



## UvA-DARE (Digital Academic Repository)

### Entanglement-assisted zero-error source-channel coding

Briët, J.; Buhrman, H.; Laurent, M.; Piovesan, T.; Scarpa, G.

**DOI**

[10.1109/TIT.2014.2385080](https://doi.org/10.1109/TIT.2014.2385080)

**Publication date**

2015

**Document Version**

Submitted manuscript

**Published in**

IEEE Transactions on Information Theory

[Link to publication](#)

**Citation for published version (APA):**

Briët, J., Buhrman, H., Laurent, M., Piovesan, T., & Scarpa, G. (2015). Entanglement-assisted zero-error source-channel coding. *IEEE Transactions on Information Theory*, 61(2), 1124-1138. <https://doi.org/10.1109/TIT.2014.2385080>

**General rights**

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

**Disclaimer/Complaints regulations**

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

# ENTANGLEMENT-ASSISTED ZERO-ERROR SOURCE-CHANNEL CODING

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 of Szegegy'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 remote state preparation.

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

## 1. INTRODUCTION

We study a problem from classical zero-error information theory: the *zero-error source-channel coding problem*, in the setting where the sender and receiver may use quantum entanglement. Viewed separately, the source coding problem asks a sender, Alice, to efficiently communicate data about which a receiver, Bob, already has some information<sup>1</sup>, 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

---

J. B. was partially supported by a Rubicon grant from the Netherlands Organization for Scientific Research (NWO). J. B. and H. B. were supported by the European Commission under the project QCS (Grant No. 255961). H. B. and T. P. were supported by the EU grant SIQS. G. S. was supported by Ronald de Wolf's Vidi grant 639.072.803 from the Netherlands Organization for Scientific Research (NWO).

<sup>1</sup>This setting is usually called the *dual* source coding problem, referring to the fact that it involves two parties. Since we only consider the two-party setting we omit the word *dual*.

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 [1] 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 [2] for an extensive survey and Lubetzky's PhD thesis [3] for more recent results). The branch of this line of research involving entanglement was started only recently by Cubitt et al. [4]. 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 [5], Bell [6] 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. Below we set the stage in detail and state our results precisely.

First, we recall some preliminary definitions needed for this paper. Throughout, all graphs are assumed to be finite, undirected and without self-loops. For a graph  $G = (V, E)$ ,  $V$  and  $E$  (or equivalently  $V(G)$  and  $E(G)$ ) denote 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 it holds that  $u_1 = v_1$  or  $\{u_1, v_1\} \in E(G)$  and that  $u_2 = v_2$  or  $\{u_2, v_2\} \in E(H)$ . 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  $\phi : V(G) \rightarrow V(H)$  such that every edge  $\{u, v\}$  in  $G$  is mapped to an edge  $\{\phi(u), \phi(v)\}$  in  $H$ . If such a map exists, we write  $G \rightarrow H$ . The complete graph on  $t$  vertices, denoted by  $K_t$ , is the graph where every pair of distinct vertices is adjacent. A *d-dimensional orthogonal representation* of a graph  $G$  is a map  $f$  from  $V(G)$

to non-zero vectors in  $\mathbb{C}^d$  such that adjacent vertices are mapped to orthogonal vectors.<sup>2</sup> The *orthogonal rank*  $\xi(G)$  of  $G$  is the minimum  $d$  such that there exists a  $d$ -dimensional orthogonal representation of  $G$ . Following [7] we define  $\xi'(G)$  to be the minimum  $d$  for which there exists a  $d$ -dimensional orthogonal representation  $f$  of  $G$  such that for each  $u \in V(G)$  the entries of the vector  $f(u)$  all have absolute value one. To indicate that a Hermitian matrix  $A$  is positive semidefinite we write  $A \succeq 0$ . Recall that  $A \succeq 0$  if and only if there exist vectors  $a_i$  such that  $A_{ij} = a_i^* a_j$ . The trace inner product of matrices  $A, B$  of equal size is defined by  $\langle A, B \rangle = \text{Tr}(A^* B)$ . Finally, 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 *source*  $\mathcal{M} = (\mathbf{X}, \mathbf{U}, P)$  consists of a finite set  $\mathbf{X}$ , a (possibly infinite) set  $\mathbf{U}$  and a probability distribution  $P$  over  $\mathbf{X} \times \mathbf{U}$ . In a source instance, Alice is given an input  $x \in \mathbf{X}$  and Bob an input  $u \in \mathbf{U}$  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.<sup>3</sup> 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 applies encoding function  $\mathbf{C} : \mathbf{X}^m \rightarrow \{0, 1\}^n$  and sends  $\mathbf{C}(\mathbf{x})$  through the binary channel by using it  $n$  times in a row. Bob, who received an input  $\mathbf{u} = (u_1, \dots, u_m) \in \mathbf{U}^m$ , then applies a decoding function  $\mathbf{D} : \mathbf{U}^m \times \{0, 1\}^n \rightarrow \mathbf{X}^m$  to the pair  $(\mathbf{u}, \mathbf{C}(\mathbf{x}))$  to get a string in  $\mathbf{X}^m$ . The scheme works if Bob always gets the string  $\mathbf{x}$ . The *cost rate* of the scheme  $(\mathbf{C}, \mathbf{D})$  is then  $n/m$ , which counts the average number of channel uses per source-input symbol.

Witsenhausen [8] and Ferguson and Bailey [9] showed that the zero-error source coding problem can be studied in graph-theoretic terms. Associated with a source  $\mathcal{M} = (\mathbf{X}, \mathbf{U}, P)$  is its *characteristic graph*  $G = (\mathbf{X}, E)$ , where  $\{x, y\} \in E$  if there exists a  $u \in \mathbf{U}$  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 cannot distinguish based on his input. It is not difficult to see that every graph is the characteristic graph of a (non-unique) source. Solving one instance of the zero-error source coding problem for  $\mathcal{M}$  is equivalent to finding a proper coloring of  $G$ . Indeed, Bob's input  $u$  reduces the list of Alice's possible inputs to the set  $\{x \in \mathbf{X} : P(x, u) > 0\}$  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 take the graph  $G^{\boxtimes m}$  (the strong product of  $m$  copies of  $G$ ), whose edges are the pairs of input sequences for Alice which Bob cannot distinguish. The *Witsenhausen rate*

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

<sup>2</sup>We stress that in our definition *orthogonality corresponds to adjacency*. Some authors prefer to demand orthogonality for non-adjacent vertices instead.

<sup>3</sup>From now on we will assume that all binary channels are noiseless.

is the minimum asymptotic cost rate of a zero-error code for a source (all logarithms in this paper are in base 2). 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<sup>4</sup> the above limit exists and is equal to the infimum:  $R(G) = \inf_m \log \chi(G^{\boxtimes m})/m$ .

A (noisy) *discrete channel*  $\mathcal{N} = (\mathbf{S}, \mathbf{V}, Q)$  consists of a finite input set  $\mathbf{S}$ , a (possibly infinite) output set  $\mathbf{V}$  and a probability distribution  $Q(\cdot|s)$  over  $\mathbf{V}$  for each  $s \in \mathbf{S}$ . Throughout the paper we consider only memoryless channels, where the probability distribution of the output depends only on the current channel input. If Alice sends an input  $s \in \mathbf{S}$  through the channel, then Bob receives the output  $v \in \mathbf{V}$  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  $\mathbf{C} : \{0, 1\}^m \rightarrow \mathbf{S}^n$  and sends  $\mathbf{C}(\mathbf{y})$  through the channel by using it  $n$  times in sequence. Bob then receives an output sequence  $\mathbf{v} = (v_1, \dots, v_n)$  on his side of the channel and applies a decoding function  $\mathbf{D} : \mathbf{V}^n \rightarrow \{0, 1\}^m$ . The coding scheme  $(\mathbf{C}, \mathbf{D})$  works if  $\mathbf{D}(\mathbf{v}) = \mathbf{y}$ . The *communication rate* of the scheme is  $m/n$ , the number of bits transmitted per channel use.

Shannon [1] showed that the zero-error channel coding problem can also be studied in graph-theoretic terms. Associated to a channel  $\mathcal{N} = (\mathbf{S}, \mathbf{V}, Q)$  is its *confusability graph*  $H = (\mathbf{S}, F)$  where  $\{s, t\} \in F$  if there exists a  $v \in \mathbf{V}$  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*

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

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

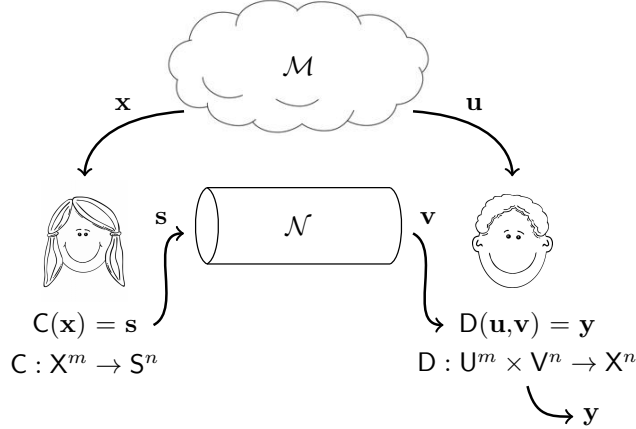
Nayak, Tuncel and Rose [10] showed that if  $\mathcal{M}$  has characteristic graph  $G$  and  $\mathcal{N}$  has confusability graph  $H$ , then a zero-error  $(m, n)$ -coding scheme is equivalent to a homomorphism from  $G^{\boxtimes m}$  to  $\overline{H^{\boxtimes n}}$ . Then, the parameter

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

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

---

<sup>4</sup>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:  $\lim_{m \rightarrow \infty} a_m/m = \inf_{m \in \mathbb{N}} a_m/m$ .



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

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  $\eta(G, H) = \lim_{m \rightarrow \infty} \eta_m(G, H)/m$  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,*

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

*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

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

Combining these inequalities applied to  $G' = G^{\boxtimes m}$  with the identity  $\overline{K_2^{\boxtimes n}} = K_{2^n}$ , we obtain

$$\eta(G, \overline{K_2}) = \lim_{m \rightarrow \infty} \frac{1}{m} \min \{ n : G^{\boxtimes m} \longrightarrow \overline{K_2^{\boxtimes n}} = K_{2^n} \} = \lim_{m \rightarrow \infty} \frac{1}{m} \log \chi(G^{\boxtimes m}) = R(G).$$

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

we get

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

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

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

and thus

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

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

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

Taking the limit of the above, we obtain the inequality  $1/c(H) \leq \lim_{m \rightarrow \infty} n(m)/m = \eta_m(K_2, H)$ .

Next, as  $\eta_m(K_2, H) = \inf_n n(m)/m$ , using the right most inequality in (4) we deduce that

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

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

**Separation theorems.** 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 [11]. But when errors cannot be tolerated, Nayak, Tuncel and Rose [10] showed that separated codes can be highly suboptimal. In terms of the above graph parameters, this says that in general  $\eta(G, H) \leq R(G)/c(H)$  holds (this is implied in [10]; see Proposition 3.2 for an explicit proof), 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.** Here we describe the information-theoretic problems introduced above, but now in the setting where the parties are allowed to use entanglement. We refer readers who are unfamiliar with quantum information theory to Section 3. We will make repeated use of the following simple lemma (already implicitly used in a similar context in [4]), which we will refer to here as the *Orthogonality Lemma*.

**Lemma 1.2** (Orthogonality Lemma). *Let  $\rho_1, \dots, \rho_\ell \in \mathbb{C}^{d \times d}$  be a collection of Hermitian positive semidefinite matrices. Then the following are equivalent:*

- (1) *We have  $\rho_i \rho_j = 0$  for every  $i \neq j \in [\ell]$ .*
- (2) *There exists a measurement consisting of positive semidefinite matrices  $P^1, \dots, P^\ell, P^\perp \in \mathbb{C}^{d \times d}$  such that  $\text{Tr}(P^i \rho_j) = \delta_{ij} \text{Tr}(\rho_j)$  and  $\text{Tr}(P^\perp \rho_j) = 0$  for every  $i, j \in [\ell]$ .*

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

(2)  $\Rightarrow$  (1): Let  $V_i$  be the column space of  $\rho_i$  and let  $W_i$  be the column space of  $P^i$ . By Hermiticity, these spaces are the complex conjugates of the respective row spaces. We claim that  $V_i \subseteq W_i$ . Indeed, expanding  $\rho_i$  in its spectral decomposition, writing  $\lambda[\ell] \in \mathbb{R}_+$  and  $a[\ell] \in \mathbb{C}^d$ ,  $\ell \in [d]$ , for its eigenvalues and eigenvectors, respectively,

$$(5) \quad \sum_{\ell=1}^d \lambda[\ell] = \text{Tr}(\rho_i) = \text{Tr}(P^i \rho_i) = \sum_{\ell=1}^d \lambda[\ell] \text{Tr}(P^i a[\ell] a[\ell]^*) = \sum_{\ell=1}^d \lambda[\ell] a[\ell]^* P^i a[\ell].$$

Since  $P^i$  has eigenvalues in  $[0, 1]$ , we have that (5) holds if and only if  $a[\ell]^* P^i a[\ell] = 1$  for each  $\ell \in [d]$ . This implies that each eigenvector  $a[\ell]$  of  $\rho_i$  is an eigenvector of  $P^i$  with eigenvalue 1. The condition  $\text{Tr}(P^i \rho_j) = 0$  if  $i \neq j$  implies that  $W_i$  is orthogonal to  $V_j$ . Hence, for  $i \neq j$ , we have  $V_i \subseteq W_i$  and  $W_i$  is orthogonal to  $V_j$  and so the row space  $V_i$  of  $\rho_i$  is orthogonal to the column space  $V_j$  of  $\rho_j$ , giving  $\rho_i \rho_j = 0$ .  $\square$

**Entanglement-assisted source coding.** We continue by describing the special case of entanglement-assisted source coding. The setup is as before, except now Alice and Bob have quantum registers  $\mathcal{A}$  and  $\mathcal{B}$ , respectively, that are initialized to be in some entangled state. If the source gives the parties inputs  $x \in \mathbf{X}$  and  $u \in \mathbf{U}$ , respectively, then their most general course of action is as follows:

- (1) After receiving her input  $x$  Alice performs a measurement on her register  $\mathcal{A}$  and communicates the measurement outcome to Bob;
- (2) After receiving both his input  $u$  and Alice's measurement outcome, Bob performs a measurement on his register  $\mathcal{B}$  and obtains a measurement outcome  $y \in \mathbf{X}$ .

If this protocol is successful, that is, Bob gets outcome  $y = x$  for each possible input pair  $(x, u)$ , then it yields a system of matrices which we use below to derive the definition of the *entangled chromatic number* (Definition 1.5). We get these matrices as follows.

Let  $\sigma$  denote the state in which the pair of registers  $(\mathcal{A}, \mathcal{B})$  is initialized before the protocol starts. The measurement Alice performs in step (1) is given by a collection of positive semidefinite matrices  $A_x^1, \dots, A_x^t$  that add up to the identity. If Alice gets outcome  $i \in [t]$ , then after step (1) Bob's register is left in a state proportional to  $\rho_x^i = \text{Tr}_{\mathcal{A}}((A_x^i \otimes I)\sigma)$ . Note that for each  $x \in \mathbf{X}$  the matrices  $(\rho_x^i)_{i \in [t]}$  sum to Bob's reduced density matrix  $\rho := \text{Tr}_{\mathcal{A}}(\sigma)$ . Observe that Bob's input  $u$  already allows him to reduce the list of Alice's possible inputs to a clique  $\mathcal{C}$  in the characteristic graph  $G$  that contains Alice's input  $x$ . After Alice does her measurement and sends the outcome  $i$ , he thus knows that his register is in one of the states proportional to  $\rho_y^i$  for  $y \in \mathcal{C}$ . The measurement he does in step (2) allows him to learn exactly which of these states his register is in. More explicitly, Bob has a measurement of positive semidefinite matrices  $(B^x)_{x \in \mathcal{C}}$  and  $B^\perp$  such that  $\text{Tr}(B^x \rho_y^i) = \delta_{xy} \text{Tr}(\rho_y^i)$  and  $\text{Tr}(B^\perp \rho_x^i) = 0$



for every  $x, y \in \mathcal{C}$ . Hence, by the Orthogonality Lemma (Lemma 1.2), the states  $\rho_x^i / \text{Tr}(\rho_x^i)$  must be pairwise orthogonal and we have

$$(6) \quad \rho_x^i \rho_y^i = 0 \quad \forall i \in [t] \text{ and every } \{x, y\} \in E(G)$$

$$(7) \quad \sum_{i \in [t]} \rho_x^i = \rho \quad \forall x \in V(G).$$

In the opposite direction we now show that such a system of matrices can be used to construct a protocol for the source-coding problem. Cubitt et al. [4] implicitly used a similar fact in the context of channel coding, albeit without proof.

**Proposition 1.3.** *Suppose that there exists a collection of positive semidefinite matrices  $\rho$  and  $\{\rho_x^i : x \in V(G), i \in [t]\}$  in  $\mathbb{C}^{d \times d}$  such that  $\text{Tr}(\rho) = 1$  and (6) and (7) hold. Then, there exists a  $t$ -message protocol for the source-coding problem.*

The proposition follows almost directly from the following well-known theorem (see e.g., [12, HJW Theorem, pp. 74], where it is attributed to Hughston, Jozsa and Wootters).

**Theorem 1.4** (Hughston–Jozsa–Wootters Theorem). *Let  $d, t$  be positive integers,  $p_1, \dots, p_t \geq 0$  satisfying  $p_1 + \dots + p_t = 1$ , and let  $\rho_1, \dots, \rho_t \in \mathbb{C}^{d \times d}$  be positive semidefinite matrices with trace 1. Then, there exists a state  $\sigma$  for a pair of registers  $(\mathcal{A}, \mathcal{B})$ , and a measurement on  $\mathcal{A}$  consisting of a collection of positive semidefinite matrices  $A^1, \dots, A^t$  that add up to the identity, such that for each  $i \in [t]$ , we have  $\text{Tr}_{\mathcal{A}}((A^i \otimes I)\sigma) = p_i \rho_i$ . Moreover,  $\sigma$  depends only on  $p_1 \rho_1 + \dots + p_t \rho_t$ .*

*Proof of Proposition 1.3.* It suffices to find an entangled state  $\sigma$  and measurement  $\{A_x^i : i \in [t]\}$  for each of Alice's inputs  $x$  such that  $\rho_x^i = \text{Tr}_{\mathcal{A}}((A_x^i \otimes I)\sigma)$ . Indeed, recall that after getting his input  $u$  and Alice's measurement outcome  $i$ , Bob knows that his register is in the state  $\rho_y^i / \text{Tr}(\rho_y^i)$  for some  $y \in \mathcal{C}$  and clique  $\mathcal{C}$  in  $G$  that contains Alice's input  $x$ . By (6) and the Orthogonality Lemma there exists a measurement  $\{B_y : y \in \mathcal{C}\} \cup \{B^\perp\}$  such that  $\text{Tr}(B_y \rho_x^i) = \delta_{xy} \text{Tr}(\rho_x^i)$ , which thus allows Bob to correctly identify Alice's input  $x$ .

For each  $x \in V(G)$ , define the nonnegative numbers  $p_x^i = \text{Tr}(\rho_x^i)$ ,  $i \in [t]$ , and trace-1 matrices  $\tilde{\rho}_x^i$  given by  $\rho_x^i / p_x^i$  if  $p_x^i$  is nonzero and an arbitrary trace-1 matrix otherwise. Then,  $p_x^1 + \dots + p_x^t = 1$  and  $p_x^1 \tilde{\rho}_x^1 + \dots + p_x^t \tilde{\rho}_x^t = \rho$ . Hence, Theorem 1.4 gives the desired state and measurements.  $\square$

An optimal entanglement-assisted protocol minimizes the number of outcomes  $t$  of Alice's measurement, since these outcomes need to be communicated. This and above considerations now justify defining the following variants of the chromatic number and Witsenhausen rate.

**Definition 1.5** (Entangled chromatic number and Witsenhausen rate). *For a graph  $G$ , define  $\chi^*(G)$  as the minimum integer  $t \in \mathbb{N}$  for which there exist  $d \in \mathbb{N}$  and positive semidefinite matrices  $\rho$  and  $\{\rho_x^i : x \in V(G), i \in [t]\}$  in  $\mathbb{C}^{d \times d}$  such that  $\text{Tr}(\rho) = 1$  and (6) and (7) hold. The entangled Witsenhausen rate is defined by*

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

The operational interpretation of the parameter  $\chi^*(G)$  makes it easy to see that it is sub-multiplicative with respect to strong graph products, that is  $\chi^*(G \boxtimes G') \leq \chi^*(G) \chi^*(G')$ . Indeed, suppose that two sources  $\mathcal{M}$  and  $\mathcal{M}'$  admit entanglement-assisted protocols involving  $t$ - and  $t'$ -outcome measurements done by Alice, respectively. For the combined source  $\mathcal{M} \otimes \mathcal{M}'$  (with characteristic graph  $G \boxtimes G'$ ), Alice and Bob get input *pairs*  $(x, x')$  and  $(u, u')$  respectively and Bob must learn  $(x, x')$ . By running the two protocols separately on an instance of  $\mathcal{M} \otimes \mathcal{M}'$  we get a protocol involving  $t \times t'$ -outcome measurements, which gives the claim. Alternatively, the sub-multiplicativity can be derived from Definition 1.5 using simple matrix manipulations. Sub-multiplicativity of  $\chi^*(G)$  implies that the entangled Witsenhausen rate is also given by the infimum:  $R^*(G) = \inf_m \log \chi^*(G^{\boxtimes m})/m$ .

**Entanglement-assisted channel coding.** In a similar fashion as above one arrives at the following entangled variants of the independence number and Shannon capacity.

**Definition 1.6** (Entangled independence number and Shannon capacity). *For a graph  $H$ , define  $\alpha^*(H)$  as the maximum integer  $M \in \mathbb{N}$  for which there exist  $d \in \mathbb{N}$  and positive semidefinite matrices  $\rho$  and  $\{\rho_i^u : i \in [M], u \in V(H)\}$  in  $\mathbb{C}^{d \times d}$  such that  $\text{Tr}(\rho) = 1$  and*

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

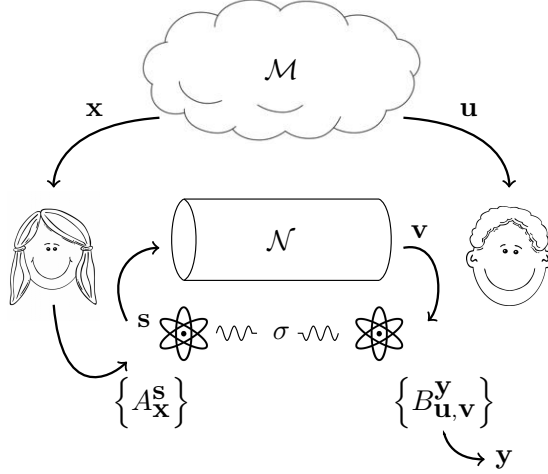
The entangled Shannon capacity is defined by

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

The parameter  $\alpha^*(H)$  was introduced by Cubitt et al. [4], who showed that it equals the maximum number of messages that can be sent without error using entanglement and a single use of a channel with confusability graph  $H$ . It follows that  $c^*(H)$  equals the maximum asymptotic communication rate of such a channel when we allow for entanglement. The parameter  $\alpha^*(H)$  is super-multiplicative (which also follows easily from its operational interpretation) and so in the definition of  $c^*(H)$  the limit can be replaced with the supremum.

In [4] it is shown that  $\alpha^*(H)$  can be strictly larger than  $\alpha(H)$ , meaning that the number of messages that can be sent with a single use of a channel can be increased with the use of entanglement (see also Mančinska, Severini and Scarpa [13]). This result was subsequently strengthened by Leung, Mančinska, Matthews, Ozols and Roy [14] and Briët, Buhrman and Gijswijt [15], who found families of graphs for which  $c^*(H) > c(H)$ .

**Entanglement-assisted source-channel coding.** Finally, we consider the general source-channel coding problem in the entanglement-assisted setting. As before, Alice and Bob receive inputs from a 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 a given length. But in addition Alice and Bob possess quantum registers  $\mathcal{A}$  and  $\mathcal{B}$ , respectively, that are initialized to be in some entangled state  $\sigma$ . The entanglement-assisted version of an  $(m, n)$ -coding scheme is as follows (Figure 2):



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

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

Using the same arguments as above one then arrives at the following variants of the cost-rate.

**Definition 1.7** (Entangled cost rate). *For graphs  $G, H$  and  $m \in \mathbb{N}$ , define  $\eta_m^*(G, H)$  as the minimum integer  $n \in \mathbb{N}$  for which there exist  $d \in \mathbb{N}$  and positive semidefinite matrices  $\rho$  and  $\{\rho_{\mathbf{x}}^{\mathbf{s}} : \mathbf{x} \in V(G^{\boxtimes m}), \mathbf{s} \in V(H^{\boxtimes n})\}$  in  $\mathbb{C}^{d \times d}$  such that  $\text{Tr}(\rho) = 1$  and*

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

The entangled cost rate is defined by

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

As for the classical counterpart, we assume throughout that both graphs  $G$  and  $\overline{H}$  contain at least one edge, thereby excluding trivial settings. It is not difficult to see that we regain the parameter  $\eta(G, H)$  if we restrict the above matrices  $\rho$  and  $\rho_{\mathbf{x}}^{\mathbf{s}}$  to be  $\{0, 1\}$ -valued scalars. Thus sharing an entangled quantum system cannot make the coding scheme worse and so  $\eta^*(G, H) \leq \eta(G, H)$ . As in the classical case, the parameter  $\eta_m^*(G, H)$  is sub-additive (as can be easily be derived by its operational interpretation or by matrix manipulations using Definition 1.7), hence the parameter  $\eta^*(G, H)$  is well defined and can equivalently be given by  $\inf_m \eta_m^*(G, H)/m$ .

We finish this subsection by noticing that the following analogue of Lemma 1.1 holds.

**Lemma 1.8.** *Let  $G$  and  $H$  be graphs such that both  $G$  and  $\overline{H}$  have at least one edge. Then,*

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

*Proof.* Notice that the graph  $\overline{K}_2^{\boxtimes n}$  has  $2^n$  vertices and no edges. It then follows from the definitions that  $\eta_m^*(G, \overline{K}_2) = \lceil \log \chi^*(G^{\boxtimes m}) \rceil$ . The identity  $R^*(G) = \eta^*(G, \overline{K}_2)$  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  $\eta_m^*(K_2, H)$  is the minimum  $n \in \mathbb{N}$  such that  $\alpha^*(H^{\boxtimes n}) \geq 2^m$  or, equivalently,  $\log \alpha^*(H^{\boxtimes n}) \geq m$ . Now we use the same techniques as in Lemma 1.1 to prove that  $1/c^*(H) = \eta^*(K_2, H)$ .  $\square$

**1.3. Further relations to previous work.** 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. [7] studied the *quantum chromatic number*  $\chi_q(G)$ , and Roberson and Mančinska [16] considered a variant of the quantum independence number  $\alpha_q(H)$ . These parameters can be obtained from the respective definitions of  $\chi^*$  and  $\alpha^*$  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. Moreover, we regain  $\chi$  and  $\alpha$  if we further restrict these matrices to be  $\{0, 1\}$ -valued scalars. It thus follows immediately that

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

It is well-known that determining the classical chromatic and independence numbers of a graph are 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 [17, 18] 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  $\chi^*$ ,  $\alpha^*$ ,  $\alpha_q$ ,  $R^*$  and  $c^*$  is an open problem. Very recently, however, Ji [19, Theorem 1 and Theorem 4] showed that it is NP-hard to decide if  $\chi_q(G) \leq 3$ .

**1.4. Outline of the paper.** In Section 1 we introduced the problems and basic notions. In Section 2 we present our main results. In Section 3 we give a brief introduction to quantum information theory and we explain quantum remote state preparation. 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 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  $\chi^*(G)$  is lower bounded by an efficiently computable graph parameter, namely a variant of the famous *theta number* introduced by Szegedy [20]. The theta number itself was originally introduced by Lovász [21] to solve a long-standing problem posed by Shannon [1]: computing the Shannon capacity of the five-cycle. Out of the many equivalent formulations of the theta number (see [22] for a survey), the following is the most appropriate for our setting:

$$(8) \quad \vartheta(G) = \min \left\{ \lambda : \exists Z \in \mathbb{R}^{V(G) \times V(G)}, Z \succeq 0, \right. \\ \left. \begin{aligned} Z(u, u) &= \lambda - 1 \text{ for } u \in V(G), \\ Z(u, v) &= -1 \text{ for } \{u, v\} \notin E(G) \end{aligned} \right\}.$$

Lovász [21] proved that  $\alpha(G) \leq \vartheta(G) \leq \chi(\overline{G})$  holds (this inequality is often referred to as the Sandwich Theorem [22]). The theta number is defined by a semidefinite program and it can be approximated to within arbitrary precision in polynomial time. It thus gives a tractable and in many cases useful bound for both  $\alpha$  and  $\chi$ .

Szegedy [20] introduced the following strengthening of the theta number, which includes an extra linear constraint:

$$(9) \quad \vartheta^+(G) = \min \left\{ \lambda : \exists Z \in \mathbb{R}^{V(G) \times V(G)}, Z \succeq 0, \right. \\ \left. \begin{aligned} Z(u, u) &= \lambda - 1 \text{ for } u \in V(G), \\ Z(u, v) &= -1 \text{ for } \{u, v\} \notin E(G), \\ Z(u, v) &\geq -1 \text{ for } \{u, v\} \in E(G) \end{aligned} \right\}.$$

Szegedy's number satisfies the chain of inequalities:  $\alpha(G) \leq \vartheta(G) \leq \vartheta^+(G) \leq \chi(\overline{G})$ . Lovász [21] proved that  $\vartheta$  is *multiplicative* under the strong graph product, that is,  $\vartheta(G \boxtimes H) = \vartheta(G)\vartheta(H)$ . Moreover Knuth [22] showed that  $\vartheta(\overline{G \boxtimes H}) = \vartheta(\overline{G})\vartheta(\overline{H})$  (recently Cubitt et al. [23] showed that this identity fails for  $\vartheta^+$ ). The identities of Lovász and Knuth give for any graph  $G$  and  $m \in \mathbb{N}$ :

$$(10) \quad \vartheta(\overline{G^{\boxtimes m}}) = \vartheta(\overline{G^{\boxtimes m}}) = \vartheta(\overline{G})^m.$$

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*

$$(11) \quad \vartheta^+(G) \leq \chi^*(\overline{G}),$$

$$(12) \quad \log \vartheta(G) \leq R^*(\overline{G}).$$

In [16] it is shown that the inequality  $\vartheta(G) \leq \chi_q(\overline{G})$  holds. Theorem 2.1 thus strengthens this bound as it gives  $\vartheta(G) \leq \vartheta^+(G) \leq \chi^*(\overline{G}) \leq \chi_q(\overline{G})$ .

Beigi [24] and Duan, Severini and Winter [25] proved that  $\vartheta(G)$  upper bounds  $\alpha^*(G)$ . The above-mentioned relations therefore imply the following sequence of inequalities:

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

Moreover, building on this work, Cubitt et al. [23] proved that  $\eta^*(G, H) \geq \log \vartheta(\overline{G}) / \log \vartheta(H)$ .

**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 [15] 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  $(k + 1)/2$ .*

The usual *orthogonality graph* has vertex set  $\{-1, 1\}^k$  (for  $k$  even) and two vertices are adjacent if they are orthogonal. The quarter-orthogonality graph is a subgraph of the orthogonality graph, which can be seen by appending a ‘1’ to the vertices of  $H_k$ . We record the following simple lower bound on the independence number of  $H_k$  results for later use.

**Lemma 2.3.** *For every odd positive integer  $k$ , 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  $(k + 1)/2$ .  $\square$

Some of our results rely on the existence of certain Hadamard matrices. A *Hadamard matrix* is a square matrix  $A \in \{-1, 1\}^{\ell \times \ell}$  that satisfies  $AA^T = \ell I$ . The size  $\ell$  of a Hadamard matrix must necessarily be 2 or a multiple of 4 and the famous Hadamard conjecture (usually attributed to Paley [26]) 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 [27] (see for example [28, 29, 30, 31, 32] for closely related constructions).

**Theorem 2.4** (Xia and Liu [27]). *Let  $q$  be a prime power such that  $q \equiv 1 \pmod{4}$ . 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 [15]). *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.

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

$$(13) \quad R^*(H_k) \leq \log(k + 1).$$

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

$$(14) \quad R(H_k) \geq 0.154k - 1.$$

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 [15], 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 [15]). *Let  $p$  be an odd prime such that there exists a Hadamard matrix of size  $4p$ . Set  $k = 4p - 1$ . Then,*

$$\begin{aligned} c^*(H_k) &\geq k - 1 - 2 \log(k + 1), \\ c(H_k) &\leq 0.846k. \end{aligned}$$

Note that here we consider the exact bounds on  $c^*(H_k)$  and  $c(H_k)$  rather than the asymptotic ones as originally written in [15]. 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  $c^*(H_k)$ , which originates from [14]. 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 [33].

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  $k \geq 11$ , we have*

$$(15) \quad c^*(H_k) \geq (k-1) \left( 1 - \frac{2 \log(k+1)}{k-3} \right).$$

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

$$(16) \quad c(H_k) \leq 0.846 k.$$

We prove Theorem 2.8 in Section 6. To prove (15) we use a technique that is based on quantum remote state preparation [34]. This proof technique appears not to have been considered before in the context of zero-error entanglement-assisted communication. The proof of (16) combines an instance of the linear algebra method due to Alon [35] 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 [36]. Roughly this combination was previously used in the context of Ramsey graphs [37].

*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,*

$$(17) \quad \eta^*(H_k, H_k) \leq \frac{\log(k+1)}{(k-1) \left( 1 - \frac{2 \log(k+1)}{k-3} \right)},$$

$$(18) \quad \eta(H_k, H_k) > \frac{0.154 k - 1}{k - 1 - \log(k+1)}.$$

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 3.3 for details. There we show that the asymptotic cost rate of a separate coding scheme is  $R^*(H_k)/c^*(H_k)$  and thus  $\eta^*(H_k, H_k) \leq R^*(H_k)/c^*(H_k)$ . The bound for the classical parameter  $\eta(H_k, H_k)$  relies on the No-Homomorphism Lemma due to Albertson and Collins [38] and the 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 [27] in Theorem 2.4, since there exist infinitely many  $(p, \ell)$ -pairs such that  $p^{\ell/2} \equiv 1 \pmod{4}$ . (For instance, for  $p = 5$  and  $\ell = 2i$  with  $i \in \mathbb{N}$ ,  $5^i = (4+1)^i \equiv 1 \pmod{4}$ .)

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

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

As shown in [10], 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,  $\eta(H_k, H_k) \geq$



$\Omega(1)$  and

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

where in the second last inequality we use that  $\log \chi(H_k) \leq \log |V(H_k)| = k-1$  and, from Lemma 2.3, that  $\log \alpha(H_k) \geq (k-3)/2$ .

### 3. ENTANGLEMENT-ASSISTED SOURCE-CHANNEL CODING

In this section we describe quantum remote state preparation and show how to obtain a coding scheme for the source-channel problem from concatenating a coding scheme for the source with one from the channel. 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 [39].

**3.1. Quantum states and operations.** Recall that a *positive semidefinite matrix* is a Hermitian matrix whose eigenvalues are non-negative. 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*  $\rho$ . A state with rank 1 is called a *pure state*. An experimenter can alter a state  $\rho \in \mathbb{C}^{d \times d}$  of his or her register by performing a unitary transformation. This is a mapping  $\rho \mapsto U\rho U^*$ , where  $U \in \mathbb{C}^{d \times d}$  satisfies  $UU^* = I$ .

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*  $\rho_A \otimes \rho_B$ . If Alice performs a unitary  $U \in \mathbb{C}^{d_A \times d_A}$  on her register  $\mathcal{A}$  then the state  $\rho$  of the system  $(\mathcal{A}, \mathcal{B})$  is mapped to  $(U \otimes I)\rho(U \otimes I)^*$ , and similarly if Bob performs a unitary on his register.

A *t-outcome measurement* is a collection  $\mathbf{M} = \{M_i \in \mathbb{C}^{d \times d} : i \in [t]\}$  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  $\mathbf{M}$  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  $\Pr[\lambda = i] = \text{Tr}(M_i \rho)$ . 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  $\text{Tr}_{\mathcal{A}}(A \otimes B) = \text{Tr}(A)B$  and  $\text{Tr}_{\mathcal{B}}(A \otimes B) = A \text{Tr}(B)$ , 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  $\mathbf{M}$  on  $\mathcal{A}$ . Then, the probability that Alice gets measurement outcome  $i$  equals

$p_i = \text{Tr}((M_i \otimes I)\rho)$ . Moreover, in the event that Alice gets measurement outcome  $i$ , Bob's register  $\mathcal{B}$  is left in the state  $\rho^i = \text{Tr}_{\mathcal{A}}((M_i \otimes I)\rho)/p_i$ . If Bob now performs an  $r$ -outcome measurement  $M'$  on  $\mathcal{B}$ , then the probability that he gets outcome  $j \in [r]$  equals  $\text{Tr}_{\mathcal{B}}(M'_j \rho^i)$ .

**3.2. Remote state preparation.** Suppose that Alice has in mind a pure state  $\rho = uu^*$  where  $u \in \mathbb{C}^d$  is some unit vector that is unknown to Bob. *Remote state preparation* [34] is a protocol that enables the parties to prepare a quantum register belonging to Bob in the state  $\rho$  using only local measurements on a pair of entangled quantum registers and classical communication from Alice to Bob. This task can be achieved using the teleportation scheme of Bennett et al. [40], which allows to remotely prepare a  $d$ -dimensional state  $\rho$  with the communication of one among  $d^2$  distinct messages. (We refer interested readers to [40] and [39, pp. 26–28] for the details of the teleportation scheme.) However, for certain states, remote state preparation can be performed with less communication. In particular, for the case where the vector  $u$  has only entries of absolute value  $d^{-1/2}$  we now describe a remote state preparation protocol that requires only sending one among  $d$  distinct messages, or equivalently  $\lceil \log d \rceil$  bits.

To remotely prepare the state  $\rho = uu^*$ , where  $u \in \mathbb{C}^d$  is a vector with entries of modulus  $d^{-1/2}$  that is unknown to Bob, in a  $d$ -dimensional register  $\mathcal{Y}$  that he possesses the two parties can use the following protocol: Assume Alice has a  $d$ -dimensional quantum register  $\mathcal{X}$ , such that  $(\mathcal{X}, \mathcal{Y})$  is in the *maximally entangled state*  $\sigma = vv^*$  defined by the vector

$$v = \frac{1}{\sqrt{d}} \sum_{\ell=1}^d e_{\ell} \otimes e_{\ell},$$

where  $e_{\ell}$  denotes the  $\ell$ th canonical basis vector.

- (1) Alice performs on her register  $\mathcal{X}$  the unitary transformation given by  $U = \sqrt{d} \text{Diag}(u)$ , where  $\text{Diag}(u)$  is the diagonal matrix whose main diagonal is vector  $u$ .
- (2) Next, Alice performs on her register the discrete Fourier transform  $F \in \mathbb{C}^{d \times d}$ , which is given by the unitary matrix  $F_{\ell,m} = e^{2\pi i(\ell-1)(m-1)/d} / \sqrt{d}$ .
- (3) She then measures in the canonical basis (i.e., she performs the measurement  $\{e_{\ell} e_{\ell}^* : \ell \in [d]\}$  on her register) and gets an outcome  $\ell \in [d]$  which she communicates to Bob.
- (4) Last, Bob performs on his register  $\mathcal{Y}$  the unitary given by the diagonal matrix whose main diagonal is the vector  $(e^{-2\pi i(\ell-1)(m-1)/d})_{m=1}^d$ .

The correctness of the protocol follows easily from the following observations. After step (2) the state of the register pair  $(\mathcal{X}, \mathcal{Y})$  is given by  $(FU \otimes I)vv^*(FU \otimes I)^*$ . Notice that if  $a_{\ell}, b_{\ell}$  are the  $\ell$ th column and row of  $FU$ , resp., then

$$(FU \otimes I)v = \frac{1}{\sqrt{d}} \sum_{\ell=1}^d a_{\ell} \otimes e_{\ell} = \frac{1}{\sqrt{d}} \sum_{\ell=1}^d e_{\ell} \otimes b_{\ell}$$

Since the  $\ell$ th row of the matrix  $FU$  is given by

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

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

**3.3. Separate coding schemes.** By concatenating a coding scheme for a source with one for a channel, one obtains a coding scheme for the combined source-channel problem. For this to work, the number of bits one can send perfectly with  $n$  uses of the channel must be at least as large as the number of bits required to solve  $m$  instances of the source problem. In other words, for a source with characteristic graph  $G$  and a channel with confusability graph  $H$ , we need the condition  $\chi(G^{\boxtimes m}) \leq \alpha(H^{\boxtimes n})$  in order to send length  $m$  source-input sequences with  $n$  uses of the channel. If this condition holds, then it follows that  $\eta_m(G, H) \leq n$ . The same type of reasoning holds also in the entanglement-assisted case. We have just proved the following simple lemma, which can alternatively be shown using the definition of a graph homomorphism and simple matrix manipulations.

**Lemma 3.1.** *Given graphs  $G, H$  and positive integers  $n, m$ , we have*

$$(19) \quad \chi(G^{\boxtimes m}) \leq \alpha(H^{\boxtimes n}) \implies \eta_m(G, H) \leq n,$$

$$(20) \quad \chi^*(G^{\boxtimes m}) \leq \alpha^*(H^{\boxtimes n}) \implies \eta_m^*(G, H) \leq n.$$

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

**Proposition 3.2.** *Let  $G$  and  $H$  be graphs and assume that both  $G$  and  $\overline{H}$  have at least one edge. Then,*

$$(21) \quad \eta(G, H) \leq \frac{R(G)}{c(H)} = \lim_{m \rightarrow \infty} \frac{1}{m} \epsilon_m(G, H),$$

$$(22) \quad \eta^*(G, H) \leq \frac{R^*(G)}{c^*(H)} = \lim_{m \rightarrow \infty} \frac{1}{m} \epsilon_m^*(G, H),$$

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

*Proof.* We show (21); we omit the proof of (22) which is analogous (and uses (20)). From (19) we have the inequality:  $\eta_m(G, H) \leq \epsilon_m(G, H)$ , which implies  $\eta(G, H) \leq \lim_{m \rightarrow \infty} \epsilon_m(G, H)/m$ . Next we show that this limit is equal to  $R(G)/c(H)$ , which concludes the proof of (21). 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  $m \rightarrow \infty$  in the right most terms we obtain that  $R(G)/c(H)$  is equal to  $\lim_{m \rightarrow \infty} \epsilon_m(G, H)/m$ .  $\square$

#### 4. LOWER BOUND ON THE ENTANGLED CHROMATIC NUMBER

In this section we prove Theorem 2.1, which we restate here for convenience.

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

$$(23) \quad \vartheta^+(G) \leq \chi^*(\overline{G}),$$

$$(24) \quad \log \vartheta(G) \leq R^*(\overline{G}).$$

We will use the following simple fact about positive semidefinite matrices with a special block form (which can be found, e.g., in [41, Lemma 2.8]).

**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 (23) and (24) hold for the graph  $\overline{G}$ . First we observe that (24) follows easily from (23). Indeed, (23) combined with the identity (10) implies  $\vartheta(\overline{G})^m = \vartheta(\overline{G^{\boxtimes m}}) \leq \chi^*(G^{\boxtimes m})$  and thus  $\log \vartheta(\overline{G}) \leq R^*(G)$  follows after taking limits.

We now prove (23) for the graph  $\overline{G}$ , i.e., we show that the inequality  $\vartheta^+(\overline{G}) \leq \chi^*(G)$  holds. 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  $\chi^*(G) = t$ . We may assume that  $\langle \rho, \rho \rangle = 1$ . 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  $\text{Sym}(t)$ , the group of permutations of  $[t]$ , we define the new (permuted) matrix  $\sigma(X) = (X_{u\sigma(i), v\sigma(j)})$ . Then we average the matrix  $X$  over the group  $\text{Sym}(t)$ , obtaining the new matrix

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

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

$$(25) \quad A_{uv} = 0 \quad \text{for all } \{u, v\} \in E(G).$$

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  $u, v \in V(G)$ :

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

implying that  $t(A_{uv} + (t-1)B_{uv}) = 1$  holds since

$$(26) \quad \sum_{i \in [t]} \sum_{j \in [t]} Y_{ui,vj} = t \sum_{j \in [t]} Y_{ui,vj} = t(A_{uv} + (t-1)B_{uv}).$$

We are now ready to define a matrix  $Z$  which is a feasible solution for the program (9) 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 (25) and (26) give  $A_{uv} = 0$  and  $t(t-1)B_{uv} = 1$  and thus  $Z_{uv} = -1$ . For a non-edge  $\{u, v\}$ , relation (26) combined with the fact that  $A_{uv} \geq 0$  implies that  $Z_{uv} \geq -1$ . Finally, for any  $u \in V(G)$ , relation (26) 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 (9) defining  $\vartheta^+(\overline{G})$ . This shows that  $\vartheta^+(\overline{G}) \leq \chi^*(G)$ , which concludes the proof.  $\square$

## 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. We repeat the formal statement of the theorem here for convenience.

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

$$(27) \quad R^*(H_k) \leq \log(k+1).$$

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

$$(28) \quad R(H_k) \geq 0.154k - 1.$$

**5.1. Upper bound on the entangled Witsenhausen rate.** Here we prove the upper bound (27) on  $R^*(H_k)$  stated in Theorem 2.6. We first prove that  $\chi^*(G) \leq \xi'(G)$ . This inequality can be derived from the fact that  $\chi^*(G) \leq \chi_q(G)$  and a result of [7] stating that  $\chi_q(G) \leq \xi'(G)$ . We give self-contained proofs of the implied bound on  $\chi^*(G)$  here.

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

We give two simple proofs of this fact, the first operational in spirit, the second combinatorial. Although the combinatorial proof is more elementary, the operational proof will prepare the reader for the slightly more involved proof of Lemma 6.1 below, for which an operational proof appears to be most natural. In addition, it follows almost immediately from the operational proof that  $\chi^*(G) \leq \xi(G)^2$ ; the only thing to change in the proof is to replace the remote state preparation scheme described in Section 3.2 by quantum teleportation.

*Proof.* Operational proof: Consider a source  $\mathcal{M} = (\mathsf{X}, \mathsf{U}, P)$  with characteristic graph  $G$ . Let  $d = \xi'(G)$  and let  $f$  be a  $d$ -dimensional orthogonal representation of  $G$  such that all the entries of the vectors have modulus one. By the discussion in Section 1.2 it suffices to find an entanglement-assisted protocol for  $\mathcal{M}$  that involves only  $d$ -outcome measurements on Alice's part. To this end, let us recall the observation that Bob's input  $u$  already allows him to reduce the list of Alice's possible inputs to a clique  $\mathcal{C}$  in  $G$  that contains Alice's actual input  $x$ . Next, notice that the set of states  $f(y)f(y)^*/d$  for  $y \in \mathcal{C}$  are pairwise orthogonal. This suggests the following protocol. First the parties perform the remote state preparation protocol (Section 3.2) to put a quantum register belonging to Bob in the state  $f(x)f(x)^*/d$ . Now Bob performs the measurement with outcomes in  $\mathcal{C} \cup \{\perp\}$  as promised to exist by the Orthogonality Lemma (Lemma 1.2) to learn which of the states  $f(y)f(y)^*/d$  his register is in, thus learning  $x$ . The result now follows because the remote state preparation involves only  $d$ -outcome measurements.

Combinatorial proof: Set  $d = \xi'(G)$ ,  $\omega_d = e^{2i\pi/d}$  and let  $h_i = [1, \omega_d^i, \omega_d^{2i}, \dots, \omega_d^{(d-1)i}]^\top \in \mathbb{C}^d$  for each  $i \in [d]$ . 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  $\text{Tr}(\rho) = 1$ . Let  $f : V(G) \rightarrow \mathbb{C}^d$  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  $\rho_u^i = (f(u) \circ h_i)(f(u) \circ h_i)^*$ , 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 vertex  $u \in V(G)$ , we have  $\sum_{i \in [d]} \rho_u^i = I/d = \rho$ . 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 1.5 and so  $\chi^*(G) \leq d$ .  $\square$

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 (27).

**Lemma 5.2.** *Let  $k$  be an odd positive integer and  $m \in \mathbb{N}$ . Then,  $\chi^*(H_k^{\boxtimes m}) \leq (k+1)^m$ . Moreover, if there exists a Hadamard matrix of size  $k+1$ , then equality holds.*

*Proof.* Consider the map  $f : V(H_k) \rightarrow \{-1, 1\}^{k+1}$  defined by  $f(u) = (u, 1)^\top$  (i.e.,  $f$  sends a vertex  $u \in \{-1, 1\}^k$  to itself and appends a "1" to it). Then clearly  $f$  is an orthogonal representation of  $H_k$  and since its vectors have entries of modulus one we get  $\xi'(H_k) \leq k+1$ . By Lemma 5.1 and sub-multiplicativity we therefore have  $\chi^*(H_k^{\boxtimes m}) \leq \chi^*(H_k)^m \leq (k+1)^m$ .

We now prove that if there exists a Hadamard matrix of size  $k+1$  then also the reverse inequality holds:  $\chi^*(H_k^{\boxtimes m}) \geq (k+1)^m$ . Recall from Proposition 2.5 that 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

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

where the third inequality uses the simple fact that if a subset  $W \subseteq V(G)$  forms a clique in a graph  $G$ , then the set  $W^m$  of  $m$ -tuples of elements from  $W$  forms a clique in  $G^{\boxtimes m}$ .  $\square$

The bound (27) now follows as a simple corollary.

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

*Proof.* By Lemma 5.2 we have

$$R^*(H_k) = \inf_m \log \chi^*(H_k^{\boxtimes m})/m \leq \log \chi^*(H_k) \leq \log(k+1).$$

□

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 (28) 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*

$$(29) \quad \alpha(H_k^{\boxtimes m}) \leq \left( \binom{k}{0} + \binom{k}{1} + \cdots + \binom{k}{p^\ell - 1} \right)^m \leq 2^{kmH(3/11)} < 2^{0.846km},$$

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 [35] (see also Gopalan [37]), 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  $((f_u, c_u))_{u \in V(G)} \subseteq \mathcal{F} \times \mathbb{F}^k$  of polynomial-point pairs to the vertices of  $G$  such that

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

**Lemma 5.6** (Alon [35]). *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  $((f_u, c_u))_{u \in V} \subseteq \mathcal{F} \times \mathbb{F}^k$  represents  $G$ , then  $\alpha(G^{\boxtimes n}) \leq \dim(\mathcal{F})^n$  for all  $n \in \mathbb{N}$ .*

We get a representation for the graph  $H_k$ , for  $k = 4p^\ell - 1$ , from the following result of Barrington, Beigel and Rudich [36] (see [42, Lemma 5.6] for the statement as it appears below).

**Lemma 5.7** (Barrington, Beigel and Rudich [36]). *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*

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

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  $|c| \equiv 0 \pmod{p^\ell}$ . Then, since  $p$  is odd and  $k < 4p^\ell$ , we have  $|c| \in \{0, 2p^\ell\}$ . Hence, if  $|c| \notin \{0, 2p^\ell\}$ , then  $|c| \not\equiv 0 \pmod{p^\ell}$ .

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  $|u \oplus v| \not\equiv 0 \pmod{p^\ell}$ .

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

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

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  $((f_u, c_u))_{u \in V(H_k)}$  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  $\binom{k}{0} + \binom{k}{1} + \dots + \binom{k}{p^\ell - 1}$ , which is the number of multilinear monomials of degree at most  $p^\ell - 1$ . Applying Lemma 5.6 we obtain that

$$(30) \quad \alpha(H_k^{\boxtimes m}) \leq \left( \binom{k}{0} + \binom{k}{1} + \dots + \binom{k}{p^\ell - 1} \right)^m.$$

We now use the well known fact that for  $q, k \in \mathbb{N}$  with  $1 < q < k/2$ ,  $\binom{k}{0} + \dots + \binom{k}{q-1} \leq 2^{kH(q/k)}$ . From this, since  $p^\ell / (4p^\ell - 1) \leq 3/11$ , we deduce that the right hand side in (30) can be upper bounded by  $2^{kmH(3/11)} < 2^{0.846km}$ .  $\square$

The bound (28) stated in Theorem 2.6 is a simple corollary of Lemma 5.5.

**Corollary 5.8.** *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  $m \rightarrow \infty$  gives the result.  $\square$



## 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. We repeat the statement of the theorem here for convenience.

**Theorem 2.8.** *For every odd integer  $k \geq 11$ , we have*

$$(31) \quad c^*(H_k) \geq (k-1) \left( 1 - \frac{2 \log(k+1)}{k-3} \right).$$

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

$$(32) \quad c(H_k) \leq 0.846 k.$$

**6.1. Lower bound on the entangled Shannon capacity.** The proof of the bound (31) on the entangled Shannon capacity is again based on remote state preparation (Section 3.2).

The idea behind the proof is to show that with  $t+1$  sequential uses of a channel with confusability graph  $H_k$ , Alice can send Bob  $|V(H_k)|^t$  distinct messages with zero probability of error, provided that  $t \leq \log \alpha(H_k) / \log \xi'(H_k)$ . More generally, we have the following lemma, from which the result (31) easily follows.

**Lemma 6.1.** *For every graph  $G$  and integer  $t \geq 1$  such that  $t \leq \log \alpha(G) / \log \xi'(G)$ , we have*

$$c^*(G) \geq \frac{\log |V(G)|}{1 + \frac{1}{t}}.$$

*Proof.* Let  $\mathcal{N} = (\mathbf{S}, \mathbf{V}, Q)$  be a channel with confusability graph  $G$ . Let  $d = \xi'(G)$  and let  $f$  be a  $d$ -dimensional orthogonal representation of  $G$  such that its vectors have entries of modulus one. For each  $x \in V(G)$  define  $\rho_x = f(x)f(x)^*/d$ . Let  $t$  be a positive integer such that

$$(33) \quad t \leq \frac{\log \alpha(G)}{\log d}.$$

By [4, Theorem 1] it suffices to find an entanglement-assisted protocol for the noiseless transmission of  $|V(G)|^t$  distinct messages based on at most  $t+1$  uses of the channel  $\mathcal{N}$ . Indeed this then implies that  $\alpha^*(G^{\boxtimes(t+1)}) \geq |V(G)|^t$  and therefore

$$c^*(G) \geq \frac{\log \alpha^*(G^{\boxtimes(t+1)})}{t+1} \geq \frac{t \log |V(G)|}{1+t}$$

as claimed. To this end, we consider the following four-step protocol for transmitting a sequence  $\mathbf{x} = (x_1, \dots, x_t) \in V(G)^t$ . First, Alice prepares  $d$ -dimensional quantum registers  $\mathcal{A}_1, \dots, \mathcal{A}_t$  to be in the states  $\rho_{x_1}, \dots, \rho_{x_t}$ , respectively. Second, Alice sends the sequence  $\mathbf{x}$  through the channel by using it  $t$  times in a row. This will result in  $t$  channel-outputs on Bob's end of the channel from which he can infer that each  $x_i$  belongs to a particular clique in  $G$ . Third, Alice and Bob execute the remote state preparation scheme described in Section 3.2  $t$  times in a row, once for each of the states  $\rho_{x_1}, \dots, \rho_{x_t}$  separately. (Recall that  $\rho_{x_i} = f(x_i)f(x_i)^*/d$  where  $f(x_i) \in \mathbb{C}^d$  has norm  $\sqrt{d}$ , so in the notation of Section 3.2

we are setting  $u = f(x_i)/\sqrt{d}$ .) This requires that Alice communicates a total of  $t\lceil\log d\rceil$  bits to Bob. To do so, Alice uses the channel one more time to send, without error, the bits required to perform the remote state preparation. This can be done if  $\log \alpha(G) \geq t\lceil\log d\rceil$ , which holds by our assumed bound (33). At this point Bob's quantum registers  $\mathcal{B}_1, \dots, \mathcal{B}_t$  are in states  $\rho_{x_1}, \dots, \rho_{x_t}$ . Moreover, for each  $x_i$  Bob knows a clique in the graph  $G$  that contains  $x_i$  and by construction elements of a clique have pairwise orthogonal states. In the last step, for every  $i \in [t]$ , Bob can perform a measurement on register  $\mathcal{B}_i$  such that he gets outcome  $x_i$  with probability one (due to Lemma 1.2). Hence, Bob can recover any sequence  $\mathbf{x} = (x_1, \dots, x_t) \in V(G)^t$  with zero probability of error, completing the proof.  $\square$

*Proof of (31).* Recall from the proof of Lemma 5.2 that  $\log \xi'(H_k) \leq \log(k+1)$ . From Lemma 2.3, we know that  $\log \alpha(H_k) \geq (k-3)/2$ . Hence, for  $k \geq 11$  we can choose  $t = \lfloor (k-3)/(2\log(k+1)) \rfloor \leq \lfloor \log \alpha(H_k) / \log \xi'(H_k) \rfloor$ . Note that we require  $k \geq 11$  to ensure that  $t \geq 1$ . We now apply Lemma 6.1 and, recalling that  $|V(H_k)| = 2^{(k-1)}$ , we obtain

$$c^*(H_k) \geq (k-1) \frac{1}{1 + \frac{1}{t}} \geq (k-1) \left( 1 - \frac{2\log(k+1)}{k-3} \right)$$

which gives the result.  $\square$

**6.2. Upper bound on the Shannon capacity.** The upper bound (32) 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  $m \rightarrow \infty$  on both sides of (29) we get the result.  $\square$

## 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 (17) for  $\eta^*$  and (18) for  $\eta$ .

**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,*

$$(34) \quad \eta^*(H_k, H_k) \leq \frac{\log(k+1)}{(k-1) \left( 1 - \frac{2\log(k+1)}{k-3} \right)},$$

$$(35) \quad \eta(H_k, H_k) > \frac{0.154k - 1}{k - 1 - \log(k+1)}.$$

The bound (34) is obtained by combining (13), (15) with Proposition 3.2. The proof of (18) relies on the No-Homomorphism Lemma due to Albertson and Collins [38].

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  $v = \pi(u)$ .

**Lemma 7.1** (No-Homomorphism Lemma, Albertson and Collins [38]). *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 [15], the graph  $H_k$  is vertex-transitive; indeed, for any  $u \in V(H_k)$ , the map  $v \mapsto u \oplus v$  is an automorphism of  $H_k$ . It is easy to see that vertex transitivity is preserved under strong products and complements. Hence,  $\overline{H_k^{\boxtimes n}}$  is vertex-transitive for any  $n \in \mathbb{N}$ .

*Proof of (35).* Recall the definition of  $\eta(H_k, H_k)$  from (3). Consider integers  $m, n \in \mathbb{N}$  for which  $H_k^{\boxtimes m} \rightarrow \overline{H_k^{\boxtimes n}}$ . Applying Lemma 7.1, we deduce that

$$(36) \quad \frac{|V(H_k^{\boxtimes m})|}{\alpha(H_k^{\boxtimes m})} \leq \frac{|V(\overline{H_k^{\boxtimes n}})|}{\alpha(\overline{H_k^{\boxtimes n}})} = \frac{|V(H_k^{\boxtimes n})|}{\omega(H_k^{\boxtimes n})}.$$

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

$$\frac{2^{(k-1)m}}{2^{km} 0.846} \stackrel{\text{Lemma 5.5}}{<} \frac{|V(H_k^{\boxtimes m})|}{\alpha(H_k^{\boxtimes m})} \stackrel{(36)}{\leq} \frac{|V(\overline{H_k^{\boxtimes n}})|}{\omega(H_k^{\boxtimes n})} = \frac{2^{(k-1)n}}{(k+1)^n}.$$

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

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

□

## 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 [10] is that, for the classical source-channel coding problem, there exist situations for which separate encoding is highly suboptimal. Does this happen also in the entanglement-assisted case? This question has a positive answer if there exists a graph  $G$  with  $R^*(G) > c^*(\overline{G})$ . In [10] 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  $\chi^*$  and  $\alpha^*$ .

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 [16] it is conjectured that  $\alpha^*(G) = \alpha_q(G)$  for every graph  $G$ . A possible approach to show that  $\chi^*$  and  $\chi_q$  are two separate quantities is to prove that the relationship between Kochen-Specker sets and  $\chi_q$  found in [13] does not hold for  $\chi^*$ . Finally, we mention that the existence of a graph  $G$  for which  $\chi^*(G) < \chi_q(G)$  or  $\alpha_q(G) < \alpha^*(G)$  would prove the existence of a non-local game such that every quantum strategy that wins with probability one does not use a maximally entangled state.

#### ACKNOWLEDGMENTS

We thank Ronald de Wolf for useful discussions and helpful comments on a preliminary version of this paper, and we thank Dion Gijswijt and Debbie Leung for helpful discussions, and the anonymous referees for helpful comments on an earlier version of this manuscript, for pointing out an error in a previous version of Lemma 1.2, and for providing the reference for Theorem 1.4.

J. B. was partially supported by a Rubicon grant from the Netherlands Organization for Scientific Research (NWO). J. B. and H. B. were supported by the European Commission under the project QCS (Grant No. 255961). H. B. and T. P. were supported by the EU grant SIQS. G. S. was supported by Ronald de Wolf's Vidi grant 639.072.803 from the Netherlands Organization for Scientific Research (NWO). This work was done while G.S. was a PhD student at CWI, Amsterdam.

#### REFERENCES

- [1] C. Shannon, "The zero error capacity of a noisy channel," *IRE Transactions on Information Theory*, vol. 2, no. 3, pp. 8–19, 1956.
- [2] J. Körner and A. Orlitsky, "Zero-error information theory," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2207–2229, 1998.
- [3] E. Lubetzky, "Graph powers and related extremal problems," Ph.D. dissertation, Tel Aviv University, 2007.
- [4] T. S. Cubitt, D. Leung, W. Matthews, and A. Winter, "Improving zero-error classical communication with entanglement," *Physical Review Letters*, vol. 104, no. 23, p. 230503, 2010.
- [5] A. Einstein, P. Podolsky, and N. Rosen, "Can quantum-mechanical description of physical reality be considered complete?" *Physical Review*, vol. 47, pp. 777–780, 1935.
- [6] J. S. Bell, "On the Einstein-Podolsky-Rosen paradox," *Physics*, vol. 1, pp. 195–200, 1964.
- [7] P. Cameron, A. Montanaro, M. Newman, S. Severini, and A. Winter, "On the quantum chromatic number of a graph," *Electronic Journal of Combinatorics*, vol. 14, no. R81, p. 1, 2007.
- [8] H. S. Witsenhausen, "The zero-error side information problem and chromatic numbers," *IEEE Transactions on Information Theory*, vol. 22, no. 5, pp. 592–593, 1976.
- [9] M. J. Ferguson and D. W. Bailey, "Zero-error coding for correlated sources," unpublished manuscript, 1975.
- [10] J. Nayak, E. Tuncel, and K. Rose, "Zero-error source-channel coding with side information," *IEEE Transactions on Information Theory*, vol. 52, no. 10, pp. 4626–4629, 2006.
- [11] S. Vembu, S. Verdú, and Y. Steinberg, "The source-channel separation theorem revisited," *IEEE Transactions on Information Theory*, vol. 41, no. 1, pp. 44–54, 1995.
- [12] R. Spekkens and T. Rudolph, "Optimization of coherent attacks in generalizations of the bb84 quantum bit commitment protocol," *Quantum Information & Computation*, vol. 2, pp. 66–96, 2002.

- [13] 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*, vol. 59, no. 6, pp. 4025–4032, 2013.
- [14] 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*, vol. 311, pp. 97–111, 2012.
- [15] J. Briët, H. Buhrman, and D. Gijswijt, “Violating the Shannon capacity of metric graphs with entanglement,” *Proceedings of the National Academy of Sciences*, 2012.
- [16] D. E. Roberson and L. Mančinska, “Graph homomorphisms for quantum players,” arXiv:1212.1724, 2012.
- [17] N. Alon, “Graph powers,” *Contemporary Combinatorics*, pp. 11–28, 2002.
- [18] N. Alon and E. Lubetzky, “The Shannon capacity of a graph and the independence numbers of its powers,” *IEEE Transactions on Information Theory*, vol. 52, no. 5, pp. 2172–2176, 2006.
- [19] Z. Ji, “Binary constraint system games and locally commutative reductions,” arXiv:1310.3794, 2013.
- [20] M. Szegedy, “A note on the  $\vartheta$  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.
- [21] L. Lovász, “On the Shannon capacity of a graph,” *IEEE Transactions on Information Theory*, vol. 25, no. 1, pp. 1–7, 1979.
- [22] D. E. Knuth, “The sandwich theorem,” *Electronic Journal of Combinatorics*, vol. 1, no. A1, 1994.
- [23] T. S. Cubitt, L. Mančinska, D. E. Roberson, S. Severini, D. Stahlke, and A. Winter, “Bounds on entanglement assisted source-channel coding via the Lovász theta number and its variants,” *IEEE Transactions on Information Theory*, vol. 60, no. 11, pp. 7330–7344, 2014.
- [24] S. Beigi, “Entanglement-assisted zero-error capacity is upper-bounded by the Lovász  $\vartheta$  function,” *Physical Review A*, vol. 82, no. 1, p. 010303, 2010.
- [25] 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*, vol. 59, no. 2, pp. 1164–1174, 2013.
- [26] R. E. A. C. Paley, “On orthogonal matrices,” *Journal of Mathematics and Physics (now called Studies in Applied Mathematics)*, vol. 12, pp. 311–320, 1933.
- [27] M. Xia and G. Liu, “An infinite class of supplementary difference sets and Williamson matrices,” *Journal of Combinatorial Theory, Series A*, vol. 58, no. 2, pp. 310–317, 1991.
- [28] —, “A new family of supplementary difference sets and Hadamard matrices,” *Journal of Statistical Planning and Inference*, vol. 51, no. 3, pp. 283–291, 1996.
- [29] R. Wilson and Q. Xiang, “Constructions of Hadamard difference sets,” *Journal of combinatorial theory. Series A*, vol. 77, no. 1, pp. 148–160, 1997.
- [30] Y. Chen, “On the existence of abelian Hadamard difference sets and a new family of difference sets,” *Finite fields and their Applications*, vol. 3, no. 3, pp. 234–256, 1997.
- [31] Q. Xiang, “Difference families from lines and half lines,” *European Journal of Combinatorics*, vol. 19, no. 3, pp. 395–400, 1998.
- [32] T. Xia, J. Seberry, and M. Xia, “New constructing of regular Hadamard matrices,” in *Proceedings of the 10th WSEAS international conference on Computers (CCOMP’06)*, pp. 1294–1299, 2006.
- [33] P. Frankl and R. Wilson, “Intersection theorems with geometric consequences,” *Combinatorica*, vol. 1, no. 4, pp. 357–368, 1981.
- [34] C. H. Bennett, D. P. DiVincenzo, P. W. Shor, J. A. Smolin, B. M. Terhal, and W. K. Wootters, “Remote state preparation,” *Physical Review Letters*, vol. 87, p. 077902, 2001.
- [35] N. Alon, “The Shannon capacity of a union,” *Combinatorica*, vol. 18, no. 3, pp. 301–310, 1998.
- [36] D. Barrington, R. Beigel, and S. Rudich, “Representing boolean functions as polynomials modulo composite numbers,” *Computational Complexity*, vol. 4, no. 4, pp. 367–382, 1994.

- [37] P. Gopalan, “Constructing Ramsey graphs from boolean function representations,” in *Proceedings of the 21st Annual IEEE Conference on Computational Complexity (CCC 2006)*, pp. 14, 2006.
- [38] M. O. Albertson and K. L. Collins, “Homomorphisms of 3-chromatic graphs,” *Discrete Mathematics*, vol. 54, no. 2, pp. 127–132, 1985.
- [39] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. New York: Cambridge University Press, 2000.
- [40] 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*, vol. 70, pp. 1895–1899, 1993.
- [41] N. Gvozdenovic and M. Laurent, “The operator  $\Psi$  for the chromatic number of a graph,” *SIAM Journal on Optimization*, vol. 19, no. 2, pp. 572–591, 2008.
- [42] S. Yekhanin, “Locally decodable codes,” *Foundations and Trends in Theoretical Computer Science*, vol. 6, no. 3, pp. 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

UNIVERSITAT AUTÒNOMA DE BARCELONA, E-08193 BELLATERRA (BARCELONA), SPAIN  
*E-mail address:* giannicolascarpa@gmail.com