Preliminaries for Haar–POVM Tomography: Groups, Permutations, and Conjugacy

Haodong Yang

August 22, 2025

Note. This write-up follows the organization of John Wright's *Quantum Learning Theory* course notes (UC Berkeley, 2024)

Sources. Standard references include Nielsen–Chuang, Watrous, and Wright's course notes [1, 2, 3].

1 Groups and Actions

Definition 1 (Group)**.** A *group* is a set G with a binary operation $(g, h) \mapsto gh$ such that (i) associativity holds, (ii) there is an identity $e \in G$ with $eg = ge = g$, and (iii) every $g \in G$ has an inverse g^{-1} with $gg^{-1} = g^{-1}g = e$.

Definition 2 (Homomorphism and isomorphism)**.** A map $\varphi : G \to H$ between groups is a *homomorphism* if $\varphi(gh) = \varphi(g)\varphi(h)$ for all $g, h \in G$. If, in addition, φ is bijective, it is an *isomorphism.*

Definition 3 (Group action (left action))**.** A *(left) action* of G on a set X is a map $G \times X \to X$, $(g, x) \mapsto g \cdot x$, such that $e \cdot x = x$ and $(gh) \cdot x = g \cdot (h \cdot x)$. We also write a homomorphism $G \to \mathrm{Sym}(X)$, $g \mapsto (x \mapsto g \cdot x)$.

Definition 4 (Orbit and stabilizer)**.** For $x \in X$, the *orbit* is $\mathcal{O}(x) = \{g \cdot x : g \in G\}$ and the *stabilizer* is $G_x = \{g \in G : g \cdot x = x\}$.

Remark 1 (Orbit–stabilizer (finite case))*.* If G is finite, then $|G| = |G_x| \cdot |\mathcal{O}(x)|$ for every $x \in X$.

2 Permutations and the symmetric group

Let $[n] = \{1, \ldots, n\}$.

Definition 5 (Permutation and S_n)**.** A *permutation* of $[n]$ is a bijection $\pi : [n] \to [n]$. The set of all permutations is the *symmetric group* S_n, with composition $(\pi\sigma)(i) = \pi(\sigma(i))$. We use two-line notation

$$\pi = \begin{pmatrix} 1 & 2 & \cdots & n \\ \pi(1) & \pi(2) & \cdots & \pi(n) \end{pmatrix} \quad \text{or cycle notation, e.g. } \pi = (1\,3\,2)(4\,5)(6).$$

Proposition 1 (Basic identities)**.** *For $\pi, \sigma \in S_n$, π^{-1} is the inverse permutation, $\pi\pi^{-1} = \pi^{-1}\pi = \mathrm{id}$, and composition is associative.*

Examples of permutation groups (subgroups of S_n)

Example 1.

$$\pi = (1\,3\,2)(4\,5)(6) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 1 & 2 & 5 & 4 & 6 \end{pmatrix}.$$

$$\pi^{-1} = (1\,2\,3)(4\,5)(6) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 1 & 5 & 4 & 6 \end{pmatrix}.$$

Example 2.

$$\mathrm{id} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 2 & 3 & 4 & 5 & 6 \end{pmatrix}.$$

$$\mathrm{id}_{S_n} = \begin{pmatrix} 1 & 2 & \cdots & n \\ 1 & 2 & \cdots & n \end{pmatrix}.$$

Example 3 (Cyclic subgroup generated by an n-cycle). If $c = (1\,2\,\ldots\,n)$, then $\langle c\rangle = \{e, c, c^2, \ldots, c^{n-1}\} \cong \mathbb{Z}_n$.

Example 4 (Dihedral group D_n). Act on the vertices of a regular n-gon labeled $1, \ldots, n$ by rotations and reflections. As a subgroup of S_n, $D_n = \langle (1\,2\,\ldots\,n),\ (2\,n)(3\,n-1)\cdots\rangle$ has order $2n$.

Example 5 (Alternating group A_n). $A_n = \{\pi \in S_n : \pi \text{ is even}\}$ is a normal subgroup of index 2.

Example 6 (Young (block) subgroups). For a partition $n = m_1 + \cdots + m_r$, the subgroup $S_{m_1} \times \cdots \times S_{m_r} \subseteq S_n$ permutes elements within each block; useful for symmetrizing tensor indices.

2.1 Permutation representation on n registers

Let $\mathcal{H} \cong \mathbb{C}^d$ and consider $\mathcal{H}^{\otimes n}$ with computational basis $\{|i_1, \ldots, i_n\rangle\}$.

Definition 6 (Unitary permutation operators)**.** For $\pi \in S_n$, define $P(\pi)$ by

$$P(\pi)\,|i_1, \ldots, i_n\rangle = |i_{\pi^{-1}(1)}, \ldots, i_{\pi^{-1}(n)}\rangle\,. \tag{1}$$

Proposition 2 (Homomorphism property)**.** $P : S_n \to \mathrm{U}(\mathcal{H}^{\otimes n})$ *is a group homomorphism(representation):* $P(\pi)P(\sigma) = P(\pi\sigma)$*,* $P(\mathrm{id}) = \mathbb{1}$*, and* $P(\pi)^{-1} = P(\pi^{-1})$.

Remark 2. For this property, we will extend to the representation theory in a later document.

Remark 3 (Symmetric subspace). The symmetric subspace is the $+1$ eigenspace of all $P(\pi)$, i.e. vectors invariant under every permutation of the n registers.

3 Conjugacy in groups and in S_n

Definition 7 (Conjugacy and conjugacy class)**.** In a group G, elements g, h are *conjugate* if there exists $x \in G$ with $h = xgx^{-1}$. The *conjugacy class* of g is $C_G(g) = \{xgx^{-1} : x \in G\}$.

Definition 8 (Cycle type in S_n)**.** Write a permutation $\pi \in S_n$ as a product of disjoint cycles. If m_ℓ denotes the number of ℓ-cycles of π (so $\sum_{\ell\geq 1} \ell\, m_\ell = n$), then the *cycle type* of π is the multiset of lengths

$$\mathrm{type}(\pi) = 1^{m_1} 2^{m_2} 3^{m_3} \cdots,$$

equivalently the partition $n = \sum_{\ell\geq 1} \ell\, m_\ell$.

Theorem 1 (Conjugacy in S_n = same cycle type)**.** *Two permutations $\pi, \sigma \in S_n$ are conjugate in S_n if and only if their cycle decompositions have the same* cycle type *(i.e. the same multiset of cycle lengths).*

Proof sketch. If $\sigma = \tau\pi\tau^{-1}$, then σ is obtained from π by relabeling symbols via τ; conjugation preserves cycle lengths, so cycle types match. Conversely, if π and σ have the same cycle type, pair each cycle of π with a cycle of σ of the same length and define a bijection τ that maps elements along corresponding positions in each cycle. Then $\tau\pi\tau^{-1} = \sigma$. $\square$

Example 7 (Same cycle type $\Rightarrow$ same conjugacy class). In S_6 let

$$\pi = (1\,3\,2)(4\,5)(6) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 1 & 2 & 5 & 4 & 6 \end{pmatrix}.$$

Its cycle type is $3^1\,2^1\,1^1$ (partition $3+2+1$). The permutation

$$\sigma = (1\,4\,2)(3)(5\,6) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 1 & 3 & 2 & 6 & 5 \end{pmatrix}$$

has the same cycle type $3^1\,2^1\,1^1$, hence π and σ are conjugate in S_6.

Example 8 (Other basic types). The identity has type 1^n. Any transposition $(a\,b)$ has type $2^1\,1^{n-2}$ (so all transpositions are conjugate). A 4-cycle (e.g. $(1\,2\,3\,4)$ in S_6) has type $4^1\,1^2$, which is *not* the same as $3^1\,2^1\,1^1$, so it lies in a different conjugacy class.

Proposition 3 (Size of a conjugacy class in S_n)**.** *Let the cycle type of $\pi \in S_n$ be specified by integers $m_\ell \geq 0$ (the number of ℓ-cycles), so that $\sum_{\ell\geq 1} \ell\, m_\ell = n$. Then*

$$|C_{S_n}(\pi)| \;=\; \frac{n!}{\prod_{\ell\geq 1} \ell^{m_\ell}\, m_\ell!}.$$

Example 9. In S_6, the type (3)(2)(1) has $m_1 = 1$, $m_2 = 1$, $m_3 = 1$. The conjugation class size is $6!/(1^1\,1! \cdot 2^1\,1! \cdot 3^1\,1!) = 720/6 = 120$.

Unitary representations

Definition 9 (Unitary representation)**.** Let G be a group and V a complex inner-product space. A *unitary representation* of G on V is a homomorphism $\mu : G \to \mathrm{U}(V)$, i.e. $\mu(gh) = \mu(g)\mu(h)$ for all $g, h \in G$, and each $\mu(g)$ is unitary.

Permutation (tensor) representation of S_n. Let $\mathcal{H} \cong \mathbb{C}^d$ and consider $\mathcal{H}^{\otimes n}$ with computational basis $\{|i_1, \ldots, i_n\rangle : i_k \in [d]\}$. For $\pi \in S_n$ define

$$P(\pi)\,|i_1, \ldots, i_n\rangle = |i_{\pi^{-1}(1)}, \ldots, i_{\pi^{-1}(n)}\rangle\,.$$

Then $P : S_n \to \mathrm{U}(\mathcal{H}^{\otimes n})$ is a unitary representation: $P(\pi)P(\sigma) = P(\pi\sigma)$, $P(\mathrm{id}) = \mathrm{Id}$, and $P(\pi)^\dagger = P(\pi^{-1})$ (so $P(\pi)$ is unitary).

Example 10 ($n = 2$: the SWAP). For $\pi = (1\,2)$, $P(\pi)\,|i, j\rangle = |j, i\rangle$; this is the usual SWAP gate. Its matrix in the basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ is

$$\mathrm{SWAP} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Symmetric states and the symmetric subspace

Definition 10 (Symmetric vector and subspace)**.** A vector $|\psi\rangle \in \mathcal{H}^{\otimes n}$ is *symmetric* if $P(\pi)|\psi\rangle = |\psi\rangle$ for all $\pi \in S_n$. The *symmetric subspace* is

$$\mathrm{Sym}^n(\mathbb{C}^d) \;=\; \{|\psi\rangle \in \mathcal{H}^{\otimes n} : P(\pi)|\psi\rangle = |\psi\rangle \;\; \forall \pi \in S_n\}.$$

Example 11 (Symmetric vectors). For any $|v\rangle \in \mathbb{C}^d$, the n-fold product $|v\rangle^{\otimes n}$ is symmetric. For $d = 2$, $n = 2$, the vectors $|00\rangle$, $|11\rangle$, and $\frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$ are symmetric. The uniform superposition $\sum_{x\in[d]^n} |x\rangle$ is also symmetric.

Type classes (histograms) and type vectors

Fix d, n. For a string $x = (x_1, \ldots, x_n) \in [d]^n$, its *type (histogram)* is $\tau(x) = (\tau_1, \ldots, \tau_d)$ where $\tau_a = \#\{k : x_k = a\}$ and $\sum_{a=1}^d \tau_a = n$. Let $T_\tau = \{x \in [d]^n : \tau(x) = \tau\}$ and define the *type vector*

$$|\tau\rangle := \frac{1}{\sqrt{|T_\tau|}} \sum_{x\in T_\tau} |x\rangle \,.$$

Proposition 4. *Each $|\tau\rangle$ is symmetric; the family $\{|\tau\rangle\}$ (over all histograms τ) is orthonormal.*

Proof. For any π, $P(\pi)$ permutes the strings inside T_τ, so $P(\pi)|\tau\rangle = |\tau\rangle$. If $\tau \neq \tau'$, then $T_\tau \cap T_{\tau'} = \varnothing$, hence $\langle\tau\rangle\,\tau' = 0$. Normalization is by the $1/\sqrt{|T_\tau|}$ factor. □

Theorem 2 (Type basis and dimension)**.** *The type vectors $\{|\tau\rangle\}$ form an orthonormal basis of* $\mathrm{Sym}^n(\mathbb{C}^d)$*. Consequently,*

$$\dim \mathrm{Sym}^n(\mathbb{C}^d) = \#\{\textit{histograms } \tau\} = \binom{n+d-1}{d-1}.$$

Idea. Any symmetric vector must assign equal amplitudes to all strings of the same type (otherwise some permutation changes the state), so it lies in the span of $\{|\tau\rangle\}$; together with Proposition 4, these vectors form an ONB. Counting histograms is the stars-and-bars argument. □

Span by product states and a Vandermonde argument

Define $S := \mathrm{span}\{|v\rangle^{\otimes n} : |v\rangle \in \mathbb{C}^d\}$. Clearly $S \subseteq \mathrm{Sym}^n(\mathbb{C}^d)$. We show $S = \mathrm{Sym}^n(\mathbb{C}^d)$ by proving that each type vector lies in S.

Case $d = 2$ (explicit). Write types as $\tau_i = (n - i, i)$, $i = 0, \ldots, n$, and $|\tau_i\rangle = \frac{1}{\sqrt{\binom{n}{i}}} \sum_{|x|=i} |x\rangle$, where $|x|$ counts 1's. For any $z \in \mathbb{C}$,

$$(|0\rangle + z\,|1\rangle)^{\otimes n} = \sum_{i=0}^{n} z^i \sqrt{\binom{n}{i}}\;|\tau_i\rangle \,.$$

Choose $K = n + 1$ distinct complex numbers $z_1, \ldots, z_{n+1}$ and consider the system

$$\sum_{j=1}^{n+1} \alpha_j \,(|0\rangle + z_j\,|1\rangle)^{\otimes n} = \sqrt{\binom{n}{i^\star}}\;|\tau_{i^\star}\rangle \,.$$

This reduces to the linear equations $\sum_j \alpha_j z_j^i = \delta_{i,i^\star}$ for $i = 0, \ldots, n$, whose coefficient matrix is the $(n+1) \times (n+1)$ Vandermonde $V = (z_j^i)$. Since the z_j are distinct, V is invertible; thus every $|\tau_i\rangle$ is a linear combination of $|v\rangle^{\otimes n}$'s, so S contains the type basis.

General d. This is a high-level understanding of the proof later.A multivariate version uses $(\sum_{a=1}^{d} z_a |a\rangle)^{\otimes n}$ and separates coefficients by choosing a finite grid of d-tuples $z^{(j)} = (z_1^{(j)}, \ldots, z_d^{(j)})$ so that the associated multivariate Vandermonde matrix is invertible; this yields each $|\tau\rangle$. Hence $S = \mathrm{Sym}^n(\mathbb{C}^d)$.

Concrete examples ($n = 2$, $d = 2$)

Type classes and type vectors:

$$\tau = (2,0): \ |\tau\rangle = |00\rangle\,, \qquad \tau = (0,2): \ |\tau\rangle = |11\rangle\,, \qquad \tau = (1,1): \ |\tau\rangle = \tfrac{1}{\sqrt{2}}(|01\rangle + |10\rangle).$$

Recovering $|\tau = (1,1)\rangle$ from product states:

$$\frac{1}{2}\,(|0\rangle + |1\rangle)^{\otimes 2} - \frac{1}{2}\,(|0\rangle - |1\rangle)^{\otimes 2} \;=\; \tfrac{1}{\sqrt{2}}(|01\rangle + |10\rangle).$$

Theorem 3 (Product-state span equals the symmetric subspace)**.** *Let $\mathcal{H} \simeq \mathbb{C}^d$. Then*

$$\mathrm{span}\{\,|v\rangle^{\otimes n} : |v\rangle \in \mathcal{H}\,\} \;=\; \mathrm{Sym}^n(\mathbb{C}^d).$$

Proof. It is clear that every $|v\rangle^{\otimes n}$ is invariant under all register permutations, so the left-hand side is contained in $\mathrm{Sym}^n(\mathbb{C}^d)$. To prove the reverse inclusion we show that the standard *type (histogram) basis* of $\mathrm{Sym}^n(\mathbb{C}^d)$ lies in the span of product states.

Step 1 (set up type vectors). For $d = 2$ write types as $\tau_i = (n-i, i)$ and define

$$|\tau_i\rangle \;=\; \frac{1}{\sqrt{\binom{n}{i}}} \sum_{\substack{x \in \{0,1\}^n \\ |x| = i}} |x\rangle\,, \qquad i = 0, 1, \ldots, n.$$

Then $\{|\tau_i\rangle\}_{i=0}^{n}$ is an orthonormal basis of $\mathrm{Sym}^n(\mathbb{C}^2)$. The binomial expansion gives, for any $z \in \mathbb{C}$,

$$(|0\rangle + z\,|1\rangle)^{\otimes n} = \sum_{i=0}^{n} z^i \sqrt{\binom{n}{i}}\;|\tau_i\rangle\,. \tag{2}$$

Step 2 (Vandermonde isolation for $d = 2$). Fix $i^\star \in \{0, \ldots, n\}$. Choose $K = n+1$ distinct complex numbers $z_1, \ldots, z_{n+1}$ and seek coefficients $\alpha_1, \ldots, \alpha_{n+1}$ such that

$$\sum_{j=1}^{n+1} \alpha_j\,(|0\rangle + z_j\,|1\rangle)^{\otimes n} = |\tau_{i^\star}\rangle\,.$$

Using (2) this is equivalent to the linear system

$$\sum_{j=1}^{n+1} \alpha_j\, z_j^{\,i} = \begin{cases} \dfrac{1}{\sqrt{\binom{n}{i^\star}}}, & i = i^\star, \\ 0, & i \neq i^\star, \end{cases} \qquad i = 0, 1, \ldots, n.$$

In matrix form $V\alpha = e_{i^\star}/\sqrt{\binom{n}{i^\star}}$, where

$$V = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ z_1 & z_2 & \cdots & z_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ z_1^n & z_2^n & \cdots & z_{n+1}^n \end{pmatrix}$$

is the $(n+1) \times (n+1)$ Vandermonde matrix. Since the z_j are distinct, V is invertible; hence such α exists and $|\tau_{i^\star}\rangle$ is a linear combination of product states. As the $|\tau_i\rangle$'s span $\mathrm{Sym}^n(\mathbb{C}^2)$, we have equality for $d = 2$.

Step 3 (general d via a univariate reduction). For $d \geq 2$, index types by $m = (m_1, \dots, m_d) \in \mathbb{N}^d$ with $|m| := \sum_a m_a = n$ and set

$$|\tau_m\rangle := \sqrt{\frac{\prod_{a=1}^d m_a!}{n!}} \sum_{\substack{x \in [d]^n \\ \mathrm{type}(x) = m}} |x\rangle ,$$

an orthonormal basis of $\mathrm{Sym}^n(\mathbb{C}^d)$. The multinomial theorem yields, for $z = (z_1, \dots, z_d) \in \mathbb{C}^d$,

$$\Big(\sum_{a=1}^d z_a |a\rangle\Big)^{\otimes n} = \sum_{|m|=n} z^m \sqrt{\frac{n!}{\prod_a m_a!}} \, |\tau_m\rangle , \qquad z^m := \prod_{a=1}^d z_a^{m_a}. \tag{3}$$

Choose a base $B := n + 1$ and distinct scalars $t_1, \dots, t_M$ with $M = \binom{n+d-1}{d-1}$. Define points $z^{(j)} \in \mathbb{C}^d$ by

$$z_a^{(j)} := t_j^{B^{a-1}}, \quad a = 1, \dots, d.$$

For $|m| = n$ the monomial evaluates to

$$\big(z^{(j)}\big)^m = \prod_{a=1}^d t_j^{m_a B^{a-1}} = t_j^{\sum_{a=1}^d m_a B^{a-1}}.$$

Because $0 \leq m_a \leq n$ and the base is $B = n+1$, the exponent $\sum_a m_a B^{a-1}$ is the base-B encoding of m; distinct m's yield distinct exponents. Thus the evaluation matrix with entries $\big(z^{(j)}\big)^m$ is a (rectangular) Vandermonde in the variables t_j with *distinct* exponents, hence has full row rank. Arguing exactly as in Step 2, we can linearly combine the product states $\big(\sum_a z_a^{(j)} |a\rangle\big)^{\otimes n}$ to isolate any fixed $|\tau_m\rangle$. Therefore, every type of vector lies in the span of product states, proving the reverse inclusion. □

Definition 11 (Symmetrizer / projector onto the symmetric subspace)**.** Let $\mathrm{Sym}^n(\mathbb{C}^d) \subset (\mathbb{C}^d)^{\otimes n}$ be the symmetric subspace. Define the *symmetrizer*

$$\Pi_{\mathrm{sym}} := \frac{1}{n!} \sum_{\pi \in S_n} P(\pi).$$

Proposition 5 (Uniform pushforward on S_n)**.** *If π is uniform on S_n and $\sigma \in S_n$ is fixed, then $\pi\sigma$ is also uniform. Equivalently, for any function $f : S_n \to \mathbb{C}$,*

$$\mathbb{E}_{\pi \sim S_n} f(\pi\sigma) = \mathbb{E}_{\pi \sim S_n} f(\pi), \qquad \textit{and} \qquad \Pr[\pi = \tau] = \frac{1}{n!} \ \forall \tau \in S_n.$$

Theorem 4 (Averaging projector)**.** *The operator* Π_{sym} *defined in Definition 11 is the* orthogonal projector *onto* $\mathrm{Sym}^n(\mathbb{C}^d)$. *In particular,*

$$\Pi_{\mathrm{sym}}^\dagger = \Pi_{\mathrm{sym}}, \qquad \Pi_{\mathrm{sym}}^2 = \Pi_{\mathrm{sym}}, \qquad \mathrm{Ran}(\Pi_{\mathrm{sym}}) = \mathrm{Sym}^n(\mathbb{C}^d).$$

Proof. **Hermitian.** Since $P(\pi)^\dagger = P(\pi^{-1})$ and the map $\pi \mapsto \pi^{-1}$ is a bijection of S_n,

$$\Pi_{\mathrm{sym}}^\dagger = \frac{1}{n!} \sum_{\pi \in S_n} P(\pi)^\dagger = \frac{1}{n!} \sum_{\pi \in S_n} P(\pi^{-1}) = \frac{1}{n!} \sum_{\pi \in S_n} P(\pi) = \Pi_{\mathrm{sym}}.$$

Idempotent. Using group multiplication and Proposition 5,

$$\Pi_{\rm sym}^2 = \Big(\tfrac{1}{n!}\sum_{\pi} P(\pi)\Big)\Big(\tfrac{1}{n!}\sum_{\sigma} P(\sigma)\Big) = \frac{1}{(n!)^2}\sum_{\pi,\sigma} P(\pi\sigma) = \frac{1}{n!}\sum_{\tau\in S_n} P(\tau) = \Pi_{\rm sym},$$

because for each fixed τ there are exactly $n!$ pairs (π,σ) with $\pi\sigma = \tau$ (take any σ and set $\pi = \tau\sigma^{-1}$).

Since $\Pi_{\rm sym}$ is Hermitian and idempotent, it is an orthogonal projector onto its range.

Range equals the symmetric subspace. (i) If $|\psi\rangle \in \mathrm{Sym}^n(\mathbb{C}^d)$, then $P(\pi)\,|\psi\rangle = |\psi\rangle$ for all π; hence $\Pi_{\rm sym}\,|\psi\rangle = \frac{1}{n!}\sum_\pi |\psi\rangle = |\psi\rangle$. Thus $\mathrm{Sym}^n(\mathbb{C}^d) \subseteq \mathrm{Ran}(\Pi_{\rm sym})$.

(ii) Conversely, for any $|\phi\rangle$ and any $\sigma \in S_n$,

$$P(\sigma)\,\Pi_{\rm sym}\,|\phi\rangle = \frac{1}{n!}\sum_{\pi} P(\sigma\pi)\,|\phi\rangle = \frac{1}{n!}\sum_{\tau} P(\tau)\,|\phi\rangle = \Pi_{\rm sym}\,|\phi\rangle\,,$$

relabeling $\tau = \sigma\pi$. Hence $\Pi_{\rm sym}\,|\phi\rangle$ is invariant under all permutations, so $\mathrm{Ran}(\Pi_{\rm sym}) \subseteq \mathrm{Sym}^n(\mathbb{C}^d)$.

Combining (i) and (ii) completes the proof. □

Example 12 ($n = 2$). Here $S_2 = \{e, (1\,2)\}$ and $P(1\,2) = \mathrm{SWAP}$. Thus

$$\Pi_{\rm sym} = \tfrac{1}{2}\big(I + \mathrm{SWAP}\big),$$

which projects onto the span of $\{|00\rangle\,,\ \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle),\ |11\rangle\}$.

4 The symmetric subspace and the unitary group action

Definition 12 (Unitary group)**.** $\mathrm{U}(d) = \{U \in \mathbb{C}^{d\times d} : U^\dagger U = UU^\dagger = I\}$.

Proposition 6. $\mathrm{U}(d)$ *is a group under matrix multiplication (associativity, identity I, and inverses $U^\dagger$).*

Definition 13 (Tensor (diagonal) action of $\mathrm{U}(d)$)**.** For $n \geq 1$ and $U \in \mathrm{U}(d)$ define the unitary on $(\mathbb{C}^d)^{\otimes n}$

$$Q(U) := U^{\otimes n}.$$

Fact 1 (Representation property). $Q : \mathrm{U}(d) \to \mathrm{U}\big((\mathbb{C}^d)^{\otimes n}\big)$ is a unitary representation since $Q(U)Q(V) = U^{\otimes n}V^{\otimes n} = (UV)^{\otimes n} = Q(UV)$.

Proposition 7 (Invariance of the symmetric subspace)**.** *Let* $\mathrm{Sym}^n(\mathbb{C}^d) \subset (\mathbb{C}^d)^{\otimes n}$ *be the symmetric subspace. Then* $Q(U)\,\mathrm{Sym}^n(\mathbb{C}^d) \subseteq \mathrm{Sym}^n(\mathbb{C}^d)$ *for every* $U \in \mathrm{U}(d)$*.*

Proof. By Theorem "product-state span = symmetric subspace", every $|\psi\rangle \in \mathrm{Sym}^n(\mathbb{C}^d)$ can be written as $|\psi\rangle = \sum_i \alpha_i\,|v_i\rangle^{\otimes n}$. Then $Q(U)\,|\psi\rangle = \sum_i \alpha_i(U\,|v_i\rangle)^{\otimes n}$, which is again a linear combination of n-fold product states, hence symmetric. □

Remark 4. The permutation representation $P : S_n \to \mathrm{U}\big((\mathbb{C}^d)^{\otimes n}\big)$ acts trivially on $\mathrm{Sym}^n(\mathbb{C}^d)$: $P(\pi)\,|\psi\rangle = |\psi\rangle$ for all $\pi \in S_n$ and $|\psi\rangle \in \mathrm{Sym}^n(\mathbb{C}^d)$.

Haar measure and Haar-random vectors

Definition 14 (Haar measure on U(d))**.** The (normalized) Haar measure μ_{Haar} is the unique probability measure on U(d) that is invariant under left and right multiplication: $\mu_{\text{Haar}}(VUW) = \mu_{\text{Haar}}(U)$ for all fixed $V, W \in \text{U}(d)$.

Fact 2 (Haar pushforward to the sphere). Fix any unit vector $|v\rangle \in \mathbb{C}^d$. If $U \sim \mu_{\text{Haar}}$, then $U|v\rangle$ is a *Haar-random unit vector* (i.e., uniformly distributed on the complex unit sphere). Conversely, a Haar-random unitary can be obtained by sampling d i.i.d. complex Gaussian vectors, applying Gram–Schmidt, and stacking them as columns.

Theorem 5 (Haar moment on the symmetric projector)**.** *Let $|v\rangle$ be a Haar-random unit vector in $\mathbb{C}^d$ and*

$$M \ := \ \mathbb{E}\big[\ |v\rangle\langle v|^{\otimes n}\ \big].$$

Then M is a scalar multiple of the symmetrizer $\Pi_{\text{sym}} = \frac{1}{n!}\sum_{\pi\in S_n} P(\pi)$, namely

$$\boxed{M \ = \ \frac{1}{\binom{n+d-1}{d-1}}\,\Pi_{\text{sym}}}\,.$$

Proof. (i) Invariance under $U(d)$. If $U \sim \mu_{\text{Haar}}$, then $U|v\rangle$ is Haar-random; hence

$$U^{\otimes n} M\, U^{\otimes n\dagger} = \mathbb{E}\big[(U|v\rangle\langle v|U^\dagger)^{\otimes n}\big] = M.$$

Thus M lies in the commutant of $Q(\text{U}(d)) = \{U^{\otimes n}\}$.

(ii) Support in the symmetric subspace. Each sample $|v\rangle\langle v|^{\otimes n}$ has range contained in $\text{Sym}^n(\mathbb{C}^d)$; therefore so does M. Hence M acts as zero on the orthogonal complement of $\text{Sym}^n(\mathbb{C}^d)$.

(iii) Proportionality to Π_{sym}. By Schur–Weyl duality (or by the fact that the only operators on the irreducible $U(d)$-module $\text{Sym}^n(\mathbb{C}^d)$ commuting with all $U^{\otimes n}$ are scalars), the restriction of M to $\text{Sym}^n(\mathbb{C}^d)$ is a scalar multiple of the identity there: $M = c\,\Pi_{\text{sym}}$ for some $c > 0$.

(iv) Determine c by traces. Since $\text{Tr}(|v\rangle\langle v|^{\otimes n}) = 1$, we have $\text{Tr}(M) = 1$. Also $\text{Tr}(\Pi_{\text{sym}}) = \dim \text{Sym}^n(\mathbb{C}^d) = \binom{n+d-1}{d-1}$. Therefore $1 = \text{Tr}(M) = c\binom{n+d-1}{d-1}$, giving $c = \binom{n+d-1}{d-1}^{-1}$. □

Example 13 ($n = 2$). Using $\Pi_{\text{sym}} = \frac{1}{2}(I + F)$ (with F the swap),

$$\mathbb{E}\big[\,|v\rangle\langle v|^{\otimes 2}\,\big] = \frac{2}{d(d+1)}\,\Pi_{\text{sym}} = \frac{I+F}{d(d+1)},$$

the familiar second-moment identity.

4.1 Toy example: the $n = 1$ case

Let $|v\rangle \in \mathbb{C}^d$ be Haar–random on the unit sphere and expand in the computational basis $|v\rangle = \sum_{i=1}^d v_i\,|i\rangle$ with $\sum_i |v_i|^2 = 1$. Then

$$\mathbb{E}[|v\rangle\langle v|] = \mathbb{E}\left[\Big(\sum_i v_i\,|i\rangle\Big)\Big(\sum_j \bar{v}_j\,\langle j|\Big)\right] = \sum_{i,j}\mathbb{E}[v_i\bar{v}_j]\ |i\rangle\langle j|\,.$$

Off–diagonals vanish. For any diagonal phase unitary $D = \text{diag}(e^{i\theta_1}, \ldots, e^{i\theta_d})$, $D|v\rangle$ is also Haar–distributed, so

$$\mathbb{E}[v_i\bar{v}_j] = \mathbb{E}[(e^{i\theta_i}v_i)(e^{-i\theta_j}\bar{v}_j)] = e^{i(\theta_i-\theta_j)}\mathbb{E}[v_i\bar{v}_j] \quad \forall\,\theta_i,\theta_j.$$

If $i \neq j$ this forces $\mathbb{E}[v_i\bar{v}_j] = 0$.

Diagonals are all equal and sum to 1**.** By permutation invariance of the Haar measure, $\mathbb{E}[|v_i|^2]$ is the same for all i; write $\mathbb{E}[|v_i|^2] = \alpha$. Taking expectations in $\sum_i |v_i|^2 = 1$ yields

$$1 = \mathbb{E}\left[\sum_{i=1}^{d} |v_i|^2\right] = \sum_{i=1}^{d} \mathbb{E}[|v_i|^2] = d\,\alpha \quad \Longrightarrow \quad \alpha = \frac{1}{d}.$$

Combining these two facts,

$$\boxed{\mathbb{E}[|v\rangle\langle v|] = \sum_{i=1}^{d} \frac{1}{d}\ |i\rangle\langle i| = \frac{1}{d}\, I}.$$

5 A potential obstruction and irreducibility

5.1 A potential obstruction

Let $|\phi\rangle = |1\rangle^{\otimes n}$ and recall

$$M \ = \ \mathbb{E}_{U\sim\text{Haar}}\big[\, Q(U)\ |\phi\rangle\langle\phi|\ Q(U)^\dagger \,\big], \qquad Q(U) = U^{\otimes n}.$$

Suppose (hypothetically) that in some fixed basis every $Q(U)$ had the same block–diagonal form $Q(U) = \begin{pmatrix} Q_1(U) & 0 \\ 0 & Q_2(U) \end{pmatrix}$. Then $Q(U)$ would never mix the two invariant subspaces, and averaging could not move a vector from one block into the other. In that case M could be at best a projector onto *one* block, rather than a multiple of the full symmetrizer Π_{sym}. This motivates the need to show that no such nontrivial decomposition exists on the symmetric subspace, i.e. the action is *irreducible.*

5.2 (Ir)reducible representations and examples

Definition 15 (Reducible / irreducible)**.** A unitary representation (μ, V) of a group G is *reducible* if there is a nontrivial proper subspace $0 \neq W \subsetneq V$ with $\mu(g)W \subseteq W$ for all $g \in G$. Otherwise it is *irreducible.* Equivalently, in some basis $\mu(g)$ is block diagonal for all g.

Examples. (1) The permutation representation $P : S_n \to \mathrm{U}((\mathbb{C}^d)^{\otimes n})$ is reducible since it preserves the symmetric subspace $\mathrm{Sym}^n(\mathbb{C}^d)$. (2) The tensor action $Q(U) = U^{\otimes n}$ on $(\mathbb{C}^d)^{\otimes n}$ is also reducible because it preserves $\mathrm{Sym}^n(\mathbb{C}^d)$. (3) On $\mathrm{Sym}^n(\mathbb{C}^d)$, the permutation action $P(\pi)$ is trivial (acts as the identity), hence "extremely" reducible.

5.3 Irreducibility of Q on the symmetric subspace

Theorem 6. *Let* $Q : \mathrm{U}(d) \to \mathrm{U}(\mathrm{Sym}^n(\mathbb{C}^d))$ *be the restricted tensor action* $Q(U) = U^{\otimes n}$*. Then* Q *is irreducible on* $\mathrm{Sym}^n(\mathbb{C}^d)$*.*

Proof. We follow the sketch from the notes.

Assume for contradiction that Q is reducible. Then there are nonzero, proper, *orthogonal* Q–invariant subspaces $X, Y \subset \mathrm{Sym}^n(\mathbb{C}^d)$ with

$$\mathrm{Sym}^n(\mathbb{C}^d) = X \oplus Y, \qquad Q(U)X \subseteq X,\ Q(U)Y \subseteq Y \ \ \forall U \in \mathrm{U}(d).$$

By the product-state span theorem (proved earlier), the set $\{|v\rangle^{\otimes n} : |v\rangle \in \mathbb{C}^d\}$ spans $\mathrm{Sym}^n(\mathbb{C}^d)$. Hence there exists a family of unit vectors $\{|v_i\rangle\}$ and an index set I such that $|v_i\rangle^{\otimes n} \in X$ for all $i \in I$, and (since $Y \neq \{0\}$) there is some $j \notin I$ with $|v_j\rangle^{\otimes n} \in Y$.

Because $\mathrm{U}(d)$ acts transitively on unit vectors, there exists a unitary $U \in \mathrm{U}(d)$ with $U\,|v_j\rangle = |v_{i_0}\rangle$ for some $i_0 \in I$. Then

$$Q(U)\ |v_j\rangle^{\otimes n} = (U\,|v_j\rangle)^{\otimes n} = |v_{i_0}\rangle^{\otimes n} \in X.$$

But $|v_j\rangle^{\otimes n} \in Y$ and Y is Q–invariant, so $Q(U)\,|v_j\rangle^{\otimes n} \in Y$ as well. Thus $|v_{i_0}\rangle^{\otimes n} \in X \cap Y$, a nonzero vector, contradicting $X \perp Y$ and $\mathrm{Sym}^n(\mathbb{C}^d) = X \oplus Y$.

Therefore no such nontrivial invariant decomposition exists, and Q is irreducible on $\mathrm{Sym}^n(\mathbb{C}^d)$. □

5.4 Proof of the Haar moment theorem via irreducibility

Recall Theorem 5: for $|v\rangle$ Haar–random, $M = \mathbb{E}\big[\,|v\rangle\langle v|^{\otimes n}\,\big] = \binom{n+d-1}{d-1}^{-1}\Pi_{\mathrm{sym}}$.

Proof (representation–theoretic). For any fixed $U \in \mathrm{U}(d)$, Haar invariance gives

$$Q(U)\,M\,Q(U)^\dagger = \mathbb{E}\big[(U\,|v\rangle\langle v|\,U^\dagger)^{\otimes n}\big] = M,$$

so M commutes with every $Q(U)$ and acts trivially on the orthogonal complement of Sym^n. By Theorem 6 and Schur's lemma(which we will show later), $M = c\,\Pi_{\mathrm{sym}}$ for some c. Taking traces, $1 = \mathrm{Tr}(M) = c\ \mathrm{Tr}(\Pi_{\mathrm{sym}}) = c\binom{n+d-1}{d-1}$, so $c = \binom{n+d-1}{d-1}^{-1}$. □

6 Pure-state tomography via the Haar POVM

Problem. Given n copies of an unknown pure state $|\psi\rangle \in \mathbb{C}^d$, we perform one collective measurement on $|\psi\rangle^{\otimes n}$ and output an estimate $|\hat{\psi}\rangle$. Our accuracy metric will be the (squared) fidelity $F := |\langle\psi|\hat{\psi}\rangle|^2$.

Equivalence for pure states. For pure states,

$$D_{\mathrm{tr}}\big(\,|\psi\rangle\langle\psi|,\ |\hat{\psi}\rangle\langle\hat{\psi}|\,\big) = \sqrt{1 - |\langle\psi|\hat{\psi}\rangle|^2}.$$

Hence a fidelity target $|\langle\psi|\hat{\psi}\rangle|^2 \geq 1 - \varepsilon^2$ is exactly the trace-distance target $D_{\mathrm{tr}} \leq \varepsilon$.

Proposition 8 (Expected trace distance)**.** *Under the Haar POVM estimator $\hat{\psi} = v$,*

$$\mathbb{E}\Big[\,D_{\mathrm{tr}}\big(|\psi\rangle\langle\psi|, |\hat{\psi}\rangle\langle\hat{\psi}|\big)\,\Big] \leq \sqrt{1 - \mathbb{E}\Big[|\langle\psi|\hat{\psi}\rangle|^2\Big]} = \sqrt{\frac{d-1}{n+d}} \ \leq\ \sqrt{\frac{d}{n}}.$$

(The inequality uses concavity of $x \mapsto \sqrt{1-x}$ and Theorem 8.)

Theorem 7 (Tail bound / error exponent in trace distance)**.** *Let $F = |\langle\psi|\hat{\psi}\rangle|^2$. Then $F \sim \mathrm{Beta}(n{+}1, d{-}1)$, so for any $\varepsilon \in (0,1)$,*

$$\Pr\big[\,D_{\mathrm{tr}} \geq \varepsilon\,\big] = \Pr\big[\,1 - F \geq \varepsilon^2\,\big] = I_{1-\varepsilon^2}(n{+}1, d{-}1) \ \leq\ \frac{(n{+}d{-}1)^{d-2}}{(d{-}2)!\,(n{+}1)}\, e^{-(n+1)\varepsilon^2}.$$

Thus the error probability decays at least like $\mathrm{poly}(n,d)\,e^{-(n+1)\varepsilon^2}$ *with exponent* $(n{+}1)\varepsilon^2$.

Proposition 9 (Samples for trace-distance target with tail $\leq \delta$)**.** *To ensure* $\Pr[D_{\mathrm{tr}} \geq \varepsilon] \leq \delta$*, it suffices to take*

$$n \ \geq\ \frac{(d-2)\,\log(n+d-1) + \log\big(\frac{1}{(d-2)!\,\delta}\big)}{\varepsilon^2} - 1 \quad = \ O\Big(\frac{d + \log(1/\delta)}{\varepsilon^2}\Big).$$

The Haar POVM on $\mathrm{Sym}^n(\mathbb{C}^d)$

By Theorem 5, for Haar–random $|v\rangle$, $\mathbb{E}[|v\rangle\langle v|^{\otimes n}] = \binom{n+d-1}{d-1}^{-1}\Pi_{\mathrm{sym}}$. This implies that the operator density

$$E(dv) \;=\; \binom{n+d-1}{d-1}\, |v\rangle\langle v|^{\otimes n}\; d\nu(v), \tag{4}$$

with $d\nu$ the normalized Haar measure on the unit sphere of $\mathbb{C}^d$, forms a valid POVM on $\mathrm{Sym}^n(\mathbb{C}^d)$. For input $|\psi\rangle^{\otimes n}$ the outcome law is

$$\Pr[dv \mid \psi] \;=\; \binom{n+d-1}{d-1}\, |\langle v\rangle\, \psi|^{2n}\, d\nu(v). \tag{5}$$

We use the simple estimator $|\hat{\psi}\rangle := |v\rangle$ (the outcome direction).

Proposition 10. *For Haar–random $|v\rangle$ and any fixed $|\psi\rangle$,*

$$\mathbb{E}_{\mathrm{Haar}}\big[\langle\psi|\hat{\psi}\rangle^{2m}\,\big] = \binom{m+d-1}{d-1}^{-1} \qquad (m \in \mathbb{N}).$$

Proof. By Theorem 5, $\mathbb{E}[|v\rangle\langle v|^{\otimes m}] = \binom{m+d-1}{d-1}^{-1}\Pi_{\mathrm{sym}}$. Taking the matrix element on $|\psi\rangle^{\otimes m}$ gives the claim since $\Pi_{\mathrm{sym}}\,|\psi\rangle^{\otimes m} = |\psi\rangle^{\otimes m}$. □

Theorem 8 (Expected fidelity)**.** *If we measure $|\psi\rangle^{\otimes n}$ with the Haar POVM* (4) *and output $|\hat{\psi}\rangle = |v\rangle$, then*

$$\mathbb{E}\Big[\,\langle\psi|\hat{\psi}\rangle^2\,\Big] = \frac{n+1}{n+d} = 1 - \frac{d-1}{n+d}.$$

Proof. From (5),

$$\mathbb{E}[\langle\psi|\hat{\psi}\rangle^2] = \binom{n+d-1}{d-1}\int |\langle\psi\rangle\, v|^{2(n+1)}\, d\nu(v) = \binom{n+d-1}{d-1}\binom{n+d}{d-1}^{-1},$$

using Lemma 10 with $m = n+1$. The ratio simplifies to $(n+1)/(n+d)$. □

Full distribution and an error exponent

Let $F := |\langle\psi\rangle\, \hat{\psi}|^2$. Combining (5) with the well-known fact that $T := |\langle v\rangle\, \psi|^2$ is $\mathrm{Beta}(1, d-1)$ under Haar measure, we find

$$F \sim \mathrm{Beta}\big(n+1,\ d-1\big) \quad \text{with density} \quad f_F(t) = \frac{t^n(1-t)^{d-2}}{B(n+1, d-1)}\ (t \in [0,1]),$$

where B is the beta function. In particular, for any $\varepsilon \in (0,1)$,

$$\Pr\big[\,1 - F \ge \varepsilon^2\,\big] = I_{1-\varepsilon^2}(n+1, d-1), \tag{6}$$

the regularized incomplete beta.

A convenient explicit upper bound (polynomial pre-factor with an *exponential* rate) is

$$\Pr\big[\,1 - F \ge \varepsilon^2\,\big] \le \frac{1}{(n+1)\, B(n+1, d-1)}\,(1-\varepsilon^2)^{\,n+1} \le \frac{(n+d-1)^{d-2}}{(d-2)!\,(n+1)}\, e^{-(n+1)\varepsilon^2}. \tag{7}$$

The first inequality integrates the density on $[0, 1-\varepsilon^2]$ and the second uses $(1-x) \le e^{-x}$ and $1/B(n+1, d-1) \le \frac{(n+d-1)^{d-2}}{(d-2)!}$. Thus, the *error exponent* is at least $(n+1)\varepsilon^2$ up to a dimension–dependent polynomial pre-factor.

Why Beta, not just "a concentration bound"? Let $T = |\langle v|\psi\rangle|^2$. For $v \sim$ Haar on $\mathbb{C}^d$,

$$T \sim \mathrm{Beta}(1, d-1),$$

because $|v_1|^2$ is the ratio of two independent Γ variables (Dirichlet on the sphere). Under the Haar POVM, the outcome density is tilted by T^n [Eq. (5)], so the posterior law of $F := |\langle\hat{\psi}|\psi\rangle|^2$ is

$$F \sim \mathrm{Beta}(n+1, d-1).$$

This one-dimensional reduction has three advantages:

1. **Exact quantities.** We get $\mathbb{E}[F] = \dfrac{n+1}{n+d}$ and, for any m, $\mathbb{E}[F^m] = \binom{m+d-1}{d-1}^{-1}$ *exactly*, and the tail is the regularized incomplete beta:
$$\Pr[1-F \geq \varepsilon^2] = I_{1-\varepsilon^2}(n+1, d-1).$$

2. **Sharp, dimension-aware tails.** From the Beta form,
$$\Pr[1-F \geq \varepsilon^2] = \frac{1}{B(n+1,d-1)} \int_0^{1-\varepsilon^2} t^n (1-t)^{d-2}\, dt \;\leq\; \frac{(1-\varepsilon^2)^{n+1}}{B(n+1,d-1)},$$
which yields the explicit error exponent $e^{-(n+1)\varepsilon^2}$ up to a polynomial pre-factor in $(n+d)$ [cf. Eq. (7)]. This captures the correct n–vs–d dependence with the best constants you can hope for from this route.

3. **No independence assumptions.** Standard tools like Hoeffding/Bernstein apply to *sums of i.i.d.* variables; here F is a single draw whose density already encodes n (via t^n). Forcing a generic concentration argument either does not apply directly or gives looser bounds.

Proposition 11. *[Sample complexity with tail $\leq \delta$] For any $\delta \in (0,1)$, it suffices to take*

$$n \;\geq\; \frac{(d-2)\,\log(n+d-1) + \log\!\left(\frac{1}{(d-2)!\,\delta}\right)}{\varepsilon^2} - 1$$

to guarantee $\Pr[\,1-F \geq \varepsilon^2\,] \leq \delta$*. In coarse scaling,* $n = O\big((d + \log(1/\delta))/\varepsilon^2\big)$.

7 Single-copy tomography with a Haar-random basis

Definition 16 (Haar-random basis). If $U \in \mathrm{U}(d)$ is Haar-random and $\{|1\rangle, \ldots, |d\rangle\}$ is a fixed orthonormal basis, then $\{\, |u_i\rangle := U\,|i\rangle \,\}_{i=1}^d$ is a *Haar-random basis.*

Algorithm (incomplete single-copy tomography).

1. Draw a Haar-random basis $\{|u_1\rangle, \ldots, |u_d\rangle\}$.
2. Measure ρ in this basis; let the outcome be $|u\rangle$.
3. Output the estimator
$$\widehat{\rho} \;:=\; (d{+}1)\,|u\rangle\langle u| - I.$$

Haar-basis measurement equals the uniform POVM

[Symmetry of outcomes in a Haar basis] For a Haar-random basis $\{|u_i\rangle\}$ and any state ρ, $\Pr[\text{outcome} = |u_1\rangle] = \cdots = \Pr[\text{outcome} = |u_d\rangle]$.

[Outcome density] Let $|u\rangle$ be any unit vector. Then

$$\Pr[\text{outcome} \in d\nu(u) \text{ around } |u\rangle] = d\ \mathrm{Tr}\big[\,|u\rangle\langle u|\ \rho\big]\ d\nu(u),$$

where $d\nu$ is the normalized Haar measure on the unit sphere of $\mathbb{C}^d$.

Definition 17 (Uniform POVM)**.** The *uniform POVM* on $\mathbb{C}^d$ has operator density

$$E(du) = d\ |u\rangle\langle u|\ d\nu(u) \qquad (\text{so } \int E(du) = I).$$

Performing the Haar-basis measurement is equivalent to applying this POVM.

Unbiased single-copy estimator

Write $Q(U) = U^{\otimes 2}$, F for SWAP, and $\Pi_{\mathrm{sym},2} = \frac{1}{2}(I + F)$. Using Theorem 5 with $n = 2$, $\int |u\rangle\langle u|^{\otimes 2}\, d\nu(u) = \dfrac{2}{d(d+1)}\,\Pi_{\mathrm{sym},2}$. Then

$$\begin{aligned}
\mathbb{E}[|u\rangle\langle u|] &= \int |u\rangle\langle u|\ \Pr[\text{outcome} \in du] = d \int |u\rangle\langle u|\ \mathrm{Tr}\big[\,|u\rangle\langle u|\ \rho\big]\, d\nu(u) \\
&= d\ \mathrm{Tr}_B\left[\Big(\int |u\rangle\langle u|^{\otimes 2}\, d\nu(u)\Big)(I \otimes \rho)\right] = \frac{2}{d+1}\ \mathrm{Tr}_B\big[\Pi_{\mathrm{sym},2}(I \otimes \rho)\big] \\
&= \frac{1}{d+1}\ \mathrm{Tr}_B\big[(I + F)(I \otimes \rho)\big] = \frac{1}{d+1}\,(I + \rho).
\end{aligned}$$

Hence:

Theorem 9 (Unbiasedness)**.** *With the uniform POVM and estimator* $\widehat{\rho} = (d{+}1)\,|u\rangle\langle u| - I$*,*

$$\mathbb{E}[\widehat{\rho}] = \rho.$$

Second moment and a variance bound

For any outcome $|u\rangle$, in a basis with $|u\rangle$ first, the matrix of $\widehat{\rho}$ is $\mathrm{diag}(d, -1, \ldots, -1)$. Thus

$$\mathrm{Tr}(\widehat{\rho}^2) = d^2 + (d-1)\cdot 1 = d^2 + d - 1 \quad (\text{independent of } |u\rangle\,).$$

Using $\mathbb{E}[\widehat{\rho}] = \rho$,

$$\begin{aligned}
\mathbb{E}\big[\|\widehat{\rho} - \rho\|_2^2\big] &= \mathbb{E}\big[\,\mathrm{Tr}(\widehat{\rho}^2) - 2\,\mathrm{Tr}(\widehat{\rho}\rho) + \mathrm{Tr}(\rho^2)\big] \\
&= \mathbb{E}[\mathrm{Tr}(\widehat{\rho}^2)] - \mathrm{Tr}(\rho^2)\ \le\ d^2 + d - 1.
\end{aligned}$$

Standard unentangled tomography (averaging n copies)

Run the single-copy procedure independently on n copies of ρ, obtaining $\widehat{\rho}_1, \ldots, \widehat{\rho}_n$, and output $\overline{\rho} := \frac{1}{n}\sum_{k=1}^n \widehat{\rho}_k$. Then

$$\mathbb{E}\big[\|\overline{\rho} - \rho\|_2^2\big] = \frac{1}{n^2}\sum_{k=1}^n \mathbb{E}\big[\|\widehat{\rho}_k - \rho\|_2^2\big] \le \frac{d^2 + d - 1}{n}.$$

Using $\|A\|_1 \le \sqrt{\mathrm{rank}(A)}\,\|A\|_2 \le \sqrt{d}\,\|A\|_2$ gives the trace-distance guarantee

$$\mathbb{E}\big[D_{\mathrm{tr}}(\overline{\rho}, \rho)\big] = \frac{1}{2}\,\mathbb{E}[\|\overline{\rho} - \rho\|_1] \le \frac{1}{2}\sqrt{d}\,\big(\mathbb{E}\|\overline{\rho} - \rho\|_2^2\big)^{1/2} \le \sqrt{\frac{d^3 + d^2 - d}{4n}}.$$

Hence $n = \Theta\big(d^3/\varepsilon^2\big)$ samples suffice to achieve $\mathbb{E}[D_{\mathrm{tr}}(\overline{\rho}, \rho)] \le \varepsilon$. (With independence, standard scalar concentration upgrades this to $n = O\big((d^3 + \log(1/\delta))/\varepsilon^2\big)$ for tail $\le \delta$.)

Acknowledgments

This note follows the organization and notation of John Wright's *Quantum Learning Theory* course notes (UC Berkeley, 2024) while adding a self-contained derivation of the Haar-POVM inversion.

References

[1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information.* Cambridge University Press, 10th anniversary ed., 2010.

[2] J. Watrous, *The Theory of Quantum Information.* Cambridge University Press, 2018.

[3] J. Wright, "Quantum learning theory (course notes)." University of California, Berkeley, 2024. Lecture notes, accessed for structure/notation.