Supplement to Fourier Bases for Solving Permutation Puzzles

Our supplement is organized as follows:

- Section 1 gives a full description of the irreducible representations of the symmetric group
- Section 2 discusses the irreducible representations of wreath product groups
- Section 3 provides additional information on our experiments

1 Irreducible Representations of the Symmetric Group

We present the construction of Young's Orthogonal Representation (YOR). The presentation here largely follows Huang et al. (2009), and Kondor (2008).

1.1 Preliminaries

A compact way of denoting permutations of $\mathbb{S}_n$ is through **cycle notation**. Permutations can be expressed as a product of disjoint cycles. The cycle $(a_1, \dots a_k)$ denotes the permutation that sends $a_1 \mapsto a_2, a_2 \mapsto a_3, \dots, a_k \mapsto a_1$. Without loss of generality, we can omit singleton cycles going forward.

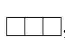
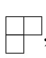
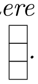
The representations of $\mathbb{S}_n$ are indexed by **partitions** of n, a set of positive integers that sum to n. The tuple of non-negative integers $\lambda = (\lambda_1, \lambda_2 \dots \lambda_k)$ is a partition of n if $\sum_{i=1}^{k} \lambda_i = n$. Conventionally, the parts of a partition are listed in weakly decreasing order: $\lambda_1 \geq \dots \geq \lambda_k$. We use the notation $\lambda \vdash n$ to indicate that λ is a partition of n.

Partitions can be visualized by **Ferrers diagrams**, patterns of unfilled left aligned boxes where row i contains λ_i boxes.

Definition 1 *A **Young tableau** is a assignment of the numbers $1, 2, \dots n$ to the n boxes of a Ferrers diagram.*

Definition 2 *A **Young standard tableau** is a Young tableau where the entries are strictly increasing along each row and across each column.*

For convenience, we will also refer to Young standard tableaux as standard tableaux. The partition underlying a Ferrers diagram or a Young tableau as also referred to as its shape.

Example 1 *For $n = 3$, there are three possible partitions $(3), (2, 1)$ and $(1, 1, 1)$, which have the respective Ferrers diagrams:* , , .

Example 2 *There are only two possible Young standard tableaux with the shape $(2, 1)$:*

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

Permutations act on the set of Young standard tableaux by permuting the entries of the boxes. For $\sigma \in \mathbb{S}_n$ and t a standard tableau, the action of σ on t is denoted $\sigma \circ t$ and is given by permuting the entries of t by σ. For example, $(2, 3) \in \mathbb{S}_3$. The action of $(2, 3)$ on the standard tableau $\begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1} \end{array}$ is

$$(2, 3) \circ \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1} \end{array} = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1} \end{array}$$

Definition 3 *The **axial distance** $d_t(i, j)$ between entries i and j in tableau t is defined to be*

$$d_t(i, j) = (col(t, j) - col(t, i)) - (row(t, j) - row(t, i))$$

where $col(t, i)$ is the column index of label i in tableau t and similarly $row(t, i)$ is the row index of label i in tableau t. Also, $d_t(i, j) = -d_t(j, i)$

Example 3 *For* $t = \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & \\ \cline{1-1}\end{array}$, *we have the axial distances:*

$$d_t(1,2) = (1-1) - (2-1) = -1 \tag{1}$$
$$d_t(1,3) = (2-1) - (1-1) = 1 \tag{2}$$
$$d_t(2,3) = (2-1) - (1-2) = 2 \tag{3}$$

1.2 Defining Young's Orthogonal Representation on Transpositions

Young's Orthogonal Representation is one instantiation of the irreps irreps of $\mathbb{S}_n$ that is relatively efficient to compute. Given a set of generators for $\mathbb{S}_n$, it suffices to define the irrep matrices on all generator elements. Then any $\sigma \in \mathbb{S}_n$ can be expressed as a product of these generators: $\sigma = g_1 g_2 \dots g_k$ and using the property of irrep matrices we can produce $\rho(\sigma)$ from $\rho(\sigma) = \rho(\sigma_1)\dots\rho(\sigma_k)$.

A **transposition** is a permutation that swaps only two elements. Any cycle $(c_1, ..., c_l)$ is the product of transpositions: $(c_1, c_2)(c_2, c_3)\dots(c_{l-1}, c_l)$. Furthermore, any transposition (i, j) can be written as a product of adjacent transpositions of the form $\tau_k = (k, k+1)$

$$(i,j) = \tau_{j-1}, \dots \tau_{i+1}\tau_i\tau_{i+1}\dots\tau_{j-1}$$

The set of transpositions: $(1,2), (2,3), \dots (n-1, n)$ generate $\mathbb{S}_n$, so we only need to define Young's Orthogonal Representation for these transpositions.

For partition λ, we we denote the Young Orthogonal Representation associated with λ as $\rho_\lambda : \mathbb{S}_n \to \mathbb{C}^{d_\lambda \times d_\lambda}$, where d_λ is the dimension of ρ_λ. It turns out that d_λ is also the number of standard tableaux of shape λ.

Suppose we are given a fixed ordering of the standard Young tableaux of shape λ: $t_1, t_2, \dots, t_{d_\lambda}$. We will refer to the rows and columns of the ρ_λ matrices using the standard tableau in the fixed ordering above. The first standard tableau t_1 refers to the first column and row, t_2 refers to the second column and row, etc. $[\rho_\lambda(\sigma)]_{t_i,t_j}$ refers to the (i,j) entry of $\rho_\lambda(\sigma)$. For a transposition $(i-1,i)$, the entries of the YOR matrix $\rho_\lambda(i-1,i)$ are defined by the following rules:

1. On the diagonal entries:
$$[\rho_\lambda(i-1,i)]_{tt} = \frac{1}{d_t(i-1,i)}$$
2. For all entries (t_j, t_k) where $t_k \neq (i-1,i) \circ t_j$:
$$[\rho_\lambda(i-1,i)]_{t_j,t_k} = 0$$
3. If $(i-1,i) \circ t_j = t_k$, then
$$[\rho_\lambda(i-1,i)]_{t_j,t_k} = \sqrt{1 - \frac{1}{d_{t_j}^2(i-1,i)}}$$

Now that we have a prescription for computing the ρ_λ YOR matrices of any transposition of the form $\tau_k = (k, k+1)$, we can use them to compute representations for any other permutation. For $\sigma \in \mathbb{S}_n$, we first decompose σ into a product of transpositions: $\sigma = \tau_{a_1}\dots\tau_{a_k}$ using bubblesort Huang et al. (2009). Then $\rho_\lambda(\sigma)$ is simply the product of the YOR matrices of σ's constituent transpositions:

$$\sigma = \tau_{a_1}\dots\tau_{a_k} \tag{4}$$
$$\rho_\lambda(\sigma) = \rho_\lambda(\tau_{a_1})\dots\rho_\lambda(\tau_{a_k}) \tag{5}$$

Example 4 $\mathbb{S}_3$ *is generated by the two transpositions:* $\tau_1 = (1,2), \tau_2 = (2,3)$. *For* $\lambda = (2,1)$ *we have exactly two standard tableaux so* ρ_λ *is a* 2×2 *irrep. The YOR irrep matrices* $\rho_{\tiny\yng(2,1)}$ *for the elements of* $\mathbb{S}_3$ *are:*

$$\rho_{\tiny\yng(2,1)}(e) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \rho_{\tiny\yng(2,1)}(1,2) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \rho_{\tiny\yng(2,1)}(2,3) = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$$
$$\rho_{\tiny\yng(2,1)}(1,3) = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}, \rho_{\tiny\yng(2,1)}(1,2,3) = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}, \rho_{\tiny\yng(2,1)}(1,3,2) = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$$

The last three permutations of $\mathbb{S}_3$ are: $(1,3),(1,2,3)$ and $(1,3,2)$. To compute their $\rho_{\boxplus}$ matrices, we express each permutation as a product of the transpositions $(1,3)$ and $(2,3)$. So $(1,2,3) = (1,2)(2,3)$, giving us $\rho_{\boxplus}(1,2,3) = \rho_{\boxplus}(1,2)\rho_{\boxplus}(2,3)$, and so on for the remaining permutations.

1.3 Dimensions of the Irreps of the Symmetric Group

The number of standard tableau of shape λ, which is also the dimension of the irrep ρ_λ, is given by the **hook length formula**. Given any box of a Ferrers diagram of shape λ, its corresponding hook is defined to be that box, the boxes in its row to its right and the boxes in its column below it. The hook length of a box is the number of boxes in its hook. The hook length formula is

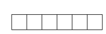
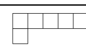
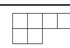
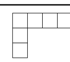
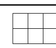
$$f^\lambda = \frac{n!}{\Pi_i l_i!}$$

where l_i is the hook length of box i, and i ranges over all boxes of the λ shaped Ferrers diagram.

Example 5 *The Ferrers diagram of shape $\lambda = (2,1)$ is $\begin{array}{|c|c|}\hline a & b \\ \hline c & \\ \cline{1-1}\end{array}$. The hook lengths of the boxes are: $l_a = 3, l_b = 1, l_c = 1$. So $f^{\boxplus} = \frac{3!}{3\cdot 1\cdot 1} = 2$.*

In Table 1, we show the hook length or irrep dimension for a few of the partitions of $n = 6$.

Table 1: Dimension of irreps for $n = 6$

λ	Ferrers Diagram	f^λ
(6)		1
$(5,1)$		5
$(4,2)$		9
$(4,1,1)$		10
$(3,3)$		5

2 Irreducible Representations of Wreath Product Groups

The description of the irreps of wreath product groups here follows Kerber (1971) and Rockmore (1995).

2.1 Irreducible Representations of C_m^n

Recall that the irreps of C_m are the complex exponentials: $\chi_j(k) = e^{ijk/m}$ for $j = 0, 1, \ldots, m-1$. The irreps of C_m^n are: $\chi_{j_1} \otimes \chi_{j_2} \otimes \ldots \otimes \chi_{j_n}$ for $j_1, ..., j_n \in \{0, \ldots m-1\}$. For $x \in C_m^n$,

$$(\chi_{j_1} \otimes \chi_{j_2} \otimes \ldots \otimes \chi_{j_n})(x) = \bigotimes_{i=1}^{n} \chi_{j_i}(x_i) = \prod_{i=1}^{n} \chi_{j_i}(x_i) \tag{6}$$

The last equality holds because the χ_j irreps are 1-dimensional.

In subsequent sections, for partitions of n into m non-negative parts, $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_m)$, $\alpha_i \geq 0$, we will use the notation χ^α to denote the irrep of C_m^n where

$$\chi^\alpha = \underbrace{\chi_0 \otimes ... \otimes \chi_0}_{\alpha_1 \text{ times}} \otimes \underbrace{\chi_1 \otimes ... \otimes \chi_1}_{\alpha_2 \text{ times}} \otimes \ldots \otimes \underbrace{\chi_{m-1} \otimes ... \otimes \chi_{m-1}}_{\alpha_m \text{ times}} \tag{7}$$

$$= \chi_0^{\otimes \alpha_1} \otimes \chi_1^{\otimes \alpha_2} \otimes \ldots \otimes \chi_{m-1}^{\otimes \alpha_m} \tag{8}$$

For $f \in C_m^n$, $\pi \in \mathbb{S}_n$, let $f = (f_1, ..., f_n)$. The permutation action of $\mathbb{S}_n$ on the C_m^n, denoted $\pi \circ f$, permutes the indices of f:

$$(\pi \circ f)_i = f_{\pi^{-1}(i)} \tag{9}$$

for $i = 1 \ldots n$.

2.2 Young Subgroups

The **Young Subgroup** corresponding to a partition $\alpha = (\alpha_1, \ldots, \alpha_m)$, where $\sum_{i=1}^{m} \alpha_i = n$, $\alpha_i \geq 0$ for $i = 1, \ldots m$ is the group

$$\mathbb{S}_\alpha = \mathbb{S}_{\{1,\ldots,\alpha_1\}} \times \mathbb{S}_{\{\alpha_1+1,\ldots,\alpha_1+\alpha_2\}} \times \ldots \times \mathbb{S}_{\{n-\alpha_m+1,\ldots,n\}}.$$

$\mathbb{S}_\alpha$ is also isomorphic to $\mathbb{S}_{\alpha_1} \times \mathbb{S}_{\alpha_2} \times \ldots \times \mathbb{S}_{\alpha_m}$.

The irreps of $\mathbb{S}_\alpha$ are constructed by taking the tensor product of irreps of each individual $\mathbb{S}_{\alpha_i}$. Suppose we have $\psi = (\psi_1, \psi_2, ..., \psi_m)$, where each ψ_i is a partition of α_i, then ψ indexes an irrep of $\mathbb{S}_\alpha$:

$$\rho_\psi = \rho_{\psi_1} \otimes \rho_{\psi_2} \otimes \ldots \otimes \rho_{\psi_m}.$$

We can compute each ρ_{ψ_i} using Young's Orthogonal Representation as detailed in Section 1.2.

Example 6 *For* $\alpha = (2, 2), \mathbb{S}_\alpha = \mathbb{S}_{\{1,2\}} \times \mathbb{S}_{\{3,4\}} = \{e, (1, 2), (3, 4), (1, 2)(3, 4)\}$. *The irreps of* $\mathbb{S}_\alpha$ *are the tensor product of the irreps of* $\mathbb{S}_2$ *by the irreps of* $\mathbb{S}_2$*:* $(\rho_{(2)} \otimes \rho_{(2)}), (\rho_{(2)} \otimes \rho_{(1,1)}), (\rho_{(1,1)} \otimes \rho_{(2)})$*, and* $(\rho_{(1,1)} \otimes \rho_{(1,1)})$.

2.3 Induced Representations

Given a group G with a subgroup H. Let $g_1, \ldots g_r$ be a set of left coset representatives of H in G. So $G = g_1 H \cup \ldots \cup g_r H$, with $r = [G : H]$. Suppose H has the representation $\rho : H \to \mathbb{C}^{n \times n}$. Then the **induced representation** $\tilde{\rho} : G \to \mathbb{C}^{nr \times nr}$ is defined as:

$$\tilde{\rho}(x) = \begin{pmatrix} \rho(g_1^{-1} x g_1) & \rho(g_1^{-1} x g_2) & \ldots & \rho(g_1^{-1} x g_r) \\ \rho(g_2^{-1} x g_1) & \rho(g_2^{-1} x g_2) & \ldots & \rho(g_2^{-1} x g_r) \\ \vdots & \vdots & \ddots & \vdots \\ \rho(g_r^{-1} x g_1) & \rho(g_r^{-1} x g_2) & \ldots & \rho(g_r^{-1} x g_r) \end{pmatrix} \tag{10}$$

where ρ is extended to G by defining $\rho(x) = 0$ if $x \notin H$. See Sagan (2013) for a proof on how this construction of $\tilde{\rho}$ results in a representation of G. We denote the induced representation of ρ from H to G as $\tilde{\rho} = \mathrm{Ind}_H^G \rho$.

An important property of induced representations is that only one entry of every row of Eq. (10) will be non-zero, and likewise for every column. This means that that the total number of non-zero blocks of $\mathrm{Ind}_H^G \rho$ will be $[G : H]$. At most, only $\frac{1}{[G:H]}$ of the entries of $\mathrm{Ind}_H^G \rho(g)$ will be non-zero for all $g \in G$, giving us a sparse matrix which can be stored in memory more efficiently.

2.4 Putting it all together

The irreps of $C_m \wr \mathbb{S}_n$ are indexed by tuples (α, ψ), where α is a partition of n into m non-negative parts: $\alpha = (\alpha_1, \ldots, \alpha_m)$, and $\psi = (\psi_1, \ldots, \psi_m)$ is a tuple of m partitions, where ψ_i is a partition of α_i. Some of the α_i's may be 0, but each of the ψ_i partitions must only contain positive integers. We denote the irrep of $C_m \wr \mathbb{S}_n$ indexed by (α, ψ) with $\rho_{(\alpha,\psi)}$.

Example 7 *The tuple* (α, ψ) *for* $\alpha = (2, 3, 3), \psi = ((2), (1, 1, 1), (2, 1)$ *indexes an irrep of* $C_3 \wr \mathbb{S}_8$.

Let a full set of coset representatives of the Young Subgroup $\mathbb{S}_\alpha$ in $\mathbb{S}_n$ be $R = \{g_1, ..., g_k\}$. For $(f, \pi) \in C_m \wr \mathbb{S}_n$, the irrep matrix $\rho_{(\alpha,\psi)}(f, \pi)$ is the block matrix, with the (i, j) block defined as

$$\begin{aligned}[\rho_{(\alpha,\psi)}(f, \pi)]_{ij} &= z(g_i, f) \cdot [\mathrm{Ind}_{\mathbb{S}_\alpha}^{\mathbb{S}_n} \rho(\pi)]_{ij} \\ &= z(g_i, f) \cdot \rho(g_i^{-1} \pi g_j)\end{aligned}$$

for $i, j = 1, \dots, k$, where $z : \mathbb{S}_n \times C_m^n \to \mathbb{C}$ is the function defined as

$$z(g, f) = \chi^{\alpha}(g^{-1} \circ f) \tag{11}$$

($g^{-1} \circ f$ is defined by Eq. (9), χ^{α} is defined by Eq.(7)) and $\rho = \bigotimes_{i=1}^{m} \rho_{\psi_i}$, is an irrep of the Young Subgroup $\mathbb{S}_{\alpha}$. The dimension of this representation is $d_{(\alpha,\psi)} = [\mathbb{S}_n : \mathbb{S}_{\alpha}] \cdot \Pi_{i=1}^{m} d_{\psi_i}$, where d_{ψ_i} is the dimension of the irrep ρ_{ψ_i} of $\mathbb{S}_{\alpha_i}$.

Recall from Eq (10) that $\mathrm{Ind}_{\mathbb{S}_{\alpha}}^{\mathbb{S}_n} \rho(\pi)$, is a block matrix where the (i, j) block is $\rho(g_i^{-1} \pi g_j)$, where

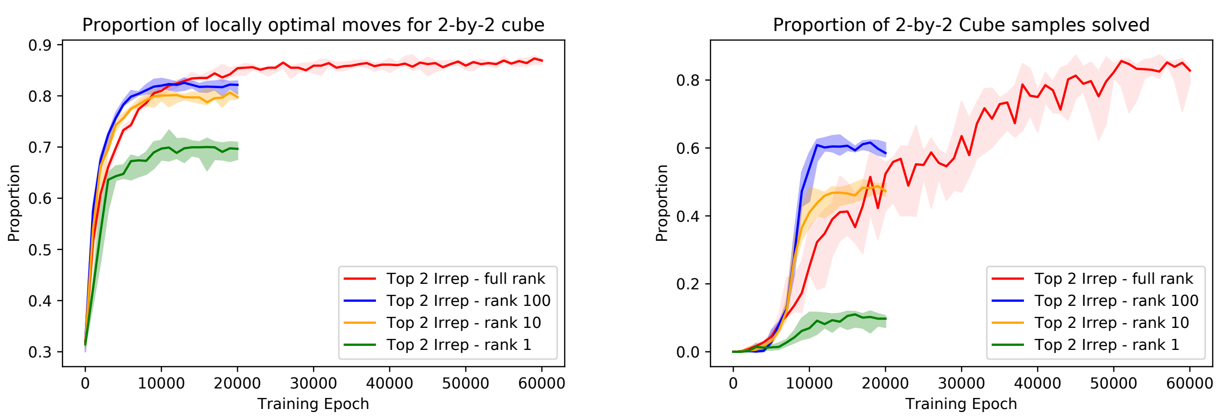
$$\rho(g_i^{-1} \pi g_j) = \begin{cases} 0 & \text{if } g_i^{-1} \pi g_j \notin \mathbb{S}_{\alpha} \\ (\rho_{\alpha_1} \otimes \ldots \otimes \rho_{\alpha_m})(g_i^{-1} \pi g_j) & \text{otherwise} \end{cases}$$

as defined shown in Eq. (10).

All of the irreps of $\mathbb{S}_8$ and Pyraminx ($C_2 \wr \mathbb{S}_6$) can be computed once and loaded into RAM when needed. The 2-by-2 cube group ($C_3 \wr \mathbb{S}_8$) is too large to load every irrep matrix into RAM so we only store and load the irrep matrices ρ of $\mathbb{S}_{\alpha}$ and $\mathrm{Ind}_{\mathbb{S}_{\alpha}}^{\mathbb{S}_n} \rho$. At runtime, we compute the scalar factors $z(g_i, f)$ to multiply with each non-zero (i, j) block of $\mathrm{Ind}_{\mathbb{S}_{\alpha}}^{\mathbb{S}_n} \rho$ to construct $\rho_{(\alpha,\psi)}$.

3 Miscellaneous Experiment Details

We trained low rank Fourier models in the basis spanned by the top 2 irreps for the 2-by-2 cube. The top two irreps comprise a total of: $560 \times 560 + 420 \times 420 = 490000$ basis functions, which happens to be a perfect square: 700×700. This allowed us to express the parameters of the low rank value function as $\theta_k = U_k V_k^{\top} \in \mathbb{C}^{700 \times 700}$: where $U_k \in \mathbb{C}^{700 \times k}$, $V_k \in \mathbb{C}^{700 \times k}$ for rank $k = 1, 10, 100$. We trained the low rank models until their performance on proportion of greedy solves and locally optimal moves converged, which happened after 20k epochs. Over the course of these 20k epochs, the low rank models saw only 165k unique cube states, or about 4.5% of all possible states. Figure 1 shows the proportion of locally optimal moves and proportion of greedy solves made by the low rank models during training plotted against the performance of the full rank top 2 irrep model. As we can see from these plots, the low rank models have limited capacity but also converge in a fraction of the training epochs required for the full rank model.



(a) Locally optimal moves

(b) Greedy solves

Figure 1: Training curve of low rank models for 2-by-2 cube

3.1 Hyperparameters

As we mentioned in the Experiments section, we chose the hyperparameters of our training procedure and baseline DVN architecture by doing a randomized grid search over the parameters listed in 2. While McAleer et al. (2018)

and Agostinelli et al. (2019) both used batch normalization layers after each layer's ReLU, we found that it slowed training time without improving the proportion of solves and locally optimal moves.

Table 2: Hyperparameters

Parameter	Value Range
seed	{0, 1, 2, 3, 4}
minibatch size	{32, 64, 128}
learning rate	[0.003, 0.006]
discount γ	{0.95, 0.99, 1}
hidden layer size	{512, 1024, 1536, 2048}
number of layers	{1, 2, 3}
weight initialization std	{0.03, 0.05, 0.1}
replay buffer capacity	{10000, 100000}
target update interval	{50, 100}
Adam weight decay	{0, 1e-1, 1e-2, 1e-3, 1e-4}
Adam β	{0.99}

3.2 Irreps

Tables 3, 4 and 5 show the irreps used for the Pyraminx, $\mathbb{S}_8$ and 2-by-2 cube experiments respectively. The irreps of $\mathbb{S}_n$ are indexed by partitions of n. The irreps of wreath product groups of the form $C_m \wr \mathbb{S}_n$ are indexed by tuples (α, ψ), where α is a partition of n and ψ is a tuple of partitions, where each $\psi_i \in \psi$ is a partition of α_i.

Finding the top irreps to use remains a nontrivial and open ended research question. There is, however, a certain symmetry among the irreps of wreath product groups that helps reduce our search space. Consider arbitrary irreps ρ_1 and ρ_2 of the 2-by-2 cube, where ρ_1 is indexed by the tuple $(3, 1, 4), ((\lambda_1, \lambda_2, \lambda_3))$ and ρ_2 is indexed by the tuple $((3, 4, 1), (\lambda_1, \lambda_3, \lambda_2))$. Then ρ_1 and ρ_2 essentially contain equivalent sets of basis functions, so we only need to consider the efficacy of one of them. This symmetry alone allowed us to ignore $\sim 40\%$ of the 270 total irreps of the 2-by-2 cube.

References

Forest Agostinelli, Stephen McAleer, Alexander Shmakov, and Pierre Baldi. Solving the Rubik's cube with deep reinforcement learning and search. *Nature Machine Intelligence*, 1(8):356–363, 2019.

Jonathan Huang, Carlos Guestrin, and Leonidas Guibas. Fourier theoretic probabilistic inference over permutations. *Journal of Machine Learning Research*, 10(5), 2009.

A. Kerber. *Representations of Permutation Groups I*, volume 240 of *Lecture Notes in Mathamatics*. Springer-Verlag Berlin Heidelberg, 1971.

Imre Risi Kondor. *Group theoretical methods in machine learning*, volume 2. Columbia University New York, 2008.

Stephen McAleer, Forest Agostinelli, Alexander Shmakov, and Pierre Baldi. Solving the rubik's cube with approximate policy iteration. In *International Conference on Learning Representations*, 2018.

Daniel N Rockmore. Fast Fourier transforms for wreath products. *Applied and Computational Harmonic Analysis*, 2(3): 279–292, 1995.

Bruce E Sagan. *The symmetric group: representations, combinatorial algorithms, and symmetric functions*, volume 203. Springer Science & Business Media, 2013.

Table 3: Top 24 irreps used for Pyraminx

α	ψ	$d_{(\alpha,\psi)}$	$d^2_{(\alpha,\psi)}$
(3, 3)	((2, 1), (3))	40	1600
(3, 3)	((2, 1), (1, 1, 1))	40	1600
(2, 4)	((1, 1), (2, 1, 1))	45	2025
(2, 4)	((2), (3, 1))	45	2025
(4, 2)	((1, 1, 1, 1), (2))	15	225
(2, 4)	((1, 1), (4))	15	225
(1, 5)	((1), (3, 1, 1))	36	1296
(5, 1)	((4, 1), (1))	24	576
(2, 4)	((1, 1), (3, 1))	45	2025
(4, 2)	((1, 1, 1, 1), (1, 1))	15	225
(4, 2)	((2, 1, 1), (2))	45	2025
(0, 6)	((), (5, 1))	5	25
(6, 0)	((1, 1, 1, 1, 1, 1), ())	1	1
(6, 0)	((4, 1, 1), ())	10	100
(6, 0)	((2, 2, 2), ())	5	25
(6, 0)	((3, 3), ())	5	25
(0, 6)	((), (2, 1, 1, 1, 1))	5	25
(1, 5)	((1), (5))	6	36
(1, 5)	((1), (2, 1, 1, 1))	24	576
(1, 5)	((1), (1, 1, 1, 1, 1))	6	36
(1, 5)	((1), (2, 2, 1))	30	900
(1, 5)	((1), (3, 2))	30	900
(2, 4)	((2), (4,))	15	225
(2, 4)	((2), (2, 2))	30	900
		Total basis functions	17621

Table 4: Top 4 irreps used for $\mathbb{S}_8$

λ	d_λ	d^2_λ
(4, 2, 2)	56	3136
(3, 2, 2, 1)	70	4900
(5, 1, 1, 1)	35	1225
(3, 3, 1, 1)	56	3136
	Total basis functions	12397

Table 5: Top 2 irreps used for the 2-by-2 Cube

α	ψ	$d_{(\alpha,\psi)}$	$d^2_{(\alpha,\psi)}$
(2, 3, 3)	((2), (1,1,1), (1,1,1))	560	313600
(4, 2, 2)	((4), (1,1), (1,1))	420	176400
		Total basis functions	490000