

# Character Theory, Schur Polynomials, and Symmetric Functions

Yuan-Pon Chen, Michael Hu, Abhinav Angirekula

UIUC MATH 595 Final Presentation

# Outline

- 1 Introduction
- 2 Character Theory Refresher, Schur Polynomials, and RSK
- 3 Application of RSK in QIT
- 4 References

# Character Theory: The Bridge to Symmetric Functions

- Characters capture essential info about a representation without needing the full matrices. They give rise to many foundational results in representation theory
- **Goal:** Use character theory to create a bridge between representation theory and algebraic combinatorics
- In particular, we introduce a special family of symmetric functions called the Schur polynomials
  - ▶ **Schur functions**  $s_\lambda$  correspond to irreducible characters of  $GL_n$ .
  - ▶ They are naturally indexed by Young diagrams and are super combinatorial.
- We introduce the **RSK correspondence**, a famous bijection in combinatorics:
- Walk backwards across our to compute representation-theoretic quantities efficiently with application in quantum information tasks.

# Outline

- 1 Introduction
- 2 Character Theory Refresher, Schur Polynomials, and RSK
- 3 Application of RSK in QIT
- 4 References

# Outline

- 1 Introduction
- 2 Character Theory Refresher, Schur Polynomials, and RSK**
- 3 Application of RSK in QIT
- 4 References

# Character Theory Review

**Definition:** Let  $G$  be a finite group and  $\rho : G \rightarrow \text{GL}(V)$  a representation of  $G$  over  $\mathbb{C}$ . The *character* of  $\rho$  is the function

$$\chi_\rho : G \rightarrow \mathbb{C}, \quad \chi_\rho(g) = \text{Tr}(\rho(g)),$$

where  $\text{Tr}$  denotes the trace.

## Immediate Properties

- $\chi_\rho(e) = \dim(V)$
- Characters are *class functions*:  $\chi_\rho(hgh^{-1}) = \chi_\rho(g)$  for all  $g, h \in G$ .
- Equivalent representations have the same character.

# Character Theory Review

**Definition:** Let  $G$  be a finite group and  $\chi, \psi$  be characters of  $G$ . Define the *inner product*:

$$\langle \chi, \psi \rangle = \frac{1}{|G|} \sum_{g \in G} \chi(g) \overline{\psi(g)}.$$

**Interpretation:** Inner products measure "overlap" of representations.

**Theorem (Character Relations of the First Kind):** If  $\chi$  and  $\psi$  are irreducible characters of  $G$ , then

$$\langle \chi, \psi \rangle = \delta_{\chi, \psi}$$

## Character Theory Review

**Corollary:** Let  $V$  be a representation of  $G$  with character  $\chi$ . Suppose

$$V = m_1 V^{(1)} \oplus m_2 V^{(2)} \oplus \cdots \oplus m_k V^{(k)},$$

where the  $V^{(i)}$  are pairwise inequivalent irreducible representations with characters  $\chi^{(i)}$ . Then:

- 1  $\chi = m_1 \chi^{(1)} + m_2 \chi^{(2)} + \cdots + m_k \chi^{(k)}$ .
- 2  $(\chi, \chi^{(j)}) = m_j$  for all  $j$ .
- 3  $(\chi, \chi) = m_1^2 + m_2^2 + \cdots + m_k^2$ .
- 4  $X$  is irreducible if and only if  $(\chi, \chi) = 1$ .
- 5 Let  $W$  be another representation of  $G$  with character  $\psi$ . Then

$$V \cong W \quad \text{if and only if} \quad \chi(g) = \psi(g) \text{ for all } g \in G.$$

# Character Theory Review

Let  $G$  be a finite group with group algebra  $\mathbb{C}[G]$ . By Maschke's theorem,

$$\mathbb{C}[G] = m_1 V^{(1)} \oplus m_2 V^{(2)} \oplus \cdots \oplus m_k V^{(k)}$$

**Proposition:** Let  $G$  be as above:

- 1  $m_i = \dim(V^{(i)})$
- 2  $\sum_i \dim(V^{(i)})^2 = |G|$
- 3 The number of  $V^{(i)}$  equals the number of conjugacy classes of  $G$

**Proposition:** The irreducible characters form an orthonormal basis for the space of class functions of  $G$ .

# Irreducible Representations of the Symmetric Group

For  $\lambda \vdash n$ , the Specht module  $S^\lambda := \mathbb{C}[S_n]e_T$  where  $e_T$  is the young symmetrizer for some tableau  $T$  of shape  $\lambda$ .

**Theorem:** The set of  $S^\lambda$  for  $\lambda \vdash n$  form a full set of irreducibles of  $S_n$  representations of  $\mathbb{C}$ .

*Proof Sketch.*

- $S^\lambda$  are irreducible (Submodule Theorem)
- Distinct partitions  $\lambda \neq \mu$  give non-isomorphic Specht modules (Majorization Lemma)
- Part 3 of previous proposition (or part 2 with dimension counting)

## $U_d$ Representations and Schur Polynomials

Recall Schur-Weyl duality

$$(\mathbb{C}^d)^{\otimes n} = \bigoplus_{\lambda \vdash dn} V_\lambda \otimes U_\lambda^d$$

Applying  $e_T$  gives us projection onto  $U_\lambda^d$ . Furthermore,  $U_\lambda^d$  has a standard basis given by semistandard young tableaux of shape  $\lambda$ .

If  $g \in \mathcal{U}_d$  is the diagonal matrix  $\text{diag}(x_1, \dots, x_d)$ , then the action on  $T \in \text{SSYT}(\lambda)$  is:

$$x^T := x_1^{\#1's} \cdots x_d^{\#d's}$$

**Definition.** For a partition  $\lambda$  with  $\ell(\lambda) \leq n$ , the *Schur polynomial* is defined:

$$s_\lambda(x_1, \dots, x_d) := \sum_{T \in \text{SSYT}(\lambda)} x^T$$

Notice that if  $g \in \mathcal{U}_d$  has eigenvalues  $x_1, \dots, x_d$ , then

$$\chi_{U_\lambda^d}(g) = s_\lambda(x_1, \dots, x_d)$$

## Example of a Schur Polynomial

For  $\lambda = (2, 1)$ , some of the tableaux with fillings 1, 2, 3 are shown below

1	1	1	2	1	1	1	2	1	3	...
2		2		3		3		2		

Hence

$$s_{\lambda}(x_1, x_2, x_3) = x_1^2 x_2 + x_1 x_2^2 + x_1^2 x_3 + 2x_1 x_2 x_3 + \dots$$

## More about Schur Polynomials

**Basis of symmetric functions:** With a clever argument (Bender-Knuth Involution), one can see that the Schur polynomials are symmetric functions. In fact,  $\{s_\lambda\}_{\lambda \vdash n}$  forms an important basis of the ring of symmetric functions, and many identities (Cauchy, Pieri, Littlewood–Richardson) are naturally expressed in this basis.

$$h_k(x)s_\lambda(x) = \sum_{\mu} s_\mu \quad s_{\lambda^{(1)}} \cdots s_{\lambda^{(r)}} = \sum_{\nu} c_{\lambda^{(1)} \dots \lambda^{(r)}}^{\nu} s_{\nu}$$

$$\prod_{i=1}^n \prod_{j=1}^m \frac{1}{1 - x_i y_j} = \sum_{\lambda} s_{\lambda}(x_1, \dots, x_n) s_{\lambda}(y_1, \dots, y_m).$$

The Robinson–Schensted–Knuth correspondence is a famous bijection in algebraic combinatorics that proves the Cauchy identity. It is a generalization of Schensted-Insertion. It gives a combinatorial rule for turning  $m \times n$  matrices with nonnegative entries into pairs of SSYT of the same shape with fillings  $\{1, \dots, n\}$  and  $\{1, \dots, m\}$ , and back.

$$\left\{ \begin{array}{l} \text{nonnegative} \\ m \times n \text{ matrices} \end{array} \right\} \longleftrightarrow \left\{ \begin{array}{l} \text{pairs of words} \\ \text{of same length} \end{array} \right\} \longleftrightarrow \left\{ \begin{array}{l} \text{pairs of SSYT of same shape} \\ \text{fillings } \{1, \dots, n\}, \{1, \dots, m\} \end{array} \right\}$$

## Pairs of Words to Matrices and Back

( $\implies$ ) Given a pair of words  $w^A$  and  $w^B$  with entries in  $\{1, \dots, n\}$  and  $\{1, \dots, m\}$  respectively with the sorting condition so that 1)  $w^A$  is sorted and 2)  $w_i^A = w_j^B$  implies  $w_i^B \leq w_j^B$ , we assign the matrix such that  $M_{ij} = \#$  of occurrences of  $\binom{i}{j}$

( $\impliedby$ ) Given a nonnegative  $m \times n$  matrix, produce  $w^A$  by including the sum of the  $i$ th number of  $i$ 's, and produce  $w^B$  so that  $M_{ij}$  determines the number of occurrences of  $j$  below  $i$ . For example:

$$\begin{pmatrix} 1 & 1 & 1 & 2 & 2 & 3 & 3 & 3 & 3 \\ 1 & 2 & 2 & 1 & 2 & 1 & 1 & 1 & 2 \end{pmatrix} \longleftrightarrow \begin{pmatrix} 1 & 2 \\ 1 & 1 \\ 3 & 1 \end{pmatrix}$$

## Pairs of Words to Pairs of Tableaux

**Row Bumping Insertion** To insert an entry  $i$  into a SSYT  $T$ :

- Find the leftmost box in the first row with filling greater than  $i$ . If no box exists, insert  $i$  as a box at the end of the first row
- Otherwise, this box has filling  $j > i$ .  $i$  now "bumps"  $j$  by usurping its box.
- Repeat row insertion with  $j$  in the next row down.

**Pairs of Words to Tableaux:** Do row insertion with entries in the bottom word. At each step, have a second "recording" tableaux which inserts the top word's entry in the new box from each insertion.

# RSK Example

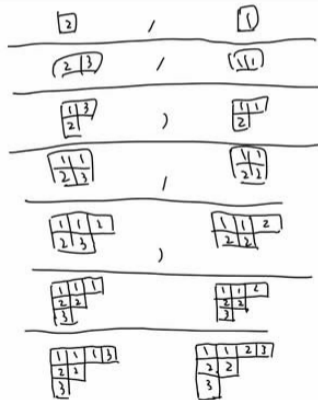
$$\begin{pmatrix} 0 & 1 & 1 \\ 2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

$\#(i) = M_{i,j}$  read rows  
 $\xrightarrow{\hspace{2cm}}$   
 $\xleftarrow{\hspace{2cm}}$   
 $M_{i,j} = \# \text{ of } (i,j)$

$$\begin{pmatrix} 1 & 1 & 2 & 2 & 2 & 3 & 3 \\ 2 & 3 & 1 & 1 & 2 & 1 & 3 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 & 2 & 2 & 2 & 3 & 3 \\ 2 & 3 & 1 & 1 & 2 & 1 & 3 \end{pmatrix}$$

row insert record  
 bottom  $\rightarrow$  top  
 $\xleftarrow{\hspace{2cm}}$   
 unwind record, unbump



# Outline

- 1 Introduction
- 2 Character Theory Refresher, Schur Polynomials, and RSK
- 3 Application of RSK in QIT**
- 4 References

## Recall: Weak Schur Sampling

### Weak Schur Sampling

Consider the quantum state  $\rho^{\otimes n}$ , where  $\rho$  is a  $d$ -dimensional qudit state with sorted spectrum  $\rho = (\rho_1, \dots, \rho_d)$ . From Schur-Weyl duality we have

$$(\mathbb{C}^d)^{\otimes n} \simeq \bigoplus_{\lambda \vdash_d n} V_\lambda \otimes U_\lambda^d.$$

For each  $\lambda \vdash_d n$  define  $P_\lambda$  to be the projector onto  $V_\lambda \otimes U_\lambda^d$ . The **weak Schur sampling** refers to the projective measurement

$$\{P_\lambda\}_{\lambda \vdash_d n}.$$

We obtain the measurement **outcome**  $\lambda$  with **probability**  $\text{tr}[P_\lambda \rho^{\otimes n}]$ .

# Weak Schur Sampling and the RSK Algorithm

- Recall: In class we **upper-bounded** the quantity  $\text{tr}[P_\lambda \rho^{\otimes n}]$  using information-theoretic arguments.
- It turns out that  $\text{tr}[P_\lambda \rho^{\otimes n}]$  describes **the same distribution** obtained by applying the **RSK algorithm** to a **random word**.
- The **combinatorial properties** of the RSK algorithm can give more insight of  $\text{tr}[P_\lambda \rho^{\otimes n}]$ .

# Weak Schur Sampling Probability Formula

## Property 1

Let  $\rho$  have (sorted) spectrum  $p = (p_1, \dots, p_d)$ . We have

$$\text{tr}[P_\lambda \rho^{\otimes n}] = \dim(\lambda) s_\lambda(p_1, \dots, p_d),$$

where  $\dim(\lambda)$  denotes the number of standard Young tableaux with shape  $\lambda$ .

- In other words, after applying the **weak Schur sampling** to  $\rho^{\otimes n}$ , the **probability** that we see the **outcome  $\lambda$**  is  $\dim(\lambda) s_\lambda(p_1, \dots, p_d)$ .

Proof sketch:

- The operator  $\rho^{\otimes n}$  is **permutation invariant**. By Schur-Weyl duality,

$$\rho^{\otimes n} = \bigoplus_{\mu \vdash dn} \mathbb{1}_\mu \otimes \omega_\mu(\rho),$$

where  $(\omega_\mu, U_\mu^d)$  are inequivalent irreps of  $\text{GL}(\mathbb{C}^d)$ .

- The projector  $P_\lambda$  kills the other subspaces. That is,  $P_\lambda \rho^{\otimes n} = \mathbb{1}_\lambda \otimes \omega_\lambda(\rho)$ .
- Taking the trace gives  $\text{tr}[P_\lambda \rho^{\otimes n}] = \text{tr}[\mathbb{1}_\lambda] \text{tr}[\omega_\lambda(\rho)] = \dim(\lambda) s_\lambda(p_1, \dots, p_d)$ .
  - ▶ The equality  $\text{tr}[\omega_\lambda(\rho)] = s_\lambda(p_1, \dots, p_d)$  comes from the fact that  $s_\lambda$  is the **character** of the irrep  $(\omega_\lambda, U_\lambda^d)$  of  $\text{GL}(\mathbb{C}^d)$ .

## RSK for a Single Word

- For any word  $w = w_1 \cdots w_n \in \{1, \dots, d\}^n$ , we can apply RSK to map the following pair of words

$$M_w = \begin{pmatrix} 1 & 2 & \cdots & n \\ w_1 & w_2 & \cdots & w_n \end{pmatrix}$$

to a pair of SSYT.

- ▶ The first row is indeed non-decreasing.
  - ▶ Any two elements in the first row are distinct.
- By an abuse of terminology, we say we **apply RSK to  $w$**  when we apply RSK to  $M_w$ .
- The record tableau is a **SSYT** using each of  $1, \dots, n$  **exactly once**, which is exactly a **SYT**.
- We then have the following **bijection**:

$$w \in \{1, \dots, d\}^n \xleftrightarrow{RSK} \bigcup_{\lambda \vdash_d n} SSYT_d(\lambda) \times SYT(\lambda),$$

where  $SSYT_d(\lambda)$  is the set of all SSYTs using  $\{1, \dots, d\}$  of shape  $\lambda$ , and  $SYT(\lambda)$  is the set of all SYTs of shape  $\lambda$ .

- In words, a **length- $n$  word** over the alphabet  $\{1, \dots, d\}$  uniquely determines a **pair** of a  **$n$ -box SSYT** using  $\{1, \dots, d\}$  and a **SYT** of the **same shape**, and vice versa.

## RSK for a Single Word (Conti'd)

- Let  $d = 3$  and identify  $\mathcal{A} = \{a, b, c\}$ , let  $n = 6$ , and consider  $w = w_1 \cdots w_6 = bcabca$ .
- For each  $i$  from 1 to 6, “bump”  $w_i$  to the **first row** and decide its position as follows:
  - ▶ If putting  $w_i$  **at the end of the first row** keeps this first row **sorted** in the alphabetical order, do so.
  - ▶ If not, find the **rightmost position** in the first row such that **replacing** that entry with  $w_i$  keeps the first row **sorted** in the alphabetical order. In this case, we “bump” the **replaced letter** (say  $\ell$ ) to the **second row**.
  - ▶ Determine the position of  $\ell$  in the **second row** in a similar way. **Repeat** this procedure for the third row (if necessary), and so on.
- The construction above yields a **SSYT** of size  $n = 6$  using  $\{a, b, c\}$ .
- During the process, maintain a **SYT** to keep track of how the **shape** of the tableau **evolved**.

# RSK for a Single Word: Example

$$w = (w_1, \dots, w_6) = bcabca.$$

$i$	$w_i$	SSYT	SYT								
1	b	<table border="1"><tr><td>b</td></tr></table>	b	<table border="1"><tr><td>1</td></tr></table>	1						
b											
1											
2	c	<table border="1"><tr><td>b</td><td>c</td></tr></table>	b	c	<table border="1"><tr><td>1</td><td>2</td></tr></table>	1	2				
b	c										
1	2										
3	a	<table border="1"><tr><td>a</td><td>c</td></tr><tr><td>b</td><td></td></tr></table>	a	c	b		<table border="1"><tr><td>1</td><td>2</td></tr><tr><td>3</td><td></td></tr></table>	1	2	3	
a	c										
b											
1	2										
3											

$i$	$w_i$	SSYT	SYT																		
4	b	<table border="1"><tr><td>a</td><td>b</td></tr><tr><td>b</td><td>c</td></tr></table>	a	b	b	c	<table border="1"><tr><td>1</td><td>2</td></tr><tr><td>3</td><td>4</td></tr></table>	1	2	3	4										
a	b																				
b	c																				
1	2																				
3	4																				
5	c	<table border="1"><tr><td>a</td><td>b</td><td>c</td></tr><tr><td>b</td><td>c</td><td></td></tr></table>	a	b	c	b	c		<table border="1"><tr><td>1</td><td>2</td><td>5</td></tr><tr><td>3</td><td>4</td><td></td></tr></table>	1	2	5	3	4							
a	b	c																			
b	c																				
1	2	5																			
3	4																				
6	a	<table border="1"><tr><td>a</td><td>a</td><td>c</td></tr><tr><td>b</td><td>b</td><td></td></tr><tr><td>c</td><td></td><td></td></tr></table>	a	a	c	b	b		c			<table border="1"><tr><td>1</td><td>2</td><td>5</td></tr><tr><td>3</td><td>4</td><td></td></tr><tr><td>6</td><td></td><td></td></tr></table>	1	2	5	3	4		6		
a	a	c																			
b	b																				
c																					
1	2	5																			
3	4																				
6																					

## RSK for a Single Word: Example (Conti'd)

- $\text{RSK}(w_1 \dots w_6) = \left( \begin{array}{|c|c|c|} \hline a & a & c \\ \hline b & b & \\ \hline c & & \\ \hline \end{array}, \begin{array}{|c|c|c|} \hline 1 & 2 & 5 \\ \hline 3 & 4 & \\ \hline 6 & & \\ \hline \end{array} \right).$

- Can we **recover**  $w = w_1 w_2 \dots w_6$ ?
- **Yes**, as what we did for recovering a pair of words after the (original-version) RSK.
- First recover  $w_6$ :
  - ▶ At the last step the third row received a letter  $c$ , which must have been bumped from the second row.
  - ▶ The second row was originally  $\begin{array}{|c|c|} \hline b & c \\ \hline \end{array}$  and received a letter  $b$ , which must have been bumped from the first row.
  - ▶ The first row was originally  $\begin{array}{|c|c|c|} \hline a & b & c \\ \hline \end{array}$  and received a letter  $a$ , which is  $w_6$ .

- We deduce that  $w_6 = a$  and  $\text{RSK}(w_1 \dots w_5) = \left( \begin{array}{|c|c|c|} \hline a & b & c \\ \hline b & c & \\ \hline \end{array}, \begin{array}{|c|c|c|} \hline 1 & 2 & 5 \\ \hline 3 & 4 & \\ \hline \end{array} \right).$

- **Repeating** the argument **recovers the input word**  $w = w_1 w_2 \dots w_6$ .

## RSK applied to a Random Word

- Let  $p = (p_1, \dots, p_d)$  be a probability distribution over  $\mathcal{A} = \{1, \dots, d\}$ .
- Let  $w = w_1 \cdots w_n$  have distribution  $p^{\otimes n}$ .
  - ▶ That is,  $w_1, \dots, w_n$  are **i.i.d.** with a common distribution  $p$ .
- For a given Young diagram  $\lambda \vdash_d n$ , what is the **probability** that  $\text{RSK}(w)$  has **shape**  $\lambda$ ?

## RSK applied to a Random Word: Example

Let  $d = 2$ ,  $\mathcal{A} = \{a, b\}$ , and  $n = 3$ . Write  $p_a, p_b$  for the probability of  $a$  and  $b$ , respectively.

$w$	SSYT	SYT	Probability								
aaa	<table border="1"><tr><td>a</td><td>a</td><td>a</td></tr></table>	a	a	a	<table border="1"><tr><td>1</td><td>2</td><td>3</td></tr></table>	1	2	3	$p_a^3$		
a	a	a									
1	2	3									
aab	<table border="1"><tr><td>a</td><td>a</td><td>b</td></tr></table>	a	a	b	<table border="1"><tr><td>1</td><td>2</td><td>3</td></tr></table>	1	2	3	$p_a^2 p_b$		
a	a	b									
1	2	3									
aba	<table border="1"><tr><td>a</td><td>a</td></tr><tr><td>b</td><td></td></tr></table>	a	a	b		<table border="1"><tr><td>1</td><td>2</td></tr><tr><td>3</td><td></td></tr></table>	1	2	3		$p_a^2 p_b$
a	a										
b											
1	2										
3											
abb	<table border="1"><tr><td>a</td><td>b</td><td>b</td></tr></table>	a	b	b	<table border="1"><tr><td>1</td><td>2</td><td>3</td></tr></table>	1	2	3	$p_a p_b^2$		
a	b	b									
1	2	3									

$w$	SSYT	SYT	Probability								
baa	<table border="1"><tr><td>a</td><td>a</td></tr><tr><td>b</td><td></td></tr></table>	a	a	b		<table border="1"><tr><td>1</td><td>3</td></tr><tr><td>2</td><td></td></tr></table>	1	3	2		$p_a^2 p_b$
a	a										
b											
1	3										
2											
bab	<table border="1"><tr><td>a</td><td>b</td></tr><tr><td>b</td><td></td></tr></table>	a	b	b		<table border="1"><tr><td>1</td><td>3</td></tr><tr><td>2</td><td></td></tr></table>	1	3	2		$p_a p_b^2$
a	b										
b											
1	3										
2											
bba	<table border="1"><tr><td>a</td><td>b</td></tr><tr><td>b</td><td></td></tr></table>	a	b	b		<table border="1"><tr><td>1</td><td>2</td></tr><tr><td>3</td><td></td></tr></table>	1	2	3		$p_a p_b^2$
a	b										
b											
1	2										
3											
bbb	<table border="1"><tr><td>b</td><td>b</td><td>b</td></tr></table>	b	b	b	<table border="1"><tr><td>1</td><td>2</td><td>3</td></tr></table>	1	2	3	$p_b^3$		
b	b	b									
1	2	3									

- Consider  $\lambda = \begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \\ \hline \end{array}$ .
- Each **SSYT** of shape  $\lambda$  appears **twice** (since there are **two SYTs** of shape  $\lambda$ ).
- Probability that the output shape is  $\lambda$  is  $2(p_a^2 p_b + p_a p_b^2)$ .

# RSK for Random Words and Weak Schur Sampling

## Property 2

For  $\rho = (\rho_1, \dots, \rho_d)$ ,  $w = w_1 \cdots w_n$  having distribution  $\rho^{\otimes n}$ , and  $\lambda \vdash_d n$ , the **probability** that the **shape** of  $\text{RSK}(w)$  is  $\lambda$  is  $|\text{SYT}(\lambda)| \sum_{T \in \text{SSYT}_d(\lambda)} \rho^T = \text{dim}(\lambda) s_\lambda(\rho_1, \dots, \rho_d)$ .

- The **same probability** as obtaining the outcome  $\lambda$  after **weak Schur sampling** over  $\rho^{\otimes n}$  with  $\rho$  having **spectrum**  $(\rho_1, \dots, \rho_d)$ .

# Application of RSK in Spectrum Estimation

## Theorem 1 [OW16, Theorem 1.1]

Let  $p = (p_1, \dots, p_d)$  be a **sorted** probability vector (spectrum) and let  $\lambda = (\lambda_1, \dots, \lambda_d)$  be a random Young diagram  $\vdash_d n$  with distribution  $\mathbb{P}(\lambda = \lambda) = \dim(\lambda) s_\lambda(p_1, \dots, p_d)$ . Then,

$$\mathbb{E} [\|\bar{\lambda} - p\|_2^2] \leq \frac{d}{n},$$

where  $\bar{\lambda} := \lambda/n$ ,  $\|x\|_2 := \sqrt{\sum_{i=1}^d x_i^2}$  denotes the  $\ell^2$ -norm of  $x = (x_1, \dots, x_d)$ , and  $\mathbb{E}$  is the expectation with respect to the probability distribution  $\mathbb{P}$ .

- Recall: In class we had for all  $\varepsilon > 0$  that  $\mathbb{P}(d_{\text{TV}}(\bar{\lambda}, p) > \varepsilon) < (n+1)^{\frac{d(d+1)}{2}} \exp(-2n\varepsilon^2)$ .
  - ▶ Use  $O(d^2/\varepsilon^2) \log(d/\varepsilon) \log(1/\delta)$  samples to have  $\mathbb{P}(d_{\text{TV}}(\bar{\lambda}, p) \leq \varepsilon) \geq 1 - \delta$ .

- From Theorem 1 we have [OW16, Corollary 1.3]

$$\mathbb{E} [d_{\text{TV}}(\bar{\lambda}, p)] \leq \frac{\sqrt{d}}{2} \sqrt{\mathbb{E} [\|\bar{\lambda} - p\|_2^2]} \leq \frac{d}{2\sqrt{n}},$$

leading to sample complexity  $O(d^2/\varepsilon^2) \log(1/\delta)$ .

- One step in the proof of Theorem 1 involved some **combinatorial properties of RSK**.

# Proof Sketch of Theorem 1

- 1 Expand by definition  $\mathbb{E} [\|\bar{\lambda} - p\|_2^2] = \mathbb{E} \left[ \sum_{i=1}^d \left( \frac{\lambda_i}{n} - p_i \right)^2 \right]$ .
- 2 Use some results about the character  $\chi_\lambda(\pi)$  (over  $S_n$ ) for  $\pi$  a transposition to get

$$\mathbb{E} \left[ \sum_{i=1}^d \left( \frac{\lambda_i}{n} - p_i \right)^2 \right] = \left( 2 - \frac{1}{n} \right) \sum_{i=1}^d p_i^2 + \frac{1}{n^2} \sum_{i=1}^d (2i - 1 - 2np_i) \mathbb{E} [\lambda_i].$$

- 3 Use the **combinatorial properties of RSK** to show that  $\mathbb{E} [\lambda] = (\mathbb{E} [\lambda_1], \dots, \mathbb{E} [\lambda_d]) \succ np = (np_1, \dots, np_d)$  (proof in the next page).
- 4 Use  $\langle x^\downarrow, u^\downarrow \rangle \leq \langle y^\downarrow, u^\downarrow \rangle$  for any  $u \in \mathbb{R}^d$  and  $x \prec y$  and the fact that  $(2i - 1 - 2np_i)$  is increasing in  $i$  to upper-bound  $\sum_{i=1}^d (2i - 1 - 2np_i) \mathbb{E} [\lambda_i] \leq \sum_{i=1}^d (2i - 1 - 2np_i) np_i$ .
- 5 Simplify.

# A Majorization Result about RSK

## Lemma 1 [OW16, Lemma 3.2]

$$\mathbb{E}[\boldsymbol{\lambda}] = (\mathbb{E}[\lambda_1], \dots, \mathbb{E}[\lambda_d]) \succ np = (np_1, \dots, np_d).$$

- For each word  $w = w_1 \cdots w_n \in \{1, \dots, d\}^n$  and  $i \in [1, d]$ , let  $H_i(w)$  be the number of occurrences of  $i$  in  $w$ . Then define the histogram of  $w$  as  $H(w) := (H_1(w), \dots, H_d(w))$ .
- Let  $\text{RSK}(w)$  have shape  $\lambda$ . For each  $i \in \{1, \dots, d-1\}$  we have  $\sum_{j=1}^i \lambda_j \geq \sum_{j=1}^i H_j(w)$ .

## Example

$w = bcabca$ . The SSYT part of  $\text{RSK}(w)$  is 

$a$	$a$	$c$
$b$	$b$	
$c$		

. The histogram of  $w$  is 

$a$	$a$
$b$	$b$
$c$	$c$

.

- ▶  $\lambda_1 = 3 \geq H_1(w) = 2$  since in the first row of the SSYT all  $a$  must have appeared.
- ▶  $\lambda_1 + \lambda_2 = 5 \geq H_1(w) + H_2(w) = 4$  since in the first two rows all  $a$  and  $b$  must have appeared.

- Taking expectations proves Lemma 1.

# Outline

- 1 Introduction
- 2 Character Theory Refresher, Schur Polynomials, and RSK
- 3 Application of RSK in QIT
- 4 References

## References

- 1 [OW16] O'Donnell, Ryan, and John Wright. "Efficient quantum tomography." Proceedings of the forty-eighth annual ACM symposium on Theory of Computing. 2016.
- 2 [OW17] O'Donnell, Ryan, and John Wright. "A Primer on the Statistics of Longest Increasing Subsequences and Quantum States." <https://people.eecs.berkeley.edu/~jswright/papers/tomography-survey.pdf>.
- 3 [Ful97] Ful97 Fulton, William. *Young Tableaux: With Applications to Representation Theory and Geometry*. Cambridge University Press, 1997.
- 4 [Sag01] Sag01 Sagan, Bruce E. *The Symmetric Group: Representations, Combinatorial Algorithms, and Symmetric Functions*. 2nd edition. Springer, 2001.
- 5 [Mac95] Mac95 Macdonald, I. G. *Symmetric Functions and Hall Polynomials*. 2nd edition. Oxford University Press, 1995. (See also: <http://www.fultonbook.com/symmetricfunctions.html>)