

Quantum advantage in zero-error function computation with side information

Ruoyu Meng and Aditya Ramamoorthy
 Department of Electrical and Computer Engineering
 Iowa State University, Ames, IA , USA
 Email: {rmeng, adityar}@iastate.edu

Abstract—We consider the problem of zero-error function computation with side information. Alice has a source X and Bob has correlated source Y and they can communicate via either classical or a quantum channel. Bob wants to calculate $f(X, Y)$ with zero error. We aim to characterize the minimum amount of information that Alice needs to send to Bob for this to happen with zero-error. In the classical setting, this quantity depends on the asymptotic growth of $\chi(G^{(m)})$, the chromatic number of an appropriately defined m -instance “confusion graph”. In this work we present structural characterizations of $G^{(m)}$ and demonstrate two function computation scenarios that have the same single-instance confusion graph. However, in one case there a strict advantage in using quantum transmission as against classical transmission, whereas there is no such advantage in the other case.

I. INTRODUCTION

In this work, we consider the problem of zero-error function computation with side information when there are two parties. This can be formally specified as follows. Alice observes a sequence of i.i.d. observations of a random variable X (taking values in a discrete alphabet \mathcal{X}). Bob has access to i.i.d. observations of a side information random variable Y (taking values in a discrete alphabet \mathcal{Y}). X and Y are correlated such that their joint probability mass function (p.m.f.) is $p_{X,Y}(x, y), x \in \mathcal{X}, y \in \mathcal{Y}$. Bob seeks to compute a function $f : \mathcal{X} \times \mathcal{Y} \rightarrow \mathcal{Z}$. The channel from Alice to Bob is assumed to be error-free and can support either classical or quantum transmission depending on the considered scenario. The aim is to understand the minimum rate at which Alice can communicate information to Bob such that the function computation is successful with probability-1, i.e., the probability of error is zero. The typical setting considers m -length blocks of the sources X_i, Y_i , for $i \in [m]$ where $[m] = \{1, \dots, m\}$, and the schemes allow the calculation of $f(X_i, Y_i)$ for $i \in [m]$. We consider m -length blocks, since it is often the case that computing multiple instances of the function at the same time allows Alice to send less information “per” computation than computing them one by one [1].

The classical version of this problem where Alice communicates classical bits has a long history [2]–[9], see [8] for an extensive survey. Witsenhausen [2] and Ferguson and Bailey [3] and Ahlswede [4] showed that for zero-error source coding with side information, the rate can be phrased in terms of graph-theoretic parameters of the appropriate graph-products. The optimal classical transmission strategy for Alice

is to color an appropriately defined confusion graph [2] and transmit the color to Bob. Alon and Orlitsky [5], Linial and Vazirani [6] exhibited examples in source coding such that the multiple-instance rate is less than the one-shot rate by a large amount. Korner and Orlitsky [8] noted that one of these examples combined with results in [10], [11] imply that for arbitrary $\chi, \epsilon > 0$, there exists a setting where the one-shot rate $\geq \log_2 \chi$ and the asymptotic rate $\leq 2 + \epsilon$. For general functions (not necessarily $f(X, Y) = X$) recent results appear in Charpenay’s PhD thesis [9]. Orlitsky and Roche [7] showed that a natural conditional entropy defined on graphs is the rate of asymptotically vanishing distortion (i.e., not a zero-error scenario) for the function computation with side information.

In this work, we consider general functions and a quantum variant where Alice can transmit quantum states to Bob. This version has received much less attention in the literature. Briët et al [12] and Stahlke [13] characterized the rate of zero-error source-channel coding with entanglement-assisted classical communication and quantum communication respectively; they did not consider general functions. In our problem setting, for function computation, Buhrman, Cleve, Wigderson [14], Gavinsky et al [15], Bar-Yossef, Jayram and Kenrenidis [16] showed exponential separation between the classical and quantum rates. However, these results are for the one-shot case, i.e., they do not consider the case of multiple instances. As pointed out in the previous paragraph, a quantum separation in one-shot case may not hold in multiple-instance case due to the classical rate’s discrepancy between one-shot and multiple-instance. The work that is most related to ours is by Gupta et al [17] who considered our setting and a specific function that demonstrates a quantum advantage. In contrast, we formulate the general version of the problem in terms of confusion graphs and present a substantially simpler example that demonstrates a quantum advantage.

Main contributions: The general m -instance confusion graph $G^{(m)}$ is sandwiched between the strong product ($G^{\boxtimes m}$) and OR-product ($G^{\vee n}$) of the single-instance graph (G). We provide necessary and sufficient conditions on the function and the joint p.m.f. such that $G^{(m)}$ equals $G^{\boxtimes m}$ or $G^{\vee n}$. We demonstrate example functions f and g and corresponding p.m.f.’s such that their single instance graph is the same. However, when considering m -instances, there is a strict quantum advantage in g while there is no quantum advantage for f .

II. PROBLEM FORMULATION

If Alice has symbols x and x' such that for any possible y that occurs with joint non-zero likelihood, i.e., $p_{XY}(x, y) > 0$ and $p_{XY}(x', y) > 0$, it holds that $f(x, y) = f(x', y)$ then from Bob's perspective, x and x' are equivalent and can be given the same description by Alice. Thus, Alice's symbols that need to be given "different" descriptions can be formalized by the following definition [2], [7].

Definition 1. *f*-confusion graph. The *f*-confusion graph of X given Y is a graph $G = (V, E)$ where $V = \mathcal{X}$ and $(x_1, x_2) \in E$ if there exists $y \in \mathcal{Y}$ such that $p_{XY}(x_1, y)p_{XY}(x_2, y) > 0$ and $f(x_1, y) \neq f(x_2, y)$.

Similarly, if we consider computations over m instances, then we can define the *f*-confusion graph over m instances denoted $G^{(m)}$ analogously. Let \mathbf{x} and \mathbf{y} denote m -length Alice and Bob sequences, respectively. The vertex set corresponds to all m -length Alice sequences; vertices $\mathbf{x}_1, \mathbf{x}_2 \in E(G^{(m)})$ if there exists \mathbf{y} such that $\prod_{i=1}^m p_{XY}(x_{1i}, y_i)p_{XY}(x_{2i}, y_i) > 0$ and there exists $j \in [m]$ such that $f(x_{1j}, y_j) \neq f(x_{2j}, y_j)$. We use shorthand $f^{(m)}(\mathbf{x}, \mathbf{y}) := \{f(x_i, y_i)\}_{i=1}^m$, $p_{\mathbf{X}\mathbf{Y}}(\mathbf{x}, \mathbf{y}) = \prod_{i=1}^m p_{XY}(x_i, y_i)$.

To better understand the notion of a *f*-confusion graph let us consider a scenario where X and Y take values in $\{0, 1, \dots, 4\}$ and be correlated such that

$$p_{XY}(x, y) = \begin{cases} \frac{1}{10}, & \text{if } y = x \text{ or } y = (x + 1) \bmod 5, \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

Suppose that Bob wishes to determine whether $X = Y$ or $X \neq Y$. We observe that Alice symbols 0 and 1 are connected since $p_{XY}(0, 1)p_{XY}(1, 1) = 1/100 > 0$ and $f(0, 1) \neq f(1, 1)$; the reasoning for the other edges is similar. It can be observed that this *f*-confusion graph is a pentagon (see Fig. 1).

Classical Setting: It can be observed that the Alice's strategy in the classical setting corresponds to coloring $G^{(m)}$ with the fewest possible colors and transmitting the color of her realization [2], [5], [8]. Thus, the rate in the classical setting appears as follows [2]–[4].

Definition 2. The rate in the classical setting is given by

$$R_{\text{classical}} = \lim_{m \rightarrow \infty} \frac{\log \chi(G^{(m)})}{m} = \inf_m \frac{\log \chi(G^{(m)})}{m}.$$

Quantum Setting: Suppose that Alice and Bob have m -length sequences \mathbf{x} and \mathbf{y} and Alice communicates to Bob through an error-free quantum channel that supports the transmission of operators on Hilbert space \mathcal{H} ; the space of operators is denoted as $\mathcal{L}(\mathcal{H})$. A quantum state $\rho \in \mathcal{L}(\mathcal{H})$ is a Hermitian, positive semi-definite, unit-trace operator. States ρ and σ are said to be orthogonal, denoted $\rho \perp \sigma$ if $\text{Tr}(\rho^\dagger \sigma) = 0$ ¹. This is equivalent to ρ and σ having orthogonal supports².

The quantum protocol operates as follows.

¹ \dagger stands for conjugate transpose

²The support of an operator is the orthogonal complement of its kernel. For Hermitian operators (as we consider) the support is its image.

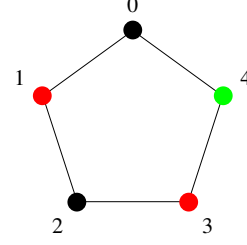


Fig. 1. The figure shows the confusion graph G of f with p.m.f. defined in (1). A coloring of G can be performed with three colors. This is shown in the figure with red, black and green. Thus, the rate can be $\log_2 3$. It can be shown that the graph over two instances $G^{(2)}$ can be colored with five colors. This leads to a per-computation rate of $\frac{1}{2} \log_2 5$. This is in fact optimal [2], [19].

- Alice picks a quantum state $\rho_{\mathbf{x}} \in \mathcal{L}(\mathcal{H})$, where $\dim \mathcal{H} = d$ that depends on the sequence \mathbf{x} and transmits it to Bob.
- Bob performs a POVM measurement [18, page 134] on the received state. The POVM is specified by $\{\Lambda_{f(\cdot, \mathbf{y})}^{\mathbf{z}}\}_{\mathbf{z} \in \mathcal{Z}^m}$, where the indices \mathbf{z} are measurement results and the matrices $\Lambda_{f(\cdot, \mathbf{y})}^{\mathbf{z}}$ are positive semi-definite matrices that sum to the identity, i.e., $\Lambda_{f(\cdot, \mathbf{y})}^{\mathbf{z}} \succeq 0$, for $\mathbf{z} \in \mathcal{Z}^m$, and $\sum_{\mathbf{z} \in \mathcal{Z}^m} \Lambda_{f(\cdot, \mathbf{y})}^{\mathbf{z}} = I$.

The protocol is deemed successful if $z_i = f(x_i, y_i)$, for $i \in \{1, \dots, m\}$ for all possible pairs (\mathbf{x}, \mathbf{y}) . Alice's rate of transmission is defined as $\frac{1}{m} \log_2 d$.

It can be shown (Section V), that this protocol is zero-error if and only if for $(\mathbf{x}, \mathbf{x}') \in G^{(m)}$ then $\rho_{\mathbf{x}} \perp \rho_{\mathbf{x}'}$. If this condition is satisfied, depending on the \mathbf{y} sequence observed by Bob, it can be shown that he can prepare the measurement that recovers the function value with zero error. Furthermore, Alice needs to find the suitable set of states in as small of a dimension d as possible, so that the task can be achieved by transmitting the fewest number of quantum bits. The quantum rate is defined formally in Section V (Theorem 4).

We point out that the classical protocol can be considered an instance of the proposed quantum protocol, simply by operating in a large enough vector space and labeling the nodes of the confusion graph by canonical basis vectors (binary vectors with all components zero except one). However, considering general unit-norm vectors provides much more flexibility and therefore the dimension d in which we need to operate in can be much lower.

III. PRELIMINARIES

For a graph G , \overline{G} denotes the complementary graph, $\alpha(G)$ its independence number, $\chi(G)$ its chromatic number. We say that G is a spanning subgraph of H , denoted by $G \subseteq H$ if $V(G) = V(H)$ and $E(G) \subseteq E(H)$. Likewise, graph union of two graphs $G \cup H$ is the graph whose vertex set is $V(G) \cup V(H)$ and edge set is $E(G) \cup E(H)$. We will deal extensively with graph products.

Definition 3. The strong product of two graphs, $G \boxtimes H$, has vertex set $V(G) \times V(H)$. $((u_1, v_1), (u_2, v_2)) \in E(G \boxtimes H)$ iff

$$\begin{aligned} & (u_1 = u_2 \text{ and } (v_1, v_2) \in E(H)) \text{ or} \\ & ((u_1, u_2) \in E(G) \text{ and } v_1 = v_2) \text{ or} \\ & ((u_1, u_2) \in E(G) \text{ and } (v_1, v_2) \in E(H)) \end{aligned}$$

The m -fold strong product is written $G^{\boxtimes m} := G \boxtimes G \boxtimes \dots \boxtimes G$.

The OR product of two graphs, $G \vee H$, has vertex set $V(G) \times V(H)$. $((u_1, v_1), (u_2, v_2)) \in E(G \vee H)$ iff

$$\begin{aligned} & (v_1, v_2) \in E(H) \text{ if } u_1 = u_2, \\ & (u_1, u_2) \in E(G) \text{ if } v_1 = v_2, \\ & (u_1, u_2) \in E(G) \text{ or } (v_1, v_2) \in E(H) \text{ if } u_1 \neq u_2, v_1 \neq v_2. \end{aligned}$$

The m -fold OR product is written $G^{\vee m} := G \vee G \vee \dots \vee G$.

We will use the following known results (proved in Appendix A for completeness). Let G, H, H_1, H_2 be simple graphs.

Proposition 1. $\overline{G \boxtimes H} = \overline{G} \vee \overline{H}$.

Proposition 2. $G \boxtimes (H_1 \cup H_2) = G \boxtimes H_1 \cup G \boxtimes H_2$.

Proposition 3. $\alpha(G \vee H) = \alpha(G)\alpha(H) \leq \alpha(G \boxtimes H)$.

Proposition 4. $\chi(G \boxtimes H) \leq \chi(G \vee H) \leq \chi(G)\chi(H)$.

IV. STRUCTURE OF CONFUSION GRAPH

We now demonstrate that the m -instance confusion graph depends strongly on the underlying joint pmf $p_{XY}(x, y)$ and the function $f(x, y)$. In particular, we demonstrate three scenarios, all of which have identical single instance confusion graphs but very different m -instance confusion graphs.

The examples appear in Table I-III. The functions are denoted f, g and h . The rows corresponding to y and columns corresponding to x . Entries represent the value of the function with input (x, y) and $*$ entries consists of all (x, y) such that $p_{X,Y}(x, y) = 0$. While we checked manually that $G_f = G_g = G_h = C_5$, we have distinct confusion graphs: $G_f^{(m)} = C_5^{\boxtimes m}, G_g^{(m)} = C_5^{\vee m}, C_5^{\boxtimes m} \neq G_h^{(m)} \neq C_5^{\vee m}$ by Theorem 1 that is shown later.

If there exists $x \in \mathcal{X}$ such that $p_{XY}(x, y) = 0$ for all $y \in \mathcal{Y}$, Alice may remove x from \mathcal{X} . Therefore, we assume the opposite holds for the rest of this paper.

Assumption 1. $\forall x \in \mathcal{X}, \exists y \in \mathcal{Y}$ s.t. $p_{XY}(x, y) > 0$.

For a confusion graph G , the following proposition holds.

Proposition 5. $G^{\boxtimes m} \subseteq G^{(m)} \subseteq G^{\vee m}$ for $m \geq 1$.

Proof. See Appendix B-A. \square

Consider two trivial cases. If G is edge-less, i.e. G is a set of isolated vertices, then $G^{\vee m}$ is edge-less and thus $G^{(m)}$ is edge-less. Similarly, if G is complete, then $G^{\boxtimes m}$ is complete and thus $G^{(m)}$ is complete. In both cases, we have $G^{\boxtimes} = G^{(m)} = G^{\vee m}$. If G is neither complete nor edge-less, we say G is nontrivial. We want to know conditions for $G^{(m)} = G^{\boxtimes m}$ and for $G^{(m)} = G^{\vee m}$ when G is nontrivial.

Let G be a nontrivial f -confusion graph. Then, for distinct $x, x' \in V(G)$, $(x, x') \notin E(G)$ can be due to either of following mutually exclusive conditions.

- $C1$: there is no $y \in \mathcal{Y}$ s.t. $p_{XY}(x, y)p_{XY}(x', y) > 0$.
- $C2$: $\exists y$ such that $p_{XY}(x, y)p_{XY}(x', y) > 0$ but for all such y , we have $f(x, y) = f(x', y)$.

Examples of $C1, C2$ above can be observed in the functions f, g, h from Table I-III. A manual check shows that all non-adjacent (x, x') s in Table I are due to $C1$. Similarly, all non-adjacent (x, x') s in Table II are due to $C2$. Differently, Table III contains both non-adjacent (x, x') of $C1$ and that of $C2$.

Theorem 1. Let $f : \mathcal{X} \times \mathcal{Y} \mapsto \mathcal{Z}$ be a function and $p_{XY}(x, y)$. Recall G and $G^{(m)}$ are f -confusion graphs over one-instance and m -instances respectively. Assume G is nontrivial.

- (a) $G^{(m)} = G^{\boxtimes m}$ for all $m \in \mathbb{N}_{\geq 1}$ iff all $(x, x') \notin E(G)$ is due to condition $C1$.
- (b) $G^{(m)} = G^{\vee m}$ for all $m \in \mathbb{N}_{\geq 1}$ iff all $(x, x') \notin E(G)$ is due to condition $C2$.

Proof. See Appendix B-B. \square

As a consequence of Theorem 1, given a finite simple graph G , we can always construct functions f and g s.t. the m -instance confusion graphs of f and g are $G^{\boxtimes m}$ and $G^{\vee m}$ respectively.

Theorem 2. Let G be a finite simple graph. Then

- (a) There exists a function f and a joint pmf $p_{XY}(x, y)$ s.t. the f -confusion graph over m -instances is $G^{\boxtimes m}$.
- (b) There exists a function g and a joint pmf $p_{XY}(x, y)$ s.t. the g -confusion graph over m -instances is $G^{\vee m}$.

Proof. See Appendix B-C. \square

V. ORTHOGONAL REPRESENTATION

We now discuss “orthogonal representations” of graphs.

Definition 4. Let G be a graph. An orthogonal representation³ of G is a mapping $\phi : V(G) \mapsto \mathbb{C}^m$ for some $m \geq 1$ s.t. each $\phi(v)$ is a unit-norm vector and **non-adjacent** vertices are assigned orthogonal vectors. The orthogonal rank of G , denoted by $\xi(G)$, is the minimum dimension m s.t. there exists an orthogonal representation of G .

We will use the following propositions of $\xi(G)$. The proof appears in Appendix C. Let G, H be finite simple graphs.

Proposition 6. If $G \subseteq H$, then $\xi(H) \leq \xi(G)$.

Proposition 7. $\xi(G \vee H) \leq \xi(G \boxtimes H) \leq \xi(G)\xi(H)$.

Proposition 8. $\xi(G) \leq \chi(\overline{G})$.

Next, we define the Lovász number over vector spaces in \mathbb{C} . This is a slightly different from [19] as they assume the field is \mathbb{R} , but these two definition are actually the same, which

³Our definition of the orthogonal representation aligns with Lovász’s number defined in [19]. The orthogonal representation is sometimes defined s.t. **adjacent** vertices are assigned orthogonal vectors. The other definition is related with ours by taking graph complement.

$f(x, y)$		x				
y	1	1	2	3	4	5
	1	1	0	*	*	*
	2	*	1	0	*	*
	3	*	*	1	0	*
	4	*	*	*	1	0
	5	0	*	*	*	1

TABLE I
FUNCTION f

$g(x, y)$		x				
y	1	1	2	3	4	5
	1	1	0	1	*	*
	2	*	1	0	1	*
	3	*	*	1	0	1
	4	1	*	*	1	0
	5	0	1	*	*	1

TABLE II
FUNCTION g

$h(x, y)$		x				
y	1	1	2	3	4	5
	1	1	0	1	*	*
	2	*	1	0	*	*
	3	*	*	1	0	*
	4	*	*	*	1	0
	5	0	*	*	*	1

TABLE III
FUNCTION h

we include a proof in Appendix D-A for completeness. For a complex number z , denote $|z| := \sqrt{z\bar{z}}$ where the overline is complex conjugate. For complex vectors $\mathbf{v}, \mathbf{u} \in \mathbb{C}^n$, denote $\langle \mathbf{v}, \mathbf{u} \rangle := \mathbf{v}^\dagger \mathbf{u}$ where \dagger stands for conjugate transpose.

Definition 5 (Lovász number). Let G be a finite simple graph. Its Lovász number $\vartheta(G)$ is

$$\vartheta(G) = \min_{\phi} \max_{c \in V(G)} \frac{1}{|\langle c, \phi(i) \rangle|^2}$$

where the minimum is taking over all orthogonal representation ϕ with field \mathbb{C} and complex unit-norm vector c , which is called the handle.

Proposition 9. If $G \subseteq H$, then $\vartheta(G) \geq \vartheta(H)$.

Proof. See Appendix D-B. \square

Lemma 1 (from [19]). $\alpha(G) \leq \vartheta(G)$.

Lemma 2 (from [19]). $\vartheta(G \vee H) = \vartheta(G \boxtimes H) = \vartheta(G)\vartheta(H)$.

In [19], Lovász proved that $\vartheta(G)\vartheta(H) \leq \vartheta(G \boxtimes H) \leq \vartheta(G)\vartheta(H)$. The first inequality is sufficient for him to compute the Shannon capacity of C_5 . Lovász noted that his proof of the first inequality can be strength to $\vartheta(G)\vartheta(H) \leq \vartheta(G \vee H)$. This can be used to show that $\vartheta(G)\vartheta(H) = \vartheta(G \vee H)$. We prove it in Appendix D-C for completeness.

Recall that in [19], the orthogonal representation is in \mathbb{R} , but the following lemma still holds in the case of orthogonal representation in \mathbb{C} . We include a proof for completeness.

Lemma 3. $\vartheta(G) \leq \xi(G)$.

Proof. See Appendix D-D. \square

Generally, it seems hard to know any information of an f -confusion graph over m -instances $G^{(m)}$. However, the following theorem allows us to calculate $\vartheta(G^{(m)})$ from $\vartheta(G)$.

Theorem 3. Let G be a f -confusion graph and $G^{(m)}$ be its f -confusion graph over m -instances. We have

$$\begin{aligned} \alpha(G^{\boxtimes m}) &\leq \vartheta(G^{\boxtimes m}) = \vartheta(G)^m = \\ \vartheta(G^{(m)}) &= \vartheta(G^{\vee m}) \leq \xi(G^{\vee m}). \end{aligned}$$

Proof. See Appendix E. \square

Lemma 4. Fix the block length m and function f and a pmf $p_{XY}(x, y)$. The optimal rate is $\frac{\log_2 \xi(\overline{G^{(m)}})}{m}$, where we recall that $\overline{G^{(m)}}$ denotes the complement of the f -confusion graph over m -instances $G^{(m)}$.

Proof. An orthogonal representation $\phi : \overline{G^{(m)}} \mapsto \mathbb{C}^{\xi(\overline{G^{(m)}})}$ induces a quantum protocol as follows. Suppose Alice and Bob get \mathbf{x} and \mathbf{y} respectively. Alice sends $|\phi(\mathbf{x})\rangle \langle \phi(\mathbf{x})|$ to Bob. Bob choose his POVM to be $\{\Pi^{\mathbf{z}} : \mathbf{z} \in \mathcal{Z}^m\}$ where $\Pi^{\mathbf{z}}$ is the projector onto $S_{\mathbf{z}} := \text{span}\{|\phi(\mathbf{x})\rangle : p_{\mathbf{X}\mathbf{Y}}(\mathbf{x}, \mathbf{y}) > 0, f^{(m)}(\mathbf{x}, \mathbf{y}) = \mathbf{z}\}$. Since ϕ is an orthogonal representation of \overline{G} , then $\Pi^{\mathbf{z}} \perp \Pi^{\mathbf{z}'}$ for all $\mathbf{z} \neq \mathbf{z}'$. If $\sum_{\mathbf{z}} \Pi^{\mathbf{z}} \neq I$, Bob adds $I - \sum_{\mathbf{z}} \Pi^{\mathbf{z}}$ to complete a POVM. The protocol has rate $\frac{1}{m} \log_2 \xi(\overline{G^{(m)}})$ as the orthogonal representation ϕ is $\xi(\overline{G^{(m)}})$ -dimensional.

Now, we show the protocol is zero-error. Let \mathbf{x} and \mathbf{y} be the same as the paragraph above. Denote $\mathbf{z} = f^{(m)}(\mathbf{x}, \mathbf{y})$. Since $|\phi(\mathbf{x})\rangle \in S_{\mathbf{z}}$ and $|\phi(\mathbf{x})\rangle \notin S_{\mathbf{z}'}$ for all $\mathbf{z} \neq \mathbf{z}'$, Bob's measurement result will be \mathbf{z} with probability 1.

Conversely, a protocol Π of rate $\frac{1}{m} \log_2 d$ induces an orthogonal representation $\psi : V(G^{(m)}) \mapsto \mathbb{C}^d$. Let $\rho_{\mathbf{x}}$ be a $d \times d$ density operator associated with input \mathbf{x} . If $\rho_{\mathbf{x}}$ is a pure state $|v\rangle \langle v|$ for some $|v\rangle \in \mathbb{C}^d$, then we set $\psi(\mathbf{x}) = |v\rangle$. If $\rho_{\mathbf{x}}$ is a mixed state with spectral decomposition $\rho_{\mathbf{x}} = \sum_{i=1}^k \lambda_i |v_i\rangle \langle v_i|$ with $\lambda_i \neq 0$ for $i = 1, \dots, k$. Then let $l \in [k]$ be arbitrary and $\psi(\mathbf{x}) = |v_l\rangle$.

Now we check it is indeed an orthogonal representation. Let $(\mathbf{x}, \mathbf{x}') \in E(G^{(m)})$. The protocol Π is zero-error, so $\rho_{\mathbf{x}} = \sum_{i=1}^k \lambda_i |v_i\rangle \langle v_i|, \rho_{\mathbf{x}'} = \sum_{i=1}^k \lambda'_i |v'_i\rangle \langle v'_i|$ must be perfectly distinguishable, i.e. $\rho_{\mathbf{x}} \perp \rho_{\mathbf{x}'}$. We have

$$\begin{aligned} \sum_{i=1}^k \lambda_i |v_i\rangle \langle v_i| \perp \sum_{i=1}^k \lambda'_i |v'_i\rangle \langle v'_i| &\Rightarrow |v_i\rangle \perp |v'_j\rangle \text{ for all } i, j \\ &\Rightarrow \psi(\mathbf{x}) \perp \psi(\mathbf{x}'). \end{aligned}$$

Since $(\mathbf{x}, \mathbf{x}') \in E(G)$ is arbitrary, it implies that ψ is an orthogonal representation. \square

Theorem 4. The rate in the quantum setting is given by

$$R_{\text{quantum}} = \inf_m \frac{\log_2 \xi(\overline{G^{(m)}})}{m}$$

Proof. It can be seen that $\xi(\overline{G^{(m+n)}}) \leq \xi(\overline{G^{(m)}})\xi(\overline{G^{(n)}})$, since a protocol for the $(m+n)$ -instance case can be obtained by putting together the m -instance and n -instance protocols. By Fekete lemma [20], the limit exists and is $\inf_m \frac{1}{m} \log \xi(\overline{G^{(m)}})$. \square

VI. DISCREPANCY OF $R_{\text{classical}}/R_{\text{quantum}}$ BETWEEN m -FOLD STRONG AND OR PRODUCTS

In this section, we consider the function f, g described in the Table I-II in detail. The m -instance confusion graphs for f

and g are $C_5^{\boxtimes m}$ and $C_5^{\vee m}$ respectively, while the one-instance confusion graphs for both are C_5 . We demonstrate that there is a quantum advantage for g while there is no quantum advantage for f . At various points in the discussion below, we use the well-known fact that C_5 is self-complementary.

Claim 1.

$$\xi(\overline{C_5}) = \chi(C_5) = 3.$$

Proof. See Appendix F-A. \square

Proposition 10.

$$R_{\text{quantum}}(f) = R_{\text{classical}}(f) = \frac{1}{2} \log_2 5.$$

Proof. See Appendix F-B. \square

Proposition 11.

$$R_{\text{quantum}}(g) = \frac{1}{2} \log_2 5 < \log_2 \frac{5}{2} \leq R_{\text{classical}}(g).$$

The rest of this section is the proof of Proposition 11. Let $G^{(m)}$ be the m -fold g -confusion graph. From Table I and Theorem 1, we know that $G^{(1)} = C_5$ and $G^{(m)} = C_5^{\vee m}$ respectively.

Claim 2.

$$\chi(C_5^{\vee m}) \geq \left(\frac{5}{2}\right)^m$$

Proof. See Appendix F-C. \square

Thus we have that $R_{\text{classical}}(g) = \inf_m \frac{1}{m} \log_2 \chi(C_5^{\vee m}) \geq \log_2 \frac{5}{2}$. Next, we will show that $R_{\text{quantum}}(g) = \frac{1}{2} \log_2 5$. We will use the following claim, whose proof is given after the proof of $R_{\text{quantum}}(g) = \frac{1}{2} \log_2 5$.

Claim 3.

$$\xi(\overline{C_5^{\vee 2}}) \leq 5.$$

Now we are ready to show $R_{\text{quantum}}(g) = \frac{1}{2} \log_2 5$, which completes the proof of Proposition 11.

Proof of $R_{\text{quantum}}(g) = \frac{1}{2} \log_2 5$. We have $\overline{C_5^{\vee m}} = \overline{C_5}^{\boxtimes m} = C_5^{\boxtimes m}$, and in particular $\overline{C_5^{\vee 2}} = C_5^{\boxtimes 2}$. Therefore, we have

$$\begin{aligned} \xi(\overline{C_5^{\vee m}}) &= \xi(C_5^{\boxtimes m}) \leq \begin{cases} \xi(C_5^{\boxtimes 2})^{m/2} & \text{if } m \text{ even} \\ \xi(C_5) \chi(C_5^{\boxtimes 2})^{(m-1)/2} & \text{if } m \text{ odd} \end{cases} \\ &\leq 3 \cdot 5^{(m-1)/2} \end{aligned} \quad (2)$$

where the first inequality follows from Proposition 7 and the second inequality follows from Claim 3, C_5 is self-complementary and Claim 1. A lower bound of $\xi(\overline{C_5^{\vee m}}) = \xi(C_5^{\boxtimes m})$ is given by

$$\xi(C_5^{\boxtimes m}) \geq \vartheta(C_5^{\boxtimes m}) = \vartheta(C_5)^m = 5^{m/2}. \quad (3)$$

where first inequality follows from Lemma 3, the first equality follows from Theorem 3, the second equality follows from $\vartheta(C_5) = \sqrt{5}$ [19].

Combining (2) and (3), we have

$$5^{m/2} \leq \xi(\overline{C_5^{\vee m}}) \leq 3 \cdot 5^{(m-1)/2}.$$

The result follows by taking logarithms, normalizing by m and taking the infimum. \square

Next, we prove Claim 3. For this we leverage the result of [21]. In order to develop this argument, we need some definitions that we now provide.

Definition 6. Let Γ be a group and $S \subseteq \Gamma$ be a subset of Γ . The Cayley graph $\text{Cay}(\Gamma, S)$ has vertex set $V = \Gamma$. $g, h \in \Gamma$, where $g \neq h$ are adjacent if we have

$$gh^{-1} \in S \text{ or } hg^{-1} \in S.$$

Definition 7. Let $\tilde{G} = \overline{\text{Cay}(\Gamma, S)}$ be the complement of a Cayley graph and $\phi : V(\tilde{G}) \mapsto \mathbb{C}^m$ be an orthogonal representation of \tilde{G} . ϕ is said to be “symmetric” if there exists $\tilde{f} : \Gamma \mapsto \mathbb{C}$ s.t. $\langle \phi(g), \phi(h) \rangle = \tilde{f}(gh^{-1})$ for all $g, h \in \Gamma$. The symmetric orthogonal rank $\xi_{\text{sym}}(\tilde{G})$ is the minimum dimension m such that there exists a symmetric orthogonal representation of \tilde{G} .

Since an orthogonal representation has less constraints than a symmetric one, we have

$$\xi(\tilde{G}) \leq \xi_{\text{sym}}(\tilde{G}). \quad (4)$$

Definition 8. Let $\mathbb{Z}_\ell = \{0, \dots, \ell - 1\}$ be the cyclic group of integers under addition modulo ℓ and $\mathbb{Z}_\ell^{\times n}$ be the direct product of n copies of \mathbb{Z}_ℓ . We define

$$H_\ell^n(d) = \text{Cay}(\mathbb{Z}_\ell^{\times n}, S) \quad (5)$$

where $S = \{x \in \mathbb{Z}_\ell^{\times n} : \sum_{i=1}^n x_i \geq d\}$ and the sum in the definition S is taken in \mathbb{N} .

Theorem 1 of [21] shows that For all positive integers ℓ, n and any $d \in [(\ell - 1)n]$ that is divisible by $\ell - 1$, we have $\xi_{\text{sym}}(H_\ell^n(d)) = \ell^{n - \frac{d}{\ell - 1}}$.

Proof of Claim 3. Since $C_5 = \text{Cay}(\mathbb{Z}_5, \{x \in \mathbb{Z}_5 : x \geq 4\})$, $C_5 \vee C_5$ can also be written as $C_5 \vee C_5 = \text{Cay}(\mathbb{Z}_5^{\times 2}, T)$ where $T = \{x \in \mathbb{Z}_5^{\times 2} : x_1 \geq 4 \text{ or } x_2 \geq 4\}$.

On the other hand, notice that $H_5^2(4) = \text{Cay}(\mathbb{Z}_5^{\times 2}, S)$ where $S = \{x \in \mathbb{Z}_5^{\times 2} : x_1 + x_2 \geq 4\}$. Since $x_1 \geq 4$ or $x_2 \geq 4$ implies $x_1 + x_2 \geq 4$, we have $E(C_5 \vee C_5) \subseteq E(H_5^2(4))$. Note $V(C_5 \vee C_5) = V(H_5^2(4)) = \mathbb{Z}_5^{\times 2}$. Thus, we conclude that

$$\overline{H_5^2(4)} \subseteq \overline{C_5 \vee C_5}. \quad (6)$$

Then we have

$$\xi(\overline{C_5 \vee C_5}) \leq \xi(\overline{H_5^2(4)}) \leq \xi_{\text{sym}}(\overline{H_5^2(4)}) = 5$$

where the first inequality follows from (6) and Proposition 6, the second inequality follows from (4) and the fact that $\overline{H_5^2(4)}$ is complement of a Cayley graph, the last equality follows from Theorem 1 of [21], i.e. $\xi_{\text{sym}}(H_5^2(4)) = 5^{2 - \frac{4}{4}} = 5$. \square

REFERENCES

- [1] C. Shannon, "The zero error capacity of a noisy channel," *IRE Transactions on Information Theory*, vol. 2, no. 3, pp. 8–19, 1956.
- [2] H. 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] M. Ferguson and D. Bailey, "Zero-error coding for correlated sources," 1975.
- [4] R. Ahlswede, "Coloring hypergraphs: A new approach to multi-user source coding," *Journal of Combinatorics*, vol. 4, no. 1, pp. 76–115, 1979.
- [5] N. Alon and A. Orlitsky, "Repeated communication and ramsey graphs," *IEEE Transactions on Information Theory*, vol. 41, no. 5, pp. 1276–1289, 1995.
- [6] N. Linial and U. Vazirani, "Graph products and chromatic numbers," in *30th Annual Symposium on Foundations of Computer Science*, 1989, pp. 124–128.
- [7] A. Orlitsky and J. Roche, "Coding for computing," *IEEE Transactions on Information Theory*, vol. 47, no. 3, pp. 903–917, 2001.
- [8] J. Körner and A. Orlitsky, "Zero-error information theory," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2207–2229, 1998.
- [9] N. Charpenay, "Zero-error network information theory: graphs, coding for computing and source-channel duality," Theses, Université de Rennes, Nov. 2023. [Online]. Available: <https://theses.hal.science/tel-04384481>
- [10] R. J. McEliece and E. C. Posner, "Hide and seek, data storage, and entropy," *The Annals of Mathematical Statistics*, vol. 42, no. 5, pp. 1706–1716, 1971.
- [11] C. Berge and M. Simonovits, "The coloring numbers of the direct product of two hypergraphs," in *Hypergraph Seminar*, C. Berge and D. Ray-Chaudhuri, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 1974, pp. 21–33.
- [12] J. Briët, H. Buhrman, M. Laurent, T. Piovosan, and G. Scarpa, "Entanglement-assisted zero-error source-channel coding," *IEEE Transactions on Information Theory*, vol. 61, no. 2, pp. 1124–1138, 2015.
- [13] D. Stahlke, "Quantum zero-error source-channel coding and non-commutative graph theory," *IEEE Transactions on Information Theory*, vol. 62, no. 1, pp. 554–577, 2016.
- [14] H. Buhrman, R. Cleve, and A. Wigderson, "Quantum vs. classical communication and computation," in *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing*, ser. STOC '98. New York, NY, USA: Association for Computing Machinery, 1998, p. 63–68. [Online]. Available: <https://doi.org/10.1145/276698.276713>
- [15] D. Gavinsky, J. Kempe, I. Kerenidis, R. Raz, and R. de Wolf, "Exponential separations for one-way quantum communication complexity, with applications to cryptography," in *Proceedings of the Thirty-Ninth Annual ACM Symposium on Theory of Computing*, ser. STOC '07. New York, NY, USA: Association for Computing Machinery, 2007, p. 516–525. [Online]. Available: <https://doi.org/10.1145/1250790.1250866>
- [16] Z. Bar-Yossef, T. S. Jayram, and I. Kerenidis, "Exponential separation of quantum and classical one-way communication complexity," in *Proceedings of the Thirty-Sixth Annual ACM Symposium on Theory of Computing*, ser. STOC '04. New York, NY, USA: Association for Computing Machinery, 2004, p. 128–137. [Online]. Available: <https://doi.org/10.1145/1007352.1007379>
- [17] S. Gupta, D. Saha, Z.-P. Xu, A. Cabello, and A. S. Majumdar, "Quantum contextuality provides communication complexity advantage," *Phys. Rev. Lett.*, vol. 130, p. 080802, Feb 2023. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.130.080802>
- [18] M. M. Wilde, *Quantum Information Theory*, 2nd ed. Cambridge University Press, 2017.
- [19] L. Lovász, "On the shannon capacity of a graph," *IEEE Transactions on Information Theory*, vol. 25, no. 1, pp. 1–7, 1979.
- [20] M. Fekete, "Über die verteilung der wurzeln bei gewissen algebraischen gleichungen mit ganzzahligen koeffizienten," *Mathematische Zeitschrift*, vol. 17, no. 1, pp. 228–249, Dec 1923. [Online]. Available: <https://doi.org/10.1007/BF01504345>
- [21] J. Briët and J. Zuiddam, "On the orthogonal rank of cayley graphs and impossibility of quantum round elimination," *Quantum Info. Comput.*, vol. 17, no. 1–2, p. 106–116, feb 2017.
- [22] D. B. West, *Introduction to Graph Theory*, 2nd edition. Prentice Hall, 2017.

APPENDIX A
PROOFS OF PROPOSITION 1-4

A. Proof of Proposition 1

Proof. We only need to show that $E(\overline{G \boxtimes H}) = E(\overline{G \vee H})$.

$$\begin{aligned} & ((v_1, u_1), (v_2, u_2)) \in E(\overline{G \boxtimes H}) \\ \Leftrightarrow & ((v_1, u_1), (v_2, u_2)) \notin E(G \boxtimes H) \\ \Leftrightarrow & (v_1, v_2) \notin E(G) \text{ or } (u_1, u_2) \notin E(H) \\ \Leftrightarrow & ((v_1, u_1), (v_2, u_2)) \in E(\overline{G \vee H}). \end{aligned}$$

We conclude that $E(\overline{G \boxtimes H}) = E(\overline{G \vee H})$. \square

B. Proof of Proposition 2

Proof.

$$\begin{aligned} V(G \boxtimes (H_1 \cup H_2)) &= V(G) \times V(H_1 \cup H_2) \\ &= V(G) \times V(H_1) \cup V(G) \times V(H_2). \end{aligned}$$

Let $u_1, u_2 \in V(G), v_1, v_2 \in V(H_1 \cup H_2)$ be arbitrary. Case $u_1 = u = u_2$ for some $u \in V(G)$:

$$\begin{aligned} & ((u, v_1), (u, v_2)) \in E(G \boxtimes (H_1 \cup H_2)) \\ \Leftrightarrow & (v_1, v_2) \in E(H_1 \cup H_2) \\ \Leftrightarrow & (v_1, v_2) \in E(H_1) \text{ or } (v_1, v_2) \in E(H_2) \\ \Leftrightarrow & ((u, v_1), (u, v_2)) \in E(G \boxtimes H_1) \\ & \text{or } ((u, v_1), (u, v_2)) \in E(G \boxtimes H_2). \end{aligned}$$

Case $v_1 = v = v_2$ for some $v \in V(H_1 \cup H_2)$:

$$\begin{aligned} & ((u_1, v), (u_2, v)) \in E(G \boxtimes (H_1 \cup H_2)) \\ \Leftrightarrow & (u_1, u_2) \in E(G) \\ \Leftrightarrow & ((u_1, v), (u_2, v)) \in E(G \boxtimes H_1) \\ & \text{or } ((u_1, v), (u_2, v)) \in E(G \boxtimes H_2). \end{aligned}$$

Case $u_1 \neq u_2, v_1 \neq v_2$:

$$\begin{aligned} & ((u_1, v_1), (u_2, v_2)) \in E(G \boxtimes (H_1 \cup H_2)) \\ \Leftrightarrow & (u_1, u_2) \in E(G) \text{ and } (v_1, v_2) \in E(H_1 \cup H_2) \\ \Leftrightarrow & (u_1, u_2) \in E(G) \\ & \text{and } ((v_1, v_2) \in E(H_1) \text{ or } (v_1, v_2) \in E(H_2)) \\ \Leftrightarrow & ((u_1, v_1), (u_2, v_2)) \in E(G \boxtimes H_1) \\ & \text{or } ((u_1, v_1), (u_2, v_2)) \in E(G \boxtimes H_2). \end{aligned}$$

It follows that $E(G \boxtimes (H_1 \cup H_2)) = E(G \boxtimes H_1) \cup E(G \boxtimes H_2)$. \square

C. Proof of Proposition 3

Proof. Let S_1 and S_2 be independent in G and H respectively. Then, $S_1 \times S_2$ is independent in $G \vee H$. Thus, $\alpha(G \vee H) \geq \alpha(G)\alpha(H)$. Now we show the other inequality. Let S be a maximum independent set in $G \vee H$. Define

$$\begin{aligned} S_g &:= \{g \in V(G) : \exists h \in V(H), (g, h) \in S\}, \\ S_h &:= \{h \in V(H) : \exists g \in V(G), (g, h) \in S\} \end{aligned}$$

The role of S_g, S_h is interchangeable, so it suffices to show S_g is independent in G . Let distinct $g_1, g_2 \in S_g$. Then there exist h_1, h_2 s.t. $(g_1, h_1), (g_2, h_2) \in S$. Since S is independent in $G \vee H$, $(g_1, h_1), (g_2, h_2) \notin E(G \vee H)$ and thus $(g_1, g_2) \notin E(G)$. It follows that S_g is independent. Therefore, $\alpha(G \vee H) \leq \alpha(G)\alpha(H)$ holds.

For $\alpha(G)\alpha(H) \leq \alpha(G \boxtimes H)$, we note that if S_1 and S_2 are independent in G and H respectively, then $S_1 \times S_2$ is independent in $G \boxtimes H$. Then, $\alpha(G \boxtimes H) \geq \alpha(G)\alpha(H)$ holds. \square

\square D. Proof of Proposition 4

Proof. $\chi(G \boxtimes H) \leq \chi(G \vee H)$ holds because $G \boxtimes H \subseteq G \vee H$.

If $c : V(G) \mapsto [\chi(G)]$ and $d : V(H) \mapsto [\chi(H)]$ are proper colorings of G and H respectively, then

$$\begin{aligned} c \times d : V(G) \times V(H) &\mapsto [\chi(G)] \times [\chi(H)], \\ (u, v) &\mapsto (c(u), d(v)) \end{aligned}$$

is a proper coloring of $G \vee H$. Therefore, $\chi(G \vee H) \leq \chi(G)\chi(H)$. \square

APPENDIX B
PROOFS FROM SECTION VI

A. Proof of Proposition 5

Proof. Since the vertex sets of $G^{\boxtimes m}, G^{(m)}, G^{\vee m}$ are the same, that is \mathcal{X}^m , it suffices to show $E(G^{\boxtimes m}) \subseteq E(G^{(m)}) \subseteq E(G^{\vee m})$.

Let $(\mathbf{x}, \mathbf{x}') \in E(G^{\boxtimes m})$ and $\mathbf{x} \neq \mathbf{x}'$. There exists i s.t. $x_i \neq x'_i$ because $\mathbf{x} \neq \mathbf{x}'$. For such i indices, $(x_i, x'_i) \in E(G)$, so there exists y_i s.t. $p_{XY}(x_i, y_i)p_{XY}(x'_i, y_i) > 0$ and $f(x_i, y_i) \neq f(x'_i, y_i)$. For the remaining i indices s.t. $x_i = x'_i$, there exists y_i s.t. $p_{XY}(x_i, y_i) > 0$ by Assumption 1. Let $\mathbf{y} = \{y_i\}_{i=1}^m$. Then we have

$$p_{\mathbf{X}\mathbf{Y}}(\mathbf{x}, \mathbf{y})p_{\mathbf{X}\mathbf{Y}}(\mathbf{x}', \mathbf{y}) > 0 \text{ and } f^{(m)}(\mathbf{x}, \mathbf{y}) \neq f^{(m)}(\mathbf{x}', \mathbf{y}).$$

Therefore, $(\mathbf{x}, \mathbf{x}') \in E(G^{(m)})$. It implies $E(G^{\boxtimes m}) \subseteq E(G^{(m)})$ as $(\mathbf{x}, \mathbf{x}') \in E(G^{\boxtimes m})$ is arbitrary.

Let $(\mathbf{x}, \mathbf{x}') \in E(G^{(m)})$ be arbitrary. Then, there exists \mathbf{y} s.t. $p_{\mathbf{X}\mathbf{Y}}(\mathbf{x}, \mathbf{y})p_{\mathbf{X}\mathbf{Y}}(\mathbf{x}', \mathbf{y}) > 0$ and $f^{(m)}(\mathbf{x}, \mathbf{y}) \neq f^{(m)}(\mathbf{x}', \mathbf{y})$. It implies that there exists $j \in [m]$ such that $p_{XY}(x_j, y_j)p_{XY}(x'_j, y_j) > 0$ and $f(x_j, y_j) \neq f(x'_j, y_j)$, i.e. $(x_j, x'_j) \in E(G)$. This implies $(\mathbf{x}, \mathbf{x}') \in E(G^{\vee m})$. Thus, $E(G^{(m)}) \subseteq E(G^{\vee m})$ as $(\mathbf{x}, \mathbf{x}') \in E(G^{(m)})$ are arbitrary. \square

B. Proof of Theorem 1

Proof of (a). \Rightarrow : We show that if there exists $(x, x') \notin E(G)$ s.t. it is due to condition C2, then $G^{(m)} \neq G^{\boxtimes m}$. In particular, we show $G^{(2)} \neq G^{\boxtimes 2}$ by showing $E(G^{(2)}) \neq E(G^{\boxtimes 2})$.

Let $(x_1, x'_1) \notin E(G)$ due to condition C2 and $(x_2, x'_2) \in E(G)$. Such (x_2, x'_2) exists as G is nontrivial. Denote $\mathbf{x} = [x_1, x_2], \mathbf{x}' = [x'_1, x'_2]$. $(\mathbf{x}, \mathbf{x}') \notin E(G^{\boxtimes 2})$ as $(x_1, x'_1) \notin E(G)$.

By condition C2, there exists y_1 s.t.

$$p_{XY}(x_1, y_1)p_{XY}(x'_1, y_1) > 0 \text{ and } f(x_1, y_1) = f(x'_1, y_1).$$

$(x_2, x'_2) \in E(G)$ implies the existence of y_2 s.t.

$$p_{XY}(x_2, y_2)p_{XY}(x'_2, y_2) > 0 \text{ and } f(x_2, y_2) \neq f(x'_2, y_2).$$

Denote $\mathbf{y} = [y_1, y_2]$. Then it follows that

$$\prod_{i=1}^2 p_{XY}(x_i, y_i)p_{XY}(x'_i, y_i) > 0 \text{ and } f^{(2)}(\mathbf{x}, \mathbf{y}) \neq f^{(2)}(\mathbf{x}', \mathbf{y}),$$

because $f(x_2, y_2) \neq f(x'_2, y_2)$. Therefore, we have $(\mathbf{x}, \mathbf{x}') \in E(G^{(2)})$, which implies $E(G^{(2)}) \neq E(G^{\boxtimes 2})$.

\Leftarrow : By Proposition 5 and the fact that $V(G^{\vee m}) = V(G^{\boxtimes m})$, it suffices to show $E(G^{(m)}) \subseteq E(G^{\boxtimes m})$.

Now let $(\mathbf{x}, \mathbf{x}') \notin E(G^{\boxtimes m})$ be arbitrary. Then, there is i s.t. $(x_i, x'_i) \notin E(G)$. Since all $(x, x') \notin E(G)$ are due to condition C1, there does not exist y_i s.t. $p_{XY}(x_i, y_i)p_{XY}(x'_i, y_i) > 0$, which implies $p_{\mathbf{X}\mathbf{Y}}(\mathbf{x}, \mathbf{y})p_{\mathbf{X}\mathbf{Y}}(\mathbf{x}', \mathbf{y}) = 0$ for all choices of \mathbf{y} . Thus, we have $(\mathbf{x}, \mathbf{x}') \notin E(G^{(m)})$. Since $(\mathbf{x}, \mathbf{x}') \notin E(G^{\boxtimes m})$ is arbitrary, this implies $E(G^{(m)}) \subseteq E(G^{\boxtimes m})$. \square

Proof of (b). \Rightarrow : Since G is nontrivial, there exists $(x_1, x'_1) \in E(G)$ and $(x_2, x'_2) \notin E(G)$. Denote $\mathbf{x} = [x_1, x_2]$, $\mathbf{x}' = [x'_1, x'_2]$ and note $(\mathbf{x}, \mathbf{x}') \in G^{\vee 2}$. By the assumption that $G^{\vee m} = G^{(m)}$ for $m \geq 1$ and setting $m = 2$, there exists $\mathbf{y} = [y_1, y_2]$ s.t. $p_{\mathbf{X}\mathbf{Y}}(\mathbf{x}, \mathbf{y})p_{\mathbf{X}\mathbf{Y}}(\mathbf{x}', \mathbf{y}) > 0$, which implies $p_{XY}(x_2, y_2)p_{XY}(x'_2, y_2) > 0$. Therefore, $(x_2, x'_2) \notin E(G)$ is not due to condition C1.

\Leftarrow : By Proposition 5 and the fact $V(G^{\vee m}) = V(G^{\boxtimes m})$, it suffices to show $E(G^{\vee m}) \subseteq E(G^{(m)})$. Let $(\mathbf{x}, \mathbf{x}') \in E(G^{\vee m})$. Since role of (x_i, x'_i) s are interchangeable, we may assume \mathbf{x}, \mathbf{x}' is s.t.

$$\begin{cases} (x_i, x'_i) \in E(G) \text{ for } i \in \{1, \dots, k_1\} \\ (x_i, x'_i) \notin E(G) \text{ for } i \in \{k_1 + 1, \dots, k_2\} \\ x_i = x'_i \text{ for } i \in \{k_2 + 1, \dots, m\} \end{cases}$$

for some $1 \leq k_1 \leq k_2 \leq m$. To show $(\mathbf{x}, \mathbf{x}') \in E(G^{(m)})$, it suffices to find $\mathbf{y} = [y_1, \dots, y_m]$ s.t. $p_{\mathbf{X}\mathbf{Y}}(\mathbf{x}, \mathbf{y})p_{\mathbf{X}\mathbf{Y}}(\mathbf{x}', \mathbf{y}) > 0$ and $f^{(m)}(\mathbf{x}, \mathbf{y}) \neq f^{(m)}(\mathbf{x}', \mathbf{y})$.

For $i \in \{1, \dots, k_1\}$, choose y_i be s.t. $p_{XY}(x_i, y_i)p_{XY}(x'_i, y_i) > 0$ and $f(x_i, y_i) \neq f(x'_i, y_i)$. Such y_i exists as $(x_i, x'_i) \in E(G)$. For $i \in \{k_1 + 1, \dots, k_2\}$, choose y_i be s.t. $p_{XY}(x_i, y_i)p_{XY}(x'_i, y_i) > 0$ and $f(x_i, y_i) = f(x'_i, y_i)$. Such y_i exists because $(x_i, x'_i) \notin E(G)$ and because all $(x, x') \notin E(G)$ is due to condition C2. For $i \in \{k_2 + 1, \dots, m\}$, choose y_i be s.t. $p_{XY}(x'_i, y_i) = p_{XY}(x_i, y_i) > 0$. Such y_i by $p_{XY}(x, y)$ satisfies Assumption 1. Then, $p_{\mathbf{X}\mathbf{Y}}(\mathbf{x}, \mathbf{y})p_{\mathbf{X}\mathbf{Y}}(\mathbf{x}', \mathbf{y}) = \prod_{i=1}^m p_{XY}(x_i, y_i)p_{XY}(x'_i, y_i) > 0$. We have $f^{(m)}(\mathbf{x}, \mathbf{y}) \neq f^{(m)}(\mathbf{x}', \mathbf{y})$ because $f(x_i, y_i) \neq f(x'_i, y_i)$ for $i \in \{1, \dots, k_1\}$.

Therefore, $(\mathbf{x}, \mathbf{x}') \in E(G^{(m)})$ holds. Since $(\mathbf{x}, \mathbf{x}') \in E(G^{\vee m})$ is arbitrary, we have $E(G^{\vee m}) \subseteq E(G^{(m)})$. \square

C. Proof of Theorem 2

Proof of (a). We construct the function f as follows. Let $\mathcal{X} = V(G)$ and $\mathcal{Y} = E(G)$. Choose any distribution p_{XY} s.t. $p_{XY}(x, y) > 0$ if and only if vertex x is incident with edge y . Then, we set $f(x, y) = x$. Denote f -confusion graph by G_f .

For $x \neq x' \in \mathcal{X}$, $(x, x') \notin E(G_f)$ if and only if there does not exist $y \in \mathcal{Y}$ s.t. $p_{XY}(x, y)p_{XY}(x', y) > 0$ if and only if $(x, x') \notin E(G)$. Therefore, $G_f = G$ and all $(x, x') \notin E(G_f)$ is due to C1 as $x \neq x' \in \mathcal{X}$ are arbitrary. By Theorem 1, we have that $G^{(m)} = G^{\boxtimes m}$ for $m \geq 1$. \square

Proof of (b). We construct the function g as follows. Let $\mathcal{X} = V(G)$ and $\mathcal{Y} = \{\{i, j\} : i \neq j \in V(G)\}$. Choose any distribution p_{XY} s.t. $p_{XY}(x, y) > 0$ if and only if vertex x is an element of set y . The function evaluation is s.t.

$$g(x, y) = \begin{cases} x, & \text{if } y \in E(G), \\ 1, & \text{if } y \notin E(G). \end{cases}$$

Denote g -confusion graph by G_g . Now we show that $G_g = G$ is g -confusion graph. For each $x \neq x' \in \mathcal{X}$, there is exactly one choice of y , that is $\{x, x'\}$, s.t. $p_{XY}(x, y)p_{XY}(x', y) > 0$. $g(x, \{x, x'\}) = g(x', \{x, x'\})$ if and only if $\{x, x'\} \notin E(G)$. This implies $G_g = G$ and all $(x, x') \notin E(G_g)$ is due to C2 as $x \neq x' \in \mathcal{X}$ are arbitrary. By Theorem 1, we have that $G^{(m)} = G^{\vee m}$ for $m \geq 1$. \square

APPENDIX C

PROOFS OF PROPOSITION 6-8

A. Proof of Proposition 6

Proof. Let ϕ be an orthogonal representation of G . $V(G) = V(H)$ as $G \subseteq H$. Since $V(G) = V(H)$, every vertex in $V(H)$ is assigned a vector. Let $u, v \in V(H)$ be non-adjacent in H . Since $G \subseteq H$, u, v is also non-adjacent in G . Thus $\phi(u) \perp \phi(v)$. Therefore, ϕ is also an orthogonal representation of H . This implies $\xi(H) \leq \xi(G)$. \square

B. Proof of Proposition 7

Proof. $\xi(G \vee H) \leq \xi(G \boxtimes H)$ holds because of Proposition 6 and the fact that $G \boxtimes H \subseteq G \vee H$.

Now we show that $\xi(G \boxtimes H) \leq \xi(G)\xi(H)$. Suppose $\phi_G : V(G) \mapsto \mathbb{C}^{m_1}$, $\phi_H : V(H) \mapsto \mathbb{C}^{m_2}$ are orthogonal representations of G, H respectively. We claim

$$\phi_G \times \phi_H : V(G) \times V(H) \mapsto \mathbb{C}^{m_1 m_2}, (g, h) \mapsto \phi_G(g) \otimes \phi_H(h)$$

where \otimes denotes tensor product, is an orthogonal representation of $G \boxtimes H$.

Indeed, let $(g_1, h_1), (g_2, h_2)$ be non-adjacent in $G \boxtimes H$. Since the role of G, H are interchangeable, we assume $(g_1, g_2) \notin E(G)$. Thus $\phi_G(g_1) \perp \phi_G(g_2)$. It follows that $\phi_G \times \phi_H(g_1, h_1) = \phi_G(g_1) \otimes \phi_H(h_1) \perp \phi_G(g_2) \otimes \phi_H(h_2) = \phi_G \times \phi_H(g_2, h_2)$. \square

C. Proof of Proposition 8

Proof. Let $c : V(G) \mapsto [k]$ be a proper coloring of \overline{G} and e_i be the i -th elementary vector in \mathbb{C}^k for $i \in [k]$. Then c induces the following k -dimensional orthogonal representation

$$\phi : V(G) \mapsto \mathbb{C}^k, \phi(v) \mapsto e_i \text{ iff } c(v) = i.$$

□

APPENDIX D

DISCUSSION ON LOVÁSZ NUMBER

A. Two definitions of Lovász number are the same

Most of our proof is the same as those in [19] except that we use conjugate transpose. We include them here for completeness.

In this subsection, let $\vartheta_C(G)$ and $\vartheta_R(G)$ be the Lovász number defined on orthogonal representation over vector spaces in \mathbb{C} and \mathbb{R} respectively. We use the shorthand PSD matrix for positive semidefinite matrix.

Since $\mathbb{R} \subseteq \mathbb{C}$, the feasible set in the minimization problem defining $\vartheta_R(G)$ is a subset of the feasible set in the minimization problem defining $\vartheta_C(G)$. Therefore, we have $\vartheta_R(G) \geq \vartheta_C(G)$. Now it suffices to show $\vartheta_R(G) \leq \vartheta_C(G)$.

Theorem 5. Let G be a graph on vertices $[n]$. Then $\vartheta_C(G)$ is minimum of the largest eigenvalue of any Hermitian matrix $(a_{ij})_{i,j=1}^n$ s.t.

$$a_{ij} = 1, \text{ if } i = j \text{ or } i, j \text{ are non-adjacent} \quad (7)$$

Proof. Let c, ϕ be s.t. the optimal handle and orthogonal representation that achieves $\vartheta_C(G)$. Define

$$\begin{aligned} a_{ij} &= 1 - \frac{\langle u_i, u_j \rangle}{\langle c, u_i \rangle \langle c, u_j \rangle}, \quad i \neq j \\ a_{ii} &= 1, \\ A &= (a_{ij})_{i,j=1}^n. \end{aligned}$$

Then (7) is satisfied and

$$\begin{aligned} -a_{ij} &= \left(c - \frac{u_i}{\langle c, u_i \rangle} \right)^\dagger \left(c - \frac{u_j}{\langle c, u_j \rangle} \right), \\ \vartheta_C(G) - a_{ii} &= \left(c - \frac{u_i}{\langle c, u_i \rangle} \right)^\dagger \left(c - \frac{u_i}{\langle c, u_i \rangle} \right) \\ &\quad + \left(\vartheta_C(G) - \frac{1}{|\langle c, u_i \rangle|^2} \right). \end{aligned}$$

These implies that $\vartheta_C(G)I - A$ is Hermitian PSD and hence the largest eigenvalue of A is at most $\vartheta_C(G)$.

Conversely, let $A = (a_{ij})_{i,j=1}^n$ be any Hermitian matrix satisfying (7) and λ be its largest eigenvalue. Then $\lambda I - A$ is Hermitian, positive semidefinite. Hence there exist vectors $\mathbf{x}_1, \dots, \mathbf{x}_n$ s.t.

$$\lambda \delta_{ij} - a_{ij} = \mathbf{x}_i^\dagger \mathbf{x}_j.$$

Let c be a unit vector orthogonal to $\mathbf{x}_1, \dots, \mathbf{x}_n$ and set

$$u_i = \frac{1}{\sqrt{\lambda}}(c + \mathbf{x}_i).$$

Then

$$\begin{aligned} u_i^\dagger u_i &= \frac{1}{\lambda}(c^\dagger c + \mathbf{x}_i^\dagger \mathbf{x}_i) = 1, \\ u_i^\dagger u_j &= \frac{1}{\lambda}(1 + \mathbf{x}_i^\dagger \mathbf{x}_j) = 0 \text{ if } (i, j) \notin E(G). \end{aligned}$$

Then u_1, \dots, u_n is an orthogonal representation of G and

$$\lambda = \frac{1}{|\langle c, u_i \rangle|^2}, \forall i \in [n].$$

This completes the proof. □

Lemma 5. Let G be a graph on vertices $[n]$, and $B = (b_{ij})_{i,j=1}^n$ range over all Hermitian PSD matrices s.t.

$$b_{ij} = 0, \text{ for all } (i, j) \in E(G), \text{ and } \text{Tr}(B) = 1.$$

Then

$$\vartheta_C(G) \geq \max_B \text{Tr } B J.$$

Proof. Let $A = (a_{ij})_{i,j=1}^n$ be a matrix satisfying (7) with largest eigenvalue $\vartheta_C(G)$ and B be any Hermitian PSD matrix satisfying the above constraints. Then, using (7),

$$\text{Tr } B J = \sum_{i,j=1}^n b_{ij} = \sum_{i,j=1}^n a_{ij} b_{ij} = \text{Tr } A B$$

and so

$$\vartheta_C(G) - \text{Tr } B J = \text{Tr } (\vartheta_C(G)I - A)B.$$

Here, both $\vartheta_C(G)I - A$ and B are PSD. Let e_1, \dots, e_n be a set of mutually orthogonal eigenvectors of B with corresponding eigenvalues $\lambda_1, \dots, \lambda_n \geq 0$. Then

$$\begin{aligned} \text{Tr } (\vartheta_C(G)I - A)B &= \sum_{i=1}^n e_i^\dagger (\vartheta_C(G)I - A)B e_i \\ &= \sum_{i=1}^n \lambda_i e_i^\dagger (\vartheta_C(G)I - A) e_i \geq 0 \end{aligned}$$

□

$\vartheta_R(G)$ has an equivalent definition as follows.

Theorem 6 (from [19]). Let G be a graph on vertices $[n]$, and $B = (b_{ij})_{i,j=1}^n$ range over all symmetric PSD matrices, which consists of Hermitian PSD matrices with real entries, s.t.

$$b_{ij} = 0, \text{ for all } (i, j) \in E(G), \text{ and } \text{Tr}(B) = 1.$$

Then

$$\vartheta_R(G) = \max_B \text{Tr } B J.$$

Since $n \times n$ symmetric PSD is already Hermitian, we have that

$$\begin{aligned} \vartheta_C(G) &\geq \max_{B \text{ is Hermitian PSD}} \text{Tr } B J \\ &\geq \max_{B \text{ is symmetric PSD}} \text{Tr } B J = \vartheta_R(G). \end{aligned}$$

Since an optimal point achieving $\vartheta_R(G)$ is an optimal point achieving $\vartheta_C(G)$, $\vartheta_C(G)$ can be always achieved by an orthogonal representation ϕ s.t. all vectors are real and real vector c .

B. Proof of Proposition 9

Proof. Since $G \subseteq H$, we have $V(G) = V(H)$ and that if $(u, v) \notin E(H)$, then $(u, v) \notin E(G)$. Therefore, an orthogonal representation ϕ of G is an orthogonal representation of H . Let ψ and d be optimal orthogonal representation and handle that achieves $\vartheta(G)$. Then, we have

$$\begin{aligned}\vartheta(G) &= \max_{i \in V(G)} \frac{1}{|\langle d, \psi(i) \rangle|^2} \\ &\geq \min_{\phi, c} \max_{i \in V(H)} \frac{1}{|\langle c, \phi(i) \rangle|^2} = \vartheta(H).\end{aligned}$$

□

C. Proof of $\vartheta(G)\vartheta(H) = \vartheta(G \vee H)$

From Theorem 7 of [19], we have that $\vartheta(G \boxtimes H) = \vartheta(G)\vartheta(H)$. Since $G \boxtimes H \subseteq G \vee H$, we have $\vartheta(G \vee H) \leq \vartheta(G \boxtimes H) = \vartheta(G)\vartheta(H)$ by Proposition 9. Therefore, it suffices to show

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

We will use the following equivalent definition of $\vartheta(G)$.

Theorem 7 (from [19]). Let G be a graph on vertices $[n]$. Let ϕ range over all orthogonal representation over \overline{G} be s.t. $\phi(i)$ is a real vector for all $i \in V(G)$ and d range over all real unit-norm vectors. Then

$$\vartheta(G) = \max_{\phi, d} \sum_{i=1}^n (d^T \phi(i))^2 \quad (8)$$

The proof is mostly the same as the one in Theorem 7 of [19]. We include it for completeness.

Proof of $\vartheta(G)\vartheta(H) \leq \vartheta(G \vee H)$. Let ϕ and c be an orthogonal representation of G and a vector in (8) that achieves $\vartheta(G)$, and ψ and d be an orthogonal representation of H and a vector in (8) that achieves $\vartheta(H)$. Using the same argument as Proposition 7, we have that $\phi \times \psi(i, j) := \phi(i) \otimes \psi(j)$ is an orthogonal representation of $\overline{G \vee H}$. Since $G \vee H = \overline{G \boxtimes H}$, we have

$$\begin{aligned}\vartheta(G \vee H) &\geq \sum_{i=1}^{|V(G)|} \sum_{j=1}^{|V(H)|} \left((d \otimes c)^T (u_i \otimes v_j) \right)^2 \\ &= \sum_{i=1}^{|V(G)|} \sum_{j=1}^{|V(H)|} (d^T u_i)^2 (c^T v_j)^2 \\ &= \sum_{i=1}^{|V(G)|} (d^T u_i)^2 \sum_{j=1}^{|V(H)|} (c^T v_j)^2 = \vartheta(G)\vartheta(H)\end{aligned}$$

□

D. Proof of Lemma 3

Proof. Let ϕ be a complex orthogonal representation of dimension $d = \xi(G)$. Then, $\psi : V(G) \mapsto \mathbb{C}^d$, $\psi(i) = \phi(i) \otimes \overline{\phi(i)}$ is also an orthogonal representation of G where $\overline{\phi(i)}$ is complex conjugate of $\phi(i)$. Let e_1, \dots, e_d be the canonical

basis (binary vectors with all components zero except one) and $b = \frac{1}{\sqrt{d}}(e_1 \otimes e_1 + \dots + e_d \otimes e_d)$. Then $\|b\|_2^2 = 1$ and

$$\begin{aligned}\langle \psi(i), b \rangle &= \frac{1}{\sqrt{d}} \sum_{i=1}^d \langle \phi(i) \otimes \overline{\phi(i)}, e_i \otimes e_i \rangle \\ &= \frac{1}{\sqrt{d}} \sum_{i=1}^d \langle \phi(i), e_i \rangle \langle \overline{\phi(i)}, e_i \rangle = \frac{1}{\sqrt{d}} \|\phi(i)\|_2^2 = \frac{1}{\sqrt{d}}\end{aligned}$$

Therefore, $\vartheta(G) \leq d$. □

APPENDIX E

PROOFS OF THEOREM 3

Proof. $\alpha(G^{\boxtimes m}) \leq \vartheta(G^{\boxtimes m})$ follows from Lemma 1. The equality $\vartheta(G^{\vee m}) = \vartheta(G^{\boxtimes m}) = \vartheta(G)^m$ holds by Lemma 2. Since $G^{\boxtimes m} \subseteq G^{(m)} \subseteq G^{\vee m}$ by Proposition 5, we have

$$\vartheta(G)^m = \vartheta(G^{\boxtimes m}) \geq \vartheta(G^{(m)}) \geq \vartheta(G^{\vee m}) = \vartheta(G)^m$$

by Proposition 9. The inequality $\vartheta(G^{\vee m}) \leq \xi(G^{\vee m})$ follows from Lemma 3. □

APPENDIX F

PROOFS OF SECTION VIII

A. Proof of Claim 1

Proof. By Proposition 6, we have $\xi(\overline{C_5}) \leq \chi(C_5) = 3$. Moreover, we know that $\overline{C_5}$ is isomorphic to C_5 and furthermore that $\vartheta(C_5) \leq \xi(C_5)$. It is well known that $\vartheta(C_5) = \sqrt{5}$ [19]. Since $\xi(C_5)$ is an integer, we must have $\xi(C_5) = 3$. □

B. Proof of Proposition 10

Proof. Note $G^{(m)} = C_5^{\boxtimes m}$. We have that

$$\xi(\overline{C_5^{\boxtimes m}}) \geq \vartheta(\overline{C_5^{\boxtimes m}}) = \vartheta(C_5^{\vee m}) = \vartheta(C_5)^m = 5^{m/2}. \quad (9)$$

The first inequality follows from Lemma 3. The first equality holds because $\overline{C_5^{\boxtimes m}} = \overline{C_5}^{\vee m} = C_5^{\vee m}$. The second equality holds by Theorem 3. The last equality holds as $\vartheta(C_5) = \sqrt{5}$ [19].

Consider $G^{(2)} = C_5^{\boxtimes 2}$. We know $\chi(C_5^{\boxtimes 2}) = 5$ by [2]. Therefore, for $m \geq 1$, we have

$$\begin{aligned}\chi(C_5^{\boxtimes m}) &\leq \begin{cases} \chi(C_5^{\boxtimes 2})^{m/2} & \text{if } m \text{ even} \\ \chi(C_5) \chi(C_5^{\boxtimes 2})^{(m-1)/2} & \text{if } m \text{ odd} \end{cases} \\ &\leq 3 \cdot 5^{(m-1)/2}\end{aligned} \quad (10)$$

where the first inequality holds by Proposition 4 and the second inequality holds by $\chi(C_5^{\boxtimes 2}) = 5$ and $\chi(C_5) = 3$. By Proposition 8, we have $\xi(\overline{C_5^{\boxtimes m}}) \leq \chi(C_5^{\boxtimes m})$. This combined with bounds in (9) and (10) gives

$$5^{m/2} \leq \xi(\overline{C_5^{\boxtimes m}}) \leq \chi(C_5^{\boxtimes m}) \leq 3 \cdot 5^{(m-1)/2}.$$

Apply log and then multiply by $\frac{1}{m}$ on each part of the inequality. We have

$$\begin{aligned}\frac{1}{2} \log_2 5 &\leq \frac{1}{m} \log_2 \xi(\overline{C_5^{\boxtimes m}}) \leq \frac{1}{m} \log_2 \chi(C_5^{\boxtimes m}) \\ &\leq \frac{1}{m} \log_2 3 + \frac{m-1}{2m} \log_2 5.\end{aligned}$$

Taking infimum over m and we have $R_{\text{quantum}}(f) = R_{\text{classical}}(f) = \frac{1}{2} \log_2 5$. □

C. Proof of $\chi(C_5^{\vee m}) \geq (\frac{5}{2})^m$

Proof. By Proposition 3, we have that $\alpha(C_5^{\vee m}) = \alpha(C_5)^m = 2^m$. Since $\alpha(G)\chi(G) \geq |V(G)|$ [22, page 193] and, we have $\chi(C_5^{\vee m}) \geq |V(C_5^{\vee m})|/\alpha(C_5^{\vee m}) = \frac{5^m}{2^m}$. \square