Bounds on Entanglement Assisted Source-channel Coding via the Lovász ϑ Number and its Variants

Toby Cubitt, Laura Mančinska, David Roberson, Simone Severini, Dan Stahlke, and Andreas Winter

***Abstract*—We study zero-error entanglement assisted source-channel coding (communication in the presence of side information). Adapting a technique of Beigi, we show that such coding requires existence of a set of vectors satisfying orthogonality conditions related to suitably defined graphs G and H. Such vectors exist if and only if $\vartheta(\overline{G}) \leq \vartheta(\overline{H})$ where ϑ represents the Lovász number. We also obtain similar inequalities for the related Schrijver ϑ^- and Szegedy ϑ^+ numbers.**

These inequalities reproduce several known bounds and also lead to new results. We provide a lower bound on the entanglement assisted cost rate. We show that the entanglement assisted independence number is bounded by the Schrijver number: $\alpha^*(G) \leq \vartheta^-(G)$. Therefore, we are able to disprove the conjecture that the one-shot entanglement-assisted zero-error capacity is equal to the integer part of the Lovász number. Beigi introduced a quantity β as an upper bound on α^* and posed the question of whether $\beta(G) = \lfloor \vartheta(G) \rfloor$. We answer this in the affirmative and show that a related quantity is equal to $\lceil \vartheta(G) \rceil$. We show that a quantity $\chi_{\text{vect}}(G)$ recently introduced in the context of Tsirelson's problem is equal to $\lceil \vartheta^+(\overline{G}) \rceil$.

In an appendix we investigate multiplicativity properties of Schrijver's and Szegedy's numbers, as well as projective rank.

***Index Terms*—Graph theory, Quantum entanglement, Quantum information, Zero-error information theory, Linear programming**

I. Introduction

The source-channel coding problem is as follows: Alice and Bob can communicate only through a noisy channel. Alice wishes to send a message to Bob, and Bob already has some side information regarding Alice's message. (Note that Alice's message may be several bits long.) Alice encodes her message and sends a transmission through the channel. Given the (noisy) channel output along with his side information, Bob must be able to deduce Alice's message with zero probability of error (we always require zero error throughout this entire paper). An (m, n)-*coding scheme* consists of encoding and decoding operations which allow sending m messages via n uses of the noisy channel (again, each of the m messages may be several bits long). The *cost rate* η is the infimum of n/m over all (m, n)-coding schemes.

There are two special cases which are particularly noteworthy. If the messages are bits and there is no side information then the inverse of the cost rate, $1/\eta$, is the *Shannon capacity* [1], the number of zero-error bits that can be transmitted per channel use in the limit of many uses of the channel. On the other hand, communication over a perfect channel with side information was considered by Witsenhausen [2]; the corresponding cost rate is known as the *Witsenhausen rate*. The general problem, with both side information and a noisy channel, was considered by Nayak, Tuncel, and Rose [3].

The Shannon capacity of a channel is very difficult to compute, and is not even known to be decidable. However, a useful upper bound on Shannon capacity is provided by the ϑ number introduced by Lovász [4]. The Lovász ϑ number also provides a lower bound on the Witsenhausen rate [3] and, in general, the cost rate.

Recently it has been of interest to study a version of this problem in which the parties may make use of an entangled quantum state, which can in certain cases increase the zero-error capacity of a classical channel [5], [6]. The Lovász ϑ number upper bounds entanglement assisted Shannon capacity, just as it does classical Shannon capacity [7], [8]. Beigi's proof [7] proceeds through a relaxation of the channel coding problem, with the relaxed constraints consisting of various orthogonality conditions imposed upon a set of vectors. We study a relaxation of the entanglement assisted source-channel coding problem inspired by this technique of Beigi. This relaxation leads to a set of constraints that are exactly characterized by monotonicity of ϑ. This has a number of consequences. Beigi defined a function β as an upper bound on entanglement assisted independence number and posed the question of whether β is equal to $\lfloor \vartheta \rfloor$. We answer this in the affirmative and show that a similarly defined quantity is equal to $\lceil \vartheta \rceil$. We show that ϑ provides a bound for the source-channel coding problem. As a special case this reproduces both Beigi's result as well as that of Briët et al. [9] in which it is shown that ϑ is a lower bound on the entanglement assisted Witsenhausen rate.

Toby Cubitt is with the Mathematics and Quantum Information group, Departamento de Análisis Matemático, Facultad de CC Matemáticas, Universidad Complutense de Madrid, 28040 Madrid, Spain and the Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Wilberforce Road, Cambridge, U.K. (e-mail: tsc25@cam.ac.uk).

Laura Mančinska is with the Institute for Quantum Computing, University of Waterloo, Canada and the Centre for Quantum Technologies, National University of Singapore, 21 Lower Kent Ridge Road, Singapore (e-mail: laura.mancinska@gmail.com).

David Roberson is with the Department of Combinatorics & Optimization, University of Waterloo, Canada and the Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, 50 Nanyang Avenue, Singapore (e-mail: davideroberson@gmail.com).

Simone Severini is with the Department of Computer Science, and Department of Physics and Astronomy, University College London, WC1E 6BT London, U.K. (e-mail: simoseve@gmail.com).

Dan Stahlke is with the Department of Physics, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA (e-mail: dan@stahlke.org)

Andreas Winter is with the ICREA & Física Teòrica: Informació i Fenomens Quàntics, Universitat Autònoma de Barcelona, ES-08193 Bellaterra (Barcelona), Spain. When the core part of the present research was done he was affiliated with the Department of Mathematics, University of Bristol, BS8 1TW Bristol, U.K. (e-mail: andreas.winter@uab.cat)

Some of the results in the present paper were presented at the Theory of Quantum Computation, Communication & Cryptography 2014 conference.

A slightly different relaxation of source-channel coding leads to three necessary conditions for the existence of a $(1, 1)$-coding scheme in terms of ϑ and two variants: Schrijver's ϑ^- and Szegedy's ϑ^+. This reproduces or strengthens results from [7], [9], [10] under a unified framework, with simpler proofs. In particular, we produce a tighter bound on the entanglement assisted independence number: $\alpha^* \leq \vartheta^-$.

The technical results, Theorems 6 and 10, should be accessible to the reader who is familiar with graph theory but not information theory or quantum mechanics, which merely provide a motivation for the problem.

II. Source-channel coding

We will make use of the following graph theoretical concepts. A *graph* G consists of a set of *vertices* $V(G)$ along with a symmetric binary relation $x \sim_G y$ among vertices (we abbreviate $x \sim y$ when the graph can be inferred from context). A pair of vertices (x, y) satisfying $x \sim y$ are said to be *adjacent*. Equivalently, it is said that there is an *edge* between x and y. Vertices are not adjacent to themselves, so $x \not\sim x$ for all $x \in V(G)$. The *complement* of a graph G, denoted $\overline{G}$, has the same set of vertices but has edges between distinct pairs of vertices which are not adjacent in G (i.e. for $x \neq y$ we have $x \sim_{\overline{G}} y \iff x \not\sim_G y$). A set of vertices no two of which form an edge is known as an *independent set*; the size of the largest independent set is the *independence number* $\alpha(G)$. A set of vertices such that every pair is adjacent is known as a *clique*; the size of the largest clique is the *clique number* $\omega(G)$. Clearly $\omega(G) = \alpha(\overline{G})$. An assignment of colors to vertices such that adjacent vertices are given distinct colors is called a *proper coloring*; the minimum number of colors needed is the *chromatic number* $\chi(G)$. A function mapping the vertices of one graph to those of another, $f : V(G) \to V(H)$, is a *homomorphism* if $x \sim_G y \implies f(x) \sim_H f(y)$. Since vertices are not adjacent to themselves it is necessary that $f(x) \neq f(y)$ when $x \sim y$. If such a function exists, we say that G *is homomorphic to* H and write $G \to H$. The *complete graph* on n vertices, denoted K_n, has an edge between every pair of vertices. It is not hard to see that $\omega(G)$ is equal to the largest n such that $K_n \to G$, and $\chi(G)$ is equal to the smallest n such that $G \to K_n$. Many other graph properties can be expressed in terms of homomorphisms; for details see [11], [12]. The *strong product* of two graphs, $G \boxtimes H$, has vertex set $V(G) \times V(H)$ and has edges

$$\begin{aligned}(x_1, y_1) \sim (x_2, y_2) \iff &(x_1 = x_2 \text{ and } y_1 \sim y_2) \text{ or} \\ &(x_1 \sim x_2 \text{ and } y_1 = y_2) \text{ or} \\ &(x_1 \sim x_2 \text{ and } y_1 \sim y_2).\end{aligned}$$

The n-fold strong product is written $G^{\boxtimes n} := G \boxtimes G \boxtimes \cdots \boxtimes G$. The *disjunctive product* $G * H$ has edges

$$(x_1, y_1) \sim (x_2, y_2) \iff x_1 \sim x_2 \text{ or } y_1 \sim y_2.$$

It is easy to see that $\overline{G * H} = \overline{G} \boxtimes \overline{H}$. The n-fold disjunctive product is written $G^{*n} := G * G * \ldots * G$.

Suppose that Alice communicates to Bob through a noisy classical channel $\mathcal{N} : S \to V$. She wishes to send a message to Bob with zero chance of error. Let $\mathcal{N}(v|s)$ denote the probability that $\mathcal{N}$ will produce symbol v when given symbol s as input, and define the graph H with vertices S and edges

$$s \sim_H t \iff \mathcal{N}(v|s)\mathcal{N}(v|t) = 0 \text{ for all } v \in V. \quad (1)$$

Bob can distinguish codewords s and t if and only if they have no chance of being mapped to the same output by $\mathcal{N}$. Therefore, Alice's set of codewords must form a clique of H; the size of the largest such set is the clique number $\omega(H)$. We will call this the *distinguishability graph* of $\mathcal{N}$. The complementary graph $\overline{H}$ is known as the *confusability graph* of $\mathcal{N}$. Note that standard convention is to denote the confusability graph by H rather than $\overline{H}$. We break convention in order to make notation in this paper much simpler. However, to minimize confusion when discussing prior results, we will follow the tradition of using the independence number when speaking of the number of codewords that Alice can send (equal to $\alpha(\overline{H}) = \omega(H)$ in our notation).

The number of bits (the base-2 log of the number of distinct codewords) that Alice can send with a single use of $\mathcal{N}$ is known as the *one-shot zero-error capacity* of $\mathcal{N}$, and is equal to $\log \alpha(\overline{H})$. The average number of bits that can be sent per channel use (again with zero error) in the limit of many uses of a channel is known as the *Shannon capacity*. With n parallel uses of $\mathcal{N}$ the distinguishability graph is H^{*n}. The Shannon capacity of $\mathcal{N}$ is therefore

$$\begin{aligned}\Theta(\overline{H}) :&= \lim_{n\to\infty} \frac{1}{n} \log \omega(H^{*n}) \\ &= \lim_{n\to\infty} \frac{1}{n} \log \alpha(\overline{H}^{\boxtimes n}).\end{aligned}$$

This quantity is in general very difficult to compute, with the capacity of the five cycle graph $\overline{H} = C_5$ having been open for over 20 years and the capacity of C_7 being unknown to this day. The capacity of C_5 was solved by Lovász [4] who introduced a function $\vartheta(\overline{H})$, the definition of which will be postponed until Section III. Lovász proved a sandwich theorem which, using the notation $\bar{\vartheta}(H) := \vartheta(\overline{H})$, takes the form

$$\alpha(\overline{H}) = \omega(H) \leq \bar{\vartheta}(H) \leq \chi(H).$$

He also showed that $\bar{\vartheta}(H^{*n}) = \bar{\vartheta}(H)^n$, therefore $\Theta(\overline{H}) \leq \log \bar{\vartheta}(H)$. This bound also applies to entanglement assisted communication [7], which we will investigate in detail, and has been generalized to quantum channels [8].

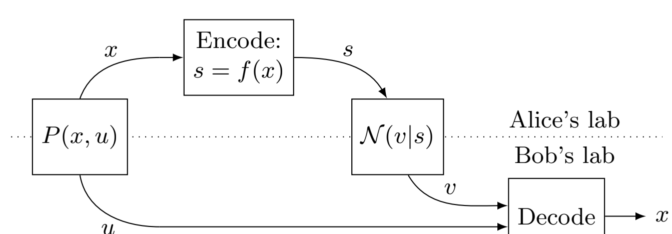

Fig. 1. A zero-error source-channel $(1, 1)$-coding scheme.

We now introduce the *source-channel coding* problem. As before, Alice wishes to send Bob a message $x \in X$, and she can only communicate through a noisy channel $\mathcal{N} : S \to V$. Now, however, Bob has some side information about Alice's message. Specifically, Alice and Bob each receive one part of a pair (x, u) drawn according to a probability distribution $P(x, u)$. This is known as a *dual source*. Alice encodes her input x using a function $f : X \to S$ and sends the result through $\mathcal{N}$. Bob must deduce x with zero chance of error using the output of $\mathcal{N}$ along with his side information u. Such a protocol is called a *zero-error source-channel* $(1, 1)$*-coding scheme*, and is depicted in Fig. 1. An (m, n)-coding scheme transmits m independent instances of the source using n copies of the channel.

Again the analysis involves graphs. Let H again be the distinguishability graph (1) and define the *characteristic graph* G with vertices X and edges

$$x \sim_G y \iff \exists u \in U \text{ such that } P(x, u)P(y, u) \neq 0.$$

In [3] it was shown that decoding is possible if and only if Alice's encoding f is a homomorphism from G to H.[1] Therefore a zero-error $(1, 1)$-coding scheme exists if and only if $G \to H$. A zero-error (m, n)-coding scheme is possible if and only if

$$G^{\boxtimes m} \to H^{*n}. \tag{2}$$

The smallest possible ratio n/m (in the limit $m \to \infty$) is called the *cost rate*, $\eta(G, \overline{H})$. More precisely, the cost rate is defined as

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

The $\bar{\vartheta}$ quantity is monotone under graph homomorphisms in the sense that $G \to H \implies \bar{\vartheta}(G) \leq \bar{\vartheta}(H)$ [13]. Consequently, a zero-error $(1, 1)$-coding scheme requires $\bar{\vartheta}(G) \leq \bar{\vartheta}(H)$. Since $\bar{\vartheta}(G^{\boxtimes m}) = \bar{\vartheta}(G)^m$ [14] and $\bar{\vartheta}(H^{*n}) = \bar{\vartheta}(H)^n$ [4], it follows that an (m, n)-coding scheme is possible only if $\log \bar{\vartheta}(G)/ \log \bar{\vartheta}(H) \leq n/m$. Thus we have the bound

$$\eta(G, \overline{H}) \geq \frac{\log \bar{\vartheta}(G)}{\log \bar{\vartheta}(H)}.$$

(Cf. [3] for the special case of the Witsenhausen rate.)

We will return to this in Section III when we prove an analogous bound for entanglement assisted zero-error source-channel coding.

When Bob has no side information (equivalently, when U is a singleton), G is the complete graph. In this case zero-error transmission of x is possible if and only if $K_n \to H$ where $n = |X|$, which in turn holds if and only if $n \leq \omega(H) = \alpha(\overline{H})$. This is the expected result, since as mentioned before $\alpha(\overline{H})$ is the number of unambiguously decodable codewords that Alice can send through $\mathcal{N}$. On the other hand, consider the case where there is side information and $\mathcal{N}$ is a noiseless channel of size $n = |S|$. Now H becomes the complete graph K_n, so x can be perfectly transmitted if and only if $G \to K_n$. This holds if and only if $n \geq \chi(G)$. These two examples provide an operational interpretation to the independence number and chromatic number of a graph. The analogous communication problems in the presence of an entangled state (which we examine shortly) define the entanglement assisted independence and chromatic numbers.

[1] Basically, G represents the information that needs to be sent and H represents the information that survives the channel. A homomorphism $G \to H$ ensures that the needed information makes it through the channel intact.

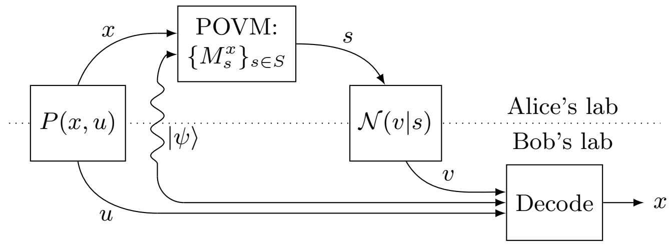

Fig. 2. An entanglement assisted zero-error source-channel $(1, 1)$-coding scheme.

If Alice and Bob share an entangled state they can use the strategy depicted in Fig. 2, which is described in greater detail in [9]. Alice, upon receiving $x \in X$, performs a POVM $\{M_s^x\}_{s\in S}$ on her half of the entanglement resource $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ and receives measurement outcome $s \in S$. Without loss of generality this can be assumed to be a projective measurement since any POVM can be converted to a projective measurement by enlarging the entangled state. So for each $x \in X$, the collection $\{M_s^x\}_{s\in S}$ consists of projectors on $\mathcal{H}_A$ which sum to the identity. Alice sends the measurement outcome s through the channel $\mathcal{N}$ to Bob, who receives some $v \in V$ such that $\mathcal{N}(v|s) > 0$. Bob then measures his half of the entangled state using a projective measurement depending on v and his side information u. An *entanglement assisted zero-error* $(1, 1)$*-coding scheme* is one in which Bob is able to determine x with zero chance of error; an *entanglement assisted zero-error* (m, n)*-coding scheme* involves sending m independent samples of the source using n copies of the channel.

After Alice's measurement, Bob's half of the entanglement resource is in the state

$$\rho_s^x = \mathrm{Tr}_A\{(M_s^x \otimes I)|\psi\rangle \langle\psi|\}.$$

An error free decoding operation exists for Bob if and only if these states are orthogonal for every $x \in X$ consistent with the information in Bob's possession (i.e. u and v). We then have the following necessary and sufficient condition [9]. Let G be the characteristic graph of the source and H be the distinguishability graph of the channel. There must be a bipartite pure state $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ for some Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, and for each $x \in X$ there must be a projective decomposition of the identity $\{M_s^x\}_{s\in S}$ on $\mathcal{H}_A$ such that

$$\rho_s^x \perp \rho_t^y \text{ for all } x \sim_G y \text{ and } s \not\sim_H t,$$

where orthogonality is in terms of the Hilbert–Schmidt inner product.

Recall that without entanglement a zero-error $(1, 1)$-coding scheme was possible if and only if $G \to H$. By analogy we say there is an *entanglement assisted homomorphism* $G \xrightarrow{*} H$ when there exists an entanglement assisted zero-error $(1, 1)$-coding scheme:

Definition 1. *Let G and H be graphs. There is an entanglement assisted homomorphism from G to H, written $G \xrightarrow{*} H$,* *if there is a bipartite state $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ (for some Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$) and, for each $x \in V(G)$, a projective decomposition of the identity $\{M_s^x\}_{s\in V(H)}$ on $\mathcal{H}_A$ such that*

$$\rho_s^x \perp \rho_t^y \text{ for all } x \sim_G y \text{ and } s \not\simeq_H t, \tag{4}$$

where

$$\rho_s^x := \mathrm{Tr}_A\{(M_s^x \otimes I)|\psi\rangle\langle\psi|\}. \tag{5}$$

Analogous to (2), there is an entanglement assisted (m,n)-coding scheme if and only if $G^{\boxtimes m} \xrightarrow{*} H^{*n}$. The entangled cost rate [9] is analogous to (3),

$$\eta^*(G,\overline{H}) = \lim_{m\to\infty} \frac{1}{m} \min\left\{n : G^{\boxtimes m} \xrightarrow{*} H^{*n}\right\}. \tag{6}$$

In the absence of side information (i.e. with U being a singleton set), G becomes the complete graph. We saw above that without entanglement and without side information, n distinct codewords can be sent error-free through a noisy channel if and only if $K_n \to H$; the largest such n is $\omega(H) = \alpha(\overline{H})$. With the help of entanglement the largest number of codewords is the largest n such that $K_n \xrightarrow{*} H$; this defines the *entanglement assisted independence number*, $\alpha^*(\overline{H})$. Since an entanglement resource never hurts, $\alpha^*(\overline{H}) \geq \alpha(\overline{H})$ always. In some cases $\alpha^*(\overline{H})$ can be strictly larger than $\alpha(\overline{H})$ [5].

We saw above that $\alpha(\overline{H}) \leq \bar{\vartheta}(H)$. Indeed, this was the original application of $\bar{\vartheta}$. Beigi showed that also $\alpha^*(\overline{H}) \leq \bar{\vartheta}(H)$ [7] (this has been generalized to quantum channels as well [8]; however, we consider here only classical channels). Beigi proved his bound by showing that if n distinct codewords can be sent through a noisy channel with zero-error using entanglement ($K_n \xrightarrow{*} H$ in our notation) then there are vectors $|w\rangle \neq 0$ and $|w_s^x\rangle$ with $x \in \{1,\ldots,n\}$ and $s \in V(H)$ such that[2]

$$\sum_s |w_s^x\rangle = |w\rangle \tag{7}$$

$$\langle w_s^x|w_t^y\rangle = 0 \text{ for all } x \neq y, s \not\simeq_H t \tag{8}$$

$$\langle w_s^x|w_t^x\rangle = 0 \text{ for all } s \neq t. \tag{9}$$

Denote by $\beta(\overline{H})$ the largest n such that vectors of this form exist. Then $\beta(\overline{H}) \geq \alpha^*(\overline{H})$. Beigi showed that the existence of such vectors implies $n \leq \bar{\vartheta}(H)$, therefore $\alpha^*(\overline{H}) \leq \beta(\overline{H}) \leq \bar{\vartheta}(H)$. Since ϑ is multiplicative under the strong graph product, $\bar{\vartheta}(H)$ is in fact an upper bound on the entanglement assisted Shannon capacity. Beigi left open the question of whether $\beta(\overline{H})$ was equal to $\lfloor\bar{\vartheta}(H)\rfloor$. We will answer this question in the affirmative (Corollary 7).

In fact, we show something more general. We generalize Beigi's vectors so that they apply to the source-channel coding problem (i.e. with G not necessarily being K_n) and give a bound in terms of the Lovász ϑ number. The conditions we will introduce can be thought of as a relaxation of the condition (4), which defines $G \xrightarrow{*} H$. A related but different relaxation will give bounds in terms of two variations of the Lovász number: the Schrijver number [15], [16] and the Szegedy number [17]. We denote the first relaxation $G \xrightarrow{B} H$ since it generalizes Beigi's condition, and denote the second $G \xrightarrow{+} H$ since it contains a positivity condition. A third relaxation, $G \xrightarrow{V} H$, is defined here but the significance is discussed later.

[2] Recall that we take $\overline{H}$ to be the confusability graph rather than H. So Beigi's definition is worded differently.

Definition 2. *Let G and H be graphs. Write $G \xrightarrow{B} H$ if there are vectors $|w\rangle \neq 0$ and $|w_s^x\rangle \in \mathbb{C}^d$ for each $x \in V(G)$, $s \in V(H)$, for some $d \in \mathbb{N}$, such that*

1) $\sum_s |w_s^x\rangle = |w\rangle$
2) $\langle w_s^x|w_t^y\rangle = 0$ *for all* $x \sim_G y$, $s \not\simeq_H t$
3) $\langle w_s^x|w_t^x\rangle = 0$ *for all* $s \neq t$.

Write $G \xrightarrow{+} H$ if there are vectors satisfying

1) $\sum_s |w_s^x\rangle = |w\rangle$
2) $\langle w_s^x|w_t^y\rangle = 0$ *for all* $x \sim_G y$, $s \not\simeq_H t$
3) $\langle w_s^x|w_t^y\rangle \geq 0$.

Write $G \xrightarrow{V} H$ if there are vectors satisfying the conditions for both $G \xrightarrow{B} H$ and $G \xrightarrow{+} H$, i.e.,

1) $\sum_s |w_s^x\rangle = |w\rangle$
2) $\langle w_s^x|w_t^y\rangle = 0$ *for all* $x \sim_G y$, $s \not\simeq_H t$
3) $\langle w_s^x|w_t^x\rangle = 0$ *for all* $s \neq t$
4) $\langle w_s^x|w_t^y\rangle \geq 0$.

Without loss of generality one could consider only real vectors, since complex vectors can be turned real via the recipe $|\hat{w}_s^x\rangle = \mathrm{Re}(|w_s^x\rangle) \oplus \mathrm{Im}(|w_s^x\rangle)$ while preserving the inner product properties required by the above definitions.

It is enlightening to consider the Gram matrices of the $|w_s^x\rangle$ vectors. In fact, it is this formulation that will be used to prove our main theorems.

Theorem 3. *$G \xrightarrow{B} H$ if and only if there is a positive semidefinite matrix $C : \mathcal{L}(\mathbb{C}^{|V(G)|}) \otimes \mathcal{L}(\mathbb{C}^{|V(H)|})$ satisfying*

$$\sum_{s,t} C_{xyst} = 1 \tag{10}$$

$$C_{xyst} = 0 \textit{ for } x \sim_G y \textit{ and } s \not\simeq_H t \tag{11}$$

$$C_{xxst} = 0 \textit{ for } s \neq t. \tag{12}$$

$G \xrightarrow{+} H$ if and only if there is a positive semidefinite matrix satisfying (10), (11), and

$$C_{xyst} \geq 0. \tag{13}$$

$G \xrightarrow{V} H$ if and only if there is a positive semidefinite matrix satisfying (10)-(13).

Proof: We prove only the $G \xrightarrow{B} H$ claim; the proofs for $G \xrightarrow{+} H$ and $G \xrightarrow{V} H$ are analogous. Suppose $G \xrightarrow{B} H$ and let $|w\rangle$ and $|w_s^x\rangle$ be the vectors described in Definition 2. Without loss of generality rescale so that $\langle w|w\rangle = 1$. Define the matrix $C : \mathcal{L}(\mathbb{C}^{|V(G)|}) \otimes \mathcal{L}(\mathbb{C}^{|V(H)|})$ with entries $C_{xyst} = \langle w_s^x|w_t^y\rangle$ for $x,y \in V(G)$ and $s,t \in V(H)$. Since C is a Gram matrix, it is positive semidefinite. Properties (10)-(12) follow directly from the three properties listed in Definition 2 for $G \xrightarrow{B} H$ (the first of these uses $\langle w|w\rangle = 1$).

For the converse, note that any positive semidefinite matrix $C : \mathcal{L}(\mathbb{C}^{|V(G)|}) \otimes \mathcal{L}(\mathbb{C}^{|V(H)|})$ is a Gram matrix of some vectors $|w_s^x\rangle$. The three properties of Definition 2 for $G \xrightarrow{B} H$ follow from (10)-(12). Only the first of these is nontrivial. We have that for all $x, y \in V(G)$,

$$1 = \sum_{st} C_{xyst} = \sum_{st} \langle w_s^x | w_t^y \rangle = \left(\sum_s \langle w_s^x | \right) \left(\sum_t | w_t^y \rangle \right).$$

For $x = y$, the above implies that $\sum_s |w_s^x\rangle$ is a unit vector for all x. These unit vectors must have unit inner product amongst themselves by the $x \neq y$ cases above, and therefore they must all be the same vector. Call this $|w\rangle$. Clearly $|w\rangle \neq 0$. ■

It is interesting to note that a matrix with properties (10), (11), and (13) (those associated with $G \overset{+}{\to} H$) can be interpreted as a conditional probability distribution, $P(s, t|x, y) = C_{xyst}$. With this interpretation, $G \overset{+}{\to} H$ if and only if there exists a conditional probability distribution $P(s, t|x, y)$ such that $C_{sx;ty} = P(s, t|x, y)$ is a positive semidefinite matrix and $P(s \sim_H t|x \sim_G y) = 1$. A similar interpretation holds for $G \overset{V}{\to} H$.

We now show that $G \overset{B}{\to} H$ and $G \overset{+}{\to} H$ are indeed relaxations of $G \overset{*}{\to} H$ (the significance of $G \overset{V}{\to} H$ will be explained later). Since Definition 2 reduces to Beigi's criteria when considering $K_n \overset{B}{\to} H$, the argument that follows provides an alternative and simpler proof of Beigi's result that $\alpha^*(\overline{H}) \leq \beta(\overline{H})$.

Theorem 4. *If $G \overset{*}{\to} H$ then $G \overset{B}{\to} H$ and $G \overset{+}{\to} H$.*

Proof: ($G \overset{*}{\to} H \implies G \overset{B}{\to} H$): Suppose that $G \overset{*}{\to} H$. Let $|\psi\rangle$ and M_s^x for $x \in V(G)$ and $s \in V(H)$ satisfy condition (4) (with ρ_s^x given by (5)). Define $|w\rangle = |\psi\rangle$ and

$$|w_s^x\rangle = (M_s^x \otimes I) |\psi\rangle .$$

Since $\{M_s^x\}_{s \in S}$ is a projective decomposition of the identity,

$$\sum_s M_s^x = I \implies \sum_s |w_s^x\rangle = |w\rangle ,$$
$$M_s^x M_t^x = 0 \implies \langle w_s^x | w_t^x \rangle = 0 \text{ for } s \neq t.$$

For all $x \sim_G y$ and $s \not\sim_H t$, condition (4) gives that the reduced density operators (tracing over $\mathcal{H}_A$) of the post-measurement states $(M_s^x \otimes I) |\psi\rangle$ and $(M_t^y \otimes I) |\psi\rangle$ are orthogonal. But this is only possible if the pure states (without tracing out $\mathcal{H}_A$) are orthogonal. So,

$$\langle \psi | (M_s^x \otimes I)^\dagger (M_t^y \otimes I) | \psi \rangle = 0 \implies \langle w_s^x | w_t^y \rangle = 0.$$

($G \overset{*}{\to} H \implies G \overset{+}{\to} H$): Suppose that $G \overset{*}{\to} H$. Let $|\psi\rangle$ and M_s^x for $x \in V(G)$ and $s \in V(H)$ satisfy condition (4) (with ρ_s^x given by (5)). Define $|w_s^x\rangle$ to be the vectorization of the post-measurement reduced density operator,

$$|w_s^x\rangle = \mathrm{vec}(\rho_s^x).$$

Since $\{M_s^x\}_{s \in S}$ sum to identity,

$$\sum_s |w_s^x\rangle = \mathrm{vec}\left(\sum_s \mathrm{Tr}_A\{(M_s^x \otimes I) |\psi\rangle \langle \psi|\} \right) = \mathrm{vec}(\mathrm{Tr}_A\{|\psi\rangle \langle \psi|\}) =: |w\rangle .$$

For all $x \sim_G y$ and $s \not\sim_H t$, condition (4) gives $\langle w_s^x | w_t^y \rangle = 0$. Density operators are positive, giving positive inner products $\langle w_s^x | w_t^y \rangle \geq 0$. ■

III. Monotonicity theorems

Our main results concern monotonicity properties of the Lovász number ϑ, Schrijver number ϑ^-, and Szegedy number ϑ^+ for graphs that are related by the generalized homomorphisms of Definition 2. These will lead to various bounds relevant to entanglement assisted zero-error source-channel coding. These three quantities are defined as follows.

Definition 5. *In this definition we use real matrices. For convenience we state the definitions in terms of the complement of a graph, since this form is used throughout the theorems.*

The Lovász number of the complement, $\bar{\vartheta}(G) := \vartheta(\overline{G})$, is given by either of the following two semidefinite programs, which are equivalent [4], [14], [18]:

$$\bar{\vartheta}(G) = \max\{\|I + T\| : I + T \succeq 0, \; T_{ij} = 0 \text{ for } i \not\sim j\}, \tag{14}$$
$$\bar{\vartheta}(G) = \min\{\lambda : \exists Z \succeq 0, Z_{ii} = \lambda - 1, \; Z_{ij} = -1 \text{ for } i \sim j\}, \tag{15}$$

where $\|\cdot\|$ denotes the operator norm (the largest singular value) and $\succeq 0$ means that a matrix is positive semidefinite. The Schrijver number of the complement, $\bar{\vartheta}^-(G) := \vartheta^-(\overline{G})$, (sometimes written ϑ') is [15], [16]

$$\bar{\vartheta}^-(G) = \min\{\lambda : \exists Z \succeq 0, Z_{ii} = \lambda - 1, \; Z_{ij} \leq -1 \text{ for } i \sim j\}. \tag{16}$$

The Szegedy number of the complement, $\bar{\vartheta}^+(G) := \vartheta^+(\overline{G})$, is [17]

$$\bar{\vartheta}^+(G) = \min\{\lambda : \exists Z \succeq 0, Z_{ii} = \lambda - 1, \; Z_{ij} = -1 \text{ for } i \sim j, \; Z_{ij} \geq -1 \text{ for all } i, j\}. \tag{17}$$

Clearly $\bar{\vartheta}^-(G) \leq \bar{\vartheta}(G) \leq \bar{\vartheta}^+(G)$.

Our first result is that $G \overset{B}{\to} H$ exactly characterizes ordering of $\bar{\vartheta}$. This will lead to a bound on entanglement assisted cost rate.

Theorem 6. $G \overset{B}{\to} H \iff \bar{\vartheta}(G) \leq \bar{\vartheta}(H)$.

Proof: ($\Longleftarrow$): Suppose $\bar{\vartheta}(G) \leq \bar{\vartheta}(H)$. We will explicitly construct a matrix $C \succeq 0$ satisfying properties (10)-(12) of Theorem 3. Let $\lambda = \bar{\vartheta}(H)$. By definition, there is a matrix T such that $\|I + T\| = \lambda$, $I + T \succeq 0$, and $T_{st} = 0$ for $s \not\sim t$. With $|\psi\rangle$ denoting the vector corresponding to the largest eigenvalue of $I + T$, and with $\circ$ denoting the Schur–Hadamard (i.e. entrywise) product, define the matrices

$$D = |\psi\rangle \langle \psi| \circ I,$$
$$B = |\psi\rangle \langle \psi| \circ (I + T).$$

With J being the all-ones matrix and $\langle \cdot, \cdot \rangle$ denoting the Hilbert–Schmidt inner product, it is readily verified that

$$\langle D, J \rangle = \langle \psi | \psi \rangle = 1,$$
$$\langle B, J \rangle = \langle \psi | I + T | \psi \rangle = \lambda.$$

Since the Schur–Hadamard product of two matrices is a principal submatrix of their tensor product, this operation preserves positive semidefiniteness. As a consequence, $B \succeq 0$ and

$$\|I+T\| = \lambda \implies \lambda I - (I+T) \succeq 0 \implies \lambda D - B \succeq 0.$$

Since $\lambda \geq \bar{\vartheta}(G)$, there is a matrix Z such that $Z \succeq 0$, $Z_{xx} = \lambda - 1$ for all x, and $Z_{xy} = -1$ for all $x \sim_G y$. Note that Definition 5 gives existence of a matrix with $\bar{\vartheta}(G) - 1$ on the diagonal, but since $\lambda \geq \bar{\vartheta}(G)$ we can add a multiple of the identity to get $\lambda - 1$ on the diagonal.

We now construct C. Define

$$C = \lambda^{-1}\left[J \otimes B + (\lambda-1)^{-1} Z \otimes (\lambda D - B)\right].$$

Since J, B, Z, and $\lambda D - B$ are all positive semidefinite, and $\lambda - 1 \geq 0$, we have that C is positive semidefinite. The other desired conditions on C are easy to verify. For all x, y we have

$$\sum_{st} C_{xyst} = \lambda^{-1}\left[\langle B, J\rangle + (\lambda-1)^{-1} Z_{xy}[\lambda\langle D, J\rangle - \langle B, J\rangle]\right] = 1.$$

Note that the J in the above equation is indexed by $V(H)$, whereas the J in the definition of C is indexed by $V(G)$. For $x \sim_G y$ and $s \not\sim_H t$,

$$\begin{aligned} C_{xyst} &= \lambda^{-1}\left[B_{st} + (\lambda-1)^{-1} Z_{xy}(\lambda D_{st} - B_{st})\right] \\ &= \lambda^{-1}\left[B_{st} + (\lambda-1)^{-1}(-1)(\lambda B_{st} - B_{st})\right] = 0. \end{aligned}$$

For all x and for $s \neq t$,

$$\begin{aligned} C_{xxst} &= \lambda^{-1}\left[B_{st} + (\lambda-1)^{-1} Z_{xx}(\lambda D_{st} - B_{st})\right] \\ &= \lambda^{-1}\left[B_{st} + (0 - B_{st})\right] = 0. \end{aligned}$$

($\Longrightarrow$): Suppose $G \xrightarrow{B} H$. By Theorem 3, there is a matrix $C \succeq 0$ satisfying properties (10)-(12). Let Z achieve the optimal value (call it λ) for the minimization (15) for $\bar{\vartheta}(H)$. We will provide a feasible solution for (15) for $\bar{\vartheta}(G)$ to show that $\bar{\vartheta}(G) \leq \lambda = \bar{\vartheta}(H)$. To this end, let $|\mathbf{1}\rangle$ be the all ones vector and define

$$Y = \big(I \otimes \langle\mathbf{1}|\big)\big[\big(J \otimes Z\big) \circ C\big]\big(I \otimes |\mathbf{1}\rangle\big).$$

Since $C \succeq 0$ and $Z \succeq 0$, and positive semidefiniteness is preserved by conjugation, we have that $Y \succeq 0$. Also note that

$$Y_{xy} = \sum_{st} Z_{st} C_{xyst}.$$

Using the fact that $Z_{ss} = \lambda - 1$ and $C_{xxst} = 0$ for $s \neq t$, we have

$$Y_{xx} = \sum_{st} Z_{st} C_{xxst} = (\lambda - 1)\sum_{st} C_{xxst} = \lambda - 1.$$

Using the fact that $Z_{st} = -1$ for $s \sim_H t$ and $C_{xyst} = 0$ for $x \sim_G y$, $s \not\sim_H t$, we have that for $x \sim_G y$,

$$\begin{aligned} Y_{xy} &= \sum_{st} Z_{st} C_{xyst} = \sum_{s \sim_H t} Z_{st} C_{xyst} = (-1)\sum_{s \sim_H t} C_{xyst} \\ &= (-1)\sum_{st} C_{xyst} = -1. \end{aligned}$$

Now define a matrix Y' consisting of the real part of Y (i.e. with coefficients $Y'_{xy} = \mathrm{Re}[Y_{xy}]$). This matrix is real, positive semidefinite,[3] and satisfies $Y_{xx} = \lambda - 1$ for all x and $Y_{xy} = -1$ for $x \sim y$. Therefore Y' is feasible for (15) with value $\lambda = \bar{\vartheta}(H)$. Since $\bar{\vartheta}(G)$ is the minimum possible value of (15), we have $\bar{\vartheta}(G) \leq \bar{\vartheta}(H)$. ■

We are now prepared to answer in the affirmative an open question posed by Beigi [7].

Corollary 7. *Let $\beta(\overline{H})$ be the largest n such that there exist vectors $|w\rangle \neq 0$ and $|w_s^x\rangle$ with $x \in \{1, \ldots, n\}$ and $s \in V(H)$ which satisfy conditions (7)-(9). Then $\beta(\overline{H}) = \lfloor\bar{\vartheta}(H)\rfloor$.*

Proof: Considering $K_n \xrightarrow{B} H$, the conditions of Definition 2 are equivalent to (7)-(9). Since $\bar{\vartheta}(K_n) = n$, Theorem 6 gives $K_n \xrightarrow{B} H \iff n \leq \bar{\vartheta}(H)$. Since $\beta(\overline{H})$ is the largest n such that $K_n \xrightarrow{B} H$, we have that $\beta(\overline{H}) = \lfloor\bar{\vartheta}(H)\rfloor$. ■

A related corollary can be formed by considering $G \xrightarrow{B} K_n$ rather than $K_n \xrightarrow{B} H$. This defines a set of vectors $|w_s^x\rangle$ satisfying conditions in some sense complementary to Beigi's (7)-(9). Now we approach ϑ from above:

Corollary 8. *Let $\beta_\chi(G)$ be the smallest n such that there exist vectors $|w\rangle \neq 0$ and $|w_s^x\rangle$ with $x \in V(G)$ and $s \in \{1, \ldots, n\}$ for which*

1) $\sum_s |w_s^x\rangle = |w\rangle$
2) $\langle w_s^x | w_s^y\rangle = 0$ *for all* $x \sim_G y$
3) $\langle w_s^x | w_t^x\rangle = 0$ *for all* $s \neq t$.

Then $\beta_\chi(G) = \lceil\bar{\vartheta}(G)\rceil$.

Proof: Considering $G \xrightarrow{B} K_n$, the conditions of Definition 2 are equivalent to the conditions stated above. Since $\bar{\vartheta}(K_n) = n$, Theorem 6 gives $G \xrightarrow{B} K_n \iff \bar{\vartheta}(G) \leq n$. Since $\beta_\chi(G)$ is the smallest n such that $G \xrightarrow{B} K_n$, we have $\beta_\chi(G) = \lceil\bar{\vartheta}(G)\rceil$. ■

Corollary 9. *The entanglement assisted cost rate is bounded as follows:*

$$\eta^*(G, \overline{H}) \geq \frac{\log\bar{\vartheta}(G)}{\log\bar{\vartheta}(H)}.$$

Proof: Since $\bar{\vartheta}(G^{\boxtimes m}) = \bar{\vartheta}(G)^m$ [14] and $\bar{\vartheta}(H^{*n}) = \bar{\vartheta}(H)^n$ [4], it follows that

$$\begin{aligned} G^{\boxtimes m} \xrightarrow{*} H^{*n} &\implies G^{\boxtimes m} \xrightarrow{B} H^{*n} && \text{(by Theorem 4)} \\ &\implies \bar{\vartheta}(G^{\boxtimes m}) \leq \bar{\vartheta}(H^{*n}) && \text{(by Theorem 6)} \\ &\implies \bar{\vartheta}(G)^m \leq \bar{\vartheta}(H)^n \\ &\implies \frac{\log\bar{\vartheta}(G)}{\log\bar{\vartheta}(H)} \leq \frac{n}{m}. \end{aligned}$$

Therefore,

$$\eta^*(G, \overline{H}) = \lim_{m\to\infty} \min_n \left\{\frac{n}{m} : G^{\boxtimes m} \xrightarrow{*} H^{*n}\right\} \geq \frac{\log\bar{\vartheta}(G)}{\log\bar{\vartheta}(H)}.$$

■

Something similar to Theorem 6 holds for the relation $G \xrightarrow{+} H$. In this case there is an inequality not just for the

[3] The entrywise complex conjugate of a positive semidefinite matrix is positive semidefinite, so $Y' = (Y + \mathrm{conj}(Y))/2 \succeq 0$.

Lovász ϑ number but also for Schrijver's ϑ^- and Szegedy's ϑ^+. Unfortunately, this will no longer be an if-and-only-if statement (but see Theorem 13 for a weakened converse, and Appendix B for a somewhat more complicated if-and-only-if involving $\bar{\vartheta}^-$).

Theorem 10. *Suppose $G \xrightarrow{+} H$. Then $\bar{\vartheta}(G) \leq \bar{\vartheta}(H)$, $\bar{\vartheta}^-(G) \leq \bar{\vartheta}^-(H)$, and $\bar{\vartheta}^+(G) \leq \bar{\vartheta}^+(H)$.*

Proof: As per Theorem 3, let C be a positive semidefinite matrix satisfying properties (10), (11), and (13). We give the proof for $\bar{\vartheta}^-(G) \leq \bar{\vartheta}^-(H)$; the others are proved in a similar way. The proof is very similar to that of Theorem 6, with slight modification due to the fact that the last condition on C is different. Let Z achieve the optimal value for the minimization program (16) for $\bar{\vartheta}^-(H)$. We will provide a feasible solution for (16) for $\bar{\vartheta}^-(G)$ to show that $\bar{\vartheta}^-(G) \leq \bar{\vartheta}^-(H)$. Specifically, let $Y_{xy} = \sum_{st} Z_{st} C_{xyst}$. Since C and Z are positive semidefinite, so is Y.

A feasible solution for (16), with value $\bar{\vartheta}^-(H)$, requires $Y_{xx} = \bar{\vartheta}^-(H) - 1$. However, it suffices to show $Y_{xx} \leq \bar{\vartheta}^-(H) - 1$ since equality can be achieved by adding a non-negative diagonal matrix to Y. We have

$$\begin{aligned} Y_{xx} &= \sum_{st} Z_{st} C_{xxst} \\ &\leq \max_{st} |Z_{st}| \sum_{st} C_{xxst} \qquad \text{(since } C_{xxst} \geq 0) \\ &\leq \max_{s} |Z_{ss}| \sum_{st} C_{xxst} \qquad \text{(since } Z \succeq 0) \\ &= \bar{\vartheta}^-(H) - 1. \end{aligned}$$

Similarly, for $x \sim_G y$ we have

$$\begin{aligned} Y_{xy} &= \sum_{st} Z_{st} C_{xyst} = \sum_{s \sim_H t} Z_{st} C_{xyst} \\ &\leq (-1) \sum_{s \sim_H t} C_{xyst} = (-1) \sum_{st} C_{xyst} = -1. \end{aligned} \tag{18}$$

Therefore Y is feasible for (16) with value $\lambda = \bar{\vartheta}^-(H)$. Since $\bar{\vartheta}^-(G)$ is the minimum possible value of (16), we have $\bar{\vartheta}^-(G) \leq \bar{\vartheta}^-(H)$.

To show $\bar{\vartheta}(G) \leq \bar{\vartheta}(H)$ or $\bar{\vartheta}^+(G) \leq \bar{\vartheta}^+(H)$, replace inequality with equality in (18). For $\bar{\vartheta}^+(G) \leq \bar{\vartheta}^+(H)$ we have $Z_{st} \geq -1$ for all s, t and need to show $Y_{xy} \geq -1$ for all x, y. This is readily verified:

$$Y_{xy} = \sum_{st} Z_{st} C_{xyst} \geq (-1) \sum_{st} C_{xyst} = -1.$$

■

It is well known that $\alpha(G) \leq \vartheta^-(G) \leq \vartheta(G) \leq \vartheta^+(G) \leq \chi(\overline{G})$. We show that similar inequalities hold for the entanglement assisted independence and chromatic numbers. Since ϑ^- and ϑ^+ are not multiplicative under the required graph products (Appendix A), these do not lead to bounds on asymptotic quantities such as entanglement assisted Shannon capacity or entanglement assisted cost rate.

Corollary 11. $\alpha^*(H) \leq \vartheta^-(H)$.

Proof: By definition $\alpha^*(H)$ is the largest n such that $K_n \xrightarrow{*} \overline{H}$. But $K_n \xrightarrow{*} \overline{H} \implies K_n \xrightarrow{+} \overline{H} \implies \bar{\vartheta}^-(K_n) \leq \vartheta^-(H)$. Since $\bar{\vartheta}^-(K_n) = n$, the conclusion follows. ■

The following corollary was already shown in [9] via a different method.

Corollary 12. *Define $\chi^*(G)$ to be the smallest n such that $G \xrightarrow{*} K_n$. Then $\chi^*(G) \geq \bar{\vartheta}^+(G)$.*

Proof: By definition, $\chi^*(G)$ is the smallest n such that $G \xrightarrow{*} K_n$. But $G \xrightarrow{*} K_n \implies G \xrightarrow{+} K_n \implies \bar{\vartheta}^+(G) \leq \bar{\vartheta}^+(K_n) = n$. ■

It would be nice to have a converse to Theorem 10, like there was with Theorem 6. Is it the case that $\bar{\vartheta}(G) \leq \bar{\vartheta}(H)$, $\bar{\vartheta}^-(G) \leq \bar{\vartheta}^-(H)$, and $\bar{\vartheta}^+(G) \leq \bar{\vartheta}^+(H)$ together imply $G \xrightarrow{+} H$? We do not know. However, it is the case that $\bar{\vartheta}^+(G) \leq \bar{\vartheta}^-(H) \implies G \xrightarrow{+} H$. In fact, something stronger can be said. We have the following theorem, the consequences of which will be further explored in Section IV.

Theorem 13. $\bar{\vartheta}^+(G) \leq \bar{\vartheta}^-(H) \implies G \xrightarrow{V} H$.

Proof: The proof mirrors that of the ($\Longleftarrow$) portion of Theorem 6, so we only describe the differences. Let $\lambda = \bar{\vartheta}^-(H)$. Theorem 29 in Appendix B gives that

$$\begin{aligned} \bar{\vartheta}^-(H) = \max\{\|I + T\| : \; & I + T \succeq 0, \\ & T_{st} = 0 \text{ for } s \not\sim t, \\ & T_{st} \geq 0 \text{ for all } s, t\}. \end{aligned}$$

So there is a matrix T such that $\|I + T\| = \lambda$, $I + T \succeq 0$, $T_{st} = 0$ for $s \not\sim t$, and $T_{st} \geq 0$ for all s, t. Since $\lambda \geq \bar{\vartheta}^+(G)$, there is a matrix Z such that $Z \succeq 0$, $Z_{xx} = \lambda - 1$ for all x, $Z_{xy} = -1$ for all $x \sim_G y$, and $Z_{xy} \geq -1$ for all x, y. Note that T and Z satisfy all conditions required by Theorem 6 plus the additional conditions $T_{st} \geq 0$ for all s, t and $Z_{xy} \geq -1$ for all x, y.

Define B and D as in Theorem 6. The eigenvector $|\psi\rangle$ corresponding to the maximum eigenvalue of $I + T$ can be chosen to be entrywise non-negative (this follows from the Perron–Frobenius theorem and the fact that $I+T$ is entrywise non-negative). It follows that B can be chosen entrywise non-negative. As before, define

$$C = \lambda^{-1} \left[J \otimes B + (\lambda - 1)^{-1} Z \otimes (\lambda D - B) \right].$$

Since T and Z satisfy all conditions needed by Theorem 6, C satisfies (10)-(12). To get $G \xrightarrow{V} H$ it remains only to show satisfaction of (13): $C_{xyst} \geq 0$ for all x, y, s, t. When $s = t$,

$$\begin{aligned} C_{xyss} &= \lambda^{-1} [D_{ss} + (\lambda - 1)^{-1} Z_{xy} (\lambda D_{ss} - D_{ss})] \\ &= \lambda^{-1} (1 + Z_{xy}) D_{ss} \geq 0. \end{aligned}$$

The last inequality follows from $Z_{xy} \geq -1$ and $D_{ss} \geq 0$. When $s \neq t$,

$$\begin{aligned} C_{xyst} &= \lambda^{-1} [B_{st} + (\lambda - 1)^{-1} Z_{xy} (0 - B_{st})] \\ &= \lambda^{-1} (\lambda - 1)^{-1} [(\lambda - 1) - Z_{xy}] B_{st} \geq 0. \end{aligned}$$

The last inequality follows from $Z_{xy} \leq \max\{Z_{xx}, Z_{yy}\} = \lambda - 1$ (since $Z \succeq 0$) and $B_{st} \geq 0$. ■

Finally, we show that the two conditions $G \xrightarrow{+} H$ and $G \xrightarrow{B} H$ are not equivalent: the second one is weaker.

Theorem 14. *If $G \xrightarrow{+} H$ then $G \xrightarrow{B} H$, but there are graphs for which the converse does not hold.*

Proof: The forward implication is an immediate consequence of Theorems 6 and 10:

$$G \xrightarrow{+} H \implies \bar{\vartheta}(G) \le \bar{\vartheta}(H) \implies G \xrightarrow{B} H.$$

To see that the converse does not hold, take H to be any graph such that $\lfloor \bar{\vartheta}^-(H) \rfloor < \lfloor \bar{\vartheta}(H) \rfloor$. For example, a graph with $\bar{\vartheta}^-(H) = 4$ but $\bar{\vartheta}(H) = 16/3 > 5$ is given at the end of [15]. Then $5 = \bar{\vartheta}(K_5) \le \bar{\vartheta}(H) \implies K_5 \xrightarrow{B} H$ but $5 = \bar{\vartheta}^-(K_5) > \bar{\vartheta}^-(H) \implies K_5 \not\xrightarrow{+} H$. ■

IV. Quantum homomorphisms

Suppose Alice and Bob share an entangled state $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ on Hilbert spaces of arbitrary dimension. A referee asks Alice a question $x \in X$ and Bob a question $y \in Y$. Based on x, Alice performs a (without loss of generality, projective) measurement $\{E^x_s\}_s$ and reports outcome $s \in S$ to the referee. Similarly, Bob performs measurement $\{F^y_t\}_t$ and reports $t \in T$. The sets X, Y, S, T are finite. The probability distribution of Alice and Bob's answer, conditioned upon the referee's question, is

$$P(s, t|x, y) = \langle \psi | E^x_s \otimes F^y_t | \psi \rangle \tag{19}$$

where $\sum_s E^x_s = I$, $\sum_t F^y_t = I$, and $\langle \psi | \psi \rangle = 1$.

The assumption that Alice and Bob's measurements take such a tensor product form is associated with non-relativistic quantum mechanics. One may alternatively consider a model in which there is only a single Hilbert space, $|\psi\rangle \in \mathcal{H}_A$ and $E^x_s, F^y_t \in \mathcal{L}(\mathcal{H}_A)$, but in which each E^x_s commutes with each F^y_t. The conditional probability distribution in this model is

$$P(s, t|x, y) = \langle \psi | E^x_s F^y_t | \psi \rangle \, . \tag{20}$$

Tsirelson's problem is the question of whether these two models differ. That is to say, is there a conditional probability distribution realizable as (20) but not as (19)? Tsirelson showed that if the Hilbert spaces are finite dimensional then the two models are the same. A simplified proof appears in [19]. In addition to its importance to quantum mechanics, Tsirelson's problem is of mathematical interest since it is closely related to Connes' embedding problem [20]. Note that any correlation of the form (19) can be written in the form (20) since $\langle \psi | E^x_s \otimes F^y_t | \psi \rangle = \langle \psi | (E^x_s \otimes I)(I \otimes F^y_t) | \psi \rangle$ and $E^x_s \otimes I$ commutes with $I \otimes F^y_t$.

Graph G is said to have a *quantum homomorphism* to H (written $G \xrightarrow{q} H$) if there is a probability distribution of the form (19), with $X = Y = V(G)$, $S = T = V(H)$, and finite dimensional $|\psi\rangle$, satisfying [10]

$$\begin{aligned} P(s \ne t | x = y) &= 0 \\ P(s \not\sim_H t | x \sim_G y) &= 0. \end{aligned} \tag{21}$$

This is called a "quantum homomorphism" because if Alice and Bob are not allowed to share an entangled state [equivalently, if $\dim(\mathcal{H}_A) = \dim(\mathcal{H}_B) = 1$] then such a conditional probability distribution is achievable if and only if $G \to H$. Although $G \xrightarrow{q} H \implies G \xrightarrow{*} H$, it is an open question whether the converse holds.

In [10] it is shown that $G \xrightarrow{q} H$ if and only if there exist projection operators (i.e., Hermitian matrices with eigenvalues in $\{0, 1\}$) E^x_s for $x \in V(G)$ and $s \in V(H)$ such that

$$\begin{aligned} \sum_s E^x_s &= I \\ E^x_s E^y_t &= 0 \text{ for all } x \sim_G y, s \not\sim_H t \\ E^x_s E^x_t &= 0 \text{ for all } s \ne t. \end{aligned}$$

Note that the first condition actually implies the third. Define $|w^x_s\rangle = \mathrm{vec}(E^x_s)$. Since $\langle w^x_s | w^y_t \rangle = \mathrm{Tr}(E^x_s E^y_t)$, this gives a set of vectors satisfying the conditions of Definition 2 for $G \xrightarrow{V} H$. So $G \xrightarrow{q} H \implies G \xrightarrow{V} H$. Since $G \xrightarrow{V} H \implies G \xrightarrow{+} H$, Theorem 10 gives the following corollary which was previously shown in [21].

Corollary 15. *Suppose $G \xrightarrow{q} H$. Then $\bar{\vartheta}(G) \le \bar{\vartheta}(H)$, $\bar{\vartheta}^-(G) \le \bar{\vartheta}^-(H)$, and $\bar{\vartheta}^+(G) \le \bar{\vartheta}^+(H)$.*

The *quantum chromatic number* $\chi_q(G)$ is the least n such that $G \xrightarrow{q} K_n$, in analogy to the chromatic number $\chi(G)$ which is the least n such that $G \to K_n$. As a means of studying Tsirelson's problem, a number of variations of χ_q were considered in [22]. For instance, they define χ_{qr} by taking the correlation model to be (20) with infinite dimensional $|\psi\rangle$ rather than (19) with finite dimensional $|\psi\rangle$ as was used to define $G \xrightarrow{q} H$ (and thus χ_q). Also, they consider a semidefinite relaxation χ_{vect}. In the language of our paper, $\chi_{\mathrm{vect}}(G)$ is the least n such that $G \xrightarrow{V} K_n$. In fact, our $G \xrightarrow{V} H$ definition was inspired by their work[4]. One could also define $\omega_{\mathrm{vect}}(H)$ as the largest n such that $K_n \xrightarrow{V} H$. Both $\chi_{\mathrm{vect}}(G)$ and $\omega_{\mathrm{vect}}(H)$ can be computed using the tools of Section III.

Corollary 16. *$\chi_{vect}(G) = \lceil \bar{\vartheta}^+(G) \rceil$ and $\omega_{vect}(H) = \lfloor \bar{\vartheta}^-(H) \rfloor$.*

Proof: Theorem 13 gives (for integer n) $\bar{\vartheta}^+(G) \le n \implies G \xrightarrow{V} K_n$ and $n \le \bar{\vartheta}^-(H) \implies K_n \xrightarrow{V} H$. Theorem 10 gives the converse, so $\bar{\vartheta}^+(G) \le n \iff G \xrightarrow{V} K_n$ and $n \le \bar{\vartheta}^-(H) \iff K_n \xrightarrow{V} H$. ■

The authors of [22] posed the question of whether $\chi_{\mathrm{vect}}(G) = \chi_q(G)$, that is to say whether χ_q is equivalent to its semidefinite relaxation. In fact, these two quantities are not equal.

Theorem 17. *There is a graph G such that $\chi_{vect}(G) < \chi_q(G)$. Therefore $G \xrightarrow{V} H$ does not imply $G \xrightarrow{q} H$.*

Proof: In light of Corollary 16, the goal is to find G such that $\lceil \bar{\vartheta}^+(G) \rceil < \chi_q(G)$. The projective rank of a graph, $\xi_f(G)$, is the infimum of d/r such that the vertices of a graph can be assigned rank-r projectors in $\mathbb{C}^d$ such that adjacent vertices have orthogonal projectors. Since $\xi_f(G) \le \chi_q(G)$ [21], it suffices to find a gap between $\lceil \bar{\vartheta}^+(G) \rceil$ and $\xi_f(G)$.

[4] The first version of the present paper was posted before [22]. We later amended this paper to address the question posed in [22].

The five cycle has $\bar{\vartheta}^+(C_5) = \sqrt{5} < \xi_f(C_5) = 5/2$ [21]. But this is not enough since $\lceil\sqrt{5}\rceil = 3 > 5/2$. Fortunately, we can amplify the difference by taking the disjunctive product with a complete graph. $\bar{\vartheta}^+$ is sub-multiplicative under disjunctive product, as feasible solutions $Z + J$ to (17) can be combined by tensor product. Theorem 27 states that ξ_f is multiplicative under the disjunctive (and lexicographical) product, so

$$\left\lceil \bar{\vartheta}^+(C_5 * K_3) \right\rceil \leq \left\lceil 3\sqrt{5} \right\rceil = 7 < 3 \cdot \frac{5}{2} = \xi_f(C_5 * K_3).$$

■

Subsequently, this result has been strengthened to $\chi_{\text{vect}}(G) < \chi_{\text{qr}}(G)$ [23].

V. Conclusion

Beigi provided a vector relaxation of the entanglement assisted zero-error communication problem, leading to an upper bound on the entanglement assisted independence number: $\alpha^* \leq \lfloor\vartheta\rfloor$ [7]. We generalized Beigi's construction to apply it to entanglement assisted zero-error source-channel coding, defining a relaxed graph homomorphism $G \overset{B}{\to} H$. This ends up exactly characterizing monotonicity of ϑ, and shows that ϑ can be used to provide a lower bound on the cost rate for entanglement assisted source-channel coding. As a corollary we answer in the affirmative an open question posed by Beigi of whether a quantity β that he defined is equal to $\lfloor\vartheta\rfloor$. Applying a Beigi-style argument to chromatic number rather than independence number yields a quantity analogous to β which is equal to $\lceil\vartheta\rceil$. We defined a similar (and stronger) relaxation, $G \overset{+}{\to} H$, which yields bounds involving Schrijver's number ϑ^- and Szegedy's number ϑ^+. This leads to a stronger bound on entanglement assisted independence number: $\alpha^* \leq \lfloor\vartheta^-\rfloor$. In addition to these new bounds, we reproduce previously known bounds from [7], [9], [10], [21]. We also answer an open question from [22] regarding the relation of the quantum chromatic number to its semidefinite relaxation.

A number of open questions remain. Since there is a graph for which $\vartheta^- < \vartheta - 1$ [15], our bound $\alpha^* \leq \lfloor\vartheta^-\rfloor$ shows a gap between one-shot entanglement assisted zero-error capacity and $\lfloor\vartheta\rfloor$. However, since ϑ^- is not multiplicative, it is still not known whether there can be a gap between the *asymptotic* capacity (i.e. the entanglement assisted Shannon capacity) and ϑ. To show such a gap requires a stronger bound on entanglement assisted Shannon capacity. Haemers provided a bound on Shannon capacity which is sometimes stronger than Lovász' bound [24]–[27]; however, this bound does not apply to the entanglement assisted case [6].

The standard notion of graph homomorphism, along with two of its quantum generalizations, and our three relaxations, form a hierarchy as outlined in Fig. 3. In some cases we do not know whether converses hold. $G \overset{*}{\to} H$ is equivalent to $G \overset{q}{\to} H$ if and only if projective measurements and a maximally entangled state always suffice for entanglement assisted zero-error source-channel coding. Equivalence between $G \overset{*}{\to} H$ and $G \overset{+}{\to} H$ seems unlikely but would have two important consequences. First, we would have a much simpler characterization (vector rather than operator) of entanglement assisted homomorphisms and, in particular, entanglement assisted zero-error communication. Second, since $G \overset{V}{\to} H \implies G \overset{+}{\to} H$, the gap that we found between $G \overset{q}{\to} H$ and $G \overset{V}{\to} H$ would give a gap between $G \overset{q}{\to} H$ and $G \overset{*}{\to} H$.

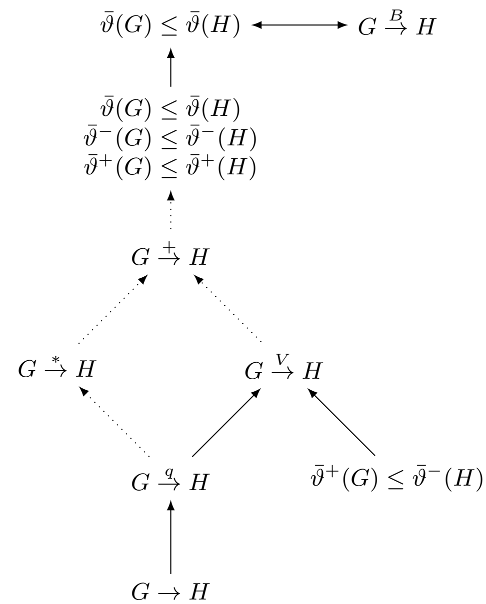

Fig. 3. Implications between various conditions discussed in this paper. Double ended arrows mean if-and-only-if, solid arrows mean the converse is known to not hold, and dotted arrows mean we do not know whether the converse holds.

After completing this work, we became aware of a previous investigation of a similar problem. Semidefinite relaxations of the homomorphism game (outside of the quantum context) were investigated in [28], and this was further developed in [29]. What they call a *hoax* corresponds to our $G \overset{V}{\to} H$, and what they call a *semi-hoax* corresponds to our $G \overset{B}{\to} H$. Though they studied the same problem, they reached a different conclusion: they showed (using our terminology)

$$G \overset{B}{\to} H \iff \bar{\vartheta}(G \circ H) = |V(G)| \tag{22}$$

$$G \overset{V}{\to} H \iff \bar{\vartheta}^-(G \circ H) = |V(G)| \tag{23}$$

where $\circ$ denotes the *hom-product* with vertices $V(G) \times V(H)$ and edges

$$(x, s) \sim (y, t) \iff (x \neq y) \text{ and } (x \sim y \implies s \sim t).$$

Combining our Theorem 6 with (22) gives $\bar{\vartheta}(G \circ H) = |V(G)| \iff \bar{\vartheta}(G) \leq \bar{\vartheta}(H)$ and combining Theorems 10 and 13 with (23) gives

$$\begin{aligned} \bar{\vartheta}^+(G) \leq \bar{\vartheta}^-(H) &\implies \bar{\vartheta}^-(G \circ H) = |V(G)| \\ &\implies \bar{\vartheta}(G) \leq \bar{\vartheta}(H), \bar{\vartheta}^-(G) \leq \bar{\vartheta}^-(H), \\ &\quad\ \ \bar{\vartheta}^+(G) \leq \bar{\vartheta}^+(H). \end{aligned}$$

Considering the special case $H = K_n$, we have $G \circ K_n = G \square K_n$ (Cartesian product) giving $\bar{\vartheta}(G \square K_n) = |V(G)| \iff \bar{\vartheta}(G) \leq n$ and $\bar{\vartheta}^-(G \square K_n) = |V(G)| \iff \bar{\vartheta}^+(G) \leq n$, reproducing Theorem 2.7 of [30].

VI. Acknowledgments

We thank Vern Paulsen and Ivan Todorov for inspiring us to define $G \xrightarrow{V} H$ and to find the gap between this and $G \xrightarrow{q} H$.

TC is supported by the Royal Society. LM is supported by the Ministry of Education (MOE) and National Research Foundation Singapore, as well as MOE Tier 3 Grant "Random numbers from quantum processes" (MOE2012-T3-1-009). DR is supported by an NTU start-up grant awarded to D.V. Pasechnik. SS is supported by the Royal Society and the British Heart Foundation. DS is supported by the National Science Foundation through Grant PHY-1068331. AW is supported by the European Commission (STREPs "QCS" and "RAQUEL"), the European Research Council (Advanced Grant "IRQUAT") and the Philip Leverhulme Trust; furthermore by the Spanish MINECO, project FIS2008-01236, with the support of FEDER funds.

References

[1] C. E. Shannon, "The zero error capacity of a noisy channel," *IRE Transactions on Information Theory*, vol. 2, no. 3, pp. 8–19, 1956.

[2] H. S. Witsenhausen, "The zero-error side information problem and chromatic numbers (corresp.)," *IEEE Transactions on Information Theory*, vol. 22, no. 5, pp. 592–593, 1976.

[3] J. Nayak, E. Tunçel, and K. Rose, "Zero-error source-channel coding with side information," *IEEE Transactions on Information Theory*, vol. 52, no. 10, pp. 4626–4629, 2006.

[4] L. Lovász, "On the Shannon capacity of a graph," *IEEE Transactions on Information Theory*, vol. 25, no. 1, pp. 1 – 7, Jan 1979. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1055985

[5] T. S. Cubitt, D. Leung, W. Matthews, and A. Winter, "Improving zero-error classical communication with entanglement," *Physical Review Letters*, vol. 104, p. 230503, Jun 2010. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRevLett.104.230503

[6] D. Leung, L. Mančinska, W. Matthews, M. Ozols, and A. Roy, "Entanglement can increase asymptotic rates of zero-error classical communication over classical channels," *Communications in Mathematical Physics*, vol. 311, no. 1, pp. 97–111, 2012.

[7] S. Beigi, "Entanglement-assisted zero-error capacity is upper-bounded by the Lovász ϑ function," *Physical Review A*, vol. 82, p. 010303, Jul 2010. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRevA.82.010303

[8] 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.

[9] J. Briët, H. Buhrman, M. Laurent, T. Piovesan, and G. Scarpa, "Zero-error source-channel coding with entanglement," 2013. [Online]. Available: http://www.arxiv.org/abs/1308.4283

[10] D. E. Roberson and L. Mančinska, "Graph homomorphisms for quantum players," 2012. [Online]. Available: http://www.arxiv.org/abs/1212.1724

[11] G. Hahn and C. Tardif, "Graph homomorphisms: structure and symmetry," in *Graph symmetry*. Springer, 1997, pp. 107–166.

[12] P. Hell and J. Nešetřil, *Graphs and Homomorphisms (Oxford Lecture Series in Mathematics and Its Applications)*. Oxford University Press, USA, 9 2004.

[13] M. K. de Carli Silva and L. Tunçel, "Optimization problems over unit-distance representations of graphs," *Electronic Journal of Combinatorics*, vol. 20, no. 1, p. P43, 2013.

[14] D. E. Knuth, "The sandwich theorem," *Electronic Journal of Combinatorics*, vol. 1, no. 1, p. A1, 1994. [Online]. Available: http://www.combinatorics.org/Volume_1/Abstracts/v1i1a1.html

[15] A. Schrijver, "A comparison of the Delsarte and Lovász bounds," *IEEE Transactions on Information Theory*, vol. 25, no. 4, pp. 425–429, 1979.

[16] R. J. McEliece, E. R. Rodemich, and H. C. Rumsey Jr., "The Lovász bound and some generalizations," *Journal of Combinatorics and Information Systems Science*, vol. 3, no. 3, pp. 134–152, 1978.

[17] M. Szegedy, "A note on the ϑ number of Lovász and the generalized Delsarte bound," in *Proc. 35th Annual Symposium on Foundations of Computer Science, 1994*, 1994, pp. 36–39.

[18] L. Lovász, "Semidefinite programs and combinatorial optimization," in *Recent Advances in Algorithms and Combinatorics*, ser. CMS Books in Mathematics / Ouvrages de mathématiques de la SMC, B. A. Reed and C. L. Sales, Eds. Springer New York, 2003, pp. 137–194. [Online]. Available: http://dx.doi.org/10.1007/0-387-22444-0_6

[19] V. B. Scholz and R. F. Werner, "Tsirelson's problem," 2008. [Online]. Available: http://www.arxiv.org/abs/0812.4305

[20] M. Junge, M. Navascues, C. Palazuelos, D. Perez-Garcia, V. B. Scholz, and R. F. Werner, "Connes' embedding problem and Tsirelson's problem," *Journal of Mathematical Physics*, vol. 52, no. 1, 2011. [Online]. Available: http://scitation.aip.org/content/aip/journal/jmp/52/1/10.1063/1.3514538

[21] D. E. Roberson, "Variations on a theme: Graph homomorphisms," Ph.D. dissertation, University of Waterloo, 2013.

[22] V. I. Paulsen and I. G. Todorov, "Quantum chromatic numbers via operator systems," 2013. [Online]. Available: http://www.arxiv.org/abs/1311.6850

[23] V. I. Paulsen, S. Severini, D. Stahlke, I. G. Todorov, and A. Winter, "Estimating quantum chromatic numbers," 2014. [Online]. Available: http://www.arxiv.org/abs/1407.6918

[24] W. H. Haemers, "An upper bound for the Shannon capacity of a graph," *Colloquia Mathematica Societatis Janos Bolyai*, vol. 25, pp. 267–272, 1978. [Online]. Available: http://arno.uvt.nl/show.cgi?fid=80314

[25] ——, "On some problems of Lovász concerning the Shannon capacity of a graph," *IEEE Transactions on Information Theory*, vol. 25, pp. 231–232, 1979. [Online]. Available: http://arno.uvt.nl/show.cgi?fid=80311

[26] N. Alon, "The Shannon capacity of a union," *Combinatorica*, vol. 18, no. 3, pp. 301–310, 1998. [Online]. Available: http://dx.doi.org/10.1007/PL00009824

[27] R. Peeters, "Orthogonal representations over finite fields and the chromatic number of graphs," *Combinatorica*, vol. 16, no. 3, pp. 417–431, 1996. [Online]. Available: http://dx.doi.org/10.1007/BF01261326

[28] U. Feige and L. Lovász, "Two-prover one-round proof systems: Their power and their problems (extended abstract)," in *Proceedings of the Twenty-fourth Annual ACM Symposium on Theory of Computing*, ser. STOC '92. New York, NY, USA: ACM, 1992, pp. 733–744. [Online]. Available: http://doi.acm.org/10.1145/129712.129783

[29] R. Bačík and S. Mahajan, "Semidefinite programming and its applications to NP problems," in *Computing and Combinatorics*, ser. Lecture Notes in Computer Science, D.-Z. Du and M. Li, Eds. Springer Berlin Heidelberg, 1995, vol. 959, pp. 566–575. [Online]. Available: http://dx.doi.org/10.1007/BFb0030878

[30] N. Gvozdenović and M. Laurent, "The operator Ψ for the chromatic number of a graph," *SIAM Journal on Optimization*, vol. 19, no. 2, pp. 572–591, 2008. [Online]. Available: http://dx.doi.org/10.1137/050648237

[31] I. Bomze, F. Frommlet, and M. Locatelli, "Gap, cosum and product properties of the θ' bound on the clique number," *Optimization*, vol. 59, no. 7, pp. 1041–1051, 2010.

[32] A. Galtman, "Spectral characterizations of the Lovász number and the Delsarte number of a graph," *Journal of Algebraic Combinatorics*, vol. 12, no. 2, pp. 131–143, 2000.

Appendix A
Multiplicativity

In Lovász' original paper [4] on the ϑ function, he proved that

$$\vartheta(G \boxtimes H) = \vartheta(G)\vartheta(H),$$

i.e. ϑ is multiplicative with respect to the strong product. To do this he proved the following two inequalities:

$$\vartheta(G)\vartheta(H) \leq \vartheta(G \boxtimes H) \leq \vartheta(G)\vartheta(H).$$

This sufficed for Lovász because it was only required to show that ϑ is multiplicative with respect to the strong product in order to prove that it was an upper bound on Shannon capacity. However, Lovász also noted that his proof of the first inequality above also proves the following stronger statement:

$$\vartheta(G)\vartheta(H) \leq \vartheta(G * H).$$

Together these inequalities imply that ϑ is multiplicative with respect to both the strong and disjunctive products. Our aim in this appendix is to show that ϑ^- is not multiplicative with respect to the strong product and ϑ^+ is not multiplicative with respect to the disjunctive product, but ϑ^- is multiplicative with respect to the disjunctive product. Also we will show multiplicativity of projective rank ξ_f.

A. Counterexamples

Some of the inequalities involving ϑ above can be proved for ϑ^- as well. Adapting Lovász' proof of the analogous statement for ϑ, it can be shown that

$$\vartheta^-(G \boxtimes H) \geq \vartheta^-(G * H) \geq \vartheta^-(G)\vartheta^-(H).$$

Similarly, it can be shown that

$$\vartheta^+(G * H) \leq \vartheta^+(G \boxtimes H) \leq \vartheta^+(G)\vartheta^+(H).$$

Therefore, in order to show that neither ϑ^- or ϑ^+ are multiplicative with respect to both the strong and disjunctive products, we must find counterexamples to both of the following inequalities:

$$\vartheta^-(G \boxtimes H) \leq \vartheta^-(G)\vartheta^-(H), \quad \vartheta^+(G * H) \geq \vartheta^+(G)\vartheta^+(H).$$

At the end of [15], Schrijver gives an example of a graph, which we refer to as G_S, that satisfies $\vartheta^-(G_S) < \vartheta(G_S)$. The vertices of G_S are the 0-1-strings of length six, and two strings are adjacent if their Hamming distance is at most three. In other words, Schrijver's graph G_S is an instance of a *Hamming graph*. Note that this graph is vertex transitive. We will see how to use the graph G_S to construct counterexamples to both of the above inequalities. To do this we will need two lemmas, the first of which is from [4].

Lemma 18. *For any graph G,*

$$\vartheta(G)\vartheta(\overline{G}) \geq |V(G)|,$$

with equality when G is vertex transitive.

An analogous statement involving ϑ^- and ϑ^+ was proved by Szegedy in [17]:

Lemma 19. *For any graph G,*

$$\vartheta^-(G)\vartheta^+(\overline{G}) \geq |V(G)|,$$

with equality when G is vertex transitive.

One easy consequence of these lemmas is that if G is a vertex transitive graph such that $\vartheta^-(G) < \vartheta(G)$, then

$$\vartheta^+(\overline{G}) = \frac{|V(G)|}{\vartheta^-(G)} > \frac{|V(G)|}{\vartheta(G)} = \vartheta(\overline{G}).$$

In particular this implies that $\vartheta^+(\overline{G_S}) > \vartheta(\overline{G_S})$.

More pertinent to our discussion are the following lemmas.

Lemma 20. *If G is a vertex transitive graph such that $\vartheta^-(G) < \vartheta(G)$, then*

$$\vartheta^-(G \boxtimes \overline{G}) > \vartheta^-(G)\vartheta^-(\overline{G}).$$

Proof: First note the vertices of the form (v, v) in $G \boxtimes \overline{G}$ form an independent set of size $|V(G)|$. Therefore, $\vartheta^-(G \boxtimes \overline{G}) \geq |V(G)|$, and we have the following:

$$\vartheta^-(G)\vartheta^-(\overline{G}) < \vartheta(G)\vartheta(\overline{G}) = |V(G)| \leq \vartheta^-(G \boxtimes \overline{G}),$$

since $\vartheta^-(\overline{G}) \leq \vartheta(\overline{G})$. ■

Since G_S satisfies the hypotheses of Lemma 20, we have the following desired corollary:

Corollary 21. *The parameter ϑ^- is not multiplicative with respect to the strong product.*

We are also able to use Lemma 20 to prove a similar lemma for ϑ^+.

Lemma 22. *If G is a vertex transitive graph such that $\vartheta^-(G) < \vartheta(G)$, then*

$$\vartheta^+(G * \overline{G}) < \vartheta^+(G)\vartheta^+(\overline{G}).$$

Proof: Suppose that G is such a graph. By Lemma 20, we have that

$$\vartheta^-(G \boxtimes \overline{G}) > \vartheta^-(G)\vartheta^-(\overline{G}).$$

Since G is vertex transitive, so is $G * \overline{G}$ and thus we can apply Lemma 19 to obtain

$$\begin{aligned}\vartheta^+(G * \overline{G}) &= \frac{|V(G)|^2}{\vartheta^-\left(\overline{G * \overline{G}}\right)} = \frac{|V(G)|^2}{\vartheta^-(\overline{G} \boxtimes G)} \\ &< \frac{|V(G)|}{\vartheta^-(\overline{G})} \frac{|V(G)|}{\vartheta^-(G)} = \vartheta^+(G)\vartheta^+(\overline{G}).\end{aligned}$$

■

Similarly to the above, this implies the following:

Corollary 23. *The parameter ϑ^+ is not multiplicative with respect to the disjunctive product.*

Even though ϑ^- is not multiplicative with respect to the strong product, nor is ϑ^+ with respect to the disjunctive product, one could ask whether they are at least multiplicative with respect to the corresponding graph powers, as this would be enough to prove an analogue of Corollary 9. It turns out that they are not, as we now show. Non-multiplicativity for ϑ^- was shown already in [31] but with a much smaller gap.

Corollary 24. *The parameter ϑ^- is not multiplicative under strong graph powers $G^{\boxtimes n}$, and ϑ^+ is not multiplicative under disjunctive graph powers G^{*n}.*

Proof: Let G_S be a vertex transitive graph such that $\vartheta^-(G_S) < \vartheta(G_S)$, whose existence was discussed above. Let $H = G_S \oplus \overline{G_S}$ where $\oplus$ denotes disjoint union. Since ϑ^- is additive under disjoint union and is super-multiplicative under the strong product,

$$\begin{aligned}\vartheta^-(H^{\boxtimes 2}) &= \vartheta^-[G_S^{\boxtimes 2} \oplus (G_S \boxtimes \overline{G_S}) \oplus (\overline{G_S} \boxtimes G_S) \oplus \overline{G_S}^{\boxtimes 2}] \\ &= \vartheta^-(G_S^{\boxtimes 2}) + \vartheta^-(G_S \boxtimes \overline{G_S}) \\ &\quad + \vartheta^-(\overline{G_S} \boxtimes G_S) + \vartheta^-(\overline{G_S}^{\boxtimes 2}) \\ &\geq \vartheta^-(G_S)^2 + \vartheta^-(G_S \boxtimes \overline{G_S}) \\ &\quad + \vartheta^-(\overline{G_S} \boxtimes G_S) + \vartheta^-(\overline{G_S})^2 \\ &> \vartheta^-(G_S)^2 + \vartheta^-(G_S)\vartheta^-(\overline{G_S}) \\ &\quad + \vartheta^-(\overline{G_S})\vartheta^-(G_S) + \vartheta^-(\overline{G_S})^2 \\ &= [\vartheta^-(G_S) + \vartheta^-(\overline{G_S}]^2 \\ &= \vartheta^-(H)^2.\end{aligned}$$

Similarly,

$$\begin{aligned}\vartheta^+(H^{*2}) &= \vartheta^+[(G_S \oplus \overline{G_S}) * (G_S \oplus \overline{G_S})] \\ &\leq \vartheta^+[G_S^{\boxtimes 2} \oplus (G_S * \overline{G_S}) \\ &\quad \oplus (\overline{G_S} * G_S) \oplus \overline{G_S}^{\boxtimes 2}] \qquad (24) \\ &\leq \vartheta^+(G_S)^2 + \vartheta^+(G_S * \overline{G_S}) \\ &\quad + \vartheta^+(\overline{G_S} * G_S) + \vartheta^+(\overline{G_S})^2 \\ &< \vartheta^+(G_S)^2 + \vartheta^+(G_S)\vartheta^+(\overline{G_S}) \\ &\quad + \vartheta^+(\overline{G_S})\vartheta^+(G_S) + \vartheta^+(\overline{G_S})^2 \\ &= [\vartheta^+(G_S) + \vartheta^+(\overline{G_S}]^2 \\ &= \vartheta^+(H)^2.\end{aligned}$$

where (24) follows from the fact that $\vartheta^+(G_1) \geq \vartheta^+(G_2)$ when G_1 is a subgraph of G_2. ■

B. ϑ^- and the disjunctive product

Though ϑ^- is not multiplicative with respect to the strong product, we are able to show that it is multiplicative with respect to the disjunctive product and the lexicographical product. The lexicographical product $G[H]$ has vertices $V(G) \times V(H)$ and edges $(x, y) \sim (x', y')$ if $x \sim_G x'$ or ($x = x'$ and $y \sim_H y'$).

Theorem 25. *Schrijver's number is multiplicative under the disjunctive and the lexicographical products:* $\vartheta^-(G * H) = \vartheta^-(G[H]) = \vartheta^-(G)\vartheta^-(H)$.

Proof: We use the following formulation for Schrijver's number:

$$\begin{aligned}\vartheta^-(G) = \max\{\langle B, J\rangle : B &\succeq 0, \mathrm{Tr}B = 1, \\ B_{ij} &\geq 0 \text{ for all } i, j, \\ B_{ij} &= 0 \text{ for } i \sim j\}. \qquad (25)\end{aligned}$$

It is easy to show that $\vartheta^-(G * H) \geq \vartheta^-(G)\vartheta^-(H)$: if B_G and B_H are optimal solutions of (25) for $\vartheta^-(G)$ and $\vartheta^-(H)$ then $B_G \otimes B_H$ is feasible for (25) for $\vartheta^-(G * H)$ with value $\vartheta^-(G)\vartheta^-(H)$. Since $G[H]$ is a subgraph of $G * H$, we have $\vartheta^-(G[H]) \geq \vartheta^-(G * H)$. It remains only to show $\vartheta^-(G[H]) \leq \vartheta^-(G)\vartheta^-(H)$.

Let B be an optimal solution for (25) for $\vartheta^-(G[H])$. This can be considered as an operator $B \in \mathcal{L}(\mathbb{C}^{|V(G)|}) \otimes \mathcal{L}(\mathbb{C}^{|V(H)|})$, and we have $\langle B, J_G \otimes J_H\rangle = \vartheta^-(G[H])$ where J_G is the all ones matrix indexed by $V(G)$ and similarly for J_H. Note that $B_{xyx'y'} = 0$ when $x \sim_G x'$ or ($x = x'$ and $y \sim_H y'$). For $x \in V(G)$ let $|x\rangle$ denote the corresponding basis vector in $\mathbb{C}^{|V(G)|}$ and define

$$B^x = (\langle x| \otimes I)B(|x\rangle \otimes I) \in \mathcal{L}(\mathbb{C}^{|V(H)|}).$$

Since $B^x \succeq 0$, $B^x_{yy'} \geq 0$ for all y, y', and $B^x_{yy'} = 0$ when $y \sim_H y'$, it holds that $B^x/\mathrm{Tr}B^x$ is feasible for (25) for $\vartheta^-(H)$. The value of this solution is $\langle B^x, J_H\rangle/\mathrm{Tr}B^x$, thus

$$\frac{\langle B^x, J_H\rangle}{\mathrm{Tr}B^x} \leq \vartheta^-(H).$$

Also, $\sum_x \mathrm{Tr}B^x = \mathrm{Tr}B = 1$, giving

$$\sum_x \langle B^x, J_H\rangle \leq \sum_x \vartheta^-(H)\mathrm{Tr}B^x = \vartheta^-(H).$$

Define $B' = (I \otimes \langle \mathbf{1}|)B(I \otimes |\mathbf{1}\rangle) \in \mathcal{L}(\mathbb{C}^{|V(G)|})$ where $|\mathbf{1}\rangle$ is the all ones vector. Note that $B' = \mathrm{Tr}_H\{(I \otimes J_H)B\}$. Since $B' \succeq 0$, $B'_{xx'} \geq 0$ for all x, x', and $B'_{xx'} = 0$ when $x \sim_G x'$, it holds that $B'/\mathrm{Tr}B'$ is feasible for (25) for $\vartheta^-(G)$. The value of this solution is $\langle B', J_G\rangle/\mathrm{Tr}B'$, thus $\langle B', J_G\rangle/\mathrm{Tr}B' \leq \vartheta^-(G)$. Finally,

$$\begin{aligned}\vartheta^-(G[H]) = \langle B, J_G \otimes J_H\rangle &= \langle B', J_G\rangle \\ &\leq \vartheta^-(G)\mathrm{Tr}B' \\ &= \vartheta^-(G)\sum_x \langle B^x, J_H\rangle \\ &\leq \vartheta^-(G)\vartheta^-(H).\end{aligned}$$

■

C. What About ϑ^+?

Based on other results concerning ϑ^- and ϑ^+, Theorem 25 seems to suggest that one should be able to prove that ϑ^+ is multiplicative with respect to the strong product. We already noted above that one of the needed inequalities, namely $\vartheta^+(G \boxtimes H) \leq \vartheta^+(G)\vartheta^+(H)$, does hold, so we would only need to show that $\vartheta^+(G \boxtimes H) \geq \vartheta^+(G)\vartheta^+(H)$ holds as well. For now, a proof of this fact eludes us, but we are able to prove the multiplicativity of ϑ^+ in the case of vertex transitive graphs using Lemma 19 and the multiplicativity of ϑ^- with respect to the disjunctive product.

Theorem 26. *If G and H are vertex transitive, then*

$$\vartheta^+(G \boxtimes H) = \vartheta^+(G)\vartheta^+(H).$$

Proof: Since G and H are vertex transitive, so is $G \boxtimes H$. Therefore

$$\begin{aligned}\vartheta^+(G \boxtimes H) &= \frac{|V(G)| \cdot |V(H)|}{\vartheta^-(\overline{G} * \overline{H})} = \frac{|V(G)| \cdot |V(H)|}{\vartheta^-(\overline{G})\vartheta^-(\overline{H})} \\ &= \vartheta^+(G)\vartheta^+(H).\end{aligned}$$

■

This seems to be pretty strong evidence that ϑ^+ is multiplicative with respect to the strong product in general.

D. Projective Rank

The *projective rank* of a graph, $\xi_f(G)$, is the infimum of d/r such that the vertices of a graph can be assigned rank-r projectors in $\mathbb{C}^d$ such that adjacent vertices have orthogonal projectors. Such an assignment is called a *d/r-representation*. The 'f' subscript in the notation for projective rank indicates that it can be thought of as a fractional version of orthogonal rank: the minimum dimension of an assignment of vectors such that adjacent vertices receive orthogonal vectors. We will show ξ_f to be multiplicative under both the disjunctive and the lexicographical products. As a reminder, the lexicographical product $G[H]$ has edges $(x, y) \sim (x', y')$ if $x \sim_G x'$ or ($x = x'$ and $y \sim_H y'$).

Theorem 27. *Projective rank is multiplicative under the disjunctive and the lexicographical products: $\xi_f(G * H) = \xi_f(G[H]) = \xi_f(G)\xi_f(H)$.*

Proof: A d_1/r_1-representation for G and a d_2/r_2-representation for H can be turned into a d_1d_2/r_1r_2-representation for $G * H$ by taking the tensor products of the projectors associated with each graph. So $\xi_f(G * H) \le \xi_f(G)\xi_f(H)$.

On the other hand, let U_{xy} for $x \in V(G), y \in V(H)$ be the subspaces associated with a d/r-representation of $G[H]$. For each x, the subspaces $\{U_{xy} : y\}$ form an r'_x/r projective representation of H where r'_x is the dimension of $\text{span}\{U_{xy} : y\}$, so it must hold that $r'_x/r \ge \xi_f(H)$. Let $r' = \min\{r'_x\}$ and for each x let V_x be an r' dimensional subspace of $\text{span}\{U_{xy} : y\}$. These form a d/r' representation of G, so $d/r' \ge \xi_f(G)$. Then, $d/r = (d/r')(r'/r) \ge \xi_f(G)\xi_f(H)$ so $\xi_f(G[H]) \ge \xi_f(G)\xi_f(H)$. Since $G[H] \subseteq G * H$ we have $\xi_f(G[H]) \ge \xi_f(G)\xi_f(H) \ge \xi_f(G * H) \ge \xi_f(G[H])$. ∎

Appendix B
An if-and-only-if for Schrijver's number

Monotonicity of Schrijver's number admits an if-and-only-if statement along the lines of Theorem 6; however, the corresponding conditions on the $|w_s^x\rangle$ vectors are a bit more complicated and there is seemingly no direct connection to entanglement assisted source-channel coding. Specifically, we have the following result:

Theorem 28. *$\bar{\vartheta}^-(G) \le \bar{\vartheta}^-(H)$ if and only if there are vectors $|w\rangle \ne 0$ and $|w_s^x\rangle \in \mathbb{C}^d$ for each $x \in V(G)$, $s \in V(H)$, for some $d \in \mathbb{N}$, such that*

1) $\sum_s |w_s^x\rangle = |w\rangle$
2) $\langle w_s^x|w_t^y\rangle = 0$ *for* $s \not\sim_H t$, $s \ne t$
3) $\langle w_s^x|w_s^y\rangle \le 0$ *for* $x \sim_G y$
4) $\langle w_s^x|w_t^x\rangle = 0$ *for* $s \ne t$
5) $\langle w_s^x|w_t^y\rangle \ge 0$ *for* $s \ne t$.

The proof is a straightforward modification of the proof for Theorem 6. Before proceeding with this, it is necessary to express $\bar{\vartheta}^-$ in a form analogous to (14). This characterization appears without proof in [32]; we give the proof below. We do not know how to provide such a formulation for $\bar{\vartheta}^+$, so it may be possible that $\bar{\vartheta}^+$ does not admit an if-and-only-if statement along the lines of Theorems 6 and 28.

Theorem 29.

$$\begin{aligned} \bar{\vartheta}^-(G) = \max\{\|I+T\| : \ & I + T \succeq 0, \\ & T_{ij} = 0 \text{ for } i \not\sim j, \\ & T_{ij} \ge 0 \text{ for all } i, j\}. \end{aligned} \tag{26}$$

Proof: The dual to the semidefinite program (16) is [15]

$$\begin{aligned} \bar{\vartheta}^-(G) = \max\{\langle B, J\rangle : \ & B \succeq 0, \\ & \text{Tr}B = 1, \\ & B_{ij} = 0 \text{ for } i \not\sim j, i \ne j, \\ & B_{ij} \ge 0 \text{ for all } i, j\}. \end{aligned} \tag{27}$$

Let T be the optimal solution for (26). We will show that this induces a feasible solution for (27) via the recipe

$$B = |\psi\rangle\langle\psi| \circ (I + T),$$

where $|\psi\rangle$ is the eigenvector corresponding to the largest eigenvalue of $I + T$. This is positive semidefinite (being the Schur–Hadamard product of two positive semidefinite matrices), and $\langle B, J\rangle = \langle\psi|I + T|\psi\rangle = \lambda$. T_{ii} vanishes, so the diagonal of B is equal to the diagonal of $|\psi\rangle\langle\psi|$; consequently $\text{Tr}B = 1$. The matrix $I + T$ has nonnegative entries so its eigenvector $|\psi\rangle$ can be chosen nonnegative, leading to to $B_{ij} \ge 0$. So B is feasible for (27) and (27) $\ge$ (26).

Conversely, suppose that B is feasible for (27) with value λ. Let D be the diagonal component of B. Let $D^{-1/2}$ be the diagonal matrix having entries $D_{ii} = 1/\sqrt{B_{ii}}$ with the convention $1/0 = 0$ (note that $D^{-1/2}$ is the Moore–Penrose pseudoinverse of $D^{1/2}$). Define

$$T = D^{-1/2}(B - D)D^{-1/2}.$$

When $i \not\sim j$, this matrix satisfies $T_{ij} = 0$. Since D and $B - D$ have nonnegative entries, T does as well. We have

$$\begin{aligned} I + T &\succeq D^{-1/2}DD^{-1/2} + T \\ &= D^{-1/2}BD^{-1/2} \\ &\succeq 0. \end{aligned} \tag{28}$$

So T is feasible for (26). Let $|\psi\rangle$ be the vector with coefficients $\psi_i = \sqrt{B_{ii}}$. Since $\text{Tr}B = 1$, this is a unit vector. Making use of (28),

$$\begin{aligned} \langle\psi|I + T|\psi\rangle &\ge \left\langle\psi\middle|D^{-1/2}BD^{-1/2}\middle|\psi\right\rangle \\ &= \sum_{\substack{ij \text{ s.t.} \\ B_{ii}B_{jj} \ne 0}} B_{ij} \\ &= \sum_{ij} B_{ij} \\ &= \langle J, B\rangle = \lambda. \end{aligned} \tag{29}$$

Equality (29) holds because B is positive semidefinite and so satisfies $B_{ij} = 0$ when $B_{ii}B_{jj} = 0$. Since T is feasible for (26),

$$(26) \ge \|I + T\| \ge \lambda = (27).$$

∎

Proof of Theorem 28: As in the proof of Theorem 6, we work with the Gram matrix of the $|w_s^x\rangle$ vectors. The existence of vectors satisfying the conditions in the theorem statement is easily seen to be equivalent to the existence of a matrix $C : \mathcal{L}(\mathbb{C}^{|V(G)|}) \otimes \mathcal{L}(\mathbb{C}^{|V(H)|})$ satisfying

$$\begin{aligned}
&C \succeq 0\\
&\sum_{st} C_{xyst} = 1\\
&C_{xyst} = 0 \text{ for } s \not\sim t, s \neq t\\
&C_{xyss} \leq 0 \text{ for } x \sim y\\
&C_{xxst} = 0 \text{ for } s \neq t\\
&C_{xyst} \geq 0 \text{ for } s \neq t
\end{aligned}$$

Using this characterization, we proceed with the proof.

($\Longrightarrow$): Suppose $\bar{\vartheta}^-(G) \leq \bar{\vartheta}^-(H)$. We will explicitly construct a matrix C having the above properties. Let $\lambda = \bar{\vartheta}^-(H)$. By Theorem 29 there is a matrix T such that $\|I + T\| = \lambda$, $I + T \succeq 0$, $T_{st} = 0$ for $s \not\sim t$, and $T_{st} \geq 0$ for all s, t. Let $|\psi\rangle$ be the vector corresponding to the largest eigenvalue of $I+T$, which can be chosen nonnegative since T is entrywise nonnegative. With $\circ$ denoting the Schur–Hadamard product, define the matrices

$$\begin{aligned}
D &= |\psi\rangle\langle\psi| \circ I,\\
B &= |\psi\rangle\langle\psi| \circ (I + T).
\end{aligned}$$

These are entrywise nonnegative. With J being the all-ones matrix and $\langle\cdot,\cdot\rangle$ denoting the Hilbert–Schmidt inner product, it is readily verified that

$$\begin{aligned}
\langle D, J\rangle &= \langle\psi|\psi\rangle = 1,\\
\langle B, J\rangle &= \langle\psi|I + T|\psi\rangle = \lambda.
\end{aligned}$$

Schur–Hadamard products between positive semidefinite matrices yield positive semidefinite matrices. As a consequence, $B \succeq 0$ and

$$\|I + T\| = \lambda \implies \lambda I - (I + T) \succeq 0 \implies \lambda D - B \succeq 0.$$

Since $\lambda \geq \bar{\vartheta}^-(G)$, there is a matrix Z such that $Z \succeq 0$, $Z_{xx} = \lambda - 1$ for all x, and $Z_{xy} \leq -1$ for all $x \sim y$. Note that (16) gives existence of a matrix with $\bar{\vartheta}^-(G) - 1$ on the diagonal, but since $\lambda \geq \bar{\vartheta}^-(G)$ we can add a multiple of the identity to get $\lambda - 1$ on the diagonal.

We now construct C. Define

$$C = \lambda^{-1}\left[J \otimes B + (\lambda - 1)^{-1} Z \otimes (\lambda D - B)\right].$$

Since J, B, Z, and $\lambda D - B$ are all positive semidefinite, and $\lambda - 1 \geq 0$, we have that C is positive semidefinite. The other desired conditions on C are easy to verify. For all x, y we have

$$\begin{aligned}
\sum_{st} C_{xyst} &= \lambda^{-1}\left[\langle B, J\rangle + (\lambda - 1)^{-1} Z_{xy}[\lambda\langle D, J\rangle - \langle B, J\rangle]\right]\\
&= 1.
\end{aligned}$$

For $s \not\sim t$, $s \neq t$, we have that $B_{st} = D_{st} = 0$ so $C_{xyst} = 0$. For $x \sim y$,

$$\begin{aligned}
C_{xyss} &= \lambda^{-1}\left[B_{ss} + (\lambda - 1)^{-1} Z_{xy}(\lambda D_{ss} - B_{ss})\right]\\
&= \lambda^{-1}\left[D_{ss} + (\lambda - 1)^{-1} Z_{xy}(\lambda D_{ss} - D_{ss})\right]\\
&= \lambda^{-1} D_{ss}\left[1 + Z_{xy}\right]\\
&\leq 0.
\end{aligned}$$

For all x and for $s \neq t$,

$$\begin{aligned}
C_{xxst} &= \lambda^{-1}\left[B_{st} + (\lambda - 1)^{-1} Z_{xx}(\lambda D_{st} - B_{st})\right]\\
&= \lambda^{-1} B_{st}\left[1 - (\lambda - 1)^{-1} Z_{xx}\right]\\
&= 0.
\end{aligned}$$

For all x, y and for $s \neq t$,

$$\begin{aligned}
C_{xyst} &= \lambda^{-1}\left[B_{st} + (\lambda - 1)^{-1} Z_{xy}(\lambda D_{st} - B_{st})\right]\\
&= \lambda^{-1} B_{st}\left[1 - (\lambda - 1)^{-1} Z_{xy}\right]\\
&\geq 0,
\end{aligned}$$

where the last inequality follows from the fact that $Z \succeq 0 \implies |Z_{xy}| \leq \max\{Z_{xx}, Z_{yy}\} = \lambda - 1$.

($\Longleftarrow$): Let Z achieve the optimal value (call it λ) for the minimization program (16) for $\bar{\vartheta}^-(H)$. We will provide a feasible solution for (16) for $\bar{\vartheta}^-(G)$ to show that $\bar{\vartheta}^-(G) \leq \bar{\vartheta}^-(H)$. Specifically, let

$$Y = \big(I \otimes \langle\mathbf{1}|\big)\big[(J \otimes Z) \circ C\big]\big(I \otimes |\mathbf{1}\rangle\big),$$

as in the proof of Theorem 6. Since $C \succeq 0$ and $Z \succeq 0$, and positive semidefiniteness is preserved by conjugation, we have that $Y \succeq 0$. Considering the entries of Y we see that $Y_{xy} = \sum_{st} Z_{st} C_{xyst}$.

Using the fact that $Z_{ss} = \lambda - 1$ and $C_{xxst} = 0$ for $s \neq t$, we have

$$Y_{xx} = \sum_{st} Z_{st} C_{xxst} = (\lambda - 1)\sum_{st} C_{xxst} = \lambda - 1.$$

For $x \sim y$ we have

$$\begin{aligned}
Y_{xy} &= \sum_{st} Z_{st} C_{xyst}\\
&= \sum_{s \sim t} \underbrace{Z_{st}}_{\leq -1} \underbrace{C_{xyst}}_{\geq 0} + \sum_{s \not\sim t, s \neq t} Z_{st} \underbrace{C_{xyst}}_{=0} + \sum_{s} \underbrace{Z_{ss}}_{\geq -1} \underbrace{C_{xyss}}_{\leq 0}\\
&\leq \sum_{s \sim t} (-1) C_{xyst} + \sum_{s \not\sim t, s \neq t} (-1) C_{xyst} + \sum_{s} (-1) C_{xyss}\\
&= \sum_{st} (-1) C_{xyst} = -1.
\end{aligned}$$

Now define a matrix Y' consisting of the real part of Y (i.e. with coefficients $Y'_{xy} = \mathrm{Re}[Y_{xy}]$). This matrix is real, positive semidefinite, and satisfies $Y_{xx} = \lambda - 1$ for all x and $Y_{xy} \leq -1$ for $x \sim y$. Therefore Y' is feasible for (16) with value $\lambda = \bar{\vartheta}^-(H)$. Since $\bar{\vartheta}^-(G)$ is the minimum possible value of (16), we have $\bar{\vartheta}^-(G) \leq \bar{\vartheta}^-(H)$. ■

By setting $G = K_n$ or $H = K_n$ it is possible to formulate corollaries analogous to Corollaries 7 and 8. We describe only the first of these here.

Corollary 30. *Let $\beta^-(H)$ be the largest n such that there are vectors $|w\rangle \neq 0$ and $|w_s^x\rangle \in \mathbb{C}^d$ for each $x \in \{1, \ldots, n\}$, $s \in V(H)$, for some $d \in \mathbb{N}$, such that*

1) $\sum_s |w_s^x\rangle = |w\rangle$
2) $\langle w_s^x | w_t^y\rangle = 0$ *for* $s \sim_H t$
3) $\langle w_s^x | w_s^y\rangle \leq 0$ *for* $x \neq y$
4) $\langle w_s^x | w_t^x\rangle = 0$ *for* $s \neq t$
5) $\langle w_s^x | w_t^y\rangle \geq 0$ *for* $s \neq t$.

Then $\beta^-(H) = \lfloor \vartheta^-(H) \rfloor$.

Toby Cubitt received his Bachelor's and Master's degrees from the University of Cambridge, U.K., in 2002. He carried out his doctoral studies at the Max Planck Institute for Quantum Optics, Garching, Germany, and received his Ph.D. degree from the Technische Universität München in 2006. He was a Research Assistant and then a Leverhulme Early Career Fellow at Bristol University until 2010, a Juan de la Cierva Fellow at the Universidad Complutense de Madrid until 2012, and now holds a Royal Society University Research Fellowship in the Department for Applied Mathematics and Theoretical Physics at the University of Cambridge, U.K.

Laura Mančinska is currently a research fellow in Stephanie Wehner's group at the Centre for Quantum Technologies, National University of Singapore. She received her PhD from the Combinatorics and Optimization department of the University of Waterloo, Canada where she studied under the guidance of Andrew Childs and Debbie Leung. In her PhD thesis she investigated a restricted model of quantum operations in which separated parties are only allowed to perform local quantum maps and communicate classically (LOCC). Her research interests also include other topics in quantum information such as nonlocal games, entanglement assisted zero-error communication etc.

David Roberson is currently a research fellow under Dmitrii Pasechnik at Nanyang Technological University. He received his PhD from the department of Combinatorics & Optimization at the University of Waterloo, Canada where he studied under the guidance of Chris Godsil. In his PhD thesis he introduced and investigated quantum homomorphisms, a class of non-local games which generalize quantum colorings. His research interests also include zero-error communication, graph parameters, and cores of graphs.

Simone Severini is a Reader (Associate Professor) in Physics of Information and a Royal Society University Research Fellow at UCL. Previously, he was a Newton International Fellow at UCL, a Postdoctoral Research Associate in the Institute for Quantum Computing at the University of Waterloo, and in the Department of Mathematics at the University of York. A visiting student at UC Berkeley, he obtained his PhD from Bristol University, under the supervision of Richard Jozsa. Among other institutions, he has been a long term visiting scientist at the Cesarea Rothschild Institute, MIT, and Shanghai Jiao-Tong University. His interests are at the interface between discrete mathematics, physics, and complex systems.

Dan Stahlke received an MS in Physics from the University of Alaska Fairbanks, in 2010, studying nonlinear dynamics under the supervision of Renate Wackerbauer, and a PhD in Physics from Carnegie Mellon University, in 2014, studying quantum information under the supervision of Robert B. Griffiths. From 2002-2010 he was a computer programmer at the Geographic Information Network of Alaska, University of Alaska Fairbanks. He will soon be a Rotation Engineer at Intel Corporation.

Andreas Winter received a Diploma degree in Mathematics from the Freie Universität Berlin, Berlin, Germany, in 1997, and a Ph.D. degree from the Fakultät für Mathematik, Universität Bielefeld, Bielefeld, Germany, in 1999. He was Research Associate at the University of Bielefeld until 2001, and then with the Department of Computer Science at the University of Bristol, Bristol, UK. In 2003, still with the University of Bristol, he was appointed Lecturer in Mathematics, and in 2006 Professor of Physics of Information. Since 2012 he is ICREA Research Professor with the Universitat Autònoma de Barcelona, Barcelona, Spain.