# ZERO-ERROR SOURCE-CHANNEL CODING WITH ENTANGLEMENT

JOP BRIËT, HARRY BUHRMAN, MONIQUE LAURENT, TERESA PIOVESAN, AND GIANNICOLA SCARPA

Abstract. We study the use of quantum entanglement in the zero-error source-channel coding problem. Here, Alice and Bob are connected by a noisy classical one-way channel, and are given correlated inputs from a random source. Their goal is for Bob to learn Alice's input while using the channel as little as possible. 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, a graph parameter defined by a semidefinite program, gives the best efficiently-computable upper bound on the Shannon capacity and it also upper bounds its entanglement-assisted counterpart. At the same time it was recently shown that the Shannon capacity can be increased if Alice and Bob may use entanglement.

Here we partially extend these results to the source-coding problem and to 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 also show that entanglement can allow for an unbounded improvement of the asymptotic rate of both classical source codes and classical source-channel codes. Our separation 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.

**Keywords: Entanglement, Shannon capacity, Witsenhausen rate, quantum teleportation, graph homomorphism, graph coloring, Lovász theta number, semidefinite programming.**

*Date*: August 21, 2013.

J. B. and H. B. were supported by the European Commission under the project QCS (Grant No. 255961). G. S. was supported by Ronald de Wolf's Vidi grant 639.072.803 from the Netherlands Organization for Scientific Research (NWO).

## Contents

## 1. Introduction

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 [30] 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 [20] for an extensive survey and Lubetzky's PhD thesis [23] for more recent results). The branch of this line of research involving entanglement was started only recently by Cubitt et al. [12]. The possibility for a pair of quantum systems to be entangled is one of the most striking features of quantum mechanics. The typical setting in which this phenomenon manifests itself is where two parties, Alice and Bob, each have a quantum system and perform on it a measurement of their choice. In a celebrated response to a paper of Einstein, Podolsky and Rosen [14], Bell [7] showed that entanglement between Alice's and Bob's systems can cause their measurement outcomes to be distributed according to probability distributions that fall outside the realm of classical physics. In particular, entanglement can give outcome pairs which do not follow a product distribution, nor any convex combination of such distributions. Entanglement therefore allows spatially separated parties to produce so-called *non-local* correlations without needing to communicate. Our main results concern *lower bounds* on the optimum rate of entanglement-assisted source codes and the *advantage* that entanglement can give in the source-channel coding problem. Next we set the stage in detail and state our results precisely.

First, let us recall some definitions of graph theory. Throughout the paper all graphs are assumed to be finite, undirected and without self-loops. For any graph $G$, we denote with $V(G)$ and $E(G)$ its vertex and edge set, respectively. The complement of $G$ is $\overline{G}$, the graph with vertex set $V(G)$ where distinct vertices are adjacent if and only if they are not adjacent in $G$. An independent set is a subset of the vertex set such that no pair is adjacent and the *independence number* $\alpha(G)$ is the maximum cardinality of an independent set in $G$. A clique is a subset of vertices in which each pair is adjacent and the *clique number* $\omega(G)$ is the maximum cardinality of a clique in $G$. Clearly $\alpha(G) = \omega(\overline{G})$. A *proper coloring* is a set of pairwise disjoint independence sets that cover $V(G)$, i.e., an assignment of a color to each vertex such that adjacent vertices receive distinct colors. The *chromatic number* $\chi(G)$ is the minimum number of colors needed for a proper coloring. Thus $\alpha(G) \cdot \chi(G) \geq |V(G)|$. The *strong product* $G \boxtimes H$ of two graphs $G$ and $H$ is the graph whose vertex set is the cartesian

product $V(G) \times V(H)$ and where two distinct vertices $(u_1, u_2)$ and $(v_1, v_2)$ are adjacent if and only if $u_1 = v_1$ or $\{u_1, v_1\} \in E(G)$ and $u_2 = v_2$ or $\{u_2, v_2\} \in E(H)$. For a graph $G$ and $m \in \mathbb{N}$, $G^{\boxtimes m}$ denotes the strong product of $m$ copies of $G$, with vertex set $V(G)^m$ and where two distinct vertices $(u_1, \dots, u_m)$ and $(v_1, \dots, v_m)$ are adjacent if, for all $i \in [m]$, either $u_i = v_i$ or $\{u_i, v_i\} \in E(G)$. A *homomorphism* from a graph $G$ to a graph $H$ is a map $\phi : V(G) \to V(H)$ such that every edge $\{u, v\} \in E(G)$ in $G$ is mapped to an edge $\{\phi(u), \phi(v)\} \in E(H)$ in $H$. If such a map exists, we write $G \longrightarrow H$. Throughout $K_t$ denotes the complete graph on $t$ vertices.
Finally, all the logarithms are in base 2 and for $n \in \mathbb{N}$ we denote $[n] = \{1, 2, \dots, n\}$.

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

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

Witsenhausen [34] and Ferguson and Bailey [15] showed that the zero-error source coding problem can be studied in graph-theoretic terms. Associated with a dual source $\mathcal{M} = (\mathsf{X}, \mathsf{U}, P)$ is its *characteristic graph* $G = (\mathsf{X}, E)$, where $\{x, y\} \in E$ if there exists a $u \in \mathsf{U}$ such that $P(x, u) > 0$ and $P(y, u) > 0$. As such, the edge set identifies the pairs of inputs for Alice which Bob may not be able to distinguish based on his input. It is not difficult to see that every graph is the characteristic graph of a (non-unique) source. Solving a single instance of the zero-error source coding problem for $\mathcal{M}$ is equivalent to finding a proper coloring of $G$. Indeed, Bob's input $u$ reduces the list of Alice's possible inputs to the set $\{x \in \mathsf{X} : P(x, u) > 0\}$ and this set forms a clique in $G$. So Bob can learn Alice's input if she sends him its color. Conversely, a length-1 block-code for $\mathcal{M}$ defines a proper coloring of $G$. To deal with length-$m$ input sequences we consider the graph $G^{\boxtimes m}$ (the strong product of $m$ copies of $G$) whose edges are precisely the pairs of input sequences on Alice's side which Bob

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

cannot distinguish. The *Witsenhausen rate*

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

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

A *discrete channel* $\mathcal{N} = (\mathsf{S}, \mathsf{V}, Q)$ consists of a finite input set $\mathsf{S}$, a (possibly infinite) output set $\mathsf{V}$ and a probability distribution $Q(\cdot|s)$ over $\mathsf{V}$ for each $s \in \mathsf{S}$. Throughout the paper we consider only memoryless channels. If Alice sends an input $s \in \mathsf{S}$ through the channel, then Bob receives the output $v \in \mathsf{V}$ with probability $Q(v|s)$. Their goal is to transmit a binary string $\mathbf{y}$ of, say, $m$ bits from Alice to Bob while using the channel as little as possible. If the parties use a block code of length $n$, then Alice has an encoding function $\mathsf{C} : \{0,1\}^m \to \mathsf{S}^n$ and sends $\mathsf{C}(\mathbf{y})$ through the channel by using it $n$ times in sequence. Bob then receives an output sequence $\mathbf{v} = (v_1, \ldots, v_n)$ on his side of the channel and applies a decoding function $\mathsf{D} : \mathsf{V}^n \to \{0,1\}^m$. The coding scheme $(\mathsf{C}, \mathsf{D})$ works if $\mathsf{D}(\mathbf{v}) = \mathbf{y}$. The *communication rate* of the scheme is $m/n$, the number of bits transmitted per channel use.

Previously, Shannon [30] showed that the zero-error channel coding problem can also be studied in graph-theoretic terms. Associated to a channel $\mathcal{N} = (\mathsf{S}, \mathsf{V}, Q)$ is its *confusability graph* $H = (\mathsf{S}, F)$ where $\{s,t\} \in F$ if there exists a $v \in \mathsf{V}$ such that both $Q(v|s) > 0$ and $Q(v|t) > 0$. The edge set identifies pairs which can lead to identical channel outputs on Bob's side. Sets of non-confusable inputs thus correspond to independent sets in $H$. Codes of block-length $n$ then allow the zero-error transmission of $\alpha(H^{\boxtimes n})$ distinct messages. The *Shannon capacity*

$$c(H) = \lim_{n\to\infty} \frac{1}{n} \log \alpha(H^{\boxtimes n}) \tag{2}$$

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

In the *source-channel coding problem* the parties receive inputs from a dual source $\mathcal{M} = (\mathsf{X}, \mathsf{U}, P)$ and get access to a channel $\mathcal{N} = (\mathsf{S}, \mathsf{V}, Q)$. Their goal is to solve the source coding problem, but now using the channel $\mathcal{N}$ instead of a binary channel. An $(m,n)$*-coding scheme* for this problem consists of an encoding function $\mathsf{C} : \mathsf{X}^m \to \mathsf{S}^n$ and a decoding function $\mathsf{D} : \mathsf{U}^m \times \mathsf{V}^n \to \mathsf{X}^m$ (see Figure 1). The *cost rate* is $n/m$.

Nayak, Tuncel and Rose [26] 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

[2] If a sequence $(a_m)_{m\in\mathbb{N}}$ is sub-additive (i.e., $a_{m+m'} \leq a_m + a_{m'}$ for all $m, m' \in \mathbb{N}$), Fekete's lemma claims that the sequence $(a_m/m)_{m\in\mathbb{N}}$ has a limit, which is equal to its infimum: $\lim_{m\to\infty} a_m/m = \inf_{m\in\mathbb{N}} a_m/m$.

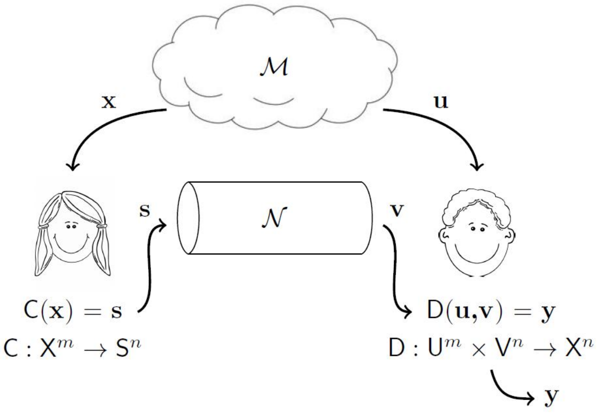
$\mathcal{M}$
$\mathbf{x}$
$\mathbf{u}$
$\mathbf{s}$
$\mathcal{N}$
$\mathbf{v}$
$\mathsf{C}(\mathbf{x}) = \mathbf{s}$
$\mathsf{C} : \mathsf{X}^m \to \mathsf{S}^n$
$\mathsf{D}(\mathbf{u},\mathbf{v}) = \mathbf{y}$
$\mathsf{D} : \mathsf{U}^m \times \mathsf{V}^n \to \mathsf{X}^n$
$\mathbf{y}$

**Figure 1.** The figure illustrates a classical source-coding instance where Alice and Bob use an $(m, n)$-coding scheme $(\mathsf{C}, \mathsf{D})$. The parties receive length-$m$ input strings $\mathbf{x}$ and $\mathbf{u}$, respectively, from a dual source $\mathcal{M} = (\mathsf{X}, \mathsf{U}, P)$ and have a one-way channel $\mathcal{N} = (\mathsf{S}, \mathsf{V}, Q)$. Using $\mathsf{C}$, Alice encodes her input into a string $\mathbf{s} \in \mathsf{S}^n$ which she sends through the channel. After receiving a channel output $\mathbf{v}$, Bob applies $\mathsf{D}$ to the pair $(\mathbf{u}, \mathbf{v})$ to get a string $\mathbf{y}$. The scheme works if $\mathbf{y} = \mathbf{x}$.

from $G^{\boxtimes m}$ to $\overline{H^{\boxtimes n}}$. Then, the parameter

$$\eta(G, H) := \lim_{m\to\infty} \frac{1}{m} \min \left\{ n \in \mathbb{N} : \, G^{\boxtimes m} \longrightarrow \overline{H^{\boxtimes n}} \right\} \tag{3}$$

gives the minimum asymptotic cost rate of a zero-error code. We will assume throughout that both $G$ and $\overline{H}$ contain at least one edge. (Indeed, if $G$ has no edge then $\eta(G, H) = 0$ for any $H$ and, if $G$ has at least one edge, then $\eta(G, H)$ is well defined only if $\overline{H}$ has at least one edge.) To see that the limit exists, observe that the parameter

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

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

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

**Lemma 1.1.** *Let $G$ and $H$ be graphs such that both $G$ and $\overline{H}$ have at least one edge. Then,*

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

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

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

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

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

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

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

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

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

and thus

$$\frac{n(m)}{\log\alpha(H^{\boxtimes n(m)})} \le \frac{n(m)}{m} < \frac{n(m)}{\log\alpha(H^{\boxtimes(n(m)-1)})}. \tag{4}$$

As $c(H) = \sup_n \log\alpha(H^{\boxtimes n})/n$, using the left most inequality in (4) we deduce that

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

for all $m$. Taking the limit, we obtain the inequality $1/c(H) \le \lim_{m\to\infty} n(m)/m = \eta_m(K_2, H)$. Next, as $\eta_m(K_2, H) = \inf_m n(m)/m$, using the right most inequality in (4) we deduce that

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

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

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 [32]. But when no errors can be tolerated at all, Nayak, Tuncel and Rose [26] showed that separated codes can be highly suboptimal. In terms of the above graph parameters, this says that in general the inequality $\eta(G, H) \le R(G)/c(H)$

holds (this inequality is implied in [26], we will give an explicit proof of it in Proposition 3.4), but that for some families of graphs there can be a large separation: $\eta(G,H) \ll R(G)/c(H)$.

1.2. **Entanglement-assisted source-channel coding.** The entanglement-assisted model of source-channel coding is roughly as follows (details are given in Section 3). After receiving an input from the source, Alice performs a measurement on her quantum system. Based on her measurement outcome she sends a (classical) message through the channel. Then, after receiving his source input and channel output, Bob performs a measurement on his quantum system. The entanglement-assisted scheme works if Bob's measurement outcome equals Alice's source input. In graph-theoretic terms this model gives the following algebraic definition of the entangled variant of $\eta(G,H)$. Recall that a *positive semidefinite matrix* is a Hermitian matrix whose eigenvalues are non-negative. To indicate that a matrix $\rho$ is positive semidefinite we use the standard notation $\rho \succeq 0$.

**Definition 1.2** (Entangled cost rate). *For graphs $G, H$ and $m \in \mathbb{N}$, define $\eta^\star_m(G,H)$ as the minimum integer $n \in \mathbb{N}$ for which there exist $d \in \mathbb{N}$ and $d \times d$ positive semidefinite matrices $\rho$ and $\{\rho^{\mathbf{s}}_{\mathbf{x}} : \mathbf{x} \in V(G^{\boxtimes m}), \mathbf{s} \in V(H^{\boxtimes n})\}$ such that $\mathsf{Tr}(\rho) = 1$ and*

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

*The* entangled cost rate *is defined by*

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

As for the classical counterpart, we assume throughout that both graphs $G$ and $\overline{H}$ contain at least one edge. We regain the parameter $\eta(G,H)$ if we restrict the above matrices $\rho$ and $\rho^{\mathbf{s}}_{\mathbf{x}}$ to be $\{0,1\}$-valued scalars. Thus sharing an entangled quantum system cannot make the coding scheme worse and $\eta^\star(G,H) \leq \eta(G,H)$. As in the classical case, the parameter $\eta^\star_m(G,H)$ is sub-additive (see Lemma 3.1), hence the parameter $\eta^\star(G,H)$ is well defined and can be equivalently written as the infimum of $\eta^\star_m(G,H)/m$.

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

**Definition 1.3** (Entangled chromatic number). *For a graph $G$, define $\chi^\star(G)$ as the minimum integer $t \in \mathbb{N}$ for which there exist $d \in \mathbb{N}$ and $d \times d$ positive semidefinite matrices $\rho$ and $\{\rho^i_u : u \in V(G),\, i \in [t]\}$ such that $\mathsf{Tr}(\rho) = 1$ and*

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

*The* entangled Witsenhausen rate *is defined by*

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

In Lemma 3.2 we show that $\chi^\star$ is sub-multiplicative and thus the entangled Witsenhausen rate can be equivalently defined as the infimum: $R^\star(G) = \inf_m \log \chi^\star(G^{\boxtimes m})/m$.

**Definition 1.4** (Entangled independence number)**.** *For a graph $H$, define $\alpha^\star(H)$ as the maximum integer $M \in \mathbb{N}$ for which there exist $d \in \mathbb{N}$ and $d \times d$ positive semidefinite matrices $\rho$ and $\{\rho_i^u : i \in [M],\, u \in V(H)\}$ such that $\mathsf{Tr}(\rho) = 1$ and*

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

*The* entangled Shannon capacity *is defined by*

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

The parameter $\alpha^\star(H)$ was introduced by Cubitt et al. [12] and it is known to be super-multiplicative. Hence, in the definition of $c^\star(H)$ the limit can be replaced with the supremum. Observe that in the above definitions it would suffice to require that $\rho$ is not identically zero (as it can then be rescaled to have trace 1).

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

**Lemma 1.5.** *Let $G$ and $H$ be graphs such that both $G$ and $\overline{H}$ have at least one edge. Then,*

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

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

Since $K_2^{\boxtimes m} = K_{2^m}$, it follows from the definitions that $\eta_m^\star(K_2, H)$ is the minimum $n \in \mathbb{N}$ such that $\alpha^\star(H^{\boxtimes n}) \geq 2^m$ or, equivalently, $\log \alpha^\star(H^{\boxtimes n}) \geq m$. Now we use the same techniques as in Lemma 1.1 to prove that $1/c^\star(H) = \eta^\star(K_2, H)$. □

In [12] it is shown that $\alpha^\star(H)$ can be strictly larger than $\alpha(H)$, meaning that the number of messages that can be sent with a single use of a channel can be increased with the use of entanglement (see also Mančinska, Severini and Scarpa [24]). This result was subsequently strengthened by Leung, Mančinska, Matthews, Ozols and Roy [21] and Briët, Buhrman and Gijswijt [9], who found families of graphs for which $c^\star(H) > c(H)$.

To the best of our knowledge, neither source nor source-channel coding were considered in the context of shared entanglement before. However, in the context of Bell inequalities, Cameron et al. [10] studied the *quantum chromatic number* $\chi_q(G)$, and Roberson and Mančinska [29] considered a variant of the quantum independence number $\alpha_q(H)$. These parameters can be obtained from the respective definitions of $\chi^\star$ and $\alpha^\star$ given above, if we require $\rho$ to be the identity matrix and if we further restrict the other positive semidefinite matrices to be orthogonal projections (matrices that satisfy $P^2 = P$). Furthermore, we regain $\chi$ and $\alpha$ if we further restrict these matrices to be $\{0,1\}$-valued scalars. It thus follows immediately that

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

It is well-known that determining the classical chromatic and independence numbers of a graph are NP-hard problems. Determining the Shannon capacity and the Witsenhausen rate appears to be even harder (we do not even know if they are computable). Despite substantial efforts, the properties of these parameters are still only partially understood (see [3, 4] and references therein). For example, the largest odd cycle for which the Shannon capacity has been determined is $C_5$ and the decidability 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 parameters $\chi^\star, \alpha^\star, \chi_q, \alpha_q, R^\star$ and $c^\star$ is an open problem.

1.3. **Outline of the paper.** In Section 1 we introduced the problems and the basic notions. In Section 2 we present our main results. In Section 3 we give a brief introduction to quantum information theory, we explain the quantum teleportation scheme and the entangled-assisted source-channel coding protocol. Some properties of the entangled parameters are also presented there. The proofs of our main results are given in Sections 4 - 7. Finally in Section 8 we summarize our results and mention a few open questions.

## 2. Our results

2.1. **The entangled chromatic number and Szegedy's number.** Here we explain our lower bound on the entangled chromatic number. We show that $\chi^\star(G)$ is lower bounded by an efficiently computable graph parameter, namely a variant of the famous *theta number* introduced by Szegedy [31]. The theta number itself was originally introduced by Lovász [22] to solve a long-standing problem posed by Shannon [30]: computing the Shannon capacity of the five-cycle. Out of the many equivalent formulations of the theta number (see [19] for a survey), the following is the most appropriate for our setting:

$$\begin{aligned} \vartheta(G) = \min\Big\{ \lambda : \quad & \exists\, Z \in \mathbb{R}^{V(G)\times V(G)},\ Z \succeq 0, \\ & Z(u,u) = \lambda - 1 \ \text{ for } u \in V(G), \\ & Z(u,v) = -1 \ \text{ for } \{u,v\} \notin E(G) \Big\}. \end{aligned} \tag{5}$$

Lovász [22] proved that $\alpha(G) \leq \vartheta(G) \leq \chi(\overline{G})$ holds (this inequality is often referred to as the Sandwich Theorem [19]). The theta number is defined by a semidefinite program, hence it can be approximated to within arbitrary precision in polynomial time and thus it gives a tractable and in many cases useful bound for both $\alpha$ and $\chi$.

Szegedy [31] introduced the following strengthening of the theta number, which includes an extra linear constraint:

$$
\begin{aligned}
\vartheta^+(G) = \min \Big\{ \lambda : \ \ & \exists\, Z \in \mathbb{R}^{V(G)\times V(G)},\ Z \succeq 0, \\
& Z(u,u) = \lambda - 1 \ \text{ for } u \in V(G), \\
& Z(u,v) = -1 \ \text{ for } \{u,v\} \notin E(G), \\
& Z(u,v) \geq -1 \ \text{ for } \{u,v\} \in E(G) \Big\}.
\end{aligned}
\tag{6}
$$

Szegedy's number satisfies $\alpha(G) \leq \vartheta(G) \leq \vartheta^+(G) \leq \chi(\overline{G})$. Lovász [22] proved that $\vartheta$ is *multiplicative* under the strong graph product, that is, $\vartheta(G \boxtimes H) = \vartheta(G)\vartheta(H)$. Moreover Knuth [19] showed that $\vartheta(\overline{G \boxtimes H}) = \vartheta(\overline{G})\vartheta(\overline{H})$; it is unknown if this identity holds for $\vartheta^+$ [25]. The identities of Lovász and Knuth give for any graph $G$ and $m \in \mathbb{N}$:

$$\vartheta(\overline{G}^{\boxtimes m}) = \vartheta(\overline{G^{\boxtimes m}}) = \vartheta(\overline{G})^m. \tag{7}$$

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

$$c(G) \leq \log \vartheta(G) \leq R(\overline{G}).$$

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

Our first main result is 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. For the proof we refer to Section 4.

**Theorem 2.1.** *For any graph $G$, we have*

$$\vartheta^+(G) \leq \chi^\star(\overline{G}), \tag{8}$$

$$\log \vartheta(G) \leq R^\star(\overline{G}). \tag{9}$$

In [29] it is shown that $\vartheta(G) \leq \chi_q(\overline{G})$ holds. Theorem 2.1 thus strengthens this bound as it gives $\vartheta(G) \leq \vartheta^+(G) \leq \chi^\star(\overline{G}) \leq \chi_q(\overline{G})$.

Beigi [6] and Duan, Severini and Winter [13] proved that $\vartheta(G)$ upper bounds $\alpha^\star(G)$. The above-mentioned relations therefore imply the following sequence of inequalities:

$$c(G) \leq c^\star(G) \leq \log \vartheta(G) \leq R^\star(\overline{G}) \leq R(\overline{G}).$$

2.2. **Lower cost rates with entanglement.** We give quantitative bounds on the advantage of sharing entanglement for the following three parameters: the Witsenhausen rate, the Shannon capacity and the cost rate of certain source-channel combinations.

2.2.1. *Quarter-orthogonality graphs and Hadamard matrices.* To show separations between the classical and entangled variants of the above-mentioned parameters, we use the following family of graphs (also considered in [9] for similar reasons).

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

$$\begin{array}{rcccc} f & : & V(H_k) & \longrightarrow & \mathbb{R}^{k+1} \\ & & u & \longmapsto & \phi(u)/\sqrt{k+1}. \end{array} \tag{10}$$

**Lemma 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. It is easy to see that $|W| = 2^{(k-3)/2}$ and that $W$ is an independent set since it does not contain pairs of strings at Hamming distance $(k+1)/2$. □

Some of our results rely on the existence of certain Hadamard matrices. A *Hadamard matrix* is a square matrix $A \in \{-1,1\}^{\ell\times\ell}$ that satisfies $AA^{\mathsf{T}} = \ell I$. The size $\ell$ of a Hadamard matrix must necessarily be 2 or a multiple of 4 and the famous Hadamard conjecture (usually attributed to Paley [28]) states that for every $\ell$ 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. We will use a family constructed by Xia and Liu [35] (see for example [36, 33, 11, 38, 37] for closely related constructions).

**Theorem 2.4** (Xia and Liu [35])**.** *Let $q$ be a prime power such that $q \equiv 1 \bmod 4$. Then, there exists a Hadamard matrix of size $4q^2$.*

We also use the following result regarding the graph $H_k$.

**Proposition 2.5** (Briët, Buhrman and Gijswijt [9])**.** *Let $k$ be a positive integer such that there exists a Hadamard matrix of size $k+1$. Then, $\omega(H_k) \geq k+1$.*

2.2.2. *Improving the Witsenhausen rate.* Our first separation result shows an exponential gap between the entangled and classical Witsenhausen rates of quarter-orthogonality graphs. An instance of the source coding problem for a source with characteristic graph $H_k$ arises in the following setting. A source has a set of messages $\mathcal{H}$ and maps each message $h \in \mathcal{H}$ to a $k$-bit codeword $C(h)$ using an error-correcting code of distance $(k+1)/2$ (recall that we assume $k$ to be odd). Next, it encrypts the codeword by "shifting" it by a random $k$-bit string $\Delta$ of even Hamming weight to get $C^{\Delta}(h) = C(h) \oplus \Delta$ (where $\oplus$ denotes entrywise addition modulo 2). Alice is given $C^{\Delta}(h)$ and somehow Bob learns the coding scheme $C$ and the shift $\Delta$. If Bob wants to learn $h$, then Alice needs to send Bob $\log \chi(H_k)$ bits in the classical case and $\log \chi^\star(H_k)$ bits in the entangled case. In the limit of many sequential messages this translates to $R(H_k)$ and $R^\star(H_k)$ bits, respectively.

**Theorem 2.6.** *For every odd integer $k$, we have*

$$R^\star(H_k) \leq \log(k+1). \tag{11}$$

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

$$R(H_k) \geq 0.154k - 1. \tag{12}$$

The proof of the theorem is given in Section 5.

2.2.3. *Improving the Shannon capacity.* Our second separation result is a strengthening of the following result of [9], 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 2.7** (Briët, Buhrman and Gijswijt [9])**.** *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 [9]. It is not known if Hadamard matrices of size $4p$ exist for infinitely many primes $p$. Theorem 2.7 requires the existence of Hadamard matrices due to the technique used to lower bound $c^\star(H_k)$, which originates from [21]. 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 [16].

Here we relax the conditions in Theorem 2.7 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 2.8.** *For every odd integer $k \geq 5$, we have*

$$c^\star(H_k) \geq (k-1)\left(1 - \frac{4\log(k+1)}{k-3}\right). \tag{13}$$

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

$$c(H_k) \leq 0.846\, k. \tag{14}$$

We prove Theorem 2.8 in Section 6. To prove (13) we use a technique that is based on the quantum teleportation scheme of Bennett et al. [8]. This proof technique appears not to have been considered before in the context of zero-error entanglement-assisted communication. The proof of (14) combines an instance of the linear algebra method due to Alon [2] with a construction of certain low-degree polynomials over a finite field for a low-degree representations of the OR-function due to Barrington, Beigel and Rudich [5]. Roughly this combination was previously used in the context of Ramsey graphs [17].

2.2.4. *Improving on source-channel codes.* Our last contribution concerns the combined source-channel problem for a source and channel that both have $H_k$ as characteristic and confusability graph, respectively. The result is the following.

**Theorem 2.9.** *Let $p$ be an odd prime and $\ell \in \mathbb{N}$ such that there exists a Hadamard matrix of size $4p^\ell$. Set $k = 4p^\ell - 1$. Then,*

$$\eta^\star(H_k, H_k) \leq \frac{\log(k+1)}{(k-1)\left(1 - \frac{4\log(k+1)}{k-3}\right)}, \tag{15}$$

$$\eta(H_k, H_k) > \frac{0.154\, k - 1}{k - 1 - \log(k+1)}. \tag{16}$$

The proof of Theorem 2.9 is given in Section 7. 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 3.4.2 for details. There we show that the asymptotic cost rate of a separate coding scheme is $R^\star(H_k)/c^\star(H_k)$ and thus $\eta^\star(H_k, H_k) \leq R^\star(H_k)/c^\star(H_k)$. The bound for the classical parameter $\eta(H_k, H_k)$ relies on the No-Homomorphism Lemma due to Albertson and Collins [1] and fact that $H_k$ is vertex-transitive. Let us point out that Theorem 2.9 holds for an infinite family of graphs. This follows from the result of Xia and Lu [35] in Theorem 2.4, since there exist infinitely many $(p, \ell)$-pairs such that $p^{\ell/2} \equiv 1 \bmod 4$. (For instance, for $p = 5$ and $\ell = 2i$ with $i \in \mathbb{N}$, $5^i = (4+1)^i \equiv 1 \bmod 4$.)

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

$$\eta^\star(H_k, H_k) \leq \frac{R^\star(H_k)}{c^\star(H_k)} \leq O\Big(\frac{\log k}{k}\Big) \;\text{ while }\; \eta(H_k, H_k) \geq \Omega(1).$$

As shown in [26], 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 2.3).

## 3. Entanglement-assisted source-channel coding

In this section we describe the quantum teleportation scheme and the model of entanglement-assisted source-channel coding. We also prove properties of the entangled parameters. We begin by describing the elements of quantum information theory that appear in the subsequent discussion. For more on quantum information theory we refer to Nielsen and Chuang [27].

### 3.1. **Quantum states and measurements.**

A *quantum register* is an abstract physical system with which experimenters (Alice and Bob) may interact. A quantum register is represented by a finite-dimensional complex vector space and the register is $d$-dimensional if this vector space is $\mathbb{C}^d$. 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 $\rho$, the quantum register $\mathcal{A}$ is said to be *in* state $\rho$. 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, and $\mathbb{C}^{d_A \times d_A} \otimes \mathbb{C}^{d_B \times d_B}$ is the tensor product space, consisting of all linear combinations of matrices of the form $\rho_A \otimes \rho_B$, where $\rho_A \in \mathbb{C}^{d_A \times d_A}$ and $\rho_B \in \mathbb{C}^{d_B \times d_B}$. The pair of systems $(\mathcal{A}, \mathcal{B})$ is said to be *entangled* if it is in a state $\rho$ which is *not a convex combination of states of the form* $\rho_A \otimes \rho_B$.

A *$t$-outcome measurement* is a collection $\mathsf{M} = \{M_i \in \mathbb{C}^{d \times d} : i \in [t]\}$ of positive semidefinite matrices $M_i$ that satisfy $\sum_{i=1}^t M_i = I$, where $I$ is the identity matrix. A measurement describes an experiment which one may perform on a $d$-dimensional quantum register. If Alice performs a $t$-outcome measurement $\mathsf{M}$ on a register $\mathcal{A}$ which is in a state $\rho$, then she will observe a random variable $\lambda$ over the set $[t]$ whose probability distribution is given by $\Pr[\lambda = i] = \mathsf{Tr}(M_i \rho)$. In the event that $\lambda = i$, we say that Alice gets measurement outcome $i$.

Below we will consider settings where Alice and Bob hold (possibly entangled) quantum registers $\mathcal{A}$ and $\mathcal{B}$, respectively, and they each perform a measurement. For this we introduce a linear operator called the *partial trace*. For matrices $A \in \mathbb{C}^{d_A \times d_A}$ 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 $\mathbb{C}^{d_A \times d_A} \otimes \mathbb{C}^{d_B \times d_B}$.

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

It is a simple fact that, for any set of states $\rho_1, \dots, \rho_t \in \mathbb{C}^{d\times d}$ that are pairwise orthogonal (i.e., $\rho_i\rho_j = 0$ if $i \neq j$), there exists a measurement $\{M_i \in \mathbb{C}^{d\times d} : i \in [t]\}$ such that $\mathsf{Tr}(M_i\rho_j) = \delta_{i,j}$. If Bob knows that his quantum register is in one of these states, say $\rho_k$, then the measurement will tell him which state it is since he will get measurement outcome $k$ with probability 1. Finally, we mention that Bob may alter a state $\rho \in \mathbb{C}^{d\times d}$ of his register by performing a unitary transformation. This is a mapping $\rho \mapsto U\rho U^*$, where $U \in \mathbb{C}^{d\times d}$ satisfies $UU^* = I$.

3.2. **Quantum teleportation.** Next we briefly explain the quantum teleportation scheme of Bennett et al. [8]. This 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. The essential features of this scheme are as follows (we refer to [8] and [27, pp. 26–28] for the details). Suppose that Alice has a local $d$-dimensional quantum register $\mathcal{A}$ in state $\rho$. Suppose in addition that Alice and Bob each have local $d$-dimensional registers $\mathcal{X}$ and $\mathcal{Y}$, respectively. For this set-up, it follows easily from the basic quantum teleportation scheme of [8] that there exist:

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

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

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

That is, at the end of the protocol,

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

3.3. **The protocol for entanglement-assisted coding.** We now explain the model of entanglement-assisted source-channel coding, also pictured by Figure 2.

Similar to classical source-channel coding, Alice and Bob receive inputs from a dual source $\mathcal{M} = (\mathsf{X}, \mathsf{U}, P)$ and Alice can send messages through a classical channel $\mathcal{N} = (\mathsf{S}, \mathsf{V}, Q)$. Their goal is for Bob to learn Alice's input, minimizing the number of channel uses per input sequence of given length. In addition Alice and Bob have a local quantum register $\mathcal{A}$ and

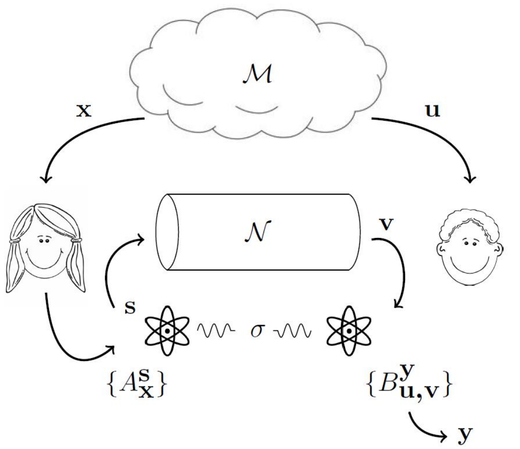

**Figure 2.** The figure illustrates the entanglement-assisted source-channel coding protocol. After receiving a source-input $\mathbf{x} \in \mathsf{X}^m$, Alice performs a measurement $\{A_{\mathbf{x}}^{\mathbf{s}} : \mathbf{s} \in \mathsf{S}^n\}$ on her part of an entangled state $\sigma$ which she shares with Bob. She sends her measurement outcome $\mathbf{s}$ through the channel, upon which Bob—who previously already received a source-input $\mathbf{u}$—receives a channel-output $\mathbf{v} \in \mathsf{V}^n$. Bob performs a measurement $\{B_{\mathbf{u},\mathbf{v}}^{\mathbf{y}} : \mathbf{y} \in \mathsf{X}^m\}$ on his part of $\sigma$ and obtains a measurement outcome $\mathbf{y} \in \mathsf{X}^m$.

$\mathcal{B}$, respectively, and they share an entangled state $\sigma$ in $(\mathcal{A}, \mathcal{B})$ on which they can perform measurements. The entanglement-assisted source-channel coding protocol goes as follows:

(1) Alice and Bob receive inputs $\mathbf{x} \in \mathsf{X}^m$ and $\mathbf{u} \in \mathsf{U}^m$, respectively, from the dual source $\mathcal{M}$;
(2) Alice performs a measurement $\{A_{\mathbf{x}}^{\mathbf{s}}\}_{\mathbf{s} \in \mathsf{S}^n}$ (which can depend on $\mathbf{x}$) on $\mathcal{A}$ and gets $\mathbf{s}$ as outcome;
(3) Alice sends $\mathbf{s}$ through the channel $\mathcal{N}$ and Bob receives $\mathbf{v} \in \mathsf{V}^n$;
(4) Bob performs a measurement $\{B_{\mathbf{u},\mathbf{v}}^{\mathbf{y}}\}_{\mathbf{y} \in \mathsf{X}^m}$ (which can depend on $\mathbf{u}$ and $\mathbf{v}$) on $\mathcal{B}$ and gets $\mathbf{y} \in \mathsf{X}^m$ as outcome.

Recall that if the two parties share no entanglement, then a zero-error $(m, n)$-coding scheme is equivalent to a homomorphism from $G^{\boxtimes m}$ to $\overline{H^{\boxtimes n}}$, i.e., a map that sends edges of $G^{\boxtimes m}$ to non-edges of $H^{\boxtimes n}$, where $G$ is the characteristic graph of $\mathcal{M}$ and $H$ is the confusability graph of $\mathcal{N}$. Analogously, the entangled-assisted protocol is successful if only if, for every edge $\{\mathbf{x}, \mathbf{y}\}$ in $G^{\boxtimes m}$ and every non-edge $\{\mathbf{s}, \mathbf{t}\}$ in $H^{\boxtimes n}$, we have that $\mathsf{Tr}_{\mathcal{A}}\left((A_{\mathbf{x}}^{\mathbf{s}} \otimes I)\sigma\right)$ is orthogonal to $\mathsf{Tr}_{\mathcal{A}}\left((A_{\mathbf{y}}^{\mathbf{t}} \otimes I)\sigma\right)$. The intuition being that undistinguishable 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. The algebraic characterization of $\eta^\star$ given in Definition 1.2 can now be derived by putting $\rho_{\mathbf{x}}^{\mathbf{s}} = \mathsf{Tr}_{\mathcal{A}}\left((A_{\mathbf{x}}^{\mathbf{s}} \otimes \mathbf{I})\sigma\right)$ and $\rho = \mathsf{Tr}_{\mathcal{A}}(\sigma)$.

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

3.4.1. *Sub-additivity of $\eta^\star_m$ and sub-multiplicativity of $\chi^\star$.*

**Lemma 3.1.** *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 $\eta^\star_{m+m'}(G,H) \le \eta^\star_m(G,H) + \eta^\star_{m'}(G,H)$.*

*Proof:* Let $\varphi, \{\varphi^{\mathbf{s}}_{\mathbf{x}} : \mathbf{x} \in V(G^{\boxtimes m}), \mathbf{s} \in V(H^{\boxtimes n})\}$ be a set of positive semidefinite matrices that witness $\eta^\star_m(G,H) = n$ (Definition 1.2) and let $\psi, \{\psi^{\mathbf{t}}_{\mathbf{y}} : \mathbf{y} \in V(G^{\boxtimes m'}), \mathbf{t} \in V(H^{\boxtimes n'})\}$ be a collection of matrices which are a solution for $\eta^\star_{m'}(G,H) = n'$. Notice that every vertex $\mathbf{w}$ of $G^{\boxtimes (m+m')}$ can be written as $\mathbf{w} = (\mathbf{x}, \mathbf{y})$ where $\mathbf{x} \in V(G^{\boxtimes m})$ and $\mathbf{y} \in V(G^{\boxtimes m'})$ and similarly any $\mathbf{r} \in V(H^{\boxtimes (n+n')})$, $\mathbf{r} = (\mathbf{s}, \mathbf{t})$ where $\mathbf{s} \in V(H^{\boxtimes n})$ and $\mathbf{t} \in V(H^{\boxtimes n'})$. We create a solution for $\eta^\star_{m+m'}(G,H)$ as follows. Let $\rho = \varphi \otimes \psi$ and for every vertex $(\mathbf{x}, \mathbf{y}) \in V(G^{\boxtimes (m+m')})$ and $(\mathbf{s}, \mathbf{t}) \in V(H^{\boxtimes (n+n')})$ define

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

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

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

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

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

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

**Lemma 3.2.** *For two graphs $G$ and $H$, $\chi^\star(G \boxtimes H) \le \chi^\star(G)\chi^\star(H)$.*

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

$$\rho^{\mathbf{k}}_{(u,v)} = \varphi^i_u \otimes \psi^j_v.$$

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

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

**Lemma 3.3.** *Given graphs $G, H$ and positive integers $n, m$, we have*

$$\chi(G^{\boxtimes m}) \leq \alpha(H^{\boxtimes n}) \Longrightarrow \eta_m(G,H) \leq n, \tag{17}$$

$$\chi^\star(G^{\boxtimes m}) \leq \alpha^\star(H^{\boxtimes n}) \Longrightarrow \eta^\star_m(G,H) \leq n. \tag{18}$$

*Proof:* If $\chi(G^{\boxtimes m}) \leq \alpha(H^{\boxtimes n})$, then there is a homomorphism from $G^{\boxtimes m}$ to $\overline{H^{\boxtimes n}}$ and thus $\eta_m(G,H) \leq n$, which shows (17). We now show (18). For this set $t = \chi^\star(G^{\boxtimes m})$ and $M = \alpha^\star(H^{\boxtimes n})$, with $t \leq M$ by assumption. Let $\varphi, \{\varphi^i_{\mathbf{x}} : \mathbf{x} \in V(G^{\boxtimes m}), i \in [t]\}$ be a collection of positive semidefinite matrices forming a solution for $\chi^\star(G^{\boxtimes m})$ and let the set of positive semidefinite matrices $\psi, \{\psi^{\mathbf{s}}_i : \mathbf{s} \in V(H^{\boxtimes n}), i \in [M]\}$ be feasible for $\alpha^\star(H^{\boxtimes n})$. We construct a solution for $\eta^\star_m(G,H)$ as follows. For $\mathbf{x} \in V(G^{\boxtimes m})$ and $\mathbf{s} \in V(H^{\boxtimes n})$ set

$$\rho^{\mathbf{s}}_{\mathbf{x}} = \sum_{i \in [t]} \varphi^i_{\mathbf{x}} \otimes \psi^{\mathbf{s}}_i \quad \text{and} \quad \rho = \varphi \otimes \psi.$$

Then, we have that $\mathsf{Tr}(\rho) = \mathsf{Tr}(\varphi \otimes \psi) = 1$ and, for every $\mathbf{x} \in V(G^{\boxtimes m})$, we get that $\sum_{\mathbf{s} \in V(H^{\boxtimes n})} \rho^{\mathbf{s}}_{\mathbf{x}} = \sum_{\mathbf{s} \in V(H^{\boxtimes n})} \sum_{i \in [t]} \varphi^i_{\mathbf{x}} \otimes \psi^{\mathbf{s}}_i = \sum_{i \in [t]} \varphi^i_{\mathbf{x}} \otimes (\sum_{\mathbf{s} \in V(H^{\boxtimes n})} \psi^{\mathbf{s}}_i)$ is equal to $\varphi \otimes \psi = \rho$. Moreover, for every $\{\mathbf{x}, \mathbf{y}\} \in E(G^{\boxtimes m})$ and every $\{\mathbf{s}, \mathbf{t}\}$ equal or adjacent in $H^{\boxtimes n}$,

$$\begin{aligned} \rho^{\mathbf{s}}_{\mathbf{x}} \rho^{\mathbf{t}}_{\mathbf{y}} &= \big(\sum_{i \in [t]} \varphi^i_{\mathbf{x}} \otimes \psi^{\mathbf{s}}_i\big)\big(\sum_{j \in [t]} \varphi^j_{\mathbf{y}} \otimes \psi^{\mathbf{t}}_j\big) = \sum_{i \in [t]} \sum_{j \in [t]} \varphi^i_{\mathbf{x}} \varphi^j_{\mathbf{y}} \otimes \psi^{\mathbf{s}}_i \psi^{\mathbf{t}}_j \\ &= \sum_{i \in [t]} \varphi^i_{\mathbf{x}} \varphi^i_{\mathbf{y}} \otimes \psi^{\mathbf{s}}_i \psi^{\mathbf{t}}_i + \sum_{i,j \in [t], i \neq j} \varphi^i_{\mathbf{x}} \varphi^j_{\mathbf{y}} \otimes \psi^{\mathbf{s}}_i \psi^{\mathbf{t}}_j = 0, \end{aligned}$$

where the last identity uses the orthogonality conditions of the matrices $\varphi^i_{\mathbf{x}}$ and $\psi^{\mathbf{s}}_i$. Hence $\rho, \{\rho^{\mathbf{s}}_{\mathbf{x}} : \mathbf{x} \in V(G^{\boxtimes m}), \mathbf{s} \in V(H^{\boxtimes n})\}$ is a feasible solution for $\eta^\star_m(G,H) \leq n$. □

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

**Proposition 3.4.** *Let $G$ and $H$ be graphs and assume that both $G$ and $\overline{H}$ have at least one edge. Then,*

$$\eta(G,H) \leq \frac{R(G)}{c(H)} = \lim_{m\to\infty} \frac{1}{m} \min\{n : \chi(G^{\boxtimes m}) \leq \alpha(H^{\boxtimes n})\}, \tag{19}$$

$$\eta^\star(G,H) \leq \frac{R^\star(G)}{c^\star(H)} = \lim_{m\to\infty} \frac{1}{m} \min\{n : \chi^\star(G^{\boxtimes m}) \leq \alpha^\star(H^{\boxtimes n})\}. \tag{20}$$

*Proof:* We show (19); we omit the proof of (20) which is analogous (and uses (18)). From (17) we have the inequality:

$$\eta_m(G,H) \leq \epsilon_m(G,H) := \min\{n : \chi(G^{\boxtimes m}) \leq \alpha(H^{\boxtimes n})\},$$

which implies $\eta(G,H) \leq \lim_{m\to\infty} \epsilon_m(G,H)/m$. Next we show that this limit is equal to $R(G)/c(H)$, which concludes the proof of (19). Setting $n = \epsilon_m(G,H)$, we have that $\alpha(H^{\boxtimes (n-1)}) < \chi(G^{\boxtimes m}) \leq \alpha(H^{\boxtimes n})$, implying

$$\frac{R(G)}{c(H)} \leq \frac{\log \chi(G^{\boxtimes m})}{m} \frac{n}{\log \alpha(H^{\boxtimes n})} \leq \frac{n}{m} < \frac{n}{n-1} \frac{\log \chi(G^{\boxtimes m})}{m} \frac{n-1}{\log \alpha(H^{\boxtimes n-1})}.$$

Taking limits as $m \to \infty$ in the right most terms we obtain that $R(G)/c(H)$ is equal to $\lim_{m\to\infty} \epsilon_m(G,H)/m$. □

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

**Proposition 3.5.** *Let $G$ and $H$ be graphs and assume that both $G$ and $\overline{H}$ have at least one edge. For every positive integer $m$, we have*

$$\eta^\star_m(G,H) \leq \left\lceil \frac{\log \chi^\star(G^{\boxtimes m})}{\log \alpha^\star(H)} \right\rceil.$$

*Proof:* Set $n = \lceil \log \chi^\star(G^{\boxtimes m}) / \log \alpha^\star(H) \rceil$. Using the super-multiplicativity of $\alpha^\star(H)$ we get

$$\log \alpha^\star(H^{\boxtimes n}) \quad \geq \quad n \log \alpha^\star(H) = \left\lceil \frac{\log \chi^\star(G^{\boxtimes m})}{\log \alpha^\star(H)} \right\rceil \log \alpha^\star(H) \geq \log \chi^\star(G^{\boxtimes m}).$$

From Lemma 3.3 it then follows that $\eta^\star_m(G,H) \leq n$. □

## 4. Lower bound on the entangled chromatic number

In this section we prove Theorem 2.1. We will use the following simple fact about positive semidefinite matrices with a special block form (which can be found, e.g., in [18]).

**Lemma 4.1.** *Let $X$ be a $t \times t$ block matrix, with a matrix $A$ as diagonal blocks and a matrix $B$ as non-diagonal blocks, of the form*

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

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

*Proof of Theorem 2.1:* We show that relations (8) and (9) hold for the graph $\overline{G}$. First we observe that (9) follows easily from (8). Indeed, relation (8) combined with the identity (7) implies $\vartheta(\overline{G})^m = \vartheta(\overline{G^{\boxtimes m}}) \leq \chi^\star(G^{\boxtimes m})$ and thus $\log \vartheta(\overline{G}) \leq R^\star(G)$ follows after taking limits.

We now prove (8) for the graph $\overline{G}$, i.e., we show the inequality $\vartheta^+(\overline{G}) \leq \chi^\star(G)$. For this let $\rho, \{\rho_u^i : u \in V(G),\, i \in [t]\}$ be a set of positive semidefinite matrices which form a solution for $\chi^\star(G) = t$. We may assume that $\langle \rho, \rho \rangle = 1$. Here, $\langle \cdot, \cdot \rangle$ is the trace inner product, defined by $\langle A, B \rangle = \mathsf{Tr}(A^* B)$ for matrices $A, B$ of the same size. Define the matrix $X$, indexed by all pairs $\{u, i\} \in V(G) \times [t]$, with entries $X_{ui,vj} := \langle \rho_u^i, \rho_v^j \rangle$. By construction, $X$ is a non-negative positive semidefinite matrix which satisfies $X_{ui,vi} = 0$ for every $\{u, v\} \in E(G)$ and $i \in [t]$.

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

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

By construction, the matrix $Y$ is invariant under any permutation of $[t]$, i.e., $\sigma(Y) = Y$ for any $\sigma \in \mathrm{Sym}(t)$. Therefore, $Y$ has the block form of Lemma 4.1 with, moreover,

$$A_{uv} = 0 \ \text{ for all } \{u, v\} \in E(G). \tag{21}$$

As each matrix $\sigma(X)$ is positive semidefinite, the matrix $Y$ is positive semidefinite as well. From Lemma 4.1, 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 $\rho_u^i$ and the invariance of $Y$, we obtain the following relation for any $u, v \in V(G)$:

$$1 = \langle \rho, \rho \rangle = \langle \sum_{i \in [t]} \rho_u^i, \sum_{j \in [t]} \rho_v^j \rangle = \sum_{i \in [t]} \sum_{j \in [t]} \langle \rho_u^i, \rho_v^j \rangle = \sum_{i \in [t]} \sum_{j \in [t]} X_{ui,vj} = \sum_{i \in [t]} \sum_{j \in [t]} Y_{ui,vj},$$

implying

$$1 = \sum_{i \in [t]} \sum_{j \in [t]} Y_{ui,vj} = t \sum_{j \in [t]} Y_{ui,vj} = t(A_{uv} + (t-1)B_{uv}). \tag{22}$$

We are now ready to define a matrix $Z$ which is a feasible solution for the program (6) 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 (21) and (22) give $A_{uv} = 0$ and $t(t-1)B_{uv} = 1$ and thus $Z_{uv} = -1$. For a non-edge $\{u, v\}$, relation (22) combined with the fact that $A_{uv} \geq 0$ implies that $Z_{uv} \geq -1$. Finally, for any $u \in V(G)$, relation (22) 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 (6) defining $\vartheta^+(\overline{G})$. This shows that $\vartheta^+(\overline{G}) \leq \chi^\star(G)$, which concludes the proof. □

## 5. Separation between classical and entangled Witsenhausen rate

Here we prove Theorem 2.6, which shows an exponential separation between the classical and entangled-assisted Witsenhausen rate for the family of graphs $H_k$ in Definition 2.2.

5.1. **Upper bound on the entangled Witsenhausen rate.** Here we prove the upper bound (11) stated in Theorem 2.6 on $R^\star(H_k)$. 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.[3] The *orthogonal rank* $\xi(G)$ of $G$ is the minimum $d$ such that there exists a $d$-dimensional orthonormal representation of $G$. Following [10] we define $\xi'(G)$ to be the minimum dimension $d$ such that there exists a $d$-dimensional orthonormal representation $f$ of $G$ such that, for every vertex $u \in V(G)$, the $d$ entries of the vector $f(u)$ all have absolute value $1/\sqrt{d}$.

The following bound on $\chi^\star(G)$ follows from the fact that $\chi^\star(G) \leq \chi_q(G)$ and a result proved in [10] stating that $\chi_q(G) \leq \xi'(G)$. We give a self-contained proof of the implied bound on $\chi^\star(G)$ for completeness.

**Lemma 5.1.** *For every graph $G$, we have $\chi^\star(G) \leq \xi'(G)$.*

*Proof:* Set $d = \xi'(G)$, $\omega_d = e^{2i\pi/d}$ and, for every $i \in [d]$, let $h_i = [\omega_d^i, \omega_d^{i+1}, \dots, \omega_d^{i+d-1}]^{\mathsf{T}} \in \mathbb{C}^d$. It is not hard to see that $\{h_1, h_2, \dots, h_d\}$ is a complete orthogonal basis for $\mathbb{C}^d$. Set $\rho = I/d$. Then $\mathsf{Tr}(\rho) = 1$.

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

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

It follows that for every $u \in V(G)$ we have $\rho_u^1 + \rho_u^2 + \cdots \rho_u^d = I/d = \rho$. Moreover, for each $\{u, v\} \in E(G)$ and $i \in [d]$, we have $\rho_u^i \rho_v^i = 0$. As the matrices $\rho, \rho_u^i$ are also positive semidefinite, they satisfy all the requirements of Definition 1.3 and so $\chi^\star(G) \leq d$. □

[3] We stress that in our definition *orthogonality corresponds to adjacency*. Some authors prefer to demand orthogonality for non-adjacent vertices instead.

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 (11).

**Lemma 5.2.** *Let $k$ be an odd positive integer and $m \in \mathbb{N}$. Then,*

$$\chi^\star(H_k^{\boxtimes m}) \leq (k+1)^m.$$

*Moreover, if there exists a Hadamard matrix of size $k+1$, then equality holds.*

*Proof:* We first prove that $\chi^\star(H_k) \leq k+1$ by using Lemma 5.1. To this end we use the map $f$ defined in (10), 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 5.1 we get $\chi^\star(H_k) \leq k+1$. Using the sub-multiplicativity of $\chi^\star$ (Lemma 3.2) we get $\chi^\star(H_k^{\boxtimes m}) \leq (k+1)^m$.

We now prove that if there exists a Hadamard matrix of size $k+1$ then also the reverse inequality holds: $\chi^\star(H_k^{\boxtimes m}) \geq (k+1)^m$. Recall from Proposition 2.5 the existence of a Hadamard matrix of size $k+1$ implies $\omega(H_k) \geq k+1$. Combining this with Theorem 2.1 and the Sandwich Theorem gives that for every positive integer $m$, we have

$$\chi^\star(H_k^{\boxtimes m}) \geq \vartheta(\overline{H_k^{\boxtimes m}}) \geq \omega(H_k^{\boxtimes m}) \geq \omega(H_k)^m \geq (k+1)^m,$$

where the second-last 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 forms a clique in $G^{\boxtimes m}$. □

The bound (11) now follows as a simple corollary.

**Corollary 5.3.** *Let $k$ be a positive integer. Then $R^\star(H_k) \leq \log(k+1)$.*

*Proof:* By Lemma 5.2 we have $R^\star(H_k) = \inf_m \log \chi^\star(H_k^{\boxtimes m})/m \leq \log \chi^\star(H_k) \leq \log(k+1)$. □

We also record the following additional corollary, which we use later in Section 7.

**Corollary 5.4.** *For every odd integer $k$ such that there is a Hadamard matrix of size $k+1$, we have $\omega(H_k^{\boxtimes m}) = (k+1)^m$.*

*Proof:* Combining Proposition 2.5 and Lemma 5.2 gives the result. □

### 5.2. **Lower bound on the classical Witsenhausen rate.**

To prove the lower bound (12) on $R(H_k)$ stated in Theorem 2.6 we use the following upper bound on the classical independence number of the graphs $H_k^{\boxtimes m}$ for certain values of $k$.

**Lemma 5.5.** *Let $p$ be an odd prime number, $\ell \in \mathbb{N}$ and set $k = 4p^\ell - 1$. Then, for every $m \in \mathbb{N}$, we have*

$$\alpha(H_k^{\boxtimes m}) \leq \left(\binom{k}{0} + \binom{k}{1} + \cdots + \binom{k}{p^\ell - 1}\right)^m \leq 2^{k\,m\,H(3/11)} < 2^{0.846\,k\,m}, \tag{23}$$

*where $H(t) = -t\log t - (1-t)\log(1-t)$ is the binary entropy function.*

The proof of this lemma is an instance of the linear algebra method due to Alon [2] (see also Gopalan [17]), which we recall below for completeness. Let $G$ be a graph and $\mathbb{F}$ be a field. Let $\mathcal{F} \subseteq \mathbb{F}[x_1, \dots, x_k]$ be a subspace of the space of $k$-variate polynomials over $\mathbb{F}$. A *representation* of $G$ over $\mathcal{F}$ is an assignment $\big((f_u, c_u)\big)_{u \in V(G)} \subseteq \mathcal{F} \times \mathbb{F}^k$ of polynomial-point pairs to the vertices of $G$ such that

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

**Lemma 5.6** (Alon [2])**.** *Let $G$ be a graph, $\mathbb{F}$ be a field, $k \in \mathbb{N}$ and $\mathcal{F}$ be a subspace of $\mathbb{F}[x_1, \dots, x_k]$. If $\big((f_u, c_u)\big)_{u \in V} \subseteq \mathcal{F} \times \mathbb{F}^k$ represents $G$, then $\alpha(G^{\boxtimes n}) \leq \dim(\mathcal{F})^n$ for all $n \in \mathbb{N}$.*

*Proof:* Let $I \subseteq V(G)^n$ be an independent set in $G^{\boxtimes n}$. For each $\mathbf{u} = (u_1, \dots, u_n) \in I$, define the polynomial $f_{\mathbf{u}} \in \mathcal{F}^{\otimes n}$, which takes as input $n$-tuples of vectors $\mathbf{y} = (y_1, \dots, y_n) \in (\mathbb{F}^k)^n$ and assumes the value $f_{\mathbf{u}}(\mathbf{y}) = f_{u_1}(y_1) \cdots f_{u_n}(y_n)$. Moreover, for $\mathbf{u} \in \mathbf{I}$, define the $n$-tuple of vectors $c_{\mathbf{u}} = (c_{u_1}, \dots, c_{u_n}) \in (\mathbb{F}^k)^n$. Then, the pair $\big((f_{\mathbf{u}}, c_{\mathbf{u}})\big)_{\mathbf{u} \in V^n}$ represents $G^{\boxtimes n}$. From this one can easily verify that the polynomials $\{f_{\mathbf{u}} : \mathbf{u} \in I\} \subseteq \mathcal{F}^{\otimes n}$ are linearly independent, which implies $\alpha(G^{\boxtimes n}) \leq \dim(\mathcal{F}^{\otimes n}) = \dim(\mathcal{F})^n$. □

We will get a representation for the graph $H_k$, for $k = 4p^\ell - 1$, from the following result of Barrington, Beigel and Rudich [5]. The proof we give here closely follows Yekhanin's [39, 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 5.7** (Barrington, Beigel and Rudich [5])**.** *Let $p$ be a prime number and let $k$, $\ell$ and $w$ be integers such that $k > p^\ell$. There exists a multilinear polynomial $f \in \mathbb{Z}_p[x_1, \dots, x_k]$ of degree $\deg(f) \leq p^\ell - 1$ such that for every $c \in \{0,1\}^k$, we have*

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

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

**Theorem 5.8** (Lucas's Theorem)**.** *Let $p$ be a prime and $a, b \in \mathbb{N}$ with $p$-ary expansions $a = \sum_i a_i p^i$ and $b = \sum_i b_i p^i$, where $0 \leq a_i, b_i < p$. Then,*

$$\binom{a}{b} \equiv \prod_i \binom{a_i}{b_i} \bmod p.$$

*Proof of Lemma 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 $|c|_0, |c|_1, \dots, |c|_{\ell-1}$ of the $p$-ary expansion of $|c|$. The $k$-variate symmetric polynomial of degree $d$ is defined by

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

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

$$P_{p^i}(c) = \binom{|c|}{p^i} \equiv \binom{|c|_i}{1} \bmod p \equiv |c|_i \bmod p,$$

where the second identity follows from Lucas's theorem and the $p$-ary expansion of $p^i$, in which the coefficient of value 1 multiplying $p^i$ is the only nonzero coefficient. Now, define the polynomial $\hat{f} \in \mathbb{Z}_p[x_1, \dots, x_k]$ by

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

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

With this we can now prove Lemma 5.5.

*Proof of Lemma 5.5:* Let $c \in \{0,1\}^k$ be a string such that its Hamming weight $|c|$ is even and satisfies $|c| \equiv 0 \bmod p^\ell$. Then, since $p$ is odd and $k < 4p^\ell$, we have $|c| \in \{0, 2p^\ell\}$. Hence, if $|c| \notin \{0, 2p^\ell\}$, then $|c| \not\equiv 0 \bmod p^\ell$.

Recall from Definition 2.2 that $H_k$ can be defined as the graph whose vertices are the strings of $\{0,1\}^k$ with an even Hamming weight and where two distinct vertices $u, v$ are adjacent if their Hamming distance $|u \oplus v|$ is equal to $(k+1)/2 = 2p^\ell$. Here $u \oplus v$ is the sum modulo 2. For $u, v \in V(H_k)$, their Hamming distance $|u \oplus v|$ is an even number. Hence if $u \neq v$ are not adjacent in $H_k$, then $|u \oplus v| \notin \{0, 2p^\ell\}$ and thus $|u \oplus v| \not\equiv 0 \mod p^\ell$.

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

$$f(c) \equiv \begin{cases} 1 & \text{if } |c| \equiv 0 \bmod p^\ell \\ 0 & \text{otherwise,} \end{cases}$$

as is promised to exist by Lemma 5.7 (applied to $w = 0$).

We use $f$ to define a representation for $H_k$. To this end define for each $u \in \{0,1\}^k$ vertex in $V(H_k)$ the polynomial $f_u \in \mathbb{Z}_p[x_1, \dots, x_k]$ obtained by replacing in the polynomial $f$ the variable $x_i$ by $1 - x_i$ if $u_i = 1$ and leaving it unchanged otherwise. For example, if $u = (1, 1, 0, \dots, 0)$, then $f_u(x_1, \dots, x_k) = f(1 - x_1, 1 - x_2, x_3, \dots, x_k)$. Moreover, associate to the vertex $u$ the point $c_u = u$ seen as a 0/1 vector in $\mathbb{Z}_p^k$. We claim that $\big((f_u, c_u)\big)_{u \in V(H_k)}$

is a representation of $H_k$. To see this, observe that $f_u(c_v) = f(u \oplus v)$ for any $u, v \in V(H_k)$, so that $f_u(c_u) = f(0) = 1$, and $f_u(c_v) = 0$ if $u, v$ are distinct and non-adjacent.

Since the polynomials $f_u$ are multilinear and have degree at most $p^\ell - 1$, they span a space of dimension at most $\binom{k}{0} + \binom{k}{1} + \cdots + \binom{k}{p^\ell - 1}$, which is the number of multilinear monomials of degree at most $p^\ell - 1$. Applying Lemma 5.6 we obtain that

$$\alpha(H_k^{\boxtimes m}) \leq \left( \binom{k}{0} + \binom{k}{1} + \cdots + \binom{k}{p^\ell - 1} \right)^m . \tag{24}$$

We now use the well known fact that for $q, k \in \mathbb{N}$ with $1 < q < k/2$, $\binom{k}{0} + \ldots + \binom{k}{q-1} \leq 2^{kH(q/k)}$. From this, since $p^\ell/(4p^\ell - 1) \leq 3/11$, we deduce that the right hand side in (24) can be upper bounded by $2^{k\,m\,H(3/11)} < 2^{0.846\,k\,m}$. □

The bound (12) stated in Theorem 2.6 is a simple corollary of Lemma 5.5.

**Corollary 5.9.** *Let $p$ be an odd prime number and $\ell \in \mathbb{N}$. Then, for $k = 4p^\ell - 1$, we have $R(H_k) \geq 0.154k - 1$.*

*Proof:* By Lemma 5.5, for every integer $m$ we have

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

Taking the logarithm, dividing by $m$ and taking the limit $m \to \infty$ gives the result. □

## 6. Separation between classical and entangled Shannon capacity

Here we prove Theorem 2.8, thus showing the existence of an infinite family of graphs for which the entangled capacity exceeds the Shannon capacity.

6.1. **Lower bound on the entangled Shannon capacity.** The proof of the bound (13) on the entangled Shannon capacity is based on quantum teleportation (see Section 3.2). 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 (10) and define $\rho_x = f(x)f(x)^\mathsf{T}$ for $x \in V(H_k)$. To transmit a sequence $\mathbf{x} = (x_1, \ldots, x_t) \in V(H_k)^t$ Alice and Bob may follow the following four-step procedure. First, Alice prepares $(k+1)$-dimensional quantum registers $\mathcal{A}_1, \ldots, \mathcal{A}_t$ to be in the states $\rho_{x_1}, \ldots, \rho_{x_t}$, respectively. Second, Alice sends the sequence $\mathbf{x}$ through the channel by using it $t$ times in a row. This will result in $t$ channel-outputs on Bob's end of the channel from which he can infer that each $x_i$ belongs to a particular clique in $H_k$. Third, Alice and Bob execute a quantum teleportation scheme after which Bob ends up with quantum registers $\mathcal{Y}_1, \ldots, \mathcal{Y}_t$ 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 now give a formal proof of (13), which we repeat below for convenience.

**Lemma 6.1.** *For every odd integer $k \geq 5$, we have*

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

*Proof:* Set $V = V(H_k)$ and let $t \in \mathbb{N}$ such that $(k+1)^{2t} \leq \alpha(H_k)$. In what follows we construct trace-1 positive semidefinite matrix $\rho$ and, for every $\mathbf{x} \in V^t$, positive semidefinite matrices $\{\rho_{\mathbf{x}}^{\mathbf{u}} : \mathbf{u} \in V^{t+1}\}$ satisfying the conditions of Definition 1.4, i.e.,

$$\sum_{\mathbf{u}\in V^{t+1}} \rho_{\mathbf{x}}^{\mathbf{u}} \;=\; \rho, \tag{25}$$

$$\rho_{\mathbf{x}}^{\mathbf{u}}\rho_{\mathbf{y}}^{\mathbf{v}} \;=\; 0 \quad \forall \mathbf{x} \neq \mathbf{y},\, \mathbf{u} = \mathbf{v} \text{ or } \{\mathbf{u},\mathbf{v}\} \in E(H_k^{\boxtimes(t+1)}). \tag{26}$$

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

Let $f : V \to \mathbb{R}^{k+1}$ be the orthonormal representation of $H_k$ defined in (10). For $x \in V$ define $\rho_x = f(x)f(x)^\mathsf{T}$ and, for $\mathbf{x} = (x_1,\ldots,x_t) \in V^t$, define $\rho_{\mathbf{x}} = \rho_{x_1} \otimes \rho_{x_2} \otimes \cdots \otimes \rho_{x_t}$. Notice that $\mathsf{Tr}(\rho_{\mathbf{x}}) = 1$ and that $\rho_{\mathbf{x}}\rho_{\mathbf{y}} = 0$ for every $\{\mathbf{x},\mathbf{y}\} \in E(H_k^{\boxtimes t})$. We now consider the quantum teleportation scheme from Section 3.2, for the setting where Alice would want to transmit the state $\rho_{\mathbf{x}}$ of a $(k+1)^t$-dimensional quantum register $\mathcal{A}$ to Bob. According to (QT1), let $\sigma$ be the maximally entangled state defined over a pair of $(k+1)^t$-dimensional quantum registers $(\mathcal{X},\mathcal{Y})$, 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 register-pair $(\mathcal{A},\mathcal{X})$ provided by (QT2), and let $U_1,\ldots,U_T$ be Bob's unitary operators on $\mathcal{Y}$ given by (QT3). Define

$$\begin{aligned}
\rho &= \mathsf{Tr}_{\mathcal{X}}(\sigma),\\
\rho_{\mathbf{x}}^i &= \mathsf{Tr}_{(\mathcal{A},\mathcal{X})}\big((M_i \otimes I)(\rho_{\mathbf{x}} \otimes \sigma)\big) \;\; \forall x \in V^t, i \in [T].
\end{aligned}$$

Since the $M_i$'s sum to the identity, for every $\mathbf{x}$, we have

$$\sum_{i=1}^{T} \rho_{\mathbf{x}}^i = \mathsf{Tr}_{(\mathcal{A},\mathcal{X})}(\rho_{\mathbf{x}} \otimes \sigma) = \mathsf{Tr}_{\mathcal{A}}(\rho_{\mathbf{x}})\ \mathsf{Tr}_{\mathcal{X}}(\sigma) = \rho. \tag{27}$$

By (QT4), we know that that the identity $U_i\rho_{\mathbf{x}}^iU_i^* = \beta_{\mathbf{x}}^i\rho_{\mathbf{x}}$ holds, where $\beta_{\mathbf{x}}^i = \mathsf{Tr}(\rho_{\mathbf{x}}^i)$. Hence, since $f$ is an orthonormal representation, for every edge $\{\mathbf{x},\mathbf{y}\} \in E(H_k^{\boxtimes t})$, we have

$$\rho_{\mathbf{x}}^i\rho_{\mathbf{y}}^i = (U_i^*\beta_{\mathbf{x}}^i\rho_{\mathbf{x}}U_i)\,(U_i^*\beta_{\mathbf{y}}^i\rho_{\mathbf{y}}U_i) = \beta_{\mathbf{x}}^i\beta_{\mathbf{y}}^iU_i^*\rho_{\mathbf{x}}\rho_{\mathbf{y}}U_i = 0. \tag{28}$$

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

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

Then, $\sum_{\mathbf{u}\in V^{t+1}} \rho_{\mathbf{x}}^{\mathbf{u}} = \sum_{u_{t+1}\in W} \rho_{\mathbf{x}}^{\phi(u_{t+1})} = \sum_{i=1}^{T} \rho_{\mathbf{x}}^{i} = \rho$ by (27). Next, let $\mathbf{x} \neq \mathbf{y} \in V^t$ and $\{\mathbf{u}, \mathbf{v}\}$ be equal or adjacent in $H_k^{\boxtimes(t+1)}$; we show that $\rho_{\mathbf{x}}^{\mathbf{u}}\rho_{\mathbf{y}}^{\mathbf{v}} = 0$. This is clear if $\mathbf{x} \neq (u_1, \dots, u_t)$, or $\mathbf{y} \neq (v_1, \dots, v_t)$, or $\{u_{t+1}, v_{t+1}\} \not\subseteq W$. So we may assume $\mathbf{u} = (\mathbf{x}, u_{t+1})$, $\mathbf{v} = (\mathbf{y}, v_{t+1})$ and $\{u_{t+1}, v_{t+1}\} \subseteq W$ and thus $\{\mathbf{u}, \mathbf{v}\} \in E(H_k^{\boxtimes(t+1)})$, $\{\mathbf{x}, \mathbf{y}\} \in E(H_k^{\boxtimes t})$ and $u_{t+1} = v_{t+1}$. Then we have that $\rho_{\mathbf{x}}^{\mathbf{u}}\rho_{\mathbf{y}}^{\mathbf{v}} = \rho_{\mathbf{x}}^{\phi(u_{t+1})}\rho_{\mathbf{y}}^{\phi(u_{t+1})} = 0$ by (28).

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

$$c^\star(H_k) \geq \frac{1}{t+1} \log \alpha^\star(H_k^{\boxtimes(t+1)}) \geq \frac{1}{t+1} t(k-1). \tag{29}$$

By Lemma 2.3 we have $\alpha(H_k) \geq 2^{(k-3)/2}$. Hence, for $k \geq 5$ we can choose the integer $t$ to be equal to $t = \lfloor (k-3)/4\log(k+1) \rfloor$. From (29) we then get

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

which gives the claimed result. □

6.2. **Upper bound on the Shannon capacity.** The upper bound (14) on the Shannon capacity of $H_k$ (for certain values of $k$) stated in Theorem 2.8 is an easy corollary of Lemma 5.5.

**Corollary 6.2.** *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 $m \to \infty$ on both sides of (23) we get the result. □

## 7. Separation between classical and entangled source-channel cost rate

7.1. **Proof of Theorem 2.9.** Now we prove Theorem 2.9, separately showing the two bounds (15) for $\eta^\star$ and (16) for $\eta$. The bound (15) is obtained by combining (11), (13) with Proposition 3.4. The proof of (16) relies on the No-Homomorphism Lemma due to Alberson and Collins [1].

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 $v = \pi(u)$. The next lemma follows easily from the fact that, if $G$ is vertex-transitive, then $|V(G)|/\alpha(G)$ is equal to its fractional chromatic number.

**Lemma 7.1** (No-Homomorphism Lemma, Albertson and Collins [1]). *Let $H$ be a vertex-transitive graph. If there is a homomorphism from $G$ to $H$, then*

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

As observed in [9], 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 taking the strong product and complement of graphs preserves vertex-transitivity. Hence, $\overline{H_k^{\boxtimes n}}$ is vertex-transitive for any $n \in \mathbb{N}$.

We prove the bound (16) which we repeat for convenience.

**Lemma 7.2.** *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:* Recall the definition of $\eta(H_k, H_k)$ from (3). Consider integers $m, n \in \mathbb{N}$ for which $H_k^{\boxtimes m} \longrightarrow \overline{H_k^{\boxtimes n}}$. Applying Lemma 7.1, we deduce that

$$\frac{|V(H_k^{\boxtimes m})|}{\alpha(H_k^{\boxtimes m})} \leq \frac{|V(\overline{H_k^{\boxtimes n}})|}{\alpha(\overline{H_k^{\boxtimes n}})} = \frac{|V(\overline{H_k^{\boxtimes n}})|}{\omega(H_k^{\boxtimes n})}. \tag{30}$$

From Corollary 5.4 we have $\omega(H_k^{\boxtimes n}) = (k+1)^n$. As $|V(H_k)| = 2^{k-1}$ and applying Lemma 5.5, we get

$$\frac{2^{(k-1)\,m}}{2^{k\,m\,0.846}} \overset{\text{Lemma 5.5}}{<} \frac{|V(H_k^{\boxtimes m})|}{\alpha(H_k^{\boxtimes m})} \overset{(30)}{\leq} \frac{|V(\overline{H_k^{\boxtimes n}})|}{\omega(H_k^{\boxtimes n})} = \frac{2^{(k-1)\,n}}{(k+1)^n}.$$

After a few elementary algebraic manipulations and taking logarithms the above inequality implies

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

This shows the lower bound from Lemma 7.2. □

7.2. **Stronger bounds based on Hadamard matrices.** The reader may have noticed that for the purpose of proving Theorem 2.9, we may assume that the integer $k$ appearing in the statement is such that there exists a Hadamard matrix of size $k+1$. The reason for this is that the bound (16) is conditional on the existence of such a matrix. With this additional assumption a stronger upper bound on $\eta^\star(H_k, H_k)$ can be proved without the use of quantum teleportation.

To prove this, we bound $\eta_1^\star(H_k, H_k)$ by the rate achievable with separate entangled coding schemes for the source-coding and channel-coding problem, respectively (see Section 3.4.2). To do so, we need a lower bound on the entangled independence number that was obtained previously in [9].

**Lemma 7.3** ([9])**.** *Let $k$ be a positive integer such that there exists a Hadamard matrix of size $k+1$. Then,*

$$\log \alpha^\star(H_k) \geq k - 1 - 2\log(k+1).$$

**Lemma 7.4.** *Let $k$ be a positive integer such that there exists a Hadamard matrix of size $k+1$. Then,*

$$\eta_1^\star(H_k,H_k) \le \left\lceil \frac{\log(k+1)}{k-1-2\log(k+1)} \right\rceil .$$

*Proof:* Putting together Proposition 3.5, Lemma 5.2 and Lemma 7.3 we have that, for every $k$ such that there exists a Hadamard matrix of size $(k+1)$,

$$\eta_1^\star(H_k,H_k) \le \left\lceil \frac{\log \chi^\star(H_k)}{\log \alpha^\star(H_k)} \right\rceil \le \left\lceil \frac{\log(k+1)}{k-1-2\log(k+1)} \right\rceil$$

which proves the claim. □

Since we have $\eta^\star(H_k,H_k) \le \eta_1^\star(H_k,H_k)$, the above result also implies an upper bound of the cost rate attainable by encoding infinitely long sequences of source inputs into single codewords.

## 8. Concluding remarks and open problems

We have shown a separation between classical and entangled-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 showed 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 [26] is that, for the classical source-channel coding problem, there exist situations for which separate encoding is highly suboptimal. Does this happen also in the entanglement-assisted case? This question has a positive answer if there exists a graph $G$ with $R^\star(G) > c^\star(\overline{G})$. In [26] 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 entangled-assisted encoding to be optimal. Are there weaker conditions that hold for the entangled case?

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

In Section 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 [29] it is conjectured that $\alpha^\star(G) = \alpha_q(G)$ for every graph $G$. A possible approach to show that $\chi^\star$ and $\chi_q$ are two separate quantities is to prove that the relationship between Kochen-Specker sets and $\chi_q$ found in [24] does not hold for $\chi^\star$. Finally, we mention that the existence of a graph $G$ for which $\chi^\star(G) < \chi_q(G)$ or $\alpha_q(G) < \alpha^\star(G)$ would prove the existence of a non-local game such that every quantum strategy that wins with probability one does not use a maximally entangled state.

## Acknowledgments

The authors thank Ronald de Wolf for useful discussions and helpful comments on a preliminary version of the paper. J.B. thanks Dion Gijswijt for helpful discussions.

## References

[1] M. O. Albertson and K. L. Collins. Homomorphisms of 3-chromatic graphs. *Discrete Mathematics*, 54(2):127-132, 1985.

[2] N. Alon. The Shannon capacity of a union. *Combinatorica*, 18(3):301–310, 1998.

[3] N. Alon. Graph powers. *Contemporary Combinatorics*, 11-28, 2002.

[4] 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.

[5] D. Barrington, R. Beigel, and S. Rudich. Representing Boolean functions as polynomials modulo composite numbers. *Computational Complexity*, 4(4):367–382, 1994.

[6] S. Beigi. Entanglement-assisted zero-error capacity is upper-bounded by the Lovász $\vartheta$ function. *Physical Review A*, 82(1):010303, 2010.

[7] J. S. Bell. On the Einstein-Podolsky-Rosen Paradox, *Physics*, 1:195-200, 1964.

[8] 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.

[9] J. Briët, H. Buhrman, and D. Gijswijt. Violating the Shannon capacity of metric graphs with entanglement. *Proceedings of the National Academy of Sciences*, 2012.

[10] P. J. Cameron, A. Montanaro, M. W. Newman, S. Severini, and A. Winter. On the quantum chromatic number of a graph. *Electronic Journal of Combinatorics*, 14(R81):1, 2007.

[11] Y. Q. 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.

[12] T. S. Cubitt, D. Leung, W. Matthews, and A. Winter. Improving zero-error classical communication with entanglement. *Physical Review Letters*, 104(23):230503, 2010.

[13] 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.

[14] A. Einstein, P. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Physical Review*, 47:777-780, 1935.

[15] M. J. Ferguson and D. W. Bailey. Zero-error coding for correlated sources. Unpublished manuscript, 1975.

[16] P. Frankl and R. M. Wilson. Intersection theorems with geometric consequences. *Combinatorica*, 1(4):357-368, 1981.

[17] P. Gopalan. Constructing Ramsey graphs from boolean function representations. *Proceedings of the 21st Annual IEEE Conference on Computational Complexity (CCC 2006)*, pp. 115-128, 2006.

[18] N. Gvozdenović and M. Laurent. The Operator $\Psi$ for the chromatic number of a graph. *SIAM Journal on Optimization*, 19(2):572-591, 2008.

[19] D. E. Knuth. The sandwich theorem. *Electronic Journal of Combinatorics*, 1:1, 1993.

[20] J. Körner and A. Orlitsky. Zero-error information theory. *IEEE Transactions on Information Theory*, 44(6):2207–2229, 1998.

[21] D. Leung, L. Mancinska, 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.

[22] L. Lovász. On the Shannon capacity of a graph. *IEEE Transactions on Information Theory*, 25(1):1–7, 1979.

[23] E. Lubetzky. *Graph Powers and Related Extremal Problems.* Ph.D. thesis, Tel Aviv University, 2007.

[24] L. Mančinska, G. Scarpa, and S. Severini. New separations in zero-error channel capacity through projective Kochen-Specker sets and quantum coloring. *IEEE Transactions on Information Theory*, 59(6):4025-4032, 2013.

[25] P. Meurdesoif. Strengthening the Lovász $\theta(\overline{G})$ bound for graph coloring. *Mathematical Programming*, 102(3):577-588, 2005.

[26] J. Nayak, E. Tuncel, and K. Rose. Zero-error source-channel coding with side information. *IEEE Transactions on Information Theory*, 52(10):4626–4629, 2006.

[27] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information.* 2000.

[28] R. E. A. C. Paley. On orthogonal matrices. *Journal of Mathematics and Physics* (now called *Studies in Applied Mathematics)*, 12:311-320, 1933.

[29] D. E. Roberson and L. Mančinska. Graph homomorphisms for quantum players. 2012. Available at arXiv:1212.1724 [quant-ph].

[30] C. Shannon. The zero error capacity of a noisy channel. *IRE Transactions on Information Theory*, 2(3):8–19, 1956.

[31] M. Szegedy. A note on the $\theta$ number of Lovász and the generalized Delsarte bound. *Proceedings of the 45th Annual IEEE Annual Symposium on Foundations of Computer Science (FOCS 2004)*, pp. 36–39, 1994.

[32] S. Vembu, S. Verdu, and Y. Steinberg. The source-channel separation theorem revisited. *IEEE Transactions on Information Theory*, 41(1):44-54, 1995.

[33] R. M. Wilson and Q. Xiang. Constructions of Hadamard difference sets. *Journal of Combinatorial Theory, Series A*, 77(1):148-160, 1997.

[34] H. S. Witsenhausen. The zero-error side information problem and chromatic numbers. *IEEE Transactions on Information Theory*, 22(5):592–593, 1976.

[35] 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.

[36] M. Xia and G. Liu. A new family of supplementary difference sets and Hadamard matrices. *Journal of Statistical Planning and Inference*, 51(3):283-291, 1996.

[37] T. Xia, J. Seberry, and M. Xia. New constructing of regular Hadamard matrices. *Proceedings of the 10th WSEAS international conference on Computers*, pp. 1294-1299, 2006.

[38] Q. Xiang. Difference families from lines and half lines. *European Journal of Combinatorics*, 19(3):395-400, 1998.

[39] S. Yekhanin. Locally decodable codes. *Foundations and Trends in Theoretical Computer Science*, 6(3):139-255, 2012.

Courant Institute, New York University, 251 Mercer Street, New York NY 10012, USA

*E-mail address*: jop.briet@cims.nyu.edu

Centrum Wiskunde & Informatica (CWI), Science Park 123, 1098 XG Amsterdam, The Netherlands

*E-mail address*: buhrman@cwi.nl

Centrum Wiskunde & Informatica (CWI), Science Park 123, 1098 XG Amsterdam, and Tilburg University, PO Box 90153, 5000 LE Tilburg, The Netherlands

*E-mail address*: m.laurent@cwi.nl

Centrum Wiskunde & Informatica (CWI), Science Park 123, 1098 XG Amsterdam, The Netherlands

*E-mail address*: t.piovesan@cwi.nl

Centrum Wiskunde & Informatica (CWI), Science Park 123, 1098 XG Amsterdam, The Netherlands

*E-mail address*: g.scarpa@cwi.nl