Approximate orthogonality of permutation operators, with application to quantum information

Citation: Letters in Mathematical Physics. 2023 Dec 16;114(1):1

As Published: https://doi.org/10.1007/s11005-023-01744-1

Publisher: Springer Netherlands

Persistent URL: https://hdl.handle.net/1721.1/153193

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Approximate orthogonality of permutation operators, with application to quantum information

Approximate orthogonality of permutation operators, with application to quantum information

Aram W. Harrow

Center for Theoretical Physics, Massachusetts Institute of Technology,
aram@mit.edu.

Abstract

Consider the $n!$ different unitary matrices that permute n d-dimensional quantum systems. If $d \geq n$ then they are linearly independent. This paper discusses a sense in which they are approximately orthogonal (with respect to the Hilbert-Schmidt inner product, $\langle A, B\rangle = \operatorname{tr} A^\dagger B / \operatorname{tr} I$) if $d \gg n^2$, or, in a different sense, if $d \gg n$. Previous work had shown pairwise approximate orthogonality of these matrices, but here we show a more collective statement, quantified in terms of the operator norm distance of the Gram matrix to the identity matrix. This simple point has several applications in quantum information and random matrix theory: (1) showing that random maximally entangled states resemble fully random states, (2) showing that Boson sampling output probabilities resemble those from Gaussian matrices, (3) improving the Eggeling-Werner scheme for multipartite data hiding, (4) proving that the product test of Harrow-Montanaro cannot be performed using LOCC without a large number of copies of the state to be tested, (5) proving that the purity of a quantum state also cannot be efficiently tested using LOCC, and (6, published separately with Brandão and Horodecki) helping prove that poly-size random quantum circuits are poly-designs.

Dedicated to the memory of Mary Beth Ruskai

1 Introduction

1.1 Permutations and quantum states

When quantum states have symmetries under permutation or collective rotation, it is possible to reduce the number of parameters in a problem. But this may come at a cost in complexity, for example if the small number of parameters label basis states in irreducible representations which lack simple constructions.

The main focus of the paper is inspired by the following point: matrices on $(\mathbb{C}^d)^{\otimes n}$ that commute with collective unitary rotations are known to be linear combinations of the permutations of the n qudits (see below for precise definitions). If $d \gg n$, then this indeed reduces the number of parameters from d^{2n} to $n!$. However, these parameters are coefficients of permutation matrices that are not quite orthogonal to one another (again, in a sense that we will clarify below). We will argue that they are *almost* orthogonal, in a manner that suffices for most applications, when $d \gg n^2$.

To make these claims more precise, we introduce some definitions. Denote the symmetric group on n elements by $\mathcal{S}_n$. This has a representation P_d on $(\mathbb{C}^d)^{\otimes n}$ in which the n qudits are permuted. Formally, if $\pi \in \mathcal{S}_n$, then

$$P_d(\pi) = \sum_{i_1,\ldots,i_n \in [d]} |i_1, \ldots, i_n\rangle \langle i_{\pi(1)}, \ldots, i_{\pi(n)}| , \tag{1}$$

where $[d] := \{1, \ldots, d\}$. The definition is chosen so that $P_d(\pi_1)P_d(\pi_2) = P_d(\pi_1\pi_2)$, that is, P_d is a representation.

Let M_d denote the set of $d \times d$ complex matrices and $\mathcal{U}_d$ the subset of unitary matrices. One place where the permutation matrices arise is when considering operators $A \in M_d^{\otimes n}$ that commute with every $X^{\otimes n}$ for $X \in \mathcal{U}_d$. Such A can be written as (see Thm 4.1.13 of [1] or Cor 4 of [2])

$$A = \sum_{\pi \in \mathcal{S}_n} a_\pi P_d(\pi), \tag{2}$$

for some coefficients $a_\pi \in \mathbb{C}$. This decomposition is useful because it reduces the number of parameters needed to describe A. However, it is inconvenient that the terms in (2) are not orthogonal. We will see this in more detail in our applications below, where (2) becomes useful precisely when we can establish an approximate orthogonality for the $P_d(\pi)$ matrices.

We will use the following normalized Hilbert-Schmidt inner product:

$$\langle A, B\rangle := \frac{\operatorname{tr} A^\dagger B}{\operatorname{tr} I} = \frac{\operatorname{tr} A^\dagger B}{d^n} = \langle \Phi_d|^{\otimes n} (I \otimes A^\dagger B) |\Phi_d\rangle^{\otimes n} , \tag{3}$$

where $|\Phi_d\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^d |i, i\rangle$ is the standard maximally entangled state. We also use the convention that $\psi := |\psi\rangle \langle\psi|$.

1.2 Overview of results

In Section 2 we will show that the $n!$ different $P_d(\pi)$ are approximately orthogonal, not only pairwise (which is rather trivial to show), but in a certain global sense as well. Specifically we define the $n!$-dimensional Gram matrix $G^{(n,d)}$ whose π_1, π_2 entry is the inner product $\langle P_d(\pi_1), P_d(\pi_2)\rangle$ and argue that it is close to the identity in operator norm. Previous work typically focused on entrywise bounds on G^{-1}, and while these often obtained much sharper results, our applications will rely on the operator norm estimates presented here.

One easy consequence of this approximate orthogonality relation is to controlling various norms of linear combinations of permutation matrices, as we discuss in Section 3. A less obvious, but still easy, application is showing that the lower moments of random bipartite states are close to the moments of random maximally entangled states, which we describe in Section 4. This is related to the well-known fact that the entries of Haar-random unitaries appear to be nearly Gaussian, again when we examine only the low moments. This in turn has application to improving the parameters in boson sampling, as we discuss in Section 5.

The next family of results involves limitations on multi-party quantum operations where the parties are connected by only classical communication. More generally, we consider measurements that remain valid when the partial transpose operator is applied to a subset of systems, and that commute with rotations of the form $U^{\otimes n}$. It turns out that these operators are severely constrained and we use this to analyze the Eggeling-Werner data hiding scheme and the complexity of purity testing in Section 7 and establish limitations on product tests in Section 8. A further application has appeared in [3], where this approximate orthogonality is used to analyze the convergence speed of the low-order moments of random unitary quantum circuits.

Two appendices explore further topics. Appendix A fleshes out some calculations used in Section 2 and Appendix B explains how replacing Haar uniform unitaries with other distributions, such as random classical reversible operations, does not yield the same structure.

2 Approximate orthogonality

2.1 Statement of results

This section gives a quantitative statement of the approximate orthogonality of permutation operators.

First we relate the inner product between a pair of permutations to a natural metric on the group of permutations. Observe that

$$\operatorname{tr} P_d(\pi) = d^{c(\pi)}, \tag{4}$$

where $c(\pi)$ counts the number of cycles of π. Let $T_n \subset S_n$ be the set of $\binom{n}{2}$ transpositions, and let $\Gamma_n := \Gamma(S_n, T_n)$ be the Cayley graph of S_n defined by this generating set; i.e. the vertices are S_n and there is an edge between π_1 and π_2 iff $\pi_1^{-1}\pi_2 \in T_n$. Define $|\pi|$ to be the minimum number of transpositions necessary to obtain π from e.

Since graph distance is invariant under multiplication by S_n, $|\cdot|$ satisfies the triangle inequality:

$$|\pi_1\pi_2| \leq |\pi_1| + |\pi_2|, \tag{5}$$

Observe also that $|\pi| = n - c(\pi)$. We now calculate

$$\langle P_d(\pi_1), P_d(\pi_2)\rangle = \frac{\operatorname{tr} P_d(\pi_1^{-1}\pi_2)}{d^n} = d^{c(\pi_1^{-1}\pi_2)-n} = d^{-|\pi_1^{-1}\pi_2|}. \tag{6}$$

Thus, $P_d(\pi_1)$ and $P_d(\pi_2)$ are approximately orthonormal when d is large and/or when π_1 and π_2 are far apart in the transposition metric.

The main goal of this paper is to extend the pairwise approximate orthogonality of (6) to a certain notion of global approximate orthogonality. In particular we will show that the $P_d(\pi)$ are close to an orthonormal basis. In general, a collection of vectors with pairwise small inner products does not have to be close to an orthonormal basis, as we will discuss further in Appendix B. The key fact we will use about the $P_d(\pi)$ matrices is that they are close to an orthonormal basis.

Define the $n! \times n!$ Gram matrix $G^{(n,d)}$ by

$$G^{(n,d)}_{\pi_1,\pi_2} = \langle P_d(\pi_1), P_d(\pi_2)\rangle = d^{-|\pi_1^{-1}\pi_2|}, \tag{7}$$

Observe that $G^{(n,d)}$ has ones on the diagonal, and positive powers of $1/d$ in every off-diagonal entry. Thus we have

$$\lim_{d\to\infty} G^{(n,d)} = I_{n!}, \tag{8}$$

corresponding to the fact that different permutations approach orthogonality as $d \to \infty$.

To make this fact useful, we need to know how quickly this limit converges as a function of n. Naively, we can observe that there are $n!-1$ off-diagonal terms per row, each $\leq 1/d$, so they add up to something small if $d \gg n!$. But much better bounds are possible.

Lemma 1 (approximate orthogonality)**.**

1. *$G^{(n,d)}$ is always positive semidefinite, has trace $n!$, and is invertible if and only if $n \leq d$.*
2. $$\frac{1}{n!}\|G^{(n,d)} - I_{n!}\|_1 \leq \sqrt{2}\frac{n}{d}. \tag{9}$$
3. $$\lambda_{\min}(G^{(n,d)}) = \prod_{j=1}^{n-1}\left(1 - \frac{j}{d}\right) \geq 1 - \frac{n(n-1)}{2d} \tag{10a}$$
$$\lambda_{\max}(G^{(n,d)}) = \prod_{j=1}^{n-1}\left(1 + \frac{j}{d}\right) \leq e^{\frac{n(n-1)}{2d}} \tag{10b}$$

Our applications will mostly rely on the following simplified bounds:

$$\|G^{(n,d)} - I_{n!}\|_\infty \leq \frac{n^2}{d} \qquad \textit{if } n^2 \leq d \tag{11}$$

$$\|G^{(n,d)} - I_{n!}\|_{1\to 1} \leq e^{\frac{n^2}{2d}} - 1, \tag{12}$$

where the 1 → 1 *norm of a matrix means the maximum sum of absolute values of entries of any row.*

We see that there are a few different regimes. If $n > d$, then $G^{(n,d)}$ is singular and is far from $I_{n!}$. Because of the qualitative difference between the $n > d$ and $n \leq d$ regimes, the $n \leq d$ case is referred to as the "stable range" in the context of Schur-Weyl duality. If $n \leq O(d)$, then the *average* eigenvalue of $G^{(n,d)}$ is close to 1, even though the top and bottom eigenvalues will be exponentially large and exponentially close to zero respectively. Finally, if $n \leq O(\sqrt{d})$, then $G^{(n,d)}$ will be close to $I_{n!}$ in operator norm.

There are two proofs of Lemma 1, both requiring some facts from representation theory. Using precise statements about the dimensions of irreps of $\mathcal{U}_d$ and $\mathcal{S}_n$, we can calculate the exact formula for the eigenvalues of G and their multiplicities. We will do this below in Lemma 2. However, part 1 of Lemma 1 and eqs. (10b) and (11) can also be proved using only a few simple facts about the symmetric and antisymmetric subspaces. We give this proof here.

First we recall some facts about the symmetric and antisymmetric subspaces. Define $\vee^n\mathbb{C}^d$ to be the symmetric subspace of $(\mathbb{C}^d)^{\otimes n}$, meaning the set of vectors that is invariant under each $P_d(\pi)$. We will also use the antisymmetric subspace $\wedge^n\mathbb{C}^d$, which is the set of vectors invariant under each $P_d(\pi)\,\mathrm{sgn}(\pi)$, where $\mathrm{sgn}(\pi)$ is defined to be the sign of π. The dimensions of these subspaces are known to be given by $\dim \vee^n\mathbb{C}^d = \binom{d+n-1}{n} =: d[n]$ and $\dim \wedge^n\mathbb{C}^d = \binom{d}{n}$. For readers unfamiliar with the properties of the symmetric subspace, Ref. [2] gives a review from a quantum-information perspective.

Proof of parts 1 and 3 of Lemma 1. For part 1, we observe that G is a Gram matrix, so is automatically positive semi-definite. It has dimension $n!$ and ones along its diagonal, so G has trace $n!$. It is invertible if and only if the matrices $P_d(\pi)$ are linearly independent. If $n \leq d$, then the linear independence of these matrices can be seen by considering their action on the state $|1\rangle \otimes |2\rangle \otimes \cdots \otimes |n\rangle \in (\mathbb{C}^d)^{\otimes n}$. To show that G is singular when $n > d$, we define the vector $|\zeta\rangle := \sqrt{\frac{d^n}{n!}} \sum_{\pi\in\mathcal{S}_n} \mathrm{sgn}(\pi)\,|\pi\rangle$. Now calculate

$$\begin{aligned}\langle\zeta|\,G\,|\zeta\rangle &= \frac{1}{n!}\sum_{\pi_1,\pi_2} d^{c(\pi_1^{-1}\pi_2)}\,\mathrm{sgn}(\pi_1)\,\mathrm{sgn}(\pi_2)\\ &= \sum_{\pi\in\mathcal{S}_n} d^{c(\pi)}\,\mathrm{sgn}(\pi)\\ &= \sum_{\pi\in\mathcal{S}_n} \mathrm{tr}\, P_d(\pi)\,\mathrm{sgn}(\pi)\end{aligned}$$

$$= \dim \wedge^n \mathbb{C}^d = \binom{d}{n}.$$

When $n > d$, this expression is 0. Since G is positive semidefinite, it follows that it must have an eigenvalue equal to 0.

For part 3, we observe that the sum of the π_1 row of G is

$$\begin{aligned}
\sum_{\pi_2 \in \mathcal{S}_n} G_{\pi_1,\pi_2} &= \sum_{\pi_2 \in \mathcal{S}_n} d^{c(\pi_1^{-1}\pi_2)-n} && \text{(13a)}\\
&= \sum_{\pi \in \mathcal{S}_n} d^{c(\pi)-n} && \text{(13b)}\\
&= d^{-n} \sum_{\pi \in \mathcal{S}_n} \operatorname{tr} P_d(\pi) && \text{(13c)}\\
&= \frac{n!}{d^n} d[n] = \frac{d+n-1!}{d! \cdot d^n} && \text{(13d)}\\
&= \prod_{j=1}^{n-1} \left(1 + \frac{j}{d}\right) && \text{(13e)}
\end{aligned}$$

Finally, we use the inequality $1 + x \leq e^x$ (which holds for all x) to upper-bound the last equation with $e^{\frac{n(n-1)}{2d}}$. This yields (12) which implies (11) and in turn (10b).

□

Remark 1. An even simpler proof of a nearly equivalent bound was found by Kevin Zatloukal. The idea is that $|\cdot|$ describes a metric on a Cayley graph of degree $\binom{n}{2}$. Thus, there are at most $\binom{n}{2}^k$ permutations with $|\pi| = k$, and we have

$$\sum_{\pi \in \mathcal{S}_n} d^{-|\pi|} \leq \sum_{k \geq 0} \binom{n}{2}^k d^{-k} = \left(1 - \frac{\binom{n}{2}}{d}\right)^{-1}.$$

Most of the rest of the paper is devoted to applications of (11). For our applications, we do not need any more precise information about the distribution of eigenvalues. However, for completeness, we will describe the exact spectrum of $G^{(n,d)}$. The answer turns out to involve the representation theory of the symmetric and unitary groups.

Lemma 2. *For each $\lambda \in \mathrm{Par}(n,d)$, $G^{(n,d)}$ has $\dim^2 \mathcal{P}_\lambda$ eigenvalues, each equal to*

$$\frac{n! \dim \mathcal{Q}_\lambda^d}{\dim \mathcal{P}_\lambda d^n} = \prod_{(i,j)\in\lambda} \left(1 + \frac{j-i}{d}\right). \tag{14}$$

Here $\mathcal{Q}_\lambda^d$ and $\mathcal{P}_\lambda$ are irreps of $\mathcal{U}_d$ and $\mathcal{S}_n$ respectively; see Appendix A for full details. From Lemma 2, we immediately obtain Parts 1 and 3 of Lemma 1. Part 2 is nontrivial, but was previously derived in Lemma 6 of [4]. That paper also gave asymptotically matching lower bounds on $\|G - I\|_1$ that we will omit here.

Lemma 2 has been proven previously [5–7] but using different techniques. In Appendix A we give a new proof using the terminology of quantum information, and based on the fact that G is a Gram matrix.

2.2 Related work

As noted above, Lemma 2 has been previously proved, in several places [5–8]. Ref. [7] used the same representation-theoretic argument, pointing out that it can be applied to any representation of any group, while [5, 6] used properties of the symmetric group to obtain a simple, nearly self-contained, calculation.

Define the *Weingarten matrix* $\mathrm{Wg}^{(n,d)}$ to be $(G^{(n,d)})^{-1}$. The Weingarten matrix was first introduced by Collins and Śniady [8, 9], although they used a different normalization convention. Their goals were to calculate matrix elements of $\mathbb{E}[(U \otimes \bar{U})^{\otimes n}]$ and to derive asymptotic properties such as freeness for related families of random matrices. The relevance of the Weingarten matrix can be seen from Cor 2.4 of Ref. [8], which gives the following exact expression for this expectation value:

$$\mathop{\mathbb{E}}_{U}[U^{\otimes n} \otimes U^{*,\otimes n}] = \sum_{\sigma,\tau\in\mathcal{S}_n} (I \otimes P_d(\sigma))\Phi_d^{\otimes n}(I \otimes P_d(\tau))^\dagger \, \mathrm{Wg}^{(n,d)}(\sigma,\tau) \, |v_\sigma\rangle \langle v_\tau| \, . \tag{15}$$

Following [9], note that $\mathrm{Wg}(\sigma,\tau)$ only depends on $\sigma^{-1}\tau$ so we can also denote this matrix element by $\mathrm{Wg}(\sigma^{-1}\tau)$ and we can refer to $\mathrm{Wg}(\cdot)$ as the Weingarten function. See [10] for an accessible recent review, and [11] for a discussion of applications.

Several papers have studied the asymptotic behavior of Wg as $d \to \infty$. Ref. [8] (in Cor 2.7) derived its leading order behavior:

$$\mathrm{Wg}^{(n,d)}(\sigma) = \mathrm{Moeb}(\sigma) d^{-n-|\sigma|} + O_n(d^{-n-|\sigma|-2}), \tag{16}$$

where $\mathrm{Moeb}(\sigma)$ is the Möbius function. If σ has c_k cycles of length k and $C_k := 2k!/k!(k+1)!$ is the k^{th} Catalan number then $\mathrm{Moeb}(\sigma) := \prod_k((-1)^{k-1}C_k)^{c_{k+1}}$. Since $\mathrm{Moeb}(e) = 1$, (16) is another way of saying that in the $d \to \infty$ limit, Wg (or equivalently G) approaches the identity matrix. (16) does not address the question of how large d needs to be as a function of n in order for the approximation to be accurate. Later works addressed this question, culminating in [12] which showed that (16) is nearly sharp when $d \gg n^{7/4}$. However, as a pointwise estimate on the entries of a rank-$n!$ matrix, this does not immediately imply bounds on the spectrum.

To get some intuition for (15) in the regime where $\mathrm{Wg} \approx I$, we will compare with the case of Gaussian random matrices. Let X be a random complex $d \times d$ Gaussian matrix whose entries are i.i.d. and satisfy $\mathbb{E}[X_{ij}] = 0$ and $\mathbb{E}[|X_{ij}|^2] = 1/d$. The following formula is known as Wick's theorem (or Isserlis' theorem):

$$\mathop{\mathbb{E}}_{X}[X^{\otimes n} \otimes X^{*,\otimes n}] = \sum_{\pi\in\mathcal{S}_n} (I \otimes P_d(\pi))\Phi_d^{\otimes n}(I \otimes P_d(\pi))^\dagger \tag{17}$$

This resembles (15) but with Wg replaced by the identity matrix. Thus (15) and the fact that $\mathrm{Wg} \approx I$ together imply that low moments of unitary matrices are close to moments of complex Gaussian matrices. This can be seen as a generalization of the Poincaré-Maxwell-Borel Lemma which states that applying a low-dimensional projector to a uniformly random point in a high-dimensional sphere yields an approximately Gaussian distribution. Indeed $\mathrm{Wg} \approx I \approx G$ precisely in the regime where submatrices of unitary matrices look Gaussian. Similar observations were made earlier by Novak [13] and Matsumoto [14]. We explore this point further in Sections 4 and 5.

3 Spectra and norms of sums of permutations

One easy consequence of our main result (Lemma 1) is that we can control various norms of sums of permutations. Suppose that $\epsilon = n^2/d \leq 1$ and consider some operator $A = \sum_\pi a_\pi P_d(\pi)$ with $a_\pi \in \mathbb{R}$. We would like to estimate various norms of A.

The 2-norm is:

$$\frac{\operatorname{tr} A^2}{d^n} = \langle a, Ga \rangle, \tag{18}$$

which is $\in [1-\epsilon/2, e^{\epsilon/2}]\|a\|_2^2$. This follows directly from the operator inequalities $(1-\epsilon/2)I \leq G \leq e^{\epsilon/2} I$.

To bound the ∞-norm of A, let $\pi = \arg\max_\pi |a_\pi|$. Then

$$\frac{|\operatorname{tr}[P_d(\pi)^\dagger A]|}{d^n} = \left| \sum_{\sigma \in \mathcal{S}_n} a_\sigma G_{\pi,\sigma} \right| \geq |a_\pi| \left(1 - \sum_{\sigma \neq \pi} G_{\pi,\sigma} \right) \geq \|a\|_\infty (1-\epsilon), \tag{19}$$

where the last inequality follows from (12). Using $\langle A, B \rangle \leq \|A\|_\infty \|B\|_1$ with $B = P_d(\pi)/d^n$, we obtain

$$\|A\|_\infty \geq (1-\epsilon/2)\|a\|_\infty. \tag{20}$$

On the other hand, the only obvious upper bound is the trivial $\|A\|_\infty \leq \|a\|_1$. This is tight when the a_π all have the same sign, or when the $a_\pi \operatorname{sgn}(\pi)$ do. Similarly we obtain

$$\|A\|_1 \geq (1-\epsilon) d^n \|a\|_\infty. \tag{21}$$

This method does not seem to yield good bounds on the 1-norm of A. The triangle inequality yields the rather weak bound $\|A\|_1 \leq d^n \|a\|_1$ which is usually improved upon by $\|A\|_1 \leq \sqrt{d^n}\|A\|_2 \leq d^n \sqrt{1+\epsilon}\|a\|_2$.

4 Random maximally entangled states

Random pure states are known to be nearly maximally entangled. This is an easy consequence of random matrix theory but was first discussed in the context of quantum states by Page [15]. Ref. [16] introduced many applications to quantum information theory were discovered [16]; see also [17] for a comprehensive review.

In this section we describe one way to formalize this intuition, by proving that the low moments of random bipartite states resemble those of random maximally entangled states.

Theorem 3. *Let $|\psi\rangle$ be a random unit vector in $\mathbb{C}^{d^2}$, U a Haar-uniform unitary in $\mathcal{U}_d$ and $|\varphi_U\rangle := (U \otimes I)\,|\Phi_d\rangle$. If $n^2 \leq d$ then*

$$\left(1 - \frac{n^2}{2d}\right) \mathop{\mathbb{E}}_{U}[\varphi_U^{\otimes n}] \leq \mathop{\mathbb{E}}_{\psi}[\psi^{\otimes n}] \leq \left(1 + \frac{n^2}{d}\right) \mathop{\mathbb{E}}_{U}[\varphi_U^{\otimes n}] \tag{22}$$

Again we use the convention that $\psi := |\psi\rangle\langle\psi|$. The proof follows immediately from the representation-theory facts in Appendix A).

Proof. Using Schur's Lemma we can derive expressions for both sides of Equation (22).

$$\mathop{\mathbb{E}}_{\psi}[\psi^{\otimes n}] = \frac{1}{\binom{d^2+n-1}{n}} \sum_{\lambda \in \mathrm{Par}(n,d)} |\lambda,\lambda\rangle\langle\lambda,\lambda| \otimes I_{\mathcal{Q}_\lambda^d}^{\otimes 2} \otimes \Phi_{\mathcal{P}_\lambda}. \tag{23}$$

$$\mathop{\mathbb{E}}_{U}[\varphi_U^{\otimes n}] = \sum_{\lambda \in \mathrm{Par}(n,d)} \frac{\dim \mathcal{P}_\lambda}{\dim \mathcal{Q}_\lambda^d \cdot d^n} |\lambda,\lambda\rangle\langle\lambda,\lambda| \otimes I_{\mathcal{Q}_\lambda^d}^{\otimes 2} \otimes \Phi_{\mathcal{P}_\lambda}. \tag{24}$$

The ratio between these coefficients, for a fixed λ, is the the same one appearing in Equation (14), namely

$$\frac{\dim \mathcal{Q}_\lambda^d n! d^n}{\dim \mathcal{P}_\lambda\, d^2 \cdots (d^2+n-1)} = \prod_{k=1}^{n-1} \left(1 + \frac{k}{d^2}\right)^{-1} \cdot \prod_{(i,j)\in\lambda} \left(1 + \frac{j-i}{d}\right). \tag{25}$$

This is again $\geq 1 - n^2/2d$ and $\leq 1 + n^2/d$, assuming $n^2 \leq d$. □

As an example, suppose that $d = d_A d_B$ corresponding to a decomposition into two systems A and B. One way to estimate entanglement is via the second moment: $\mathbb{E}[\operatorname{tr} \psi_A^2]$. While an exact expression for this is already known, a simple corollary of Theorem 3 yields the bounds

$$\mathbb{E}[\operatorname{tr} \psi_A^2] \leq (1 + 4d)\, \mathbb{E}[\operatorname{tr}(\varphi_U)_A^2] = \frac{1 + 4/d}{d_A}. \tag{26}$$

By comparison, an exact calculation yields $\mathbb{E}[\operatorname{tr} \psi_A^2] = \frac{d_A + d_B}{d(d+1)}$ which differs only in sub-leading terms. Similar bounds apply to higher moments of $\operatorname{tr}\left[\psi_A^2\right]$ or to related quantities.

Applying a random $U^{\otimes n}$ to half of $|\Phi_d\rangle^{\otimes n}$ yields Equation (24). A dual question is applying a random permutation $P_d(\pi)$. We will discuss this further in Remark 2 in Appendix A.

5 Boson sampling anticoncentration

Boson Sampling is the process of sending n photons through an array of beam-splitters that couple m optical modes and then measuring each mode. It was introduced as a computational task in [18] and is significant because it appears to not be universal for quantum computing while remaining hard to simulate classically, assuming some plausible conjectures. This gives a plausible route to quantum computational supremacy using current technology; see [19] for a recent demonstration.

To understand the output distribution of Boson Sampling, first observe that the beam-splitters define a unitary $U \in \mathcal{U}_m$, often taken to be Haar random. Suppose the $n \ll m$ photons are input into n modes corresponding to a set $T \subset [m]$. Then the probability of finding them into n output modes $S \subset [m]$ is

$$\Pr[S] = |\operatorname{Per}(U_{S,T})|^2, \tag{27}$$

where $U_{S,T}$ denotes the submatrix of U with rows corresponding to S and columns corresponding to T. (There is also a $O(n^2/m)$ probability of finding two or more photons in the same mode. In this case S becomes a multiset, we interpret $U_{S,T}$ to allow repeated rows, and the RHS of Equation (27) is divided by $s_1! \dots s_m!$ where s_i is the number of photons in mode i. We avoid considering this case by choosing $n \ll \sqrt{m}$.) Recall that the permanent of a matrix V is

$$\operatorname{Per}(V) = \sum_{\pi \in S_n} \prod_{i=1}^{n} V_{i,\pi(i)}. \tag{28}$$

Several steps in the analysis of Boson Sampling are simplified by approximating the submatrices $V := U_{S,T}$ by a Gaussian matrix X. We define X to be an $n \times n$ matrix of i.i.d. complex Gaussians such that

$$\begin{aligned} \mathbb{E}[X_{i,j}] = \mathbb{E}[V_{i,j}] = \mathbb{E}[X_{i,j}^2] = \mathbb{E}[V_{i,j}^2] = 0 \\ \mathbb{E}[|X_{i,j}|^2] = \mathbb{E}[|V_{i,j}|^2] = \frac{1}{m} \end{aligned} \tag{29}$$

By definition these moments of V and X match, but what about higher moments? In Section 2.2 we argued that higher moments are close as well in the regime where $G \approx I$. In this section we will show how this implies that low moments of the permanent are also close. Note that the notation in this section is chosen to be consistent with the boson sampling literature and (n, m) here will turn out to correspond to (n, d) in the rest of the paper.

Theorem 4. *If $n^2t^2 \le 2m$ and V, X are defined as above then*

$$1 - \frac{n^2t^2}{m} \le \frac{\mathbb{E}[|\operatorname{Per}(V)|^{2t}]}{\mathbb{E}[|\operatorname{Per}(X)|^{2t}]} \le 1 + \frac{n^2t^2}{m}. \tag{30}$$

Section 5.1 of [18] establishes a similar but incomparable result, finding that the distribution of unitary submatrices of size $m^{1/6}$ are close in variational distance to an i.i.d. Gaussian distribution. Theorem 4 by contrast works for submatrices with dimension as large as $O(m^{1/2})$ but controls only low moments and not the entire distribution. However, for some applications, such as the "anticoncentration" conjecture, this can be enough.

Nezami [20] used representation theory to give formulas for the moments of both $|\operatorname{Per}(X)|^2$ and $|\operatorname{Per}(V)|^2$. These can be used to establish bounds similar to (30) but slightly stronger[1]. The contribution of this work then is an independent and somewhat simpler proof of a nearly equivalent result.

Proof. Define

$$|S_n\rangle = \frac{1}{n!}\sum_{\pi\in S_n} |\pi\rangle \qquad \text{where} \qquad |\pi\rangle = |\pi(1)\rangle \otimes \cdots \otimes |\pi(n)\rangle \in (\mathbb{C}^n)^{\otimes n}. \tag{32}$$

Then $\operatorname{Per}(V) = \langle S_n| V^{\otimes n} |S_n\rangle$. Since the distribution of V is invariant under $V \mapsto e^{i\phi}V$, $\mathbb{E}[\operatorname{Per}(V)] = 0$ and similarly for X. Thus we will focus instead on

$$\mathbb{E}[|\operatorname{Per}(V)|^{2t}] = \mathbb{E}[\langle S_n|^{\otimes 2t} V^{\otimes nt} \otimes V^{*,\otimes nt} |S_n\rangle^{\otimes 2t}]. \tag{33}$$

If we interpret $|S_n\rangle$ as being a vector in $(\mathbb{C}^m)^{\otimes n}$ then we can replace V with U in the RHS of (33), obtaining

$$\mathbb{E}[|\operatorname{Per}(V)|^{2t}] = \mathbb{E}[\langle S_n|^{\otimes 2t} U^{\otimes nt} \otimes U^{*,\otimes nt} |S_n\rangle^{\otimes 2t}]. \tag{34}$$

Now apply (15) to evaluate the expectation over U and obtain

$$\mathbb{E}[|\operatorname{Per}(V)|^{2t}] = \langle S_n|^{\otimes 2t} \sum_{\sigma,\tau\in\mathcal{S}_{nt}} (I\otimes P_m(\sigma))\Phi_m^{\otimes nt}(I\otimes P_m(\tau))^\dagger \operatorname{Wg}^{(nt,m)}(\sigma,\tau) |S_n\rangle^{\otimes 2t}. \tag{35}$$

By contrast, for the moments of a complex Gaussian,

$$\mathbb{E}[|\operatorname{Per}(X)|^{2t}] = \langle S_n|^{\otimes 2t} \sum_{\pi\in\mathcal{S}_{nt}} (I\otimes P_m(\pi))\Phi_m^{\otimes nt}(I\otimes P_m(\pi))^\dagger |S_n\rangle^{\otimes 2t}. \tag{36}$$

For $\pi \in \mathcal{S}_{nt}$ we need to evaluate

$$\alpha_\pi := \langle \Phi_m|^{\otimes nt} (I\otimes P_m(\pi)) |S_n\rangle^{\otimes 2t} = \frac{1}{\sqrt{m^{nt}}} \langle S_n|^{\otimes t} P_m(\pi) |S_n\rangle^{\otimes t} = \frac{1}{n!^t\sqrt{m^{nt}}} \sum_{\sigma,\sigma'\in\mathcal{S}_n^t} 1_{\sigma'=\pi\sigma}. \tag{37}$$

Before evaluating α_π, we can make some observations about the moments of the permanent. Substituting into (35) and (36) we have

$$\mathbb{E}[|\operatorname{Per}(V)|^{2t}] = \langle\alpha| \operatorname{Wg}^{(nt,m)} |\alpha\rangle \quad \text{and} \quad \mathbb{E}[|\operatorname{Per}(X)|^{2t}] = \langle\alpha|\alpha\rangle. \tag{38}$$

[1] Specifically eqns (9), (13) and (19) from [20], along with the fact that $\rho_\lambda(RCRC) \geq 0$, directly imply that

$$1 - \frac{n^2t^2}{2m} \approx \prod_{i=0}^{nt-1}\left(1+\frac{i}{m}\right)^{-1} \leq \frac{\mathbb{E}[|\operatorname{Per}(V)|^{2t}]}{\mathbb{E}[|\operatorname{Per}(X)|^{2t}]} \leq \prod_{i=1}^{\min(n,t)}\prod_{j=1}^{\max(n,t)}\left(1+\frac{j-i}{m}\right)^{-1} \approx 1 - \frac{nt|n-t|}{m} \tag{31}$$

This observation is due to Sepehr Nezami.

In the $n^2t^2 \ll m$ regime, our control of the spectrum of Wg lets us relate these quantities. Indeed

$$1 - \frac{n^2t^2}{m} \leq \lambda_{\max}(G^{(nt,m)})^{-1} \leq \frac{\mathbb{E}[|\operatorname{Per}(V)|^{2t}]}{\mathbb{E}[|\operatorname{Per}(G)|^{2t}]} \leq \lambda_{\min}(G^{(nt,m)})^{-1} \leq 1 + \frac{n^2t^2}{m}, \quad (39)$$

where the outer inequalities are valid when $n^2t^2 \leq 2m$. □

We can also use these formulas to calculate some moments of Gaussian matrices. It is not trivial since α_π will depend on π for $t > 1$. However, the case $t = 2$ is relatively quick. Then (37) will depend on the parameter $\ell = w(\pi) := |\pi(\{1, \ldots, n\}) \cap \{n+1, \ldots, 2n\}|$. Let's fix $\pi \in \mathcal{S}_{2n}$, choose $\sigma \in \mathcal{S}_n^2$ at random and calculate the probability that $\pi\sigma \in \mathcal{S}_n \times \mathcal{S}_n$. This is $\binom{n}{\ell}^{-1}$. Thus we find

$$\alpha_\pi = \frac{1}{\binom{n}{w(\pi)} m^{nt/2}}. \quad (40)$$

We also want to calculate $|w^{-1}(\ell)|$. This is given by a hypergeometric distribution.

$$|w^{-1}(\ell)| = \binom{n}{\ell}^4 \ell!^2 (n-\ell)!^2 = n!^2 \binom{n}{\ell}^2. \quad (41)$$

To apply this to the Gaussian case, we substitute into(36).

$$\mathbb{E}[|\operatorname{Per}(X)|^4] = \sum_\ell n!^2 \binom{n}{\ell}^2 \binom{n}{\ell}^{-2} m^{-2n} = (n+1)\frac{n!^2}{m^{2n}}. \quad (42)$$

This yields an alternate proof of Lemma 56 of [18].

6 Partial transposes of permutation operators

This section will introduce some mathematical tools that will be relevant to applications involving multipartite quantum systems and specifically the proofs in Sections 7 and 8.

A frequently used tool in understanding locality is the PPT (Positive Partial Transpose) restriction [21, 22]. The PPT criteria for seperability of states and measurements is useful in part because it has an efficient semidefinite program and these same attributes also make it more amenable to proofs. In this section we study the spectrum of permutation operators with the partial transpose applied to some of the subsystems. The goal is to establish lemmas that will be later used in the applications.

If $\{M, I - M\}$ is a two-outcome measurement that can be implemented by LOCC (local operations and classical communication [23]), then a useful relaxation is to require that M and $I - M$ remain positive semi-definite whenever any collection of subsystems is partially transposed. We call the measurements satisfying this condition "PPT", meaning that measurement operators are *P*ositive under *P*artial

Transposition[2]. Let M act on n systems, and let $S \subseteq [n]$. Then we let M^{Γ_S} denote M with the indices in S transposed. In this notation, an equivalent characterization of the PPT condition is that

$$0 \preceq M^{\Gamma_S} \preceq I \qquad \forall S \subseteq [n] \tag{43}$$

In this section, we discuss partial transposes of the operators $P_d(\pi)$.

The relevance of the PPT constraint is that taking the partial transpose of part of a permutation matrix can result in the largest eigenvalue increasing dramatically. Thus, if Γ denotes the partial transpose, then requiring that $0 \leq M^{\Gamma} \leq I$ can be a potent constraint in addition to the usual $0 \leq M \leq I$.

We will consider taking the partial transpose of an arbitrary set $S \subset [n]$ and will denote this operation ${}^{\Gamma_S}$. We also define $\bar{S} := [n] - S$, so that $(S, \bar{S})$ partition $[n]$.

Lemma 5. *For any $\pi \in \mathcal{S}_n$, let $k = |S \cap \pi(\bar{S})|$. Then $P_d(\pi)^{\Gamma_S}$ has d^{n-2k} non-zero singular values, each equal to d^k.*

This is a generalization of the well-known fact that $\mathcal{F}_{1,2}^{\Gamma_2} = d\Phi$, where Φ is a projection on the maximally entangled state. In fact, we can say somewhat more about the structure of $P_d(\pi)^{\Gamma_S}$ (see [24, 25]), but Lemma 5 is all we need for our argument.

Proof of Lemma 5. Let $X = (P_d(\pi)^{\Gamma_S})^{\dagger} P_d(\pi)^{\Gamma_S} = P_d(\pi)^{\Gamma_{\bar{S}}} P_d(\pi)^{\Gamma_S}$. Then the square of the singular values of $P_d(\pi)^{\Gamma_S}$ are the eigenvalues of X. To represent tensor products of n systems, we will use a superscript ${}^{(i)}$ to indicate that a system should be placed in the i^{th} position, so that we can list the systems in an order that is more convenient. We now calculate

$$\begin{aligned}
X &= \sum_{\substack{x_1,\ldots,x_n \in [d] \\ y_1,\ldots,y_n \in [d]}} \bigotimes_{i \in S} |x_i\rangle \langle x_{\pi(i)}|y_{\pi(i)}\rangle \langle y_i|^{(i)} \otimes \bigotimes_{i \in \bar{S}} |x_{\pi(i)}\rangle \langle x_i|y_i\rangle \langle y_{\pi(i)}|^{(i)} \\
&= \sum_{\substack{x_1,\ldots,x_n \in [d] \\ y_1,\ldots,y_n \in [d]}} \left(\prod_{i \in \pi(S) \cup \bar{S}} \delta_{x_i,y_i} \right) \bigotimes_{i \in S} |x_i\rangle \langle y_i|^{(i)} \otimes \bigotimes_{i \in \bar{S}} |x_{\pi(i)}\rangle \langle y_{\pi(i)}|^{(i)} \\
&= \sum_{\substack{x_1,\ldots,x_n \in [d] \\ y_1,\ldots,y_n \in [d]}} \left(\prod_{i \in \pi(S) \cup \bar{S}} \delta_{x_i,y_i} \right) \bigotimes_{i \in S} |x_i\rangle \langle y_i|^{(i)} \otimes \bigotimes_{i \in \pi(\bar{S})} |x_i\rangle \langle y_i|^{(\pi^{-1}(i))}
\end{aligned}$$

We see that a δ_{x_i,y_i} appears for all i in $\pi(S) \cup \bar{S}$, or equivalently, all i not contained in $\pi(\bar{S}) \cap S$. Additionally we see that each $|x_i\rangle \langle y_i|$ appears zero times for $i \in \bar{S} \cap \pi(S)$, twice for $i \in \pi(\bar{S}) \cap S$ and once otherwise; i.e. for $i \in (S \cap \pi(S)) \cup (\bar{S} \cap \pi(\bar{S}))$. (To justify these arguments, recall that $(S, \bar{S})$ and $(\pi(S), \pi(\bar{S}))$ both partition $[n]$.)

We now consider the partition of $[n]$ into $\bar{S} \cap \pi(S)$, $(S \cap \pi(S)) \cup (\bar{S} \cap \pi(\bar{S}))$ and $S \cap \pi(\bar{S})$ and determine the contributions from each. Note that since π is a permutation, we have $|\bar{S} \cap \pi(S)| = |S \cap \pi(\bar{S})| (= k)$,

[2] In some cases one might want to constrain only the yes or no operators to being PPT. In this paper we will always take PPT to mean that all measurement outcomes are PPT.

- For $i \in \bar{S} \cap \pi(S)$, we have an appearance of δ_{x_i,y_i}, but not of $|x_i\rangle\langle y_i|$. Thus this term contributes the scalar multiple d.
- For $i \in (S \cap \pi(S)) \cup (\bar{S} \cap \pi(\bar{S}))$, we have a δ_{x_i,y_i} constraint as well as a $|x_i\rangle\langle y_i|$ term. Thus, we have one appearance of the $d \times d$ identity operator $I_d = \sum_{x_i \in [d]} |x_i\rangle\langle x_i|$ at position i.
- Finally, for $i \in \pi(\bar{S}) \cap S$, there is no $x_i = y_i$ constraint and the total contribution is

$$\sum_{x_i,y_i \in [d]} |x_i\rangle\langle y_i|^{(i)} \otimes |x_i\rangle\langle y_i|^{\pi^{-1}(i)} = d\Phi^{(i,\pi^{-1}(i))}.$$

Together, we conclude that

$$X = d^{2k} \bigotimes_{i \in S \cap \pi(\bar{S})} \Phi^{(i,\pi^{-1}(i))} \otimes \bigotimes_{i \in (S \cap \pi(S)) \cup (\bar{S} \cap \pi^{-1}(\bar{S}))} I_d^{(i)},$$

which has the claimed eigenvalues. □

Since our bounds are often in terms of $|\pi|$, it is convenient to express Lemma 5 using this quantity. This is possible because we often are free to choose S arbitrarily. In some cases, we will need to choose a single S that works for multiple permutations. This too is straightforward but yields a weaker bound.

Lemma 6. *For any $\pi_1, \ldots, \pi_k \in \mathcal{S}_n$ there exists $S \subseteq [n]$ such that*

$$\sum_{i=1}^{k} |\pi_i(S) \cap \bar{S}| \geq \frac{1}{4} \sum_{i=1}^{k} |\pi_i|. \tag{44}$$

In the special case where we have a single $\pi \in \mathcal{S}_n$ we can find S such that

$$|\pi(S) \cap \bar{S}| \geq \frac{|\pi|}{2}. \tag{45}$$

Proof. Suppose S is chosen uniformly at random from the subsets of $[n]$. For each $i \in [k]$, let m_i denote the number of derangements of of π_i, i.e. the number of x such that $\pi_i(x) \neq x$. For such x, the probability that $x \in \pi_i(S) \cap \bar{S}$ is $1/4$. By linearity of expectation, the expectation of $|\pi_i(S) \cap \bar{S}|$ is $m_i/4$. Now suppose that π_i has c_1 1-cycles, c_2 2-cycles, and so on. Then since a single cycle of length $j \geq 2$ has j derangements,

$$m_i = n - c_1 = \sum_{j \geq 2} j c_j \quad \text{and} \quad |\pi_i| = \sum_{j \geq 2} (j-1) c_j. \tag{46}$$

Together this implies that $m_i \geq |\pi_i|$. Thus (44) holds in expectation, and also therefore holds for at least one choice of S.

For the $k = 1$ case we will choose S based on the cycle decomposition of π. For a cycle containing elements $x_1, x_2, \ldots, x_j$ we put $x_1, x_3, x_5, \ldots$ into S. A cycle of length j then contributes $\lfloor j/2 \rfloor$ to $|\pi(S) \cap S|$. Since $\lfloor j/2 \rfloor \geq (j-1)/2$ we obtain (45). □

Let $M = \sum_\pi m_\pi P_d(\pi)$ satisfy the PPT condition (43) and assume that $n \leq d^{1/2}$. From the bound $\|M\| \leq 1$ and eq. (20) we have $|m_\pi| \leq 1/(1-n^2/2d) \leq 1 + \frac{n^2}{d}$. Using the PPT condition and the stronger condition $n \leq d^{1/4}$ we can show a much stronger bound when π is far from e.

Lemma 7. *If $M = \sum_\pi m_\pi P_d(\pi)$ satisfies (43) and $\frac{n^2}{\sqrt{d}} \leq 1$ then*

$$|m_\pi| \leq \left(1 + \frac{n^2}{\sqrt{d}}\right) d^{-|\pi|/2} \tag{47}$$

Proof. Let

$$\pi := \arg\max_\pi |m_\pi| d^{|\pi|/2}. \tag{48}$$

Use Lemma 5 to choose S so that $|\pi(S) \cap \bar{S} \geq |\pi|/2$ and thus

$$\|P_d(\pi)^{\Gamma_S}\|_1 = d^{n-|\pi(S)\cap\bar{S}|} \leq d^{n-|\pi|/2}. \tag{49}$$

Then

$$\begin{align}
1 &\geq \left| \operatorname{tr} \frac{P_d(\pi)^{\Gamma_S}}{d^{n-|\pi|/2}} M^{\Gamma_S} \right| && \text{from Hölder, (43) and (49)} \tag{50}\\
&= d^{|\pi|/2} \left| \langle M^{\Gamma_S}, P_d(\pi)^{\Gamma_S} \rangle \right| \tag{51}\\
&= d^{|\pi|/2} \left| \langle M, P_d(\pi) \rangle \right| \tag{52}\\
&\geq d^{|\pi|/2} |m_\pi| - d^{|\pi|/2} \sum_{\pi' \neq \pi} |m_{\pi'}| G_{\pi,\pi'} \tag{53}\\
&\geq d^{|\pi|/2} |m_\pi| (1 - \sum_{\pi' \neq \pi} d^{\frac{|\pi|-|\pi'|}{2}} d^{-|\pi^{-1}\pi'|}) && \text{by (48)} \tag{54}\\
&\geq d^{|\pi|/2} |m_\pi| (1 - \sum_{\pi' \neq \pi} d^{-|\pi^{-1}\pi'|/2}) && \text{by the triangle inequality, (5)} \tag{55}\\
&\geq d^{|\pi|/2} |m_\pi| (2 - e^{n^2/2\sqrt{d}}) && \text{by Equation (13)} \tag{56}\\
&\geq d^{|\pi|/2} |m_\pi| / (1 + n^2/\sqrt{d}) && \text{using } n^2 \leq \sqrt{d}. \tag{57}
\end{align}$$

□

7 Multipartite data hiding

Let ρ_0, ρ_1 be density matrices on n d-dimensional systems that commute with all $U^{\otimes n}$. If $\|\rho_0 - \rho_1\|_1$ is large, then of course, given ρ_b for $b \in \{0,1\}$, there is some (global) measurement that can estimate b with some non-negligible bias. Here we will argue that, on the other hand, LOCC measurements, or even PPT measurements, cannot learn anything about b. This data-hiding scheme is due to Eggeling and Werner[26] who shows that it was secure when n is fixed and $d \to \infty$. Our contribution is to extend their analysis to the case when n is up to $O(\sqrt{d})$ by using our approximate orthogonality relationship.

Theorem 8. *Let ρ_0, ρ_1 be any density matrices on $(\mathbb{C}^d)^{\otimes n}$ that commute with all $U^{\otimes n}$, and let $\{M, I - M\}$ be a PPT measurement. If $n \leq d^{1/4}$ then*

$$|\operatorname{tr} M(\rho_0 - \rho_1)| \leq \frac{6n^2}{\sqrt{d}}. \tag{58}$$

Proof. Let M_0, M_1 be the PPT measurement operators corresponding to guessing $0, 1$ respectively. Let $M = M_0 - M_1$. Then

$$-I \leq M^{\Gamma_S} \leq I \tag{59}$$

for any $S \subseteq [n]$, where Γ_S corresponds to taking the partial transpose of indices S. Let $\Delta = \rho_0 - \rho_1$, so that the bias achieved by the measurement is $\operatorname{tr} M\Delta$. Observe that $\operatorname{tr} \Delta = 0$ and that $[\Delta, U^{\otimes n}] = 0$ for all U.

We can assume WLOG that $[M, U^{\otimes n}] = 0$ as well. This is because

$$\operatorname{tr} M\Delta = \mathop{\mathbb{E}}_{U} \operatorname{tr} MU^{\otimes n}\Delta(U^\dagger)^{\otimes n} = \mathop{\mathbb{E}}_{U} \operatorname{tr}\big((U^\dagger)^{\otimes n} M U^{\otimes n}\big)\Delta.$$

Thus, we can write $M = \sum_{\pi \in \mathcal{S}_n} m_\pi P_d(\pi)$.

The bias is now bounded by

$$\begin{aligned}
\operatorname{tr} M\Delta &= \sum_{\pi \neq e} m_\pi \operatorname{tr}[P_d(\pi)\Delta] && \text{because } \operatorname{tr}\Delta = 0 && (60)\\
&\leq \sum_{\pi \neq e} |m_\pi| \|P_d(\pi)\|_\infty \|\Delta\|_1 && \text{triangle inequality and Hölder} && (61)\\
&= \sum_{\pi \neq e} |m_\pi| && && (62)\\
&\leq 2\sum_{\pi \neq e} d^{-|\pi|/2} && \text{using Lemma 7} && (63)\\
&\leq 3(e^{n^2/\sqrt{d}} - 1) && && (64)\\
&\leq 6n^2/\sqrt{d} && \text{using } n^2 \leq \sqrt{d}. && (65)
\end{aligned}$$

□

Theorem 8 applies to any two states ρ_0, ρ_1 satisfying the symmetry condition, although it is only interesting when $\|\rho_0 - \rho_1\|_1$ is large. Coming up with one such pair is straightforward, but how many can be constructed simultaneously? Here we can use Schur-Weyl duality (c.f. Appendix A) to show that any state commuting with all $U^{\otimes n}$ must be of the form

$$\sum_{\lambda \in \operatorname{Par}(n,d)} p_\lambda \rho_\lambda \otimes \tau_{\mathcal{Q}^d_\lambda}, \tag{66}$$

where $\tau_{\mathcal{Q}^d_\lambda} := I_{\mathcal{Q}^d_\lambda} / \dim \mathcal{Q}^d_\lambda$. This permits $N = \sum_{\lambda \in \operatorname{Par}(n,d)} \dim P_\lambda$ perfectly orthogonal states. Since $d \geq n$, $\operatorname{Par}(n,d)$ includes all partitions of n, and thus

$\sum_{\lambda\in\mathrm{Par}(n,d)} \dim P_\lambda^2 = n!$. As a result, $N \geq \sqrt{n!}$. On the other hand, $\mathrm{Par}(n,n) \leq e^{2c\sqrt{n}}$ for $c \approx 1.28$, so $N \leq \sqrt{n!}e^{c\sqrt{n}}$. This analysis also implies that $\frac{1}{2}\log(n!)$ qubits can be hidden in such states. If we are content with pairwise approximate distinguishability then exponentially more states can be hidden [27].

Another application concerns the distinguishability of n copies of the same random state from n copies of independently random states. As density matrices, these correspond to $\mathbb{E}[\psi^{\otimes n}]$ and $(I/d)^{\otimes n}$ respectively. If collective measurements are allowed then projecting onto the symmetric subspace will almost perfectly distinguish these states. But the situation is different with LOCC measurement.

Corollary 9 (Local purity tests)**.** *If a PPT measurement is used to distinguish $\mathbb{E}[\psi^{\otimes n}]$ from $(I/d)^{\otimes n}$ it will achieve bias $\leq O(n^2/\sqrt{d})$.*

Recently and independently of this work, sharper upper and lower bounds were found by Chen, Cotler, Huang and Li [28] who showed that $n = \Theta(\sqrt{d})$ copies are necessary and sufficient for local purity testing.

8 Limitations of local product tests

Suppose we are given n copies of a k-partite pure state $|\psi\rangle \in (\mathbb{C}^d)^{\otimes k}$. We would like to know if $|\psi\rangle$ is close to being a product state $|\psi_1\rangle \otimes \cdots \otimes |\psi_k\rangle$ or far from any such state. A natural test for this is to project all n copies of each of the k subsystems onto the n-fold symmetric subspace $\vee^n\mathbb{C}^d$. If all the projections succeed, output "product", otherwise output "not product". This test was proposed by [29] and analyzed by [30]. The test can be easily shown to be optimal among a reasonable class of such product tests (see Section 5 of [30]), but the projections require entangling operations across the n copies.

How effective can be make product tests without such entangling operations? If an LOCC test existed, then it would imply that $\mathsf{QMA} = \mathsf{QMA}(2)$ [31], and, depending on the accuracy of the test, this might falsify the Unique Games Conjecture [32] or the Exponential Time Hypothesis [30]. In [30] it was proved that such a test cannot exist for $n = 2$. Here we show it cannot exist even for larger values of n, and even in the easiest case where $k = 2$.

To be more precise we say that a product test consists of a two-outcome measurement $\{M, I - M\}$, corresponding to outcomes "product" and "not product." The completeness c is $\min \mathrm{tr}[M\psi^{\otimes n}]$ over all product states ψ while the soundness s is $\max \mathrm{tr}[M\psi^{\otimes n}]$ over all states ψ with overlap $\leq 1/2$ with any product state. (The constant $1/2$ is arbitrary, however note that no state is orthogonal to all product states.) Define the *bias* to be $b = c - s$. The standard product test from [29] was proved in [30] to have bias $\geq \Omega(1)$ with $n = 2$ and k arbitrary. However, we will see that this cannot be achieved by a PPT test unless n grows with d.

Theorem 10. *If $\{M, I - M\}$ is a PPT product test for $k = 2$ acting on n copies of a state, then its bias b is $\leq O(n^2/d^{1/4})$.*

Our relation between n and d is tight up to polynomial factors, since when $n \gg d^2$ then state tomography can be carried out even with no communication between subsystems.

Proof of Theorem 10. Assume that $n \leq d^{1/8}$ since otherwise the theorem holds trivially. Let

$$\Delta = \mathop{\mathbb{E}}_{|\psi_A\rangle,|\psi_B\rangle\in\mathbb{C}^d}[\psi_A^{\otimes n} \otimes \psi_B^{\otimes n}] - \mathop{\mathbb{E}}_{|\psi\rangle\in\mathbb{C}^{d^2}}[\psi^{\otimes n}]. \tag{67}$$

Our goal is to show that $\operatorname{tr} M\Delta$ is small for any PPT measurement $\{M, I - M\}$.

First, we observe that Δ commutes with $U^{\otimes n} \otimes V^{\otimes n}$, and so without loss of generality we can assume that M does as well. Thus, the arguments leading to (2) imply that

$$M = \sum_{\pi_A,\pi_B\in\mathcal{S}_n} m_{\pi_A,\pi_B} P_d(\pi_A) \otimes P_d(\pi_B). \tag{68}$$

For convenience, we will refer to the pair (π_A, π_B) as a single permutation $\pi \in \mathcal{S}_{2n}$. Formally, we can embed $\mathcal{S}_n \times \mathcal{S}_n$ into $\mathcal{S}_{2n}$ as the set of permutations that does not mix $\{1, \dots, n\}$ and $\{n + 1, \dots, 2n\}$.

We will need to develop a variant of Lemma 7 to show that

$$m_\pi \leq 2d^{-|\pi|/4}. \tag{69}$$

This will imply our desired result as follows:

$$\begin{aligned}
\operatorname{tr} M\Delta &= \sum_{\pi_A,\pi_B\in\mathcal{S}_n} m_{\pi_A,\pi_B} \operatorname{tr}(P_d(\pi_A) \otimes P_d(\pi_B))\Delta && (70)\\
&= \sum_{(\pi_A,\pi_B)\neq(e,e)} m_{\pi_A,\pi_B} \operatorname{tr}(P_d(\pi_A) \otimes P_d(\pi_B))\Delta \qquad \text{since } \operatorname{tr}\Delta = 0 && (71)\\
&\leq \sum_{(\pi_A,\pi_B)\neq(e,e)} |m_{\pi_A,\pi_B}| && (72)\\
&\leq 2 \sum_{(\pi_A,\pi_B)\neq(e,e)} d^{-\frac{|\pi_A|+|\pi_B|}{4}} && (73)\\
&\leq 2\left(\left(\sum_{\pi\in\mathcal{S}_n} d^{-\frac{|\pi|}{4}}\right)^2 - 1\right) && (74)\\
&\leq 2((e^{n^2/2d^{1/4}})^2 - 1) && (75)\\
&\leq \frac{4n^2}{d^{1/4}} && (76)
\end{aligned}$$

Now we return to the proof of (69), which essentially repeats the proof of Lemma 7 but uses the multiple-permutation version of Lemma 6. The new feature of this setting is that the locality constraint here is between $A_1B_1 : A_2B_2 : \cdots : A_nB_n$ while the permutations π_A and π_B act on $A_1 \dots A_n$ and $B_1 \dots B_n$ respectively. Thus our PPT condition is that $\|M^{\Gamma_S}\| \leq 1$ where Γ_S is a shorthand for the transpose of systems $\bigcup_{i\in S}\{A_i, B_i\}$.

Following the proof of Lemma 7, let

$$\pi := \arg \max_{\pi=\pi_A\times\pi_B} |m_\pi| d^{|\pi|/4} \tag{77}$$

and use Lemma 6 to find $S \subseteq [n]$ such that

$$|S \cap \pi_A(\bar{S})| + |S \cap \pi_B(\bar{S})| \geq \frac{|\pi|}{4} = \frac{|\pi_A| + |\pi_B|}{4} \tag{78}$$

The rest of the proof is almost identical.

$$\begin{aligned}
\|P_d(\pi)^{\Gamma_S}\|_1 &= d^{n-|S\cap\pi_A(\bar{S})|} \cdot d^{n-|S\cap\pi_B(\bar{S})|} \leq d^{2n-|\pi|/4}. && (79)\\
1 &\geq \left|\operatorname{tr} \frac{P_d(\pi)^{\Gamma_S}}{d^{2n-|\pi|/4}} M^{\Gamma_S}\right| && (80)\\
&= d^{|\pi|/4}\,|\langle M, P_d(\pi)\rangle| && (81)\\
&\geq d^{|\pi|/4}|m_\pi| - d^{|\pi|/4} \sum_{\substack{\pi'\in\mathcal{S}_n\times\mathcal{S}_n\\ \pi'\neq\pi}} |m_{\pi'}|G_{\pi,\pi'} && (82)\\
&\geq d^{|\pi|/4}|m_\pi|(1 - \sum_{\substack{\pi'\in\mathcal{S}_n\times\mathcal{S}_n\\ \pi'\neq\pi}} d^{\frac{|\pi|-|\pi'|}{4}} d^{-|\pi^{-1}\pi'|}) && (83)\\
&\geq d^{|\pi|/4}|m_\pi|(1 - \sum_{\substack{\pi'\in\mathcal{S}_n\times\mathcal{S}_n\\ \pi'\neq\pi}} d^{-\frac{3}{4}|\pi^{-1}\pi'|}) && \text{by the triangle inequality, (5)}\\
& && (84)\\
&\geq d^{|\pi|/4}|m_\pi|\left(2 - \left(e^{\frac{n^2}{2d^{3/4}}}\right)^2\right) && \text{by Equation (12)}\\
& && (85)\\
&\geq \frac{1}{2} d^{|\pi|/4}|m_\pi| && (86)
\end{aligned}$$

□

A Full spectrum of the Gram matrix

In this appendix we give a self-contained proof of Lemma 2. The idea is to decompose the permutation action P_d into irreps of $\mathcal{S}_n$. We begin with some terminology from representation theory.

Let $\mathrm{Par}(n,d)$ denote the set of partitions of n into d parts; that is $\lambda \in \mathrm{Par}(n,d)$ if $\lambda = (\lambda_1,\ldots,\lambda_d) \in \mathbb{Z}_+^d$ with $\lambda_1 \geq \cdots \geq \lambda_d \geq 0$ and $\sum_{i=1}^d \lambda_i$. We also identify λ with the set of $(i,j) \in \mathbb{N}^2$ with $j \leq \lambda_i$. Schur-Weyl duality states that

$$(\mathbb{C}^d)^{\otimes n} \cong \bigoplus_{\lambda\in\mathrm{Par}(n,d)} \mathcal{Q}_\lambda^d \otimes \mathcal{P}_\lambda, \tag{87}$$

where $\mathcal{Q}_\lambda^d$ labels an irrep of $\mathcal{U}_d$ and $\mathcal{P}_\lambda$ labels an irrep of $\mathcal{S}_n$. Let $\mathbf{q}_\lambda^d(U)$ and $\mathbf{p}_\lambda(\pi)$ denote the corresponding group actions of $\mathcal{U}_d$ and $\mathcal{S}_n$. Assume for convenience that

$\mathbf{p}_\lambda(\pi)$ is always a real orthogonal matrix. We let U_{Sch} denote the unitary isomorphism mapping the LHS of (87) to the RHS; however, we generally abuse notation and omit writing U_{Sch}.

We will need to make use of the following formulas for the dimensions of these irreps. Define $\tilde{\lambda} := \lambda + (d-1, d-2, \ldots, 1, 0)$. Then [1, 33]

$$\dim \mathcal{Q}_\lambda^d = \frac{\prod_{1 \le i < j \le d} (\tilde{\lambda}_i - \tilde{\lambda}_j)}{\prod_{m=1}^{d-1} m!} \tag{88}$$

$$\dim \mathcal{P}_\lambda = \frac{n!}{\tilde{\lambda}_1! \tilde{\lambda}_2! \cdots \tilde{\lambda}_d!} \prod_{1 \le i < j \le d} (\tilde{\lambda}_i - \tilde{\lambda}_j) \tag{89}$$

We will need the ratio of these dimensions. One can directly calculate (and see also [34])

$$\frac{n! \dim \mathcal{Q}_\lambda^d}{\dim \mathcal{P}_\lambda} = \prod_{i=1}^{d} \frac{\lambda_i + d - i!}{d - i!} = \prod_{i=1}^{d} \prod_{j=1}^{\lambda_i} d - i + j \tag{90}$$

This last double product can be abbreviated as the product over $(i,j) \in \lambda$, where λ is overloaded to mean both the partition $\lambda_1, \ldots, \lambda_d$ and the set $\{(i,j) : 1 \le j \le \lambda_i\}$.

Proof of Lemma 2. Let $\{|\pi\rangle : \pi \in \mathcal{S}_n\}$ denote a set of orthonormal vectors indexed by the permutations and define $|v_\pi\rangle = (I \otimes P_d(\pi)) |\Phi_d\rangle^{\otimes n}$. We also define the maximally entangled states $|\Phi_{\mathcal{P}_\lambda}\rangle \in \mathcal{P}_\lambda \otimes \mathcal{P}_\lambda$ and $\left|\Phi_{\mathcal{Q}_\lambda^d}\right\rangle \in \mathcal{Q}_\lambda^d \otimes (\mathcal{Q}_\lambda^d)^*$ to be unit vectors that are invariant respectively under $\mathbf{p}_\lambda(\pi) \otimes \mathbf{p}_\lambda(\pi)$ for all $\pi \in \mathcal{S}_n$ and $\mathbf{q}_\lambda^d(U) \otimes \mathbf{q}_\lambda^d(U)^*$ for all $U \in \mathcal{U}_d$. (We can omit the $*$ for $\mathcal{P}_\lambda$ because we have taken $\mathbf{p}_\lambda(\pi)$ to be real orthogonal matrices.) By Schur's Lemma, these conditions specify $|\Phi_{\mathcal{P}_\lambda}\rangle$ and $\left|\Phi_{\mathcal{Q}_\lambda^d}\right\rangle$ uniquely, up to a phase. To set this phase, let $|\Phi_d\rangle^{\otimes n} := \sum_{\lambda \in \text{Par}(n,d)} \sqrt{\frac{\dim \mathcal{Q}_\lambda^d \dim \mathcal{P}_\lambda}{d^n}} |\lambda, \lambda\rangle \left|\Phi_{\mathcal{Q}_\lambda^d}\right\rangle |\Phi_{\mathcal{P}_\lambda}\rangle$. Thus

$$|v_\pi\rangle = \sum_{\lambda \in \text{Par}(n,d)} \sqrt{\frac{\dim \mathcal{Q}_\lambda^d \dim \mathcal{P}_\lambda}{d^n}} |\lambda, \lambda\rangle \left|\Phi_{\mathcal{Q}_\lambda^d}\right\rangle (I \otimes \mathbf{p}_\lambda(\pi)) |\Phi_{\mathcal{P}_\lambda}\rangle. \tag{91}$$

Observe that $\langle v_{\pi_1} | v_{\pi_2} \rangle = \langle P_d(\pi_1), P_d(\pi_2) \rangle$. Define the matrix $K^{(n,d)} := \sum_{\pi \in \mathcal{S}_n} |\pi\rangle \langle v_\pi|$, and observe that $G^{(n,d)} = K^{(n,d)} (K^{(n,d)})^\dagger$. Thus $G^{(n,d)}$ is isospectral to

$$\begin{aligned}(K^{(n,d)})^\dagger K^{(n,d)} &= \sum_{\pi \in \mathcal{S}_n} |v_\pi\rangle \langle v_\pi| \qquad (92)\\ &= n! \sum_{\lambda \in \text{Par}(n,d)} \frac{\dim \mathcal{Q}_\lambda^d \dim \mathcal{P}_\lambda}{d^n} |\lambda, \lambda\rangle \langle \lambda, \lambda| \otimes \left|\Phi_{\mathcal{Q}_\lambda^d}\right\rangle \left\langle \Phi_{\mathcal{Q}_\lambda^d}\right| \otimes \frac{I_{\mathcal{P}_\lambda}}{\dim \mathcal{P}_\lambda} \otimes \frac{I_{\mathcal{P}_\lambda}}{\dim \mathcal{P}_\lambda}\end{aligned}$$

□

Remark 2. The matrix $\sum_\pi |v_\pi\rangle\langle v_\pi|$ in Equation (92) has another interpretation. If we apply a random $P_d(\pi)$ to half of $|\Phi_d\rangle^{\otimes n}$ then we obtain the state

$$\frac{1}{n!}\sum_\pi |v_\pi\rangle\langle v_\pi| = \frac{(K^{(n,d)})^\dagger K^{(n,d)}}{n!}, \tag{93}$$

which is isospectral to $G/n!$.

B Partitions are not approximately orthogonal

Most conclusions in this paper do not depend strongly on the properties of $\mathcal{S}_n$ or $\mathcal{U}_d$. As noted in Remark 1, to show that $\|G-I\| \le n^2/2d$, we need only that $G_{x,y} = d^{-\text{dist}(x,y)}$ where $\text{dist}(\cdot,\cdot)$ is the graph distance on a graph of degree $\le n^2/2$. Could we replace $\mathcal{S}_n$ with other sets?

Of course for general N-dimensional vectors, one can have $\exp\bigl(O(N\epsilon^2)\bigr)$ vectors with pairwise inner product at most ϵ, but they must be collectively far from an orthonormal basis. So pairwise distance certainly does not guarantee any kind of approximate orthogonality in the collective sense we have discussed.

There is one natural analogue of $\mathcal{S}_n$ where approximate orthogonality also turns out to fail. This example is due to Kevin Zatloukal. Let $\mathcal{P}_n$ be the set of partitions of the set $[n]$. For example, $\mathcal{P}_3$ consists of five partitions: $\{\{1\},\{2\},\{3\}\}$, $\{\{1,3\},\{2\}\}$, $\{\{1\},\{2,3\}\}$, $\{\{1,2\},\{3\}\}$, and $\{\{1,2,3\}\}$. Given a partition Π, define $[d]^\Pi$ to be the set of strings $x_1,\ldots,x_n \in [d]^n$ where $x_i = x_j$ whenever i,j are in the same block of Π. The corresponding quantum state is

$$|E_\Pi\rangle := d^{-\frac{\text{number of blocks of }\Pi}{2}} \sum_{x\in[d]^\Pi} |x\rangle\,. \tag{94}$$

These states were used in 0811.2597.

Let $G[\mathcal{P}_n]$ denote the Gram matrix of $\{|E_\Pi\rangle\}$ states, while we use $G[\mathcal{S}_n]$ to denote the Gram matrix studied in the rest of the paper. Concretely $G[\mathcal{P}_n]_{\Pi_1,\Pi_2} = |\langle\Pi_1|\Pi_2\rangle|^2$. In both cases we have 1 on the diagonal and positive powers of $1/d$ for each off-diagonal entry. In both cases, the dimension is exponential in n. (The number of partitions is given by the Bell numbers, which are $\le n^n$.) However the interpretation in terms of distances in a low-degree graph does not exist. Indeed, if $\Pi_0 = \{\{1,2,\ldots,n\}\}$ and $\Pi_S = \{S,[n]-S\}$ for some nonempty $S \subset [n]$, then $|\langle\Pi_0|\Pi_S\rangle|^2 = 1/d$ and there are 2^n-2 choices of S. As a result the norm of $G[\mathcal{P}_n]$ is large unless $d \gg 2^n$.

Acknowledgments

Thanks to Ashley Montanaro for many helpful discussions especially about the product test and boson sampling; to Kevin Zatloukal for his observations in Lemma 1 and appendix B; to Fernando Brandão for discussions about applications to k-designs; to Sepehr Nezami for discussions surrounding (31); to Benoît Collins and Jon Novak for helping me understand the math literature on this topic. Section 5 benefited from helpful discussions with Scott Aaronson, Raul Garcia-Patron and Dominik Hangleiter.

Funding is from NSF grants CCF-1452616, CCF-1729369, PHY-1818914 and the NSF QLCI program through grant number OMA-2016245 as well as NTT (Grant AGMT DTD 9/24/20).

Conflict of Interest

On behalf of all authors, the corresponding author states that there is no conflict of interest.

Data availability statement

There is no data associated with this paper, aside from the paper itself.

References

[1] Goodman, R., Wallach, N.R.: Symmetry, Representations and Invariants. Springer, New York (2009)

[2] Harrow, A.W.: The church of the symmetric subspace (2013) arXiv:1308.6595

[3] Brandão, F.G.S.L., Harrow, A.W., Horodecki, M.: Local random quantum circuits are approximate polynomial-designs. Commun. Math. Phys. **346**(2), 397–434 (2016) https://doi.org/10.1007/s00220-016-2706-8 1208.0692

[4] Childs, A.M., Harrow, A.W., Wocjan, P.: Weak Fourier-Schur sampling, the hidden subgroup problem, and the quantum collision problem. In: Proc. of STACS. LNCS, vol. 4393, pp. 598–609 (2007)

[5] Novak, J.: Complete homogeneous symmetric polynomials in Jucys-Murphy elements and the Weingarten function 0811.3595 (2008)

[6] Zinn-Justin, P.: Jucys–Murphy elements and Weingarten matrices. Letters in Mathematical Physics **91**, 119–127 (2010) 0907.2719

[7] Brandão, F.G.S.L., Ćwikliński, P., Horodecki, M., Horodecki, P., Korbicz, J., Mozrzymas, M.: Convergence to equilibrium under a random Hamiltonian 1108.2985 (2011)

[8] Collins, B., Śniady, P.: Integration with respect to the Haar measure on unitary, orthogonal and symplectic group. Commun. Math. Phys. **264**, 773–795 (2006) math-ph/0402073

[9] Collins, B.: Moments and cumulants of polynomial random variables on unitary groups, the Itzykson-Zuber integral and free probability. Int. Math. Res. Not. **17**, 953–982 (2003) math-ph/0205010

[10] Collîns, B., Matsumoto, S., Novak, J.: The Weingarten calculus. Notices of the AMS **69**(5), 734–745 (2022) https://doi.org/10.1090/noti2474 2109.14890

[11] Collins, B.: Moment Methods on compact groups: Weingarten calculus and its applications arXiv:2207.08418 (2022)

[12] Collins, B., Matsumoto, S.: Weingarten calculus via orthogonality relations: new applications. ALEA. Latin American Journal of Probability and Mathematical Statistics **14**, 631–656 (2017) 1701.04493

[13] Novak, J.: Truncations of random unitary matrices and Young tableaux. the electronic journal of combinatorics **14**(R21), 1 (2007) math/0608108

[14] Matsumoto, S.: Moments of a single entry of circular orthogonal ensembles and Weingarten calculus. Letters in Mathematical Physics **103**(2), 113–130 (2013) 1104.3614

[15] Page, D.N.: Average entropy of a subsystem. Physical review letters **71**(9), 1291 (1993)

[16] Hayden, P., Leung, D.W., Winter, A.: Aspects of generic entanglement. Commun. Math. Phys. **265**, 95 (2006) quant-ph/0407049

[17] Aubrun, G., Szarek, S.J.: Alice and Bob Meet Banach. Mathematical Surveys and Monographs. American Mathematical Society, United States XXX (2017)

[18] Aaronson, S., Arkhipov, A.: The computational complexity of linear optics. Theory of Computing **9**(4), 143–252 (2013) 1011.3245

[19] Zhong, H.-S., Wang, H., Deng, Y.-H., Chen, M.-C., Peng, L.-C., Luo, Y.-H., Qin, J., Wu, D., Ding, X., Hu, Y., Hu, P., Yang, X.-Y., Zhang, W.-J., Li, H., Li, Y., Jiang, X., Gan, L., Yang, G., You, L., Wang, Z., Li, L., Liu, N.-L., Lu, C.-Y., Pan, J.-W.: Quantum computational advantage using photons. Science **370**(6523), 1460–1463 (2020) https://doi.org/10.1126/science.abe8770 2012.01625

[20] Nezami, S.: Permanent of random matrices from representation theory: moments, numerics, concentration, and comments on hardness of boson-sampling arXiv:2104.06423 (2021)

[21] Peres, A.: Separability criterion for density matrices. Phys. Rev. Lett. **77**(8), 1413–1415 (1996) https://doi.org/10.1103/PhysRevLett.77.1413

[22] Horodecki, M., Horodecki, P., Horodecki, R.: Separability of mixed states: necessary and sufficient conditions. Physics Letters A **223**(1–2), 1–8 (1996) https://doi.org/10.1016/S0375-9601(96)00706-2 quant-ph/9605038

[23] Chitambar, E., Leung, D., Mančinska, L., Ozols, M., Winter, A.: Everything you always wanted to know about LOCC (but were afraid to ask). Communications

in Mathematical Physics **328**(1), 303–326 (2014) 1210.4583

[24] Eggeling, T.: On multipartite symmetric states in quantum information theory. PhD thesis, Technische Universität Braunschweig (2003)

[25] Zhang, Y., Kauffman, L.H., Werner, R.F.: Permutation and its partial transpose. Int. J. Quant. Inf. **5**, 469–507 (2007) quant-ph/0606005

[26] Eggeling, T., Werner, R.F.: Hiding classical data in multipartite quantum states. Phys. Rev. Lett. **89**, 097905 (2002) quant-ph/0203004

[27] Winter, A.: Quantum and classical message identification via quantum channels. Quantum Inf. Comput. **4**(6&7), 563–578 (2004) quant-ph/0401060

[28] Chen, S., Cotler, J., Huang, H.-Y., Li, J.: Exponential Separations Between Learning With and Without Quantum Memory. in preparation (2021)

[29] Mintert, F., Kuś, M., Buchleitner, A.: Concurrence of mixed multipartite quantum states. Phys. Rev. Lett. **95**(26), 260502 (2005) quant-ph/0411127

[30] Harrow, A.W., Montanaro, A.: Testing product states, quantum Merlin-Arthur games and tensor optimization. J. ACM **60**(1), 3–1343 (2013) https://doi.org/10.1145/2432622.2432625 1001.0017

[31] Brandão, F.G.S.L., Christandl, M., Yard, J.: Faithful squashed entanglement. Commun. Math. Phys. **306**(3), 805–830 (2011) 1010.1750

[32] Barak, B., Brandão, F.G.S.L., Harrow, A.W., Kelner, J., Steurer, D., Zhou, Y.: Hypercontractivity, sum-of-squares proofs, and their applications. In: Proceedings of the 44th Symposium on Theory of Computing. STOC '12, pp. 307–326 (2012)

[33] Stanley, R.P., Fomin, S.: Enumerative Combinatorics. Cambridge Studies in Advanced Mathematics, vol. 2. Cambridge University Press, Cambridge (1999). https://doi.org/10.1017/CBO9780511609589

[34] Stanley, R.P.: Theory and application of plane partitions. Studies in Appl. Math. **1**, 167–187259279 (1971)