \lim_{m\to\infty} \frac{1}{m} \min\{n : H^{\boxtimes m} \longrightarrow \overline{K_2^{\boxtimes n}} = K_{2^n}\} \\ &= \lim_{m\to\infty} \frac{1}{m} \log \chi(H^{\boxtimes m}) \\ &= R(G). \end{aligned}$$

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

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

Setting $n_m = \eta_m(K_2, G)$, this implies

$$\log \alpha(G^{\boxtimes(n_m-1)}) < m \leq \log \alpha(G^{\boxtimes n_m})$$

and thus

$$\frac{n_m}{\log \alpha(G^{\boxtimes n_m})} \leq \frac{n_m}{m} < \frac{n_m}{\log \alpha(G^{\boxtimes(n_m-1)})}. \tag{8.3}$$

As $c(G) = \sup_n \log \alpha(G^{\boxtimes n})/n$, using the left most inequality in (8.3) we deduce that for all m

$$\frac{1}{c(G)} \leq \frac{n_m}{\log \alpha(G^{\boxtimes n_m})} \leq \frac{n_m}{m}.$$

Taking the limit, we obtain $1/c(G) \leq \lim_{m\to\infty} n_m/m = \eta(K_2, G)$. Next, as $\eta(K_2, G) = \inf_m n_m/m$, using the right most inequality in (8.3) we get that

$$\eta(K_2, G) \leq \frac{n_m}{m} < \frac{n_m}{\log \alpha(G^{\boxtimes(n_m-1)})} = \frac{n_m - 1}{\log \alpha(G^{\boxtimes(n_m-1)})} \frac{n_m}{n_m - 1}.$$

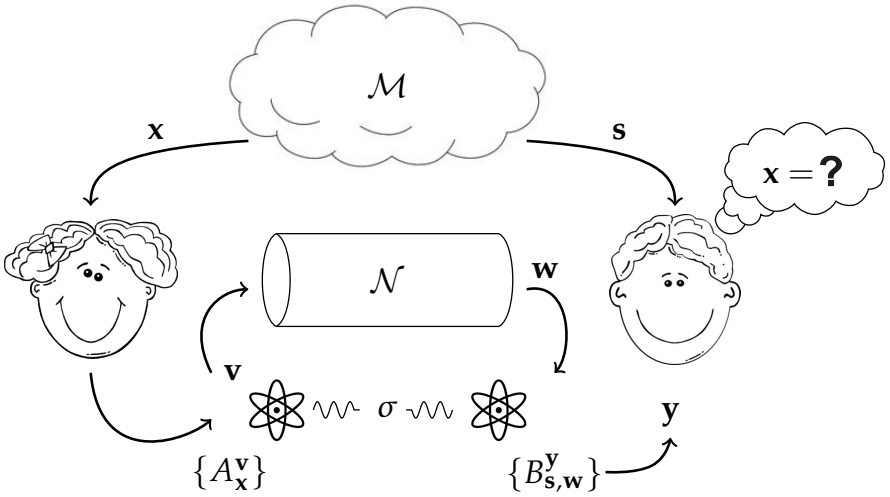

Figure 8.2: An entanglement-assisted source-channel coding protocol.

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 $1/c(G)$, which shows the reverse inequality $\eta(K_2, G) \leq 1/c(G)$. Thus the equality $\eta(K_2, G) = 1/c(G)$ holds. □

8.2.1 Entanglement-assisted source-channel coding

Suppose that now Alice and Bob possess quantum registers $\mathcal{A}$ and $\mathcal{B}$, respectively, that are initialized to be in some entangled state σ. The entanglement-assisted version of an (m, n)-coding scheme is as follows (see also Figure 8.2):

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

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

3. Alice sends $\mathbf{v}$ through the channel $\mathcal{N}$ after which Bob receives some sequence $\mathbf{w} \in \mathsf{W}^n$;

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

Using the same arguments as in Section 6.2, one then arrives at the following variants of the cost-rate.

8.2.2. DEFINITION. [Entangled source-channel cost rate] For graphs G, H and number $m \in \mathbb{N}$, define $\eta^\star_m(H, G)$ as the minimum integer $n \in \mathbb{N}$ for which there exist $d \in \mathbb{N}$ and positive semidefinite matrices ρ and $\{\rho^{\mathbf{u}}_{\mathbf{x}} : \mathbf{x} \in V(H^{\boxtimes m}), \mathbf{u} \in V(G^{\boxtimes n})\}$ in $\mathbb{C}^{d\times d}$ such that $\mathsf{Tr}(\rho) = 1$ and

$$\begin{aligned} \rho^{\mathbf{u}}_{\mathbf{x}}\rho^{\mathbf{v}}_{\mathbf{y}} = 0 \quad & \forall \mathbf{x}, \mathbf{y} \text{ such that } \{\mathbf{x}, \mathbf{y}\} \in E(H^{\boxtimes m}) \text{ and} \\ & \forall \mathbf{u}, \mathbf{v} \text{ such that } \mathbf{u} = \mathbf{v} \text{ or } \{\mathbf{u}, \mathbf{v}\} \in E(G^{\boxtimes n}), \\ \sum_{\mathbf{u}\in V(G^{\boxtimes n})} \rho^{\mathbf{u}}_{\mathbf{x}} = \rho \quad & \forall \mathbf{x} \in V(H^{\boxtimes m}). \end{aligned}$$

The *entangled source-channel cost rate* is defined by

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

As for the classical counterpart, we assume throughout that both graphs H and $\overline{G}$ contain at least one edge, thereby excluding trivial settings. It is not difficult to see that we regain the parameter $\eta(H, G)$ if we restrict the above matrices ρ and $\rho^{\mathbf{u}}_{\mathbf{x}}$ to be $\{0, 1\}$-valued scalars. Sharing an entangled quantum system cannot make the coding scheme worse and so $\eta^\star(H, G) \leq \eta(H, G)$. As in the classical case, the parameter $\eta^\star_m(H, G)$ is sub-additive (as can easily be derived by its operational interpretation or by matrix manipulations using Definition 8.2.2), hence the parameter $\eta^\star(H, G)$ is well-defined and can equivalently be given by $\inf_m \eta^\star_m(H, G)/m$.

Furthermore, building up on the work presented in this chapter, Cubitt et al. [CMR$^+$14] proved the following bound: $\eta^\star(H, G) \geq \log \vartheta(\overline{H})/\log \vartheta(G)$.

As one would expect, an analog of Lemma 8.2.1 holds.

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

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

PROOF: As $\overline{K_2}^{\boxtimes n}$ is the empty graph on 2^n vertices, one can derive from the definitions that $\eta^\star_m(H, \overline{K}_2) = \lceil \log \chi^\star(H^{\boxtimes m})\rceil$. The identity $R^\star(H) = \eta^\star(H, \overline{K}_2)$ then 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 $\eta^\star_m(K_2, G)$ is the minimum $n \in \mathbb{N}$ such that $\alpha^\star(G^{\boxtimes n}) \geq 2^m$ or, equivalently, $\log \alpha^\star(G^{\boxtimes n}) \geq m$. We can then use the same techniques as in Lemma 8.2.1 to prove the identity $1/c^\star(G) = \eta^\star(K_2, G)$. □

8.2.2 Separate coding schemes for the source-channel problem

Intuitively one can obtain a source-channel coding scheme by concatenating a coding scheme for a source with one for a channel. This is actually an optimal strategy in the setting of asymptotically vanishing error probability [VVS95], meaning that source and channel code design can be dealt separately without asymptotic loss in the code rate in the limit of large block lengths. But when errors cannot be tolerated, 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 $\eta(H,G) \leq R(H)/c(G)$ holds (see Proposition 8.2.5 below), but that for some families of graphs there can be a large separation: $\eta(H,G) \ll R(H)/c(G)$.

To be able to concatenate a source coding scheme with one for a channel, 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 H and a channel with confusability graph G, we need the condition $\chi(H^{\boxtimes m}) \leq \alpha(G^{\boxtimes n})$ in order to send length-m source-input sequences with n uses of the channel. If this condition holds, then it follows that $\eta_m(H,G) \leq n$. The same type of reasoning holds also in the entanglement-assisted case. We have just proved the following simple lemma, which can alternatively be shown using the definition of a graph homomorphism and simple matrix manipulations.

8.2.4. LEMMA. *Given graphs G, H and positive integers n, m, we have*

$$\chi(H^{\boxtimes m}) \leq \alpha(G^{\boxtimes n}) \implies \eta_m(H,G) \leq n, \tag{8.4}$$

$$\chi^\star(H^{\boxtimes m}) \leq \alpha^\star(G^{\boxtimes n}) \implies \eta^\star_m(H,G) \leq n. \tag{8.5}$$

We now relate the minimum cost rate to the ratio of the Witsenhausen rate and the Shannon capacity in both classical and entangled cases.

8.2.5. PROPOSITION. *Let G and H be graphs and assume that both G and $\overline{H}$ have at least one edge. Then,*

$$\eta(H,G) \leq \frac{R(H)}{c(G)} = \lim_{m\to\infty} \frac{1}{m} \min\{n : \chi(H^{\boxtimes m}) \leq \alpha(G^{\boxtimes n})\}, \tag{8.6}$$

$$\eta^\star(H,G) \leq \frac{R^\star(H)}{c^\star(G)} = \lim_{m\to\infty} \frac{1}{m} \min\{n : \chi^\star(H^{\boxtimes m}) \leq \alpha^\star(G^{\boxtimes n})\}. \tag{8.7}$$

PROOF: We show (8.6); we omit the proof of (8.7) which is analogous (and uses (8.5)). Let define $\varepsilon_m(H,G) = \min\{n : \chi(H^{\boxtimes m}) \leq \alpha(G^{\boxtimes n})\}$. From (8.4) we have the inequality $\eta_m(H,G) \leq \varepsilon_m(H,G)$, which in turn implies

$\eta(H,G) \leq \lim_{m\to\infty} \varepsilon_m(H,G)/m$. Next we show that this limit is equal to $R(H)/c(G)$, which concludes the proof of (8.6). Setting $n = \varepsilon_m(H,G)$, we have that $\alpha(G^{\boxtimes(n-1)}) < \chi(H^{\boxtimes m}) \leq \alpha(G^{\boxtimes n})$, implying

$$\frac{R(H)}{c(G)} \leq \frac{\log\chi(H^{\boxtimes m})}{m}\frac{n}{\log\alpha(G^{\boxtimes n})} \leq \frac{n}{m} < \frac{n}{n-1}\frac{\log\chi(H^{\boxtimes m})}{m}\frac{n-1}{\log\alpha(G^{\boxtimes n-1})}.$$

Taking limits as m tends to infinity, in the right most terms we obtain that $R(H)/c(G)$ is equal to $\lim_{m\to\infty} \varepsilon_m(H,G)/m$. □

8.2.3 Separation between classical and entangled source-channel cost rate

We exhibit a family of source-channel instances where an entanglement-assisted strategy allows to reduce the cost rate. We again use properties of the quarter-orthogonality graphs H_k (Definition 6.3.1).

8.2.6. THEOREM. *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,*

$$\eta^\star(H_k, H_k) \leq \frac{\log(k+1)}{(k-1)\left(1 - \frac{2\log(k+1)}{k-3}\right)}, \tag{8.8}$$

$$\eta(H_k, H_k) > \frac{0.154\,k - 1}{k - 1 - \log(k+1)}. \tag{8.9}$$

Let us point out that Theorem 8.2.6 holds for an infinite family of graphs. This follows from the result of Xia and Lu [XL91] in Theorem 6.3.3, 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 $i \in \mathbb{N}$, $5^i = (4+1)^i \equiv 1 \bmod 4$.)

Thus, 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

$$\eta^\star(H_k, H_k) \leq O\Big(\frac{\log k}{k}\Big) \text{ while } \eta(H_k, H_k) \geq \Omega(1).$$

In Section 8.2.2 we mentioned that there are graphs for which a large separation $\eta(H,G) \ll R(H)/c(G)$ is possible [NTR06]. This is however 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 6.3.6).

Lower bound on the entangled source-channel cost rate. The bound (8.8) is obtained as direct application of Proposition 8.2.5.

8.2.7. COROLLARY. *Let k be an odd integer with $k \geq 11$, then*

$$\eta^\star(H_k, H_k) \leq \frac{\log(k+1)}{(k-1)\left(1 - \frac{2\log(k+1)}{k-3}\right)}.$$

PROOF: From Proposition 8.2.5 we know that $\eta^\star(H_k, H_k) \leq R^\star(H_k)/c^\star(H_k)$. We now only have to apply the upper bound on $R^\star(H_k)$ given in Corollary 8.1.5 and the lower bound on $c^\star(H_k)$ of Corollary 6.3.10. □

Upper bound on the source-channel cost rate. The proof of (8.9) relies on the following result, commonly known as the No-Homomorphism Lemma, due to Albertson and Collins [AC85].

8.2.8. LEMMA (ALBERTSON–COLLINS [AC85]). *Let G be a vertex-transitive graph. If there is a homomorphism from H to G, then*

$$\frac{|V(H)|}{\alpha(H)} \leq \frac{|V(G)|}{\alpha(G)}.$$

As observed in [BBG12], the graph H_k is vertex-transitive; indeed, for any $u \in V(H_k)$, the map $v \mapsto u \oplus v$ is an automorphism of H_k. It is easy to see that vertex transitivity is preserved under strong products and complements. Hence, $\overline{H_k^{\boxtimes n}}$ is vertex-transitive for any $n \in \mathbb{N}$.

We will also need the following result about the graphs H_k.

8.2.9. COROLLARY. *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: The statement can be easily derived by combining Proposition 6.3.7 with Lemma 8.1.4. □

8.2.10. LEMMA. *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,*

$$\eta(H_k, H_k) > \frac{0.154\,k - 1}{k - 1 - \log(k+1)}.$$

PROOF: Consider integers $m, n \in \mathbb{N}$ for which $H_k^{\boxtimes m} \longrightarrow \overline{H_k^{\boxtimes n}}$. Applying Lemma 8.2.8, we deduce that

$$\frac{|V(H_k^{\boxtimes m})|}{\alpha(H_k^{\boxtimes m})} \leq \frac{|V(\overline{H_k^{\boxtimes n}})|}{\alpha(\overline{H_k^{\boxtimes n}})} = \frac{|V(\overline{H_k^{\boxtimes n}})|}{\omega(H_k^{\boxtimes n})}. \tag{8.10}$$

From Lemma 6.3.11 we have $\alpha(H_k^{\boxtimes m}) < 2^{0.846\,k\,m}$. Using Corollary 8.2.9 and the fact that $|V(H_k)| = 2^{k-1}$, we get

$$\frac{2^{(k-1)\,m}}{2^{k\,m\,0.846}} < \frac{|V(H_k^{\boxtimes m})|}{\alpha(H_k^{\boxtimes m})} \overset{(8.10)}{\leq} \frac{|V(\overline{H_k^{\boxtimes n}})|}{\omega(H_k^{\boxtimes n})} = \frac{2^{(k-1)\,n}}{(k+1)^n}.$$

After a few elementary algebraic manipulations and taking logarithms, the above inequality implies

$$\frac{n}{m} > \frac{0.154\,k - 1}{k - 1 - \log(k+1)}.$$

□

Chapter 9

Round elimination in communication complexity

In this chapter we tackle two problems arising in communication complexity: the *promise equality* and the *list* problems. We will study their classical and quantum exact communication complexity, making a distinction between one-round protocols (where the communication is unilateral) and multi-round ones (where back and forth communication is allowed). In a promise equality problem (Section 9.2), Alice and Bob must decide if their inputs are equal or not. We give an explicit instance that exhibits an exponential gap between the one- and two-round exact quantum communication complexities, while in the classical scenario one-round protocols are optimal.

In a list problem (Section 9.3), Bob gets a subset of some finite universe, Alice gets an element from Bob's subset, and their goal is for Bob to learn which element Alice was given. We prove that quantum protocols for list problems resist *round elimination*, a phenomenon that works trivially in the classical case (Theorem 9.3.8).

The content of this chapter is based on joint work with Jop Briët, Harry Buhrman, Debbie Leung, and Florian Speelman [BBL+15b].

9.1 Communication complexity

Since its introduction by Yao [Yao79] communication complexity has become a standard model in computational complexity that enjoys a wide variety of connections to other areas in theoretical computer science [KN97]. Here two parties, Alice and Bob, receive inputs x, y from sets $\mathcal{X}, \mathcal{Y}$ (respectively) and need to compute the value $f(x, y)$ of a two-variable function f known to them in advance. Usually each party has insufficient information to solve the problem alone, meaning that they have to exchange information about each others' inputs. (A communication complexity protocol is depicted in Figure 9.1.) The

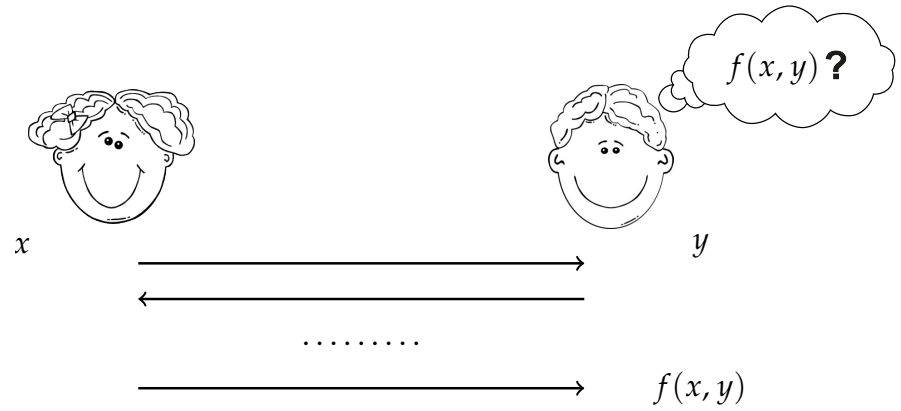

Figure 9.1: A communication complexity protocol.

idea that communication is expensive motivates the study of the *communication complexity* of f, which counts the minimum number of bits that the parties must exchange on worst-case inputs. Throughout, we consider only exact (deterministic) communication protocols, meaning that no error is allowed, and we will omit the word *exact* from now on. Of particular importance here is the distinction between *one-round* protocols, where all communication flows from Alice to Bob, and *multi-round* protocols, where they take turns in sending messages from one party to the other.

In yet another celebrated paper, Yao [Yao93] introduced *quantum communication complexity*, where to compute the value $f(x, y)$ the parties are allowed to transmit qubits back and forth. The study of this model has also become a well-established discipline in theoretical computer science and quantum information theory. The most basic question that arises when considering the classical and quantum models is whether they are actually substantially different. An upper bound on the possible difference between these models was proved by Kremer [Kre95, Theorem 4].[1]

9.1.1. THEOREM (KREMER [KRE95]). *Any quantum protocol that uses ℓ qubits of communication can be turned into a $2^{O(\ell)}$-bit one-round classical protocol for the same problem.*

The first large gap between exact classical and quantum communication complexity was demonstrated by Buhrman, Cleve, and Wigderson [BCW98],

[1] The result stated here is actually a slight generalization of Kremer's result (which focuses on Boolean functions) that can be proved in the same way; for completeness we give a proof in Section 9.4. Moreover, this statement (as well as Kremer's original formulation) holds in the bounded-error model of communication complexity, not only in the exact one.

who gave a problem admitting a one-round quantum protocol that is exponentially more efficient than any (multi-round) classical protocol. In the next section we will consider a generalization of the problem studied in [BCW98].

9.2 Promise equality

In a *promise equality problem*, Alice and Bob each receive an input from a set $\mathcal{X}$ with the promise that their inputs either are equal or come from a subset $\mathcal{D}$ of $\binom{\mathcal{X}}{2}$, where $\mathcal{D}$ is known to them in advance. We denote these problems by the pair $(\mathcal{X}, \mathcal{D})$. The players' goal is to decide whether their inputs are equal or different.

The main result of this section is given by the following theorem.

9.2.1. THEOREM. *There exist absolute constants $c, C \in (0, \infty)$, an infinite sequence of promise equality problems $(\{0,1\}^n, \mathcal{D}_n)_{n \in S}$ with $S \subset \mathbb{N}$ such that for each problem $(\{0,1\}^n, \mathcal{D}_n)$:*

(i) The classical communication complexity is attained with a single round and is at least cn.

(ii) The one-round quantum communication complexity is at least cn.

(iii) There is a two-round quantum protocol using at most $C \log n$ qubits.

The problem we consider is simple. Let n be a positive integer multiple of 8. Alice and Bob are given n-bit strings x and y, respectively, that are either equal or differ in exactly $n/4$ coordinates and they must distinguish between the two cases. We denote this problem by EQ-$\binom{n}{n/4}$. Similar promise equality problems were studied before in [BCW98, GQZ14] and we will briefly mention the known results in Section 9.2.4. An easy observation is that the problem EQ-$\binom{n}{d}$ where n and d have different parities is trivial: Alice can just send the parity bit of her string to Bob. For this reason, here and in the above mentioned works both n and d are assumed to be even numbers.

We will prove Theorem 9.2.1 in Section 9.2.2. Before doing that we observe some useful properties of the promise equality problems.

9.2.1 General properties

To any promise equality problem we associate the graph $G = (\mathcal{X}, \mathcal{D})$ where, as before, $\mathcal{X}$ is the input set and the promise is that either the inputs are equal or they come from the subset $\mathcal{D}$ of $\binom{\mathcal{X}}{2}$.

As it was observed by de Wolf [dW01, Theorem 8.5.1], the one-round classical communication of this problem equals $\lceil \log \chi(G) \rceil$. Indeed, a feasible strategy is that the players agree upon an optimal coloring of the graph beforehand and Alice communicates the color associated to her input, then Bob compares it with the color associated to his input. At the same time, any deterministic strategy gives a coloring of the graph. For general communication problems using more rounds of communication can decrease the total communication. This is for example the case for the Pointer Jumping Problem, where for every positive integer m there is an instance for which any m-round protocol requires exponentially more communication than the best $(m+1)$-round protocol [KN97, Section 4.2]. However, we show that this is not true for promise equality problems, meaning that for such problems the chromatic number not only characterizes the one-round complexity, but their overall communication complexity.

9.2.2. LEMMA. *For any promise equality problem, the classical communication complexity is attained with a single round of communication.*

PROOF: We show how to transform a k-round communication protocol into a one-round protocol that uses the same amount of bits. In a nutshell, the idea is that Alice mimics all the rounds of communication assuming that her input is equal to Bob's, and sends them in one-round. He then checks whether the message received is consistent with his input. If this is not the case, he then knows that the two strings are different, otherwise he completes the protocol.

More formally, fix a protocol Π that requires k rounds, where $k \geq 2$. Suppose that Alice has input x and Bob has y. We assume that the first round of communication is from Alice to Bob, but the same reasoning applies in the other case. For i odd, let a_i be the message that Alice would send to Bob on the i-th round of communication if she followed protocol Π and used both the knowledge of the messages exchanged in the previous rounds and of her input x. Similarly, for i even, let $\hat{b}_i$ be the message that Bob would send to Alice on the i-th round of communication if he had $y = x$ as input, followed the protocol Π and used the knowledge derived by the previous rounds. Using the protocol Π, Alice can mimic Bob's rounds of communication under the assumption that Bob's input is equal to x. Alice uses her input x to produce the string $a_1\hat{b}_2a_3 \dots a_i\hat{b}_{i+1} \dots a_k$ and sends it to Bob in one round. From his input y, Bob constructs the messages b_i that he would have produced during the protocol Π, with the knowledge of Alice's messages a_ℓ and his messages b_ℓ for all $\ell < i$. If there exists an index i such that $b_i \neq \hat{b}_i$, then x must be different from y. Otherwise, Bob uses the string $a_1\hat{b}_2a_3 \dots a_i\hat{b}_{i+1} \dots a_k$ to finish the protocol and either outputs $x = y$ or $x \neq y$. We have constructed a one-round communication protocol Π' that works as the original protocol Π does and that in the worst-case uses at most as many bits as the protocol Π. Therefore if Π is an optimal protocol, so is Π'. □

Let's now consider quantum communication protocols. De Wolf [dW01, Theorem 8.5.2] observed that the one-round quantum communication complexity is characterized by the orthogonal rank of the associated graph. For completeness we include a proof below. Contrary to the classical case, Theorem 9.2.1 shows that allowing additional rounds of quantum communication can be beneficial.

9.2.3. THEOREM (DE WOLF [DW01]). *Consider a promise equality problem defined by the sets $\mathcal{X}$ and $\mathcal{D}$. Then its one-round quantum communication complexity is equal to $\lceil \log \xi(G) \rceil$ where $G = (\mathcal{X}, \mathcal{D})$.*

PROOF: Let Π be an optimal one-round protocol for the considered promise equality problem and let ρ_x be the state that Alice sends on input $x \in \mathcal{X}$. We associate to the state ρ_x a vector $|\phi_x\rangle$ with the property that $|\phi_x\rangle\langle\phi_x|$ is a pure state in the spectral value decomposition of ρ_x. For any pair $(x, y) \in \mathcal{D}$, ρ_x and ρ_y have to be perfectly distinguishable and therefore, in view of Lemma 2.4.1, they must be orthogonal. Equivalently, $|\phi_x\rangle$ and $|\phi_y\rangle$ have to be orthogonal and we can without loss of generality assume that the protocol uses only the pure states $|\phi_x\rangle$. Hence, the map $\phi : \mathcal{X} \to \mathbb{C}^d$ where $\phi(x) = |\phi_x\rangle$ is a d-dimensional orthogonal representation of $G = (\mathcal{X}, \mathcal{D})$ and $\xi(G) \leq d$.

On the other hand, let ϕ be a d-dimensional orthogonal representation of the graph $G = (\mathcal{X}, \mathcal{D})$ and consider the one-round quantum protocol that transmits the normalized vector $\phi(x)/\|\phi(x)\| \in \mathbb{C}^d$ on input $x \in \mathcal{X}$. This uses $\log d$-qubits of communication. From Lemma 2.4.1 we know that Bob can use his input y to perform a quantum measurement that allows him to learn whether his input is equal or not to Alice's. Thus, the one-round quantum communication complexity of this equality problem is at most $\lceil \log \xi(G) \rceil$. □

9.2.2 Proof of Theorem 9.2.1

This section will be devoted to the proof of Theorem 9.2.1, which shows that there is a family of promise equality problems where using two rounds of quantum communication is exponentially more efficient than a single round. The problem that exhibits this separation is EQ-$\binom{n}{n/4}$, where Alice and Bob each receive a n-bit string that are either equal or differ in exactly $n/4$ positions (with n multiple of 8). We denote by $H(n, n/4)$ the graph associated with this problem. In general, with $H(n, d)$ we denote the graph that has $\{0,1\}^n$ as vertex set and where two n-bit strings are adjacent if they differ exactly in d positions. Equivalently, $H(n, d)$ is the graph with vertex set $\{-1, 1\}^n$ where two vertices are adjacent if their inner product is equal to $n - 2d$. We will also use the notion of *adjacency matrix* of a graph G, which is the $|V(G)| \times |V(G)|$ symmetric matrix where the (i, j)-th entry is equal to 1 if $ij \in E(G)$ and to 0 otherwise.

We split the proof in two parts: firstly we bound the one-round quantum communication complexity and secondly we give a two-round protocol.

The classical communication complexity, Theorem 9.2.1 (*i*), can be deduced by combining Lemma 9.2.2 together with the following theorem due to Frankl and Rödl [FR87, Theorem 1.10].

9.2.4. THEOREM (FRANKL AND RÖDL [FR87]). *Let $\alpha \in (0,1)$ and $\alpha n, n$ be even numbers. Then the stability number of the graph $H(n, \alpha n)$ is at most equal to $(2-\varepsilon)^n$ for some positive constant ε.*

9.2.5. COROLLARY. *Let $\alpha \in (0,1)$ and $\alpha n, n$ be even numbers. The classical communication complexity of* EQ-$\binom{n}{\alpha n}$ *is at least $\Omega(n)$.*

PROOF: We get a lower bound on the chromatic number of the graph $H(n, \alpha n)$ using Theorem 9.2.4. Indeed, we have $\chi(H(n,\alpha n)) \geq \frac{|V(H(n,\alpha n))|}{\alpha(H(n,\alpha n))} \geq (\frac{2}{2-\varepsilon})^n$. Taking the logarithm and using Lemma 9.2.2, we can conclude. □

One-round quantum communication complexity of EQ-$\binom{n}{n/4}$

Here we prove the following result, which gives Theorem 9.2.1 (*ii*) as a special case.

9.2.6. THEOREM. *Let $\alpha \in (0,1/2)$ and $\alpha n, n$ be even numbers. The one-round quantum communication complexity of* EQ-$\binom{n}{\alpha n}$ *is at least $\Omega(n)$.*

We obtain this statement by lower bounding the Lovász theta number which itself is a lower bound for the orthogonal rank: $\vartheta(\overline{G}) \leq \xi(G)$ (Lemma 3.2.13). We prove the desired bound in two steps: first, we use structural properties of the graph $H(n,d)$ together with known properties of the Lovász theta number to reformulate this bound in terms of the eigenvalues of the adjacency matrix of this graph; second, we bound the eigenvalues to get the desired result.

Step 1: Eigenvalue bound on the Lovász theta number. We show that the Lovász theta number of the graph $H(n,d)$ can be expressed in terms of the eigenvalues of its adjacency matrix. For the remainder of this step, by the eigenvalues of a graph we mean the eigenvalues of its adjacency matrix.

Lovász [Lov79, Theorems 8 and 9] showed that if a graph is both vertex- and edge-transitive, then the Lovász theta number is given by a simple formula involving its eigenvalues.

9.2.7. LEMMA (LOVÁSZ [LOV79]). *For a positive integer n, let G be an n-vertex graph with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$. If G is both vertex- and edge-transitive, then $\vartheta(\overline{G}) = 1 - \lambda_1/\lambda_n$.*

We now observe that the graph $H(n,d)$ is both vertex- and edge-transitive. We start by showing that $H(n,d)$ is vertex-transitive. Given any pair of vertices $u, v \in \{0,1\}^n$ of $H(n,d)$, consider the automorphism of the graph $H(n,d)$ that maps $x \mapsto x \oplus u \oplus v$ where $\oplus$ is the bit-wise addition. This map preserves the Hamming distance and, therefore, the adjacencies between the vertices. Moreover, it sends $u \mapsto v$ and we can conclude that $H(n,d)$ is vertex-transitive.

To show that $H(n,d)$ is edge-transitive, fix any two edges uv and st and let $p = u \oplus v$, $q = s \oplus t$. Noting that the n-bit strings p and q have the same Hamming weight d, let π be a permutation of the indices such that $\pi(p) = q$. We define ν to be an automorphism that sends a vertex x to $\pi(x \oplus u) \oplus s$. The map ν preserves the edges of $H(n,d)$ and, since the permutation π maps the all-zero string to itself and p to q, we have that $\nu(u) = s$ and $\nu(v) = t$. Thus $H(n,d)$ is edge-transitive.

The following corollary is then a direct application of Lemma 9.2.7.

9.2.8. COROLLARY. *Let $n \in \mathbb{N}$ and $d \in [n]$. Then $\vartheta(\overline{H(n,d)}) = 1 - \binom{n}{d}/\lambda_{\text{MIN}}$ holds where λ_{MIN} is the smallest eigenvalue of $H(n,d)$.*

PROOF: We are only left to observe that, since the largest eigenvalue of a vertex-transitive graph is equal to its degree, we have $\lambda_1(H(n,d)) = \binom{n}{d}$. □

Step 2: Bound on the smallest eigenvalue of $H(n,d)$. We prove an upper bound on the magnitude of the smallest eigenvalue of $H(n,d)$.

9.2.9. LEMMA. *Let n and d be even positive integers such that $d < n/2$. Then, the smallest eigenvalue λ_{MIN} of the graph $H(n,d)$ is a negative number such that*

$$|\lambda_{\text{MIN}}| \leq \sqrt{\frac{2^n \binom{n}{d}}{\binom{n}{n/2-\sqrt{d(n-d)}}}}.$$

The proof of the lemma uses the following facts from coding theory that can be found in the survey of Delsarte and Levenshtein [DL98]. The eigenvalues of $H(n,d)$ play a fundamental role in the theory of Hamming association schemes, where they are expressed in terms of a set of orthogonal polynomials known as the (binary) *Krawtchouk polynomials*. For a positive integer n and $d \in \{0,1,\ldots,n\}$ the Krawtchouk polynomial $K_d^n \in \mathbb{R}[x]$ is a degree-d polynomial that is uniquely defined by

$$K_d^n(x) = \sum_{j=0}^{d} (-1)^j \binom{x}{j}\binom{n-x}{d-j}, \qquad x = 0,1,\ldots,n.$$

When n and d are even, then K_d^n is symmetric about the point $x = n/2$. Moreover, these polynomials satisfy the orthogonality relation

$$\sum_{x=0}^{n} \binom{n}{x} K_d^n(x) K_{d'}^n(x) = \delta_{d,d'} \binom{n}{d} 2^n. \tag{9.1}$$

The set of distinct eigenvalues of $H(n,d)$ turns out to be equal to the set of integer evaluations $\{K_d^n(0), K_d^n(1), \ldots, K_d^n(n)\}$ of the polynomial K_d^n. Crucial to our proof of Lemma 9.2.9 then is the following result of Levenshtein [Lev95, Theorem 6.1] characterizing the smallest roots of the Krawtchouk polynomials.

9.2.10. THEOREM (LEVENSHTEIN [LEV95]). *Let n be a positive integer and $d \in [n]$. Then, K_d^n has exactly d distinct roots and its smallest root is given by*

$$n/2 - \max_{z} \Big(\sum_{i=0}^{d-2} z_i z_{i+1} \sqrt{(i+1)(n-i)} \Big), \tag{9.2}$$

where the maximum is over all vectors $z = (z_0, \ldots, z_{d-1})$ on the real Euclidean unit sphere.

This implies the following general bound on the location of the smallest root of K_d^n. The bound is stated for instance in [KL01] without a proof, we include one here for completeness.

9.2.11. COROLLARY. *Let n and d be positive integers such that $d < n/2$. Then, the smallest root of K_d^n lies in the interval $\big[n/2 - \sqrt{(n-d)d}, n/2\big]$.*

PROOF: Clearly (9.2) is upper bounded by $n/2$. We focus on the lower bound. To this end, let $z = (z_0, \ldots, z_{d-1})$ be a real unit vector achieving the maximum in (9.2). We define $a_i = z_i \sqrt{n-i}$ for any $i \in \{0, 1, \ldots, d-1\}$ and set $b_i = z_{i+1}\sqrt{i+1}$ for any $i \in \{0, 1, \ldots, d-2\}$. Then, we can rewrite the sum as $\sum_{i=0}^{d-2} z_i z_{i+1} \sqrt{(i+1)(n-i)} = \sum_{i=0}^{d-2} a_i b_i$. By the Cauchy-Schwarz inequality,

$$\begin{aligned}
\Big(\sum_{i=0}^{d-2} a_i b_i \Big)^2 &\le \Big(\sum_{i=0}^{d-2} a_i^2 \Big) \Big(\sum_{j=0}^{d-2} b_j^2 \Big) = \Big(\sum_{i=0}^{d-2} a_i^2 \Big) \Big(\sum_{j=1}^{d-1} b_{j-1}^2 \Big) \\
&\le \Big(\sum_{i=0}^{d-1} a_i^2 \Big) \Big(\sum_{j=1}^{d-1} b_{j-1}^2 \Big) \le \Big(\sum_{i=0}^{d-1} z_i^2 (n-i) \Big) \Big(\sum_{j=0}^{d-1} z_j^2 j \Big) \\
&= \Big(n - \sum_{i=0}^{d-1} z_i^2 i \Big) \Big(\sum_{j=0}^{d-1} z_j^2 j \Big),
\end{aligned} \tag{9.3}$$

where in the last equality we used the fact that z is a unit vector. Observe that the sum $\sum_{i=0}^{d-1} z_i^2 i$ lies in the interval $[0, d-1]$. Hence, since $d < n/2$, (9.3) is at most $\max\{(n-t)t : t \in [0, d-1]\} = (n-(d-1))(d-1) \le (n-d)d$. □

PROOF OF LEMMA 9.2.9: As the trace of a matrix equals the sum of its eigenvalues and the trace of an adjacency matrix is zero, it follows that $\lambda_{\text{MIN}} < 0$.

Recall that the eigenvalues of the graph $H(n,d)$ belong to the set $\{K_d^n(x) : x = 0,1,\ldots,n\}$. Moreover, since by assumption n and d are even, the polynomial K_d^n is symmetric about the point $n/2$. Also observe that $K_d^n(0) > 0$ and hence the first time this polynomial assumes a negative value is somewhere beyond its smallest root; i.e., the smallest x for which $K_d^n(x) < 0$ lies in between the smallest root and $n/2$. It therefore follows from Corollary 9.2.11 and from the fact that K_d^n is symmetric about the point $n/2$ that $\lambda_{\text{MIN}} = K_d^n(x^\star)$ for some integer $x^\star \in [n/2 - \sqrt{(n-d)d}, n/2]$.

Clearly (9.1) implies that

$$\sum_{x=0}^{n} \binom{n}{x} K_d^n(x)^2 = \binom{n}{d} 2^n.$$

Hence,

$$\binom{n}{x^\star} K_d^n(x^\star)^2 \leq \binom{n}{d} 2^n$$

and we can conclude that

$$|\lambda_{\text{MIN}}|^2 = |K_d^n(x^\star)|^2 \leq \frac{2^n \binom{n}{d}}{\binom{n}{x^\star}} \leq \frac{2^n \binom{n}{d}}{\binom{n}{n/2-\sqrt{(n-d)d}}}.$$

□

Putting everything together. We are almost ready to prove Theorem 9.2.6. The only missing piece is the following property of the binary entropy function H, which is defined as $H(p) = -p \log p - (1-p)\log(1-p)$ for $p \in [0,1]$.

9.2.12. LEMMA. *For any $p \in (0, 1/2)$, $H(p) + H(1/2 - \sqrt{(1-p)p}) - 1 > 0$.*

The proof of the lemma uses the following lower bound for the function H.

9.2.13. LEMMA. *For any $p \in [0,1]$, we have $H(p) \geq 1 - (1-2p)^2$. Moreover, equality holds if and only if $p \in \{0, 1/2, 1\}$.*

PROOF: The Taylor series of the binary entropy function around the point $1/2$ gives that

$$1 - H(p) = \frac{1}{2\ln 2} \sum_{n=1}^{\infty} \frac{(1-2p)^{2n}}{n(2n-1)} \leq \frac{(1-2p)^2}{2\ln 2} \sum_{n=1}^{\infty} \frac{1}{n(2n-1)} = (1-2p)^2,$$

where the first inequality is due to the fact that $|1-2p| \leq 1$ and therefore that $(1-2p)^{2n} \leq (1-2p)^2$; the last one uses the identity $2\ln 2 = \sum_{n\geq 1} \frac{1}{n(2n-1)}$.

Indeed, the Taylor series for $\ln 2$ around 0 (also known as Mercator series) gives that $\ln 2 = \sum_{n\geq 1} \frac{(-1)^{n+1}}{n} = \sum_{n\geq 1} \frac{1}{2n(2n-1)}$, and multiplying both sides by 2 gives the wanted result.

Therefore, we deduce that $H(p) \geq 1 - (1-2p)^2$. Moreover, equality holds only at the points where $(1-2p)^{2n} = (1-2p)^2$ for every $n \in \mathbb{N}$, which are $p \in \{0, 1/2, 1\}$. □

PROOF OF LEMMA 9.2.12: Using Lemma 9.2.13 and elementary algebraic manipulations, for any $p \in (0, 1/2)$ we have that $H(p) > 4p(1-p)$ and that $H(1/2 - \sqrt{(1-p)p}) > 1 - 4p(1-p)$. But now the statement follows immediately: $H(p) + H(1/2 - \sqrt{(1-p)p}) - 1 > 0$. □

PROOF OF THEOREM 9.2.6: We combine Lemma 3.2.13, Corollary 9.2.8 and Lemma 9.2.9 to obtain

$$\xi(H(n,d)) \geq \vartheta\big(\overline{H(n,d)}\big) \geq 1 - \binom{n}{d} \Big/ \lambda_{\text{MIN}} \geq 1 + \sqrt{\frac{\binom{n}{d}\binom{n}{n/2-\sqrt{(n-d)d}}}{2^n}}. \quad (9.4)$$

We take the logarithm in the above equation and use Stirling's approximation: $\log \binom{n}{k} = \big(H(k/n) + o(1)\big)n$, where H is the binary entropy function and the $o(1)$ term goes to zero as $n \to \infty$ (see for example [SF14, pp. 64]). Then, for $\alpha = d/n$, the logarithm of (9.4) is at least

$$\frac{1}{2}\log\left(\frac{\binom{n}{d}\binom{n}{n/2-\sqrt{(n-d)d}}}{2^n}\right) = \frac{n}{2}\left(H(\alpha) + H\left(1/2 - \sqrt{(1-\alpha)\alpha}\right) - 1 + o(1)\right).$$

By Lemma 9.2.12, $H(\alpha) + H(1/2 - \sqrt{(1-\alpha)\alpha}) - 1 > 0$ for any $\alpha \in (0, 1/2)$ and therefore $\log \xi(H(n, \alpha n)) \geq \Omega(n)$. □

9.2.3 Two-round quantum communication of EQ-$\binom{n}{n/4}$

Using a distributed version of Grover's search algorithm, we find a quantum protocol that solves EQ-$\binom{n}{n/4}$ with a logarithmic number of qubits, which gives Theorem 9.2.1 (*iii*).

9.2.14. THEOREM. *The two-round quantum communication complexity of* EQ-$\binom{n}{n/4}$ *is at most* $2\lceil \log n \rceil + 1$ *qubits.*

PROOF: Let x and y be the inputs of Alice and Bob, respectively, and $z = x \oplus y$ be their bit-wise addition. The promise ensures that either $|z| = 0$ if $x = y$ or $|z| = n/4$ in the case where $x \neq y$.

If a bit string $z \in \{0,1\}^n$ is known to contain exactly $n/4$ entries that are 1, Grover's algorithm [Gro96] is able to find one of these entries without error [BBHT98], needing only a single query to the string z. For any string we define the query unitary $U_z = \sum_{i=1}^n (-1)^{z_i} |i\rangle\langle i|$ and we define $|s\rangle = \frac{1}{\sqrt{n}} \sum_{i=1}^n |i\rangle$ to be the uniform superposition of all basis states. Then $G = 2|s\rangle\langle s| - I$ is a unitary operation known as the Grover diffusion operator.

The quantum communication protocol can be viewed as combining Grover's algorithm with a special case of the simulation theorem given in [BCW98, Theorem 2.1]. We want to perform the algorithm on the effective string $z = x \oplus y$, using the fact that performing a single query U_z is the same as performing the operations U_x and U_y in sequence; i.e., $U_z = U_x U_y = U_y U_x$.

At the start of the protocol, Bob creates the state $U_y|s\rangle$ and sends this state over to Alice using $\lceil \log n \rceil$ qubits. Alice first applies U_x to the incoming state and then applies the Grover operator G. The final state of Grover's algorithm is $\frac{1}{\sqrt{n/4}} \sum_{i \text{ s.t. } z_i=1} |i\rangle$ if $|z| = n/4$. That is, in the case that $x \neq y$, Grover's algorithm has produced a superposition over all indices i such that $x_i \neq y_i$. Alice measures the state, obtaining some index i^* such that $x_{i^*} \neq y_{i^*}$ if $x \neq y$. Then she sends i^* and the value x_{i^*} over to Bob using $\lceil \log n \rceil + 1$ qubits. He outputs 'equal' if and only if $x_{i^*} = y_{i^*}$. The total communication cost of the protocol is then $2\lceil \log n \rceil + 1$ qubits. □

The above protocol can be extended to efficiently solve EQ-$\binom{n}{\alpha n}$ for $\alpha < 1/2$ in a constant number of rounds, by using a more general exact version of the Grover search algorithm. This construction is described in the next section.

9.2.4 Communication complexity of EQ-$\binom{n}{d}$

The promise equality problems were first introduced by Buhrman, Cleve, and Wigderson [BCW98] to show an exponential gap between classical and quantum communication. They used the problem EQ-$\binom{n}{n/2}$, where Alice and Bob get n-bit strings that are either equal or differ in exactly half of the entries (for n multiple of 4). One can easily check that the map $\phi : \{0,1\}^n \to \mathbb{C}^n$ with $\phi(x) = \frac{1}{\sqrt{n}} \sum_{i \in [n]} (-1)^{x_i} e_i$ (where e_i is the i-th canonical basis vector of $\mathbb{C}^n$) is an orthogonal representation of $H(n, n/2)$ and therefore, by Theorem 9.2.3, the one-round quantum communication complexity is at most $\log(n)$ qubits. At the same time, Corollary 9.2.5 says that the classical communication complexity is at least $\Omega(n)$.

Similar results were shown by Gruska, Qiu, and Zheng [GQZ14] for the analogous problem EQ-$\binom{n}{\alpha n}$ for constant $\alpha > 1/2$. Corollary 9.2.5 still applies giving that the classical communication complexity is at least $\Omega(n)$. Moreover, the map $\phi : \{0,1\}^n \to \mathbb{C}^{n+1}$ with $\phi(x) = \sqrt{\frac{1-\gamma^2}{n}} \sum_{i \in [n]} (-1)^{x_i} e_i + \gamma e_{n+1}$, where

$\gamma = 1 - \frac{1}{2\alpha}$, is an orthogonal representation of $H(n, \alpha n)$ and thus the one-round quantum communication complexity is at most $\log(n+1)$ qubits.

The authors of [GQZ14] posed as open problem to determine the quantum communication complexity of EQ-$\binom{n}{\alpha n}$ when $\alpha < 1/2$. (The classical communication complexity is known and again given by Corollary 9.2.5.) Here we prove that for any of these promise equality problems, there is a quantum multi-round protocol that is exponentially more efficient than any single round one. The lower bound on the one-round quantum communication complexity follows from Theorem 9.2.6, while next we give the multi-round protocol.

Multi-round quantum protocols for EQ-$\binom{n}{\alpha n}$ **with** $\alpha < 1/2$

We split the situation in two cases. If $\alpha \in (1/4, 1/2)$ we can simply pad an appropriate number of zeros to both inputs such that the new strings are either equal or differ in exactly $1/4$-th of the positions. Then we simply run the protocol of Theorem 9.2.14.

9.2.15. THEOREM. *Let $\alpha \in (1/4, 1/2)$. The two-round quantum communication complexity of* EQ-$\binom{n}{\alpha n}$ *is at most $2\lceil \log n \rceil + 2\lceil \log(4\alpha) \rceil + 1$ qubits.*

PROOF: Let x and y be Alice's and Bob's inputs. Both of the players pad the input received with $k = 4d - n$ zeros. Hence, the new bit strings x' and y' have length $n' = n + k = 4d$ and they are either equal or differ in $n'/4$ positions. Alice and Bob can now run the communication protocol described in the proof of Theorem 9.2.14 on the new inputs $x', y' \in \{0,1\}^{n'}$. The communication cost is $2\lceil \log n' \rceil + 1 = 2\lceil \log(4\alpha n) \rceil + 1 \leq 2\lceil \log n \rceil + 2\lceil \log(4\alpha) \rceil + 1$ qubits. □

If $\alpha \in (0, 1/4)$, we need to introduce some technicalities to ensure an exact version of Grover's search algorithm.

9.2.16. THEOREM. *Let $\alpha \in (0, 1/4)$. The quantum communication complexity of* EQ-$\binom{n}{\alpha n}$ *is at most $O(\log n)$ qubits and the protocol uses $O(\frac{1}{\sqrt{\alpha}})$ rounds.*

PROOF: If a n-bit string z is known to contain exactly d entries that are 1, Grover's algorithm can be modified such that it finds an index for one of them with certainty [BHMT02, Theorem 16] (see also [BHT98, Amb04]). The number of queries ℓ that the exact version of Grover's algorithm needs in this case is given by

$$\ell = \left\lceil \frac{\pi}{4 \arcsin \sqrt{\frac{d}{n}}} - \frac{1}{2} \right\rceil < \frac{\pi}{4}\sqrt{\frac{n}{d}} + 1 \,.$$

The exact version of Grover's algorithm is the same as the original algorithm except for an adapted final step, which uses a parametrized diffusion operator

$G(\phi)$ and partial query $V_z(\varphi)$ where ϕ and φ are angles that depend on the Hamming distance d. As these angles do not have a nice closed formula, we refer the reader to [BHMT02, Equation (12)] for the relation that ϕ and φ must satisfy. Here

$$V_z(\varphi)|j\rangle = \begin{cases} |j\rangle & \text{if } z_j = 0 \\ e^{i\varphi}|j\rangle & \text{if } z_j = 1 \end{cases}$$

and

$$G(\phi) = F_n V_0(\phi) F_n^* ,$$

where F_n is the $n \times n$ discrete quantum Fourier transform.

Take $x, y \in \{0,1\}^n$ to be the input strings of Alice and Bob, let $z = x \oplus y$ and $d = \alpha n$. As in the proof of the $n/4$ case of Theorem 9.2.14, we turn this search algorithm into a quantum communication protocol by writing a single query $U_z = U_x U_y = U_y U_x$. We can use the commutativity of U_x and U_y to save rounds: The exact Grover's algorithm is performed by executing the operations

$$G(\phi)V_z(\varphi)\underbrace{GU_z \ldots GU_z}_{\ell-1 \text{ times}}$$

on starting state $|s\rangle = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}|i\rangle$. Since we can write two alternations as $GU_zGU_z = GU_xU_yGU_yU_x$, alternating whether Alice or Bob executes the query first that round, only $\ell - 1$ rounds are needed for the $\ell - 1$ ordinary Grover iterations. Alice starts the protocol if ℓ is even, and Bob sends the first message if ℓ is odd.

For the final step, the players need to simulate a query $V_z(\varphi)$ by local operations that depend only on x or y. At this point in the protocol it is Alice's turn to communicate. She currently holds the state

$$|\psi\rangle = \underbrace{GU_z \ldots GU_z}_{\ell-1 \text{ times}}|s\rangle .$$

Now Alice adds an auxiliary qubit that starts in state $|0\rangle$. Define the unitary operation Q_x by its action on the computational basis states as

$$Q_x|j\rangle|b\rangle = |j\rangle|b \oplus x_j\rangle$$

and the (diagonal) unitary matrix $R_y(\varphi)$ as

$$R_y(\varphi)|j\rangle|b\rangle = e^{i\varphi(b\oplus y_j)}|j\rangle|b\rangle .$$

Now Alice first applies Q_x on the state $|\psi\rangle|0\rangle$, sends this state to Bob who performs $R_y(\varphi)$, sending the state back to Alice who again performs Q_x. It is easy to check that $Q_xR_y(\varphi)Q_x|\psi\rangle|0\rangle = (V_z(\varphi)\otimes I)|\psi\rangle|0\rangle$, therefore Alice now

discards the auxiliary qubit and applies $G(\phi)$ to finish the simulation of the exact version of Grover's algorithm.

The final state of the exact Grover's algorithm is $\frac{1}{\sqrt{d}} \sum_{i \text{ s.t. } z_i=1} |i\rangle$ if $|z| = d$. Once Alice has this state in her possession, she performs a measurement in the computational basis, obtaining an index i^* such that $x_{i^*} \neq y_{i^*}$ if $x \neq y$. Then she sends i^* and the value x_{i^*} over to Bob, who outputs 'equal' if and only if $x_{i^*} = y_{i^*}$. This final message consists of $\lceil \log n \rceil + 1$ qubits. By the correctness of the exact Grover's algorithm, this protocol correctly outputs 'not equal' if the Hamming distance between x and y is the fixed value d. Therefore we turned an ℓ-query execution of the exact version of Grover's algorithm into a protocol that uses $(\ell + 2)\lceil \log n \rceil + 2$ qubits of communication in $\ell + 2$ rounds. □

Distances close to $n/2$

The one-round quantum communication of the problem EQ-$\binom{n}{\alpha n}$ is $O(\log n)$ for $\alpha \geq 1/2$ [BCW98, GQZ14], while it is at least $\Omega(n)$ for $\alpha \in (0, 1/2)$. One may wonder whether $1/2$ is exactly the threshold where this exponential jump sits. We show that this is not the case. When α is strictly smaller than $1/2$ but very close to it, the one-round quantum communication complexity of EQ-$\binom{n}{\alpha n}$ still requires only a logarithmic number of qubits.

9.2.17. LEMMA. *Let $d = n/2 - \ell$ with $\ell \leq O(\log n)$ and n, d be even numbers. The one-round quantum communication complexity of* EQ-$\binom{n}{d}$ *is at most $O(\log n)$.*

PROOF: We start by making the following easy observation. Suppose Alice sends to Bob the first 2ℓ bits of her input. If this 2ℓ-bit string differ from Bob's initial part of the input, he knows that the answer is 'not equal'. Otherwise Alice and Bob have to exchange information about the remaining part of their inputs, which have length $n' = n - 2\ell$ and they are either equal or differ in exactly $d' = d = n/2 - \ell = n'/2$ positions.

More formally, consider the map $\phi : \{0,1\}^n \to \mathbb{C}^k$ where $k = 2^{2\ell} n'$ and that sends $x \mapsto x_1 \otimes x_2 \otimes \cdots \otimes x_{2\ell} \otimes \frac{1}{\sqrt{n'}} \sum_{i=1}^{n'} (-1)^{x_{i+2\ell}} e_i$. This is an orthogonal representation of the graph $H(n, d)$. As $\log k$ is $O(\log n)$, the result now follows from Theorem 9.2.3. □

9.3 The list problem

In the *list problem*, inputs are picked from a subset $\mathcal{D} \subseteq \mathcal{X} \times \mathcal{Y}$ and the goal is for Bob to learn Alice's input. The reason for the name "list problem" is that Bob's input y may just as well be given to him as the list (subset) of all of Alice's possible inputs x satisfying $(x, y) \in \mathcal{D}$. A list problem can thus equivalently

be given by a family $\mathcal{L} \subseteq 2^{\mathcal{X}}$ of lists, where Bob gets a list $L \in \mathcal{L}$, Alice gets an element $x \in L$, and Bob must learn x. We refer to this communication problem as $\mathcal{L}$-LIST.

The best general lower bound (due to Orlitsky [Orl90]) and upper bound (due to Naor, Orlitsky, and Shor [NOS93]) on the classical communication complexity of such problems differ only by a constant factor. We exhibit an example showing that, somewhat surprisingly, the four-round protocol used in the bound of Naor et al. [NOS93] can in fact be optimal (Theorem 9.3.4). Furthermore, we show that a phenomenon which works trivially in the classical case does not have a quantum counterpart (Theorem 9.3.8).

9.3.1 Classical communication complexity of list problems

Notice that if one allows only one-round classical protocols, this problem is equivalent to solving one instance of a zero-error source coding problem where the input pair is an element of $\mathcal{D} \subseteq \mathcal{X} \times \mathcal{Y}$. Indeed, Witsenhausen [Wit76] observed that the one-round classical communication complexity of the list problem is characterized by the chromatic number of the graph with vertex set $\mathcal{X}$ and whose edge set consists of the pairs of distinct elements appearing together in some list $L \in \mathcal{L}$. Denoting this graph by $G_{\mathcal{L}}$, the one-round communication complexity equals $\lceil \log \chi(G_{\mathcal{L}}) \rceil$. The multi-round communication complexity of the list problem has also been studied. Orlitsky [Orl90, Corollary 3 and Lemma 3] proved the following lower bound in terms of the chromatic number of $G_{\mathcal{L}}$, and the cardinality of the largest list, denoted

$$\omega(\mathcal{L}) = \max\{|L| : L \in \mathcal{L}\}$$

(not to be confused with the cardinality of the largest clique $\omega(G_{\mathcal{L}})$, which can be larger).

9.3.1. THEOREM (ORLITSKY [ORL90]). *For every family $\mathcal{L} \subseteq 2^{\mathcal{X}}$, the classical communication complexity of $\mathcal{L}$-LIST is at least* $\max\{\log\log \chi(G_{\mathcal{L}}), \log \omega(\mathcal{L})\}$.

The basic idea behind the above result is that any multi-round protocol can be simulated by a one-round protocol with at most an exponential difference in communication, and that Alice must send sufficient information for Bob to be able to distinguish among $\omega(\mathcal{L})$ elements. In the same work, Orlitsky [Orl90, Theorem 4] gave a two-round classical protocol based on perfect hashing functions that nearly achieves the above lower bound.

9.3.2. THEOREM (ORLITSKY [ORL90]). *For any $\mathcal{L} \subseteq 2^{\mathcal{X}}$, the two-round classical communication complexity of $\mathcal{L}$-LIST is at most* $\log\log \chi(G_{\mathcal{L}}) + 3\log \omega(\mathcal{L}) + 4$.

It thus follows from Witsenhausen's observation and Theorem 9.3.2 that list problems have exponentially more efficient two-round protocols than one-round protocols, provided that $\omega(\mathcal{L}) \leq poly(\log \chi(G_{\mathcal{L}}))$. But Theorem 9.3.1 shows that—in stark contrast with the Pointer Jumping Problem—using more than two rounds cannot decrease the total communication by more than a factor of 4, as $\log\log\chi(G_{\mathcal{L}}) + 3\log\omega(\mathcal{L}) \leq 4\max\{\log\log\chi(G_{\mathcal{L}}), \log\omega(\mathcal{L})\}$. Furthermore, in a follow up work, Orlitsky [Orl91] showed that in general two-round protocols are not sufficient to reach the communication complexity. The natural question that thus arises is: Can the lower bound of Theorem 9.3.1 be attained by using more than two rounds of communication?

Towards answering this question Naor, Orlitsky, and Shor [NOS93, Corollary 1] slightly improved on Theorem 9.3.2 and showed that the four-round communication complexity gets to within a factor of about 3 of the lower bound.

9.3.3. THEOREM (NAOR–ORLITSKY–SHOR [NOS93]). *For every family $\mathcal{L} \subseteq 2^{\mathcal{X}}$, the four-round classical communication complexity of the $\mathcal{L}$-*LIST *problem is at most* $\log\log\chi(G_{\mathcal{L}}) + 2\log\omega(\mathcal{L}) + 3\log\log\omega(\mathcal{L}) + 7$.

Our contribution to this line of work is to show that, perhaps surprisingly, for some list problems the four-round protocol of Naor, Orlitsky, and Shor is in fact asymptotically optimal, thus answering the above question in the negative.

9.3.4. THEOREM. *For any $\varepsilon > 0$ there exist a set $\mathcal{X}$ and a family $\mathcal{L} \subseteq 2^{\mathcal{X}}$ such that the classical communication complexity of $\mathcal{L}$-*LIST *is at least*

$$\log\log\chi(G_{\mathcal{L}}) + (2-\varepsilon)\log\omega(\mathcal{L}).$$

Moreover, there exists such an $(\mathcal{X}, \mathcal{L})$ pair for which $\omega(\mathcal{L}) = \log\chi(G_{\mathcal{L}})$.

In particular, our result gives a family of list problems with communication complexity at least $(3-\varepsilon)\max\{\log\log\chi(G_{\mathcal{L}}), \log\omega(\mathcal{L})\}$ for any $\varepsilon > 0$.

Proof of Theorem 9.3.4. The list problem that we use for the proof of Theorem 9.3.4 is simple. For positive integers k, N such that $2 \leq k \leq N$, we consider the list problem $\mathcal{L} = \binom{[N]}{k}$, where the family of lists consists of all k-element subsets of $[N]$. Note that for this $\mathcal{L}$, $G_{\mathcal{L}}$ is the complete graph on N vertices, giving $\chi(G_{\mathcal{L}}) = N$, and we have $\omega(\mathcal{L}) = k$ (not to be confused with $\omega(G_{\mathcal{L}}) = N$). Hence, Theorem 9.3.3 gives a four-round protocol using at most $\log\log N + 2\log k + O(\log\log k)$ bits of communication.

9.3.5. THEOREM. *The classical communication complexity of $\binom{[N]}{k}$-*LIST *is at least*

$$\log\log N + 2\log(k-1) - \log\log(k-1) - O(1).$$

To see that this implies Theorem 9.3.4 note that the above bound can be written as $\log\log\chi(G_{\mathcal{L}}) + (2 - o(1))\log\omega(\mathcal{L})$, where the term $o(1)$ goes to zero as k tends to infinity. Choosing $k = \log N$ then gives the second part of the theorem.

To prove Theorem 9.3.5, we use a bound on the size of cover-free families due to Dýachkov and Rykov [DR82]; see [Rus94, Für96] for simplified proofs (in English).

9.3.6. DEFINITION. Let r be a positive integer and $\mathcal{S}$ be a finite set. A family $\mathcal{F} \subseteq 2^{\mathcal{S}}$ of at least $r+1$ subsets is *r-cover-free* if every subfamily of $r+1$ distinct sets $F_0, F_1, \ldots, F_r \in \mathcal{F}$ satisfies $F_0 \nsubseteq F_1 \cup \cdots \cup F_r$.

9.3.7. THEOREM (DÝACHKOV–RYKOV [DR82]). *There exists an absolute constant $c > 0$ such that the following holds. Let r, N be positive integers such that $r \geq 2$ and $N \geq r + 1$. Let $\mathcal{S}$ be a finite set and $\mathcal{F} \subseteq 2^{\mathcal{S}}$ be an r-cover free family consisting of N sets. Then,*

$$|\mathcal{S}| \geq \frac{cr^2 \log N}{\log r}.$$

PROOF OF THEOREM 9.3.5: For a positive integer C, suppose that the communication complexity of $\binom{[N]}{k}$-LIST is C. Fix an optimal protocol Π. For every possible input pair (x, L) in the $\binom{[N]}{k}$-LIST problem, define the transcript $T_{x,L} \in \{0,1\} \cup \{0,1\}^2 \cup \cdots \cup \{0,1\}^C$ to be the concatenation of the parties' messages in the order they are sent during their conversation on input (x, L). Let $\mathcal{T}$ be the set of said transcripts.

For each transcript $T \in \mathcal{T}$, denote by T^A the sequence of Alice's messages in T, to be understood as a sequence of strings indexed by her rounds in the conversation. Let $\mathcal{F} = \{F_x\}_{x \in \mathcal{X}} \subseteq 2^{\mathcal{T}}$ be the family where each F_x is the collection of transcripts $T \in \mathcal{T}$ that is consistent with x being Alice's input and that agrees on T^A. We claim that $\mathcal{F}$ is a $(k-1)$-cover free family. To see this, take any k sets of $\mathcal{F}$, say $F_{x_0}, \ldots, F_{x_{k-1}}$, and let $\widetilde{L}$ be the corresponding k-element list $\{x_0, \ldots, x_{k-1}\}$. Consider the transcript $T_{x_0,\widetilde{L}}$ related to the input pair $(x_0, \widetilde{L})$. Clearly, $T_{x_0,\widetilde{L}} \in F_{x_0}$. We show that $T_{x_0,\widetilde{L}} \notin F_{x_i}$ for each $i \in \{1, \ldots, k-1\}$, which gives the claim as this implies that $F_{x_0} \nsubseteq F_{x_1} \cup \cdots \cup F_{x_{k-1}}$. Suppose that $T_{x_0,\widetilde{L}} \in F_{x_i}$ holds for some $i \in \{1, \ldots, k-1\}$. This means that Alice sends identical message sequences on inputs x_0 and x_i and therefore that Bob is not able to distinguish between these two cases for the input pair $(x_0, \widetilde{L})$, contradicting our assumption that we started with a functional protocol.

We also claim that $\mathcal{F}$ consists of at least N sets. Indeed, for every pair $x, y \in [N]$, there is a list $L \in \binom{[N]}{k}$ containing both x and y. Since we must have that $T^A_{x,L} \neq T^A_{y,L}$ in order for Bob to be able to distinguish between x and y on input L, the inputs x and y induce distinct transcript sets.

It thus follows from Theorem 9.3.7 that the total number of distinct transcripts is at least

$$|\mathcal{T}| \geq \frac{c(k-1)^2 \log N}{\log(k-1)},$$

for some absolute constant $c > 0$. Now as $\mathcal{T} \subseteq \{0,1\} \cup \{0,1\}^2 \cup \cdots \cup \{0,1\}^C$, we have

$$\frac{2^{C+1}-1}{2-1} = \sum_{l=0}^{C} 2^l \geq \frac{c(k-1)^2 \log N}{\log(k-1)}.$$

Taking logarithms we get the claim. □

9.3.2 Quantum communication complexity of list problems and quantum round elimination

We show that list problems have the interesting property of resisting a quantum analog of *round elimination*. This peculiar phenomenon is also shown by Briët and Zuiddam [BZ16] using a similar reasoning.

In classical communication complexity, round elimination reduces the number of rounds of a given protocol by having the parties send some extra information instead. Consider the following basic example, where we start with a two-round $(\log n + 1)$-bit protocol in which Bob starts by sending Alice a single bit and Alice replies with a $\log n$-bit string. This protocol can easily be turned into a *one-round* $2 \log n$-bit protocol by having Alice directly send Bob two $\log n$-bit strings, one corresponding to the case where Bob sends a 0 in the two-round protocol and another for if he sends a 1. Then Bob can just pick the string corresponding to the bit he would have sent based on his input and solve the problem.

Surprisingly a quantum analog of this phenomenon does not hold.

9.3.8. THEOREM. *There exist an absolute constant $c \in (0, \infty)$, an infinite sequence of list problems $(\mathcal{K}_n)_{n \in S}$ with $S \subseteq \mathbb{N}$ such that for each $\mathcal{K}_n$-*LIST *problem:*

(i) The one-round quantum communication complexity is at least cn.

(ii) There is a two-round quantum protocol where one single qubit is transmitted in the first round and the second round consists of a $(\log n + 1)$-qubit message.

The sequence of lists that we consider is simple. For an even positive integer n and $d \in [n]$, let $\mathcal{L}_d \subseteq 2^{\{0,1\}^n}$ be the family of lists $L \subseteq \{0,1\}^n$ of maximal cardinality such that each pair of strings in L have Hamming distance exactly d. Consider the family of lists given by $\mathcal{K}_n = \mathcal{L}_{n/2} \cup \cdots \cup \mathcal{L}_n$.

Before proving Theorem 9.3.8, we observe some useful properties of quantum communication complexities for list problems.

General properties. We previously mentioned that Witsenhausen [Wit76] observed that the chromatic number of an appropriate graph characterizes the one-round communication complexity of the list problem. Similarly, the one-round quantum communication complexity of a list problem is characterized in terms of the orthogonality dimension of its associated graph. This result is similar to Theorem 9.2.3 and, indeed, the proof is along the same lines.

9.3.9. LEMMA. *For every family $\mathcal{L} \subseteq 2^{\mathcal{X}}$, the one-round quantum communication complexity of $\mathcal{L}$-*LIST *equals to $\lceil \log \xi(G_{\mathcal{L}}) \rceil$.*

PROOF: Consider an optimal one-round protocol. With the same reasoning as in Theorem 9.2.3, we can assume, without loss of generality, that Alice sends to Bob a pure state $|\phi_x\rangle \in \mathbb{C}^d$ on input $x \in \mathcal{X}$. Then, given a list $L \in \mathcal{L}$, Bob has a measurement that allows him to distinguish the states $\{|\phi_x\rangle : x \in L\}$. It thus follows from Lemma 2.4.1 that these states must be orthogonal. In particular, since for every list $L \in \mathcal{L}$, each pair of distinct elements $x, y \in L$ forms an edge in $G_{\mathcal{L}}$, the vectors $|\phi_x\rangle$, $x \in \mathcal{X}$, form a d-dimensional orthogonal representation. Hence, $\xi(G_{\mathcal{L}}) \leq d$.

Conversely, let $\phi : V(G_{\mathcal{L}}) \to \mathbb{C}^d$ be an orthogonal representation of $G_{\mathcal{L}}$. Then, for every list $L \in \mathcal{L}$, the vectors $\{\phi(x) : x \in L\}$ are pairwise orthogonal. If Bob gets a list $L \in \mathcal{L}$ and Alice gets an element $x \in L$, it follows from Lemma 2.4.1 that there is a quantum measurement allowing Bob to uniquely identify x when Alice sends $\phi(x)$ using $\log d$-qubits. Hence, the one-round quantum communication complexity is at most $\lceil \log \xi(G_{\mathcal{L}}) \rceil$. □

For multi-round protocols, a quantum analog of Theorem 9.3.1 also holds.

9.3.10. LEMMA. *For every family $\mathcal{L} \subseteq 2^{\mathcal{X}}$, the quantum communication complexity of $\mathcal{L}$-*LIST *is at least* $\max\{\Omega(\log\log \chi(G_{\mathcal{L}})), \log \omega(\mathcal{L})\}$.

PROOF: Kremer's Theorem (Theorem 9.1.1) shows that there is at most an exponential difference between the (multi-round) quantum and one-round classical communication complexity. Hence, by Witsenhausen's result, the former is at least $\Omega(\log\log \chi(G_{\mathcal{L}}))$. Moreover, on the worst input Bob has to be able to distinguish among $\omega(\mathcal{L})$ different elements. Hence, $\log \omega(\mathcal{L})$ bits of information must be communicated and Holevo's Theorem [Hol73] says that to retrieve $\log \omega(\mathcal{L})$ bits of information $\log \omega(\mathcal{L})$ qubits are necessary. □

Proof of Theorem 9.3.8. Recall that we are considering the following family of lists. For an even positive integer n and $d \in [n]$, let $\mathcal{L}_d \subseteq 2^{\{0,1\}^n}$ be the family of all lists $L \subseteq \{0,1\}^n$ of maximal cardinality such that all strings in L have pairwise Hamming distance d. We denote by $\mathcal{K}_n$ the union $\mathcal{L}_{n/2} \cup \cdots \cup \mathcal{L}_n$. In other words, this is the union of lists for which, individually, there is a one-round $O(\log n)$-qubit protocol (see Section 9.2.4).

We start by proving Theorem 9.3.8 (*i*). In view of Lemma 9.3.9, we have to lower bound the orthogonal rank of the graph G_n, where $G_n = (\{0,1\}^n, E)$ and E is given by all pairs of strings with Hamming distance in $\{n/2, \ldots, n\}$. We derive that $\log \xi(G_n) \geq \Omega(n)$ by combining the fact that $\xi(G) \geq \vartheta(\overline{G})$ (Lemma 3.2.13) together with the following lower bound on $\vartheta(\overline{G_n})$ proven by Samorodnitsky in an unpublished note [Sam98, Lemma 3.3].

9.3.11. THEOREM (SAMORODNITSKY [SAM98]). *For the graph G_n, we have that $\vartheta(\overline{G_n}) \geq 2^{(1-H(1/4))n-o(n)}$, where H is the binary entropy function.*

Indeed, taking the logarithm, we obtain the following chain of inequalities: $\log \xi(G_n) \geq \log \vartheta(\overline{G}_n) \geq (1 - H(1/4) - o(1))n \approx (0.189 - o(1))n$ where the term $o(1)$ goes to zero as $n \to \infty$. (In the above mentioned work of Briët and Zuiddam [BZ16], the authors give an independent and easier proof of a lower bound on $\vartheta(\overline{G_n})$ analogous to the one of Theorem 9.3.11.) Thus the one-round quantum communication of $\mathcal{K}_n$-LIST problem is at least $\Omega(n)$ and Theorem 9.3.8 (*i*) follows.

Secondly, we give a simple two-round protocol for $\mathcal{K}_n$-LIST which implies Theorem 9.3.8 (*ii*).

9.3.12. THEOREM. *For $\mathcal{K}_n = \mathcal{L}_{n/2} \cup \cdots \cup \mathcal{L}_n$, there exists a two-round protocol for $\mathcal{K}_n$-LIST where Bob sends to Alice a single qubit and Alice replies with a $(\log n + 1)$-qubit message.*

PROOF: Let $\ell = \lceil \log n \rceil$ and U be the $(\ell+1)$-qubit unitary matrix which satisfies $U|0\rangle|0\rangle^{\otimes \ell} = |0\rangle|0\rangle^{\otimes \ell}$ and $U|1\rangle|0\rangle^{\otimes \ell} = \frac{1}{\sqrt{n}}|1\rangle \sum_{i=1}^{n} |i\rangle$. Moreover, for any 2^ℓ-bit string z, we define the conditional query unitary U_z which acts on the computational basis states as $U_z|0\rangle|i\rangle = |0\rangle|i\rangle$ and $U_z|1\rangle|i\rangle = (-1)^{z_i}|1\rangle|i\rangle$ for any $i \in [2^\ell]$. For a small technicality if n is not a power of 2, i.e., $\ell > \log n$, we will map any n-bit string to a 2^ℓ-bit string obtained by padding zeros to the original string. We can now explain the protocol.

Fix an input pair (x, L) where $L \in \mathcal{L}_d$ for some $d \in \{n/2, \ldots, n\}$. Bob looking at the list L learns d and sends to Alice the single qubit $\gamma|0\rangle + \sqrt{1-\gamma^2}|1\rangle$ where $\gamma^2 = 1 - \frac{n}{2d} \geq 0$. Alice pads the state $|0\rangle^{\otimes \ell}$ to the one she received and then applies in sequence the unitaries U and U_x, obtaining the state

$$\begin{aligned} |\phi_x\rangle &= U_x U \left((\gamma|0\rangle + \sqrt{1-\gamma^2}|1\rangle)|0\rangle^{\otimes \ell} \right) \\ &= \gamma|0\rangle|0\rangle^{\otimes \ell} + \sqrt{\frac{1-\gamma^2}{n}} \sum_{i=1}^{n} (-1)^{x_i}|1\rangle|i\rangle. \end{aligned}$$

She sends this to Bob using $\lceil \log n \rceil + 1$ qubits. Notice that if $x, y \in \{0,1\}^n$ differ in exactly d positions, then the states $|\phi_x\rangle$ and $|\phi_y\rangle$ are orthogonal to each other. Hence, by Lemma 2.4.1, using the list L Bob can perform a measurement that allows him to learn Alice's input x. □

9.3.3 Entanglement-assisted and non-signaling communication complexity of the list problem

We end the chapter by presenting two results about the list problem when the players can only exchange classical bits but they are allowed to share non-classical correlations.

Quantum correlations. If Alice and Bob can share an entangled state and communicate classical bits, then they can use the teleportation protocol of Bennett et al. [BBC$^+$93] to simulate the quantum communication with a factor of 2 overhead. However, there may be more efficient protocols. In particular, we show for the $\mathcal{K}_n$-LIST problem, where $\mathcal{K}_n$ is the union $\mathcal{L}_{n/2} \cup \cdots \cup \mathcal{L}_n$, a two-round entanglement-assisted protocol that uses only $\lceil \log n \rceil + 3$ classical bits.

9.3.13. LEMMA. *For $\mathcal{K}_n = \mathcal{L}_{n/2} \cup \cdots \cup \mathcal{L}_n$, there exists a two-round entanglement-assisted protocol for $\mathcal{K}$-LIST where Bob sends Alice a single bit and Alice replies with $\lceil \log n \rceil + 2$ bits of communication.*

PROOF: Let (x, L) be Alice and Bob's input pair where $L \in \mathcal{L}_d$. Consider the conditional query unitary U_z, where z is a n-bit string, which acts on the computational basis states as $U_z|0\rangle|i\rangle = |0\rangle|i\rangle$ and $U_z|1\rangle|i\rangle = (-1)^{z_i}|1\rangle|i\rangle$ for any $i \in [n]$. We show a two-round communication protocol that uses as shared entanglement the state $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} |i\rangle_A |i\rangle_B$ together with two EPR pairs. We use the subscript A (respectively B) to specify Alice's part of the state (respectively Bob's).

From the list L, Bob learns the distance d and uses one EPR pair and one bit of communication to remote state prepare the qubit $\gamma|0\rangle + \sqrt{1-\gamma^2}|1\rangle$, where $\gamma^2 = 1 - \frac{n}{2d} \geq 0$ [Pat00, Lo00]. Now Alice and Bob are sharing the entangled state: $\left(\gamma|0\rangle_A + \sqrt{1-\gamma^2}|1\rangle_A\right) \frac{1}{\sqrt{n}} \sum_{i=1}^{n} |i\rangle_A |i\rangle_B$. Using her input x, Alice performs the unitary U_x followed by the unitary $U = |0\rangle\langle 0| \otimes F_n + |1\rangle\langle 1| \otimes F_n$ where F_n is the $n \times n$ discrete quantum Fourier transform. The entangled state is now:

$$\gamma|0\rangle_A \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \omega_n^{ij} |j\rangle_A |i\rangle_B + (\sqrt{1-\gamma^2})|1\rangle_A \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} (-1)^{x_i} \omega_n^{ij} |j\rangle_A |i\rangle_B,$$

where ω_n is the n-th root of unity. Alice measures in the computational basis her second n-qubit register and gets an outcome $\hat{j}$. She sends $\hat{j}$ to Bob using $\lceil \log n \rceil$ classical bits. Moreover, Alice teleports the qubit $\gamma|0\rangle + \sqrt{1-\gamma^2}|1\rangle$ to Bob using the protocol of Bennett et al. [BBC$^+$93]. This requires two classical bits and an EPR pair.

He can then use $\hat{j}$ to perform a unitary $|0\rangle\langle 0| \otimes U_{\hat{j}} + |1\rangle\langle 1| \otimes U_{\hat{j}}$, where $U_{\hat{j}} = \sum_{i=1}^{n} \omega_n^{i\hat{j}*} |i\rangle\langle i|$, that will put his register in the state

$$|\psi_x\rangle = \gamma|0\rangle \frac{1}{\sqrt{n}} \sum_{i=1}^{n} |i\rangle + (\sqrt{1-\gamma^2})|1\rangle \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (-1)^{x_i} |i\rangle.$$

At last, we notice that if $x, y \in \{0,1\}^n$ differ in exactly d bits then $|\psi_x\rangle$ is orthogonal to $|\psi_y\rangle$. Using the elements of the list L, by Lemma 2.4.1, Bob can construct a measurement that allows him to learn Alice's input. In total the protocol required entanglement and $\lceil \log n \rceil + 3$ bits of classical communication. □

Non-signaling correlations. If the two parties can share non-signaling correlations (Definition 2.4.5), every list problem becomes trivial.

9.3.14. LEMMA. *For every family $\mathcal{L} \subseteq 2^{\mathcal{X}}$, there is a one-round non-signaling protocol that uses only $\lceil \log \omega(\mathcal{L}) \rceil$ bits of communication and this is optimal.*

PROOF: Fix a list $L \in \mathcal{L}$ and an element $x \in L$. Uniquely label the elements of L with numbers in $\mathbb{Z}_{\omega(\mathcal{L})}$ and let i be the label assigned to x. Let $P(\cdot,\cdot|x,L)$ be the probability distribution over $\mathbb{Z}_{\omega(\mathcal{L})} \times \mathbb{Z}_{\omega(\mathcal{L})}$ that assigns probability $1/\omega(\mathcal{L})$ to each pair in $\{(a, a+i) : a \in \mathbb{Z}_{\omega(\mathcal{L})}\}$ and vanishes on all other pairs. Clearly this distribution is non-signaling. Similarly define non-signalling distributions for every other pairs (x', L') in the list problem.

Consider the following protocol. Upon receiving $x \in \mathcal{X}$ and $L \in \mathcal{L}$ such that $x \in L$, Alice and Bob sample from the distribution $P(\cdot,\cdot|x,L)$ as explained above and get a and $a+i \in \mathbb{Z}_{\omega(\mathcal{L})}$, respectively. Next, Alice sends a to Bob, using at most $\lceil \log \omega(\mathcal{L}) \rceil$ bits of communication. Finally, Bob subtracts Alice's message from his input, getting $(a+i) - a = i$, which tells him Alice's input.

At last, we notice that any functional protocol has to communicate at least $\lceil \log \omega(\mathcal{L}) \rceil$ bits and hence the above protocol is an optimal one. Indeed, there is an instance of the problem where Bob has to distinguish Alice's input from a list of $\omega(\mathcal{L})$ different elements. □

9.4 Kremer's Theorem

Here we prove Kremer's Theorem (Theorem 9.1.1), which we restate for convenience. The original proof by Kremer [Kre95] applied to Boolean functions; we give a slight generalization of the statement so that it applies to functions with arbitrary range. It is important to notice that the statements in this section hold for general communication protocols, not only exact ones.

9.4.1. THEOREM. *Let ℓ be a positive integer, $X, Y, \mathcal{R}$ be finite sets and $\mathcal{D} \subseteq X \times Y$. Let $f : \mathcal{D} \to \mathcal{R}$ be a function and suppose that f admits an ℓ-qubit quantum protocol. Then, there exists a one-round $2^{O(\ell)}$-bit classical protocol for f.*

The proof uses the following lemma of Yao [Yao93] and Kremer [Kre95]. To reduce the amount of notation needed in the proof we assume that the parties use the following general protocol. At any point during the protocol, both Alice and Bob have a private quantum register. If it is Alice's turn to communicate, say ℓ qubits, she appends a fresh ℓ-qubit register to her existing register, applies a unitary to both registers and sends the ℓ-qubit register over to Bob, who then absorbs the ℓ-qubit register into his private register. If it's his turn to communicate, Bob operates similarly. This assumption will allow us to deal more easily with protocols in which different numbers of qubits are sent in each round.

9.4.2. LEMMA (YAO–KREMER). *Let ℓ be a positive integer, $X, Y, \mathcal{R}$ be finite sets and $\mathcal{D} \subseteq X \times Y$. Suppose that there exists an r-round quantum protocol for a function $f : \mathcal{D} \to \mathcal{R}$, where ℓ_i qubits are communicated in round $i \in [r]$. Then, the final state of the protocol on input $(x, y) \in \mathcal{D}$ can be written as*

$$\sum \alpha_{\mathbf{u}}(x)\beta_{\mathbf{u}}(y)|A_{\mathbf{u}}(x)\rangle|B_{\mathbf{u}}(y)\rangle,$$

where the sum is over all $\mathbf{u} \in \{0,1\}^{\ell_1} \times \cdots \times \{0,1\}^{\ell_r}$, the $\alpha_{\mathbf{u}}(x)$, $\beta_{\mathbf{u}}(y)$ are complex numbers and the $|A_{\mathbf{u}}(x)\rangle$, $|B_{\mathbf{u}}(y)\rangle$ are complex unit vectors.

PROOF: By induction on r. The base case $r = 1$ is trivial, since then Alice sends Bob an ℓ-qubit state. For some $i \in \{2, 3, \ldots, r\}$, suppose that after $i - 1$ rounds the state is given by

$$\sum \alpha_{\mathbf{v}}(x)\beta_{\mathbf{v}}(y)|A_{\mathbf{v}}(x)\rangle|B_{\mathbf{v}}(y)\rangle,$$

where the sum is over all $\mathbf{v} \in \{0,1\}^{\ell_1} \times \cdots \times \{0,1\}^{\ell_{i-1}}$. Assume that the i-th round is Alice's turn (the case of Bob's turn is handled similarly). She appends a fresh ℓ_i-qubit register to her current register, causing the state to become

$$\sum \alpha_{\mathbf{v}}(x)\beta_{\mathbf{v}}(y)|A_{\mathbf{v}}(x)\rangle \underbrace{|0\rangle \ldots |0\rangle}_{\ell_i \text{ times}} |B_{\mathbf{v}}(y)\rangle.$$

Next, she applies a unitary over both of her registers, turning the state into

$$\sum \alpha_{\mathbf{v}}(x)\beta_{\mathbf{v}}(y) \left(\sum_{\mathbf{w} \in \{0,1\}^{\ell_i}} \gamma_{\mathbf{w}}|A_{\mathbf{v},\mathbf{w}}(x)\rangle|\mathbf{w}\rangle \right) |B_{\mathbf{v}}(y)\rangle,$$

where $\gamma_{\mathbf{w}}$ is a complex number (which might depend on x) and for some unit vectors $|A_{\mathbf{v},\mathbf{w}}(x)\rangle$. Now define

$$\alpha_{\mathbf{v},\mathbf{w}}(x) = \alpha_{\mathbf{v}}(x)\gamma_{\mathbf{w}}, \qquad \beta_{\mathbf{v},\mathbf{w}}(y) = \beta_{\mathbf{v}}(y) \qquad \text{and} \qquad |B_{\mathbf{v},\mathbf{w}}(y)\rangle = |\mathbf{w}\rangle|B_{\mathbf{v}}(y)\rangle,$$

so that after the i-th round, after Alice has sent the ℓ_i-qubit register to Bob, the state equals

$$\sum_{\mathbf{v},\mathbf{w}} \alpha_{\mathbf{v},\mathbf{w}}(x)\beta_{\mathbf{v},\mathbf{w}}(y)|A_{\mathbf{v},\mathbf{w}}(x)\rangle|B_{\mathbf{v},\mathbf{w}}(y)\rangle.$$

After r rounds the state thus looks like as claimed in the lemma. □

PROOF OF THEOREM 9.1.1: Assume that the protocol proceeds in r rounds and that ℓ_i qubits are communicated during round $i \in [r]$. By Lemma 9.4.2 the final state of the protocol can be written as

$$\sum \alpha_{\mathbf{u}}(x)\beta_{\mathbf{u}}(y)|A_{\mathbf{u}}(x)\rangle|B_{\mathbf{u}}(y)\rangle,$$

To produce his output, Bob performs a measurement $\{M_1, \ldots, M_k\}$ on his register. For each pair $\mathbf{u}, \mathbf{v} \in \{0,1\}^{\ell_1} \times \cdots \times \{0,1\}^{\ell_r}$ and $j \in [k]$ we define the complex numbers

$$\begin{aligned} a_{\mathbf{u},\mathbf{v}}(x) &= \overline{\alpha_{\mathbf{u}}(x)}\alpha_{\mathbf{v}}(x)\langle A_{\mathbf{u}}(x)|A_{\mathbf{v}}(x)\rangle \\ b^j_{\mathbf{u},\mathbf{v}}(x) &= \overline{\beta_{\mathbf{u}}(y)}\beta_{\mathbf{v}}(y)\langle B_{\mathbf{u}}(y)|M_j|B_{\mathbf{v}}(y)\rangle. \end{aligned}$$

Then, the probability that Bob gets measurement outcome j equals

$$p_j(x,y) = \sum_{\mathbf{u},\mathbf{v}} a_{\mathbf{u},\mathbf{v}}(x)b^j_{\mathbf{u},\mathbf{v}}(y).$$

The classical one-round protocol works in the following way. Let ℓ be the total communication of the protocol and define $\tilde{a}_{\mathbf{u},\mathbf{v}}(x)$ as an approximation of $a_{\mathbf{u},\mathbf{v}}(x)$ using $2\ell+4$ bits for the real part and $2\ell+4$ bits for the imaginary part, so that $|\tilde{a}_{\mathbf{u},\mathbf{v}}(x) - a_{\mathbf{u},\mathbf{v}}(x)| \leq 2^{-2\ell-3}$. Alice's message consists of all $2^{2\ell}$ numbers $\tilde{a}_{\mathbf{u},\mathbf{v}}(x)$, making the total communication cost $O(\ell 2^{2\ell})$ bits. Bob calculates his approximation of the probability of getting outcome j as

$$\tilde{p}_j(x,y) = \sum_{\mathbf{u},\mathbf{v}} \tilde{a}_{\mathbf{u},\mathbf{v}}(x)b^j_{\mathbf{u},\mathbf{v}}(y).$$

We can bound the difference between this approximation and the acceptance probability of the original quantum protocol by

$$\begin{aligned} |\tilde{p}_j(x,y) - p_j(x,y)| &= \Big|\sum_{\mathbf{u},\mathbf{v}} \big(\tilde{a}_{\mathbf{u},\mathbf{v}}(x) - a_{\mathbf{u},\mathbf{v}}(x)\big)b^j_{\mathbf{u},\mathbf{v}}(y)\Big| \\ &\leq \sum_{\mathbf{u},\mathbf{v}} \big|\tilde{a}_{\mathbf{u},\mathbf{v}}(x) - a_{\mathbf{u},\mathbf{v}}(x)\big|\,\big|b^j_{\mathbf{u},\mathbf{v}}(y)\big| \\ &\leq 2^{-2\ell-3}2^{2\ell} \leq \frac{1}{8}. \end{aligned}$$

Therefore, given a quantum protocol with sufficiently high success probability, here in particular probability 1, Bob can (deterministically) choose the unique outcome j for which $\tilde{p}_j(x, y)$ is strictly greater than $\frac{1}{2}$, and this outcome j is equal to the function value $f(x, y)$, by correctness of the original quantum protocol. □

Bibliography

[AC85] M. O. Albertson and K. L. Collins. Homomorphisms of 3-chromatic graphs. *Discrete Mathematics*, 54(2):127–132, 1985.

[ADR82] A. Aspect, J. Dalibard, and G. Roger. Experimental test of Bell's inequalities using time-varying analyzers. *Physical Review Letters*, 49(25):1804–1807, 1982.

[AHKS06] D. Avis, J. Hasegawa, Y. Kikuchi, and Y. Sasaki. A quantum protocol to win the graph colouring game on all Hadamard graphs. *IEICE Transactions Fundamentals of Electronics, Communications and Computer Sciences*, 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. *IEEE Transactions on Information Theory*, 52(5):2172–2176, 2006.

[AL07] N. Alon and E. Lubetzky. Privileged users in zero-error transmission over a noisy channel. *Combinatorica*, 27(6):737–743, 2007.

[Alo98] N. Alon. The Shannon capacity of a union. *Combinatorica*, 18(3):301–310, 1998.

[Amb04] A. Ambainis. Quantum search algorithms. *SIGACT News*, 35(2):22–35, 2004.

[ARU97] S. Ayupov, A. Rakhimov, and S. Usmanov. *Jordan, real and Lie structures in operator algebras*, volume 418 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1997.

[Bar01] F. Barioli. Chains of dog-ears for completely positive matrices. *Linear Algebra and its Applications*, 330(1–3):49–66, 2001.

[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. *Physical Review Letters*, 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*, 110(48):19227–19232, 2012.

[BBHT98] M. Boyer, G. Brassard, P. Høyer, and A. Tapp. Tight bounds on quantum searching. *Fortschritte der Physik*, 46(4–5):493–505, 1998.

[BBL+15a] J. Briët, H. Buhrman, M. Laurent, T. Piovesan, and G. Scarpa. Entanglement-assisted zero-error source-channel coding. *IEEE Transactions on Information Theory*, 61(2):1124–1138, 2015.

[BBL+15b] J. Briët, H. Buhrman, D. Leung, T. Piovesan, and F. Speelman. Round elimination in exact quantum communication complexity. In *Proceedings of the 10th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2015)*, pages 206–225, 2015.

[BBR94] D. Barrington, R. Beigel, and S. Rudich. Representing Boolean functions as polynomials modulo composite numbers. *Computational Complexity*, 4(4):367–382, 1994.

[BCW98] H. Buhrman, R. Cleve, and A. Wigderson. Quantum vs. classical communication and computation. In *Proceedings of the 30th Annual ACM Symposium on Theory of Computing (STOC 1998)*, pages 63–68, 1998.

[BdK02] I. M. Bomze and E. de Klerk. Solving standard quadratic optimization problems via linear, semidefinite and copositive programming. *Journal of Global Optimization*, 24:163–185, 2002.

[BDKS14] S. Burgdorf, K. Dykema, I. Klep, and M. Schweighofer. Correction of a proof in "Connes' embedding conjecture and sums of hermitian squares". *Advances in Mathematics*, 252:805–811, 2014.

[BDVS+01] C. H. Bennett, D. P. Di Vincenzo, P. W. Shor, J. A. Smolin, B. M. Terhal, and W. K. Wootters. Remote state preparation. *Physical Review Letters*, 87:077902, 2001.

[Bei10] S. Beigi. Entanglement-assisted zero-error capacity is upper-bounded by the Lovász theta function. *Physical Review A*, 82:10303–10306, 2010.

[Bel64] J. S. Bell. On the Einstein-Podolsky-Rosen paradox. *Physics*, 1:195–200, 1964.

[BFS15] M. Berta, O. Fawzi, and V. B. Scholz. Quantum bilinear optimization. *arXiv:1506.08810*, 2015.

[BHMT02] G. Brassard, P. Høyer, M. Mosca, and A. Tapp. Quantum amplitude amplification and estimation. In *Quantum computation and information*, volume 305 of *Contemporary Mathematics*, pages 53–74. American Mathematical Society, Providence (RI), 2002.

[BHT98] G. Brassard, P. Høyer, and A. Tapp. Quantum counting. In *Proceedings of the 25th International Colloquium on Automata, Languages and Programming (ICALP'98)*, pages 820–831, 1998.

[BLP15] S. Burgdorf, M. Laurent, and T. Piovesan. On the closure of the completely positive semidefinite cone and linear approximations to quantum coloring. In *Proceedings of the 10th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2015)*, pages 127–146, 2015.

[BO08] N. P. Brown and N. Ozawa. *C^*-algebras and finite-dimensional approximations*, volume 88 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence (RI), 2008.

[BSM03] A. Berman and N. Shaked-Monderer. *Completely Positive Matrices*. World Scientific Publishing, River Edge (NJ), 2003.

[BSST02] C. H. Bennett, P. Shor, J. A. Smolin, and A. V. Thapliyal. Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem. *IEEE Transactions on Information Theory*, 48(10):2637–2655, 2002.

[BTN01] A. Ben-Tal and A. Nemirovski. *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. MPS-SIAM Series on Optimization. SIAM, 2001.

[Bur11] S. Burgdorf. *Trace-positive polynomials, sums of hermitian squares and the tracial moment problem*. PhD thesis, Universtät Konstanz and Université de Rennes 1, 2011.

[BV04] S. P. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[BZ16] J. Briët and J. Zuiddam. On the orthogonal rank of Cayley graphs and impossibility of quantum round elimination. *arXiv:1608.06113*, 2016.

[CD08] B. Collins and K. Dykema. A linearization of Connes' embedding problem. *New York Journal of Mathematics*, 14:617–641, 2008.

[Che97] Y. Chen. On the existence of abelian Hadamard difference sets and a new family of difference sets. *Finite fields and their Applications*, 3(3):234–256, 1997.

[CHSH69] J. F. Claurer, M. A. Horne, A. Shimony, and R. A. Holt. Proposed experiment to test local hidden-variable theories. *Physical Review Letters*, 23(15):880–884, 1969.

[CHTW04] R. Cleve, P. Høyer, B. Toner, and J. Watrous. Consequences and limits of nonlocal strategies. In *Proceedings of the 19th Annual IEEE Conference on Computational Complexity (CCC 2004)*, pages 236–249, 2004.

[Chv73] V. Chvátal. Edmonds polytopes and a hierarchy of combinatorial problems. *Discrete Mathematics*, 4(4):305–337, 1973.

[CKS90] G. Cohen, J. Körner, and G. Symonyi. *Zero-error capacities and very different sequences*, pages 144–155. Springer-Verlag, 1990.

[CLMW10] T. S. Cubitt, D. Leung, W. Matthews, and A. Winter. Improving zero-error classical communication with entanglement. *Physical Review Letters*, 104(23):230503, 2010.

[CLMW11] T. S. Cubitt, D. Leung, W. Matthews, and A. Winter. Zero-error channel capacity and simulation assisted by non-local correlations. *IEEE Transactions on Information Theory*, 57(8):5509–5523, 2011.

[CMN$^+$07] P. J. Cameron, A. Montanaro, M. W. Newman, S. Severini, and A. Winter. On the quantum chromatic number of a graph. *The Electronic Journal of Combinatorics*, 14:R81, 2007.

[CMR$^+$14] T. S. Cubitt, L. Mančinska, D. E. Roberson, S. Severini, D. Stahlke, and A. Winter. Bounds on entanglement assisted source-channel coding via the Lovász theta number and its variants. *IEEE Transactions on Infomation Theory*, 60(11):7330–7344, 2014.

[Con76] A. Connes. Classification of injective factors. Cases $\Pi_1, \Pi_\infty, \Pi_\lambda$, $\lambda \neq 1$. *Annals of Mathematics*, 104(2):73–115, 1976.

[Dia62] P. H. Diananda. On nonnegative forms in real variables some or all of which are nonnegative. *Mathematical Proceedings of the Cambridge Philosophical Society*, 58(1):17–25, 1962.

[DJL94] J. H. Drew, C. R. Johnson, and R. Loewy. Completely positive matrices associated with M-matrices. *Linear and Multilinear Algebra*, 37(4):303–310, 1994.

[dKLP05] E. de Klerk, M. Laurent, and P. Parrilo. On the equivalence of algebraic approaches to the minimization of forms on the simplex. In *Positive Polynomials in Control*, volume 312 of *Lecture Notes in Computer Science*, pages 121–133. Springer, 2005.

[dKLP06] E. de Klerk, M. Laurent, and P. Parrilo. A PTAS for the minimization of polynomials of fixed degree over the simplex. *Theoretical Computer Science*, 361(2–3):210–225, 2006.

[dKP02] E. de Klerk and D. Pasechnik. Approximation of the stability number of a graph via copositive programming. *SIAM Journal on Optimization*, 12(4):875–892, 2002.

[DL98] P. Delsarte and V. I. Levenshtein. Association schemes and coding theory. *IEEE Transactions on Information Theory*, 44(6):2477–2504, 1998.

[DR82] A. G. Dýachkov and V. V. Rykov. Bounds on the length of disjunctive codes. *Problemy Peredachi Informatsii*, 18(3):7–13, 1982. In Russian.

[DR10] I. Dukanovic and F. Rendl. Copositive programming motivated bounds on the stability and the chromatic numbers. *Mathematical Programming*, 121(2):249–268, 2010.

[DSW13] R. Duan, S. Severini, and A. Winter. Zero-error communication via quantum channels, noncommutative graphs, and a quantum Lovász number. *IEEE Transactions on Information Theory*, 59(2):1164–1174, 2013.

[dW01] R. de Wolf. *Quantum Computing and Communication Complexity*. PhD thesis, Universiteit van Amsterdam, 2001.

[EPR35] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Physical Review*, 47:777–780, 1935.

[FB75] M. J. Ferguson and D. W. Bailey. Zero-error coding for correlated sources. Unpublished manuscript, 1975.

[FGP+15] H. Fawzi, J. Gouveia, P. Parrilo, R. Z. Robinson, and R. R. Thomas. Positive semidefinite rank. *Mathematical Programming*, 153(1):133–177, 2015.

[FHS14] I. Farah, B. Hart, and D. Sherman. Model theory of operator algebras II: model theory. *Israel Journal of Mathematics*, 201(1):477–505, 2014.

[FILG11] J. Fukawa, H. Imai, and F. Le Gall. Quantum coloring games via symmetric SAT games. *Asian Conference on Quantum Information Science*, 2011.

[FK98] U. Feige and J. Kilian. Zero knowledge and the chromatic number. *Journal of Computer and System Sciences*, 57(2):187–199, 1998.

[FR87] P. Frankl and V. Rödl. Forbidden intersections. *Transactions of the American Mathematical Society*, 300(1):259–286, 1987.

[Fri12] T. Fritz. Tsirelson's problem and Kirchberg's conecture. *Reviews in Mathematical Physics*, 24(5):1250012, 2012.

[Für96] Z. Füredi. On r-cover-free families. *Journal of Combinatorial Theory, Series A*, 73(1):172–173, 1996.

[FW81] P. Frankl and R. Wilson. Intersection theorems with geometric consequences. *Combinatorica*, 1(4):357–368, 1981.

[FW14] P. E. Frenkel and M. Weiner. On vector configurations that can be realized in the cone of positive matrices. *Linear Algebra and its Applications*, 459:465–474, 2014.

[GKV94] L. Gargano, J. Körner, and U. Vaccaro. Capacities: from information theory to extremal set theory. *Journal of Combinatorial Theory, Series A*, 68(2):296–316, 1994.

[GL08] N. Gvozdenović and M. Laurent. The operator Ψ for the chromatic number of a graph. *SIAM Journal on Optimization*, 19(2):572–591, July 2008.

[GLS88] M. Grötschel, L. Lovász, and A. Schrijver. *Geometric Algorithms and Combinatorial Optimization*. Springer, 1988.

[Gop06] P. Gopalan. Constructing Ramsey graphs from Boolean function representations. In *Proceedings of the 21th Annual IEEE Conference on Computational Complexity (CCC 2006)*, pages 14–128, 2006.

[GQZ14] J. Gruska, D. Qiu, and S. Zheng. Generalizations of the distributed Deutsch-Jozsa promise problem. *arXiv:1402.7254*, 2014.

[Gro96] L. Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the 28th Annual ACM Symposium on Theory of Computing (STOC 1996)*, pages 212–219, 1996.

[Hae78] W. Haemers. An upper bound for the Shannon capacity of a graph. *Colloquia Mathematica Societatis János Bolyai*, 25:267–272, 1978.

[Hal86] M. Hall. *Combinatorial Theory*. Wiley, second edition, 1986.

[Hås99] J. Håstad. Clique is hard to approximate within $n^{1-\varepsilon}$. *Acta Mathematica*, 182(1):105–142, 1999.

[HBD+15] B. Hensen, H. Bernien, A. E. Dreau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. L. Vermeulen, R. N. Schouten, C. Abellan, W. Amaya, V. Pruneri, M. W. Mitchell, M. Markham, D. J. Twitchen, D. Elkouss, S. Wehner, T. H. Taminiau, and R. Hanson. Loophole-free bell inequality violation using electron spins separated by 1.3 kilometres. *Nature*, 526(7575):682–686, 2015.

[HIK11] R. Hammack, W. Imrich, and S. Klavžar. *Handbook of Product Graphs*. Discrete Mathematics and its Applications. CRC Press, Boca Raton (FL), second edition, 2011.

[HJ12] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, second edition, 2012.

[HJL96] C. L. Hamilton-Jester and C.-K. Li. Extreme vectors of doubly nonnegative matrices. *Rocky Mountain Journal of Mathematics*, 26(4):1371–1383, 1996.

[Hol73] A. S. Holevo. Bounds for the quantity of information transmitted by a quantum communication channel. *Problems of Information Transmission*, 9:177–183, 1973.

[Ji13] Z. Ji. Binary constraint system games and locally commutative reductions. *arXiv:1310:3794*, 2013.

[JNP+11] M. Junge, M. Navascués, C. Palazuelos, V. B. Pérez-Garcia, D. Scholz, and F. Werner. Connes' embedding problem and Tsirelson's problem. *Journal of Mathematical Physics*, 52(1):012102, 2011.

[Kar72] R. Karp. Reducibility among combinatorial problems. In *Complexity of Computer Computations*, pages 85–103. Springer US, 1972.

[KB93] N. Kogan and A. Berman. Characterization of completely positive graphs. *Discrete Mathematics*, 114(1–3):297–304, 1993.

[KL01] I. Krasikov and S. Litsyn. *Survey of binary Krawtchouk polynomials*, pages 199–211. American Mathematical Society, 2001.

[KN97] E. Kushilevitz and N. Nisan. *Communication Complexity*. Cambridge University Press, 1997.

[Knu94] D. E. Knuth. The Sandwich Theorem. *The Electronic Journal of Combinatorics*, 1(1):1–48, 1994.

[KO98] J. Körner and A. Orlitsky. Zero-error information theory. *IEEE Transactions on Information Theory*, 44(6):2207–2229, 1998.

[Kre95] I. Kremer. Quantum communication. Master's thesis, Hebrew University, Computer Science Department, 1995.

[KS08] I. Klep and M. Schweighofer. Connes' embedding conjecture and sums of hermitian squares. *Advances in Mathematics*, 217(4):1816–1837, 2008.

[Lau14] M. Laurent. Private communications, 2014.

[Lev95] V. I. Levenshtein. Krawtchouk polynomials and universal bounds for codes and designs in Hamming spaces. *IEEE Transactions on Infomation Theory*, 41(5):1303–1321, 1995.

[Li03] B. Li. *Real Operator Algebras*. World Scientific Publishing, River Edge (NJ), 2003.

[LMM+12] D. Leung, L. Mančinska, W. Matthews, M. Ozols, and A. Roy. Entanglement can increase asymptotic rates of zero-error classical communication over classical channels. *Communications in Mathematical Physics*, 311:97–111, 2012.

[Lo00] H.-K. Lo. Classical-communication cost in distributed quantum-information processing: a generalization of quantum-communication complexity. *Physical Review A*, 62(1):012313, 2000.

[Lov72] L. Lovász. Normal hypergraphs and the perfect graph conjecture. *Discrete Mathematics*, 2(3):253–267, 1972.

[Lov75] L. Lovász. On the ratio of optimal integral and fractional covers. *Discrete Mathematics*, 13:383–390, 1975.

[Lov78] L. Lovász. Kneser's conjecture, chromatic number, and homotopy. *Journal of Combinatorial Theory, Series A*, 25(3):319–324, 1978.

[Lov79] L. Lovász. On the Shannon capacity of a graph. *IEEE Transactions on Information Theory*, 25(1):1–7, 1979.

[LP15] M. Laurent and T. Piovesan. Conic approach to quantum graph parameters using linear optimization over the completely positive semidefinite cone. *SIAM Journal on Optimization*, 25(4):2461–2493, 2015.

[LPU95] M. Larsen, J. Propp, and D. Ullman. The fractional chromatic number of Mycielski's graphs. *Journal of Graph Theory*, 19:411–416, 1995.

[Lub07] E. Lubetzky. *Graph powers and related extremal problems*. PhD thesis, Tel Aviv University, 2007.

[LY94] C. Lund and M. Yannakakis. On the hardness of approximating minimization problems. *Journal of the ACM*, 41(5):960–981, 1994.

[MAG06] Ll. Masanes, A. Acin, and N. Gisin. General properties of nonsignaling theories. *Physical Review A*, 73:012112, 2006.

[MM62] J. E. Maxfield and H. Minc. On the Matrix Equation X'X = A. *Proceedings of the Edinburgh Mathematical Society (Series 2)*, 13:125–129, 12 1962.

[MR14] L. Mančinska and D. E. Roberson. Note on the correspondence between quantum correlations and the completely positive semidefinite cone. 2014. Available online at http://quantuminfo.quantumlah.org/memberpages/laura/corr.pdf.

[MR15] L. Mančinska and D. E. Roberson. Private communications, 2015.

[MR16] L. Mančinska and D. E. Roberson. Quantum homomorphisms. *Journal of Combinatorial Theory, Series B*, 118(C):228–267, 2016.

[MRR78] R. J. McEliece, E. R. Rodemich, and H. C. Rumsey. The Lovász bound and some generalizations. *Journal of Combinatorics, Information & System Sciences.*, 3(3):134–152, 1978.

[MS65] T. S. Motzin and E. G. Straus. Maxima for graphs and a new proof of a theorem of Turán. *Canadian Journal of Mathematics*, 17:533–540, 1965.

[MSS13] L. Mančinska, G. Scarpa, and S. Severini. New separations in zero-error channel capacity through projective Kochen-Specker and quantum coloring. *IEEE Transactions on Information Theory*, 59(6):4025–4032, 2013.

[MvN36] F. J. Murray and J. von Neumann. On rings of operators. *Annals of Mathematics*, 37(1):116–229, 1936.

[NC00] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[NOS93] M. Naor, A. Orlitsky, and P. Shor. Three results on interactive communication. *IEEE Transactions on Information Theory*, 39:1608–1615, 1993.

[NPA08] M. Navascués, S. Pironio, and Acín. A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations. *New Journal of Physics*, 10:073013, 2008.

[NTR06] J. Nayak, E. Tuncel, and K. Rose. Zero-error source-channel coding with side information. *IEEE Transactions on Infomation Theory*, 52(10):4626–4629, 2006.

[Orl90] A. Orlitsky. Worst-case interactive communication I: Two messages are almost optimal. *IEEE Transactions on Information Theory*, 36(5):1111–1126, 1990.

[Orl91] A. Orlitsky. Worst-case interactive communication II: Two messages are not optimal. *IEEE Transactions on Information Theory*, 37(4):995–1005, 1991.

[Oza13] N. Ozawa. About the Connes embedding conjecture. *Japanese Journal of Mathematics*, 8(1):147–193, 2013.

[Pal33] R. E. A. C. Paley. On orthogonal matrices. *Journal of Mathematics and Physics (now called Studies in Applied Mathematics)*, 12:311–320, 1933.

[Par00] P. Parrilo. *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization*. PhD thesis, California Institute of Technology, 2000.

[Pat00] A. Pati. Minimum cbits for remote state preparation and measurement of a qubit. *Physical Review A*, 63(1):014302, 2000.

[Pee96] M. J. P. Peeters. Orthogonal representations over finite fields and the chromatic number of graphs. *Combinatorica*, 16(3):417–431, 1996.

[PLM+11] R. Prevedel, Y. Lu, W. Matthews, R. Kaltenbaek, and K. J. Resch. Entanglement-enhanced classical communication over a noisy classical channel. *Physical Review Letters*, 106:110505, 2011.

[PSS15] T. Piovesan, G. Scarpa, and C. Schaffner. Multiparty zero-error classical channel coding with entanglement. *IEEE Transactions on Information Theory*, 61(2):1113–1123, 2015.

[PSS+16] V. I. Paulsen, S. Severini, D. Stahlke, I. G. Todorov, and A. Winter. Estimating quantum chromatic numbers. *Journal of Functional Analysis*, 270(6):2188–2222, 2016.

[PT15] V. I. Paulsen and I. G. Todorov. Quantum chromatic numbers via operator systems. *The Quarterly Journal of Mathematics*, 66(2):677–692, 2015.

[Răd99] F. Rădulescu. Convex sets associated with von Neumann algebras and Connes' approximate embedding problem. *Mathematical Research Letters*, 6:229–236, 1999.

[Rob13] D. E. Roberson. *Variations on a theme: Graph Homomorphisms*. PhD thesis, University of Waterloo, 2013.

[Rus94] M. Ruszinkó. On the upper bound of the size of the r-cover-free families. *Journal of Combinatorial Theory, Series A*, 66(2):302–310, 1994.

[Sam98] A. Samorodnitsky. Extremal properties of solutions for Delsarte's linear program. Unpublished manuscript, 1998.

[Sch79] A. Schrijver. A comparison of the Delsarte and Lovász bounds. *IEEE Transactions on Information Theory*, 25(4):425–429, 1979.

[Sch91] K. Schmüdgen. The K-moment problem for compact semi-algebraic sets. *Mathematische Annalen*, 289:203–206, 1991.

[Sch03] A. Schrijver. *Combinatorial Optimization: Polyhedra and Efficiency*. Springer-Verlag, 2003.

[SF14] J. Spencer and L. Florescu. *Asymptopia*. Student Mathematical Library. American Mathematical Society, 2014.

[Sha48] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379—423 and 623—656, 1948.

[Sha56] C. E. Shannon. The zero error capacity of a noisy channel. *IRE Transactions on Information Theory*, 2(3):8–19, 1956.

[Sin09] B. Sinaimeri. *Structures of diversity*. PhD thesis, Sapienza University, Rome, 2009.

[Slo16] W. Slofstra. Tsirelson's problem and an embedding theorem for groups arising from non-local games. *arXiv:1606:03140*, 2016.

[SR02] R. W. Spekkens and T. Rudolph. Optimization of coherent attacks in generalizations of the BB84 quantum bit commitment protocol. *Quantum Information & Computation*, 2(1):66–96, 2002.

[SS12] G. Scarpa and S. Severini. Kochen-Specker sets and the rank-1 quantum chromatic number. *IEEE Transactions on Information Theory*, 58(4):2524–2529, 2012.

[SV15] J. Sikora and A. Varvitsiotis. Linear conic formulations for two-party correlations and values of nonlocal games. *arXiv:1506.07297*, 2015.

[SW74] D. Slepian and J. Wolf. Noiseless coding of correlated information sources. *IEEE Transactions on Information Theory*, 19(4):471–480, 1974.

[Sze94] M. Szegedy. A note on the theta number of Lovász and the generalized Delsarte bound. In *Proceedings of the 35th Annual Symposium on Foundations of Computer Science (FOCS 1994)*, pages 36–39, 1994.

[Tak03] M. Takesaki. *Theory of Operator Algebras II*, volume 125 of *Encyclopaedia of Mathematical Sciences on Operator Algebras and Noncommutative Geometry*. Springer-Verlag, Berlin, 2003.

[Viz63] V. G. Vizing. The cartesian product of graphs. *Viycisl. Sistemy*, 9:30–43, 1963. In Russian.

[VVS95] S. Vembu, S. Verdu, and Y. Steinberg. The source-channel separation theorem revisited. *IEEE Transactions on Information Theory*, 41(1):44–54, 1995.

[Wat11] J. Watrous. Lecture notes on Theory of Quantum Information. Available online at https://cs.uwaterloo.ca/ watrous/LectureNotes.html, 2011.

[WCD08] S. Wehner, M. Christandl, and A. C. Doherty. A lower bound on the dimension of a quantum system given measured data. *Physical Review A*, 78:062112, 2008.

[Wit76] H. S. Witsenhausen. The zero-error side information problem and chromatic numbers. *IEEE Transactions on Information Theory*, 22(5):592–593, 1976.

[Xia96] M. Xia. A new family of supplementary difference sets and Hadamard matrices. *Journal of Statistical Planning and Inference*, 51(3):283–291, 1996.

[XL91] M. Xia and G. Liu. An infinite class of supplementary difference sets and Williamson matrices. *Journal of Combinatorial Theory, Series A*, 58(2):310–317, 1991.

[XSX06] T. Xia, J. Seberry, and M. Xia. New constructing of regular Hadamard matrices. In *Proceedings of the 10th WSEAS international conference on Computers (CCOMP 2006)*, pages 1294–1299, 2006.

[Yao79] A. C.-C. Yao. Some complexity questions related to distributive computing (preliminary report). In *Proceedings of the 11th Annual ACM Symposium on Theory of Computing (STOC 1979)*, pages 209–213, 1979.

[Yao93] A. C.-C. Yao. Quantum circuit complexity. In *Proceedings of the 34th Annual Symposium on Foundations of Computer Science (FOCS 1993)*, pages 352–361, 1993.

[Yek12] S. Yekhanin. Locally decodable codes. *Foundations and Trends in Theoretical Computer Science*, 6(3):139–255, 2012.

[Yil12] E. A. Yildirim. On the accuracy of uniform polyhedral approximations of the copositive cone. *Optimization Methods and Software*, 27(1):155–173, 2012.

[ZZ02] B. Zeng and P. Zhang. Remote-state preparation in higher dimension and the parallelizable manifold S^{n-1}. *Physical Review A*, 65(2):022316, 2002.

Index

List of symbols

$\mathbb{N}$	Set of positive integers.
$[n]$	Set of the elements $\{1, 2, \ldots, n\}$.
$\mathcal{P}(\mathbb{N})$	Collection of all the subsets of $\mathbb{N}$.
$\Pi(n)$	Symmetric group over $[n]$.
δ_{ij}	Kronecker delta function which is equal to 0 if $i \neq j$ and to 1 otherwise.
log	Logarithm in base 2.
$O(f(n))$	Function that is asymptotically upper bounded by f (up to a constant factor).
$\Omega(f(n))$	Function that is asymptotically lower bounded by f (up to a constant factor).
$\Theta(f(n))$	Function that is asymptotically both upper and lower bounded by f (up to a constant factor).
$o(f)$	Function that is asymptotically upper bounded by f.
$\omega(f)$	Function that is asymptotically lower bounded by f.

Vectors

$\mathbb{R}^n$	Set of real n-vectors.
$\mathbb{R}^n_+$	Set of nonnegative real n-vectors.
$\mathbb{C}^n$	Set of complex n-vectors.
e	Vector of all ones.
e_i	Vector with i-th entry equal to 1 and all others equal to 0.
$\langle x, y \rangle$	Inner product between vectors $x, y \in \mathbb{C}^n$, $\langle x, y \rangle = x^* y = \sum_i \overline{x(i)} y(i)$.
$\|x\|$	Norm of a vector $x \in \mathbb{C}^n$, $\|x\| = \sqrt{\langle x, x \rangle}$.
Δ_n	Standard simplex $\Delta_n = \{x \in \mathbb{R}^n_+ : \sum_{i=1}^n x_i = 1\}$

Matrices

A^T	Transpose of the matrix A.
A^*	Conjugate transpose of the matrix A.
I	Identity matrix.
J	Matrix whose entries are all ones.
E_{ij}	Matrix with a 1 in positions (i, j) and (j, i) and 0 elsewhere.
$\mathrm{Diag}(x)$	Diagonal matrix whose main diagonal is vector x.
$A[I]$	Let $A \in \mathcal{S}^n$, $I \subset [n]$, then $A[I]$ is the matrix obtained from A by removing the rows and columns not indexed by I.
$\mathrm{vec}(A)$	Vector obtained from A by stacking its columns on top of each other.
$A \succeq 0$	Matrix A is positive semidefinite.
$\mathsf{Tr}(A)$	Trace of a matrix A, $\mathsf{Tr}(A) = \sum_i A_{ii}$.
$\langle A, B\rangle$	Inner product between two matrices $A, B \in \mathbb{C}^{n\times n}$, $\langle A, B\rangle = \mathsf{Tr}(A^*B) = \sum_{ij} \overline{A_{ij}} B_{ij}$.
$\|A\|_{\mathrm{F}}$	Frobenius norm of A, $\|A\|_{\mathrm{F}} = \sqrt{\langle A, A\rangle}$.
$\|A\|_2$	Spectral norm of A, $\|A\|_2 = \sqrt{\lambda_{\max}(A^*A)}$ where $\lambda_{\max}$ is the largest eigenvalue.
$\|A\|_{\mathrm{op}}$	Operator norm of A, $\|A\|_{\mathrm{op}} = \sup\{\|Ax\| : x \text{ is a unit vector}\}$.
$A \oplus B$	Direct sum of matrices A and B.
$A \circ B$	Entrywise product of matrices A and B, where $(A \circ B)_{ij} = A_{ij}B_{ij}$.
$A \otimes B$	Tensor product (also known as Kronecker product) of matrices A and B.
$\mathbb{R}_+^{n\times n}$	Set of $n \times n$ entrywise nonnegative matrices.
$\mathcal{S}^n$	Set of $n \times n$ symmetric matrices.
$\mathcal{S}_+^n$	Set of $n \times n$ positive semidefinite matrices.
$\mathcal{DNN}^n$	Set of $n \times n$ doubly nonnegative matrices.
$\mathcal{CP}^n$	Set of $n \times n$ completely positive matrices.
$\mathcal{COP}^n$	Set of $n \times n$ copositive matrices.
$\mathrm{cone}(\mathcal{C})$	For a set $\mathcal{C}$, define $\mathrm{cone}(\mathcal{C}) = \{\alpha x : \alpha \in \mathbb{R}_+, x \in \mathcal{C}\}$.
$\mathsf{cl}(\mathcal{C})$	Closure of the set $\mathcal{C}$.
$\mathrm{int}(\mathcal{C})$	Interior of the set $\mathcal{C}$.

Quantum information theory

$\|\phi\rangle$	Dirac notation for a vector $\phi \in \mathbb{C}^d$.
$\langle\phi\|$	Conjugate transpose of vector $\|\phi\rangle$.
$\|i\rangle$	Vector with ith entry equal to 1 and 0 elsewhere.
$\mathcal{H}$	Hilbert space
$\mathsf{Tr}_{\mathcal{H}}$	Partial trace over the Hilbert space $\mathcal{H}$.
$\mathcal{Q}$	Set of quantum correlations

Graphs

$\chi(G)$	Chromatic number.
$\chi_{\mathsf{f}}(G)$	Fractional chromatic number.
$\chi_q(G)$	Quantum chromatic number.
$\chi^\star(G)$	Entangled chromatic number.
$\alpha(G)$	Stability number.
$\alpha_q(G)$	Quantum stability number.
$\alpha^\star(G)$	Entangled stability number.
$\omega(G)$	Clique number.
$\Theta(G)$	Shannon capacity.
$\Theta^\star(G)$	Entangled shannon capacity.
$\vartheta(G)$	Lovász theta number.
$\xi(G)$	Orthogonal rank.
$\xi'(G)$	Minimum dimension of an orthogonal representation where the entries of the vectors have all absolute value one.
K_n	Complete graph on n vertices.
C_n	Cycle of length n.
$G + H$	Disjoint union of graphs G and H.
$G \square H$	Cartesian product of graphs G and H.
$G \boxtimes H$	Strong graph product of graphs G and H.
$G * H$	Disjunctive product of graphs G and H.

Samenvatting

Kwantummechanica is een natuurkundig model dat het gedrag van kleine deeltjes beschrijft. Een zeer eigenaardig kenmerk van deze theorie is dat het het bestaan van kwantumverstrengeling voorspelt, zoals voor het eerste ontdekt door Einstein, Podolsky en Rosen [EPR35] in 1935. Het idee is dat ruimtelijk gescheiden kwantumsystemen met elkaar verstrengeld kunnen zijn en dat lokale operaties op één systeem invloed hebben op de toestanden van de andere systemen. In het bijzonder zegt dit dat deeltjes gecorreleerd kunnen zijn op een niet-klassieke manier. In 1964 presenteerde Bell [Bel64] een experiment om te testen of dit niet-klassieke fenomeen in de natuur voorkomt en in de laatste dertig jaar hebben we steeds overtuigendere implementaties gezien van dit experiment [ADR82, HBD+15] die de niet-klassieke aard aantonen van de wereld waar wij in leven.

In deze scriptie bestuderen we de gevolgen van kwantumverstrengeling in *nonlokale spelen* en communicatieproblemen in *zero-error informatietheorie*

Een nonlokaal spel is een spel met twee samenwerkende spelers, die niet met elkaar mogen communiceren, maar die wel contact hebben met een scheidsrechter. Ze ontvangen ieder een vraag van de scheidsrechter waarop ze moeten reageren met een antwoord. Aan de hand van een bij de spelers bekend predicaat, dat afhangt van de twee vragen en de twee antwoorden, bepaalt de scheidsrechter of de spelers het spel hebben gewonnen of verloren. Het doel van de spelers is om hun winkans te maximaliseren door op de een of andere manier hun strategieën te coördineren. Klassiek is de optimale werkwijze om voor iedere vraag een vast antwoord te hebben. Echter, als de spelers toegang hebben tot een verstrengeld systeem, dan kunnen geavanceerdere strategieën worden gebruikt: elke speler antwoordt aan de hand van de uitkomst van een experiment uitgevoerd op een persoonlijk systeem. Zulke strategieën kunnen ervoor zorgen dat de spelers antwoorden geven die gecorreleerd zijn op niet-klassieke wijze.

Zero-error informatietheorie is een gebied in de wiskunde dat zich richt op

verscheidene communicatieproblemen waar geen fouten worden getolereerd. Bijvoorbeeld, in het zero-error kanaalcoderingsprobleem wil een verzender berichten sturen over een kanaal met ruis op een manier dat de ontvanger in staat is om het bericht perfect te reconstrueren. Voor dit probleem en andere problemen onderzoeken we of verstrengeling het mogelijk maakt om beter-dan-klassieke communicatieprotocollen te maken.

De verbindende schakel tussen de verscheidene problemen die we bestuderen in deze scriptie is hun combinatorische karakter. De meerderheid heeft inderdaad een graaftheoretische formulering, voornamelijk betreffende het chromatisch getal en het stabiliteitsgetal en enige kwantumgeneraliseringen daarvan.

Een van de voornaamste bijdragen in deze scriptie is een nieuwe benadering voor het bestuderen van deze *kwantum-graafparameters* door het gebruik van *conisch optimaliseren*. Dit moet worden gezien in analogie met het klassieke geval, waar het chromatisch getal en het stabiliteitsgetal kunnen worden geformuleerd als lineaire optimalisatieprogramma's over een geschikte convexe kegel, die de compleet positieve kegel wordt genoemd. We introduceren de completely positive semidefinite-kegel $\mathcal{CS}_+^n$, een nieuwe matrixkegel die bestaat uit alle $n \times n$ symmetrische matrices die een Gram-representatie hebben in positief semi-definiete matrices. Naast het bestuderen van enkele structurele eigenschappen en het leggen van verbanden met andere welbekende kegels, gebruiken we het om kwantumvarianten te formuleren van graafparameters als lineaire optimalisatieprogramma's over de completely positive semidefinite-kegel.

In het tweede deel van deze scriptie richten we ons op het probleem van het vinden van scheidingen tussen klassieke strategieën en kwantumstrategieën in enkele standaardproblemen uit de zero-error informatietheorie. We bestuderen het *kanaalcoderingsprobleem*, wat vraagt dat een verzender data betrouwbaar verstuurd naar een ontvanger in de aanwezigheid van ruis, en twee generaliseringen van dit probleem in het meerpartijenmodel. We beschouwen het *broncoderingsprobleem*, wat vraagt dat een verzender efficiënt data verstuurt waarover een ontvanger al wat informatie heeft, en het *bron-kanaalcoderingsprobleem* wat een combinatie is van het broncoderingsprobleem en het kanaalcoderingsprobleem: de verzender moet een kanaal met ruis gebruiken om de data naar de ontvanger te communiceren.

Daarnaast bestuderen we de complexiteit van twee communicatieproblemen: het *promise equality* probleem en het *lijstprobleem*. In het promise equality probleem moeten twee spelers (Alice en Bob) beslissen of hun invoeren gelijk zijn of niet. In het lijstprobleem krijgt Bob een lijst en krijgt Alice een element in Bobs lijst. Hun doel is dat Bob het element van Alice te weten komt. Voor beide problemen zijn we geïnteresseerd in het minimum aantal klassieke berichten, of kwantumberichten, dat moet worden uitgewisseld tussen Alice en Bob zo-

dat het probleem kan worden opgelost zonder fouten, in het bijzonder wanneer we onderscheid maken tussen de één-ronde communicatiecomplexiteit, waar de communicatie altijd loopt van Alice naar Bob, en de multi-ronde communicatiecomplexiteit, waar de spelers om beurten berichten sturen.

Abstract

Quantum mechanics is a physical model that describes the behavior of small particles. One very peculiar feature of this theory is that it predicts the existence of *quantum entanglement*, as it was first discovered by Einstein, Podolsky, and Rosen [EPR35] in 1935. The idea is that spatially separated quantum system can be entangled with each others and local operations on one system can have an influence on the states of the other systems. In particular, this says that particles can be correlated in a non-classical way. In 1964 Bell [Bel64] proposed an experiment to test whether this non-classical phenomenon occurs in nature and in the last three decades we have seen increasingly convincing implementations of such experiment [ADR82, HBD+15] showing the non-classical nature of the world we live in.

In this thesis we study the effects of quantum entanglement in *nonlocal games* and communication problems in *zero-error information theory*.

In a nonlocal game two cooperating players, who are not allowed to communicate with each others, interact with a referee. They each receive a question from the referee to which they have to reply with some answer. According to some known predicate, which depends on the two questions and on the two answers, the referee determines whether the players have won or lost the game. The players' goal is to maximize their chances of winning by somehow coordinating their strategy. Classically, the optimal course of action is to fix an answer to each question. However, if the players have access to a entangled physical system, more sophisticated strategies can be used: each player answers according to the outcome of an experiment performed on a private system. Such strategies can cause the players to produce answers that are correlated in a non-classical way.

Zero-error information theory is a mathematical field that studies various communication problems where no error is tolerated. For instance, in the zero-error channel coding problem, a sender wants to communicate messages over a noisy channel in a way that allows the receiver to perfectly reconstruct the

message. For this and other problems, we investigate whether entanglement allows for better-than-classical communication schemes.

The unifying link among the various problems that we study in this thesis is their combinatorial nature. Indeed, the majority of them will have a graph theoretical formulation, mainly concerning the chromatic and stability numbers and some quantum generalizations thereof.

One of the main contributions of this thesis is a novel approach to the study of these *quantum graph parameters* using the paradigm of *conic optimization*. This should be seen in analogy with the classical case, where the chromatic and stability numbers can be reformulated as linear optimization programs over an appropriate convex cone, called the completely positive cone. Here we introduce the completely positive semidefinite cone $\mathcal{CS}_+^n$, a new matrix cone consisting of all $n \times n$ symmetric matrices that admit a Gram representation by positive semidefinite matrices. Beside studying some of its structural properties and drawing connections with other well-known cones, we use it to formulate quantum variants of graph parameters as linear optimization programs over the completely positive semidefinite cone.

In the second part of the thesis, we focus on the problem of finding separations between classical and quantum strategies in some standard problems from zero-error information theory. We study the *channel coding problem*, which asks a sender to transmit data reliably to a receiver in the presence of noise, and two generalizations of this problem to the multiparty setting. We consider the *source coding problem*, where a sender has to efficiently communicate data about which a receiver has already some information, and the *source-channel coding problem* which is a combination of the source and the channel coding problem: the sender can only use a noisy channel to communicate the data to the receiver.

Moreover, we study two communication complexity problems: the *promise equality* and the *list* problems. In the promise equality problem, two parties (Alice and Bob) must decide whether their inputs are equal or not. In the list problem, Bob gets a list and Alice gets an element from Bob's list. Their goal is for Bob to learn Alice's element. In both of these problems we are interested in the minimum number of classical, or quantum, messages that have to be exchanged between Alice and Bob to be able to solve the problem without error, especially when making the distinction between one-round communication complexity, where the communication flows from Alice to Bob, and multi-round communication complexity, where the parties take turns in the transmission of the messages.

Publications

The content of this thesis is based on the following publications. The authors are listed in the alphabetical order and co-authorship is equally shared.

- M. Laurent and T. Piovesan. Conic approach to quantum graph parameters using linear optimization over the completely positive semidefinite cone. *SIAM Journal on Optimization*, 25(4):2461–2493, 2015.

- S. Burgdorf, M. Laurent, and T. Piovesan. On the closure of the completely positive semidefinite cone and linear approximations to quantum coloring. In *Proceedings of the 10th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2015)*, pages 127–146, 2015.

- J. Briët, H. Buhrman, M. Laurent, T. Piovesan, and G. Scarpa. Entanglement-assisted zero-error source-channel coding. *IEEE Transactions on Information Theory*, 61(2):1124–1138, 2015.

- T. Piovesan, G. Scarpa, and C. Schaffner. Multiparty zero-error classical channel coding with entanglement. *IEEE Transactions on Information Theory*, 61(2):1113–1123, 2015.

- J. Briët, H. Buhrman, D. Leung, T. Piovesan, and F. Speelman. Round elimination in exact quantum communication complexity. In *Proceedings of the 10th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2015)*, pages 206–225, 2015.

The author contributed to one additional paper during her time as PhD candidate. The authors are listed in descending order of contribution.

- E. Quaeghebeur, C. Wesseling, E. Bauxis-Aussalet, T. Piovesan, and T. Sterkenburg. Eliciting sets of acceptable gambles – The CWI World Cup Competition. In *Proceedings of the 9th International Symposium on Imprecise Probabilities: Theory and Applications (ISIPTA 2015)*, 2015.

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

ILLC DS-2014-01: **Machiel Keestra**
Sculpting the Space of Actions. Explaining Human Action by Integrating Intentions and Mechanisms

ILLC DS-2014-02: **Thomas Icard**
The Algorithmic Mind: A Study of Inference in Action

ILLC DS-2014-03: **Harald A. Bastiaanse**
Very, Many, Small, Penguins

ILLC DS-2014-04: **Ben Rodenhäuser**
A Matter of Trust: Dynamic Attitudes in Epistemic Logic

ILLC DS-2015-01: **María Inés Crespo**
Affecting Meaning. Subjectivity and evaluativity in gradable adjectives.

ILLC DS-2015-02: **Mathias Winther Madsen**
The Kid, the Clerk, and the Gambler - Critical Studies in Statistics and Cognitive Science

ILLC DS-2015-03: **Shengyang Zhong**
Orthogonality and Quantum Geometry: Towards a Relational Reconstruction of Quantum Theory

ILLC DS-2015-04: **Sumit Sourabh**
Correspondence and Canonicity in Non-Classical Logic

ILLC DS-2015-05: **Facundo Carreiro**
Fragments of Fixpoint Logics: Automata and Expressiveness

ILLC DS-2016-01: **Ivano A. Ciardelli**
Questions in Logic

ILLC DS-2016-02: **Zoé Christoff**
Dynamic Logics of Networks: Information Flow and the Spread of Opinion

ILLC DS-2016-03: **Fleur Leonie Bouwer**
What do we need to hear a beat? The influence of attention, musical abilities, and accents on the perception of metrical rhythm

ILLC DS-2016-04: **Johannes Marti**
Interpreting Linguistic Behavior with Possible World Models

ILLC DS-2016-05: **Phong Lê**
Learning Vector Representations for Sentences - The Recursive Deep Learning Approach

ILLC DS-2016-06: **Gideon Maillette de Buy Wenniger**
Aligning the Foundations of Hierarchical Statistical Machine Translation

ILLC DS-2016-07: **Andreas van Cranenburgh**
Rich Statistical Parsing and Literary Language

ILLC DS-2016-08: **Florian Speelman**
Position-based Quantum Cryptography and Catalytic Computation

ILLC DS-2016-09: **Teresa Piovesan**
Quantum entanglement: insights via graph parameters and conic optimization