Electronic Colloquium on Computational Complexity, Report No. 148 (2023)

One Clean Qubit Suffices for Quantum Communication Advantage

Srinivasan Arunachalam* Uma Girish† Noam Lifshitz‡

Abstract

We study the one-clean-qubit model of quantum communication where one qubit is in a pure state and all other qubits are maximally mixed. We demonstrate a partial function that has a quantum protocol of cost $O(\log N)$ in this model, however, every interactive randomized protocol has cost $\Omega(\sqrt{N})$, settling a conjecture of Klauck and Lim. In contrast, all prior quantum versus classical communication separations required at least $\Omega(\log N)$ clean qubits. The function demonstrating our separation also has an efficient protocol in the quantum-simultaneous-with-entanglement model of cost $O(\log N)$. We thus recover the state-of-the-art separations between quantum and classical communication complexity. Our proof is based on a recent hypercontractivity inequality introduced by Ellis, Kindler, Lifshitz, and Minzer, in conjunction with tools from the representation theory of compact Lie groups.

1 Introduction

A central goal in complexity theory is to understand the power of different computational resources. In the past four decades, communication complexity has provided a successful toolbox to establish several results in theoretical computer science in circuit complexity [KW90, KRW95], streaming algorithms [KKS14], property testing [BBM12], extension complexity [FMP+15], data structures [MNSW95], proof complexity [HN12]. In the standard two-player model of communication complexity introduced by Yao [Yao79] there are two parties Alice and Bob whose goal is to compute a partial function $F : \mathcal{X} \times \mathcal{Y} \to \{-1, 1, \star\}$. Alice receives $x \in \mathcal{X}$ and Bob receives $y \in \mathcal{Y}$ and their goal is to compute $F(x, y)$ for all $(x, y) \in F^{-1}(1) \cup F^{-1}(-1)$, while minimizing the number of bits of communication. One variant of this is when the players are allowed to send quantum messages. Quantum versus classical separations in communication complexity have a long and rich history. In a sequence of works [BCW98, BJK08, BCWW01, Raz99, GKRW06, GKK+07, RK11, Gav20, GRT22], it has been shown that quantum communication can exponentially outperform classical communication.[1] The state-of-the-art result among these is due to [Gav16, GRT22], who give a separation between quantum simultaneous communication complexity (where Alice and Bob share entanglement) and interactive randomized. This result subsumes most previous results and also shows that a rather weak and restricted model of quantum communication (simultaneous with entanglement) can exponentially outperform a rather strong classical model (interactive randomized). One of the benefits of proving quantum versus classical separations in communication complexity is that they are *unconditional*. The motivation for our work is two-fold:

*IBM Quantum, Almaden Research Center. Email: srinivasan.arunachalam@ibm.com

†Princeton University. Email: ugirish@cs.princeton.edu

‡Hebrew University of Jerusalem. Email: noamlifshitz@gmail.com

[1] Proving such separations for total functions is a major open question, however, we know several examples of partial functions such that quantum provides provable exponential speedups. Total functions are defined on all possible inputs, while partial functions are defined on a subset of inputs.

ISSN 1433-8092

Near-term implementations. Given that we are finally in an era of small noisy quantum devices, there have been proposals to use these communication separations to show experimental demonstrations of quantum advantage. To this end, Kumar et al. [KKD19] experimentally *demonstrated* a quantum communication advantage for the Hidden-matching problem defined by Bar-Yossef et al. [BJK08]. More recently, Aaronson coined the term "*quantum information supremacy*" wherein the goal is to show a task is solvable using a quantum resources that is exponentially more efficient than classical resources, with the benefit that this quantum advantage would be *unconditional* unlike sampling-based proposals. Aaronson et al. [ABK23] again used the communication problem of [BJK08] and showed a separation between quantum and classical complexity classes, and proposed an experimental implementation [Sco23]. Inspired by these recent works, we ask

What is the minimum quantum resource sufficient for a quantum communication speedup?

DQC_1 versus BPP. Knill and Laflamme [KL98] introduced the one-clean qubit model of quantum computation, also known as DQC_1. In this model, there is one qubit in a pure state and all other qubits are maximally mixed. The motivation of this model is two-fold: (*i*) The idea is to study the power of models of quantum computing in which the quantum memory is weak, but the control of this memory is good, in contrast to studying quantum computation, where the underlying memory is good, but the control is weak (such as Boson sampling) (*ii*) The primary motivation of [KL98] was the NMR approach to quantum computing where the initial state may be highly mixed. Despite how noisy these states are, DQC_1 is powerful and provides exponential speedups compared to the best known classical models [SJ08, KL98, CM18]. It was shown [MFF14] that DQC_1 is not efficiently classically simulable, unless the polynomial hierarchy collapses to the second level. While these results provide strong evidence that DQC_1 can exponentially outperform classical computation, proving this unconditionally has been a long-standing open question in complexity theory (in particular, what is the relation between DQC_1 and BPP). All known hardness results rely on complexity theoretic assumptions. When it comes to unconditional separations between quantum and classical, there are very few settings where quantum models provably exponentially outperform classical models. Communication complexity is a striking example of such a setting and a natural question is

Is $\mathsf{DQC}_1 \subseteq \mathsf{BPP}$ *in the communication world unconditionally?*

In the DQC_1 model of communication (first defined by Klauck and Lim [KL19]), Alice and Bob exchange quantum states such that the first qubit is in a pure state and all other qubits are maximally mixed. The players have no additional private memory, they simply take turns applying unitary operators on the state. (See Section 4.1 for a formal definition.) This is a rather restrictive model of quantum communication; all the aforementioned quantum versus classical separations require the quantum protocol to have at least $\Omega(\log N)$ clean qubits. They proposed a natural communication problem that is solvable using only one clean qubit. We call this the ABCD problem. Here, A, B, C, D are $N \times N$ special unitary matrices, Alice gets as input A, C explicitly and Bob gets as input B, D explicitly and their goal is to decide if $\mathrm{Tr}(ABCD) \geq 0.9N$ or $\mathrm{Tr}(ABCD) \leq 0.1N$ promised one of them is the case. This problem was shown to have a protocol with $O(\log N)$ qubits in the one-clean-qubit model of quantum communication [KL19].[2] They also conjectured that the interactive randomized communication complexity of this problem is $\Omega(\sqrt{N})$. The main contribution of this paper is to prove the conjecture of Klauck and Lim [KL19].

[2] Although they state their protocol for the ABC problem, the protocol trivially extends to the ABCD problem.

Main Result. We show that the ABCD problem can be computed with cost $O(\log N)$ with just *one-clean qubit*, however, every interactive randomized protocol has cost $\Omega(\sqrt{N})$. This separates DQC_1 and BPP unconditionally in the communication world. As far as we are aware, all prior quantum versus classical communication separations required at least $\Omega(\log N)$ clean qubits.

We also study the communication complexity of the ABCD problem in the quantum-simultaneous-with-entanglement model. In this model, Alice and Bob share entanglement, each apply a quantum operation on their part of the shared state and send everything to Charlie, who applies a projective measurement and announces the outcome as the answer. Interestingly, there is a protocol of cost $O(\log N)$ in this model for the ABCD problem. As a result, we show an exponential separation between quantum-simultaneous-with-entanglement and interactive randomized communication complexity and thus recover many of the best known separations, including [Gav19, GRT22]. Our quantum simultaneous protocol for the ABCD problem has the additional nice property that it is an entangled-*fingerprinting* protocol, i.e., a type of simultaneous protocol where Charlie essentially just performs a swap test. We describe this in more detail now.

Gavinsky et al. [GKdW06] introduced the quantum *fingerprinting* model in the SMP model: here Alice and Bob on input x, y respectively, send $U_x |0^n\rangle, V_y |0^n\rangle$ to Charlie who performs a swap test between $U_x |0^n\rangle$ and $V_y |0^n\rangle$. They repeat this process a few times before Charlie obtains the swap-test statistics and computes F on inputs x, y. Surprisingly, it was shown [GKdW06] that this model is efficiently simulable in the classical randomized simultaneous model of communication, thereby showing that quantum states are no stronger than classical states for the fingerprinting model. We consider the *entangled-fingerprinting* model where Alice and Bob share a few EPR pairs and on input x, y, apply U_x, V_y on their part of the shared state and send it to Charlie, who still performs a swap test between Alice's A-register and Bob's B-register. (See Section 4.3 for a formal definition of this model.) They repeat this in parallel with a fresh copy of $|\psi\rangle_{AB}$ and based on the swap test statistics, Charlie computes $F(x, y)$. A natural question is, *are entangled fingerprints stronger than just randomized fingerprints, and if so how much stronger?* In this work, we show that in contrast to the standard fingerprinting model, the entangled-fingerprinting model can *exponentially* outperform randomized fingerprinting and even outperform the strongest *interactive* classical model of communication.

Techniques. Given the definition of the ABCD problem, compact Lie groups such as the special unitary group $\mathsf{SU}(N)$ arise naturally. Our work draws inspiration from the study of quasirandom groups. A group G is said to be D-*quasirandom* if every D-dimensional representation $\rho\colon G \to \mathrm{GL}_D(\mathbb{C})$ is trivial, i.e., constantly equal to the identity. Group quasirandomness plays a central role in number theory [SX91], group theory [BG08] and combinatorics [Gow08]. Gowers and Viola [GV15] showcased how quasirandomness transcends its origins in pure mathematics, employing it as a pivotal tool in proving lower bounds for a variety of communication protocols over groups. For our purpose, it turns out that the quasirandomness of $\mathsf{SU}(N)$ alone is not sufficient. To overcome this, we instead introduce a set of new deep mathematical tools and concepts from the study of product free sets in $\mathsf{SU}(n)$ [KLM22, EKLM23] into communication complexity theory.

Our proof uses representation theory, Fourier analysis and hypercontractivity on $\mathsf{SU}(N)$. Fourier analysis on the Boolean cube and level-k inequalities have been important to prove quantum vs. classical separations [GKK+07, GRT22, Gav19, Mon10, DM20, BRW08, SWY12]. However, our problem is defined on $\mathsf{SU}(N)$ and it is unclear if the Fourier-analytic techniques on the Boolean cube extend to the special unitary group $\mathsf{SU}(N)$. A recent breakthrough work [EKLM23] studied product-free sets in quasirandom compact Lie groups, and showed hypercontractive inequalities for $\mathsf{SU}(N)$. Although not immediate, their hypercontractive inequality is pivotal for the R2 lower

bound. Our work appears to be the first application of this inequality to quantum computing and we believe that hypercontractivity on $\mathsf{SU}(N)$ will be of great interest to a broader quantum audience.

1.1 Main Theorem

Definition 1.1 (ABCD problem)**.** *Let A, B, C, D be $N \times N$ special unitary matrices. Alice is given A, C explicitly and Bob is given B, D respectively. Their goal is to output 1 if $\mathrm{Tr}(ABCD) \geq 0.9N$ and 0 if $\mathrm{Tr}(ABCD) \leq 0.1N$, promised that one of these is true.*

Our main theorem is as follows.

Theorem 1.2. *The ABCD problem has communication complexity*

1. $O(\log N)$ *in the one-clean qubit quantum model.*
2. $O(\log N)$ *in the quantum-simultaneous-with-entanglement model.*
3. $\Omega(\sqrt{N})$ *in the interactive classical randomized model.*

We remark that the classical lower bound is tight as shown in [KL19].[3]

1.2 Proof Overview

We now describe our classical lower bound for the ABCD problem, the main technical contribution. Our proof is based on a combination of three main ingredients. The first two of which are dimensional lower bounds for irreducible representations and formulas for convolution; these involve the representation theory of compact Lie groups. The third ingredient is the level-d inequality of [EKLM23], which shows that the Fourier spectrum of an indicator of a small subset of $\mathsf{SU}(n)$ is concentrated on the high dimensional representations.

Translating our lower bound to an analytic statement. To prove our lower bound, we define two distributions: Alice's inputs A, C and Bob's input B are chosen uniformly from $\mathsf{SU}(N)$ and in the YES distribution, Bob is given $D = (ABC)^{-1}$ and in the NO distribution, Bob is given uniformly random D from $\mathsf{SU}(N)$. We show that distinguishing between these two distributions requires $\Omega(\sqrt{N})$ communication. It is well-known, if we consider the matrix with rows and columns indexed by inputs of Alice and Bob respectively, then a classical cost-c communication protocol partitions this matrix into 2^c combinatorial rectangles. So, for a cost c, a typical rectangle in this partition has measure $\approx 2^{-c}$. Thus, to prove that a certain function requires $\Omega(\sqrt{N})$ classical communication cost, it suffices to show that rectangles of measure $\approx 2^{-\sqrt{N}}$ cannot distinguish the YES and NO instances of the function with sufficient advantage. Translating this to our setting, the main technical heart of our paper is the following lemma about large rectangles $f \times g$ in $\mathsf{SU}(N)^2 \times \mathsf{SU}(N)^2$ (think of $f \subseteq \mathsf{SU}(n)^2$ (resp. g) as indicators of Alice (resp. Bobs) inputs in that rectangle).

Lemma 1.3 (Main Lemma)**.** *Let $f, g : \mathsf{SU}(N)^2 \to \{0, 1\}$ be indicator functions such that $\mathbb{E}[f] = \alpha$ and $\mathbb{E}[g] = \beta$ for $\alpha, \beta \geq e^{-c'\sqrt{N}}$ for a sufficiently small global constant $c' > 0$. Then,*

$$\mathbb{E}\left[f(A, C) \cdot g(B, (ABC)^{-1})\right] \approx \alpha\beta \cdot (1 \pm 0.1)$$

where A, B, C, D are chosen independently according to the Haar probability measure over $\mathsf{SU}(N)$.

[3] Although they state their protocol for the ABC problem, the protocol trivially extends to the ABCD problem.

We now sketch the proof of this lemma.

Applying convolution formulas from nonabelian Fourier analysis Firstly, we study the difference between $\mathbb{E}[f(A,C)\cdot g(B,(ABC)^{-1})]$ and $\alpha\beta := \mathbb{E}[f(A,C)\cdot g(B,D)]$ and derive an expression for this in terms of the product of Fourier coefficients of f and g.

Claim 1.4 (Main Claim)**.** *Let $G = \mathsf{SU}(N)$ and $f, g : G \times G \to \{0,1\}$ be the indicator functions with $\alpha = \mathbb{E}[f]$ and $\beta = \mathbb{E}[g]$. Then,*

$$\Delta := \mathbb{E}\left[f(A,C)g(B,(ABC)^{-1})\right] - \alpha\beta = \sum_{\emptyset \neq \pi \in \widehat{G}} \frac{1}{\dim(\pi)} \left\langle \widehat{f}(\pi,\pi), \widehat{g}(\pi,\pi) \right\rangle.$$

where $\widehat{G}$ denotes the equivalence class of irreps of G and $\widehat{f}(\pi,\pi) \in \mathbb{C}^{\dim(\pi,\pi)\times\dim(\pi,\pi)}$ denotes the (matrix) Fourier coefficient corresponding to π.

The proof of this uses Fourier analysis, especially facts about Fourier coefficients of convolutions of functions. Loosely speaking,

- Taking an expectation of $f \times g$ over the distribution induced by $(A,B,C,(ABC)^{-1})$ has the effect of taking the convolution of f and g, which in the Fourier basis, translates to a taking a product of Fourier coefficients of f and g, divided by the square root of the dimension.
- The term $(ABC)^{-1}$ has the effect of zeroing out Fourier coefficients corresponding to (π,σ) where π and σ are inequivalent representations of G.

Utilizing the Fourier concentration on the high dimensions

We now describe how to upper bound the R.H.S. of Claim 1.4. To do this, we will use the degree-decomposition of f and g. It turns out that the space $L^2(G) = \{f : G \to \mathbb{C}, \mathbb{E}[|f|^2] < \infty\}$ can be expressed as $\oplus_{d\in\mathbb{N}} V_d \oplus V_0$, where V_0 consists of constant functions, V_d essentially captures polynomials of "pure-degree" d, furthermore, each V_d is a sub-representation of G [EKLM23]. We group the terms in the R.H.S. of Claim 1.4 based on this degree decomposition to obtain

$$\Delta = \sum_{d=1}^{\infty} \sum_{\pi \in \widehat{V_d}} \frac{1}{\dim(\pi)} \left\langle \widehat{f}(\pi,\pi), \widehat{g}(\pi,\pi)) \right\rangle$$

We now apply Cauchy-Schwarz to upper bound $\left\langle \widehat{f}(\pi,\pi), \widehat{g}(\pi,\pi)) \right\rangle$ by $\|\widehat{f}(\pi,\pi)\|_2 \cdot \|\widehat{g}(\pi,\pi)\|_2$. We again apply Cauchy-Schwarz over terms $\pi \in \widehat{V_d}$ to obtain

$$\Delta \leq \sum_{d=1}^{\infty} \sqrt{\sum_{\pi\in\widehat{V_d}} \|\widehat{f}(\pi,\pi)\|^2} \cdot \sqrt{\sum_{\pi\in\widehat{V_d}} \|\widehat{g}(\pi,\pi)\|^2} \cdot \max_{\pi\in\widehat{V_d}} \left(\frac{1}{\dim(\pi)}\right)$$

Observe that $\widehat{f}(\pi,\pi)$ and $\widehat{g}(\pi,\pi)$ correspond to the degree-$2d$ component. Thus,

$$\Delta \leq \sum_{d=1}^{\infty} \left\|f^{=2d}\right\| \cdot \left\|g^{=2d}\right\| \cdot \left(\min_{\pi\in\widehat{V_d}} \dim(\pi)\right)^{-1} \tag{1}$$

where $f^{=2d}, g^{=2d}$ denote the projection of f, g onto the degree $2d$ part. We now use the main results of [EKLM23]. Two important contributions of [EKLM23] are the following. Firstly, the dimensions of irreps of $\widehat{V_d}$ grow fast, roughly as $\gtrapprox N^d$; secondly, an analogue of the level-k inequality holds for $\mathsf{SU}(N)$ and its variants:

Lemma 1.5 (Implied by [EKLM23])**.** *There exists universal constants $c, C > 0$ such that the following holds. Let $f : \mathsf{SU}(N)^2 \to \{0,1\}$, $\alpha = \mathbb{E}[f]$ and $d \leq \min\{c\sqrt{N}, \log(1/\alpha)/2\}$. Then*

$$\left\|f^{=d}\right\|_2^2 \leq (C/d)^d \, \alpha^2 \log^d(1/\alpha).$$

Since $\mathbb{E}[f], \mathbb{E}[g] \geq e^{-c'\sqrt{N}}$ for a sufficiently small constant c', this lemma essentially implies that

$$\left\|f^{=2d}\right\|_2^2, \left\|g^{=2d}\right\|_2^2 \ll \alpha\beta \cdot N^d \cdot 11^{-d}.$$

As mentioned earlier, we have $\min_{\pi \in \widehat{V_d}} \dim(\pi) \gtrapprox N^d$ and thus $\min_{\pi \in \widehat{V_d}} \dim(\pi)$ grows much faster than $\left\|f^{=2d}\right\| \cdot \left\|g^{=2d}\right\|$. Plugging this in Eq. (1) implies the desired result:

$$\mathbb{E}[f(A,C) \cdot g(B,(ABC)^{-1})] - \alpha\beta \triangleq \Delta \ll \sum_{d=1}^{\infty} \alpha\beta \cdot N^d \cdot 11^{-d} \cdot 1/N^d \leq \alpha\beta \cdot \sum_{d=1}^{\infty} 11^{-d} \leq \alpha\beta/10.$$

Using standard techniques in communication complexity, we use the above inequality to show that every protocol of cost $\ll \sqrt{N}$ succeeds in solving the ABCD problem with probability $\leq 1/10$, completing the proof sketch.

Organization. We describe notation in Section 2 and some important results about special unitary matrices in Section 3. We describe the quantum communication models in Section 4. As mentioned before, Item 1 in Theorem 1.2 was essentially proved in [KL19], but we reprove it in Section 4.4 for completeness. We prove Item 2 in Section 4.5 and Item 3 in Section 5.

Acknowledgements. This work was done in part while the authors were visiting the Simons Institute for the Theory of Computing. We thank Vojtech Havlicek, Tarun Kathuria, Ran Raz and Makrand Sinha for the many valuable discussions.

2 Notation

Let $\mathbb{N} = \{1, 2, \ldots, \}$. We use $\mathbb{1}[E]$ to denote the indicator of an event E. Let $N = 2^n$ for $n \in \mathbb{N}$. For a vector space V, we use $\mathcal{L}(V)$ to denote the space of *endomorphisms* of V (i.e., set of linear maps from V onto itself) and $\mathcal{GL}(V)$ to denote the set of invertible endomorphisms in $\mathcal{L}(V)$. For every compact group G equipped with a Haar measure, we use $L^2(G)$ to denote the space of all square-integrable functions acting on G quotiented by the equivalence relation $f \sim g$ if f equals g almost everywhere with respect to μ. Mathematically, we write $L^2(G)$ as

$$L^2(G) = \{f : G \to \mathbb{C} \,|\, \mathbb{E}[|f|^2] < \infty\}/\sim,$$

where the expectation is with respect to the Haar measure. One can also view $L^2(G)$ as a Hilbert space equipped with a natural inner product

$$\langle f_1, f_2 \rangle = \mathbb{E}_g[f_1(g)\overline{f_2(g)}]$$

for $f_1, f_2 \in L^2(G)$, where g is Haar random.

3 Properties of Special Unitary Matrices

Special Unitary Group. We use $\mathsf{SU}(N)$ to denote the special unitary group of $N \times N$ matrices. Let $\mathcal{H}$ denote the Haar measure on $\mathsf{SU}(N)$. The Haar measure is the unique measure on $\mathsf{SU}(N)$ that is invariant under right-multiplication and left-multiplication by $\mathsf{SU}(N)$.

3.1 Representation Theory of $\mathsf{SU}(N)$

We now describe representations of groups. For a group G, a representation (π, V) of G is a *group homomorphism* $\pi : G \to \mathcal{GL}(V)$, i.e., a map from G to non-singular complex matrices satisfying $\pi(gh) = \pi(g)\pi(h)$ for all $g, h \in G$ and $\pi(1) = \mathbb{I}$. Throughout this paper, we will assume that our group G will be compact and our representations are finite dimensional. For notational convenience, we refer to the representation (π, V) simply as V or as π. We also abuse notation by writing $\pi(g)v$ as gv for $g \in G, v \in V$. For any finite dimensional representation (π, V), there is a basis for V according to which $\pi(g)$ is unitary for all $g \in G$ and we will typically work with such a basis. A *G-morphism* between irreps is a map $\varphi\colon V \to U$, satisfying $\varphi(gv) = g\varphi(v)$ (i.e., $\varphi(\pi(g)v) = \pi(g)\varphi(v)$.) for all $g \in G$ and $v \in V$

Irreducible representations and Schur's Lemma Two representations $(\pi, V), (\rho, U)$ are said to be *isomorphic* (which we denote by $\pi \sim \rho$) if there exists an invertible G-morphism between them. Otherwise they are *non-isomorphic* (denoted $\pi \nsim \rho$). We denote $U \leq V$ to be a *subrepresentation* if $gu \in U$ for all $g \in G$ and $u \in U$. A representation is said to be an *irreducible representation* (or *irrep*) if its only subrepresentations are 0 and itself, i.e., if it cannot be decomposed as the direct sum of two non-trivial representations. We use $\widehat{G}$ to denote a complete set of irreps of G, that is, every irrep of G is isomorphic to some irrep in $\widehat{G}$. We will also make use of Schur's lemma which we state below.

Lemma 3.1 (Schur's lemma)**.** *Let $(\pi, V), (\rho, U)$ be two irreps of G. Let φ be a G-morphism between π and ρ. If $\pi \nsim \rho$, then φ is 0 and if $\pi = \rho$, then φ is a scalar multiple of identity.*

Matrix Coefficients Let $(\pi, V) \in \widehat{G}$ be an irrep of G. We use $M_\pi \subseteq L^2(G)$ to denote the space spanned by functions $f_{u,v} : G \to \mathbb{R}$ of the form $g \to \langle u, \pi(g)v\rangle$ for $u, v \in V$. We refer to M_π as the space of *matrix coefficients* associated to the representation π. For any $i, j \in [\dim(V)]$, let $\widetilde{\pi}_{i,j}$ be $\sqrt{\dim(V)}\pi_{i,j}$ where $\pi_{i,j}$ is the function defined as

$$\pi_{i,j} : g \mapsto \pi(g)_{i,j} = \langle e_i, \pi(g)e_j\rangle$$

for all $g \in G$. We now state Schur's orthogonality relation, a corollary of Schur's lemma.

Fact 3.2 (Schur's Orthogonality Relations)**.** *Let $(\pi, V), (\sigma, W) \in \widehat{G}$. Then,*

$$\mathbb{E}_g\left[\widetilde{\pi}(g)_{i,j} \cdot \overline{\widetilde{\sigma}(g)_{k,\ell}}\right] = \mathbb{1}[\sigma = \pi, i = k, j = \ell], \qquad \forall\, i, j \in [\dim(V)], k, \ell \in [\dim(W)].$$

In other words, $\{\widetilde{\pi}_{i,j} : i, j \in \dim(V)\}$ is an orthonormal basis for M_π and, M_π, M_σ are orthogonal for $\pi \nsim \sigma$.

Peter-Weyl Theorem The Peter Weyl theorem states that the space of all matrix coefficients is dense in $L^2(G)$. In other words, $L^2(G)$ can be decomposed as an orthogonal direct sum of matrix coefficients $\{M_\pi : \pi \in \widehat{G}\}$.

Theorem 3.3 (Peter-Weyl Theorem)**.** *If G is a group equipped with the Haar measure, then*

$$L^2(G) = \bigoplus_{\pi \in \widehat{G}} M_\pi.$$

Furthermore, every closed subspace of $W \subseteq L^2(G)$, that commutes with the action of G from both sides can be written as $W = \bigoplus_{\pi \in L_W} M_\pi$ for some $L_W \subseteq \widehat{G}$.

This provides a very natural basis to study $L^2(G)$ and an analogue of Fourier analysis for G.

3.2 Fourier Coefficients

For any function $f \in L^2(G)$ and for any $(\pi, V) \in \widehat{G}$, define the Fourier coefficient $\widehat{f}(\pi) \in \mathcal{L}(V)$ as

$$\widehat{f}(\pi) = \mathbb{E}_g[f(g) \cdot \widetilde{\pi}(g^{-1})],$$

where the expectation is with respect to the Haar measure on G. The Peter-Weyl theorem implies that the Fourier decomposition of f can be written as

$$f(g) = \sum_{\substack{\pi \in \widehat{G} \\ i,j \in [\dim(\pi)]}} \widehat{f}(\pi)_{i,j} \cdot \widetilde{\pi}(g)_{j,i}$$

for all $g \in G$. Define $\|f\|_2^2 = \mathbb{E}_g[|f(g)|^2]$. Similarly to classical Boolean function analysis, one can define the Plancharel's theorem and convolutions of functions in $L^2(G)$ which we describe now.

Theorem 3.4 (Plancharel's Theorem)**.** *For every $f, h \in L^2(G)$, we have that*

$$\mathbb{E}_g[f(g)\overline{h(g)}] = \sum_{\substack{\pi \in \widehat{G} \\ i,j \in [\dim(\pi)]}} \widehat{f}(\pi)_{i,j} \cdot \overline{\widehat{h}(\pi)_{i,j}}.$$

In particular, $\mathbb{E}_g[|f(g)|^2] = \sum_{\substack{\pi \in \widehat{G} \\ i,j \in [\dim(\pi)]}} \left|\widehat{f}(\pi)_{i,j}\right|^2.$

This follows from the following calculation.

$$\begin{aligned}
\mathbb{E}_g[f(g)\overline{h(g)}] &= \sum_{\substack{\pi,\sigma \in \widehat{G} \\ i,j \in [\dim(\pi)] \\ k,\ell \in [\dim(\sigma)]}} \mathbb{E}_g\left[\widehat{f}(\pi)_{i,j} \cdot \widetilde{\pi}(g)_{j,i} \cdot \overline{\widehat{h}(\sigma)_{k,\ell}} \cdot \overline{\widetilde{\sigma}(g)_{\ell,k}}\right] \\
&= \sum_{\substack{\pi,\sigma \in \widehat{G} \\ i,j \in [\dim(\pi)] \\ k,\ell \in [\dim(\sigma)]}} \widehat{f}(\pi)_{i,j} \cdot \overline{\widehat{h}(\sigma)_{k,\ell}} \cdot \mathbb{E}_g[\widetilde{\pi}(g)_{j,i} \cdot \overline{\widetilde{\sigma}(g)_{\ell,k}}] = \sum_{\substack{\pi \in \widehat{G} \\ i,j \in [\dim(\pi)]}} \widehat{f}(\pi)_{i,j} \cdot \overline{\widehat{h}(\pi)_{i,j}}.
\end{aligned}$$

Convolution of Functions For every $f_1, f_2 \in L^2(G)$, define their convolution $f_1 * f_2 \in L^2(G)$ by

$$(f_1 * f_2)(g) = \mathbb{E}_h \left[f_1(gh^{-1}) f_2(h) \right]$$

for all $g \in G$. We make use of the following formula for the convolutions of two elements of our orthonormal basis of $L^2(G)$.

Fact 3.5. *Let $(\pi, V), (\sigma, W) \in \widehat{G}$. For all $i, j \in [\dim(V)], k, \ell \in [\dim(W)]$, we have that*

$$\widetilde{\pi}_{i,j} * \widetilde{\sigma}_{k,\ell} = \frac{\mathbb{1}[j = k, \pi \sim \sigma]}{\sqrt{\dim(V)}} \cdot \widetilde{\pi}_{i,\ell}$$

3.3 Hypercontractivity on $\mathsf{SU}(N)$

Consider the group $G = \mathsf{SU}(N)$. For $X \in \mathsf{SU}(N)$ and $d \geq 1$, define $V_{\leq d}$ to consist of functions representable as degree d multilinear polynomials in the formal variables $\{\mathsf{Re}(X_{ij}), \mathsf{Im}(X_{ij}) : i, j \in [N]\}$ where $X \in \mathsf{SU}(N)$. For every $d \in \{0, \ldots, N/2 - 1\}$, define $V_{=d} := V_{\leq d} \cap (V_{\leq d-1})^{\perp}$ as the "degree-d" part and $V_{\geq N/2} = (V_{<N/2})^{\perp}$. The space $L^2(G)$ can be decomposed as

$$\bigoplus_{d=0}^{N/2-1} V_{=d} \bigoplus V_{\geq N/2}.$$

Furthermore, since $V_{=d} \subseteq L^2(G)$ is closed and commutes with the action of G from both sides (i.e., linear combination of degree-d polynomials remains degree-d), by the Peter-Weyl theorem we have

$$V_{=d} = \bigoplus_{\pi \in L_d} M_\pi$$

for some $L_d \subseteq \widehat{G}$. One important contribution of [EKLM23] is in proving important properties of this decomposition, which we discuss now. For every $f \in L^2(G)$ and $0 \leq d \leq N/2-1$, let $f^{=d}$ denote the projection of f onto $V_{=d}$, i.e., $f^{=d} = \arg\min\{\langle g, f\rangle : g \in V_{=d}\}$. Let $Q_d = \min_{\pi \in L_d}\{\dim(\pi)\}$ be the minimal dimension of any non-trivial subrepresentation of $V_{=d}$. Let $f^{\geq N/2}$ denote the projection of f onto $V_{\geq N/2}$ and $Q_{\geq N/2}$ be the minimal dimension of any representation of $V_{\geq N/2}$. With this, we are now ready to state the main results we use from [EKLM23].

Theorem 3.6 (Implied by [EKLM23, Theorems 3.7, 4.5])**.** *There exists universal constants $c, C > 0$ such that the following holds. Let $f : \mathsf{SU}(N) \times \mathsf{SU}(N) \to \{0, 1\}$, $\alpha = \mathbb{E}[f]$ and $d \in \mathbb{N}$ such that $d \leq \min\{c\sqrt{N}, \log(1/\alpha)/2\}$. Then*

$$\left\| f^{=d} \right\|_2^2 \leq (C/d)^d \, \alpha^2 \log^d(1/\alpha).$$

Furthermore, one can bound the dimension of the irreps occuring in $V_{=d}$ as follows.

Theorem 3.7 ([EKLM23, Theorem 3.3])**.** *Let $G = \mathsf{SU}(N)$, $d \leq N/2 - 1$. Let $c > 0$ be a universal constant. Then*

- $Q_d \geq \left(\frac{cN}{d}\right)^d$ *if* $d < cN/(1 + c)$,
- $Q_d \geq (1 + c)^{cN/(1+c)}$ *if* $d \geq cN/(1 + c)$.

Furthermore, every irrep of $V_{\geq N/2}$ has dimension at least $(1 + c)^{cN/(1+c)}$.

4 Quantum Models & Quantum Upper Bound

We begin by describing the quantum models of communication and then presenting the upper bounds.

4.1 Clean-Qubit Model

This model was first defined by Klauck and Lim in [KL19]. A k-clean-qubit quantum protocol consists of k qubits in the state $|0\rangle$ and m qubits that are unentangled from these and in the totally mixed state. There is no other private memory for the players. The players communicate as in a standard quantum protocol, that is, they apply unitary operators on the $m+k$ qubits and exchange them back and forth. At the end of the computation, an arbitrary projective measurement (independent of the inputs) is performed. The outcome of the measurement is declared as the output of the protocol. As in standard quantum protocols, the cost of such a protocol is the total number of qubits exchanged, which in this case is the number of rounds times $(m+k)$. Note that the clean and mixed qubits are allowed to have correlations between them, and the clean qubit can act as a control over the mixed qubits.

4.2 Quantum Simultaneous with Entanglement

In this model, in addition to Alice and Bob, there is a third party Charlie. Alice and Bob initially share an entangled state (that is independent of their inputs) and each apply a quantum channel (dependent on their inputs) on their part of the entangled state and send all the qubits to Charlie. Charlie applies a projective measurement (independent of the inputs) is performed. The outcome of the measurement is declared as the output of the protocol. The cost of the protocol is the total number of qubits sent to Charlie.

4.3 Entangled Fingerprinting Model

An entangled-fingerprinting protocol is a simple type of quantum simultaneous protocol with entanglement where essentially, Charlie just performs a swap test. In more detail, Alice and Bob use their entanglement to prepare a state of the form

$$\frac{1}{\sqrt{2N}} \sum_{i\in[N]} \left(|0_A, 0_B\rangle \, |u_i\rangle_A \, |v_i\rangle_B + |1_A, 1_B\rangle \, |u_i'\rangle_A \, |v_i'\rangle_B \right),$$

where $|u_i\rangle_A, |u_i'\rangle_A$ are quantum states prepared by Alice and $|v_i'\rangle_B, |v_i\rangle_B$ are states prepared by Bob and the index denotes the player to which the qubit belongs. The players send this entire state to Charlie. Charlie first "uncomputes" the second qubit to obtain

$$\frac{1}{\sqrt{2N}} \sum_{i\in[N]} \left(|0\rangle \, |u_i\rangle \, |v_i\rangle + |1\rangle \, |u_i'\rangle \, |v_i'\rangle \right),$$

and then performs a swap test on this state. In more detail, she swaps the last few registers controlled on the first qubit to obtain

$$\frac{1}{\sqrt{2N}} \sum_{i\in[N]} \left(|0\rangle \, |u_i\rangle \, |v_i\rangle + |1\rangle \, |v_i'\rangle \, |u_i'\rangle \right),$$

and then measures the first qubit in the Hadamard basis and returns 1 iff the outcome is $|+\rangle$.

The motivation for calling this the "entangled-fingerprinting" model is as follows. In the standard quantum fingerprinting model of communication [GKdW06], Alice and Bob send quantum states $|u\rangle$ and $|v\rangle$ respectively to Charlie, who then performs a swap test and returns 1 with probability $\frac{1}{2}\left(1+\langle u|v\rangle^2\right)$. It was shown [GKdW06] that this model is efficiently simulable in the classical randomized simultaneous model of communication. The entangled-fingerprinting model can be viewed as a variant of the standard quantum fingerprinting model where Alice and Bob are allowed to share entanglement. In contrast to the standard fingerprinting model, we show that the entangled-fingerprinting model can exponentially outperform even interactive classical communication.

4.4 Quantum Upper Bound with One Clean Qubit

As we mentioned in the introduction, the ABCD problem was shown to have a simple quantum communication protocol with $O(\log N)$ qubits of communication using one clean qubit [KL19]. For completeness, we include a proof of this here.

Theorem 4.1. *There is a quantum protocol of cost $O(\log N)$ for the ABCD problem in the one-clean-qubit model such that* YES *instances are accepted with probability at least* 0.95 *and the* NO *instances are accepted with probability at most* 0.55.

This gap of 0.5 between YES and NO instances can be amplified to an arbitrary constant by repeating the protocol $O(1)$ times (and using $O(1)$ clean qubits).

Proof of Theorem 4.1. Consider the (mixed) state on $\log N + 1$ qubits identified by the density matrix $\frac{1}{N}\begin{bmatrix}\mathbb{I} & 0\\ 0 & 0\end{bmatrix}$. This state can be viewed as a probability mixture over pure states $|0\rangle \otimes |v\rangle$ where $v \in \mathbb{C}^N$ is a uniformly random unit vector. The protocol starts by Alice first applying the Hadamard operator on the first qubit to produce the uniform mixture over $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes |v\rangle$ over a random unit vector $v \in \mathbb{C}^N$. Alice applies $A^\dagger$ controlled on the first qubit being $|1\rangle$ and sends the entire state to Bob, who applies $B^\dagger$ controlled on the first qubit being one and sends it to Alice. They similarly apply $C^\dagger$ and $D^\dagger$. This produces the uniform mixture over

$$\frac{1}{\sqrt{2}}|0\rangle \otimes |v\rangle + \frac{1}{\sqrt{2}}|1\rangle \otimes D^\dagger C^\dagger B^\dagger A^\dagger |v\rangle$$

for a random unit vector $v \in \mathbb{C}^N$. Alice now measures the first qubit in the Hadamard basis and returns YES if and only if the outcome is $|+\rangle$. The probability of outcome $|+\rangle$ is precisely

$$\begin{aligned}
&\frac{1}{4}\left\| |v\rangle + D^\dagger C^\dagger B^\dagger A^\dagger |v\rangle \right\|_2^2 \\
&= \frac{1}{4}\big(\langle v| + \langle v| ABCD\big)\big(|v\rangle + D^\dagger C^\dagger B^\dagger A^\dagger |v\rangle\big) \\
&= \frac{1}{2} + \frac{1}{4}\left(\langle v|ABCD|v\rangle + \langle v|(ABCD)^\dagger|v\rangle\right).
\end{aligned}$$

If $\mathrm{Tr}(ABCD) \geq 0.9N$, then the average of the above quantity (over a random unit vector v) quantity is at least 0.95 and if $\mathrm{Tr}(ABCD) \leq 0.1N$, then the average of this quantity at most 0.55. □

4.5 Quantum Upper Bound with Entangled Fingerprints

In this section, we show that the ABCD problem can be solved in the entangled-fingerprinting model in the SMP communication model.

Theorem 4.2. *There is a quantum protocol of cost $O(\log N)$ in the entangled-fingerprinting model for the ABCD problem when Alice and Bob share $\Theta(\log N)$ EPR pairs such that the* YES *instances are accepted with probability ≥ 0.95 and the* NO *instances are accepted with probability ≤ 0.55.*

Proof of Theorem 4.2. The protocol is as follows. We express the initial state shared by Alice and Bob as follows.

$$\frac{1}{\sqrt{2N}}\Big(|0_A 0_B\rangle \sum_{i=1}^{N} |i_A, i_B\rangle + |1_A 1_B\rangle \sum_{i=1}^{N} |i_A, i_B\rangle \Big)$$

where the subscript A, B denote the registers that Alice and Bob have respectively. Alice applies the map $\begin{bmatrix} A & 0 \\ 0 & C \end{bmatrix}$ which maps $|0_A\rangle |i_A\rangle$ to $|0_A\rangle A |i_A\rangle$ and $|1_A\rangle |i_A\rangle$ to $|1_A\rangle C |i_A\rangle$. Bob applies the map $\begin{bmatrix} B^\dagger & 0 \\ 0 & D^\dagger \end{bmatrix}$ which maps $|0_B\rangle |i_B\rangle$ to $|0_B\rangle B^\dagger |i_B\rangle$ and $|1_B\rangle |i_B\rangle$ to $|1_B\rangle D^\dagger |i_B\rangle$. This produces the state

$$\frac{1}{\sqrt{2N}}\Big(|0_A, 0_B\rangle \sum_{i=1}^{N} A |i_A\rangle B^\dagger |i_B\rangle + |1_A 1_B\rangle \sum_{i=1}^{N} C |i_A\rangle D^\dagger |i_B\rangle \Big).$$

They send the above state to Charlie. Charlie first uncomputes the second qubit to obtain

$$\frac{1}{\sqrt{2N}}\Big(|0\rangle \sum_{i=1}^{N} A |i\rangle B^\dagger |i\rangle + |1\rangle \sum_{i=1}^{N} C |i\rangle D^\dagger |i\rangle \Big).$$

Charlie then does a swap test on this state. That is, she applies a controlled swap between the last two registers controlled on the first register and then measures the first qubit in the Hadamard basis. If the outcome is $|+\rangle$ then she outputs YES, else outputs NO. The probability of outcome $|+\rangle$ is precisely

$$\begin{aligned}
&\frac{1}{8N}\left\| \sum_{i=1}^{N} A |i\rangle B^\dagger |i\rangle + \sum_{i=1}^{N} D^\dagger |i\rangle C |i\rangle \right\|_2^2 \\
&= \frac{1}{8N}\Big(\sum_{j=1}^{N} \langle j| A^\dagger \langle j| B + \sum_{j=1}^{N} \langle j| D \langle j| C^\dagger \Big)\Big(\sum_{i=1}^{N} A |i\rangle B^\dagger |i\rangle + \sum_{i=1}^{N} D^\dagger |i\rangle C |i\rangle \Big) \\
&= \frac{1}{8N}\Big(2N + \sum_{i,j=1}^{N} \langle j| A^\dagger D^\dagger |i\rangle \langle j| BC |i\rangle + \sum_{i,j=1}^{N} \langle j| DA |i\rangle \langle j| C^\dagger B^\dagger |i\rangle \Big) \\
&= \frac{1}{8N}\Big(2N + \sum_{i,j=1}^{N} \langle j| A^\dagger D^\dagger |i\rangle \langle i| C^\dagger B^\dagger |j\rangle + \sum_{i,j=1}^{N} \langle j| DA |i\rangle \langle i| BC |j\rangle \Big) \\
&= \frac{1}{4} + \frac{1}{8N}\Big(\operatorname{Tr}\left((ABCD)^\dagger\right) + \operatorname{Tr}\left(ABCD\right)\Big)
\end{aligned}$$

If $\operatorname{Tr}(ABCD) \geq 0.9N$, then the above quantity is at least 0.475 and the protocol is correct with probability at least 0.95 whereas if $\operatorname{Tr}(ABCD) \leq 0.1N$, the above quantity is at most 0.275 and the protocol returns YES with probability at most 0.55. □

5 Classical Lower Bound

In this section we will prove the following theorem.

Theorem 5.1. *The randomized communication complexity of the ABCD problem is* $\Omega(\sqrt{N})$.

The main technical lemma of this paper is the following.

Lemma 5.2. *Let* $f, g : \mathsf{SU}(N) \times \mathsf{SU}(N) \to \{0,1\}$ *be such that* $\mathbb{E}[f], \mathbb{E}[g] \geq e^{-c'\sqrt{N}}$ *for a sufficiently small global constant* $c' > 0$. *Then,*

$$\left|\mathbb{E}[f(A,C)g(B,(ABC)^{-1})] - \mathbb{E}[f(A,C)g(B,D)]\right| \leq \mathbb{E}[f(A,C)g(B,D)]/30,$$

where all the expectations are taken with respect to the Haar measure over $\mathsf{SU}(N)$.

The proof of the main theorem essentially follows by adding this inequality over all rectangles in the protocol. We first prove the theorem assuming the lemma.

Proof of Theorem 5.1 from Lemma 5.2. Let c' be the global constant as in Lemma 5.2. We will show that any randomized communication protocol of cost $c'\sqrt{N}/2$ has advantage at most $1/10$ in distinguishing YES and NO instances of the ABCD problem. To this end, consider the *hard* distributions defined as follows.

Definition 5.3 (Hard Distributions)**.** *For the* YES *distribution, Alice gets* $A, C \sim \mathcal{H}$, *Bob gets* $B, D \sim \mathcal{H}$. *For the* NO *distribution, Alice gets* $A, C \sim \mathcal{H}$ *and Bob gets* $B \sim \mathcal{H}$ *and* $D = (ABC)^{-1}$.

Let $\mathcal{P} : \mathsf{SU}(N)^4 \to \{0,1\}$ denote the (probabilistic) output of any such randomized communication protocol. Here, we view the output of the protocol as a probabilistic bit in $\{0,1\}$. By Yao's lemma and the triangle inequality, it suffices to bound the difference

$$\mathbb{E}[\mathcal{P}(A,B,C,(ABC)^{-1})] - \mathbb{E}[\mathcal{P}(A,B,C,D)] \tag{2}$$

for *deterministic* protocols. Fix any deterministic protocol $\mathcal{P} : \mathsf{SU}(N)^4 \to \{0,1\}$ of cost $c'\sqrt{N}/2$. This defines a partition of the input space into rectangles, where a typical rectangle has measure $2^{-c'\sqrt{N}/2}$ under the NO distribution. Let $\mathcal{R}$ denote the set of rectangles of NO-measure at least $2^{-c'\sqrt{N}}$. Observe that NO-measure of $\mathcal{R}$ is at least $1 - 2^{c'\sqrt{N}/2} \cdot 2^{-c'\sqrt{N}} \geq 1 - 2^{-c'\sqrt{N}/2}$. We will analyze the contribution of each rectangle in Eq. (2) separately, based on whether it is in $\mathcal{R}$ or not.

Fix any rectangle $f \times g$ in $\mathcal{R}$ where $f, g : \mathsf{SU}(N)^2 \to \{0,1\}$ are the indicator functions of Alice's and Bob's sets. Since the NO-measure of the rectangle is precisely $\mathbb{E}[f] \cdot \mathbb{E}[g]$, we have $\mathbb{E}[f], \mathbb{E}[g] > e^{-c'\sqrt{N}}$. We now apply Lemma 5.2 to conclude that

$$\left|\mathbb{E}[f(A,C)g(B,(ABC)^{-1})] - \mathbb{E}[f(A,C)g(B,D)]\right| \leq \mathbb{E}[f(A,C)g(B,D)]/30. \tag{3}$$

We already argued that the NO-measure of $\mathcal{R}$ is at least $1 - 2^{-c'\sqrt{N}}$. We now add Eq. (3) over all rectangles in $\mathcal{R}$ to conclude that the YES-measure of $\mathcal{R}$ is at least $1 - 2^{-c'\sqrt{N}/2} - 1/30 \geq 1 - 1/20$. In particular, the total YES-measure of rectangles *not in* $\mathcal{R}$ is at most $1/20$. Hence, the total contribution of such rectangles to Eq. (2) is at most $1/20 + 2^{-c'\sqrt{N}/2}$. We now consider the contribution of rectangles in $\mathcal{R}$. We again use Eq. (3) and add up over all one-rectangles in $\mathcal{R}$ to conclude that the total contribution of such rectangles to Eq. (2) is at most $1/30$. Overall, we have

$$\mathbb{E}[\mathcal{P}(A,B,C,(ABC)^{-1})] - \mathbb{E}[\mathcal{P}(A,B,C,D)] \leq 1/30 + 1/20 + 2^{-c'\sqrt{N}/2} < 1/10.$$

Here, we used the facts that the protocol is constant within a rectangle and the output of the protocol is in $\{0,1\}$. This completes the proof of Theorem 5.1. □

It remains to prove Lemma 5.2 which we do in the next section.

5.1 Proof of Lemma 5.2

In order to prove this lemma, we need two claims: the first one decomposes the main expression we need to bound in Lemma 5.2 in terms of the Fourier coefficients and the next claim is a corollary of Theorem 3.6.

Claim 5.4. *Let $G = \mathsf{SU}(N)$ and let $f, g : G \times G \to \{0,1\}$ be the indicator functions. Then,*

$$\mathbb{E}\left[f(A,C)g(B,(ABC)^{-1})\right] = \sum_{\pi \in \widehat{G}} \frac{1}{\dim(\pi)} \sum_{\substack{i,j \in [\dim(\pi)] \\ k,\ell \in [\dim(\pi)]}} \widehat{f}(\pi,\pi)_{k,j,\ell,i} \cdot \widehat{g}(\pi,\pi)_{i,k,j,\ell}.$$

Proof. This follows by expanding f, g in the Fourier basis. To simplify the analysis, we introduce the diagonal probability measure μ_{diag}. It is the measure obtained by sampling $x \sim \mathsf{SU}(n)$ and outputting (x, x^{-1}). More formally it is the push-forward of the Haar measure with respect to the map $x \mapsto (x, x^{-1})$. Observe that

$$\begin{aligned}
\mathbb{E}\left[f(A,C)g(B,(ABC)^{-1})\right] &= \mathbb{E}\left[f(A,C)g(B,D) \,\middle|\, AB = (CD)^{-1}\right] \\
&= \mathbb{E}\left[(f*g)(AB,CD) \,\middle|\, AB = (CD)^{-1}\right] \\
&= \langle \mu_{\mathsf{diag}}, f * g \rangle,
\end{aligned}$$

where the second equality used the definition of convolution to get $\mathbb{E}[f(A,C)g(B,(ABC)^{-1}] = \mathbb{E}[(f*g)(AB,(AB)^{-1})]$. We can identify $\widehat{G \times G}$ with the tensor product $\widehat{G} \otimes \widehat{G}$ and accordingly, we will index the irreps of $G \times G$ by $\pi \otimes \sigma$ for $\pi, \sigma \in \widehat{G}$ and we refer to the corresponding Fourier coefficients of a function $h : G \times G \to \mathbb{R}$ by $\widehat{h}(\pi,\sigma)$. Using Plancharel's theorem, we now rewrite the above as

$$\langle \mu_{\mathsf{diag}}, f * g \rangle = \sum_{\pi,\sigma \in \widehat{G}} \sum_{\substack{i,j \in [\dim(\pi)] \\ k,\ell \in [\dim(\sigma)]}} \overline{\widehat{\mu_{\mathsf{diag}}}(\pi,\sigma)_{i,j,k,\ell}} \cdot \widehat{f*g}(\pi,\sigma)_{i,j,k,\ell} \tag{4}$$

We now compute the Fourier coefficients of μ_{diag} as well as those of $f * g$. We first write out the Fourier coefficients of μ_{diag}. Let $\pi, \sigma \in \widehat{G}$ and let $i, j \in [\dim(\pi)], k, \ell \in [\dim(\sigma)]$. We have

$$\widehat{\mu_{\mathsf{diag}}}(\pi,\sigma)_{i,j,k,\ell} = \mathbb{E}\left[\widetilde{\pi}_{i,j}(X)\widetilde{\sigma}_{k,\ell}(X^{-1})\right] = \mathbb{E}\left[\widetilde{\pi}_{i,j}(X)\overline{\widetilde{\sigma}_{\ell,k}(X)}\right],$$

where we used $\widehat{\mu_{\mathsf{diag}}}(\pi,\sigma) = \mathbb{E}_{g_1,g_2}[\mu_{\mathsf{diag}}(g_1,g_2)\pi(g_1^{-1}) \otimes \sigma(g_2^{-1})]$ in the first equality and $\sigma(X^{-1}) = \overline{\sigma(X)^T}$ in the second equality. We now use Schur's orthogonality relations in Fact 3.2 to conclude that the RHS above is 1 if $\pi = \sigma$, $i = \ell$, $j = k$ and 0 otherwise, hence we get that

$$\widehat{\mu_{\mathsf{diag}}}(\pi,\sigma)_{i,j,k,\ell} = \begin{cases} 0 & \pi \neq \sigma \\ \mathbb{1}[i = \ell, j = k] & \pi = \sigma. \end{cases} \tag{5}$$

So it suffices to consider the terms $\pi = \sigma$ and $i = \ell, j = k$ in Eq. (4). We next write out the Fourier coefficients of $\widehat{(f * g)}(\pi, \pi)$. Similar to the convolution property in Fact 3.5, we have that

$$\begin{aligned}
&\widehat{(f * g)}(\pi, \pi)_{i,j,j,i} \\
&= \mathbb{E}\left[(f * g)(X, Y)\widetilde{\pi}_{i,j}(X^{-1})\widetilde{\pi}_{j,i}(Y^{-1})\right] \\
&= \mathbb{E}\left[f(A, C)g(B, D)\widetilde{\pi}_{i,j}(B^{-1}A^{-1})\widetilde{\pi}_{j,i}(D^{-1}C^{-1})\right] \\
&= \dim(\pi) \cdot \mathbb{E}\left[f(A, C)g(B, D) \cdot \pi(B^{-1}A^{-1})_{i,j} \cdot \pi(D^{-1}C^{-1})_{j,i}\right] \\
&= \frac{1}{\dim(\pi)} \cdot \mathbb{E}\left[f(A, C)g(B, D)\left(\sum_{k\in[\dim(\pi)]} \widetilde{\pi}_{i,k}(B^{-1})\widetilde{\pi}_{k,j}(A^{-1})\right)\left(\sum_{\ell\in[\dim(\pi)]} \widetilde{\pi}_{j,\ell}(D^{-1})\widetilde{\pi}_{\ell,i}(C^{-1})\right)\right] \\
&= \frac{1}{\dim(\pi)} \cdot \sum_{k,\ell\in[\dim(\pi)]} \widehat{f}(\pi, \pi)_{k,j,\ell,i} \cdot \widehat{g}(\pi, \pi)_{i,k,j,\ell}.
\end{aligned}$$

Putting together the above equality and Eq. (5) into Eq. (4), we get that

$$\begin{aligned}
\langle \mu_{\mathsf{diag}}, f * g\rangle &= \sum_{\pi,\sigma\in\widehat{G}} \sum_{\substack{i,j\in[\dim(\pi)]\\ k,\ell\in\dim(\sigma)}} \overline{\widehat{\mu_{\mathsf{diag}}}(\pi, \sigma)_{i,j,k,\ell}} \cdot \widehat{f * g}(\pi, \sigma)_{i,j,k,\ell} \\
&= \sum_{\pi\in\widehat{G}} \sum_{i,j\in[\dim(\pi)]} \widehat{f * g}(\pi, \pi)_{i,j,j,i} = \sum_{\pi\in\widehat{G}} \frac{1}{\dim(\pi)} \sum_{\substack{i,j\in[\dim(\pi)]\\ k,\ell\in[\dim(\pi)]}} \widehat{f}(\pi, \pi)_{k,j,\ell,i} \cdot \widehat{g}(\pi, \pi)_{i,k,j,\ell},
\end{aligned}$$

hence proving the claim statement. □

We next prove the following claim which follows from Theorem 3.6 and holds for all $d \in \mathbb{N}$.

Claim 5.5. *There exists global constants $c, C > 0$ such that the following holds. Let $f : \mathsf{SU}(N) \times \mathsf{SU}(N) \to \{0, 1\}$ and $\alpha = \mathbb{E}[f]$ be such that $\alpha \geq e^{-c\sqrt{N}}$. Then, for all $d \in \mathbb{N}$, we have*

$$\left\|f^{=d}\right\|_2^2 \leq (6C)^d\alpha^2 + \left(2C/d \cdot \log(1/\alpha)\right)^d\alpha^2$$

Proof. Fix global constants C, c as in Theorem 3.6. We show $C' = 6C$ and $c' \leq c$ that satisfy Claim 5.5. Let $d \leq c\sqrt{N}$. Let $K = \lceil e^{2d}\rceil$. Observe that $d \leq \log(K/\alpha)/2$ for all $\alpha \in [0, 1]$. We express $f : G \to \{0, 1\}$ as a sum of indicator functions $\{f_i : G \to \{0, 1\}\}_{i\in[K]}$ where each f_i satisfies $\mathbb{E}[f_i] = \alpha/K$. This can be done by partitioning the set corresponding to f into K parts, each of measure α/K. We now apply Theorem 3.6 to each f_i of level d to conclude that

$$\left\|f_i^{=d}\right\|_2^2 \leq \frac{1}{K^2d^d}C^d\alpha^2\log^d(K/\alpha). \tag{6}$$

We now add this inequality for all $i \in [K]$. We obtain the bound for levels $d \leq c\sqrt{N}$ from the

following calculation.

$$\begin{aligned}\left\|f^{=d}\right\|_2^2 &\leq K \cdot \sum_{i\in[K]} \|f_i^{=d}\|_2^2 \\ &\leq K^2 \cdot \frac{1}{K^2 d^d} C^d \alpha^2 \log^d(K/\alpha) \\ &\leq \frac{C^d}{d^d}\alpha^2 \left(3d + \log(1/\alpha)\right)^d \\ &\leq \frac{C^d}{d^d}\alpha^2 2^d \left((3d)^d + \log(1/\alpha)^d\right) \leq (6C)^d\alpha^2 + \left(2C/d \cdot \log(1/\alpha)\right)^d \alpha^2,\end{aligned}$$

where the first inequality Cauchy-Schwarz and second inequality used Eq. (6). For levels $d > c\sqrt{N}$, first observe that $\|f^{=d}\|_2^2 \leq \alpha$ by Parsevals identity. Now choose $c' \leq c$ to be a small enough constant so that when $\alpha \geq e^{-c'\sqrt{N}}$ we have $\alpha \leq (6C)^{c\sqrt{N}}\alpha^2$. This implies that we can upper bound $\|f^{=d}\|_2^2 \leq \alpha$ by $(6C)^d\alpha^2$ for levels $d > c\sqrt{N}$. This proves the claim statement. □

Using these two claims, we are now ready to prove our main Lemma 5.2.

Proof of Lemma 5.2. Consider the global constant c as in Claim 5.5 and let c' be sufficiently smaller than c. For $0 \leq d \leq N/2 - 1$, let L_d denote the set of all irreps of $V_{=d}$ and let $L_{N/2}$ denote the set of all the irreps of $V_{\geq N/2}$. From Claim 5.4, we have

$$\begin{aligned}\Delta &:= \mathbb{E}[f(A,C)g(B,(ABC)^{-1})] - \mathbb{E}[f(A,C)g(B,D)] \\ &= \sum_{\emptyset\neq\pi\in\widehat{G}} \frac{1}{\dim(\pi)} \sum_{\substack{i,j\in[\dim(\pi)]\\ k,\ell\in[\dim(\pi)]}} \widehat{f}(\pi,\pi)_{k,j,\ell,i} \cdot \widehat{g}(\pi,\pi)_{i,k,j,\ell}\end{aligned}$$

We now break up the contribution of various π in the above summation depending on the L_d to which they belong.

$$\Delta \leq \sum_{1\leq d\leq N/2,\, \emptyset\neq\pi\in L_d} \frac{1}{\dim(\pi)} \sum_{\substack{i,j\in[\dim(\pi)]\\ k,\ell\in[\dim(\pi)]}} \widehat{f}(\pi,\pi)_{k,j,\ell,i} \cdot \widehat{g}(\pi,\pi)_{i,k,j,\ell}$$

Since Q_d is the minimum dimension of a representation $\pi \in L_d$, we can lower bound $\dim(\pi) \geq Q_d$ above. Applying the Cauchy Schwarz on terms $\pi \in L_d$ we get that

$$\begin{aligned}\Delta &\leq \sum_{1\leq d\leq N/2,} \frac{1}{Q_d} \sum_{\emptyset\neq\pi\in L_d} \sum_{\substack{i,j\in[\dim(\pi)]\\ k,\ell\in[\dim(\pi)]}} \left|\widehat{f}(\pi,\pi)_{k,j,\ell,i}\right| \cdot |\widehat{g}(\pi,\pi)_{i,k,j,\ell}| \qquad (7)\\ &\leq \sum_{1\leq d\leq N/2,} \frac{1}{Q_d} \sqrt{\sum_{\emptyset\neq\pi\in L_d} \sum_{\substack{i,j\in[\dim(\pi)]\\ k,\ell\in[\dim(\pi)]}} \left|\widehat{f}(\pi,\pi)_{k,j,\ell,i}\right|^2} \cdot \sqrt{\sum_{\emptyset\neq\pi\in L_d} \sum_{\substack{i,j\in[\dim(\pi)]\\ k,\ell\in[\dim(\pi)]}} |\widehat{g}(\pi,\pi)_{i,k,j,\ell}|^2} \qquad (8)\\ &\leq \sum_{1\leq d\leq N/2} \frac{1}{Q_d} \left\|f^{=2d}\right\|_2 \left\|g^{=2d}\right\|_2 \qquad (9)\end{aligned}$$

We now analyze the expression above by considering the two cases $d \leq c'\sqrt{N}/2$ or $d > c'\sqrt{N}/2$. For notational convenience, let $\alpha = \mathbb{E}[f]$ and $\beta = \mathbb{E}[g]$.

Contribution from levels $d \le c'\sqrt{N}/2$. We will show that the contribution is at most $\alpha\beta/30$ by applying the Level-k Inequality in Claim 5.5. Since $\alpha, \beta \ge e^{-c'\sqrt{N}}$ and $c' \ll c$, we can apply Claim 5.5 to f, g. This implies that for all $d \in \mathbb{N}$, we have

$$\left\|f^{=2d}\right\|_2^2 \le C^{2d}\alpha^2 + \frac{C^{2d}}{(2d)^{2d}}\alpha^2(c'\sqrt{N})^{2d} \quad \text{and} \quad \left\|g^{=2d}\right\|_2^2 \le C^{2d}\beta^2 + \frac{C^{2d}}{(2d)^{2d}} \cdot \beta^2(c'\sqrt{N})^{2d}.$$

Since $2d \le c'\sqrt{N}$, the second term in the above inequalities dominates. We now use the fact that $Q_d \ge (cN/d)^d$ from Theorem 3.7. Thus, we have

$$\frac{1}{Q_d}\left\|f^{=2d}\right\| \cdot \left\|g^{=2d}\right\| \le \frac{d^d}{(cN)^d} \cdot \alpha\beta \cdot 2C^{2d}\frac{(c'\sqrt{N})^{2d}}{(2d)^{2d}} \le \alpha\beta/(50)^d$$

since c' is a sufficiently small constant. Thus, the contribution from the levels $d \le c'\sqrt{N}$ is at most $\alpha\beta \cdot \sum_d 50^{-d} \le \alpha\beta/30$.

Contribution from levels $d > c'\sqrt{N}/2$. We will show that the contribution is at most $2\alpha\beta/30$. For these levels, we use the trivial bound $\|f^{=d}\|_2 \le \sqrt{\alpha}, \|g^{=d}\|_2 \le \sqrt{\beta}$ from Parseval's identity. We will need to handle $d \le cN/10$ and $d > cN/10$ separately. For $d \le cN/10$, Theorem 3.7 implies that $Q_d \ge (cN/d)^d$ and since $\alpha\beta \ge e^{-c'\sqrt{N}} \gg 10^{-c'\sqrt{N}}$, we have

$$\frac{1}{Q_d}\sqrt{\alpha\beta} \le \sqrt{\alpha\beta} \cdot \left(\frac{d}{cN}\right)^d \le \sqrt{\alpha\beta} \cdot \left(\frac{1}{10}\right)^{c'\sqrt{N}} \le \alpha\beta/(30N).$$

For $d \ge cN/10$, Theorem 3.7 implies that[4] $Q_d \ge (1+c)^{cd/(1+c)} \ge e^{-\Omega(N)}$. Since $\sqrt{\alpha\beta} \ge e^{-c'\sqrt{N}}$, we have

$$\frac{1}{Q_d}\sqrt{\alpha\beta} \le \alpha\beta/(30N).$$

The same calculation works for levels $\ge N/2$ and since $\|f^{\ge N/2}\|_2 \le \sqrt{\alpha}, \|g^{\ge N/2}\|_2 \le \sqrt{\beta}$, we have

$$\frac{1}{Q_{\ge N/2}}\sqrt{\alpha\beta} \le \alpha\beta/(30N)$$

Adding this over all possible $d \ge c'\sqrt{N}/2$, it follows that the contribution from levels $d \ge c'\sqrt{N}/2$ is at most $2\alpha\beta/30$. Substituting these in Eq. (7), we have $\Delta \le \alpha\beta/30$. □

References

[ABK23] Scott Aaronson, Harry Buhrman, and William Kretschmer. A qubit, a coin, and an advice string walk into a relational problem. *Electron. Colloquium Comput. Complex.*, TR23, 2023.

[BBM12] Eric Blais, Joshua Brody, and Kevin Matulef. Property testing lower bounds via communication complexity. *computational complexity*, 21(2):311–358, 2012.

[4] This is because for levels $d \le cN/(1+c)$, we have $(cN/d)^d \ge (1+c)^d \ge (1+c)^{cd/(1+c)}$ and for levels $d > cN/(1+c)$ we have $(1+c)^{cN/(1+c)} \ge (1+c)^{cd/(1+c)}$.

[BCW98] Harry Buhrman, Richard Cleve, and Avi Wigderson. Quantum vs. classical communication and computation. In *Proceedings of the thirtieth annual ACM symposium on Theory of computing*, pages 63–68, 1998.

[BCWW01] Harry Buhrman, Richard Cleve, John Watrous, and Ronald de Wolf. Quantum fingerprinting. *Physical Review Letters*, 87(16):167902, 2001.

[BG08] Jean Bourgain and Alex Gamburd. Uniform expansion bounds for cayley graphs of. *Annals of Mathematics*, pages 625–642, 2008.

[BJK08] Ziv Bar-Yossef, Thathachar S Jayram, and Iordanis Kerenidis. Exponential separation of quantum and classical one-way communication complexity. *SIAM Journal on Computing*, 38(1):366–384, 2008.

[BRW08] Avraham Ben-Aroya, Oded Regev, and Ronald de Wolf. A hypercontractive inequality for matrix-valued functions with applications to quantum computing and ldcs. In *2008 49th Annual IEEE Symposium on Foundations of Computer Science*, pages 477–486. IEEE, 2008.

[CM18] Chris Cade and Ashley Montanaro. The quantum complexity of computing schatten p-norms. In *13th Conference on the Theory of Quantum Computation, Communication and Cryptography, TQC 2018, July 16-18, 2018, Sydney, Australia*, volume 111 of *LIPIcs*, pages 4:1–4:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2018.

[DM20] João F Doriguello and Ashley Montanaro. Exponential quantum communication reductions from generalizations of the boolean hidden matching problem. *arXiv preprint arXiv:2001.05553*, 2020.

[EKLM23] David Ellis, Guy Kindler, Noam Lifshitz, and Dor Minzer. Product mixing in compact Lie groups, 2023. Preprint on electronic colloquium on computational complexity: TR23-133.

[FMP+15] Samuel Fiorini, Serge Massar, Sebastian Pokutta, Hans Raj Tiwary, and Ronald de Wolf. Exponential lower bounds for polytopes in combinatorial optimization. *Journal of the ACM (JACM)*, 62(2):1–23, 2015.

[Gav16] Dmitry Gavinsky. Entangled simultaneity versus classical interactivity in communication complexity. In Daniel Wichs and Yishay Mansour, editors, *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2016, Cambridge, MA, USA, June 18-21, 2016*, pages 877–884. ACM, 2016.

[Gav19] Dmitry Gavinsky. Quantum versus classical simultaneity in communication complexity. *IEEE Transactions on Information Theory*, 65(10):6466–6483, 2019.

[Gav20] Dmitry Gavinsky. Bare quantum simultaneity versus classical interactivity in communication complexity. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, pages 401–411, 2020.

[GKdW06] Dmytro Gavinsky, Julia Kempe, and Ronald de Wolf. Strengths and weaknesses of quantum fingerprinting. 2006.

[GKK+07] Dmitry Gavinsky, Julia Kempe, Iordanis Kerenidis, Ran Raz, and Ronald 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*, pages 516–525, 2007.

[GKRW06] Dmitry Gavinsky, Julia Kempe, Oded Regev, and Ronald de Wolf. Bounded-error quantum state identification and exponential separations in communication complexity. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of Computing*, pages 594–603, 2006.

[Gow08] W. T. Gowers. Quasirandom Groups. *Combin. Probab. Comput.*, 17:363–387, 2008.

[GRT22] Uma Girish, Ran Raz, and Avishay Tal. Quantum versus randomized communication complexity, with efficient players. *Comput. Complex.*, 31(2):17, 2022.

[GV15] Timothy Gowers and Emanuele Viola. The communication complexity of interleaved group products. In *Proceedings of the forty-seventh annual ACM symposium on Theory of Computing*, pages 351–360, 2015.

[HN12] Trinh Huynh and Jakob Nordstrom. On the virtue of succinct proofs: Amplifying communication complexity hardness to time-space trade-offs in proof complexity. In *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*, pages 233–248, 2012.

[KKD19] Niraj Kumar, Iordanis Kerenidis, and Eleni Diamanti. Experimental demonstration of quantum advantage for one-way communication complexity surpassing best-known classical protocol. *Nature communications*, 10(1):4152, 2019.

[KKS14] Michael Kapralov, Sanjeev Khanna, and Madhu Sudan. Streaming lower bounds for approximating max-cut. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1263–1282. SIAM, 2014.

[KL98] E. Knill and R. Laflamme. Power of one bit of quantum information. *Phys. Rev. Lett.*, 81:5672–5675, Dec 1998.

[KL19] Hartmut Klauck and Debbie Lim. The power of one clean qubit in communication complexity. 2019.

[KLM22] Peter Keevash, Noam Lifshitz, and Dor Minzer. On the largest product-free subsets of the alternating groups. *arXiv preprint arXiv:2205.15191*, 2022.

[KRW95] Mauricio Karchmer, Ran Raz, and Avi Wigderson. Super-logarithmic depth lower bounds via the direct sum in communication complexity. *Computational Complexity*, 5(3):191–204, 1995.

[KW90] Mauricio Karchmer and Avi Wigderson. Monotone circuits for connectivity require super-logarithmic depth. *SIAM J. Discret. Math.*, 3(2):255–265, 1990.

[MFF14] Tomoyuki Morimae, Keisuke Fujii, and Joseph F. Fitzsimons. Hardness of classically simulating the one-clean-qubit model. *Phys. Rev. Lett.*, 112:130502, Apr 2014.

[MNSW95] Peter Bro Miltersen, Noam Nisan, Shmuel Safra, and Avi Wigderson. On data structures and asymmetric communication complexity. In *Proceedings of the twenty-seventh annual ACM symposium on Theory of computing*, pages 103–111, 1995.

[Mon10] Ashley Montanaro. A new exponential separation between quantum and classical one-way communication complexity. *arXiv preprint arXiv:1007.3587*, 2010.

[Raz99] Ran Raz. Exponential separation of quantum and classical communication complexity. In *Proceedings of the thirty-first annual ACM symposium on Theory of computing*, pages 358–367, 1999.

[RK11] Oded Regev and Bo'az Klartag. Quantum one-way communication can be exponentially stronger than classical communication. In *Proceedings of the 43rd ACM Symposium on Theory of Computing, San Jose, CA, USA, 6-8 June 2011*, pages 31–40. ACM, 2011.

[Sco23] Aaronson Scott. Talk: Verifiable quantum supremacy: What i hope will be done, 2023. https://www.youtube.com/watch?v=A6YPAQlGejo&ab_channel=SimonsInstitute.

[SJ08] Peter W. Shor and Stephen P. Jordan. Estimating jones polynomials is a complete problem for one clean qubit. *Quantum Info. Comput.*, 8(8):681–714, sep 2008.

[SWY12] Yaoyun Shi, Xiaodi Wu, and Wei Yu. Limits of quantum one-way communication by matrix hypercontractive inequality. 2012.

[SX91] P. Sarnak and X. Xue. Bounds for multiplicities of automorphic representations. *Duke Math. J.*, 64(1):207–227, 1991.

[Yao79] Andrew Chi-Chih Yao. Some complexity questions related to distributive computing (preliminary report). In *Proceedings of the eleventh annual ACM symposium on Theory of computing*, pages 209–213, 1979.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il