CHAPTER 5

Sample-Optimal Algorithm for State Tomography

In the previous lecture we saw the most basic algorithms for learning quantum states. While their analysis was simple, the algorithm incurs a dependence on the Hilbert space dimension d which is far from optimal. Indeed, the total number of parameters describing an arbitrary mixed state is d^2, so intuitively we expect that the correct sample complexity for state tomography is $\Theta(d^2/\epsilon)$, where ϵ is the target infidelity.

In this lecture, we show using more sophisticated algebraic tools that with a different measurement scheme, one can indeed achieve this optimal sample complexity, up to logarithmic factors. Notably, the algorithm will measure *all copies of the unknown state ρ at once*, rather than simply measuring every qubit of every copy one at a time. So the question becomes: what is the POVM $\{M_\sigma\}$ that one can perform on $\rho^{\otimes N}$ that maximally extracts information about ρ?

1. Some Forced Moves

There are two symmetries we can exploit in the problem. Firstly, there is the trivial permutation symmetry: the "dataset" of copies of ρ being measured is invariant under swapping different copies around. Additionally, the measurement should be agnostic to the eigenbasis of ρ, because we are not making any assumptions about it: if the algorithm achieves some level of statistical efficient for states in some eigenbasis, they had better be equally statistically efficient in any other eigenbasis. Taken together, these two points imply two things about the POVM $\{M_\sigma\}$ we perform:

- **Permutation invariance**: Elements of the POVM should be invariant under conjugation by any permutation operator.
- **Rotation *equi*variance**: If M_σ is the POVM element corresponding to outputting σ, then $M_{U\sigma U^\dagger} = (U^\dagger)^{\otimes N} M_\sigma U^{\otimes N}$ for any $U \in U(d)$.

We will implement such a POVM $\{M_\sigma\}$ in two stages:

(1) Measure $\rho^{\otimes N}$ with a POVM $\{\tilde{M}_\sigma\}$ which is both permutationally and rotationally *in*variant in order to learn the rotationally invariant information about ρ, i.e., its eigenvalues
(2) Apply a suitable rotationally *equi*variant POVM to the post-measurement state to learn the eigenbasis of ρ

There are different ways of implementing the latter step, but the former step is rather generic and useful in other quantum learning tasks as well. It goes by the name of **weak Schur sampling**, and as we will see, its construction is entirely predetermined by the symmetries above.

Remark 61. *We will assume for convenience throughout this lecture that ρ is full rank. This is just for convenience of prose, as we will be talking about representations of the general linear group, but the reasoning below can be extended to states of degenerate rank by taking appropriate limits.*

2. Representation Theory Toolkit

The key ingredient behind weak Schur sampling is **Schur-Weyl duality**, a fundamental algebraic result that, very roughly speaking, ensures that there is a unitary U_{schur} for which

$$U^{\dagger}_{\mathsf{schur}}\rho^{\otimes N}U_{\mathsf{schur}}$$

is block-diagonal with very particular structure in each diagonal block. As such, we may assume without loss of generality that our algorithm first performs a projective measurement to project to the subspace corresponding to one of these diagonal blocks. Weak Schur sampling is precisely this initial projective measurement, which we spell out in detail below.

2.1. Basic Notions

Let $\mathcal{H} \triangleq (\mathbb{C}^d)^{\otimes N}$. Here we introduce just enough representation theory to be able to present Schur-Weyl duality and the full learning algorithm. Representation theory is the study of groups by associating their elements with linear transformations. Throughout this lecture, we work exclusively with representations over complex vector spaces. Given a vector space V, let $\mathrm{GL}(V)$ denote the group of invertible linear transformations on V.

Definition 62. *Given a group G, a (finite-dimensional)* ***representation*** *is given by a vector space V and a group homomorphism $\mu : G \to \mathrm{GL}(V)$, i.e., a map satisfying $\mu(gh) = \mu(g)\mu(h)$ for all $g, h \in G$. We say that μ is a G-*representation *over V, or a* representation over V *if G is clear from context. The* dimension *of the representation, denoted $\dim(\mu)$, is the dimension of V.*

We will often refer to a representation by the vector space on which G acts, with the homomorphism μ being implicitly understood from context. Similarly, we may write $g \cdot v$ to denote $\mu(g)v$.

In any introductory text on representation theory, one can find a laundry list of examples, for instance the trivial representation, the standard representation of GL_d, the regular representation of any finite group, the Fourier character representation of $\mathbb{Z}_2^n$, etc. The following two representations are most relevant to this lecture:

Example 63. *$\mathcal{S}_N$ and GL_d admit the following representations over $(\mathbb{C}^d)^{\otimes N}$, call them $P : \mathcal{S}_N \to \mathrm{GL}((\mathbb{C}^d)^{\otimes N})$ and $Q : \mathrm{GL}_d \to \mathrm{GL}((\mathbb{C}^d)^{\otimes N})$ respectively. $P(\pi)$ is the permutation operator on N qudits associated to $\pi \in \mathcal{S}_N$, and $Q(M) = M^{\otimes d}$ for $M \in \mathrm{GL}_d$.*

Note that these clearly commute with each other, so we can also define a representation of $\mathcal{S}_N \times \mathrm{GL}_d$, which we will denote by μ_{SW}, over $\mathcal{H}$ via

$$\mu_{\mathsf{SW}}(\pi, M) = P(\pi)Q(M)\,.$$

Definition 64. *A representation (μ, V) is* ***irreducible*** *if there does not exist an* invariant subspace*, i.e. a proper subspace $U \subsetneq V$ for which $\mu(g)U = U$ for all $g \in G$. We sometimes call (μ, V) an* ***irrep*** *for short.*

Like the vector spaces over which they act, representations can be stitched together using standard operations like direct sum and tensor product.

Definition 65. *The* direct sum *of representations $(\mu_1, V_1), \ldots, (\mu_m, V_m)$ of G with multiplicities $a_1, \ldots, a_m$ is the representation $(\mu, V_1^{\oplus a_1} \oplus \cdots \oplus V_m^{\oplus a_m})$ for which $\mu(g)$ is given by the block diagonal matrix whose diagonal blocks consist of a_i copies of $\mu_i(g)$ for all i. We will sometimes write $\mu = \bigotimes_{i=1}^m a_i \cdot \mu_i$ and $V \cong \sum_{i=1}^m a_i \cdot V_i$.*

The tensor product *of representations (μ, V) and (ν, W) of G is the representation $(\mu \otimes \nu, V \otimes W)$ for which*

$$\mu \otimes \nu(g \cdot h) = \mu(g) \otimes \nu(h)\,.$$

Somewhat confusingly, representations which are reducible need not decompose into a direct sum of irreps in general. Those that do are said to be **semisimple**. While not all representations are semisimple, all of the ones we will care about in this lecture are. For instance, it can be shown that any representation of $\mathcal{S}_N$ or U_d is semisimple – the former is **Maschke's theorem** (see e.g. [**FH13**, Corollary 1.6]), and the latter is immediate from the fact that if U_d preserves some subspace W, then it also preserves the orthogonal complement $W^\perp$. For the rest of this lecture, we will work with the implicit understanding that all representations discussed are semisimple.

Definition 66 (Hom spaces)**.** *Given two G-representations μ, ν over spaces V, W, a linear map $f : V \to W$ is a* ***G-linear map*** *if it commutes with the action of G, that is, if $\nu(g)f(v) = f(\mu(g)v)$ for all $v \in V$. We denote by $\mathrm{Hom}_G(V, W)$ the space of all G-linear maps $V \to W$. We say that representations μ and ν are* ***isomorphic****, which we denote by $V \cong W$ when μ and ν are clear from context, if there is a G-linear map $f : V \to W$ which is an isomorphism of vector spaces.*

The following is an elementary but extremely useful fact about G-linear maps:

Lemma 67 (Schur's lemma)**.** *Let V and W be irreps of G.*

(1) If $V \ncong W$, then $\mathrm{Hom}_G(V, W)$ consists of only the zero map.
(2) If $V = W$, then $\mathrm{Hom}_G(V, W)$ consists of all scalar multiples of the identity map.

Proof. For the first part, suppose to the contrary that there is a nonzero G-linear map $f : V \to W$. If it has a nontrivial kernel $V' \subsetneq W$, then note that $f(g \cdot v') = g \cdot f(v') = 0$, so the kernel is stable under the action of G, contradicting the assumption that V is an irrep. So f is injective. An identical argument for the image of f shows that f is surjective, completing the proof of the first part.

For the second part, we need to prove that apart from scalar multiples of the identity, there are no other elements in $\mathrm{Hom}_G(V, V)$. Let $f \in \mathrm{Hom}_G(V, V)$, and consider $f' \triangleq f - \lambda \mathrm{Id}_V$ for any eigenvalue λ of f. As the sum of G-linear maps is G-linear, f' is G-linear, and furthermore it has nontrivial kernel as it vanishes on the eigenvector of f associated to λ. But its kernel cannot be a proper subspace of V or, as in the proof of the first part, this would contradict the fact that V is an irrep. So the kernel of f' must be all of V, implying that $f' \equiv 0$ and thus that $f = \lambda \mathrm{Id}_V$ as claimed. ☐

Finally, an important object in the study of representations is their associated *characters*, which basic information about the representation like their dimension.

Definition 68 (Characters)**.** *Given a G-representation over V, its* ***character*** *is the map $\chi : G \to \mathbb{C}$ given by $\chi(g) = \mathrm{tr}(\mu(g))$. Note that if g is the identity element* id*, then $\chi(\mathrm{id}) = \dim(V)$.*

2.2. Representation Theory of the Symmetric Group

Here we give a complete classification of all irreps of the symmetric group.

We begin with a useful shift in perspective. An equivalent way to think about representations of finite groups G like the symmetric group is in terms of *modules* over the *group algebra* associated to G.

Definition 69 (Group algebras)**.** *Given a finite group G, the associated* ***group algebra*** $\mathbb{C}[G]$ *is the vector space over $\mathbb{C}$ of formal linear combinations $\sum_{g\in G} a_g g$ for $a_g \in \mathbb{C}$, additionally equipped with the* multiplication *operation $(\sum_{g\in G} a_g g)(\sum_{h\in G} b_h h) = \sum_{g,h\in G} a_g b_h \cdot gh$.*

Any G-representation μ over V naturally gives rise to an algebra homomorphism $\mathbb{C}[G] \to \mathrm{GL}(V)$ by extending linearly; the latter equips the vector space V with the structure of a $\mathbb{C}[G]$-module*, and sub-representations of μ then correspond to* sub-modules *of V. This identification also goes in the reverse direction: any such algebra homomorphism gives rise to a G-representation by restricting to the elements $G \subset \mathbb{C}[G]$.*

We now define the central objects in the representation theory of the symmetric group.

Definition 70 (Young tableaux)**.** *A* ***Young diagram*** *is a sequence of rows of boxes of nonincreasing length, like the following:*

Its shape *is the tuple of row lengths; for example, the shape of the above is $(3, 3, 1)$.*

A ***Young tableau*** *T with entries in $[d]$ is a Young diagram where each box is labeled with a number from $[d]$, e.g.*

1	5	3
4	2	2
5		

The Young tableau with ***canonical labeling*** *is given by labeling the entries from left to right and top to bottom with the numbers $1, 2, \cdots$ in increasing order.*

A Young tableau is said to be ***semi-standard*** *if every row consists of entries in non-decreasing order, and if every column consists of entries in strictly increasing order, e.g.,*

1	1	3
2	3	4
5		

It is further said to be ***standard*** *if the rows are also strictly increasing, though we will not use this notion in this lecture.*

Any Young diagram is naturally associated with a ***partition*** *via its shape: for instance* $\lambda = (3, 3, 1)$ *is a partition of* $N = 7$. *We denote such a partition with the notation* $\lambda \vdash [N]$. *If the partition/shape consists of* m *entries and we refer to the* j*-th entry of* λ *for* $j > m$, *by default we set* $\lambda_j = 0$.

Given a partition $\lambda = (\lambda_1, \ldots, \lambda_m) \vdash [N]$, *we can associate a probability distribution over* $[m]$ *which places mass* λ_i/N *on element* i. *We will refer to the vector of probabilities* $(\lambda_1/N, \ldots, \lambda_m/N)$ *by* $\overline{\lambda}$.

Definition 71 (Young symmetrizer)**.** *Let* T *be a Young tableau. Define* $P_T \subseteq \mathcal{S}_N$ *(resp.* $Q_T \subseteq \mathcal{S}_N$*) to consist of permutations which preserve the rows (resp. columns) of* T. *Define the group algebra elements* $a_T, b_T \in \mathbb{C}[\mathcal{S}_N]$ *by* $a_T \triangleq \sum_{p \in P_T} p$ *and* $b_T \triangleq \sum_{q \in Q_T} \operatorname{sgn}(q) q$, *and define the corresponding* ***Young symmetrizer*** *to be* $c_T \triangleq a_T b_T$.

If T *is the canonical labeling for shape* λ, *we use* $a_\lambda, b_\lambda, c_\lambda$ *to denote* a_T, b_T, c_T.

The Young symmetrizer is ultimately just some cleverly chosen linear combination of permutation operators, but it will play a central role not just in the classification of the irreps of the symmetric group, but also in our characterization of the representation μ_{SW} from Example 63.

Definition 72 (Specht module)**.** *Define the* ***Specht module***

$$V_\lambda \triangleq \mathbb{C}[\mathcal{S}_N] c_\lambda \,.$$

In this section, our goal is to show that the irreps of $\mathcal{S}_N$ are precisely the Specht modules for different $\lambda \vdash [N]$. First, let us get some intuition for what they look like:

Example 73. *Consider the case of* $N = 3$, *for which there are three possible partitions:* $(3), (2, 1), (1, 1, 1)$. *When the partition is* (3), *the corresponding Young tableau with canonical labeling is*

$$\begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline \end{array}$$

and a_T *is a sum over all permutations in* $\mathcal{S}_3$, *while* b_T *is the identity. The Young symmetrizer in this case is the sum over all permutations, and the Specht module is simply the line spanned by* $\sum_\pi \pi$. *One can see that this is simply the 1-dimensional* trivial representation *that maps all group elements to the identity.*

When the partition is $(1, 1, 1)$, *the corresponding Young tableau with canonical labeling*

$$\begin{array}{|c|}\hline 1 \\ \hline 2 \\ \hline 3 \\ \hline \end{array}$$

and the Young symmetrizer is given by $\sum_\pi \operatorname{sgn}(\pi)\pi$. *The Specht module is again just the line spanned by this element, and one can see that this is the 1-dimensional* sign representation *that maps all group elements to their sign.*

When the partition is $(2, 1)$, *the corresponding Young diagram is*

$$\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & \multicolumn{1}{c}{} \\ \cline{1-1} \end{array}$$

and a_T is the sum of the identity permutation id *and τ_{12}, the transposition of elements 1 and 2, while b_T is the difference between the identity and τ_{13}, the transposition of elements 1 and 3. The Young symmetrizer is given by*

$$c_\lambda = (\mathrm{id} + \tau_{12})(\mathrm{id} - \tau_{13}) = \mathrm{id} + \tau_{12} - \tau_{13} - \pi\,,$$

where π is the permutation $(1 \mapsto 2, 2 \mapsto 3, 3 \mapsto 1)$. A calculation shows that the Specht module in this case is two-dimensional, spanned by c_λ and $\tau_{13}c_\lambda$, and is in fact isomorphic to the following representation, sometimes called the standard representation. *Consider the two-dimensional subspace of $\mathbb{R}^3$ given by vectors with coordinates summing to zero. There is a natural action of $\mathcal{S}_3$ on this space, i.e., permuting the coordinates of such a vector keeps it in that subspace.*

One would be hard pressed to come up with other irreps of $\mathcal{S}_3$, and indeed there are none: as we will show, the Specht modules make up all irreps of $\mathcal{S}_3$ up to isomorphism!

Lemma 74. *Let T be a Young tableau. For any $g \in \mathbb{C}[\mathcal{S}_N]$, $a_T g b_T$ is a multiple of c_T. In particular, $c_T\mathbb{C}[\mathcal{S}_N]c_T \subseteq \mathbb{C}c_T$.*

Proof. It suffices to show this for $g = \pi$ for $\pi \in \mathcal{S}_N$. We have

$$a_T \pi b_T = \sum_{p \in P_T, q \in Q_T} \mathrm{sgn}(q) p \pi q\,.$$

If $\pi = pq$ for $p \in P_T, q \in Q_T$, then $a_T \pi b_T = \mathrm{sgn}(q) c_T$ as desired.

We will show that for $\pi \notin P_T Q_T$, $a_T \pi b_T = 0$. To show this, we show there is a *transposition* $p' \in P_T$ for which $q' \triangleq \pi^{-1} p' \pi \in Q_T$. In this case, $\pi = p'\pi q'$, so

$$a_T \pi b_T = \sum_{p \in P_T, q \in Q_T} \mathrm{sgn}(q) p p' \pi q' q = - \sum_{p \in P_T, q \in Q_T} \mathrm{sgn}(q) p \pi q = -a_T \pi b_T = 0\,,$$

and we would be done (note that in the penultimate step we used the fact that $\mathrm{sgn}(qq') = \mathrm{sgn}(q)\mathrm{sgn}(q') = \mathrm{sgn}(q)\mathrm{sgn}(p') = -\mathrm{sgn}(q)$ as p' is a transposition).

To show the existence of the transposition p', define T' by mapping every entry x in T to $\pi(x)$. We want to show there are two numbers x, y which appear in the same row of T and the same column of T'. If to the contrary such x, y do not exist, then we can shuffle the entries in each column of T' so that the first row of T' is the same as the first row of T up to permutation. Then we can proceed to shuffle the entries in each column of T', keeping the first row fixed, so that the second row agrees with that of T up to permutation. Continuing in this fashion, we end up with permutations $p \in P_T$ and $q' \in Q_{T'} = \pi Q_T \pi^{-1}$ for which $p \cdot T = q' \cdot T'$. Writing $q' = \pi q \pi^{-1}$ for $q \in Q_T$, we find that $p \cdot T = \pi q \pi^{-1} T' = \pi q \cdot T$, so $\pi = pq^{-1} \in P_T Q_T$, a contradiction. □

Next, we re-use the proof strategy in the last two paragraphs of the proof of Lemma 74 to show the following:

Lemma 75. *Let λ, μ be partitions for which $\lambda > \mu$ in lexicographic ordering. Then $a_\lambda g b_\mu = 0$ for all $g \in \mathbb{C}[\mathcal{S}_N]$. In particular, $c_\lambda \mathbb{C}[\mathcal{S}_N] c_\mu = 0$.*

Proof. As before, it suffices to show this for $g = \pi$. Let T and $\tilde{T}$ denote the tableaux corresponding to λ, μ. By the reasoning in Lemma 74, letting T' denote the result of mapping every entry x in $\tilde{T}$ to $\pi(x)$. As in the proof of Lemma 74, we want to show there are two numbers x, y which appear in the same row of T and the same column of T'. Suppose to the contrary.

If $\lambda_1 > \mu_1$, then because there are strictly fewer than λ_1 columns in T', some column must contain two numbers from the first row of T, a contradiction. If $\lambda_1 = \mu_1$, then as in the proof of Lemma 74, we can shuffle the entries in the columns of T' and in the first row of T so that T and T' have identical first row, and we can recurse on the subsequent rows until we reach a row i for which $\lambda_i > \mu_i$, inducing the desired contradiction. $\square$

This yields the following characterization for the irreps of $\mathcal{S}_N$.

Lemma 76. *(i) V_λ is an irrep of $\mathcal{S}_N$ for any $\lambda \vdash [N]$.*
(ii) For any distinct partitions $\lambda, \mu \vdash [N]$, V_λ and V_μ are not isomorphic.
(iii) Any irrep of $\mathcal{S}_N$ is isomorphic to some V_λ.

Proof. See pset 2 for the proofs of (i) and (ii).

The proof of (iii) follows from the standard fact that the number of irreps of any finite group is equal to the number of conjugacy classes (see e.g. [FH13, Proposition 2.30]), and the fact that the conjugacy classes of $\mathcal{S}_N$ are in one-to-one correspondence with the partitions $\lambda \vdash [N]$. $\square$

Lemma 77 (Hook length formula)**.** *Given an entry in a Young tableau T of shape λ, let its* ***hook length*** *denote the number of boxes either directly below it or directly to its right. Then $\dim(V_\lambda)$ is $n!$ divided by the product of the hook lengths of all entries T. In particular,*

$$\dim(V_\lambda) \le e^{NH(\overline{\lambda})}.$$

Proof. The first part is a standard fact whose proof would take us too far afield. Recall from Lemma 74 that c_T is idempotent up to scaling, that is, $c_T^2 = n_T c_T$ for some constant n_T. The hook length formula is equivalent to the claim that n_T is given by the product of hook lengths. A proof of this can be found, e.g., in [Gri25, Section 5.11.2].

As for the inequality in the second part of the lemma, note that the hook length of the entry in the i-th row and j-th column is at least $\lambda_i - j + 1$, so the product of hook lengths in row i is at least $\lambda_i!$. We thus have

$$\dim(V_\lambda) \le \frac{N!}{\prod_i \lambda_i!} \le e^{NH(\overline{\lambda})},$$

where in the last step we used the elementary fact that the logarithm of a multinomial coefficient $\binom{N}{\lambda_1 \cdots \lambda_m}$ is upper bounded by N times the entropy of the distribution with probability mass function given by $\overline{\lambda}$, see e.g. [CS⁺04, Lemma 2.2]. $\square$

In short, the irreps of the symmetric group are in one-to-one correspondence with the Young diagrams $\lambda \vdash [N]$. Remarkably, this is also the case in a certain sense for $\mathrm{GL}(\mathbb{C}^d)$, and furthermore these irreps of $\mathrm{GL}(\mathbb{C}^d)$ are intimately tied to those of a $\mathcal{S}_N$. This is the content of Schur-Weyl duality, which we discuss in the next section.

2.3. Schur-Weyl Duality

Recall the representation μ_{SW} defined in Example 63. The following decomposition result is the key ingredient behind the learning algorithm we will describe in the next section.

Lemma 78 (Isotypic decomposition)**.** *The representation μ_{SW} of $\mathcal{S}_N \times \mathrm{GL}_d$ over $\mathcal{H}$ decomposes as*

$$\mathcal{H} \cong \bigoplus_{\lambda \vdash [N]} V_\lambda \otimes \mathbb{S}_\lambda V_\lambda \,,$$

*where the **Schur functor** $\mathbb{S}_\lambda V_\lambda \triangleq \mathrm{Hom}_{\mathcal{S}_N}(V_\lambda, \mathcal{H})$ is equipped with the natural GL_d-action of composition.*

Proof. The proof is essentially just symbol pushing, but the intuition is that every irrep V_λ appears with a certain multiplicity in the decomposition of μ_{SW} as a $\mathcal{S}_N$-representation, and the spaces $\mathbb{S}_\lambda V_\lambda$ are simply there to track these multiplicities. This argument is not specific to $\mathcal{S}_N, \mathrm{GL}_d$ and only uses the fact that their actions commute.

First note that $\mathbb{S}_\lambda V_\lambda$ is a valid representation of GL_d: given $f \in \mathrm{Hom}_{\mathcal{S}_N}(V_\lambda, \mathcal{H})$ and $g \in \mathrm{GL}_d$, we have $g \cdot f = g \circ f \in \mathrm{Hom}_{\mathcal{S}_N}(V_\lambda, \mathcal{H})$ as

$$(g \circ f)(\pi \cdot v) = g(\pi \cdot f(v)) = \pi \cdot (g \circ f)(v)$$

for any $v \in V_\lambda, \pi \in \mathcal{S}_N$, where the first step follows by the fact that f is $\mathcal{S}_N$-linear, and the second step follows by the fact that the actions of $\mathcal{S}_N$ and GL_d on $\mathcal{H}$ commute.

Now consider the map $\Phi : \oplus_\lambda V_\lambda \otimes \mathbb{S}_\lambda V_\lambda \to \mathcal{H}$ given by linearly extending the maps $v \otimes f \mapsto f(v)$ for $v \in V_\lambda, f \in \mathbb{S}_\lambda V_\lambda$. One can readily check that this map is $\mathcal{S}_N \otimes \mathrm{GL}_d$-linear. It remains to show it is bijective.

To show surjectivity: if we consider the decomposition of $\mathcal{S}_N$-representations $\mathcal{H} \cong \oplus_\lambda V_\lambda^{\oplus m_\lambda}$, then by Schur's lemma, $\mathrm{Hom}_{\mathcal{S}_N}(V_\lambda, \mathcal{H})$ is spanned by the embeddings $\iota_{\lambda_1}, \ldots, \iota_{\lambda_{m_\lambda}} : V_\lambda \to \mathcal{H}$ of V_λ into the m_λ copies of V_λ within $\mathcal{H}$. Now observe that $\Phi(V_\lambda \otimes \{\iota_{\lambda_j}\}) = \iota_{\lambda_j}(V_\lambda) = V_\lambda$, so the image of Φ contains every component in the decomposition $\mathcal{H} \cong \oplus_\lambda V_\lambda^{\oplus m_\lambda}$, establishing surjectivity.

Finally, Φ is bijective because the domain and image have the same dimension $\sum_\lambda m_\lambda \mathrm{dim}(V_\lambda)$. □

The following gives a simple characterization of the Schur functors in terms of Young symmetrizers, using the fact that Young symmetrizers are idempotent up to a scaling.

Lemma 79. $\mathbb{S}_\lambda V_\lambda \cong c_\lambda \cdot \mathcal{H}$

Proof. See pset 2. □

Remark 80 (Schur-Weyl duality)**.** *A beautiful fact is that these GL_d-representations $\mathbb{S}_\lambda(V_\lambda)$ are actually* irreps*; this is the content of the famous **Schur-Weyl duality**. This is a consequence of two facts that we will not prove:*

- *Not only do the actions of $\mathcal{S}_N$ and GL_d on $\mathcal{H}$ commute with each other, but any operator that commutes with $Q(M)$ for all $M \in \mathrm{GL}_d$ must be a linear combination of the operators $\{P(\pi)\}_{\pi \in \mathcal{S}_N}$.*
- *The **double commutant theorem**: in situations where one has an isotypic decomposition like in Lemma 78, this states that one action being the* commutant *of the other in the sense of the above bullet point allows one to further deduce that the Schur functors in the isotypic decomposition are irreps.*

The first fact has a remarkably elementary proof, see `https://math.univ-lyon1.fr/~aubrun/recherche/schur-weyl.pdf`. *The second fact can be found in any standard representation theory text, e.g.* [**EGH+09**, Theorem 5.18.1].