Quantum entanglement in non-local games, graph parameters and zero-error information theory
Scarpa, G.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 4

Zero-error information theory

This chapter is based on the paper “Zero-error source-channel coding with entanglement”, by J. Briët, H. Buhrman, M. Laurent, T. Piovesan and the author. The paper was presented at the Eurocomb conference in September 2013. An extended abstract was published in the conference proceedings.

4.1 Introduction

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

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

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

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

4.2 Preliminaries

4.2.1 Classical source-channel coding

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

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

---

1From now on we will assume that all binary channels are noiseless.
4.2. Preliminaries

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 $x = (x_1, \ldots, x_m)$, Alice applies encoding function $C : X^m \to \{0, 1\}^n$ and sends $C(x)$ through the binary channel by using it $m$ times in a row. Bob, who received an input $u = (u_1, \ldots, u_m) \in U^m$, then applies a decoding function $D : U^m \times \{0, 1\}^n \to X^m$ to the pair $(u, C(x))$ to get a string in $X^m$. The scheme works if Bob always gets the string $x$. The cost rate of the scheme $(C, D)$ is then $\frac{n}{m}$, which counts the average number of channel uses per source-input symbol.

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

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

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

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

---

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

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

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

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

(4.2)

is the maximum communication rate of a zero-error coding scheme. Similarly to the Witsenhausen rate, we can replace the above limit with the supremum:

$$c(H) = \sup_n \log \alpha(H^{\otimes n})/n.$$

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

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

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

(4.3)

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

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

is equal to

$^{3}$ An intuition for this is that in the source-channel problem, parties need to map confusable pairs of source inputs to non-confusable pairs of channel inputs, and that is what an homomorphism from $G$ to the complement of $H$ does.
is sub-additive and apply Fekete’s lemma, which shows that
\[ \eta(G, H) = \lim_{m \to \infty} \frac{\eta_m(G, H)}{m} \]
is also equal to the infimum \( \inf_m \frac{\eta_m(G, H)}{m} \).

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

**Lemma 4.2.1.** Let \( G \) and \( H \) be graphs such that both \( G \) and \( \overline{H} \) have at least one
edge. Then,
\[ R(G) = \eta(G, \overline{K}_2) \text{ and } 1/c(H) = \eta(K_2, H). \]

**Proof.** For the proof of the identity \( R(G) = \eta(G, \overline{K}_2) \) we use the following fact:
for a graph \( G' \) and \( t \in \mathcal{N} \), there exists a homomorphism from \( G' \) to \( K_t \) if and
only if \( \chi(G') \leq t \), which implies
\[ \log \chi(G') \leq \min\{n : G' \to K_{2^n}\} < \log \chi(G') + 1. \]
Combining these inequalities applied to \( G' = G^{2m} \) with the identity \( \overline{K}_2^{2m} = K_{2^n} \),
we obtain
\[ \eta(G, \overline{K}_2) = \lim_{m \to \infty} \frac{1}{m} \min\{n : G^{2m} \to K_{2^n}\} = \lim_{m \to \infty} \frac{1}{m} \log \chi(G^{2m}) = R(G). \]
The proof of the identity $1/c(H) = \eta(K_2, H)$ uses the fact that, for a graph $H'$ and $t \in \mathcal{N}$, there exists a homomorphism from $K_t$ to $H'$ if and only if $\alpha(H') \geq t$. Since $K_2^{2m} = K_2m$, we get

\[
\eta_m(K_2, H) = \min \{ n : K_2^{2m} = K_2m \rightarrow H^{2n} \} = \min \{ n : \alpha(H^{2n}) \geq 2m \} = \min \{ n : \log \alpha(H^{2n}) \geq m \}.
\]

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

\[
\log \alpha(H^{2(n(m)-1)}) < m \leq \log \alpha(H^{2n(m)})
\]

and thus

\[
\frac{n(m)}{\log \alpha(H^{2(n(m)-1)})} \leq \frac{n(m)}{m} \leq \frac{n(m)}{\log \alpha(H^{2(n(m)-1)})}.
\]

As $c(H) = \sup_n \log \alpha(H^{2n})/n$, using the left most inequality in (4.4) we deduce

\[
\frac{1}{c(H)} \leq \frac{n(m)}{\log \alpha(H^{2(n(m)-1)})} \leq \frac{n(m)}{m}
\]

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

\[
\eta_m(K_2, H) \leq \frac{n(m)}{m} < \frac{n(m)}{\log \alpha(H^{2(n(m)-1)})} = \frac{n(m) - 1}{\log \alpha(H^{2(n(m)-1)})} \frac{n(m)}{n(m) - 1}.
\]

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

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

\[\square\]
4.2. Preliminaries

4.2.2 Quarter-orthogonality graphs

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

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

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

$$f : V(H_k) \rightarrow \mathbb{R}^{k+1}, \quad u \mapsto \phi(u)/\sqrt{k+1}, \quad (4.5)$$

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

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

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

“It seems probable that, whenever $m$ is divisible by 4, it is possible to construct an orthogonal matrix of order $m$ composed of $\pm 1$, but the general theorem has every appearance of difficulty.”
Indeed, the conjecture has remained unproved despite sustained efforts. However, many infinite families of Hadamard matrices are known. We will use a family constructed by Xia and Liu [XL91] (see for example [XL96, WX97, Che97, Xia98, XSX06] for closely related constructions).

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

We also use the following result regarding the graph \( H_k \).

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

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

### 4.2.3 Quantum teleportation

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

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

1. **(QT1)** a state \( \sigma \) of the pair \((X, Y)\) (known as the **maximally entangled state**),
2. **(QT2)** a measurement \( M = \{M_i \in \mathbb{C}^{d \times d} \otimes \mathbb{C}^{d \times d} : i \in [d^2]\} \) (which is independent of \( \rho \)) and
3. **(QT3)** for every \( i \in [d^2] \), a unitary operation \( U_i \in \mathbb{C}^{d \times d} \)

with which Alice and Bob can transfer (“teleport”) the state \( \rho \) of Alice’s system \( A \) to Bob’s system \( Y \). To achieve this, the parties may follow the following protocol:

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

That is, at the end of the protocol

$$(QT4) \text{ Bob's system } \mathcal{Y} \text{ is in state proportional to } U_i \text{Tr}_{A,\mathcal{X}} \left( (M_i \otimes I)(\rho \otimes \sigma) \right) U_i^\dagger,$$

which is guaranteed to be equal to $\rho$.

### 4.3 Source-channel coding with entanglement

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

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

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

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

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

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

Recall that if the two parties share no entanglement, then a zero-error $(m, n)$-coding scheme is equivalent to a homomorphism from $V(G^{\otimes m})$ to $V(H^{\otimes n})$, i.e., a map that sends edges of $G^{\otimes m}$ to non-edges of $H^{\otimes n}$, where $G$ is the characteristic graph of $\mathcal{M}$ and $H$ is the confusability graph of $\mathcal{N}$. Analogously, the entanglement-assisted protocol is successful if and only if, for every edge $\{x, y\}$ in $G^{\otimes m}$ and every non-edge $\{s, t\}$ in $H^{\otimes n}$, we have that $\text{Tr}_A((A^x_s \otimes I)\sigma)$ is orthogonal to $\text{Tr}_A((A^y_t \otimes I)\sigma)$. The intuition is that indistinguishable pairs of Alice’s inputs must be related to channel inputs that will not create confusion in Bob’s measurement, thus allowing him to output correctly. We will see in the next section how this requirement gives rise to the algebraic definition of the entangled source-channel rate.
Figure 4.2: The figure illustrates the entanglement-assisted source-channel coding protocol. After receiving a source-input $x \in X^m$, Alice performs a measurement $\{A_s^x : s \in S^m\}$ on her part of an entangled state $\sigma$ which she shares with Bob. She sends her measurement outcome $s$ through the channel, upon which Bob—who previously already received a source-input $u$—receives a channel-output $v \in V^n$. Bob performs a measurement $\{B_{u,v}^y : y \in X^m\}$ on his part of $\sigma$ and obtains outcome $y \in X^m$.

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

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

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

$$
\rho_s^x \rho_t^y = 0 \ \forall x,y,s,t : \{x,y\} \in E(G^\otimes m), s = t \text{ or } \{s,t\} \in E(H^\otimes n)
$$

$$
\sum_{s \in V(H^\otimes n)} \rho_s^x = \rho \ \forall x \in V(G^\otimes m).
$$

The entangled cost rate is defined by

$$
\eta^*(G,H) = \lim_{m \to \infty} \frac{1}{m} \eta^*_m(G,H).
$$
4.3. Source-channel coding with entanglement

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

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

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

$$\rho_i^u \rho_i^v = 0 \quad \forall i, u, v : i \in [t], \{u, v\} \in E(G)$$

$$\sum_{i \in [t]} \rho_i^u = \rho \quad \forall u \in V(G).$$

The entangled Witsenhausen rate is defined by

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

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

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

$$\rho_i^u \rho_v^j = 0 \quad \forall i, j, u, v : i \neq j, u = v \text{ or } \{u, v\} \in E(H)$$

$$\sum_{u \in V(H)} \rho_i^u = \rho \quad \forall i \in [M].$$

The entangled Shannon capacity is defined by

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

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

Analogous to the classical setting, we can reformulate the entangled variants of the Witsenhausen rate and Shannon capacity as follows.
Lemma 4.3.4. Let $G$ and $H$ be graphs such that both $G$ and $\overline{H}$ have at least one edge. Then,

$$R^*(G) = \eta^*(G, K_2) \quad \text{and} \quad 1/c^*(H) = \eta^*(K_2, H).$$

Proof. Since the graph $K_2^{\otimes n}$ has $2^n$ vertices and no edges, it follows from the definitions that $\eta^*_m(G, K_2) = \lceil \log^* (G^{\otimes m}) \rceil$. The identity $R^*(G) = \eta^*(G, K_2)$ follows by dividing by $m$ and letting $m$ go to infinity.

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

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

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

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

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

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

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

Sub-additivity of $\eta^*_m$ and sub-multiplicativity of $\chi^*$

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

$$
\eta^*_{m+m'}(G, H) \leq \eta^*_m(G, H) + \eta^*_{m'}(G, H).
$$

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

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

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

$$
\sum_{(s,t) \in V(H^{\otimes (n+n')})} \rho^{(s,t)}_{(x,y)} = \sum_{s \in V(H^{\otimes n})} \sum_{t \in V(H^{\otimes n'})} \varphi^s_x \otimes \psi^t_y
$$

$$
= \left( \sum_{s \in V(H^{\otimes n})} \varphi^s_x \right) \otimes \left( \sum_{t \in V(H^{\otimes n'})} \psi^t_y \right) = \varphi \otimes \psi = \rho.
$$

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

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

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

Lemma 4.3.6. For two graphs $G$ and $H$, $\chi^*(G \boxtimes H) \leq \chi^*(G)\chi^*(H)$. 
Lemma 4.3.7. Given graphs $\eta$ follows that uses of the channel and shared entanglement. If this condition holds, then it $x \in \chi G \text{tic graph instances of the source problem. In other words, for a source with characteris-}$

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

$$\rho_k^{(u,v)} = \varphi_i^u \otimes \psi_j^v.$$ 

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

Separate coding schemes

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

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

$$\chi(G^{\otimes m}) \leq \alpha(H^{\boxtimes n}) \implies \eta_m(G, H) \leq n,$$  \hspace{1cm} (4.6)

$$\chi^*(G^{\otimes m}) \leq \alpha^*(H^{\boxtimes n}) \implies \eta_m^*(G, H) \leq n.$$ \hspace{1cm} (4.7)

Proof. If $\chi(G^{\otimes m}) \leq \alpha(H^{\boxtimes n})$, then there is a homomorphism from $G^{\otimes m}$ to $H^{\boxtimes n}$ and thus $\eta_m(G, H) \leq n$, which shows (4.6). We now show (4.7). For this set $t = \chi^*(G^{\otimes m})$ and $M = \alpha^*(H^{\boxtimes n})$, with $t \leq M$ by assumption. Let $\varphi, \{\varphi_i^x : x \in V(G^{\otimes m}), i \in [t]\}$ be a collection of positive semidefinite matrices forming a solution for $\chi^*(G^{\otimes m})$ and let the set of positive semidefinite matrices $\psi, \{\psi_j^s : s \in V(H^{\boxtimes n}), i \in [M]\}$ be feasible for $\alpha^*(H^{\boxtimes n})$. We construct a solution for $\eta_m^*(G, H)$ as follows. For $x \in V(G^{\otimes m})$ and $s \in V(H^{\boxtimes n})$ set

$$\rho_s^x = \sum_{i \in [t]} \varphi_i^x \otimes \psi_s^i \quad \text{and} \quad \rho = \varphi \otimes \psi.$$ 

Then, we have that $\text{Tr}(\rho) = \text{Tr}(\varphi \otimes \psi) = 1$ and, for every $x \in V(G^{\otimes m})$, we get that

$$\sum_{s \in V(H^{\boxtimes n})} \rho_s^x = \sum_{s \in V(H^{\boxtimes n})} \sum_{i \in [t]} \varphi_i^x \otimes \psi_s^i \quad \text{is equal to}$$
ϕ ⊗ ψ = ρ. Moreover, for every \( \{x, y\} \in E(G^{2m}) \) and \( \{s, t\} \in V(H^{2n}) \cup E(H^{2n}) \),
\[
\rho_s^x \rho_t^y = \left( \sum_{i \in [t]} \varphi_i^x \otimes \psi_i^y \right) \left( \sum_{j \in [t]} \varphi_j^y \otimes \psi_j^t \right) = \sum_{i \in [t]} \sum_{j \in [t]} \varphi_i^x \varphi_j^y \otimes \psi_i^y \psi_j^t = 0,
\]
where the last identity uses the orthogonality conditions of the matrices \( \varphi_i^x \) and \( \psi_i^y \). Hence \( \rho, \{\rho_s^x : x \in V(G^{2m}), s \in V(H^{2n})\} \) is a feasible solution for \( \eta_m^*(G, H) \leq n \).

We now relate the minimum cost rate to the ratio of the Witsenhausen rate and the Shannon capacity in both classical and entangled assisted cases.

**Proposition 4.3.8.** Let \( G \) and \( H \) be graphs and assume that both \( G \) and \( \overline{H} \) have at least one edge. Then,
\[
\eta(G, H) \leq \frac{R(G)}{c(H)} = \lim_{m \to \infty} \frac{1}{m} \min \{n : \chi(G^{2m}) \leq \alpha(H^{2n})\}, \tag{4.8}
\]
\[
\eta^*(G, H) \leq \frac{R^*(G)}{c^*(H)} = \lim_{m \to \infty} \frac{1}{m} \min \{n : \chi^*(G^{2m}) \leq \alpha^*(H^{2n})\}. \tag{4.9}
\]

**Proof.** We show (4.8); we omit the proof of (4.9) which is analogous (and uses (4.7)). From (4.6) we have the inequality:
\[
\eta_m(G, H) \leq \epsilon_m(G, H) := \min \{n : \chi(G^{2m}) \leq \alpha(H^{2n})\},
\]
which implies \( \eta(G, H) \leq \lim_{m \to \infty} \epsilon_m(G, H)/m \). Next we show that this limit is equal to \( R(G)/c(H) \), which concludes the proof of (4.8). Setting \( n = \epsilon_m(G, H) \), we have that \( \alpha(H^{2n-1}) < \chi(G^{2m}) \leq \alpha(H^{2n}) \), implying
\[
\frac{R(G)}{c(H)} \leq \frac{\log \chi(G^{2m})}{m} \leq \frac{n}{m} \leq \frac{n}{n-1} \frac{\log \chi(G^{2m})}{c(H)} \leq \frac{n-1}{m} \frac{\log \alpha(H^{2n})}{\log \alpha(H^{2n-1})}.
\]
Taking limits as \( m \to \infty \) in the right-most terms we obtain that
\[
\frac{R(G)}{c(H)} = \lim_{m \to \infty} \frac{\epsilon_m(G, H)}{m}.
\]

We also record the following bound, which we use later.

**Proposition 4.3.9.** For graphs \( G \) and \( H \) and positive integer \( m \), we have
\[
\eta_m^*(G, H) \leq \left\lfloor \frac{\log \alpha^*(G^{2m})}{\log \alpha^*(H)} \right\rfloor.
\]
Proof. Set \( n = \lceil \log \chi^*(G^{\boxtimes m}) / \log \alpha^*(H) \rceil \). Using the super-multiplicativity of \( \alpha^*(H) \) we get

\[
\log \alpha^*(H^{\boxtimes n}) \geq n \log \alpha^*(H) = \left\lceil \frac{\log \chi^*(G^{\boxtimes m})}{\log \alpha^*(H)} \right\rceil \log \alpha^*(H) \geq \log \chi^*(G^{\boxtimes m}).
\]

From Lemma 4.3.7 it then follows that \( \eta_m^*(G, H) \leq n \).

4.4 Szegedy’s number lower bound on the entangled chromatic number

Here we explain our lower bound on the entangled chromatic number. We show that \( \chi^*(G) \) is lower bounded by an efficiently computable graph parameter, namely a variant of the famous theta number introduced by Szegedy [Sze94].

Szegedy [Sze94] introduced the following strengthening of the theta number, which includes an extra linear constraint to the formulation (3.2).

\[
\vartheta^+(G) = \min \left\{ \lambda : \exists Z \in \mathbb{R}^{V(G) \times V(G)}, Z \succeq 0, \begin{align*}
Z(u, u) &= \lambda - 1 \text{ for } u \in V(G), \\
Z(u, v) &= -1 \text{ for } \{u, v\} \notin E(G), \\
Z(u, v) &\geq -1 \text{ for } \{u, v\} \in E(G) \right\}.
\]

(4.10)

Szegedy’s number satisfies \( \vartheta^*(G) \leq \vartheta^+(G) \) and \( \alpha(G) \leq \vartheta^+(G) \leq \chi(G) \). Recall that Lovász proved that \( \vartheta \) is multiplicative under the strong graph product, that is, \( \vartheta(G \boxtimes H) = \vartheta(G)\vartheta(H) \). Moreover Knuth [KD93] showed that \( \vartheta(G^{\boxtimes m}) = \vartheta(G)^m \). It is unknown if the latter identity holds for \( \vartheta^+ \) [Meu05].

The identities of Lovász and Knuth give for any graph \( G \) and \( m \in \mathbb{N} \):

\[
\vartheta(G^{\boxtimes m}) = \vartheta(G^{\boxtimes m}) = \vartheta(G)^m.
\]

(4.11)

Combining these properties of \( \vartheta \) with the Sandwich Theorem shows that

\[
c(G) \leq \log \vartheta(G) \leq R(G).
\]

These inequalities capture the best known efficiently computable bounds for the Shannon capacity and the Witsenhausen rate.

Now we prove that the parameter \( \vartheta^+ \) (and thus \( \vartheta \) as well) lower bounds the entangled chromatic number and hence \( \log \vartheta \) lower bounds the entangled Witsenhausen rate.

Theorem 4.4.1. For any graph \( G \), we have

\[
\vartheta^+(G) \leq \chi^*(\overline{G}),
\]

(4.12)

\[
\log \vartheta(G) \leq R^*(\overline{G}).
\]

(4.13)
In [RM12] it is observed that $\vartheta(G) \leq \chi_q(G)$ holds. Theorem 4.4.1 thus strengthens this bound as it gives $\vartheta(G) \leq \vartheta^+(G) \leq \chi^*(G) \leq \chi_q(G)$.

Beigi [Bei10] and Duan, Severini and Winter [DSW13] proved that $\vartheta(G)$ upper bounds $\alpha^*(G)$. The above-mentioned relations therefore imply the following sequence of inequalities.

$$c(G) \leq c^*(G) \leq \log \vartheta(G) \leq R^*(G) \leq R(G).$$

We will use the following result about positive semidefinite matrices with a special block form (which can be found, e.g., in [GL08]).

**Lemma 4.4.2.** Let $X$ be a $t \times t$ block matrix, with a matrix $A$ as diagonal blocks and a matrix $B$ as non-diagonal blocks, of the form

$$X = \begin{pmatrix} A & B & \ldots & B \\ B & A & \ldots & B \\ \vdots & \vdots & \ddots & \vdots \\ B & B & \ldots & A \end{pmatrix}_{t \text{ blocks}}.$$

Then, $X \succeq 0$ if and only if $A - B \succeq 0$ and $A + (t - 1)B \succeq 0$.

**Proof of Theorem 4.4.1.** We show that relations (4.12) and (4.13) hold for the graph $\overline{G}$. First we observe that (4.13) follows from (4.12). Indeed, relation (4.12) combined with the identity (4.11) implies $\vartheta(G)^m = \vartheta(G^\overline{G}) \leq \chi^*(G^\overline{G})$ and thus $\log \vartheta(G) \leq R^*(G)$ follows after taking limits.

We now prove (4.12) for the graph $G$, i.e., we show the inequality $\vartheta^+(G) \leq \chi^*(G)$. For this let $\rho, \{\rho_u : u \in V(G), i \in [t]\}$ be a set of positive semidefinite matrices which form a solution for $\chi^*(G) = t$. We may assume that $\langle \rho, \rho \rangle = 1$. Here, $\langle \cdot, \cdot \rangle$ is the trace inner product, defined by $\langle A, B \rangle = \text{Tr}(A^*B)$ for matrices $A, B$ of the same size. Define the matrix $X$, indexed by all pairs $\{u, i\} \in V(G) \times [t]$, with entries $X_{ui,vj} := \langle \rho_u^i, \rho_v^j \rangle$. By construction, $X$ is a non-negative positive semidefinite matrix which satisfies $X_{ui,vj} = 0$ for every $\{u, v\} \in E(G)$ and $i \in [t]$.

For any element $\sigma$ of $\text{Sym}(t)$, the group of permutations of $[t]$, we define the new (permuted) matrix $\sigma(X) = (X_{u_\sigma(i),v_\sigma(j)})$. Then we average the matrix $X$ over the group $\text{Sym}(t)$, obtaining the new matrix

$$Y = \frac{1}{|\text{Sym}(t)|} \sum_{\sigma \in \text{Sym}(t)} \sigma(X).$$

By construction, the matrix $Y$ is invariant under any permutation of $[t]$, i.e., $\sigma(Y) = Y$ for any $\sigma \in \text{Sym}(t)$. Therefore, $Y$ has the block form of Lemma 4.4.2 with, moreover,

$$A_{uv} = 0 \text{ for all } \{u, v\} \in E(G).$$  \hspace{1cm} (4.14)
As each matrix $\sigma(X)$ is positive semidefinite, the matrix $Y$ is positive semidefinite as well. From Lemma 4.4.2, this implies that $A - B$ and $A + (t - 1)B$ are positive semidefinite matrices. Using the definition of the matrix $X$ combined with the properties of the matrices $\rho_i^u$ and the invariance of $Y$, we obtain the following relation for any $u, v \in V(G)$:

$$1 = \langle \rho, \rho \rangle = \sum_{i \in \llbracket t \rrbracket} \sum_{j \in \llbracket t \rrbracket} \langle \rho_i^u, \rho_j^v \rangle = \sum_{i \in \llbracket t \rrbracket} \sum_{j \in \llbracket t \rrbracket} X_{ui,vj} = \sum_{i \in \llbracket t \rrbracket} \sum_{j \in \llbracket t \rrbracket} Y_{ui,vj},$$

implying

$$1 = \sum_{i \in \llbracket t \rrbracket} \sum_{j \in \llbracket t \rrbracket} Y_{ui,vj} = t \sum_{j \in \llbracket t \rrbracket} Y_{ui,vj} = t (A_{uv} + (t - 1)B_{uv}).$$

We are now ready to define a matrix $Z$ which is a feasible solution for the program (4.10) defining $\vartheta^+(G)$. Namely, set $Z = t(t - 1)(A - B)$. Then, $Z$ is a positive semidefinite matrix. For any edge $\{u, v\} \in E(G)$, the relations (4.14) and (4.15) give $A_{uv} = 0$ and $t(t - 1)B_{uv} = 1$ and thus $Z_{uv} = -1$. For a non-edge $\{u, v\}$, relation (4.15) combined with the fact that $A_{uv} \geq 0$ implies that $Z_{uv} \geq -1$. Finally, for any $u \in V(G)$, relation (4.15) combined with the fact that $B_{uu} \geq 0$ implies that $Z_{uu} \leq t - 1$. Define the vector $c$ with entries $c_u = t - 1 - Z_{uu} \geq 0$ for $u \in V(G)$, the diagonal matrix $D(c)$ with $c$ as diagonal, and the matrix $Z' = Z + D(c)$. Then, $Z'$ is positive semidefinite and satisfies all the conditions of the program (4.10) defining $\vartheta^+(G)$. This shows that $\vartheta^+(G) \leq \chi^*(G)$, which concludes the proof.

### 4.5 Separation between classical and entangled Witsenhausen rate

Our first separation result shows an exponential gap between the entangled and classical Witsenhausen rates of quarter-orthogonality graphs (Definition 4.2.2).

**Theorem 4.5.1.** For every odd $k$ integer, we have

$$R^*(H_k) \leq \log(k + 1).$$

Moreover, if $k = 4p^\ell - 1$ where $p$ is an odd prime and $\ell \in \mathbb{N}$, then

$$R(H_k) \geq 0.154k - 1.$$
4.5.1 Upper bound on the entangled Witsenhausen rate

Here we prove the upper bound (4.16) stated in Theorem 4.5.1 on $R^* (H_k)$. Recall that a $d$-dimensional orthonormal representation of a graph $G$ is a map $f$ from $V(G)$ to the unit sphere in $\mathbb{C}^d$, having the property that adjacent vertices are mapped to orthogonal vectors.\footnote{We stress that in our definition orthogonality corresponds to adjacency. Some authors prefer to demand orthogonality for non-adjacent vertices instead.} Also, recall that the orthogonal rank $\xi(G)$ of $G$ is the minimum $d$ such that there exists a $d$-dimensional orthonormal representation of $G$. Following [CMN+07] we define $\xi'(G)$ to be the minimum dimension $d$ such that there exists a $d$-dimensional orthonormal representation $f$ of $G$ such that, for every vertex $u \in V(G)$, the $d$ entries of the vector $f(u)$ all have absolute value $1/\sqrt{d}$.

The following bound on $\chi^*(G)$ follows from the fact that $\chi^*(G) \leq \chi_q(G)$ and a result proved in [CMN+07] stating that $\chi_q(G) \leq \xi'(G)$. We give a self-contained proof of the implied bound on $\chi^*(G)$ for completeness.

**Lemma 4.5.2.** For every graph $G$, we have $\chi^*(G) \leq \xi'(G)$.

**Proof.** Set $d = \xi'(G)$, $\omega_d = e^{2i \pi / d}$ and let $h_j = [\omega_d^j, \omega_d^{j+1}, \ldots, \omega_d^{j+d-1}]^T \in \mathbb{C}^d$ for every $j \in [d]$. One can see that $\{h_1, h_2, \ldots, h_d\}$ is a complete orthogonal basis for $\mathbb{C}^d$. Set $\rho = I/d$. Then $\text{Tr}(\rho) = 1$.

Let $f : V(G) \to \mathbb{C}^d$ be an orthonormal representation of $G$ where each vector $f(u)$ is such that $(f(u)^*) f(u)_i = 1/d$ for every $i \in [d]$, as guaranteed to exist by the fact that $\xi'(G) = d$. For every $u \in V(G)$ and $i \in [d]$ define $\rho_i^u = |f(u) \circ h_i \rangle \langle f(u) \circ h_i|$, where $\circ$ denotes the entrywise product. We have

$$
\langle f(u) \circ h_i, f(v) \circ h_j \rangle = \begin{cases} 
\langle h_i, h_j \rangle/d & \text{if } u = v \\
\langle f(u), f(v) \rangle & \text{if } i = j.
\end{cases}
$$

It follows that for every $u \in V(G)$ we have $\rho_i^u + \rho_i^u + \cdots + \rho_i^u = I/d = \rho$. Moreover, for each $\{u, v\} \in E(G)$ and $i \in [d]$, we have $\rho_i^u \rho_i^v = 0$. As the matrices $\rho, \rho_i^u$ are also positive semidefinite, they satisfy all the requirements of Definition 4.3.2 and so $\chi^*(G) \leq d$.

The above lemma gives a bound on the entangled chromatic number of powers of $H_k$ from which it will be easy to get the upper bound on $R(H_k)$ given in (4.16).

**Lemma 4.5.3.** Let $k$ be an odd positive integer and $m \in \mathbb{N}$. Then,

$$
\chi^* (H_k^{\otimes m}) \leq (k+1)^m.
$$

Moreover, if there exists a Hadamard matrix of size $k+1$, then equality holds.
Proof. We first prove that $\chi^*(H_k) \leq k + 1$ by using Lemma 4.5.2. To this end we use the map $f$ defined in (4.5), which is an orthonormal representation from $V(H_k)$ to $\mathbb{R}^{k+1}$ where the representing vectors have entries with equal moduli. We conclude that $\xi'(H_k) \leq k+1$ and so by Lemma 4.5.2 we get $\chi^*(H_k) \leq k+1$. Using the sub-multiplicativity of $\chi^*$ (Lemma 4.3.6) we get $\chi^*(H_k^{\otimes m}) \leq (k+1)^m$.

We now prove that if there exists a Hadamard matrix of size $k + 1$ then also the other direction of the inequality holds. Recall from Proposition 4.2.5 the existence of a Hadamard matrix of size $k + 1$ implies $\omega(H_k) \geq k + 1$. Combining this with Theorem 4.4.1 and the Sandwich Theorem [KD93] gives that for every positive integer $m$, we have

$$\chi^*(H_k^{\otimes m}) \geq \vartheta(H_k^{\otimes m}) \geq \omega(H_k^{\otimes m}) \geq \omega(H_k)^m \geq (k + 1)^m,$$

(4.18)

where the second-last inequality uses the fact that if a subset $W \subseteq V(G)$ forms a clique in a graph $G$, then the set $W^m$ of $m$-tuples forms a clique in $G^{\otimes m}$. \qed

The desired result now follows as a corollary.

Proof of (4.16). Combining Lemmas 4.3.6 and 4.5.3 gives

$$R^*(H_k) \leq \log \chi^*(H_k) \leq \log(k + 1).$$

\[\]

We also record the following additional corollary, which we use in Section 4.7.

Corollary 4.5.4. For every odd integer $k$ such that there is a Hadamard matrix of size $k + 1$, we have $\omega(H_k^{\otimes m}) = (k + 1)^m$.

Proof. Combining Proposition 4.2.5, Lemma 4.5.3 and (4.18) gives the result. \qed

4.5.2 Lower bound on the classical Witsenhausen rate

To prove the lower bound (4.17) on $R(H_k)$ stated in Theorem 4.5.1 we use the following upper bound on the classical independence number of the graphs $H_k^{\otimes m}$ for certain values of $k$.

Lemma 4.5.5. Let $p$ be an odd prime number, $\ell \in \mathbb{N}$ and set $k = 4p^\ell - 1$. Then, for every $m \in \mathbb{N}$, we have

$$\alpha(H_k^{\otimes m}) \leq \left( \binom{k}{0} + \binom{k}{1} + \cdots + \binom{k}{p^\ell - 1} \right)^m \leq 2^{k m H(3/11)} < 2^{0.846 km},$$

(4.19)

where $H(t) = -t \log t - (1 - t) \log(1 - t)$ is the binary entropy function.
4.5. Separation between classical and entangled Witsenhausen rate

The proof of this lemma is an instance of the linear algebra method due to Alon [Alo98] (see also Gopalan [Gop06]), which we recall below for completeness. Let $G$ be a graph and $\mathbb{F}$ be a field. Let $\mathcal{F} \subseteq \mathbb{F}[x_1, \ldots, x_k]$ be a subspace of the space of $k$-variate polynomials over $\mathbb{F}$. A representation of $G$ over $\mathcal{F}$ is an assignment $((f_u, c_u))_{u \in V(G)} \subseteq \mathcal{F} \times \mathbb{F}^k$ of polynomial-point pairs to the vertices of $G$ such that

$$f_u(c_u) \neq 0 \quad \forall u \in V(G), \quad f_u(c_v) = 0 \quad \forall v \in V(G) \text{ with } \{u, v\} \notin E(G).$$

**Lemma 4.5.6** (Alon [Alo98]). *Let $G$ be a graph, $\mathbb{F}$ be a field, $k \in \mathbb{N}$ and $\mathcal{F}$ be a subspace of $\mathbb{F}[x_1, \ldots, x_k]$. If $((f_u, c_u))_{u \in V} \subseteq \mathcal{F} \times \mathbb{F}^k$ represents $G$, then $\alpha(G^{\otimes n}) \leq \dim(\mathcal{F})^n$ for all $n \in \mathbb{N}$.***

**Proof.** Let $I \subseteq V^n$ be an independent set in $G^{\otimes n}$. Each element $u = (u_1, \ldots, u_n)$ of $I$ is a $n$-tuple of vertices of $G$ and for every distinct pair $u, v \in I$ there is at least one index $i \in [n]$ such that $u_i$ and $v_i$ are neither equal nor adjacent in $G$.

For each $u \in I$, define the polynomial $f_u = \otimes_{i=1}^n f_{u_i} \in \mathcal{F}^{\otimes n}$, which takes as input $n$-tuples of vectors $c = (c_1, \ldots, c_n) \in (\mathbb{F}^k)^n$ and assumes the value $f_u(c) = f_{u_1}(c_1) \cdots f_{u_n}(c_n)$. Now define the $n$-tuple of vectors $c_u = (c_{u_1}, \ldots, c_{u_n})$. For all $u$ we have $f_{u_1}(c_{u_1}) \cdots f_{u_n}(c_{u_n}) \neq 0$ and for all distinct $u, v \in I$ we have $f_{u_i}(c_{v_i}) = 0$ for at least one $i \in [n]$. It follows that the pairs $((f_u, c_u))_{u \in V^n}$ represent $G^{\otimes n}$.

Now let $(a_u)_{u \in I} \in \mathbb{F}^I$ be a sequence of scalars and consider the polynomial

$$f = \sum_{u \in I} a_u f_u.$$

Then, by definition of a representation, for every $v \in I$ such that $a_v \neq 0$, we have

$$f(c_v) = \sum_{u \in I} a_u f_u(c_v) = a_v f_v(c_v) \neq 0.$$

It follows that $f$ can only be the zero polynomial if $a_u$ are zero for all $u$ and hence the polynomials $f_u$, for $u \in I$, are linearly independent. This implies that $\alpha(G^{\otimes n}) \leq \dim(\mathcal{F}^{\otimes n}) = \dim(\mathcal{F})^n$.  

We will get a representation for the graph $H_k$, for $k = 4p^\ell - 1$, from the following result of Barrington, Beigel and Rudich [BBRR94]. The proof we give here closely follows Yekhanin’s [Yek12, Lemma 5.6] but is slightly more explicit. Below, a multilinear polynomial is a polynomial in which the degree of each variable is at most 1.

**Lemma 4.5.7** (Barrington, Beigel and Rudich [BBRR94]). *Let $p$ be a prime number and let $k, \ell$ 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

\[
f(c) \equiv \begin{cases} 1 & \text{if } c_1 + c_2 + \cdots + c_k \equiv w \mod p^\ell \\ 0 & \text{otherwise.} \end{cases}
\]

The proof of this lemma relies on Lucas’ Theorem from number theory.

**Theorem 4.5.8** (Lucas’ Theorem). Let \( p \) be a prime and \( a, b \in \mathbb{N} \) with \( p \)-ary expansions \( a = \sum_i a_i p^i \) and \( b = \sum_i b_i p^i \), where \( 0 \leq a_i, b_i < p \). Then,

\[
\binom{a}{b} \equiv \prod_i \binom{a_i}{b_i} \mod p.
\]

**Proof of Lemma 4.5.7.** For \( c \in \{0, 1\}^k \), note that the value modulo \( p^\ell \) of the Hamming weight \( |c| \) depends only on the first \( \ell \) coefficients \( |c|_0, |c|_1, \ldots, |c|_{\ell-1} \) of the \( p \)-ary expansion of \( |c| \). The \( k \)-variate symmetric polynomial of degree \( d \) is defined by

\[
P_d(x_1, \ldots, x_k) = \sum_{S \in \binom{[k]}{d}} \prod_{i \in S} x_i.
\]

For every \( c \in \{0, 1\}^k \), we have

\[
P_{p^\ell}(c) = \binom{|c|}{p^\ell} \equiv \binom{|c|}{1} \mod p \equiv |c|, \mod p,
\]

where the second identity follows from Lucas’ theorem and the \( p \)-ary expansion of \( p^\ell \), in which the coefficient of value 1 multiplying \( p^i \) is the only nonzero coefficient.

Now, define the polynomial \( \hat{f} \in \mathbb{Z}_p[x_1, \ldots, x_k] \) by

\[
\hat{f}(x_1, \ldots, x_k) = \prod_{i=0}^{\ell-1} \left( 1 - \left( P_{p^\ell}(x) - w_i \right)^{p-1} \right),
\]

where \( w_i \) are the coefficients in \( p \)-ary expansion of \( w \). For \( c \in \{0, 1\}^k \), we have \( \hat{f}(c) \equiv 1 \mod p \) if \( |c|_i \equiv w_i \) for every \( i = 0, 1, \ldots, \ell - 1 \) (i.e., if \( |c| \equiv w \mod p^\ell \)) and \( \hat{f}(c) \equiv 0 \mod p \) otherwise. Here, we have used Fermat’s Little Theorem, which states that, for \( p \) prime and \( a \in \mathbb{N} \), \( a^{p-1} \equiv 1 \mod p \). Clearly the polynomial \( \hat{f} \) has only integer coefficients. Now let \( f \) be the multilinear polynomial obtained from \( \hat{f} \) by replacing each monomial \( x_1^{d_1} \cdots x_k^{d_k} \) by \( x_1^{i_1} \cdots x_k^{i_k} \) where \( i_h = \min\{d_h, 1\} \) for every \( h \in [k] \). Then, the degree of the polynomial \( f \) is bounded by \( \deg(f) \leq \deg(\hat{f}) \leq (p-1)(1 + p + p^2 + \cdots + p^{\ell-1}) = p^\ell - 1 \). Moreover, \( f \) agrees with \( \hat{f} \) on \( \{0, 1\}^k \) and satisfies the conditions of the lemma.

With this we can now prove Lemma 4.5.5.
Proof of Lemma 4.5.5. Let \( c \in \{0, 1\}^k \) be a string such that its Hamming weight \(|c|\) is even and satisfies \(|c| \equiv 0 \mod p^f\). Then, since \( p \) is odd and \( k < 4p^f \), we have \(|c| \in \{0, 2p^f\}\). Hence, if \(|c| \notin \{0, 2p^f\}\), then \(|c| \equiv 0 \mod p^f\).

Recall from Definition 4.2.2 that \( H_k \) can be defined as the graph whose vertices are the strings of \( \{0, 1\}^k \) with an even Hamming weight and where two distinct vertices \( u, v \) are adjacent if their Hamming distance \(|u \oplus v|\) is equal to \((k + 1)/2 = 2p^f\). Here \( u \oplus v \) is the sum modulo 2. For \( u, v \in V(H_k) \), their Hamming distance \(|u \oplus v|\) is an even number. Hence if \( u \neq v \) are not adjacent in \( H_k \), then \(|u \oplus v| \notin \{0, 2p^f\}\) and thus \(|u \oplus v| \equiv 0 \mod p^f\).

Let \( f \in \mathbb{Z}_p[x_1, \ldots, x_k] \) be a multilinear polynomial of degree at most \( p^f - 1 \) such that for every \( c \in \{0, 1\}^k \), we have

\[
f(c) = \begin{cases} 1 & \text{if } |c| \equiv 0 \mod p^f \\ 0 & \text{otherwise}, \end{cases}
\]

as is promised to exist by Lemma 4.5.7 (applied to \( w = 0 \)).

We use \( f \) to define a representation for \( H_k \). To this end, define for each \( u \in \{0, 1\}^k \) vertex in \( V(H_k) \) the polynomial \( f_u \in \mathbb{Z}_p[x_1, \ldots, x_k] \) 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 \( ((f_u, c_u))_{u \in V(H_k)} \) is a representation of \( H_k \). To see this, observe that \( f_u(c_u) = f(u \oplus v) \) 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 \( p^f - 1 \), they span a space of dimension at most \( \binom{k}{0} + \binom{k}{1} + \cdots + \binom{k}{p^f - 1} \), which is the number of multilinear monomials of degree at most \( p^f - 1 \). Applying Lemma 4.5.6 we obtain that

\[
\alpha(H_k^m) \leq \left( \binom{k}{0} + \binom{k}{1} + \cdots + \binom{k}{p^f - 1} \right)^m.
\]

We now use the well known fact that for \( q, k \in \mathbb{N} \) with \( 1 < q < k/2 \), \( \binom{k}{0} + \cdots + \binom{k}{q - 1} \leq 2^{kH(q/k)} \). From this, since \( p^f/(4p^f - 1) \leq 3/11 \), we deduce that the right hand side in (4.20) can be upper bounded by \( 2^{km H(3/11)} < 2^{0.846km} \).

The bound (4.17) stated in Theorem 4.5.1 is a corollary of Lemma 4.5.5.

Proof of (4.17). By Lemma 4.5.5, for every integer \( m \) we have

\[
\chi(H_k^m) \geq \frac{|V(H_k^m)|}{\alpha(H_k^m)} \geq \frac{2^{(k-1)m}}{2^{0.846km}} = 2^{[0.154k-1)m}.
\]

Taking the logarithm, dividing by \( m \) and taking the limit \( m \to \infty \) gives that for \( k = 4p^f - 1 \), we have

\[
R(H_k) \geq 0.154k - 1.
\]

\[\square\]
4.6 Separation between classical and entangled Shannon capacity

Our second separation result is a strengthening of the following result of [BBG12], which shows that for some values of $k$, the entangled Shannon capacity of $H_k$ can be strictly larger than its (classical) Shannon capacity.

**Theorem 4.6.1** (Briët, Buhrman and Gijswijt [BBG12]). Let $p$ be an odd prime such that there exists a Hadamard matrix of size $4p$. Set $k = 4p - 1$. Then,

$$c^*(H_k) \geq k - 1 - 2 \log(k + 1)$$

$$c(H_k) \leq 0.846k.$$  \hfill (4.21)

Note that here we consider the exact bounds on $c^*(H_k)$ and $c(H_k)$ rather than the asymptotic ones as originally written in [BBG12]. It is not known if Hadamard matrices of size $4p$ exist for infinitely many primes $p$. Theorem 4.6.1 requires the existence of Hadamard matrices due to the technique used to lower bound $c^*(H_k)$, which originates from [LMM+12]. It also requires that $k$ is of the form $rp - 1$ for some odd prime $p$ and positive integer $r \geq 4$ due to the technique used to upper-bound $c(H_k)$, which is based on a result of Frankl and Wilson [FW81].

Here we relax the conditions in Theorem 4.6.1 and our result does not rely anymore on the existence of a Hadamard matrix. We show the existence of an infinite family of quarter-orthogonality graphs whose entangled capacity exceeds their Shannon capacity.

**Theorem 4.6.2.** For every odd integer $k \geq 5$, we have

$$c^*(H_k) \geq (k - 1) \left(1 - \frac{4 \log(k + 1)}{k - 3}\right).$$  \hfill (4.21)

Moreover, if $k = 4p^\ell - 1$ where $p$ is an odd prime and $\ell \in \mathbb{N}$, then

$$c(H_k) \leq 0.846k.$$  \hfill (4.22)

In the next sections we prove Theorem 4.6.2.

4.6.1 Lower bound on the entangled Shannon capacity

The proof of the bound (4.21) on the entangled Shannon capacity is based on quantum teleportation (see Section 4.2.3). In operational terms the proof can be interpreted as showing that with $t + 1$ sequential uses of a channel with confusability graph $H_k$, Alice can send Bob $|V|^t$ distinct messages with zero probability of error provided that $t \leq \log \alpha(H_k)/(2 \log(k + 1))$. To give some intuition we explain this operational interpretation before moving on to the proof.
Let $f$ be the map defined in (4.5) and define $\rho^x = f(x)f(x)^T$ for $x \in V(H_k)$. To transmit a sequence $x = (x_1, \ldots, x_t) \in V(H_k)^t$ Alice and Bob may follow the following four-step procedure. First, Alice prepares $(k+1)$-dimensional quantum systems $A_1, \ldots, A_t$ to be in the states $\rho^{x_1}, \ldots, \rho^{x_t}$, respectively. Second, Alice sends the sequence $x$ through the channel by using it $t$ times in a row. This will result in $t$ channel-outputs on Bob’s end of the channel from which he can infer that each $x_i$ belongs to a particular clique in $H_k$. Third, Alice and Bob execute a quantum teleportation scheme after which Bob ends up with quantum systems $Y_1, \ldots, Y_t$ in states $\rho^{x_1}, \ldots, \rho^{x_t}$, respectively. The teleportation step requires that Alice communicates a total of $2t \lceil \log(k+1) \rceil$ bits to Bob. We are now ready to prove the lower bound.

Proof of (4.21). Set $V = V(H_k)$ and let $t \in \mathbb{N}$ be such that $(k+1)^2t \leq \alpha(H_k)$. (Note that this choice of $t$ follows from the fact that Alice needs to use the channel classically to send the measurement outcomes of the teleportation to Bob.) In what follows we construct a trace-1 positive semidefinite matrix $\rho$ and, for every $x \in V^t$, positive semidefinite matrices $\{\rho^x_u : u \in V^{t+1}\}$ satisfying the conditions of Definition 4.3.3, i.e.,

$$\sum_{u \in V^{t+1}} \rho^x_u = \rho \quad \forall x, \quad \rho^x_u \rho^x_v = 0 \quad \forall x \neq y, \{u, v\} \in V^{t+1} \cup E(H_k^{\otimes (t+1)}). \quad (4.23)$$

This implies that $\alpha^*(H_k^{\otimes (t+1)}) \geq |V|^t$.

Let $f : V \to \mathbb{R}^{k+1}$ be the orthonormal representation of $H_k$ defined in (4.5). For $x \in V$ define $\rho^x = f(x)f(x)^T$ and, for $x = (x_1, \ldots, x_t) \in V^t$, define $\rho^x = \rho^{x_1} \otimes \rho^{x_2} \otimes \cdots \otimes \rho^{x_t}$. Notice that $\text{Tr}(\rho^x) = 1$ and that $\rho^x \rho^y = 0$ for every $\{x, y\} \in E(V^{t+1})$. We now consider the quantum teleportation scheme from Section 4.2.3, for the setting where Alice would want to transmit the state $\rho^x$ of a $(k+1)^t$-dimensional quantum system $A$ to Bob. According to (QT1), let $\sigma$ be the maximally entangled state defined over a pair of $(k+1)^t$-dimensional quantum systems $(\mathcal{X}, \mathcal{Y})$, where $\mathcal{X}$ belongs to Alice and $\mathcal{Y}$ to Bob. With $T = (k+1)^{2t}$, let $\{M_i : i \in [T]\}$ be Alice’s measurement on the system $(A, \mathcal{X})$ provided by (QT2), and let $U_1, \ldots, U_T$ be Bob’s unitary operators on $\mathcal{Y}$ given by (QT3). Define

$$\rho = \text{Tr}_A(\sigma), \quad \rho^x_i = \text{Tr}_{(A, \mathcal{X})}(M_i \otimes I)(\rho^x \otimes \sigma) \quad \forall x \in V^t, i \in [T].$$

Since the $M_i$ sum to the identity, for every $x$, we have

$$\sum_{i=1}^T \rho^x_i = \text{Tr}_{(A, \mathcal{X})}(\rho^x \otimes \sigma) = \text{Tr}_A(\rho^x) \text{Tr}_\mathcal{X}(\sigma) = \rho. \quad (4.25)$$
By (QT4), we know that the identity $U_i \rho_i^x U_i^\dagger = \beta_i^x \rho^x$ holds, where $\beta_i^x = \text{Tr}(\rho_i^x)$. Hence, since $f$ is an orthonormal representation, for every edge $\{x, y\} \in E(H^{2k}_k)$, we have

$$\rho_i^x \rho_i^y = (U_i^\dagger \beta_i^x \rho^x U_i)(U_i^\dagger \beta_i^y \rho^y U_i) = \beta_i^x \beta_i^y \rho^x U_i \rho^y U_i = 0. \quad (4.26)$$

Let $W \subseteq V$ be an independent set in $H_k$ with cardinality $|W| = T$ and let $\phi : W \to [T]$ be some bijection. For every $u \in V^{t+1}$ and $x \in V^t$ define

$$\rho_u^x = \begin{cases} \rho_{\phi(u_{t+1})}^x & \text{if } (u_1, \ldots, u_t) = x \text{ and } u_{t+1} \in W \\ 0 & \text{otherwise.} \end{cases}$$

Then, $\sum_{u \in V^{t+1}} \rho_u^x = \sum_{u_{t+1} \in W} \rho_u^x = \sum_{i=1}^T \rho_i^x = \rho$ by (4.25). Next, let $x \neq y \in V^t$ and $\{u, v\} \in V^{t+1} \cup E(H^{2k(t+1)}_k)$; we show that $\rho_u^x \rho_v^y = 0$. This is clear if $x \neq (u_1, \ldots, u_t)$, or $y \neq (v_1, \ldots, v_t)$, or $\{u_{t+1}, v_{t+1}\} \not\subseteq W$. So we may assume $u = (x, u_{t+1})$, $v = (y, v_{t+1})$ and $\{u_{t+1}, v_{t+1}\} \subseteq W$ and thus $\{u, v\} \in E(H^{2k(t+1)}_k)$, $\{x, y\} \in E(H^{2t}_k)$ and $u_{t+1} = v_{t+1}$. Then we have that $\rho_u^x \rho_v^y = \rho_{\phi(u_{t+1})}^x \rho_{\phi(v_{t+1})}^y = 0$ by (4.26).

Hence, for $t$ such that $(k+1)^{2t} \leq \alpha(H_k)$, we have $\alpha^*(H^{2k(t+1)}_k) \geq |V|^t = 2^{(k-1)t}$. This implies

$$c^*(H_k) \geq \frac{1}{t+1} \log \alpha^*(H^{2k(t+1)}_k) \geq \frac{1}{t+1} t(k-1). \quad (4.27)$$

By Lemma 4.2.3 we have $\alpha(H_k) \geq 2^{(k-3)/2}$. Hence, for $k \geq 5$ we can choose the integer $t$ to be equal to $t = \lceil (k-3)/2 \rceil$. From (4.27) we then get

$$c^*(H_k) \geq \frac{4 \log(k+1)}{k-3} \left( \frac{k-3}{4 \log(k+1)} - 1 \right) (k-1) (k-1) \left( 1 - \frac{4 \log(k+1)}{k-3} \right)$$

which gives the claimed result. \qed

### 4.6.2 Upper bound on the Shannon capacity

The upper bound (4.22) on the Shannon capacity of $H_k$ (for certain values of $k$) stated in Theorem 4.6.2 is a corollary of Lemma 4.5.5.

**Proof of (4.22).** By taking the logarithm, dividing by $m$ and taking the limit $m \to \infty$ on both sides of (4.19) we get that for $p$ odd prime, $\ell \in \mathbb{N}$ and $k = 4p^\ell - 1$,

$$c(H_k) \leq 0.846k.$$ \qed
4.7 Separation between classical and entangled source-channel cost rate

Our last contribution concerns the combined source-channel problem for a source that has $H_k$ as characteristic and a channel that has $H_k$ as confusability graph. The result is the following.

**Theorem 4.7.1.** Let $p$ be an odd prime and $\ell \in \mathbb{N}$ such that there exists a Hadamard matrix of size $4p^\ell$. Set $k = 4p^\ell - 1$. Then,

$$\eta^*(H_k, H_k) \leq \frac{\log(k + 1)}{(k - 1) \left(1 - \frac{4\log(k+1)}{k-3}\right)},$$

$$\eta(H_k, H_k) > \frac{0.154k - 1}{k - 1 - \log(k + 1)}.$$  

The proof of Theorem 4.7.1 is given below. The bound on the entangled source-channel cost rate is obtained by concatenating an entanglement-assisted coding scheme for a source with one for a channel. In this way, one obtains a “separated” coding scheme for the source-channel problem, see Section 4.3.2 for details. There we show that the asymptotic cost rate of a separate coding scheme is $R^*(H_k) / c^*(H_k)$ and thus $\eta^*(H_k, H_k) \leq R^*(H_k) / c^*(H_k)$. The bound for the classical parameter $\eta(H_k, H_k)$ relies on properties of the fractional chromatic number and the fact that $H_k$ is vertex-transitive. Let us point out that Theorem 4.7.1 holds for an infinite family of graphs. This follows from the result of Xia and Lu [XL91] in Theorem 4.2.4, since there exist infinitely many $(p, \ell)$-pairs such that $p^{\ell/2} \equiv 1 \mod 4$. (For instance, for $p = 5$ and $\ell = 2i$ with $i \in \mathbb{N}$, $5^i = (4 + 1)^i \equiv 1 \mod 4$.)

Hence, for any $k$ satisfying the condition of the theorem, we have an exponential separation between the entangled and the classical source-channel cost rate as

$$\eta^*(H_k, H_k) \leq \frac{R^*(H_k)}{c^*(H_k)} \leq O\left(\frac{\log k}{k}\right) \ll \Omega(1) \leq \eta(H_k, H_k).$$

As shown in [NTR06], a large separation $\eta(G, H) \ll R(G) / c(H)$ exists for some graphs. But this is not the case for our source-channel combination using $G = H = H_k$. Indeed,

$$\Omega(1) \leq \eta(H_k, H_k) \leq \frac{R(H_k)}{c(H_k)} \leq \frac{\log \chi(H_k)}{\log \alpha(H_k)} \leq \frac{2(k - 1)}{k - 3} \leq O(1),$$

where in the second last inequality we use that $\log \chi(H_k) \leq \log |V(H_k)| = k - 1$ and that $\log \alpha(H_k) \geq (k - 3)/2$ (Lemma 4.2.3).
Now we prove Theorem 4.7.1, separately showing the two bounds (4.28) for $\eta^*$ and (4.29) for $\eta$. The bound (4.28) is obtained by combining (4.16), (4.21) with Proposition 4.3.8. The proof of (4.29) relies on some basic properties of the fractional chromatic number of vertex-transitive graphs, which we now recall.

Let $G$ be a graph and let $\mathcal{I}_G$ denote the collection of its independent sets. For $I \subseteq V(G)$, $\chi^I \in \{0, 1\}^{V(G)}$ denotes its characteristic vector, with $\chi^I_u = 1$ if and only if $u \in I$. We let $\mathbf{1}$ denotes the all-ones vector. The fractional chromatic number $\chi_f(G)$ is a lower bound to the chromatic number, defined by

$$\chi_f(G) = \min \left\{ \sum_{I \in \mathcal{I}_G} \lambda_I : \sum_{I \in \mathcal{I}_G} \lambda_I \chi^I \geq \mathbf{1}, \lambda_I \geq 0 \forall I \in \mathcal{I}_G \right\}.$$ 

An automorphism of $G$ is a permutation $\pi$ of $V(G)$ preserving edges, i.e., $\{\pi(u), \pi(v)\} \in E(G)$ if and only if $\{u, v\} \in E(G)$. The graph $G$ is vertex-transitive if, for any $u, v \in V(G)$, there exists an automorphism $\pi$ of $G$ such that $v = \pi(u)$. We use the following well known facts.

**Lemma 4.7.2.** (see e.g. [GR01, Corollaries 7.4.2, 7.5.2])

(i) For graphs $G$ and $H$, if $G \rightarrow H$ then $\chi_f(G) \leq \chi_f(H)$.

(ii) For a graph $G$, $\chi_f(G) \geq |V(G)|/\alpha(G)$, with equality if $G$ is vertex-transitive.

**Corollary 4.7.3.** Let $G$ and $H$ be vertex-transitive graphs. If there is a homomorphism from $G$ to $H$, then

$$\frac{|V(G)|}{\alpha(G)} \leq \frac{|V(H)|}{\alpha(H)}.$$

As observed in [BBG12], the graph $H_k$ is vertex-transitive; indeed, for any $u \in V(H_k)$, consider the map $v \mapsto u \oplus v$. One can see that taking the strong product and complement of graphs preserves vertex-transitivity. Hence, $H_k^{2m}$ is vertex-transitive for any $n \in \mathbb{N}$.

We are now ready to prove the bound (4.29).

**Proof of (4.29).** Recall the definition of $\eta(H_k, H_k)$ from (4.3). Consider integers $m, n \in \mathbb{N}$ for which $H_k^{2m} \rightarrow H_k^{2n}$. Applying Corollary 4.7.3, we deduce that

$$\frac{|V(H_k^{2m})|}{\alpha(H_k^{2m})} \leq \frac{|V(H_k^{2n})|}{\alpha(H_k^{2n})} = \frac{|V(H_k^{2n})|}{\omega(H_k^{2n})}.$$ 

From Corollary 4.5.4 we have $\omega(H_k^{2n}) = (k+1)^n$. As $|V(H_k)| = 2^{k-1}$ and applying Lemma 4.5.5, we get

$$\frac{2^{(k-1)n}}{2k m H(3/11)} \leq \frac{|V(H_k^{2m})|}{\alpha(H_k^{2m})} \leq \frac{|V(H_k^{2n})|}{\omega(H_k^{2n})} = \frac{2^{(k-1)n}}{(k+1)^n}.$$
After a few algebraic manipulations and taking logarithms the above inequality reduces to
\[
\frac{n}{m} \geq \frac{k(1 - H(3/11)) - 1}{k - 1 - \log(k + 1)} > \frac{0.154k - 1}{k - 1 - \log(k + 1)}.
\]
This gives directly the lower bound (4.29).

### 4.7.1 Stronger bounds based on Hadamard matrices

The reader may have noticed that for the purpose of proving Theorem 4.7.1, we may assume that the integer \( k \) appearing in (4.28) is such that there exists a Hadamard matrix of size \( k + 1 \). The reason for this is that the lower bound on \( \eta(H_k, H_k) \) given in (4.29) is conditional on the existence of such a matrix. With this additional assumption a stronger upper bound on \( \eta^*(H_k, H_k) \) can be proved without the use of quantum teleportation.

To prove this, we bound \( \eta^*_1(H_k, H_k) \) by the rate achievable with separate entangled coding schemes for the source-coding and channel-coding problem, respectively (see Section 4.3.2). To do so, we need a lower bound on the entangled independence number that was obtained previously in [BBG12].

**Lemma 4.7.4 ([BBG12]).** Let \( k \) be a positive integer such that there is a Hadamard matrix of size \( k + 1 \). Then,
\[
\log \alpha^*(H_k) \geq k - 1 - 2 \log(k + 1).
\]

**Lemma 4.7.5.** Let \( k \) be a positive integer such that there is a Hadamard matrix of size \( k + 1 \). Then,
\[
\eta^*_1(H_k, H_k) \leq \left[ \frac{\log(k + 1)}{k - 1 - \log(k + 1)} \right].
\]

**Proof.** Putting together Proposition 4.3.9, Lemma 4.5.3 and Lemma 4.7.4 we have that, for every \( k \) such that there exists a Hadamard matrix of size \( (k + 1) \),
\[
\eta^*_1(H_k, H_k) \leq \left[ \frac{\log \chi^*(H_k)}{\log \alpha^*(H_k)} \right] \leq \left[ \frac{\log(k + 1)}{k - 1 - 2 \log(k + 1)} \right]
\]
which proves the claim.

Since we have \( \eta^*(H_k, H_k) \leq \eta^*_1(H_k, H_k) \), the above result also implies an upper bound of the cost rate attainable by encoding infinitely long sequences of source inputs into single codewords.
4.8 Concluding remarks and open problems

We have shown a separation between classical and entanglement-assisted coding for the zero-error source-channel, source and channel problems. Note that these separations do not hold if asymptotically vanishing error is allowed. We have presented an infinite family of instances for which there is an exponential saving in the minimum asymptotic cost rate of communication for the source-channel and the source coding problems. Moreover, for the channel coding problem we have shown an infinite family of channels for which the entangled Shannon capacity exceeds the classical Shannon capacity by a constant factor. It would be interesting to find a family of channels with a larger separation.

The main result in [NTR06] is that, for the classical source-channel coding problem, there exist situations for which separate encoding is highly suboptimal. Does this happen also in the entanglement-assisted case? This question has a positive answer if there exists a graph $G$ with $R^*(G) > c^*(G)$. In [NTR06] a sufficient condition for a separate encoding to be optimal is also proven, namely that the characteristic or the confusability graph is a perfect graph. It is straightforward to see that this is also a sufficient condition for a separate entanglement-assisted encoding to be optimal. Are there stronger conditions that hold for the entangled case?

One of the most interesting open questions in zero-error classical information theory is the computational complexity of the Witsenhausen rate and of the Shannon capacity. The same question is also open for the entangled counterparts, as well as for the parameters $\chi^*$ and $\alpha^*$.

In Section 4.1, we have seen that the entangled chromatic and independence number generalize the parameters $\chi_q$ and $\alpha_q$ which arise in the context of Bell inequalities and non-local games. In [RM12] it is conjectured that $\alpha^*(G) = \alpha_q(G)$ for every graph $G$. A possible approach to prove that entangled chromatic number and quantum chromatic number are distinct quantities is to prove that the relationship between Kochen-Specker sets and $\chi_q$ found in Chapter 3 does not hold for $\chi^*$.

Additionally, we mention that the existence of a graph $G$ for which $\chi^*(G) < \chi_q(G)$ or $\alpha_q(G) < \alpha^*(G)$ would prove the existence of a non-local game such that every quantum strategy that wins with probability one does not use a maximally entangled state. This is because the source-channel settings studied in this chapter can be seen as non-local games as follows. Alice receives input $x \in X^m$ and Bob receives inputs $u, v \in U^m \times V^n$. Alice produces as output $s \in S^n$, and Bob produces output $y \in X^m$. The winning condition is the following: if $s, v$ is an invalid channel input-output pair, then players win, otherwise players win if $x = y$. By setting the parameters $n, m$ according to the entangled source-channel rate, it is possible to obtain a pseudo-telepathy game (although the classical value can be very close to 1, due to the many winning instances).
Chapter acknowledgments I thank the coauthors of the paper on which this chapter is based. Together, we thank Ronald de Wolf for useful discussions and helpful comments on a preliminary version of the paper, and Dion Gijswijt for helpful discussions.