Signal Processing on the Permutahedron: Tight Spectral Frames for Ranked Data Analysis

Yilin Chen · **Jennifer DeJong** · **Tom Halverson** · **David I Shuman**

Abstract Ranked data sets, where m judges/voters specify a preference ranking of n objects/candidates, are increasingly prevalent in contexts such as political elections, computer vision, recommender systems, and bioinformatics. The vote counts for each ranking can be viewed as an $n!$ data vector lying on the permutahedron, which is a Cayley graph of the symmetric group with vertices labeled by permutations and an edge when two permutations differ by an adjacent transposition. Leveraging combinatorial representation theory and recent progress in signal processing on graphs, we investigate a novel, scalable transform method to interpret and exploit structure in ranked data. We represent data on the permutahedron using an overcomplete dictionary of atoms, each of which captures both smoothness information about the data (typically the focus of spectral graph decomposition methods in graph signal processing) and structural information about the data (typically the focus of symmetry decomposition methods from representation theory). These atoms have a more naturally interpretable structure than any known basis for signals on the permutahedron, and they form a Parseval frame, ensuring beneficial numerical properties such as energy preservation. We develop specialized algorithms and open software that take advantage of the symmetry and structure of the permutahedron to improve the scalability of the proposed method, making it more applicable to the high-dimensional ranked data found in applications.

Keywords ranked data analysis · graph signal processing · tight frame · permutahedron · representation theory · Schreier graph · symmetric group

1 Introduction

Ranked data consist of m judges/voters specifying a preference ranking of n objects/candidates. While methods for analyzing such rankings date back to the late 18th century in the context of social choice theory (e.g., [11,23, 4]), ranked data are increasingly prevalent in contexts such as computer vision [45,44], recommender systems [86,98], image processing [99,8], crowdsourced subjective labeling [9,16,92], peer grading [78], metasearch [1], sports analytics [27], computational geometry [48], and bioinformatics [12,26,61,96] (see [90, Sec 2.2] for an excellent, comprehensive overview of application areas). Moreover, an increasing number of cities, states, colleges and universities, organizations, and corporations are using ranked choice voting for elections [35].

Department of Mathematics, Statistics, and Computer Science
Macalester College
Saint Paul, MN 55105 USA
E-mail: {ychen2,jdejong,halverson,dshuman1}@macalester.edu

The vote counts for each possible ranking of n objects is an $n!$-dimensional data vector in $\mathbb{R}^{n!}$. We view this vector as lying on the *permutahedron*, denoted $\mathbb{P}_n$ and also referred to by some as the *permutation polytope* [94]. The permutahedron has vertices labeled by permutations and an edge when two permutations differ by transposing adjacent entries in the permutation. For example, in $\mathbb{P}_5$, the permutation 25134 corresponds to ranking candidate 2 first, candidate 5 second, candidate 1 third, and so on, and it is connected by an edge to each of 52134, 21534, 25314, and 25143. The permutahedron $\mathbb{P}_n$ is the Cayley graph of the symmetric group $\mathbb{S}_n$ induced by the generating set of adjacent transpositions (see Sec. 3.1), and a signal on the permutahedron $\mathbb{P}_n$ is a function $f : \mathbb{S}_n \to \mathbb{R}$. In this context, $f(\sigma)$ equals the number of votes for the permutation σ.

To deal with the scale of this data (factorial in the number of candidates), it is critical to construct efficient and meaningful representations that highlight salient features of the ranking tallies. Specifically, we follow the common signal processing approach of constructing a dictionary of atoms and representing a signal on the permutahedron as a linear combination of these atoms. For audio signals and images, as well as signals residing on more general weighted graphs, Fourier, time-frequency, curvelet, shearlet, bandlet, and other dictionaries have led to resounding successes in visual analysis of data, statistical analysis of data, compression, and as regularizers in machine learning and ill-posed inverse problems such as inpainting, denoising, and classification (see, e.g., [82, Sec. II] for an excellent historical overview of dictionary design methods and signal transforms).

In general, desirable properties when designing dictionaries include: (i) the dictionaries comprise an orthonormal basis or tight frame for the signal space, ensuring that the contribution of each atom can be computed via an inner product with the signal, and the energy of the signal is equal to the energy of the transform coefficients; (ii) the atoms have an interpretable structure, so that the inner products between the signal and each atom are informative; (iii) it is numerically efficient to apply the dictionary analysis and synthesis operators (forward and inverse transforms); and (iv) signals of certain mathematical classes can be represented exactly or approximately as sparse linear combinations of a subset of the dictionary atoms [87].

The main contributions of this work are as follows. First, we leverage techniques and ideas from both signal processing on graphs and combinatorial representation theory to propose a novel dictionary construction that can be used to transform high-dimensional ranked data in order to find, interpret, and exploit structural patterns in the rankings. Each of the atoms in the overcomplete dictionaries we propose captures both smoothness information about the data (typically the focus of spectral graph decomposition methods in graph signal processing) and structural information about the data (typically the focus of symmetry decomposition methods from representation theory). Second, we prove that the proposed dictionaries comprise tight Parseval frames and therefore preserve the energy of the signal (Thm. 1 and Thm. 2 in Sec. 4). Third, we demonstrate the application of the proposed transform methods and show how the interpretable structure of the atoms can lead to insights on real ranked data sets (Sec. 5). Fourth, we investigate numerical challenges, and propose novel algorithms that take advantage of the symmetry and structure of the permutahedron to enhance the scalability of our implementations for applying the analysis operators arising from our proposed dictionaries (Sec. 6). Fifth, we relate the proposed transform methods to related methods from combinatorial representation theory, graph signal processing, and statistical modeling for ranked data (Sec. 3 and Sec. 7).

2 Example Data Sets

We use the following three ranked data sets in running examples throughout this article.

2.1 1980 American Psychological Association Presidential Election

In Fig. 1 (top), on the permutahedron $\mathbb{P}_5$ we plot the vote tallies of the 5738 American Psychological Association (APA) members who ranked all five candidates for APA president in 1980 (out of the 15449 total ballots cast) [14], [29, Tab. 1]. Under the instant runoff (Hare) voting system in which the votes for the candidate with the fewest first place votes in each iteration are transferred to the next ranked candidate on those ballots, candidate 1 was the winner.

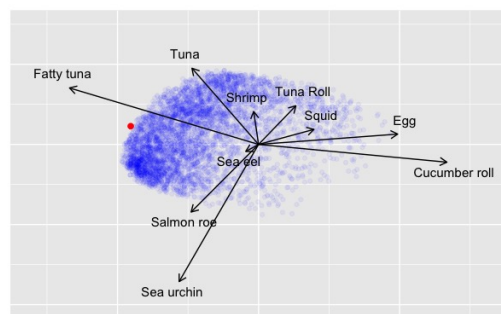### 2.2 2017 Minneapolis City Council Ward 3 Election

In Fig. 1 (bottom), on the permutahedron $\mathbb{P}_4$ we plot the vote tallies of the 5055 voters who ranked at least three of the four candidates for the Minneapolis City Council Ward 3 seat in 2017 (out of 9578 total valid votes cast) [73]. If a voter ranked three candidates, we assume that the unranked candidate was the voter's fourth choice. Candidates 1 to 4 are Ginger Jentzen (Socialist-Alternative), Samantha Pree-Stinson (Green), Steve Fletcher (Democratic-Farmer-Labor), and Tim Bildsoe (Democratic-Farmer-Labor), respectively. Pree-Stinson began as a Democratic-Farmer-Labor (DFL) candidate, but was later endorsed by the Green Party. Fletcher was endorsed by the DFL Party. Jentzen received the most first place votes, but Fletcher won the election under the instant runoff (Hare) voting system utilized by Minneapolis. A discussion of the candidates' views and an interesting analysis of the voting results by geographical regions within the ward is contained in [101].

Fig. 1 Top: The signal $\mathbf{f}$ on $\mathbb{P}_5$ represents the vote tallies of 5738 complete rankings of five candidates for APA president in 1980. Bottom: Ranked choice voting data $\mathbf{g}$ for the 2017 Minneapolis City Council Ward 3 election, which featured four candidates, shown on $\mathbb{P}_4$.

2.3 Sushi Preference Data

For a data set with more candidates ($n = 10$), we use Kamishima's type A set of sushi preferences [50,51], in which 5000 people provide complete preference rankings for the ten different types of sushi listed in Fig. 2(a). Since plotting a signal on the permutahedron $\mathbb{P}_{10}$ (a 9-dimensional object) is not particularly informative, we show in Fig. 2(b) the projection of the data into a two-dimensional space via Gabriel's biplot [39], [22], [67, Sec. 2.2], which is similar to a principal components projection except with a different choice of center. Chen [15, Sec. 5.3] notes that there is a full Condorcet ranking:

Index	Sushi Type
1	Shrimp
2	Sea eel
3	Tuna
4	Squid
5	Sea urchin
6	Salmon roe
7	Egg
8	Fatty tuna
9	Tuna roll
10 (0)	Cucumber roll

(a)

(b)

Fig. 2 (a) Candidate items in the sushi preference data set [50,51]. (b) Projection of sushi preference signal, $\mathbf{h}$, onto two dimensions using Gabriel's biplot. Each blue dot represents a single voter (of the 5000 total), and the transparency is indicative of the number of voters, with darker shades indicating a higher density of voters. The poles associated with the popular sushi items tend to fall towards the left on the fatty tuna-cucumber roll axis. The projection of the Condorcet ranking is plotted in red.

Fatty tuna $\succ$ Tuna $\succ$ Salmon roe $\succ$ Shrimp $\succ$ Sea eel $\succ$ Sea urchin $\succ$ Squid $\succ$ Tuna roll $\succ$ Egg $\succ$ Cucumber roll.

That is, in a preference comparison of any pair of two sushi items, the majority of voters would prefer the item that falls earlier in this ranking; i.e., the majority of voters prefer fatty tuna to any other item, the majority of voters prefer tuna to any item besides fatty tuna, and so forth.
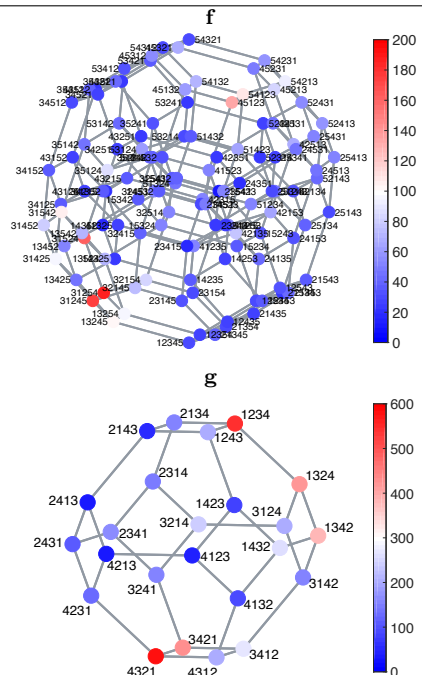
3 Related Work

Analysis of ranked data has a long history in the mathematical psychology and statistics literatures. Approaches have largely focused on parametric statistical models including order statistic models, distance-based models, and pairwise comparison models (see, e.g., [62], [63], [67], [90, Sec. 2.5], and [100] for excellent overviews of these models). We focus our attention here on linear transforms for ranked data that attempt to identify structure in the data by taking inner products between the $n!$-dimensional vote tally (signal on the permutahedron) and building block signals that have some interpretable structure.

3.1 The Fourier Analysis on the Symmetric Group Approach

The vertices of the permutahedron $\mathbb{P}_n$ are labeled by the symmetric group $\mathbb{S}_n$ of permutations of $\{1, \ldots, n\}$. For example $\sigma = \left(\begin{smallmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 1 & 3 & 4 \end{smallmatrix}\right) = 25134 \in \mathbb{S}_5$ is the bijective function $1 \mapsto 2, 2 \mapsto 5$, etc.. We let (i, j) denote the transposition (in cycle notation) that exchanges i and j and fixes the other entries, and the adjacent transpositions are $(i, i+1), 1 \leq i \leq n-1$. The group operation in $\mathbb{S}_n$ is composition of functions, and right multiplication by an adjacent transposition $(i, i+1)$ exchanges the candidates in positions i and $i+1$. For example, $\left(\begin{smallmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 1 & 3 & 4 \end{smallmatrix}\right)(2, 3) = \left(\begin{smallmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 5 & 3 & 4 \end{smallmatrix}\right)$. In this context, the permutahedron $\mathbb{P}_n$ is the Cayley graph with vertices labeled by $\mathbb{S}_n$ and edges $\{(\sigma, \sigma s) \mid \sigma \in \mathbb{S}_n, s \in S\}$, where $S = \{(i, i+1) \mid 1 \leq i \leq n-1\}$. Each transposition is its own inverse, so the generating set S is closed under inverses and $\mathbb{P}_n$ is a simple graph that is $(n-1)$-regular, since $|S| = n-1$.

Using right multiplication of transpositions is natural when studying ranked data, because two permutations are adjacent if and only if they differ by transposing adjacent candidates in the ranking. Under left multiplication, two rankings are adjacent if and only if they differ by swapping the candidates ranked ith and $(i+1)$st. For example, $(2, 3) \cdot \left(\begin{smallmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 1 & 3 & 4 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 1 & 2 & 4 \end{smallmatrix}\right)$. The adjacent transpositions S satisfy the Coxeter relations [83, 2.12.10] and give $\mathbb{S}_n$ the structure of a reflection group. If one uses the full set of transpositions $\{(i, j) \mid 1 \leq i < j \leq n\}$, then the generating set is a full conjugacy class in $\mathbb{S}_n$. In that case, the Cayley graph is said to be quasi-abelian, and the Laplacian eigenvalues and eigenvectors are especially nice, but they are less interpretable in the context of ranked data as we discuss in Sec. 3.2.

Data on $\mathbb{P}_n$ lives in the underlying real vector space $\mathbb{R}[\mathbb{S}_n]$ (or, equivalently $\mathbb{R}^{n!}$) spanned by the symmetric group $\mathbb{S}_n$, which is called the *group algebra* of $\mathbb{S}_n$ and has a canonical basis $\{\mathbf{e}_\sigma\}_{\sigma \in \mathbb{S}_n}$.[1] A signal on $\mathbb{S}_n$ is a function $f : \mathbb{S}_n \to \mathbb{R}$, which we view as a vector in $\mathbb{R}[\mathbb{S}_n]$ by $\mathbf{f} = \sum_{\sigma \in \mathbb{S}_n} f(\sigma)\mathbf{e}_\sigma$. Non-commutative harmonic analysis attempts to find structure in ranked data by decomposing them into subspaces that are set-wise invariant under the relabeling (left multiplication) and/or re-ranking (right multiplication) of the candidates [29]. Specifically, a permutation $\sigma \in \mathbb{S}_n$ acts on the right and on the left, respectively, of a signal $\mathbf{f}$, by

$$\mathbf{f}\sigma = \sum_{\tau \in \mathbb{S}_n} f(\tau)\mathbf{e}_{\tau\sigma} = \sum_{\eta \in \mathbb{S}_n} f(\eta\sigma^{-1})\mathbf{e}_\eta \qquad \text{and} \qquad \sigma\mathbf{f} = \sum_{\tau \in \mathbb{S}_n} f(\tau)\mathbf{e}_{\sigma\tau} = \sum_{\beta \in \mathbb{S}_n} f(\sigma^{-1}\beta)\mathbf{e}_\beta. \tag{1}$$

The vector space $\mathbb{R}[\mathbb{S}_n]$ decomposes into the direct sum of orthogonal subspaces, called *isotypic components*

$$\mathbb{R}[\mathbb{S}_n] \cong \bigoplus_{\gamma \vdash n} W_\gamma, \tag{2}$$

where the sum is over all integer partitions $\gamma = [\gamma_1, \gamma_2, \ldots, \gamma_\ell]$ of n, denoted by $\gamma \vdash n$. The subspaces W_γ are invariant under both relabeling and reindexing by $\mathbb{S}_n$; that is, if $\mathbf{w} \in W_\gamma$, then $\mathbf{w}\sigma \in W_\gamma$ and $\sigma\mathbf{w} \in W_\gamma$ for any permutation $\sigma \in \mathbb{S}_n$. Thus, W_γ is both a left and a right submodule of $\mathbb{R}[\mathbb{S}_n]$. We refer to the integer partition γ as the *shape* or *symmetry type* (we use these terms interchangeably) of W_γ.

The isotypic components further decompose into a direct sum of *irreducible* left and right $\mathbb{S}_n$-submodules, respectively, as

$$\mathbb{R}[\mathbb{S}_n] \cong \bigoplus_{\gamma \vdash n} W_\gamma \cong \bigoplus_{\gamma \vdash n} \bigoplus_{i=1}^{d_\gamma} V_{\gamma, i} \cong \bigoplus_{\gamma \vdash n} \bigoplus_{i=1}^{d_\gamma} V_{\gamma, i}^*, \tag{3}$$

---

[1] The results of this paper are true over the complex numbers; however, we use $\mathbb{R}[\mathbb{S}_n]$ as ranked data are real valued.
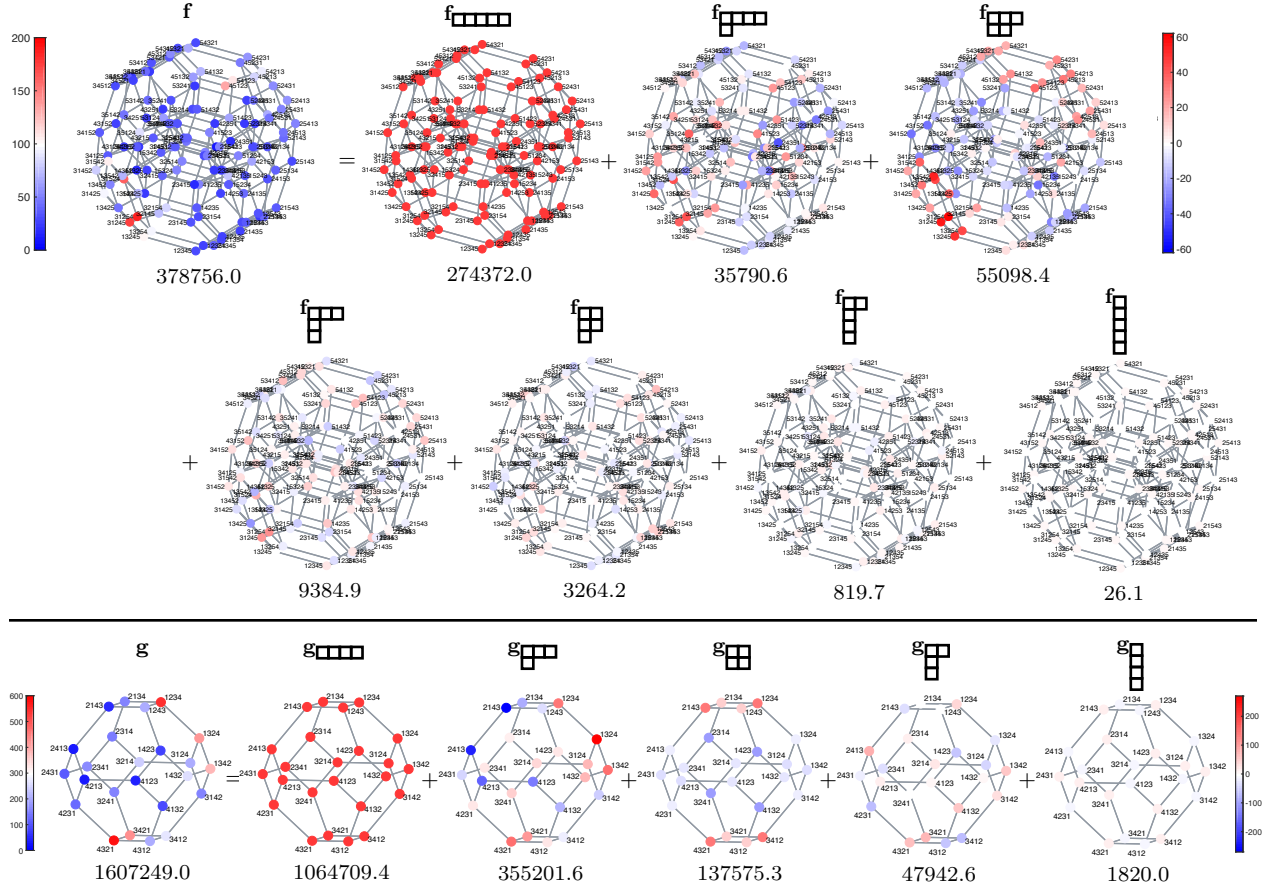
Fig. 3 Decomposition of the signals of Fig. 1 into orthogonal vectors in each of their respective isotypic components. The numbers below the signals show how the energy $\|\mathbf{f}\|^2$ splits into the energies $\|\mathbf{f}_\gamma\|^2$. The first of these energies is a function of the number of voters (e.g., $274372 = 5738^2/120$ for the APA data of [29]). The second is a measure of the first order effect given by the number of voters who rank candidate i in position j. The third and fourth capture the unordered and ordered second order (pair) effects, respectively, net of the first order marginal effects.

where $V_{\gamma,i} \cong V_{\gamma,j}$ and $V_{\gamma,i}^* \cong V_{\gamma,j}^*$ are isomorphic as left and right $\mathbb{S}_n$-modules, respectively, and $V_{\gamma,i} \not\cong V_{\rho,j}$ and $V_{\gamma,i}^* \not\cong V_{\rho,j}^*$ if $\gamma \neq \rho$. The modules $V_{\gamma,i}$ are left invariant, meaning $\sigma\mathbf{v} \in V_{\gamma,i}$ for all $\mathbf{v} \in V_{\gamma,i}$ and $\sigma \in \mathbb{S}_n$. The modules $V_{\gamma,i}^* = \{f : V_{\gamma,i} \to \mathbb{R}\}$ are the dual vector spaces of linear functionals on $V_{\gamma,i}$ and are invariant under the right action $(f\sigma)(\mathbf{v}) = f(\sigma\mathbf{v})$. The left action of $\sigma \in \mathbb{S}_n$ on permutations replaces i with $\sigma(i)$ and the right action replaces the entry in the ith position with the entry in the $\sigma(i)$th position, so the left modules $V_{\gamma,i}$ are invariant under relabeling the candidates and the right modules are invariant under re-indexing the candidates. For shorthand, we write $\bigoplus_{i=1}^{d_\gamma} V_{\gamma,i}$ as $V_\gamma^{\oplus d_\gamma}$ and $\bigoplus_{i=1}^{d_\gamma} V_{\gamma,i}^*$ as $(V_\gamma^*)^{\oplus d_\gamma}$. The famous isomorphism in (3) was proved around 1900 in the work of Frobenius and Burnside. It was extended in the 1920s to hold for topological groups in the Peter-Weyl theorem. This decomposition has special properties: (i) the dimension $d_\gamma = \dim(V_\gamma)$ equals the multiplicity of V_γ in W_γ; (ii) d_γ also equals the number of standard Young tableaux of shape λ and can be computed using the hook formula (see, e.g., [83, 3.10]); (iii) $\dim(W_\gamma) = d_\gamma^2$; and (iv) the only left submodules of $\mathbb{R}[\mathbb{S}_n]$ isomorphic to V_γ appear in W_γ. The same properties hold for the right action with the dual spaces.

In Fig. 3, we display the APA data and Minneapolis City Council election data from Fig. 1 as the sum of projections onto the orthogonal subspaces W_γ.

How can this approach be used to find structure in ranked data? The isotypic decomposition (2) has a close relation with marginal statistics. The first order marginals of ranked data, shown in Fig. 4 for an election with four

candidates, capture how many voters placed candidate i in ranking position j. There are two types of second order marginals. The unordered type, shown in Fig. 5(a) for the same election, capture how many voters placed candidates i and i' in ranking positions j and j', in either order. The ordered type, shown in Fig. 5(b), capture how many voters placed

Candidate	First Choice	Second Choice	Third Choice	Fourth Choice
Ginger Jentzen	1871	704	922	1558
Samantha Pree-Stinson	656	1307	1744	1348
Steve Fletcher	1455	1878	1277	445
Tim Bildsoe	1073	1166	1112	1704

Fig. 4 First order marginals for the 2017 Minneapolis City Council Ward 3 election.

candidate i in ranking position j and candidate i' in ranking position j'. Higher order marginals that capture how many voters placed $k > 2$ specific candidates in k specific ranking positions may be totally unordered, totally ordered, or partially ordered (e.g., candidates i and i' are in positions j and j' in either order, and candidate i'' is in position j''). In all of these marginal tables, each row and each column sum to the total number of votes (5055 in Fig. 4 and Fig. 5).

Spectral analysis through projections onto the isotypic components captures order effects related to a specific marginal, net of the structure found in the "less complicated" marginals [29], [67, Sec. 2.6.1]. For example, Fig. 4 shows that of all first order effects, two of the most significant are that Steve Fletcher was far more likely than would be the case under a uniform distribution to be listed as the second choice, and far less likely to be listed as the fourth choice. These first order effects also appear in the second order marginals of Fig. 5, as well as higher order marginals. However, the projection of the signal $\mathbf{g}$ from Fig. 1 onto $W_{\boxplus}$ captures information about the second order unordered marginals net of the number of voters (zero order) and first order marginals. Similarly, the projection of $\mathbf{g}$ onto $W_{\boxminus}$ captures information about the second order ordered marginals, net of the zero and first order marginals and the second order unordered marginals.

More generally, the indices of orthogonal subspaces W_γ in (2) have a natural partial ordering, referred to as *dominance ordering*. The shape $\nu = [\nu_1, \nu_2, \ldots, \nu_\ell]$ strictly dominates shape $\gamma = [\gamma_1, \gamma_2, \ldots, \gamma_{\ell'}] \neq \nu$, denoted by $\nu \rhd \gamma$, if $\sum_{i=1}^{j} \nu_i \geq \sum_{i=1}^{j} \gamma_i$ for each $j = 1, 2, \ldots, \max\{\ell, \ell'\}$. For example, with $n = 6$, $\boxplus$ dominates

Candidates	First and Second	First and Third	First and Fourth	Second and Third	Second and Fourth	Third and Fourth
Ginger Jentzen and Samantha Pree-Stinson	951	673	903	525	535	1468
Ginger Jentzen and Steve Fletcher	1157	1307	862	588	837	304
Ginger Jentzen and Tim Bildsoe	467	813	1664	513	890	708
Samantha Pree-Stinson and Steve Fletcher	708	890	513	1664	813	467
Samantha Pree-Stinson and Tim Bildsoe	304	837	588	862	1307	1157
Steve Fletcher and Tim Bildsoe	1468	535	525	903	673	951

(a)

Candidate1	Candidate2	First-Second	First-Third	First-Fourth	Second-Third	Second-Fourth	Third-Fourth
Ginger Jentzen	Samantha Pree-Stinson	742	498	631	250	245	472
Ginger Jentzen	Steve Fletcher	797	793	281	242	102	62
Ginger Jentzen	Tim Bildsoe	332	580	959	212	357	388
Samantha Pree-Stinson	Ginger Jentzen	209	175	272	275	290	996
Samantha Pree-Stinson	Steve Fletcher	306	264	86	669	236	123
Samantha Pree-Stinson	Tim Bildsoe	141	217	298	363	781	625
Steve Fletcher	Ginger Jentzen	360	514	581	346	735	242
Steve Fletcher	Samantha Pree-Stinson	402	626	427	995	577	344
Steve Fletcher	Tim Bildsoe	693	315	447	537	566	691
Tim Bildsoe	Ginger Jentzen	135	233	705	301	533	320
Tim Bildsoe	Samantha Pree-Stinson	163	620	290	499	526	532
Tim Bildsoe	Steve Fletcher	775	220	78	366	107	260

(b)

Fig. 5 Second order marginals for the 2017 Minneapolis City Council Ward 3 election. (a) Second order unordered marginals. (b) Second order ordered marginals. Each row in (a) is the sum of the two rows in (b) with the same pair of candidates. In (b), we are only displaying half of the columns in the table, as, e.g., the "Second-First" column contains the same entries as the "First-Second" column in a different order. This is why the row sums are not equal to the number of voters in the displayed table.

⊞┐ and ⊞, but the latter two shapes are incomparable (neither one dominates the other). The projection of a signal onto an isotypic component W_γ captures information about the marginals corresponding to shape γ, net of the marginals corresponding to all shapes strictly preceding γ in dominance order.

With this justification, returning to the APA example in Fig. 3, Diaconis [29, Sec. 2C] interprets the relatively large contribution on isotypic component $W_{⊞}$, corresponding to the two-row shape $[3, 2]$, as the contribution of unordered pair (second order) effects, and he argues that this is related to the fact that the APA divides into two groups: academics and clinicians. This Fourier analysis approach has also found application in "Q-sort" data in psychology [28, 5B], balanced incomplete block designs [80], multiple object tracking [58], finding graph invariants [56], the quadratic assignment problem [54], computational geometry [48], genomic data analysis [96], and sports analytics [27].

What are the limitations of this approach? As shown in Fig. 6, two signals with different structure and support can have exactly the same energy decomposition into isotypic components, limiting the amount of information that can be extracted from this energy decomposition alone.

A more refined approach is to use the Fourier transform of $\mathbf{f}$ on the symmetric group, which is defined as the set of matrices $\{\hat{f}(\gamma) \mid \gamma$ a partition of $n\}$, where for each integer partition γ, $\hat{f}(\gamma)$ is defined as the sum $\sum_{\sigma \in \mathbb{S}_n} f(\sigma)\rho_\gamma(\sigma)$ [28]. Here, $\rho_\gamma(\sigma)$ is the matrix of the permutation σ as a linear transformation on the $\mathbb{S}_n$-invariant subspace $V_{\gamma,i}$ (for any i). This set $\{\hat{f}(\gamma) \mid \gamma$ a partition of $n\}$ has a total of $n!$ matrix entries, which are viewed as the Fourier coefficients of $\mathbf{f}$. Unfortunately,

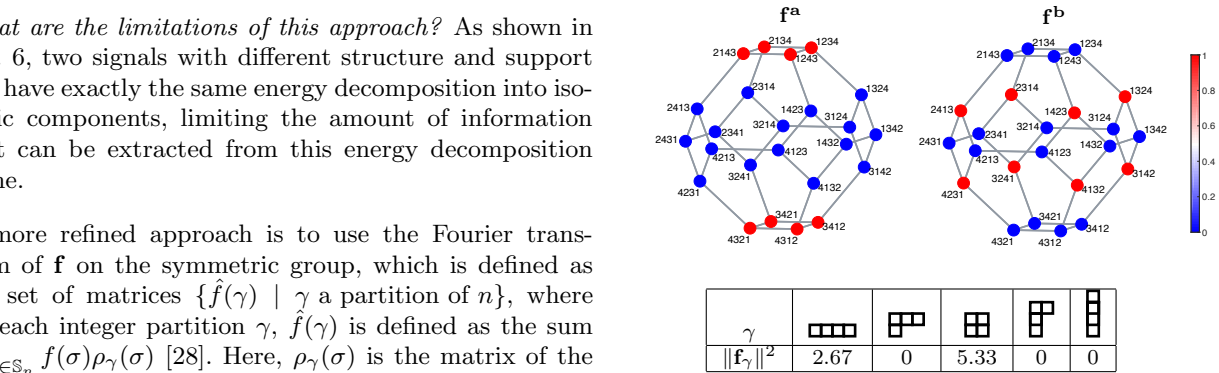| γ	⊞⊞⊞⊞	⊞┐	⊞	⊞┐	⊟
$\|\mathbf{f}_\gamma\|^2$	2.67	0	5.33	0	0

Fig. 6 Two signals on $\mathbb{P}_4$ with the same energy decomposition into isotypic components; i.e., $\|\mathbf{f}^{\mathbf{a}}_\gamma\| = \|\mathbf{f}^{\mathbf{b}}_\gamma\| := \|\mathbf{f}_\gamma\|$ for all γ.

there is no natural choice of basis of the irreducible components $V_{\gamma,i}$ (see for example Diaconis [29, p. 955]), and therefore ad-hoc methods are used to interpret these values. The standard choices are Young's seminormal basis or Young's orthogonal basis [28, 8A], [19], [30], [57], which are used largely because they are well-adapted to fast computation, but they lack interpretability.

Alternatively, [29, Sec. 2C] applies Mallow's method: construct an overcomplete spanning set for W_γ by projecting interpretable functions that capture kth order effects ($k = \sum_{i=2}^{\ell} \gamma_i$) onto W_γ, and then take inner products between these projections and the signal. In Fig. 7, we show two of the thirty-six interpretable second order functions and their projections in the overcomplete spanning set for $W_{⊞}$. In general, there are m_γ^2 of these spanning vectors for the d_γ^2-dimensional space W_γ, where $m_\gamma = \frac{n!}{\prod_{l=1}^{\ell} \gamma_l!}$.
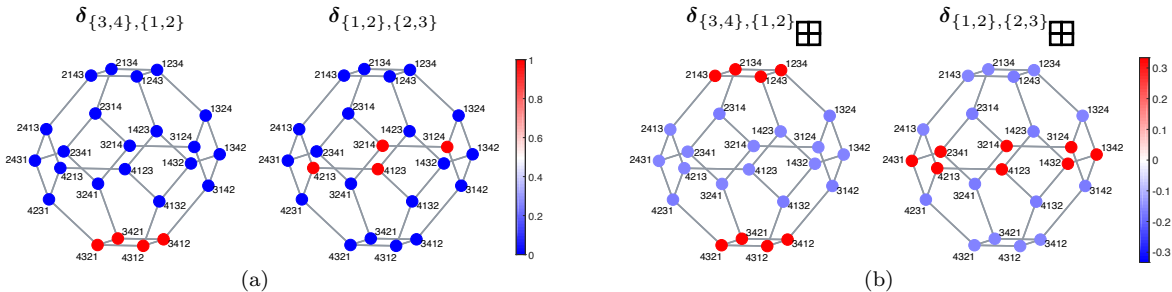

Fig. 7 Examples of Mallows' interpretable second order (unordered) functions [29, Sec. 2C]. (a) Two (of the thirty-six) functions of the form $\boldsymbol{\delta}_{\{i,i'\},\{j,j'\}}$, that are equal to 1 on the vertices where candidates i and i' are in ranking positions j and j' (in either order), and equal to 0 elsewhere. (b) The orthogonal projections of these interpretable functions onto the isotypic component $W_{⊞}$. Diaconis [29, Sec. 2C] suggests to take the inner product between each of these projected interpretable functions and the signal to identify second order effects in the data, net of the zero order and first order effects found in $W_{⊞⊞⊞⊞}$ and $W_{⊞┐}$.

For this particular isotypic component $W_{\boxplus}$, nine projected functions each appear four times in the thirty-six vectors; for example,

$$\boldsymbol{\delta}_{\{3,4\},\{1,2\}_{\boxplus}} = \boldsymbol{\delta}_{\{1,2\},\{3,4\}_{\boxplus}} = \boldsymbol{\delta}_{\{3,4\},\{3,4\}_{\boxplus}} = \boldsymbol{\delta}_{\{1,2\},\{1,2\}_{\boxplus}}. \tag{4}$$

For the 2017 Minneapolis City Council Ward 3 election data $\mathbf{g}$ shown in Fig. 1, the largest inner products

$$\langle \mathbf{g}, \boldsymbol{\delta}_{\{i,i'\},\{j,j'\}_{\boxplus}} \rangle = \langle \mathbf{g}_{\boxplus}, \boldsymbol{\delta}_{\{i,i'\},\{j,j'\}} \rangle \tag{5}$$

are the ones between the signal and $\boldsymbol{\delta}_{\{3,4\},\{1,2\}_{\boxplus}}$ and the three identical projections listed in (4). This can be seen visually by thinking about the inner products in the form listed in the right-hand side of (5) and then examining $\mathbf{g}_{\boxplus}$ in Fig. 3. Again, an interpretation of these inner products is that net of zero and first order marginal effects (i.e., starting from $\mathbf{g} - \mathbf{g}_{\boxminus} - \mathbf{g}_{\boxminus}$), the pairs of candidates $\{1,2\}$ and $\{3,4\}$ are likely to appear together in ranking positions $\{1,2\}$ or $\{3,4\}$. See [96] for an additional application of this method to genomic data.

The new approach we propose in Sec. 4 also yields an overcomplete spanning set, but we directly construct each spanning vector as an interpretable function in $V_{\gamma,i}$ (and in fact in a more precisely defined eigenspace that is a strict subset of $V_{\gamma,i}$).

Finally, it is noteworthy that the non-commutative Fourier analysis approach does not make any direct use of the permutahedron or any other underlying graph structure for that matter. Rather, it is completely independent of the choice of generating set of the symmetric group. We return to this point at the end of Sec. 3.2.

3.2 The Graph Signal Processing Approach

Within the last ten years, researchers in the field of graph signal processing [88, 75] have developed new methods to identify and exploit structure in data residing on the vertices of weighted or unweighted graphs. Signals on the unweighted permutahedron, as shown in Fig. 1, fit into this framework. While, to our knowledge, such ranked data on the permutahedron have not been analyzed with graph signal processing techniques, the natural first method to apply would be the *graph Fourier transform*. The graph Laplacian matrix is defined as $\mathcal{L} := \mathbf{D} - \mathbf{A}$, where $\mathbf{A}$ is the adjacency matrix of the permutahedron and the diagonal degree matrix $\mathbf{D}$ is equal to $(n-1)\mathbf{I}_{n!}$, where $\mathbf{I}_k$ is a $k \times k$ identity matrix, since the permutahedron is an $(n-1)$-regular graph. The matrix $\mathcal{L}$ is symmetric with nonnegative eigenvalues, and each eigenvalue λ is associated with an orthogonal eigenspace U_λ. The signal space $\mathbb{R}[\mathbb{S}_n]$ decomposes into a direct sum of these orthogonal eigenspaces,

$$\mathbb{R}[\mathbb{S}_n] \cong \bigoplus_\lambda U_\lambda. \tag{6}$$

A common definition of the graph Fourier transform is $\hat{f}(\lambda_\ell) := |\langle \mathbf{f}, \mathbf{u}_\ell \rangle|$, where $\mathbf{u}_\ell$ is the eigenvector associated with the ℓth eigenvalue of $\mathcal{L}$ [88]. Since this definition depends on the particular choice of the Laplacian eigenvectors in the case of repeated eigenvalues (which occur in $\mathbb{P}_n$), we define the graph Fourier transform here as $\hat{f}(\lambda) := \|\mathbf{f}_\lambda\|$, where $\mathbf{f}_\lambda$ is the orthogonal projection of $\mathbf{f}$ onto the eigenspace U_λ. In the absence of repeated eigenvalues, these two definitions coincide.
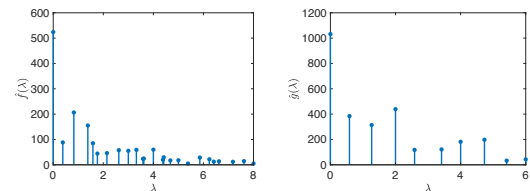

Fig. 8 Graph Fourier transforms of the two signals from Fig. 1. For each, more energy is concentrated on the smoother eigenvectors associated with the lower Laplacian eigenvalues of the corresponding permutahedron.

How can the graph Fourier transform and other graph signal processing techniques be used to find structure in ranked data? For each unit norm Laplacian eigenvector $\mathbf{u}_\lambda$ associated with eigenvalue λ, we have

$$\lambda = \mathbf{u}_\lambda^\top \mathcal{L} \mathbf{u}_\lambda = \sum_{(i,j)\in\mathcal{E}} [u_\lambda(i) - u_\lambda(j)]^2, \tag{7}$$
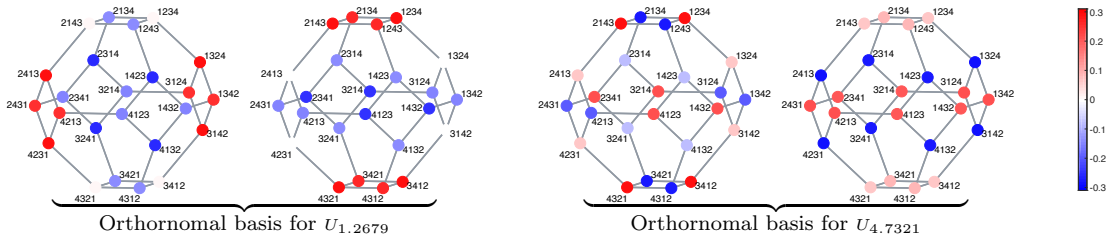
Fig. 9 Orthonormal bases for the graph Laplacian eigenspaces of $\mathbb{P}_4$ associated with the eigenvalues $\lambda = 1.2679$ and $\lambda = 4.7321$. Both of these eigenspaces are two-dimensional subspaces of the 24-dimensional signal space. Note that the Laplacian eigenvectors associated with the lower eigenvalue are smoother in the sense that the values vary less across neighboring vertices (see, e.g., the jumps in value from 1234 to 1243 in the third vector above or from 1324 to neighboring 1342 in the fourth vector).

where $\mathcal{E}$ are the edges of the permutahedron. Therefore, the eigenvectors associated with lower Laplacian eigenvalues are *smoother* in the sense that the values vary less across neighboring vertices, as shown in Fig. 9. The graph Fourier transform provides a decomposition of the energy of the signal into the energy in each Laplacian eigenspace, yielding information about the signal's smoothness, as shown in Fig. 8.

What are the limitations of applying the standard graph Fourier transform to signals on the permutahedron? First, because the graph Laplacian eigenvalues and eigenvectors of the permutahedron are not known in closed form, it is computationally intensive to compute the graph Fourier transform ($\mathcal{O}(n!^3)$), and not tractable for ranked data with more than seven or eight candidates. Second, there is not a natural orthonormal basis for each eigenspace that preserves the structure and symmetry of the graph. All but the first and last Laplacian eigenvalues of the permutahedron are repeated multiple times, and thus there are infinitely many choices of orthonormal bases for the associated eigenspaces $\{U_\lambda\}_{\lambda \notin \{0, 2(n-1)\}}$. As shown in Fig. 9 for two Laplacian eigenspaces of $\mathbb{P}_4$, the numerical computation of a basis is not guaranteed to preserve any sort of symmetry, leading to less interpretable basis vectors. Moreover, and third, different isotypic components may contain the same Laplacian eigenvalue, and thus, it is not even guaranteed that the numerically computed basis vectors live in a single isotypic component.

Why is the permutahedron the right graph to represent the underlying data domain? An alternative choice is the Cayley graph of the symmetric group induced by the generating set of all transpositions (not just neighboring transpositions), which is shown in Fig. 10 and which we denote by Γ_n. This graph has some nice mathematical properties: (i) the isotypic components W_γ are each spanned by eigenvectors associated with a single Laplacian eigenvalue of the Cayley graph Γ_n [58, Prop. 1]; (ii) the Laplacian eigenvalues and eigenvectors of Γ_n are known in closed form [81, Thm. 1.1], [58, Prop. 2], [40, Thm. III.1]; and, (iii) moreover, the Laplacian eigenvalues to which the isotypic components correspond increase according to dominance ordering. So if $\nu \rhd \gamma$, then any vector $\mathbf{f}_\nu$ in the isotypic component W_ν is smoother with respect to the Cayley graph Γ_n (recall the definition of smoothness in (7)) than any vector $\mathbf{f}_\gamma$ in W_γ with

Fig. 10 The Cayley graph Γ_n on the symmetric group induced by the generating set all transpositions, for $n = 4$.

the same norm as $\mathbf{f}_\nu$, since $\frac{\mathbf{f}_\nu^\top \mathcal{L}_{\Gamma_n} \mathbf{f}_\nu}{\mathbf{f}_\nu^\top \mathbf{f}_\nu} = \lambda_\nu < \lambda_\gamma = \frac{\mathbf{f}_\gamma^\top \mathcal{L}_{\Gamma_n} \mathbf{f}_\gamma}{\mathbf{f}_\gamma^\top \mathbf{f}_\gamma}$. In this sense, the isotypic components provide a notion of frequency, with vectors residing in isotypic components later in the dominance ordering representing "more complex" (less smooth) functions, *with respect to the Cayley graph Γ_n induced by the generating set of all transpositions* [58].
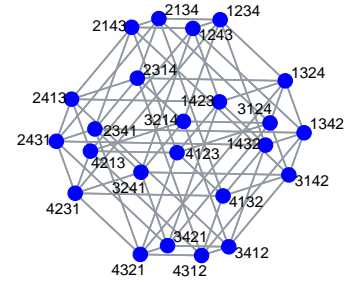
Despite these nice mathematical properties, the permutahedron is the more appropriate domain on which to develop new techniques for analyzing the structure of most ranked data sets, due to the different notions of distance the two underlying graphs capture [55]. The structure of Γ_n captures an appropriate notion of distance between the permutations in applications such as multi-object tracking, where, e.g., the object trajectories (slots) are continuously visible on radar or camera, but the object identities (candidates) associated with each trajectory are only revealed at certain time instances (e.g., pilot reports by radio, observations captured by a security camera) [58]. In this situation, there is not necessarily a physically-meaningful

linear order to the trajectories (slots), and therefore it may be likely that objects jump from one trajectory to a crossing trajectory whose label is not adjacent or even similar. This information is typically captured by a corresponding noise model (e.g., [58, Sec. 3.1]). However, in most ranked data applications, the ranking positions represent a linear ordering, and therefore permutations that swap the candidates in the first and last ranking positions, e.g., are not close from a voter's viewpoint. To illustrate this with a specific example, 1234 and 4231 are adjacent in the Cayley graph of Fig. 10, but they are far apart from a voter's perspective and they are far apart in the permutahedron $\mathbb{P}_4$, whereas 1234 and 2134 are adjacent in both graphs. Kondor [55, Sec. 3] distinguishes these cases in terms of invariance, with our notion being only right-invariant and the other notion being bi-invariant (right-invariant and left-invariant). As shown in Fig. 11, the graph structures induced by these two different distance metrics (generating sets) yield different notions of signal smoothness, as captured by the respective graph Fourier transforms. To summarize the key takeaways, when we choose an underlying graph data domain, we are defining a notion of distance between permutations; and the distance induced

Fig. 11 Graph Fourier transforms of the two signals from Fig. 6, on Cayley graphs of the symmetric group induced by two different choices of the generating set. On the Cayley graph Γ_4 of Fig. 10, the graph Fourier transforms of the two signals $\mathbf{f^a}$ and $\mathbf{f^b}$ are the same, as they have the same energy decomposition across isotypic components (c.f., Fig. 6). On the other hand, the graph Fourier transforms in the top row show that $\mathbf{f^a}$ is smoother than $\mathbf{f^b}$ with respect to the permutahedron $\mathbb{P}_4$.

by the permutahedron structure is most appropriate in ranking applications where permutations should be considered closest if the candidate swap occurs across neighboring ranking slots.
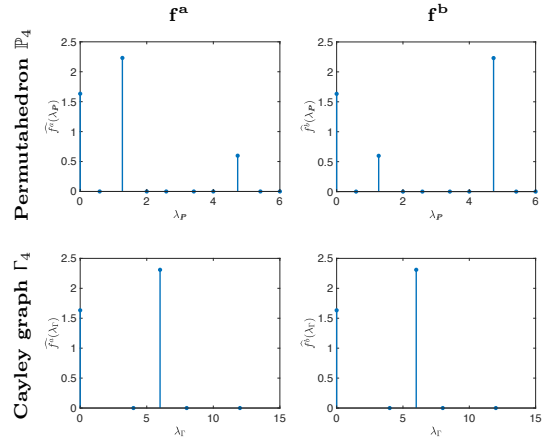
3.3 Other Related Transforms for Ranked Data

We briefly mention other related linear transforms for ranked data. Nested orthogonal contrasts [66, 67] compare the rankings of two or more groups of candidates, ignoring the relative ranks within each group of candidates. Inversions [70, 71, 67, 42] project the data onto linear subspaces based on the relative rankings of subsets (pairs, triplets, etc.) of the candidates, net of the effects of lower order subsets.

Other ranked data transforms consider the underlying graph to be a quasi-Abelian Cayley graph; i.e., the set of generators of the Cayley graph is the union of conjugacy classes. Rockmore [81] investigates fast Fourier transforms for data on quasi-Abelian Cayley graphs. Ghandehari et al. [40] extend Rockmore's work by developing tight windowed Fourier frames for data residing on quasi-Abelian Cayley graphs. Since the full set of transpositions is a conjugacy class in $\mathbb{S}_n$, the Cayley graph Γ_n induced by the generating set of all transpositions is quasi-Abelian, but the permutahedron $\mathbb{P}_n$ does not fall into this class (except for $\mathbb{P}_2$). Kondor's left-invariant coset-based multiresolution analysis [57] yields an orthogonal basis of wavelet and scaling atoms that are localized in the vertex and spectral domains of Γ_n, but not necessarily in either domain when the underlying graph is taken to be the permutahedron $\mathbb{P}_n$.

Finally, taking inspiration more from one-dimensional signal processing than from the literature on signal processing on graphs, Kakarala [49] further decomposes the coefficient matrices $\hat{f}(\gamma)$ of the Fourier transform on the symmetric group into the product of a positive semidefinite "magnitude" matrix and an orthogonal "phase" matrix.

4 Tight Spectral Frames for Ranked Data

In this section, we present a new approach to generate dictionaries for ranked data by combining the symmetry decomposition method (3) from combinatorial representation theory with the spectral graph decomposition method (6) from graph signal processing, in order to capture two different kinds of information about the data.

These two approaches are connected in the following way, which is at the crux of our method. The graph Laplacian matrix is in the regular representation of the symmetric group algebra, because it can be written as the linear combination

$$\mathcal{L} = (n-1)\rho_R(\mathbf{1}) - \sum_{s \in S} \rho_R(s), \tag{8}$$

where $\rho_R(s)$ is the matrix of $s \in S$ acting as a linear transformation on the right of $\mathbb{R}[\mathbb{S}_n]$ and $\mathbf{1} \in \mathbb{S}_n$ is the identity element so that $\rho_R(\mathbf{1}) = \mathbf{I}_{n!}$. Since our generators are involutions ($s^2 = 1$), the Laplacian is symmetric, and (8) is also true if the right regular representation is replaced with the left regular representation and $\mathcal{L}$ acts on the left. As a result of (8), the isotypic components in (3) decompose into Laplacian eigenspaces, and therefore the group algebra decomposes as follows.

Proposition 1

$$\mathbb{R}[\mathbb{S}_n] \cong \bigoplus_{\gamma \vdash n} \bigoplus_{\lambda \in \Lambda_\gamma} Z_{\gamma,\lambda}, \qquad where \quad Z_{\gamma,\lambda} = W_\gamma \cap U_\lambda, \tag{9}$$

and Λ_γ is the set of Laplacian eigenvectors λ such that $W_\gamma \cap U_\lambda \neq \emptyset$.

Proof The group algebra decomposes into a direct sum of vector spaces in two ways,

$$\mathbb{R}[\mathbb{S}_n] \cong \bigoplus_{\lambda \in \Lambda} U_\lambda \qquad and \qquad \mathbb{R}[\mathbb{S}_n] \cong \bigoplus_{\gamma \vdash n} W_\gamma.$$

On the left we use the fact that $\mathcal{L}$ is symmetric to decompose $\mathbb{R}[\mathbb{S}_n]$ into a direct sum of Laplacian eigenspaces $U_\lambda, \lambda \in \Lambda$, where Λ is the set of eigenvalues of $\mathcal{L}$. On the right is the decomposition (3) into isotypic components for $\mathbb{S}_n$. The isotypic components are closed under both the left and right action of $\mathbb{S}_n$ and $\mathcal{L}$ is a linear combination (8) of group elements, so the isotypic components W_γ are closed under the action of $\mathcal{L}$.

If $\mathbf{f} \in \mathbb{R}[\mathbb{S}_n]$, then there is a unique decomposition $\mathbf{f} = \sum_{\lambda \in \Lambda} \mathbf{f}_\lambda$ with $\mathbf{f}_\lambda \in U_\lambda$; namely, $\mathbf{f}_\lambda$ is the orthogonal projection of $\mathbf{f}$ onto U_λ. For each $\lambda \in \Lambda$, there is a unique decomposition, $\mathbf{f}_\lambda = \sum_{\{\gamma \vdash n: \lambda \in \Lambda_\gamma\}} \mathbf{f}_{\gamma,\lambda}$, of $\mathbf{f}_\lambda$ into vectors $\{\mathbf{f}_{\gamma,\lambda}\}$ in the corresponding isotypic components $\{W_\gamma\}$. Multiplying $\mathbf{f}_\lambda$ by λ gives

$$\lambda \mathbf{f}_\lambda = \sum_{\gamma \vdash n: \lambda \in \Lambda_\gamma} \lambda \mathbf{f}_{\gamma,\lambda}, \tag{10}$$

and multiplying $\mathbf{f}_\lambda$ by $\mathcal{L}$ gives

$$\lambda \mathbf{f}_\lambda = \mathcal{L} \mathbf{f}_\lambda = \sum_{\gamma \vdash n: \lambda \in \Lambda_\gamma} \mathcal{L} \mathbf{f}_{\gamma,\lambda}. \tag{11}$$

Since W_γ is closed under multiplication by $\mathcal{L}$, each term $\mathcal{L}\mathbf{f}_{\gamma,\lambda}$ in the summation in (11) is in W_γ. By the uniqueness of the decomposition of $\mathbf{f}$ into isotypic components, we conclude from (10) and (11) that $\mathcal{L}\mathbf{f}_{\gamma,\lambda} = \lambda \mathbf{f}_{\gamma,\lambda}$ for each $\gamma \vdash n$ such that $\lambda \in \Lambda_\gamma$. Thus, $\mathbf{f}_{\gamma,\lambda}$ is a Laplacian eigenvector of eigenvalue λ and $\mathbf{f}_{\gamma,\lambda} \in Z_{\gamma,\lambda} = W_\gamma \cap U_\lambda$. In summary, for any $\mathbf{f} \in \mathbb{R}[\mathbb{S}_n]$, there is a unique decomposition $\mathbf{f} = \sum_{\gamma \vdash n} \sum_{\lambda \in \Lambda_\gamma} \mathbf{f}_{\gamma,\lambda}$, with $\mathbf{f}_{\gamma,\lambda} \in Z_{\gamma,\lambda}$. $\square$

If $\lambda \in \Lambda_\gamma$, we say that the eigenvalue λ has *symmetry type* or *shape* γ. Eigenvalues may have multiple symmetry types. For example, the Laplacian eigenvalue $\lambda = 3$ on $\mathbb{P}_6$ is repeated 15 times (i.e., U_3 is a 15-dimensional space). This eigenvalue appears 5 times in the $\gamma = $ ⊞⊞⊞⊞ component and 10 times in the $\gamma = $ ⊞⊞⊞ component (e.g., $W_{⊞⊞⊞} \cap U_3$ is a 10-dimensional space), so $\lambda = 3$ has two symmetry types. On the other hand, $\lambda = 2$ has a single symmetry type as it only appears in the $\gamma = $ ⊞⊞⊞⊞⊞ component.

Our objective is to find a spanning set $\{\boldsymbol{\varphi}_{\gamma,\lambda,k,\pi}\}$ of dictionary atoms for each space $Z_{\gamma,\lambda}$ such that

(i) the overall dictionary analysis operator preserves the energy in the signal; that is, $\sum_{\gamma,\lambda,k,\pi} |\langle \mathbf{f}, \boldsymbol{\varphi}_{\gamma,\lambda,k,\pi} \rangle|^2 = \|\mathbf{f}\|_2^2$, or equivalently, $\|\boldsymbol{\Phi}^\top \mathbf{f}\|_2 = \|\mathbf{f}\|_2$, where the atoms $\{\boldsymbol{\varphi}_{\gamma,\lambda,k,\pi}\}$ comprise the columns of the matrix $\boldsymbol{\Phi}$;

(ii) the atoms are interpretable (i.e., they have a particular structure that makes the inner products useful in identifying structure in the data); and

(iii) we can efficiently compute the inner products between these dictionary atoms and the signal on the permutahedron.

4.1 Preliminaries: Schreier Graphs and Equitable Partitions

We start by detailing (i) how to construct, for each integer partition γ of n, a graph $\mathbb{P}_\gamma$, called a Schreier graph, which is isomorphic to a quotient of the permutahedron $\mathbb{P}_n$ [38,84]; and (ii) the relation between the spectral decompositions of the Schreier graphs and the spectral decomposition of the permutahedron.

If $\gamma = [\gamma_1, \ldots, \gamma_\ell]$ is an integer partition of n, then a set partition $\pi = \{C_1, \ldots, C_\ell\}$ has shape γ if its blocks have size $|C_i| = \gamma_i$. There are $m_\gamma = \frac{n!}{\prod_{i=1}^\ell \gamma_i!}$ different ordered set partitions π of $\{1, 2, \ldots, n\}$ of shape γ, and we refer to this collection of ordered partitions as Π_γ.[2] A permutation $\sigma \in \mathbb{S}_n$ acts on an ordered set partition $\pi \in \Pi_\gamma$ by permuting the entries of π. For example if $\sigma = \left(\begin{smallmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 4 & 3 & 1 \end{smallmatrix}\right)$ and $\pi = \{\{2, 4, 5\}, \{1, 3\}\}$ then $\sigma(\pi) = \{\{1, 3, 5\}, \{2, 4\}\}$.

Definition 1 The *Schreier graph* $\mathbb{P}_\gamma$ is the graph with vertex set Π_γ and edge set $\{(\pi, s\pi) \mid \pi \in \Pi_\gamma, s \in S\}$, where $S \subseteq \mathbb{S}_n$ is the subset of adjacent transpositions defined in Sec. 1.
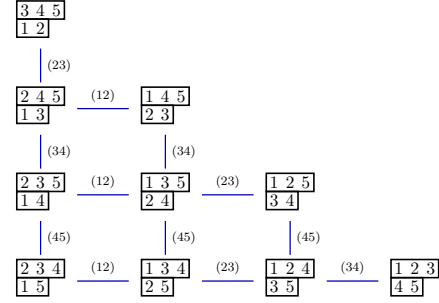

Fig. 12 The Schreier graph $\mathbb{P}_{[3,2]}$. In this graph, the vertices are the $m_{[3,2]} = 5!/(3!2!) = 10$ ordered set partitions of $\{1, 2, 3, 4, 5\}$ into subsets of size 3 and 2. We use row-strict tableaux to denote set partitions, and interpret the labels as groupings of rankings. For example, $\{\{1, 2, 4\}, \{3, 5\}\}$ is denoted by $\boxed{\begin{smallmatrix} 1 & 2 & 4 \\ 3 & 5 \end{smallmatrix}}$ and groups the third and fifth place rankings together, and the other three places together. The edge labels correspond to adjacent transpositions. For example, the (34) label on the edge between $\boxed{\begin{smallmatrix} 1 & 2 & 4 \\ 3 & 5 \end{smallmatrix}}$ and $\boxed{\begin{smallmatrix} 1 & 2 & 3 \\ 4 & 5 \end{smallmatrix}}$ represents a swapping of the third and fourth place rankings in the partition blocks. We omit self loops when drawing Schreier graphs, as they do not affect the corresponding graph Laplacian operators.

Each Schreier graph (i) is undirected since adjacent transpositions are involutions; (ii) is $(n-1)$-regular since $|S| = n - 1$; and (iii) has a self-loop at vertex π for each pair $i, i+1$ that is in the same block of π, since then the transposition $s = (i, i+1)$ fixes π. Two ordered set partitions $\pi_1 \neq \pi_2$ are connected by the edge labeled by $s = (i, i+1)$ if and only if π_1 and π_2 are identical except with i and $i+1$ switched. Thus, non-loop edge weights are equal to 1. We view the vertices of $\mathbb{P}_\gamma$ as representing groupings of rankings (first place, second place, etc.), not specific candidates. Fig. 12 shows the example of $\mathbb{P}_\gamma$ with $\gamma = [3, 2]$.

Next, we show that each ordered set partition π induces an equivalence relation on $\mathbb{S}_n$, the vertices of the permutahedron $\mathbb{P}_n$, and under this equivalence relation, the Schreier graphs are isomorphic to quotient graphs of the permutahedron $\mathbb{P}_n$.

Definition 2 Let $\pi \in \Pi_\gamma$ be an ordered set partition of shape γ, and let $\sigma, \tau \in \mathbb{S}_n$. The equivalence relation $\sim_\pi$ is given by identifying $\sigma \sim_\pi \tau$ if and only if $\sigma^{-1}(\pi) = \tau^{-1}(\pi)$. The equivalence classes under $\sim_\pi$ are the sets $\mathcal{V}_{\pi,\mu} = \{\sigma \in \mathbb{S}_n \mid \sigma(\mu) = \pi\}$ for each $\mu \in \Pi_\gamma$. These are the permutations that place candidates π in the positions given by μ.

For example, if $\pi = \{\{2, 4, 5\}, \{1, 3\}\}$ (or $\{245|13\}$ for shorter notation), then $34512 \sim_\pi 12435$ because each group of candidates, $\{2, 4, 5\}$ and $\{1, 3\}$, is in the same set of positions in the two permutations: $3\underline{45}1\underline{2}$ and $1\underline{2}4\underline{35}$. Furthermore, the equivalence class containing these two permutations is the set $\mathcal{V}_{\{245|13\},\{235|14\}}$ consisting of all permutations with $\{2,4,5\}$ and $\{1,3\}$ in positions $\{2,3,5\}$ and $\{1,4\}$, respectively.

Proposition 2 *The partition of the vertices of $\mathbb{P}_n$ into equivalence classes $\{\mathcal{V}_{\pi,\mu} \mid \mu \in \Pi_\gamma\}$ induced by $\sim_\pi$ is an* equitable partition, *meaning that for every pair of (not necessarily distinct) ordered set partitions $\mu, \nu \in \Pi_\gamma$, there is a nonnegative integer $\mathbf{K}_\pi(\mu, \nu)$ such that each vertex $\sigma \in \mathcal{V}_{\pi,\mu}$ has exactly $\mathbf{K}_\pi(\mu, \nu)$ neighbors in $\mathcal{V}_{\pi,\nu}$ [41, Sec. 9.3]. In fact, when $\mu \neq \nu$, $\mathbf{K}_\pi(\mu, \nu) = 1$ if there exists $s \in S$ such that $s(\mu) = \nu$ and equals 0 otherwise, and $\mathbf{K}_\pi(\mu, \mu)$ equals the number of $s \in S$ such that $s(\mu) = \mu$.*

---

[2] By ordered set partitions, we mean that changing the ordering of the blocks results in a different partition, but changing the ordering within blocks does not. For example, $\pi = \{\{1, 2\}, \{3, 4\}\}$ and $\pi' = \{\{3, 4\}, \{1, 2\}\}$ are distinct elements of $\Pi_{[2,2]}$. However, $\pi'' = \{\{2, 1\}, \{3, 4\}\}$ is equivalent to π.
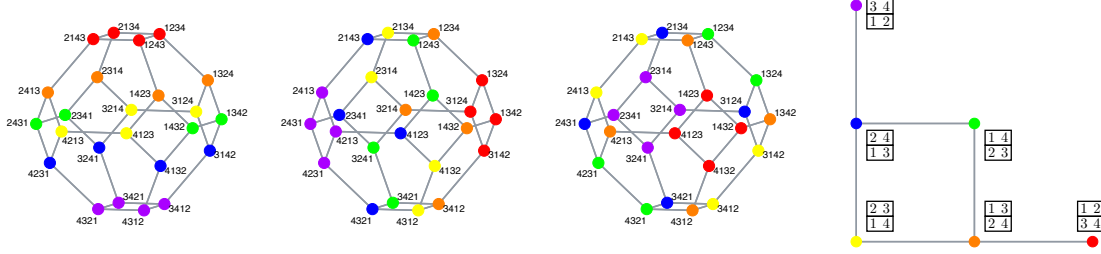
Fig. 13 Right: The Schreier graph $\mathbb{P}_{\boxplus} = \mathbb{P}_{[2,2]}$ (shown without self loops). Recall from Fig. 12 that we view the vertices of the Schreier graph as ordered set partitions of *ranking slots*. Left: Three different equitable partitions of $\mathbb{P}_4$ that yield this Schreier graph, corresponding to the ordered set partitions $\pi_1 = \{\{1,2\},\{3,4\}\}$, $\pi_2 = \{\{1,3\},\{2,4\}\}$, and $\pi_3 = \{\{1,4\},\{2,3\}\}$, respectively, from left to right. We view the three equitable partitions on the left as set partitions of *candidates*. For each equitable partition, all vertices of the permutahedron with the same color are mapped to the node of the corresponding color on the Schreier graph; this mapping is encoded in the rows of the characteristic matrix $\mathbf{B}_{\pi_i}$. For each equitable partition, the two pairs of candidates defined by the ordered set partition π_i are in the ranking slots defined by the two rows of the tableaux with the corresponding color in the Schreier graph on the right. For example, in the middle equitable partition π_2, the red vertices correspond to rankings where candidates 1 and 3 are in ranking positions 1 and 2, and candidates 2 and 4 are in positions 3 and 4.

Proof When the equivalence class $\mathcal{V}_{\pi,\mu} = \{\sigma \in \mathbb{S}_n \mid \sigma(\mu) = \pi\}$ is right multiplied by $\tau \in \mathbb{S}_n$, we have $\mathcal{V}_{\pi,\mu}\tau = \{\sigma\tau \in \mathbb{S}_n \mid \sigma(\mu) = \pi\} = \{\sigma\tau \in \mathbb{S}_n \mid \sigma\tau(\tau^{-1}(\mu)) = \pi\} = \mathcal{V}_{\pi,\tau^{-1}(\mu)}$. It follows that if $s \in S$ and $\sigma \in \mathcal{V}_{\pi,\mu}$ then $\sigma s \in \mathcal{V}_{\pi,s(\mu)}$. Multiplication by a group element s is a bijection, so each element in $\mathcal{V}_{\pi,\mu}$ is connected by an edge in $\mathbb{P}_n$ to exactly one element in $\mathcal{V}_{\pi,s(\mu)}$. If there is not an adjacent transposition s such that $s(\mu) = \nu$, then there are no edges in $\mathbb{P}_n$ between the vertices in $\mathcal{V}_{\pi,\mu}$ and those in $\mathcal{V}_{\pi,\mu}$. □

	$\begin{smallmatrix}1\ 2\\3\ 4\end{smallmatrix}$	$\begin{smallmatrix}1\ 3\\2\ 4\end{smallmatrix}$	$\begin{smallmatrix}2\ 3\\1\ 4\end{smallmatrix}$	$\begin{smallmatrix}1\ 4\\2\ 3\end{smallmatrix}$	$\begin{smallmatrix}2\ 4\\1\ 3\end{smallmatrix}$	$\begin{smallmatrix}3\ 4\\1\ 2\end{smallmatrix}$
1234	·	1	·	·	·	·
1243	·	·	·	1	·	·
1324	1	·	·	·	·	·
1342	1	·	·	·	·	·
1423	·	·	·	1	·	·
1432	·	1	·	·	·	·
2134	·	·	1	·	·	·
2143	·	·	·	·	1	·
2314	·	·	1	·	·	·
2341	·	·	·	·	1	·
2413	·	·	·	·	·	1
2431	·	·	·	·	·	1
3124	1	·	·	·	·	·
3142	1	·	·	·	·	·
3214	·	1	·	·	·	·
3241	·	·	·	1	·	·
3412	·	1	·	·	·	·
3421	·	·	·	1	·	·
4123	·	·	·	·	1	·
4132	·	·	1	·	·	·
4213	·	·	·	·	·	1
4231	·	·	·	·	·	1
4312	·	·	1	·	·	·
4321	·	·	·	·	1	·

Fig. 14 The characteristic matrix $\mathbf{B}_{\pi_2}$ of the ordered set partition $\pi_2 = \{\{1,3\},\{2,4\}\}$ from Fig. 13. In the second row, e.g., candidates 1, 3 are in ranking slots 1, 4.

Proposition 3 *For each ordered set partition $\pi \in \Pi_\gamma$, the quotient graph $\mathbb{P}_n/\sim_\pi$ is isomorphic to $\mathbb{P}_\gamma$.*

Proof Since $\sim_\pi$ induces an equitable partition of $\mathbb{P}_n$ with equivalence classes $\mathcal{V}_{\pi,\mu}$, the quotient $\mathbb{P}_n/\sim_\pi$ is a well-defined graph with vertices $\{\mathcal{V}_{\pi,\mu} \mid \mu \in \Pi_\gamma\}$ and edges $\{(\mathcal{V}_{\pi,\mu}, \mathcal{V}_{\pi,s(\mu)}) \mid \mu \in \Pi_\gamma, s \in S\}$. Comparing Prop. 2 with Def. 1 shows that these graphs are isomorphic by identifying the vertex $\mathcal{V}_{\pi,\mu}$ in $\mathbb{P}_n/\sim_\pi$ with μ in $\mathbb{P}_\gamma$. □

Definition 3 The *characteristic matrix of an equitable partition π* of shape γ, denoted by $\mathbf{B}_\pi$, is the $n! \times m_\gamma$ $(0,1)$-matrix whose (σ, μ)th element, for $\sigma \in \mathbb{S}_n$ and $\mu \in \Pi_\gamma$, is equal to 1 if and only if $\sigma(\mu) = \pi$; that is σ places the candidates from the jth block of π into the positions corresponding to the jth block of μ, for each j.

Each row of $\mathbf{B}_\pi$ contains exactly one 1, and each column contains exactly $\frac{n!}{m_\gamma}$ 1s. Therefore $\mathbf{B}_\pi^\top \mathbf{B}_\pi = (\frac{n!}{m_\gamma})\mathbf{I}_{m_\gamma}$. This is illustrated in Fig. 14. Furthermore, $\mathbf{B}_\pi$ has the following symmetry property, which we use in our frame construction (15) and Thm. 1.

Proposition 4 *For each $\sigma \in \mathbb{S}_n$ and $\pi \in \Pi_\gamma$ we have $\rho_L(\sigma)\mathbf{B}_\pi = \mathbf{B}_{\sigma(\pi)}$, where $\rho_L(\sigma)$ is the matrix of σ in the left regular representation of $\mathbb{S}_n$ on $\mathbb{R}[\mathbb{S}_n]$.*

Proof For $\sigma, \tau \in \mathbb{S}_n$ and $\mu \in \Pi_\gamma$, the (τ, μ) entry of $\rho_L(\sigma)\mathbf{B}_\pi$ equals 1 if and only if $\sigma^{-1}\tau(\mu) = \pi$, which is true if and only if $\tau(\mu) = \sigma(\pi)$, and this is exactly the condition for the (τ, μ) entry of $\mathbf{B}_{\sigma(\pi)}$. □

A key property of characteristic matrices of equitable partitions is that they can be used to lift eigenvectors of the Schreier graphs $\mathbb{P}_\gamma$ to eigenvectors of the permutahedron $\mathbb{P}_n$.

Proposition 5 *If $\pi \in \Pi_\gamma$ and $\mathbf{v}_{\gamma,\lambda}$ is a graph Laplacian eigenvector of the Schreier graph $\mathbb{P}_\gamma$ with eigenvalue λ, then $\mathbf{w} = \mathbf{B}_\pi \mathbf{v}_{\gamma,\lambda}$ is a graph Laplacian eigenvector of the permutahedron $\mathbb{P}_n$ with the same eigenvalue λ.*

Proof Let $\mathbf{A}_{\mathbb{P}_n}$ be the adjacency matrix of the permutahedron, and $\mathbf{A}_{\mathbb{P}_\gamma} = \mathbf{A}_{\mathbb{P}_n/\sim_\pi}$ be the adjacency matrix of the Schreier graph $\mathbb{P}_\gamma$. Since π induces an equitable partition $\sim_\pi$ of $\mathbb{P}_n$, $\mathbf{A}_{\mathbb{P}_n} \mathbf{B}_\pi = \mathbf{B}_\pi \mathbf{A}_{\mathbb{P}_\gamma}$ [41, Lem. 9.3.1], and thus,

$$\begin{aligned} \boldsymbol{\mathcal{L}}_{\mathbb{P}_n} \mathbf{B}_\pi &= (n-1)\mathbf{I}_{n!}\mathbf{B}_\pi - \mathbf{A}_{\mathbb{P}_n}\mathbf{B}_\pi \\ &= (n-1)\mathbf{B}_\pi \mathbf{I}_{m_\gamma} - \mathbf{B}_\pi \mathbf{A}_{\mathbb{P}_\gamma} = \mathbf{B}_\pi \boldsymbol{\mathcal{L}}_{\mathbb{P}_\gamma}, \end{aligned} \tag{12}$$

as both $\mathbb{P}_n$ and $\mathbb{P}_\gamma$ are (n-1)-regular graphs. Thus,

$$\boldsymbol{\mathcal{L}}_{\mathbb{P}_n} \mathbf{B}_\pi \mathbf{v}_{\gamma,\lambda} = \mathbf{B}_\pi \boldsymbol{\mathcal{L}}_{\mathbb{P}_\gamma} \mathbf{v}_{\gamma,\lambda} = \mathbf{B}_\pi(\lambda \mathbf{v}_{\gamma,\lambda}) = \lambda(\mathbf{B}_\pi \mathbf{v}_{\gamma,\lambda}),$$

where the first equality follows from (12).   □

For each shape γ we view signals on the Schreier graph $\mathbb{P}_\gamma$ as vectors in the vector space $\mathbb{R}[\Pi_\gamma]$ with canonical basis $\{\mathbf{e}_\pi\}_{\pi \in \Pi_\gamma}$. This vector space has a natural right $\mathbb{S}_n$-action defined on a basis element $\mathbf{e}_\pi$ and a permutation $\sigma \in \mathbb{S}_n$ by $\mathbf{e}_\pi \sigma = \mathbf{e}_{\sigma^{-1}(\pi)}$, where permutations act on set partitions in Π_γ by permuting the entries as described just before Def 2. The $\mathbb{S}_n$-module $\mathbb{R}[\Pi_\gamma]$ is known as the (right) "permutation module" for $\mathbb{S}_n$ and is often denoted by M^γ (see for example [83, Ch. 2]).

For each shape $\gamma \vdash n$, the permutation module $\mathbb{R}[\Pi_\gamma]$ decomposes into irreducible (right) $\mathbb{S}_n$-submodules according to

$$\mathbb{R}[\Pi_\gamma] \cong V_\gamma^* \oplus \bigoplus_{\nu \rhd \gamma} (V_\nu^*)^{\oplus K_{\gamma,\nu}}, \tag{13}$$

where again $\nu \rhd \gamma$ means that ν strictly dominates γ. Thus $\mathbb{R}[\Pi_\gamma]$ contains exactly one copy of the irreducible V_γ^* and $K_{\gamma,\nu}$ copies of each irreducible module that comes before it in dominance order on partitions. The multiplicities $K_{\gamma,\nu}$ are known as Kostka numbers (see [83, Sec. 2.11] and Fig. 29). Furthermore, V_γ^* does not appear as a submodule of $\mathbb{R}[\Pi_\pi]$ for partitions π whose shapes come before γ in dominance order, which is beneficial for the computational algorithms that we explore in Sec. 6.

The following proposition says that when a vector in $\mathbb{R}[\Pi_\gamma]$ that lives entirely in the submodule V_γ^* is lifted to the permutahedron by the characteristic matrix of an equitable partition of shape γ, the resulting vector resides in the single isotypic component $W_\gamma \subseteq \mathbb{R}[\mathbb{S}_n]$ defined in (2).

Proposition 6 *If $\pi \in \Pi_\gamma$ and $\mathbf{x} \in V_\gamma^* \subseteq \mathbb{R}[\Pi_\gamma]$, then $\mathbf{B}_\pi \mathbf{x} \in W_\gamma$.*

Proof The map $\mathbb{R}[\Pi_\gamma] \to \mathbb{R}[\mathbb{S}_n]$ given by $\mathbf{x} \mapsto \mathbf{B}_\pi \mathbf{x}$ is an injective, right $\mathbb{S}_n$-module homomorphism (i.e., it commutes with the right $\mathbb{S}_n$ action on the the two spaces). It is injective, since $\mathbf{B}_\pi$ has rank m_γ, and is an $\mathbb{S}_n$-module homomorphism, since for $\tau \in \mathbb{S}_n$ and $\pi, \mu \in \Pi_\gamma$ we have

$$\mathbf{B}_\pi(\mathbf{e}_\mu \tau) = \mathbf{B}_\pi(\mathbf{e}_{\tau^{-1}(\mu)}) = \sum_{\sigma, \sigma\tau^{-1}(\mu)=\pi} \mathbf{e}_\sigma = \sum_{\eta, \eta(\mu)=\pi} \mathbf{e}_{\eta\tau} = \sum_{\eta, \eta(\mu)=\pi} \mathbf{e}_\eta \tau = (\mathbf{B}_\pi \mathbf{e}_\mu)\tau.$$

Upon restriction to the irreducible submodule V_γ^*, by Schur's lemma, the map must be an isomorphism. Thus the image of V_γ^* under $\mathbf{x} \mapsto \mathbf{B}_\pi \mathbf{x}$ is an isomorphic copy of V_γ^*. The isotypic component W_γ contains all copies of V_γ^* in $\mathbb{R}[\mathbb{S}_n]$ so $\mathbf{B}_\pi \mathbf{x} \in W_\gamma$.   □

An important implication of Prop. 5 and Prop. 6 is that we can compute, visualize, and interpret Laplacian eigenvectors on the lower-dimensional Schreier graphs, and then lift them up to the higher-dimensional permutahedron graph in different manners – assigning different groups of candidates to groups of ranking slots – in order to generate vectors that reside in specific spaces $Z_{\gamma,\lambda}$ (i.e., have certain symmetry types and smoothness levels). Next, we show that scaled versions of vectors generated in this manner constitute a tight frame for the space $\mathbb{R}[\mathbb{S}_n]$ of all possible signals on the permutahedron $\mathbb{P}_n$.

4.2 Tight Frame Construction

Our strategy is to construct a tight Parseval frame for each nonempty space $Z_{\gamma,\lambda}$, and then let the dictionary $\boldsymbol{\Phi}$ be the union of these tight frames, so that $\boldsymbol{\Phi}$ is a tight Parseval frame for $\mathbb{R}[\mathbb{S}_n]$. A set of vectors $\{\boldsymbol{\varphi}_j\}$ is a *frame* for a Hilbert space $\mathcal{H}$ if there exists frame bounds (constants) $A, B > 0$ such that $A\|\mathbf{f}\|_2^2 \leq \sum_j |\langle \mathbf{f}, \boldsymbol{\varphi}_j \rangle|^2 \leq B\|\mathbf{f}\|_2^2$, $\forall \mathbf{f} \in \mathcal{H}$. A frame is said to be *tight* if $A = B$, and a *Parseval frame* if $A = B = 1$. For finite dimensional Hilbert spaces (such as our $Z_{\gamma,\lambda}$ spaces), a (finite) frame is simply a set of spanning vectors for the space. A frame is a *group frame* if there exits a finite group G that acts as linear transformations on $\mathcal{H}$ such that $\{\boldsymbol{\varphi}_j\} = \{g\boldsymbol{\varphi}_1\}_{g \in G}$; that is, the frame is generated by rotating a single frame vector $\boldsymbol{\varphi}_1$ by the group G. For more background on frames and group frames, and their use in signal processing and machine learning, see [17, 59, 60, 97].

Analogous to equation (8), the Laplacian matrix $\boldsymbol{\mathcal{L}}_\gamma$ on the Schreier graph $\mathbb{P}_\gamma$ equals

$$\boldsymbol{\mathcal{L}}_\gamma = (n-1)\rho_\gamma(\mathbf{1}) - \sum_{s \in S} \rho_\gamma(s), \tag{14}$$

where ρ_γ is the representation of $\mathbb{S}_n$ on $\mathbb{R}[\Pi_\gamma]$, which decomposes as in (13). It follows that V_γ^* is closed under multiplication by $\boldsymbol{\mathcal{L}}_\gamma$, and therefore, from the identical argument as in the proof of Prop.1, V_γ^* decomposes into eigenspaces for $\boldsymbol{\mathcal{L}}_\gamma$.

For each shape $\gamma \vdash n$ and eigenvalue $\lambda \in \Lambda_\gamma$, to generate a tight Parseval frame $\Phi_{\gamma,\lambda}$ for $Z_{\gamma,\lambda}$, we

(i) construct an orthonormal basis $\{\mathbf{v}_{\gamma,\lambda,k}\}_{k=1}^{\kappa_{\gamma,\lambda}}$ for the graph Laplacian eigenspace of the Schreier graph $\mathbb{P}_\gamma$ that is associated with the eigenvalue λ in V_γ^*, and then
(ii) lift each eigenvector $\mathbf{v}_{\gamma,\lambda,k}$ in the basis back to the permutahedron $\mathbb{P}_n$ in m_γ different ways.

Specifically, we define $\boldsymbol{\varphi}_{\gamma,\lambda,k,\pi} := c_\gamma \mathbf{B}_\pi \mathbf{v}_{\gamma,\lambda,k}$, and

$$\Phi_{\gamma,\lambda} := \bigcup_{k=1}^{\kappa_{\gamma,\lambda}} \bigcup_{\pi \in \Pi_\gamma} \boldsymbol{\varphi}_{\gamma,\lambda,k,\pi} = \bigcup_{k=1}^{\kappa_{\gamma,\lambda}} \bigcup_{\pi \in \Pi_\gamma} c_\gamma \mathbf{B}_\pi \mathbf{v}_{\gamma,\lambda,k}, \tag{15}$$

where the constant $c_\gamma := \sqrt{\frac{d_\gamma}{n!}}$ with $d_\gamma = \dim(V_\gamma)$. It may be the case that some of the liftings $\mathbf{B}_\pi \mathbf{v}_{\gamma,\lambda,k}$ are equal for different $\pi \in \Pi_\gamma$, but we keep these multiple copies in $\Phi_{\gamma,\lambda}$ (viewing it as a multiset), so that $|\Phi_{\gamma,\lambda}| = \kappa_{\gamma,\lambda} m_\gamma$. We remove some of these redundancies in the reduced frame $\bar{\Phi}_{\gamma,\lambda}$ in (22) below.

Remark 1 It is often but not always the case that $\kappa_{\gamma,\lambda} = 1$, so that the basis $\{\mathbf{v}_{\gamma,\lambda,k}\}_{k=1}^{\kappa_{\gamma,\lambda}}$ consists of a single vector. When $\kappa_{\gamma,\lambda} = 1$, we shorten the notation from $\mathbf{v}_{\gamma,\lambda,k}$ to $\mathbf{v}_{\gamma,\lambda}$, and from $\boldsymbol{\varphi}_{\gamma,\lambda,k,\pi}$ to $\boldsymbol{\varphi}_{\gamma,\lambda,\pi}$. To make interpretations more consistent, in our implementations, we always choose the eigenvectors of the Schreier graphs to have norm 1 and a positive coefficient on the vertex associated with ordered set partition μ that is last in lexicographic order (e.g., $\boxed{\begin{smallmatrix} 3 & 4 \\ 1 & 2 \end{smallmatrix}}$ in Fig. 15).

Theorem 1 *For $\gamma \vdash n$ and $\lambda \in \Lambda_\gamma$, the collection of atoms $\Phi_{\gamma,\lambda}$ defined in (15) is the union of of $\kappa_{\gamma,\lambda}$ orthogonal tight Parseval frames and, as such, is a tight Parseval frame for $Z_{\gamma,\lambda}$. The set of atoms $\mathcal{D} := \bigcup_{\gamma \vdash n} \bigcup_{\lambda \in \Lambda_\gamma} \Phi_{\gamma,\lambda}$ is a tight Parseval frame for $\mathbb{R}[\mathbb{S}_n]$.*

Proof Let $\mathbf{v} \in V_\gamma^* \subseteq \mathbb{R}[\Pi_\gamma]$ be a unit Laplacian eigenvector of $\mathbb{P}_\gamma$ of eigenvalue λ and symmetry type γ. For $\pi \in \Pi_\gamma$, the lifted vector $\mathbf{w} = \mathbf{B}_\pi \mathbf{v} \in \mathbb{R}[\mathbb{S}_n]$ is a Laplacian eigenvector of $\mathbb{P}_n$ of eigenvalue λ by Prop. 5, and $\mathbf{w} \in W_\gamma$ (i.e., it has symmetry type γ) by Prop. 6. Define the left $\mathbb{S}_n$-module,

$$V_{\gamma,\mathbf{v}} := \mathbb{R}[\mathbb{S}_n]\mathbf{w} = \mathbb{R}[\mathbb{S}_n]B_\pi \mathbf{v}, \qquad \text{which is spanned by } \{\rho_L(\sigma)\mathbf{w}\}_{\sigma \in \mathbb{S}_n}, \tag{16}$$

where $\rho_L(\sigma)$ is the matrix of σ in the left regular representation of $\mathbb{S}_n$. Therefore, $\{\rho_L(\sigma)\mathbf{w}\}_{\sigma \in \mathbb{S}_n}$ is a group frame for $V_{\gamma,\mathbf{v}}$ if we view $\{\rho_L(\sigma)\mathbf{w}\}_{\sigma \in \mathbb{S}_n}$ as a multiset of size $n!$ with (potentially many) repetitions.

Since the left and right actions of $\mathbb{S}_n$ commute and the Laplacian $\mathcal{L}$ is constructed (8) using the right representation, $V_{\gamma,\mathbf{v}}$ is a space of Laplacian eigenvectors of $\mathbb{P}_n$ of eigenvalue λ. Moreover, by the double centralizer theorem (see, for example, [34, Thm. 5.18.1]), $W_\gamma \cong V_\gamma \otimes V_\gamma^*$, as an $(\mathbb{S}_n, \mathbb{S}_n)$ bimodule, where V_γ and V_γ^* are irreducible left and right $\mathbb{S}_n$ modules indexed by γ, respectively. Using this isomorphism, we write $\mathbf{w} = w_1 \otimes w_2 \in V_\gamma \otimes V_\gamma^*$, and then $V_{\gamma,\mathbf{v}} = \mathbb{R}[\mathbb{S}_n]\mathbf{w} = (\mathbb{R}[\mathbb{S}_n]w_1) \otimes w_2 \cong V_\gamma \otimes w_2 \cong V_\gamma$, isomorphic as left $\mathbb{S}_n$-modules. Therefore, $V_{\gamma,\mathbf{v}}$ is an irreducible $\mathbb{S}_n$-module and, by [97, Thm 10.5], $\{\rho_L(\sigma)\mathbf{w}\}_{\sigma \in \mathbb{S}_n}$ is a tight group frame.

We remove some of the repetition in $\{\rho_L(\sigma)\mathbf{w}\}_{\sigma \in \mathbb{S}_n}$ by using Prop 4:

$$\{\rho(\sigma)\mathbf{w} \mid \sigma \in \mathbb{S}_n\} = \{\rho(\sigma)\mathbf{B}_\pi \mathbf{v} \mid \sigma \in \mathbb{S}_n\} = \{\mathbf{B}_{\sigma(\pi)}\mathbf{v} \mid \sigma \in \mathbb{S}_n\} = \{\mathbf{B}_\mu \mathbf{v} \mid \mu \in \Pi_\gamma\}, \tag{17}$$

where the third equality comes from the fact that $\mathbb{S}_n$ acts transitively on Π_γ. Eqn. (17) tells us that $V_{\gamma,\mathbf{v}}$ is independent of the specific lifting π that we use. Let $\mathbb{S}_\pi = \{\sigma \in \mathbb{S}_n \mid \sigma(\pi) = \pi\}$ be the stabilizer subgroup of $\pi \in \Pi_\gamma$, and for $\mu \in \Pi_\gamma$, let $\tau_\mu \in \mathbb{S}_n$ be a permutation such that $\tau_\mu(\pi) = \mu$. Then the left coset $\tau_\mu \mathbb{S}_\pi$ consists of all permutations that send π to μ. By [97, Thm 10.5], for any $\mathbf{f} \in \mathbb{R}[\mathbb{S}_n]$, we have

$$\mathbf{f} = \frac{d_\gamma}{n!} \frac{1}{\langle B_\pi \mathbf{v}, B_\pi \mathbf{v}\rangle} \sum_{\sigma \in \mathbb{S}_n} \langle \mathbf{f}, \rho_L(\sigma)B_\pi \mathbf{v}\rangle \rho_L(\sigma)B_\pi \mathbf{v} = \frac{d_\gamma}{n!} \frac{1}{\langle B_\pi \mathbf{v}, B_\pi \mathbf{v}\rangle} \sum_{\sigma \in \mathbb{S}_n} \langle \mathbf{f}, B_{\sigma(\pi)}\mathbf{v}\rangle B_{\sigma(\pi)}\mathbf{v}$$

$$= \frac{d_\gamma}{n!} \frac{1}{\langle B_\pi \mathbf{v}, B_\pi \mathbf{v}\rangle} \sum_{\mu \in \Pi_\gamma} \sum_{\sigma \in \mathbb{S}_\pi} \langle \mathbf{f}, B_{\tau_\mu \sigma(\pi)}\mathbf{v}\rangle B_{\tau_\mu \sigma(\pi)}\mathbf{v} = \frac{d_\gamma}{n!} \sum_{\mu \in \Pi_\gamma} \langle \mathbf{f}, B_\mu \mathbf{v}\rangle B_\mu \mathbf{v},$$

where the last equality follows from $|\mathbb{S}_\pi| = n!/m_\gamma$ and $\langle B_\pi \mathbf{v}, B_\pi \mathbf{v}\rangle = \mathbf{v}^\top B_\pi^\top B_\pi \mathbf{v} = n!/m_\gamma \mathbf{v}^\top \mathbf{v} = n!/m_\gamma$. It follows that $\Phi_{\gamma,\mathbf{v}} := \{\sqrt{d_\gamma/n!}B_\mu \mathbf{v} \mid \mu \in \Pi_\gamma\}$ is a tight Parseval frame for $V_{\gamma,\mathbf{v}}$, again viewing $\Phi_{\gamma,\mathbf{v}}$ as a multiset that can have repetition (as seen in Lem. 1 below).

Now suppose that $\{\mathbf{v}_i^*\}_{i=1}^{d_\gamma}$ is an orthonormal basis for $V_\gamma^* \subseteq \mathbb{R}[\Pi_\gamma]$. Then $\Phi_{\gamma,\mathbf{v}_i}$ is a tight Parseval frame for $V_{\gamma,\mathbf{v}_i^*}$ for each $1 \le i \le d_\gamma$. Moreover, $V_{\gamma,\mathbf{v}_i}$ and $V_{\gamma,\mathbf{v}_j}$ are orthogonal subspaces for $i \ne j$. To see this, identify $V_{\gamma,\mathbf{v}_i^*}$ with $V_\gamma \otimes \mathbf{v}_i^*$ in $W_\gamma \cong V_\gamma \otimes V_\gamma^*$. Since V_γ and V_γ^* are irreducible, they carry a unique (up to scalar multiple) $\mathbb{S}_n$-invariant inner product, and therefore an $\mathbb{S}_n$-invariant inner product on $W_\gamma \subseteq \mathbb{R}[\mathbb{S}_n]$ equals $\langle \mathbf{w}_1 \otimes \mathbf{v}_i, \mathbf{w}_2 \otimes \mathbf{v}_j\rangle = \langle \mathbf{w}_1, \mathbf{w}_2\rangle\langle \mathbf{v}_i, \mathbf{v}_j\rangle$, up to a scalar. The orthogonality of the spaces $\{V_{\gamma,\mathbf{v}_i^*}\}_{i=1}^{d_\gamma}$ follows from the orthogonality of $\{\mathbf{v}_i^*\}_{i=1}^{d_\gamma}$.

Finally, if $\lambda \in \Lambda_\gamma$ and $\{\mathbf{v}_{\gamma,\lambda,k}\}_{k=1}^{\kappa_{\gamma,\lambda}}$ is an orthonormal basis for the graph Laplacian eigenspace of $V_\gamma^* \subseteq \mathbb{R}[\Pi_\gamma]$ of eigenvalue λ, then $\Phi_{\gamma,\lambda} = \cup_{k=1}^{\kappa_{\gamma,\lambda}}\Phi_{\gamma,\mathbf{v}_{\gamma,\lambda,k}}$ is a union of orthogonal tight Parseval frames, and therefore is a tight Parseval frame for $Z_{\gamma,\lambda}$. Furthermore, $\Phi_\gamma = \cup_{\lambda \in \Lambda_\gamma}\Phi_{\gamma,\lambda}$ a union of orthogonal tight Parseval frames, and therefore is a tight Parseval frame for the isotypic component W_γ. Since isotypic components are orthogonal (e.g., [97, Thm. 10.7]), $\mathcal{D} = \cup_{\gamma \vdash n}\Phi_\gamma$ is a tight Parseval frame for $\mathbb{R}[\mathbb{S}_n]$. $\square$

Remark 2 (Frames for subspaces of $\mathbb{R}[\mathbb{S}_n]$) From the proof of Thm. 1 we see that the proposed method can be leveraged to construct a tight Parseval frame for any union of the $Z_{\gamma,\lambda}$ subspaces, not just $\mathbb{R}[\mathbb{S}_n]$. This property can be beneficial for computational reasons, and is explored further in Sec. 6.4. In particular, for $\gamma \vdash n, \lambda \in \Lambda_\gamma$, and $1 \le k \le \kappa_{\gamma,\lambda}$, we have:

1. $\Phi_{\gamma,\lambda,k} = \{\boldsymbol{\varphi}_{\gamma,\lambda,k,\pi}\}_{\pi \in \Pi_\gamma}$ is a tight Parseval frame for $\mathbb{R}[\mathbb{S}_n]\boldsymbol{\varphi}_{\gamma,\lambda,k,\pi} \cong V_\gamma$ (for any $\pi \in \Pi_\gamma$).
2. $\Phi_{\gamma,\lambda} = \{\boldsymbol{\varphi}_{\gamma,\lambda,k,\pi}\}_{1 \le k \le \kappa_{\gamma,\lambda},\pi \in \Pi_\gamma}$ is a tight Parseval frame for the subspace $Z_{\gamma,\lambda}$.
3. $\Phi_\gamma = \{\boldsymbol{\varphi}_{\gamma,\lambda,k,\pi}\}_{\lambda \in \Lambda_\gamma, 1 \le k \le \kappa_{\gamma,\lambda}, \pi \in \Pi_\gamma}$ is a tight Parseval frame for the isotypic component W_γ.

Remark 3 (Equal norms) Since the frame $\Phi_{\gamma,\lambda,k} = \{\boldsymbol{\varphi}_{\gamma,\lambda,k,\pi}\}_{\pi \in \Pi_\gamma}$ is generated by a group action, the frame vectors have equal norms. In fact, for any $\boldsymbol{\varphi}_{\gamma,\lambda,k,\pi} \in \Phi_{\gamma,\lambda,k}$,

$$\langle \boldsymbol{\varphi}_{\gamma,\lambda,k,\pi}, \boldsymbol{\varphi}_{\gamma,\lambda,k,\pi}\rangle = c_\gamma^2 \langle \mathbf{B}_\pi \mathbf{v}_{\gamma,\lambda,k}, \mathbf{B}_\pi \mathbf{v}_{\gamma,\lambda,k}\rangle = \mathbf{v}_{\gamma,\lambda,k}^\top \mathbf{B}_\pi^\top \mathbf{B}_\pi \mathbf{v}_{\gamma,\lambda,k} = c_\gamma^2 \frac{n!}{m_\gamma}\mathbf{v}_{\gamma,\lambda,k}^\top \mathbf{v}_{\gamma,\lambda,k} = c_\gamma^2 \frac{n!}{m_\gamma} = \frac{d_\gamma}{m_\gamma}.$$

Remark 4 (*Frame angles*) The Gram matrix (or Gramian) of $\Phi_{\gamma,\lambda,k}$ is the matrix of inner products,

$$\mathbf{G}_{\Phi_{\gamma,\lambda,k}} = [\langle \varphi, \psi \rangle]_{\varphi, \psi \in \Phi_{\gamma,\lambda,k}}.$$

Since $\Phi_{\gamma,\lambda,k}$ is a tight Parseval frame for the irreducible module $\mathbb{R}[\mathbb{S}_n]\varphi_{\gamma,\lambda,k,\pi} \cong V_\gamma$, the Gram matrix $\mathbf{G}_{\Phi_{\gamma,\lambda,k}}$ equals the $m_\gamma \times m_\gamma$ matrix that projects $\mathbb{R}[\Pi_\gamma]$ onto V_γ (see [97, Cor. 10.2, Thm. 13.1]). This projection has a well-known description (e.g., [97, (13.19)]) as the matrix of the following operator in the center of the group algebra $\mathbb{R}[\mathbb{S}_n]$:

$$p_\gamma = \frac{d_\gamma}{n!} \sum_{\sigma \in \mathbb{S}_n} \chi_\gamma(\sigma^{-1})\sigma, \tag{18}$$

where the coefficients $\chi_\gamma(\sigma^{-1})$ are given by the irreducible character χ_γ corresponding to γ. Applying p_γ to the basis $\{\mathbf{e}_\mu\}_{\mu \in \Pi_\gamma}$ of $\mathbb{R}[\Pi_\gamma]$ gives

$$p_\gamma(\mathbf{e}_\mu) = \frac{d_\gamma}{n!} \sum_{\sigma \in \mathbb{S}_n} \chi_\gamma(\sigma^{-1})e_\mu\sigma = \frac{d_\gamma}{n!} \sum_{\sigma \in \mathbb{S}_n} \chi_\gamma(\sigma^{-1})e_{\sigma^{-1}(\mu)} = \frac{d_\gamma}{n!} \sum_{\sigma \in \mathbb{S}_n} \chi_\gamma(\sigma)e_{\sigma(\mu)}.$$

For $\pi, \mu \in \Pi_\gamma$, the entry $\langle \boldsymbol{\varphi}_{\gamma,\lambda,k,\pi}, \boldsymbol{\varphi}_{\gamma,\lambda,k,\mu} \rangle$ of $\mathbf{G}_{\Phi_{\gamma,\lambda,k}}$ is the coefficient of $\mathbf{e}_\pi$ in $p_\gamma(\mathbf{e}_\mu)$; namely,

$$\langle \boldsymbol{\varphi}_{\gamma,\lambda,k,\pi}, \boldsymbol{\varphi}_{\gamma,\lambda,k,\mu} \rangle = \frac{d_\gamma}{n!} \sum_{\sigma \in \mathcal{V}_{\pi,\mu}} \chi_\gamma(\sigma), \tag{19}$$

where $\mathcal{V}_{\pi,\mu} = \{\sigma \in \mathbb{S}_n \mid \sigma(\mu) = \pi\}$ is one of the equivalence classes of Def. 2 and is a right coset of the stabilizer $\mathcal{V}_{\pi,\pi} = \{\sigma \in \mathbb{S}_n \mid \sigma(\pi) = \pi\}$. The characters of the symmetric group are integers, so, $\sum_{\sigma \in \mathcal{V}_{\pi,\mu}} \chi_\gamma(\sigma) \in \mathbb{Z}$. Frame vectors corresponding to different values of λ and k are orthogonal (as seen in the proof of Thm. 1). Therefore the Gram matrix $\mathbf{G}_{\Phi_\gamma} = \bigoplus_{\lambda \in \Lambda_\gamma} \bigoplus_{k=1}^{\kappa_{\gamma,\lambda}} \mathbf{G}_{\Phi_{\gamma,\lambda,k}}$ for the isotypic component W_γ is the direct sum of d_γ matrices, each of the form (19).

For shapes γ with multiple blocks of the same size, there is redundancy in the frame $\Phi_{\gamma,\lambda}$ that can be removed. Let $\gamma = [\gamma_1, \ldots, \gamma_\ell]$ and suppose that two parts of γ are equal; that is, $\gamma_i = \gamma_j$. For $\pi \in \Pi_\gamma$, let $\pi' \in \Pi_\gamma$ be the ordered set partition obtained from π by swapping row i and row j. For example, if $\gamma = [4, 3, 3, 2]$, then

$$\pi = \begin{array}{|c|} \hline \begin{array}{|c c c c|} \hline 1 & 4 & 7 & 12 \\ \hline \end{array} \\ \begin{array}{|c c c|} \hline 2 & 5 & 8 \\ \hline \end{array} \\ \begin{array}{|c c c|} \hline 3 & 6 & 11 \\ \hline \end{array} \\ \begin{array}{|c c|} \hline 9 & 10 \\ \hline \end{array} \end{array} \qquad \text{and} \qquad \pi' = \begin{array}{|c|} \hline \begin{array}{|c c c c|} \hline 1 & 4 & 7 & 12 \\ \hline \end{array} \\ \begin{array}{|c c c|} \hline 3 & 6 & 11 \\ \hline \end{array} \\ \begin{array}{|c c c|} \hline 2 & 5 & 8 \\ \hline \end{array} \\ \begin{array}{|c c|} \hline 9 & 10 \\ \hline \end{array} \end{array} \tag{20}$$

satisfy this condition. The liftings from $V_\gamma \subseteq \mathbb{R}[\Pi_\gamma]$ to $\mathbb{R}[\mathbb{S}_n]$ via π and π' are related according to the following lemma.

Lemma 1 *Let* $\gamma = [\gamma_1, \ldots, \gamma_\ell] \vdash n$ *with* $t = \gamma_i = \gamma_j$ *and let* $\pi, \pi' \in \Pi_\gamma$ *be equal after swapping rows i and j in* π. *If* $\mathbf{v} \in V_\gamma \subseteq \mathbb{R}[\Pi_\gamma]$, *then* $\mathbf{B}_\pi \mathbf{v} = (-1)^t \mathbf{B}_{\pi'} \mathbf{v}$.

Proof Let $\gamma = [\gamma_1, \ldots, \gamma_\ell]$ with $t = \gamma_i = \gamma_j$. For $\rho \in \Pi_\gamma$, let ρ' be the same set partition as ρ except with rows i and $i+1$ swapped, as illustrated in (20). Suppose that $\mathbf{v} \in \mathbb{R}[\Pi_\gamma]$ is expressed in the canonical basis as $\mathbf{v} = \sum_{\rho \in \Pi_\gamma} c_\rho \mathbf{e}_\rho$. Suppose further that for each ρ we have $c_\rho = (-1)^t c_{\rho'}$, which we call the symmetry property. Then,

$$\mathbf{B}_\pi \mathbf{v} = \sum_{\rho \in \Pi_\gamma} c_\rho \mathbf{B}_\pi \mathbf{e}_\rho = \sum_{\rho \in \Pi_\gamma} c_\rho \sum_{\substack{\sigma \in \mathbb{S}_n \\ \sigma(\rho) = \pi}} \mathbf{e}_\sigma = (-1)^t \sum_{\rho \in \Pi_\gamma} c_{\rho'} \sum_{\substack{\sigma \in \mathbb{S}_n \\ \sigma(\rho) = \pi}} \mathbf{e}_\sigma = (-1)^t \sum_{\rho \in \Pi_\gamma} c_{\rho'} \sum_{\substack{\sigma \in \mathbb{S}_n \\ \sigma(\rho') = \pi'}} \mathbf{e}_\sigma = (-1)^t \mathbf{B}_{\pi'} \mathbf{v},$$

since $\sigma(\rho) = \pi$ if and only if $\sigma(\rho') = \pi'$. Thus, the proposition is proved if we show that every $\mathbf{v}$ in the submodule $V_\gamma \subseteq \mathbb{R}[\Pi_\gamma]$ has the symmetry property $c_\rho = (-1)^t c_{\rho'}$.

The submodule $V_\gamma^* \subseteq \mathbb{R}[\Pi_\gamma]$ is spanned by the following set of vectors, called polytabloids (see [83, 2.3]),

$$q_\pi = \sum_{\beta \in C_\pi} \mathsf{sign}(\beta)\mathbf{e}_\pi \beta = \sum_{\beta \in C_\pi} \mathsf{sign}(\beta)\mathbf{e}_{\beta^{-1}(\pi)}, \qquad \pi \in \Pi_\gamma, \tag{21}$$

where $C_\pi \subseteq \mathbb{S}_n$ is the column group of π, that is, the permutations that stabilize the columns of π, and $\mathsf{sign}(\beta)$ is the sign of the permutation β. Let $\tau_\pi \in C_\pi$ be the permutation that is the product of the t disjoint transpositions (not necessarily adjacent) that swap an entry in row i of π with the corresponding entry in row j. For example, in (20), $\tau_\pi = (2,3)(5,6)(8,11)$. Then $\mathsf{sign}(\tau_\pi) = (-1)^t$ and $\pi' = \tau_\pi(\pi)$.

Since $\tau_\pi \in C_\pi$, we have $\tau_\pi^{-1} C_\pi \tau_\pi = C_\pi$ and $\mathsf{sign}(\tau_\pi^{-1}\beta\tau_\pi) = \mathsf{sign}(\beta)$. Moreover, $\tau_\pi^{-1} = \tau_\pi$, so we have

$$q_\pi \tau_\pi = \sum_{\beta \in C_\pi} \mathsf{sign}(\beta)\mathbf{e}_\pi \beta \tau_\pi = \sum_{\beta \in C_\pi} \mathsf{sign}(\beta)\mathbf{e}_\pi \tau_\pi \tau_\pi^{-1} \beta \tau_\pi = \sum_{\beta \in C_\pi} \mathsf{sign}(\beta)\mathbf{e}_{\tau_\pi(\pi)}\tau_\pi^{-1}\beta\tau_\pi = q_{\tau_\pi(\pi)} = q_{\pi'},$$

$$q_\pi \tau_\pi = \sum_{\beta \in C_\pi} \mathsf{sign}(\beta)\mathbf{e}_\pi \beta \tau_\pi = \sum_{\eta \in C_\pi} \mathsf{sign}(\tau_\pi \eta)\mathbf{e}_\pi \eta = \mathsf{sign}(\tau)\sum_{\eta \in C_\pi} \mathsf{sign}(\eta)\mathbf{e}_\pi \eta = (-1)^t q_\pi,$$

and thus $q_\pi = (-1)^t q_{\pi'}$. It follows that q_π has the symmetry property $c_\rho = (-1)^t c_{\rho'}$. Since this is true of each vector of the spanning set $q_\pi, \pi \in \Pi_\gamma$, it is true for all of V_γ^*, and the result is proved. $\qquad\square$

Lem. 1 tells us that, in the case where γ has repeated parts, many of the atoms in (15) are identical or are the negatives of others. We then can lift fewer vectors to generate a tight Parseval frame for $Z_{\gamma,\lambda}$, which leads to a more computationally efficient implementation without sacrificing any interpretability. Define $z_\gamma := \frac{m_\gamma}{\prod_i k_i!}$, where if $\gamma = [\gamma_1, \ldots, \gamma_\ell]$, k_i is the multiplicity of i in γ. For example, if $\gamma = [4,2,2,2,1]$, $m_\gamma = \frac{11!}{4!2!2!2!1!}$ and $z_\gamma = \frac{m_\gamma}{3!}$, as $i = 2$ appears three times in γ. Identifying ordered set partitions in Π_γ that feature the same groupings of candidates yields a smaller set of z_γ (unordered) set partitions, which we denote by $\bar{\Pi}_\gamma$. For example, the ordered set partitions $\{\{1,2,3,4\},\{5,6\},\{7,8\},\{9,10\},\{11\}\}$, $\{\{1,2,3,4\},\{7,8\},\{5,6\},\{9,10\},\{11\}\}$, and four others are all identified to a single set partition in $\bar{\Pi}_{[4,2,2,2,1]}$. For $\bar{\pi} \in \bar{\Pi}_\gamma$, define $\boldsymbol{\varphi}_{\gamma,\lambda,k,\bar{\pi}} := \bar{c}_\gamma \mathbf{B}_{\bar{\pi}} \mathbf{v}_{\gamma,\lambda,k}$, and define the reduced frame

$$\bar{\Phi}_{\gamma,\lambda} := \bigcup_{k=1}^{\kappa_{\gamma,\lambda}} \bigcup_{\bar{\pi} \in \bar{\Pi}_\gamma} \boldsymbol{\varphi}_{\gamma,\lambda,k,\bar{\pi}} := \bigcup_{k=1}^{\kappa_{\gamma,\lambda}} \bigcup_{\bar{\pi} \in \bar{\Pi}_\gamma} \bar{c}_\gamma \mathbf{B}_{\bar{\pi}} \mathbf{v}_{\gamma,\lambda,k}, \tag{22}$$

where the constant $\bar{c}_\gamma := \sqrt{\frac{d_\gamma m_\gamma}{n! z_\gamma}}$. In Thm. 2 we show that $\bar{\Phi}_{\gamma,\lambda}$ remains a tight Parseval frame for $Z_{\gamma,\lambda}$.

Theorem 2 *For $\gamma \vdash n$ and $\lambda \in \Lambda_\gamma$, the collection of atoms $\bar{\Phi}_{\gamma,\lambda}$ defined in (22) is a tight Parseval frame for $Z_{\gamma,\lambda}$, and the set of atoms $\bar{\mathcal{D}} := \bigcup_{\gamma \vdash n} \bigcup_{\lambda \in \Lambda_\gamma} \bar{\Phi}_{\gamma,\lambda}$ is a tight Parseval frame for $\mathbb{R}[\mathbb{S}_n]$.*

Proof By Lem. 1, $\Phi_{\gamma,\lambda} = \bigcup_{i=1}^{m_\gamma/z_\gamma} \pm\bar{\Phi}_{\gamma,\lambda}$. Therefore, from Thm. 1, for any $\mathbf{f} \in \mathbb{R}[\mathbb{S}_n]$, we have

$$\mathbf{f} = \sum_{\phi \in \Phi_{\gamma,\lambda}} \langle \mathbf{f}, \phi \rangle \phi = \sum_{i=1}^{m_\gamma/z_\gamma} \sum_{\pm\phi \in \pm\bar{\Phi}_{\gamma,\lambda}} \langle \mathbf{f}, \pm\phi \rangle (\pm\phi) = \sum_{i=1}^{m_\gamma/z_\gamma} \sum_{\phi \in \bar{\Phi}_{\gamma,\lambda}} \langle \mathbf{f}, \phi \rangle \phi = \frac{m_\gamma}{z_\gamma} \sum_{\phi \in \bar{\Phi}_{\gamma,\lambda}} \langle \mathbf{f}, \phi \rangle \phi.$$

By observing that $\bar{c}_\gamma = \sqrt{\frac{d_\gamma m_\gamma}{n! z_\gamma}} = \sqrt{\frac{m_\gamma}{z_\gamma}} c_\gamma$, we see that $\bar{\Phi}_{\gamma,\lambda}$ is a tight Parseval frame for $Z_{\gamma,\lambda}$. $\qquad\square$
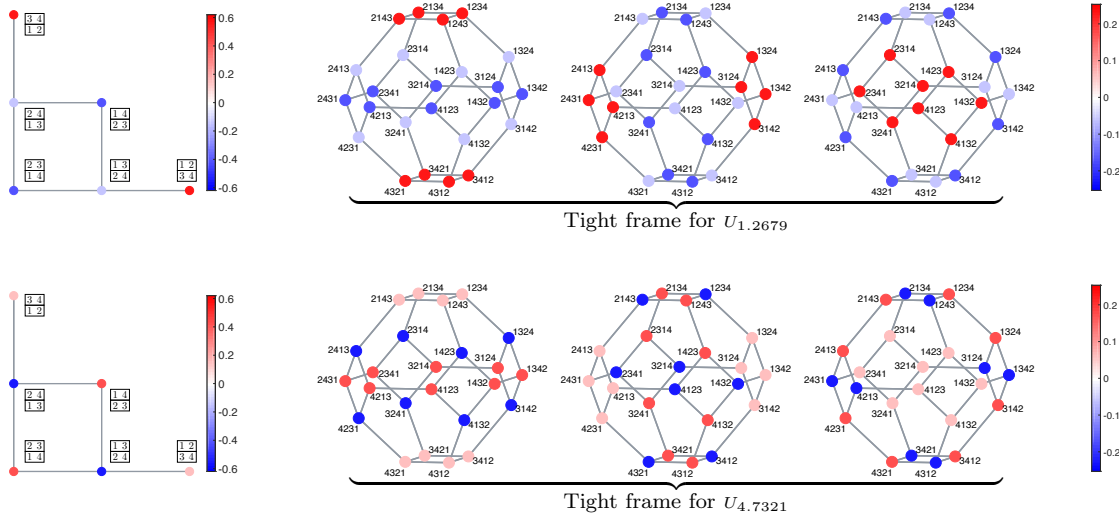
Fig. 15 Left: The graph Laplacian eigenvectors $\mathbf{v}_{[2,2],\lambda}$ of the Schreier graph $\mathbb{P}_{[2,2]}$ associated with eigenvalues $\lambda = 1.2679$ (top) and $\lambda = 4.7321$ (bottom). Right: The dictionary atoms $\{\boldsymbol{\varphi}_{\gamma,\lambda,z}\} = \{\bar{c}_\gamma \mathbf{B}_{\pi_z} \mathbf{v}_{[2,2],\lambda}\}$ are generated by lifting the Schreier Laplacian eigenvectors according to the three different ordered set partitions $\pi_z \in \bar{\Pi}_{[2,2]}$ shown in Fig. 13. The resulting sets of three atoms form tight frames for the corresponding spaces $Z_{[2,2],\lambda}$. In this case, $Z_{[2,2],1.2679} = U_{1.2679}$ and $Z_{[2,2],4.7321} = U_{4.7321}$, since the eigenvalues $\lambda = 1.2679$ and $\lambda = 4.7321$ only appear in the [2,2] irreducible.

In Fig. 15, for $\gamma = [2,2]$, we show two different eigenvectors of the Schreier graph $\mathbb{P}_\gamma$ lifted back to the permutahedron according to the $z_\gamma = 3$ different ordered set partitions in Fig. 13, yielding tight frames $\bar{\Phi}_{[2,2],1.2679}$ and $\bar{\Phi}_{[2,2],4.7321}$ with three vectors each for the two-dimensional spaces $Z_{[2,2],1.2679}$ and $Z_{[2,2],4.7321}$, respectively. *An important point of emphasis here is that compared to the orthonormal bases for the same spaces in Fig. 9, the frame vectors in Fig. 15 maintain interpretable symmetry properties.*

In this case, were we to include all $m_\gamma = 6$ liftings generated from the ordered set partitions in Π_γ, the resulting frames $\Phi_{\gamma,\lambda}$ would have two copies of each of these three atoms, all scaled by the constant factor $\frac{c_\gamma}{\bar{c}_\gamma} = \sqrt{\frac{z_\gamma}{m_\gamma}}$. In cases where the repeated parts have odd length (e.g., tight frames for $Z_{[8,1,1],\lambda}$ generated by lifting the eigenvector in Fig. 23(b)), the removed atoms would have the opposite sign in each entry. Unless specified, our default in the remainder of the paper is to use the tight frames with fewer elements defined in (22).

Returning to the three objectives outlined at the beginning of this section, the proposed dictionary atoms comprise a tight Parseval frame, as shown in Thm. 1 and Thm. 2, and therefore satisfy the first objective of preserving the energy of the signal. Each atom $\boldsymbol{\varphi}_{\gamma,\lambda,k,\bar{\pi}}$ belongs to the space $Z_{\gamma,\lambda} = W_\gamma \cap U_\lambda$ and therefore inherits known symmetry and smoothness properties from W_γ and U_λ, respectively. In the next section, we investigate the interpretation of specific frame analysis coefficients $\langle \mathbf{f}, \boldsymbol{\varphi}_{\gamma,\lambda,k,\bar{\pi}} \rangle$ in the context of several example data sets. The third objective of identifying methods to efficiently compute these inner products is the focus of Sec. 6, and this computational question actually gives rise to additional interesting theoretical questions.

5 Interpretation of the Analysis Coefficients

The inner products between a signal on the permutahedron and the atoms of the tight spectral frame $\bar{\mathcal{D}}$ from Thm. 2 (or $\mathcal{D}$ from Thm. 1), referred to as *analysis coefficients*, are useful in identifying structure in voting data, such as popular candidates, polarizing candidates, and clusters of candidates commonly ranked similarly by subgroups of voters (e.g., political parties). The analysis coefficients have the form:

$$\alpha_{\gamma,\lambda,k,\bar{\pi}} := \langle \mathbf{f}, \boldsymbol{\varphi}_{\gamma,\lambda,k,\bar{\pi}} \rangle = \langle \mathbf{f}, \bar{c}_\gamma \mathbf{B}_{\bar{\pi}} \mathbf{v}_{\gamma,\lambda,k} \rangle = \bar{c}_\gamma \langle \mathbf{B}_{\bar{\pi}}^\top \mathbf{f}, \mathbf{v}_{\gamma,\lambda,k} \rangle. \tag{23}$$

In some instances, it is beneficial for interpretation purposes to view these analysis coefficients as inner products between the signal $\mathbf{f}$ and the atoms (the first two terms in (23)). All of these signals reside on the permutahedron $\mathbb{P}_n$ (see Fig. 15 or the second row from the bottom in Fig. 18 for illustrations of example atoms). In other instances, it is helpful to view the same quantity via the last term in (23): a constant times an inner product between the signal projected down to the Schreier graph $\mathbb{P}_\gamma$ in a specific manner, and a Laplacian eigenvector of that Schreier graph.

Recalling that the energy of the signal is equal to the energy of the analysis coefficients, i.e.,

$$\|\mathbf{f}\|^2 = \sum_{\gamma,\lambda,k,\bar{\pi}} |\langle \mathbf{f}, \boldsymbol{\varphi}_{\gamma,\lambda,k,\bar{\pi}} \rangle|^2, \tag{24}$$

we first investigate what information can be garnered from the decomposition of the energy of the analysis coefficients on the right-hand side of (24) (i) across shapes γ, (ii) across eigenvalues λ within a fixed shape γ, and (iii) across atoms within a shape-eigenvalue pair γ, λ. We then examine the interpretation of specific analysis coefficients.

5.1 Energy Decomposition

The squared magnitudes $|\langle \mathbf{f}, \boldsymbol{\varphi}_{\gamma,\lambda,k,\bar{\pi}} \rangle|^2$ in the summand of (24) can be aggregated and plotted in different ways to identify structural patterns in the ranking tallies, $\mathbf{f}$. First, for each shape γ, the sum $\sum_{\lambda,k,\bar{\pi}} |\langle \mathbf{f}, \boldsymbol{\varphi}_{\gamma,\lambda,k,\bar{\pi}} \rangle|^2$ is equal to $\|\mathbf{f}_\gamma\|^2$, the energy of the projection of the signal onto the corresponding isotypic component (compare, e.g., the bottom image in Fig. 3 to the top table in Fig. 18). Each of these quantities can be further decomposed across the eigenspaces associated with the eigenvalues in Λ_γ via the sums $\|\mathbf{f}_{\gamma,\lambda}\|^2 = \sum_{k,\bar{\pi}} |\langle \mathbf{f}, \boldsymbol{\varphi}_{\gamma,\lambda,k,\bar{\pi}} \rangle|^2$. For example, as shown in Fig. 18, for the 2017 Minneapolis City Council Ward 3 election data $\mathbf{g}$,

$$355201.6 = \|\mathbf{g}_{\boxplus}\|^2 = \|\mathbf{g}_{\boxplus,0.586}\|^2 + \|\mathbf{g}_{\boxplus,2}\|^2 + \|\mathbf{g}_{\boxplus,3.414}\|^2$$
$$= \sum_{\bar{\pi}} \left| \left\langle \mathbf{g}, \boldsymbol{\varphi}_{\boxplus,0.586,\bar{\pi}} \right\rangle \right|^2 + \sum_{\bar{\pi}} \left| \left\langle \mathbf{g}, \boldsymbol{\varphi}_{\boxplus,2,\bar{\pi}} \right\rangle \right|^2 + \sum_{\bar{\pi}} \left| \left\langle \mathbf{g}, \boldsymbol{\varphi}_{\boxplus,3.414,\bar{\pi}} \right\rangle \right|^2$$
$$= 147617.5 + 192845.1 + 14739.0.$$

Through plots such as those shown in Fig. 16 and Fig. 17, we can visualize the full decomposition of energy into shape-eigenvalue pairs. For example, in Fig. 16 (or the top row of Fig. 18), we see that most of the energy from the 2017 Minneapolis City Council Ward 3 election data $\mathbf{g}$ shown in Fig. 1 falls in the spaces $Z_{[4],0}$ (which just conveys information about the total number of voters), $Z_{[3,1],2}$, $Z_{[3,1],0.586}$, and $Z_{[2,2],1.268}$. As discussed in Sec. 3.2, typically occurring rankings are smooth with respect to the underlying permutahedron structure, and we therefore expect to see a decay in the energies $\{\|\mathbf{f}_{\gamma,\lambda}\|^2\}$ as λ increases, as is the case, e.g., in the sushi data in Fig. 17. While in the examples shown Fig. 16 and Fig. 17 the bar at each eigenvalue is comprised of a single color representing the corresponding shape, this need not be the case. For example, when $n = 6$, the eigenvalue $\lambda = 3$ appears in two shapes ($\boxplus$ and $\boxplus$), and therefore the energy of a signal at this eigenvalue would be comprised of two bars of different colors stacked on top of each other. Additionally, when the graph Laplacian eigenspace of the Schreier graph $\mathbb{P}_\gamma$ that is associated with λ in V_γ has dimension greater
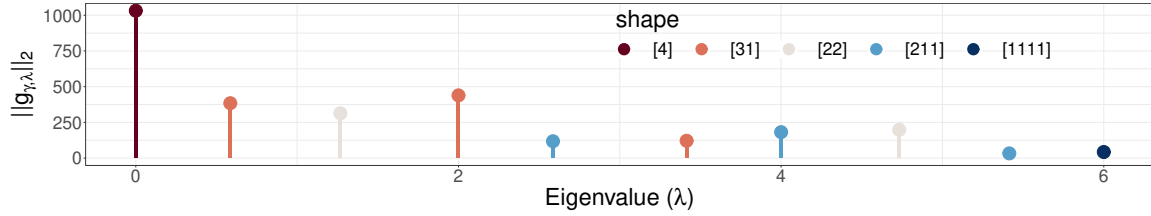
Fig. 16 Decomposition of the energy of the 2017 Minneapolis City Council Ward 3 election data $\mathbf{g}$ into the shape-eigenvalue spaces $\{Z_{\gamma,\lambda}\}$.
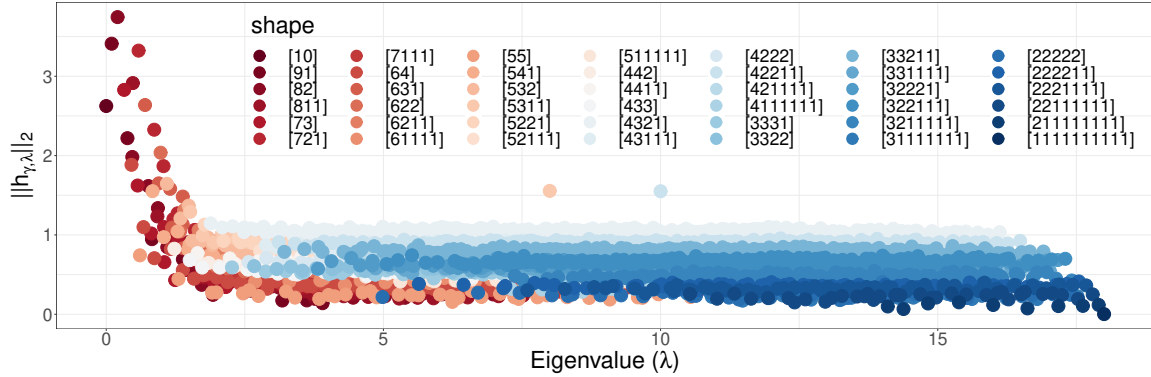

Fig. 17 The sushi data $\mathbf{h}$ is a smooth signal on the permutahedron $\mathbb{P}_{10}$, and the energies of $\mathbf{h}$ in each shape-eigenvalue space generally decay as the eigenvalues increase, with the exception of two values for which $\kappa_{\gamma,\lambda} > 1$ (these outliers are expected and do not reflect any structure of interest in this particular data).

than one (i.e., $\kappa_{\gamma,\lambda} > 1$), the energies of the analysis coefficients resulting from all atoms generated from the basis $\{\mathbf{v}_{\gamma,\lambda,k}\}_{k=1}^{\kappa_{\gamma,\lambda}}$ are stacked on top of one another and shown as a single bar. For example, the apparent outliers at $\lambda = 8$ and $\lambda = 10$ in Fig. 17 are only due to the fact that $\kappa_{\gamma,\lambda} > 1$ for these shape-eigenvalue pairs, as opposed to some structure of interest in the sushi data; a plot of the energies of the analysis coefficients of a pure noise signal yields similar outliers.

We can further decompose the energy in any shape-eigenvalue pair by examining the sequence of energies, $\{|\langle \mathbf{f}, \boldsymbol{\varphi}_{\gamma,\lambda,k,\bar{\pi}} \rangle|^2\}_{1 \le k \le \kappa_{\gamma,\lambda}, \bar{\pi} \in \bar{\Pi}_{\gamma}}$, associated with each atom in the frame. In particular, for any shape-eigenvalue pair with a significant amount of energy, we can identify the ordered set partitions $\bar{\pi}$ used to lift the associated eigenvectors to generate the atoms that yield the inner products with the highest magnitudes. For example, for the sushi data signal $\mathbf{h}$, Fig. 19 shows the shape-eigenvalue-lifting triplets associated with largest magnitude analysis coefficients.

5.2 Interpretation of Specific Analysis Coefficients

How can we use the magnitudes of the analysis coefficients to identify structure in the ranked data? The general methodology for each analysis coefficient is to look at the structure of the corresponding eigenvector on the Schreier graph of the identified shape and the projection of the signal onto that Schreier graph via $\mathbf{B}_{\bar{\pi}}^{\top}$, as the analysis coefficient is the inner product of those two signals on P_{γ} (see (23)). One key advantage of this method is that it allows us to create visualizations of high-dimensional data on much lower-dimensional graphs. Some of these eigenvectors are more easily interpretable than others, but we are fortunate that the most interpretable eigenvectors are often the ones associated with the largest magnitude analysis coefficients, particularly for the smooth signals that commonly arise in ranking applications.

γ	▯▯▯▯	⬚⬚			⊞		⬚			⬚
λ	0	0.586	2	3.414	1.268	4.732	2.586	4	5.414	6
$\sum_{\bar{\pi}} \lvert\langle \mathbf{g}, \boldsymbol{\varphi}_{\gamma,\lambda,\bar{\pi}}\rangle\rvert^2$	1064709.4	147617.5	192845.1	14739.0	98412.8	39162.5	13878.0	32979.6	1085.0	1820.0

γ	⬚⬚								⊞		
λ	0.586				2				1.268		
$\mathbf{v}_\lambda$	$\frac{234}{1}$	$\frac{134}{2}$	$\frac{124}{3}$	$\frac{123}{4}$	$\frac{234}{1}$	$\frac{134}{2}$	$\frac{124}{3}$	$\frac{123}{4}$			
$\bar{\pi}$	{234|1}	{134|2}	{124|3}	{123|4}	{234|1}	{134|2}	{124|3}	{123|4}	{12|34}	{13|24}	{14|23}
$\boldsymbol{\varphi}_{\gamma,\lambda,\bar{\pi}}$											
$\langle \mathbf{g}, \boldsymbol{\varphi}_{\gamma,\lambda,\bar{\pi}}\rangle$	51.4	-201.6	290.8	-140.6	318.7	-185.1	-221.9	88.2	239.0	-39.9	-199.3

Fig. 18 Top: The energies of the projections of $\mathbf{g}$ onto the spaces $Z_{\gamma,\lambda}$ associated with each shape-Laplacian eigenvalue pair. Bottom: Inner products between $\mathbf{g}$ and a subset of the dictionary atoms of the form $\boldsymbol{\varphi}_{\gamma,\lambda,\bar{\pi}}$, each of which is created by lifting the Schreier eigenvector $\mathbf{v}_\lambda$ back to the permutahedron according to the partition $\bar{\pi}$.

γ	$\bar{\pi}$	λ	$\langle \mathbf{h}, \boldsymbol{\varphi}_{\gamma,\lambda,\bar{\pi}}\rangle$	$\lvert\langle \mathbf{h}, \boldsymbol{\varphi}_{\gamma,\lambda,\bar{\pi}}\rangle\rvert^2$
	{1234567890}	0	2.6248	6.8893
	{123456789|0}	0.0979	-2.1513	4.6280
	{123456790|8}	0.0979	1.9978	3.9912
	{12345679|80}	0.2047	-1.7150	2.9413
	{12345689|70}	0.2047	1.6543	2.7369
	{12345679|8|0}	0.4799	1.3471	1.8147
	{12456790|38}	0.2047	1.3304	1.7699
	{123456890|7}	0.0979	-1.1896	1.4150
	{123456780|9}	0.3820	-1.1006	1.2112
	{1234569|780}	0.3227	1.0659	1.1362
	{12345690|78}	0.2047	-1.0400	1.0817
	{123456790|8}	0.3820	1.0392	1.0800
	{12345689|70}	0.4700	-1.0046	1.0093
	{123467890|5}	0.3820	0.9604	0.9223

Fig. 19 The 14 tight spectral frame analysis coefficients with the largest magnitudes for the sushi preference data.

5.2.1 The Shape $\gamma = [n]$: Number of Votes

The only eigenvalue in $\Lambda_{[n]}$ is $\lambda = 0$, and there is just a single atom associated with the shape $\gamma = [n]$: $\varphi_{[n],0,\{\{1,2,\ldots,n\}\}} = \frac{1}{\sqrt{n!}}\mathbf{1}_{n!}$, where $\mathbf{1}_{n!}$ is a constant vector of ones of length $n!$. This atom is a basis for the isotypic component $W_{[n]}$, and the analysis coefficient $\langle \mathbf{f}, \varphi_{[n],0,\{\{1,2,\ldots,n\}\}}\rangle = \hat{f}(0) = \|\mathbf{f}_{[n]}\|$ is just equal to the total number of votes (rankings) divided by $n!$ (c.f., Fig. 3).

5.2.2 The Shape $\gamma = [n-1,1]$: Individual Popularity and Polarization

The Laplacian eigenvalues $\Lambda_{[n-1,1]}$ and associated eigenvectors of the Schreier graph $\mathbb{P}_{[n-1,1]}$ are the same as the $n-1$ nonzero Laplacian eigenvalues and associated eigenvectors of a path graph with n vertices. These are known in closed form.

Lemma 2 (see, e.g., [93]) *Let Q_n denote the path graph on n vertices. The Laplacian eigenvalues of Q_n are*

$$\lambda_\ell^{Q_n} = 2 - 2\cos\left(\frac{\pi\ell}{n}\right), \ \ell = 0, 1, \ldots, n-1,$$

and the associated Laplacian eigenvectors are

$$\mathbf{v}_0^{Q_n}(i) = \frac{1}{\sqrt{n}}, \ i = 1, 2, \ldots, n, \ and$$

$$\mathbf{v}_\ell^{Q_n}(i) = \frac{2}{\sqrt{n}}\cos\left(\frac{\pi\ell(i-0.5)}{n}\right), \ \ell = 1, 2, \ldots, n-1, i = 1, 2, \ldots, n.$$

In Fig. 20, we show the first two eigenvalues $\Lambda_{[9,1]}$ and their associated eigenvectors on $\mathbb{P}_{[9,1]}$.

Atoms Generated from the Laplacian Eigenvector $\mathbf{v}_{[n-1,1],2-2\cos(\frac{\pi}{n})}$: How Popular Is Each Candidate?

For $z = 1, 2, \ldots, n$, let $\bar{\pi}_z$ be the set partition that places candidate z in one block and all other candidates in the other block (e.g. with $n = 4$, $\bar{\pi}_2 = \{\{1,3,4\},\{2\}\}$). Then the atom $\varphi_{[n-1,1],2-2\cos(\frac{\pi}{n}),\bar{\pi}_z} = \bar{c}_{[n-1,1]}\mathbf{B}_{\bar{\pi}_z}\mathbf{v}_{[n-1,1],2-2\cos(\frac{\pi}{n})}$ is equal to $\bar{c}_{[n-1,1]}\frac{2}{\sqrt{n}}\cos(\frac{\pi(i-0.5)}{n})$ on each vertex of the permutahedron $\mathbb{P}_n$ associated with a ranking in which candidate z is ranked in place i (see the second row

$\mathbf{v}_{[9,1],0.0979}$ $\mathbf{v}_{[9,1],0.3820}$

Fig. 20 The Laplacian eigenvectors associated with the first two eigenvalues in $\Lambda_{[n-1,1]}$ on the Schreier graph $\mathbb{P}_{[9,1]}$. The index on each vertex corresponds to the element in the second block of the set partition associated with the vertex; e.g., "3" is shorthand notation for $\boxed{\begin{smallmatrix}1\ 2\ 4\ 5\ 6\ 7\ 8\ 9\ 0\\ 3\end{smallmatrix}}$.

from the bottom in Fig. 18 for illustrations of such atoms). Since the eigenvector $\mathbf{v}_{[n-1,1],2-2\cos(\frac{\pi}{n})} = \frac{2}{\sqrt{n}}\cos(\frac{\pi(i-0.5)}{n})$ decreases from $i = 1$ to $i = n$ (see, Fig. 20), the analysis coefficient $\alpha_{[n-1,1],2-2\cos(\frac{\pi}{n}),\bar{\pi}_z}$ conveys a notion of general favorability of candidate z. As shown in Fig. 18, for the 2017 Minneapolis City Council Ward 3 election data, the largest analysis coefficient in this shape-eigenvalue pair is the 290.8 associated with $\bar{\pi}_3$, followed by $\bar{\pi}_1$, $\bar{\pi}_4$, and $\bar{\pi}_2$, indicating that candidate 3 is generally popular whereas candidate 2 is generally not popular. In the sushi preference data, the order of the candidate popularities according to this metric, from most popular to least popular, is 8, 3, 1, 6, 2, 5, 4, 9, 7, 10. This ordering is the same as the Condorcet ranking listed in Sec. 2.3, except with shrimp and salmon roe swapped in third and fourth place.
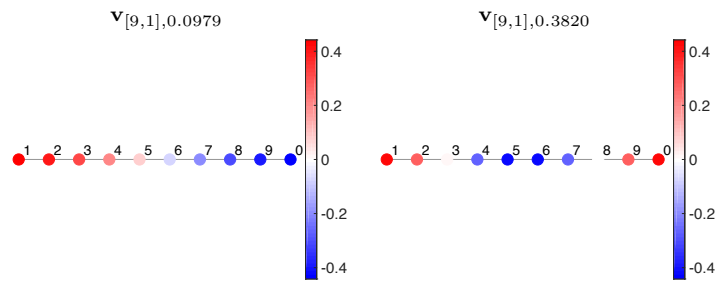
Remark 5 (Relation with rank aggregation and the Borda count) The problem of mapping ranking data into a single consensus ranking of the candidates is referred to as *rank aggregation* and dates back to social choice theory in the 18th century [11] (see [62], [100, Sec. 5] for more recent surveys). Borda's original method [11] is to assign n points to each first place vote, $n-1$ points to each second place vote, and so forth, until 1 point for each last place vote. The consensus ranking is then formed according to the total number of points each candidates receives. The ranking of the candidates according to the analysis coefficients $\alpha_{[n-1,1],2-2\cos(\frac{\pi}{n}),\bar{\pi}_z}$ demonstrated above is equivalent to a Borda count aggregate ranking that uses point values

$$\frac{(n-1)\cos\left(\frac{\pi(i-0.5)}{n}\right)}{2\cos\left(\frac{\pi}{2n}\right)} + \frac{n+1}{2}, \ i = 1, 2, \ldots, n$$

for ith place instead of the more common linearly spaced point values from n to 1. For example, with 10 candidates, the ranking of the analysis coefficients generated from this eigenvector is equivalent to a weighted Borda count where the points assigned for each vote, from first place to last place, are 10.00, 9.56, 8.72, 7.57, 6.21, 4.79, 3.43, 2.28, 1.44, and 1.00.

Atoms Generated from the Laplacian Eigenvector $\mathbf{v}_{[n-1,1],2-2\cos(\frac{2\pi}{n})}$*: How Polarizing Is Each Candidate?*

The eigenvector elements $\mathbf{v}_{[n-1,1],2-2\cos(\frac{2\pi}{n})}(i) = \frac{2}{\sqrt{n}}\cos(\frac{2\pi(i-0.5)}{n})$ decrease from $i=1$ to $i = \lceil\frac{n}{2}\rceil$, and then increase again from $i = \lfloor\frac{n}{2}\rfloor + 1$ to $i = n$ (see, e.g., Fig. 20). The atom

$$\boldsymbol{\varphi}_{[n-1,1],2-2\cos(\frac{2\pi}{n}),\bar{\pi}_z} = \bar{c}_{[n-1,1]}\mathbf{B}_{\bar{\pi}_z}\mathbf{v}_{[n-1,1],2-2\cos(\frac{2\pi}{n})}$$

therefore features positive values on the vertices of $\mathbb{P}_n$ associated with rankings in which candidate z is ranked towards the top or bottom, and negative values on vertices associated with rankings in which candidate z is ranked in the middle. Accordingly, the analysis coefficient $\alpha_{[n-1,1],2-2\cos(\frac{2\pi}{n})\bar{\pi}_z}$ is large when many voters feel strongly (either positively or negatively) about candidate z. For example, looking at the summary of the analysis coefficients for the 2017 Minneapolis City Council Ward 3 election data in the bottom row of the column for eigenvalue $\lambda = 2 - 2\cos(\frac{2\pi}{4}) = 2$ in Fig. 18, the largest analysis coefficient is the 318.7 associated with the set partition $\{\{2,3,4\},\{1\}\}$, indicating that candidate 1 (Ginger Jentzen) is often ranked in either first or last place. The corresponding coefficients for candidates 2 and 3 are negative, indicating they are often ranked in positions 2 and 3, and are less polarizing. For the sushi preference data, the largest positive analysis coefficients in the $\gamma = [9,1]$, $\lambda = 0.382$ pair (see Fig. 19) indicate items 8 and 5 (fatty tuna and sea urchin) are often ranked quite highly or quite lowly, while item 9 (tuna roll) is often ranked in the middle. The key takeaway is that the combination of a near-zero value of $\alpha_{[n-1,1],2-2\cos(\frac{\pi}{n}),\bar{\pi}_z}$ and a relatively large value of $\alpha_{[n-1,1],2-2\cos(\frac{2\pi}{n}),\bar{\pi}_z}$ indicates that the candidate identified by the singleton in the set partition $\bar{\pi}_z$ is highly polarizing. This is the case in the sushi preference data for the sea urchin item, for which $\alpha_{[n-1,1],2-2\cos(\frac{2\pi}{n}),\bar{\pi}_z} = -0.016$ (lowest magnitude of any item) and $\alpha_{[n-1,1],2-2\cos(\frac{2\pi}{n}),\bar{\pi}_z} = 0.960$ (second highest of the items), indicating that the voting population is roughly split between strongly liking and strongly disliking sea urchin.

5.2.3 The Shapes $\gamma = [n-2,2]$ and $\gamma = [n-2,1,1]$: Pairwise Co-Occurence

Net of their individual popularities, when are two candidates likely to be ranked similarly (either positively or negatively) by voters? Given a voter's first choice candidate, are there other candidates the voter is likely to feel positively, negatively, or neutral about? These are the types of pairwise co-occurence questions that can be answered with the second-order marginal information found in the shapes $\gamma = [n-2,2]$ and $\gamma = [n-2,1,1]$.

For the 2017 Minneapolis City Council Ward 3 election data, the largest analysis coefficient in this shape, shown in bottom-right of Fig. 18, is the 239.0 associated with eigenvalue $\lambda = 1.268$ and the set partition $\pi_z = \{12|34\}$, indicating candidates 3 and 4 are often ranked together in the first two positions or last two positions. This is not surprising as these candidates belong to the same political party. For a small number of candidates, such as $n = 4$ in this case, it is possible to visually inspect the atoms of the form $\boldsymbol{\varphi}_{\gamma,\lambda,\bar{\pi}}$ shown in Fig. 18 in order to interpret the analysis coefficients.
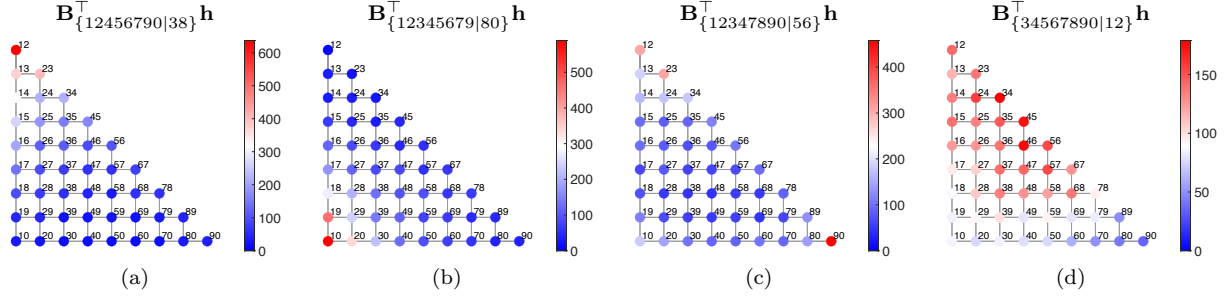
Fig. 21 The sushi preference signal $\mathbf{h}$ projected down from the permutahedron $\mathbb{P}_{10}$ to the Schreier graph $\mathbb{P}_{[8,2]}$ via four different projection matrices that correspond to four different ordered set partitions in $\bar{\Pi}_{[8,2]}$.

For larger values of n, however, it is more convenient to think about the inner products defined in (23) as $\alpha_{\gamma,\lambda,\bar{\pi}} = \bar{c}_\gamma \langle \mathbf{B}_{\bar{\pi}}^\top \mathbf{f}, \mathbf{v}_{\gamma,\lambda} \rangle$. Here, $\mathbf{B}_{\bar{\pi}}^\top \mathbf{f}$ is a projection of the signal from $\mathbb{P}_n$ down to the Schreier graph $\mathbb{P}_\gamma$, the same structure on which the eigenvector $\mathbf{v}_{\gamma,\lambda}$ resides. In Fig. 21, we show four such projections of $\mathbf{h}$ onto $\mathbb{P}_{[8,2]}$ that capture the joint placement of four different pairs of items: 3 and 8 (two favorites), 8 and 10 (one favorite and one of the least preferable items), 5 and 6 (two more polarizing items that are often ranked together near the top or together at the bottom), and 1 and 2 (two generally well liked items that are often ranked towards the top but not at the top). In each of the projections shown in Fig. 21, the value at each vertex is equal to the number of voters who ranked the selected pair of items in the ranking positions contained in the vertex labels; for example, Fig. 21(a) shows that 637 of the 5000 voters placed items 3 and 8 (tuna and fatty tuna) in their top two ranking slots.

The other halves of the inner products $\alpha_{[8,2],\lambda,\bar{\pi}} = \bar{c}_{[8,2]} \langle \mathbf{B}_{\bar{\pi}}^\top \mathbf{f}, \mathbf{v}_{[8,2],\lambda} \rangle$ are the Laplacian eigenvectors of the module $V_{[8,2]}$, the first two of which we show in Fig. 22(c)-(d). The coefficients $\{\alpha_{[8,2],0.2047,\bar{\pi}}\}$ associated with the eigenvector in Fig. 22(c) capture a notion of pairwise proximity. Specifically, a large positive coefficient indicates the two items/candidates grouped by the set partition $\bar{\pi}_z$ are likely to be close to the first two or last two ranking positions, suggesting they may have similar features or belong to the same political party in the case of an election. For the sushi preference data, as shown in Fig. 19, the two largest positive coefficients $\bar{c}_{[8,2]} \langle \mathbf{B}_{\bar{\pi}}^\top \mathbf{f}, \mathbf{v}_{[8,2],\lambda} \rangle$ are the 1.3304 associated with the set partition $\bar{\pi} = \{12456790|38\}$, which groups the two overall favorites fatty tuna and tuna together, and the 1.6543 associated with $\bar{\pi} = \{12345689|70\}$, which groups the two overall least favorites – and the only two vegetarian options – egg and cucumber roll together. In ranked voting elections where there are more than two candidates in the same political party, we might expect to see large positive coefficients on these $[n-2,2]$ analysis coefficients associated with each of the pairs of candidates from the same party. A negative coefficient with large magnitude, on the other hand, indicates the two items are frequently ranked at opposite ends. In the sushi data, the most negative coefficient $\bar{c}_{[8,2]} \langle \mathbf{B}_{\bar{\pi}}^\top \mathbf{f}, \mathbf{v}_{[8,2],\lambda} \rangle$ is the -1.7150 associated with the set partition $\bar{\pi} = \{12345679|80\}$, which groups the overall favorite (fatty tuna) and overall least favorite (cucumber roll) items together.
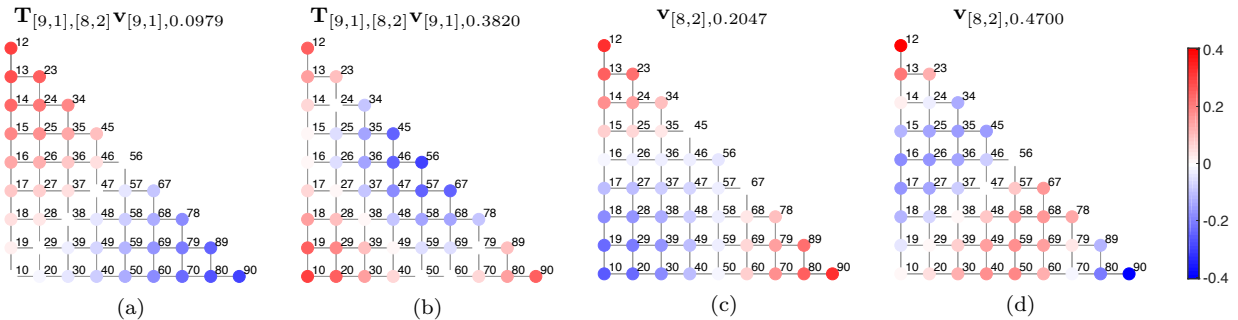

Fig. 22 (a)-(b) The Laplacian eigenvectors $\mathbf{v}_{[9,1],\lambda}$, for $\lambda = 0.0979$ and $\lambda = 0.3820$, lifted from $\mathbb{P}_{[9,1]}$ to $\mathbb{P}_{[8,2]}$. (c)-(d) The Laplacian eigenvectors $\mathbf{v}_{[8,2],\lambda}$, for $\lambda = 0.2047$ and $\lambda = 0.4700$, are orthogonal to each other and to the two vectors in (a) and (b) on the Schreier graph $\mathbb{P}_{[8,2]}$.

Let us now consider the set partition $\bar{\pi} = \{34567890|12\}$, for which the corresponding projection onto $\mathbb{P}_{[8,2]}$ is shown in Fig. 21(d). Of the 45 analysis coefficients $\{\alpha_{[8,2],0.2047,\bar{\pi}}\}$, $\alpha_{[8,2],0.2047,\{34567890|12\}} = -0.0231$ is the third smallest in magnitude, and of the 45 analysis coefficients $\{\alpha_{[8,2],0.4700,\bar{\pi}}\}$, $\alpha_{[8,2],0.4700,\{34567890|12\}} = -0.0088$ is the smallest in magnitude. There are two reasons these coefficients are closer to 0. First, the energy of the projection $\mathbf{B}_{\{34567890|12\}}^\top \mathbf{h}$ in Fig. 21(d) is more evenly spread across the vertices of $\mathbb{P}_{[8,2]}$ than the other projections in Fig. 21(a)-(c) even though they all have the same sum; thus, the norm $\|\mathbf{B}_{\{34567890|12\}}^\top \mathbf{h}\|$ is smaller. Second, from a graph signal processing viewpoint, the energy $\|\mathbf{B}_{\{34567890|12\}}^\top \mathbf{h}\|^2$ decomposes across the Laplacian eigenspaces of the Schreier graph $\mathbb{P}_{[8,2]}$. In this case, most of that energy falls into the one-dimensional spaces spanned by the eigenvectors shown in Fig. 22(a)-(b). In addition to being Laplacian eigenvectors of $\mathbb{P}_{[8,2]}$, these vectors can be viewed as liftings of eigenvectors from the module $V_{[9,1]}$ to $\mathbb{P}_{[8,2]}$. More formally and generally, if $\nu \rhd \gamma$, then for any $\xi \in \bar{\Pi}_\nu$ and $\pi \in \bar{\Pi}_\gamma$, we define the linear mapping $\mathbf{T}_{\xi,\pi} : \mathbb{R}[\Pi_\nu] \to \mathbb{R}[\Pi_\gamma]$ by $\mathbf{T}_{\xi,\pi}\mathbf{x} := \mathbf{B}_\pi^\top \mathbf{B}_\xi \mathbf{x}$ (see Prop. 8). If $K_{\gamma,\nu} = 1$, then $\frac{\mathbf{T}_{\xi,\pi}\mathbf{x}}{\|\mathbf{T}_{\xi,\pi}\mathbf{x}\|} = \pm \frac{\mathbf{T}_{\xi',\pi'}\mathbf{x}}{\|\mathbf{T}_{\xi',\pi'}\mathbf{x}\|}$ for all $\xi, \xi' \in \bar{\Pi}_\nu$, $\pi, \pi' \in \bar{\Pi}_\gamma$, and $\mathbf{x} \in V_\nu$, and we identify all of these transformations under the notation $\mathbf{T}_{\nu,\gamma}$.

Next, we examine the ordered second-order marginals, discussed in Sec. 3.1 and captured in the $[n-2,1,1]$ isotypic component. Returning to the negative coefficient $\alpha_{[8,2],0.2047,\{12345679|80\}} = -1.7150$, potential causes for the large magnitude could be that (i) the first selected candidate (8) is commonly ranked first and the second selected candidate (10) is commonly ranked last, (ii) the first selected candidate is commonly ranked last and the second selected candidate is commonly ranked first, or (iii) both candidates are polarizing, but ranked at opposite ends of the spectrum by different sets of voters. Which other coefficients provide structural information that can be combined with $\alpha_{[8,2],0.2047,\{12345679|80\}}$ to inform which of these

Fig. 23 (a) The Laplacian eigenvector $\mathbf{v}_{[8,2],0.2047}$, lifted from $\mathbb{P}_{[8,2]}$ to $\mathbb{P}_{[8,1,1]}$. (b) The Laplacian eigenvector $\mathbf{v}_{[8,1,1],0.4799}$ is orthogonal to the vector in (a) on the Schreier graph $\mathbb{P}_{[8,1,1]}$.

scenarios might be occurring? In this case, we know from the coefficients $\alpha_{[9,1],0.0979,\{123456789|0\}} = -2.1513$ and $\alpha_{[9,1],0.0979,\{123456790|8\}} = 1.9978$ that item 8 (fatty tuna) is the popular one and item 10 (cucumber roll) is the unpopular one. More generally, we can examine the coefficient $\alpha_{[8,1,1],0.4799,\{12345679|8|0\}}$, which is equal to 1.3471 in this case. From the eigenvector $\mathbf{v}_{[8,1,1],0.4799}$ in Fig. 23(b), we see that a positive value implies the first selected candidate is more often ranked towards the top (the case here), a negative value implies the second selected candidate is more often ranked towards the top, and a coefficient of small magnitude indicates that both candidates are roughly evenly located at the two ends (scenario (iii) above).

An analysis coefficient associated with the first eigenvector of the module $V_{[n-2,1,1]}^*$ can also provide some insight into the interpretation of a large positive analysis coefficient associated with the first eigenvector of the module $V_{[n-2,2]}^*$. Namely, a positive value of the coefficient indicates the second selected candidate is more popular than the first, and vice versa. For example, consider the pairs of items 7/10 and 9/10. From the fact that $\alpha_{[8,2],0.2047,\{12345689|70\}} = 1.6543$ and $\alpha_{[8,2],0.2047,\{12345678|90\}} = 0.7055$ are positive, we conclude these pairs of items often appear together at the top or bottom of the rankings. We know from the fact that the coefficients $\alpha_{[9,1],0.0979,\{123456789|0\}} = -2.1513$, $\alpha_{[9,1],0.0979,\{123456890|7\}} = -1.1896$, and $\alpha_{[9,1],0.0979,\{123456780|9\}} = -0.3733$ are all negative that all three items are unpopular and these pairs therefore appear together more often at the bottom of the rankings. Is one item in each pair more likely to be ranked last? We see from the projections onto $\mathbb{P}_{[8,1,1]}$ in Fig. 24(b) and (d) that when ranked in the last two slots, item 10 is only slightly more likely than item 7 to be ranked last, but item 9 is much more frequently ranked higher than item 10 in these pairwise frequencies. This difference is reflected in the coefficients $\alpha_{[8,1,1],0.4799,\{12345689|7|0\}} = -0.4050$ and $\alpha_{[8,1,1],0.4799,\{12345678|9|0\}} = -0.9136$.
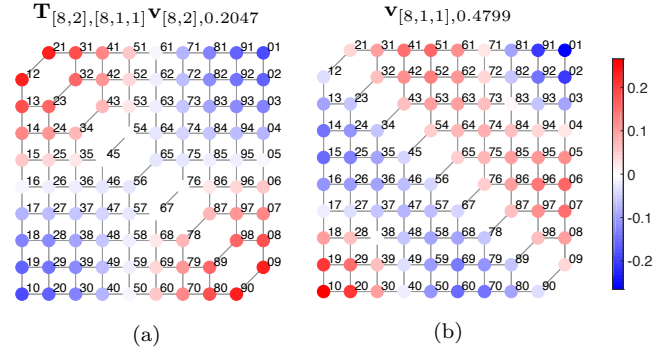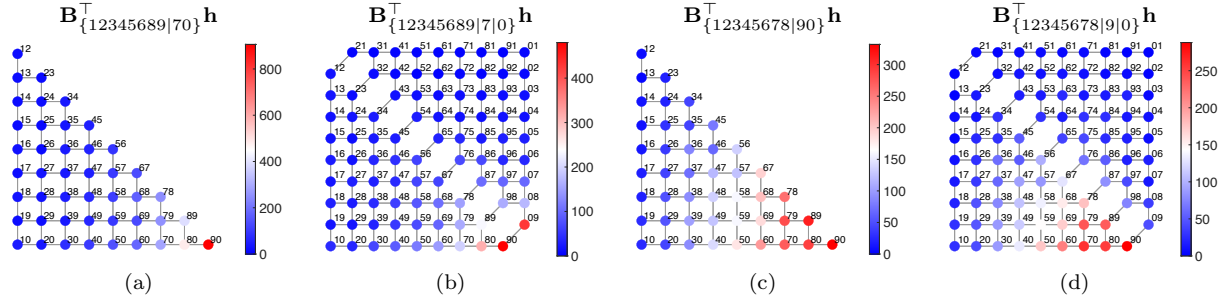
(a)   (b)   (c)   (d)

Fig. 24 Projections capturing the second order ordered and unordered marginals for the pairs of sushi items 7/10 and 9/10.

5.2.4 Shapes with $\gamma_1 < n - 2$

For subsequent shapes, we can interpret the analysis coefficient in a similar manner, as the inner products between projections of the signal down to the Schreier graph and Laplacian eigenvectors $\mathbf{v}_{\gamma,\lambda}$ in V_γ^*. Once again, the most interpretable eigenvectors and therefore informative analysis coefficients are usually those associated with lower eigenvalues (i.e., the first couple from each new irreducible V_γ^*).

As a final example, we consider the analysis coefficients $\alpha_{[7,3],\lambda,\bar{\pi}}$ generated by the values λ equal to 0.3227, 0.5660, and 0.8122 (the associated eigenvectors of which are shown in the bottom row of Fig. 26) and the set partitions $\bar{\pi}$ equal to {1234790|568} and {3567890|124}. The first projection of the data in Fig. 25, corresponding to $\bar{\pi} = \{1234790|568\}$, shows that items 5, 6, and 8 (sea urchin, salmon roe, and fatty tuna) most frequently occur in the top three rankings or rankings 1, 2, and 4, but it is also common that voters ranked two of the three at the top and one at the bottom,

Fig. 25 The sushi preference signal $\mathbf{h}$ projected down from the permutahedron $\mathbb{P}_{10}$ to the Schreier graph $\mathbb{P}_{[7,3]}$ via two different projection matrices that correspond to two of the 120 different ordered set partitions in $\bar{\Pi}_{[7,3]}$.

or one at the top and two at the bottom. Due to the high concentration of the projection on the vertices of the Schreier $\mathbb{P}_{[7,3]}$ labeled by 123 and 124, the inner products with the three eigenvectors in Fig. 26(d)-(f) have the second, seventh, and tenth largest magnitudes of the 9000 analysis coefficients $\{\alpha_{[7,3],\lambda,\bar{\pi}}\}_{\lambda \in \Lambda_{[7,3]}, \bar{\pi} \in \bar{\Pi}_{[7,3]}}$, and the second, first, and first largest magnitudes of the 120 analysis coefficients in each of the subsets $\{\alpha_{[7,3],0.3227,\bar{\pi}}\}_{\bar{\pi} \in \bar{\Pi}_{[7,3]}}$, $\{\alpha_{[7,3],0.5660,\bar{\pi}}\}_{\bar{\pi} \in \bar{\Pi}_{[7,3]}}$, and $\{\alpha_{[7,3],0.8122,\bar{\pi}}\}_{\bar{\pi} \in \bar{\Pi}_{[7,3]}}$, respectively. That is to say, there is a strong third order effect of these three items being ranked together and highly, net of the first order and second order effects, which are captured, e.g., in the lifted eigenvectors in Fig. 26(a)-(c).
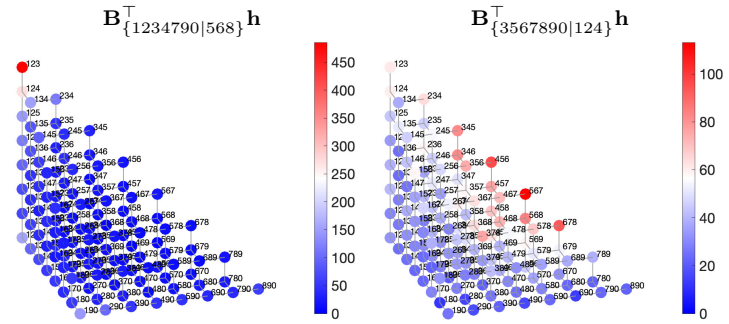
The second projection of the data in Fig. 25 captures the joint ranking positions of three sushi items with cooked fish: 1, 2, and 4 (shrimp, sea eel, and squid). Our suspicion that these three items would be ranked similarly by many voters is confirmed by the projection $\mathbf{B}_{\{3567890|124\}}^\top \mathbf{h}$, which shows the items are most commonly ranked in slots 4/5/6, 5/6/7, or 6/7/8. Despite the frequent closeness of the rankings of these three items, the magnitudes of the inner products with the three eigenvectors in Fig. 26(d)-(f) are quite small (all less than 0.075 and ranking 88th, 72nd, and 52nd out of the magnitudes of the analysis coefficients associated with 120 different set partitions for the respective eigenvalues). The reason for this is that the projection $\mathbf{B}_{\{3567890|124\}}^\top \mathbf{h}$ is quite close (up to a sign change) to $\mathbf{T}_{[9,1],[7,3]}\mathbf{v}_{[9,1],0.3820}$ in Fig. 26(b), and the eigenvectors in Fig. 26(d)-(f) are orthogonal to $\mathbf{T}_{[9,1],[7,3]}\mathbf{v}_{[9,1],0.3820}$. That is, the third order effect of these three cooked fish sushi items being ranked together is weak once the first order effects found in the $[9,1]$ shape have been accounted for.
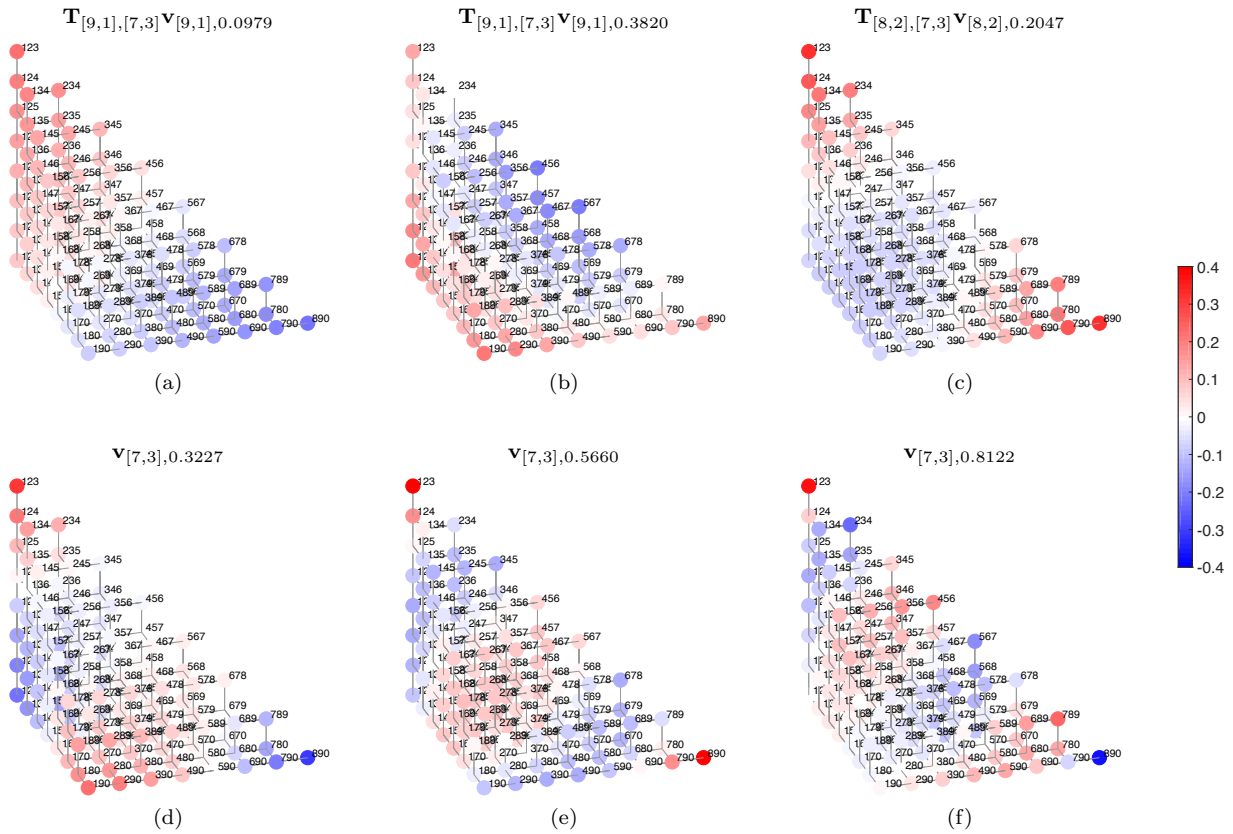
Fig. 26 The three Laplacian eigenvectors $\mathbf{v}_{[7,3],\lambda}$ of $\mathbb{P}_{[7,3]}$ shown in (d)-(f) (and the 72 others not shown) are orthogonal to each other and to the liftings to $\mathbb{P}_{[7,3]}$ of all Laplacian eigenvectors of $P_{[10]}$, $P_{[9,1]}$, and $P_{[8,2]}$ (the Schreier graphs corresponding to the shapes that strictly precede $[7,3]$ in dominance ordering), including the three liftings shown in (a)-(c).

6 Computationally Efficient Algorithms

In this section, we develop efficient algorithms for the proposed transform, and discuss details of our openly available software that implements the transform and helps users find structure in ranked data by visualizing the analysis coefficients and/or projecting the data into lower dimensional spaces to visualize.

What are the initial computational challenges? First, naively computing the eigenvectors of $\mathbb{P}_n$ is not feasible for n above 7 or 8, as it has complexity $\mathcal{O}([n!]^3)$. Second, even naively computing the eigenvectors of the required Schreier graphs is not feasible (complexity $\mathcal{O}(\left[\frac{n!}{\lceil n/2 \rceil!}\right]^3)$). Third, any method that explicitly computes and stores all dictionary atoms of length $n!$ on the permutahedron in order to take inner products with the ranked data will quickly run into memory issues as n grows. Fourth, the number of dictionary atoms in $\mathcal{D}$ from Thm. 1 is $\sum_{\gamma \vdash n} d_\gamma m_\gamma$, which grows faster than $n!$; i.e., the redundancy of the dictionary increases as n increases.

In order to circumvent these issues, we need specialized algorithms (i.e., not standard signal processing or numerical linear algebra techniques) that take advantage of the symmetry and structure present in the permutahedron.

Our general approach is to include as much of the computation as possible into a setup portion that is independent of the data and can therefore be performed just once, offline and ahead of time, for each n. This setup portion consists of three phases: the dynamic constructions of the adjacency matrix and the characteristic matrix $\mathbf{B}_\gamma$ for each Schreier graph $\mathbb{P}_\gamma$ (Sec. 6.1), the computation of the Laplacian eigenvalues and eigenvectors of the Schreier graphs (Sec. 6.2), and the computation of a path from each reading set partition to every other

vertex in the Schreier graph (Sec. 6.3). With all of this information stored for a given n, the data-dependent analysis portion of the code (Sec. 6.3) needs to be executed for each new ranked data vector $\mathbf{f}$.

6.1 Construction of the Schreier Adjacency Matrices and a Lifting Matrix from Each Schreier to the Permutahedron

The Schreier graphs for $\mathbb{S}_n$ can be constructed in an iterative dynamic manner from those for $\mathbb{S}_{n-1}$. If $\gamma = [\gamma_1, \gamma_2, \ldots, \gamma_\ell] \vdash n$, then for each $1 \leq i \leq \ell$, let $\Pi_{\gamma,i} \subseteq \Pi_\gamma$ be the subset of ordered set partitions μ with n in the ith row of μ and let $\mathbb{P}_{\gamma,i}$ be the subgraph of $\mathbb{P}_\gamma$ induced by $\Pi_{\gamma,i}$. In this way, $\mathbb{P}_\gamma$ partitions into a disjoint union of subgraphs $\{\mathbb{P}_{\gamma,i}\}_{i=1}^\ell$ with edges between vertices in different subgraphs as follows: $\mu \in \mathbb{P}_{\gamma,i}$ is connected to $\nu \in \mathbb{P}_{\gamma,j}$ $(i \neq j)$ by an edge labeled by $(n-1, n)$ if and only if $\mu = (n-1, n)(\nu)$. Note that $\mathbb{P}_{\gamma,i} \cong \mathbb{P}_{\gamma'}$ where γ' is the integer partition given by subtracting 1 from γ_i and, if necessary, sorting the parts so they are nonincreasing. For example, Fig. 27 gives the decomposition of $\mathbb{P}_{[4,1,1]}$ into $\mathbb{P}_{[3,1,1]}$, $\mathbb{P}_{[4,0,1]}$, and $\mathbb{P}_{[4,1,0]}$ with $\mathbb{P}_{[4,0,1]} \cong \mathbb{P}_{[4,1,0]} \cong \mathbb{P}_{[4,1]}$. This decomposition allows one to construct the Schreier graphs of $\mathbb{S}_n$ dynamically from those for $\mathbb{S}_{n-1}$. Since the permutahedron is a Schreier graph, $\mathbb{P}_n = \mathbb{P}_{[1,1,\ldots,1]}$, we also construct it dynamically using this method.
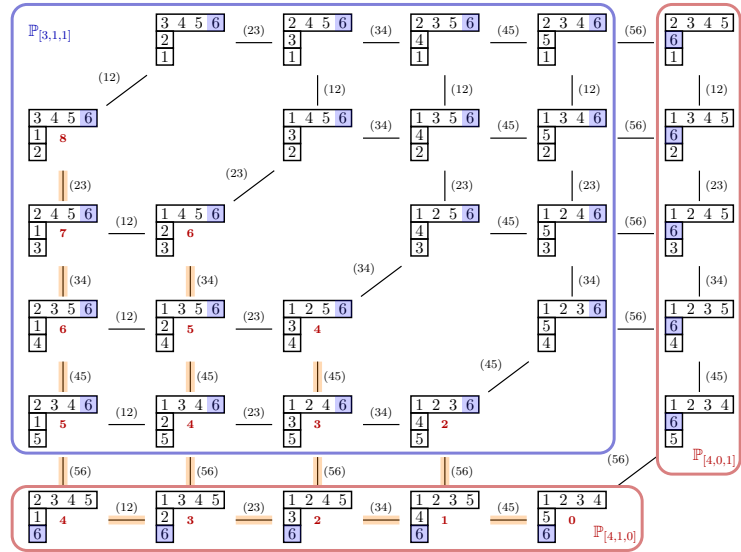

Fig. 27 The decomposition of the Schreier graph $\mathbb{P}_{[4,1,1]}$ into $\mathbb{P}_{[3,1,1]}, \mathbb{P}_{[4,1,0]},$ and $\mathbb{P}_{[4,0,1]}$. Note that $\mathbb{P}_{[4,1,0]}$ and $\mathbb{P}_{[4,0,1]}$ are each isomorphic to $\mathbb{P}_{[4,1]}$. The red numbers indicate the number of adjacent transpositions to get to the reading-order set partition π_1, and the highlighted paths are minimal paths from each $\bar{\pi} \in \bar{\Pi}_{[4,1,1]}$ to π_1, constructed via a breadth-first search.

Let $\mathbf{B}_\gamma := \mathbf{B}_{\pi_1}$ denote the characteristic matrix corresponding to the reading ordered set partition $\pi_1 \in \Pi_\gamma$, which is a matrix that lifts Schreier eigenvectors from the Schreier $\mathbb{P}_\gamma$ to the permutahedron $\mathbb{P}_n$. Then $\mathbf{B}_\gamma$ has a recursive structure that respects the decomposition of $\mathbb{P}_\gamma$, and, as we show in Sec. 6.3, this is the only lifting matrix that we need to compute for shape γ. If $\gamma = [\gamma_1, \ldots, \gamma_\ell] \vdash n$, then for $1 \leq i \leq n$, let $\mathbb{S}_{n,i} := \{\sigma \in \mathbb{S}_n \mid \sigma(n) = i\}$. The decomposition of $\mathbf{B}_\gamma$ is given in Prop. 7 and illustrated in Ex. 1.

Proposition 7 *If* $\gamma = [\gamma_1, \ldots, \gamma_\ell] \vdash n$, *then the characteristic matrix* $\mathbf{B}_\gamma$ *decomposes into block sub-matrices,*

$$(\mathbf{B}_\gamma)_{i,j} = \begin{cases} \mathbf{B}_{\gamma(j)}, & \text{if } i \text{ is in the } j\text{th row of } \pi_1, \\ \mathbf{0}, & \text{otherwise,} \end{cases} \qquad 1 \leq i \leq n, \quad 1 \leq j \leq \ell,$$

where $\gamma(j)$ *is obtained from* γ *by subtracting 1 from* γ_j.

Proof For $\sigma \in \mathbb{S}_n$ and $\mu \in \Pi_\gamma$, the (σ, μ)-entry of $\mathbf{B}_\gamma$ equals 1 if and only if $\sigma(\mu) = \pi_1$. If n is in the jth block of μ and $\sigma(n) = i$, then this entry can be nonzero only if i is in the jth row of π_1. The nonzero blocks equal $\mathbf{B}_{\gamma(j)}$ by definition (after ignoring n). $\qquad\square$

Example 1 In Fig. 28, we show the recursive structure of the characteristic matrix $\mathbf{B}_{[2,2]}$, which corresponds to the ordered set partition $\pi_1 = \{12|34\}$ (compare with Fig. 14, which gives this same matrix with the permutations in lexiocgraphic order). Similarly, the recursive structure of $\mathbf{B}_{[4,1,1]}$, which corresponds to the

$$\mathbf{B}_{[2,2]} = \mathbf{B}_{\boxed{\begin{smallmatrix}1&2\\3&4\end{smallmatrix}}} = \begin{array}{c}\mathbb{S}_{4,4}\\\mathbb{S}_{4,3}\\\mathbb{S}_{4,2}\\\mathbb{S}_{4,1}\end{array}\begin{bmatrix}\Pi_{[2,2],2} & \Pi_{[2,2],1}\\ \mathbf{B}_{[2,1]} & \cdot\\ \mathbf{B}_{[2,1]} & \cdot\\ \cdot & \mathbf{B}_{[1,2]}\\ \cdot & \mathbf{B}_{[1,2]}\end{bmatrix} = $$

	$\begin{smallmatrix}1\,2\\3\,4\end{smallmatrix}$	$\begin{smallmatrix}1\,3\\2\,4\end{smallmatrix}$	$\begin{smallmatrix}2\,3\\1\,4\end{smallmatrix}$	$\begin{smallmatrix}1\,4\\2\,3\end{smallmatrix}$	$\begin{smallmatrix}2\,4\\1\,3\end{smallmatrix}$	$\begin{smallmatrix}3\,4\\1\,2\end{smallmatrix}$
1234	1	.	.	.	.	.
2134	1	.	.	.	.	.
1324	.	1	.	.	.	.
3124	.	.	1	.	.	.
2314	.	1	.	.	.	.
3214	.	.	1	.	.	.
1243	1	.	.	.	.	.
2143	1	.	.	.	.	.
1423	.	1	.	.	.	.
4123	.	.	1	.	.	.
2413	.	1	.	.	.	.
4213	.	.	1	.	.	.
1342	.	.	.	1	.	.
3142	.	.	.	.	1	.
1432	.	.	.	1	.	.
4132	.	.	.	.	1	.
3412	.	.	.	.	.	1
4312	.	.	.	.	.	1
2341	.	.	.	1	.	.
3241	.	.	.	.	1	.
2431	.	.	.	1	.	.
4231	.	.	.	.	1	.
3421	.	.	.	.	.	1
4321	.	.	.	.	.	1

Fig. 28 The recursive structure of the characteristic matrix $\mathbf{B}_{[2,2]}$.

ordered set partition $\pi_1 = \{1234|5|6\}$, is given by

$$\mathbf{B}_{[4,1,1]} = \mathbf{B}_{\boxed{\begin{smallmatrix}1\,2\,3\,4\\5\\6\end{smallmatrix}}} = \begin{array}{c}\mathbb{S}_{6,6}\\\mathbb{S}_{6,5}\\\mathbb{S}_{6,4}\\\mathbb{S}_{6,3}\\\mathbb{S}_{6,2}\\\mathbb{S}_{6,1}\end{array}\begin{bmatrix}\Pi_{[4,1,1],3} & \Pi_{[4,1,1],2} & \Pi_{[4,1,1],1}\\ \mathbf{B}_{[4,1,0]} & \cdot & \cdot\\ \cdot & \mathbf{B}_{[4,0,1]} & \cdot\\ \cdot & \cdot & \mathbf{B}_{[3,1,1]}\\ \cdot & \cdot & \mathbf{B}_{[3,1,1]}\\ \cdot & \cdot & \mathbf{B}_{[3,1,1]}\\ \cdot & \cdot & \mathbf{B}_{[3,1,1]}\end{bmatrix}.$$

6.2 Computation of the Schreier Eigenvalues and Eigenvectors Via Deflation

Frame vectors $\varphi_{\gamma,\lambda,k,\bar{\pi}}$ of eigenvalue λ and shape γ are computed as lifts from the λ-eigenspace of the Schreier graph corresponding to the irreducible component V_γ^* inside of $\mathbb{R}[\Pi_\gamma]$. The multiplicity of V_γ^* in $\mathbb{R}[\Pi_\gamma]$ is equal to one and the other irreducible components V_ν^* that appear in $\mathbb{R}[\Pi_\gamma]$ correspond to shapes ν with $\nu \rhd \gamma$ in dominance order. See (13) and Fig. 29. By computing the Laplacian eigenvectors in any order that respects dominance, we will have already computed the eigenvectors $\mathbf{v}_{\nu,\varrho,k}$ for $\nu \rhd \gamma$. To single out V_γ^* inside $\mathbb{R}[\Pi_\gamma]$, we lift the Laplacian eigenvectors from $\mathbb{P}_\nu$ up to $\mathbb{P}_\gamma$, and find the eigenvectors associated with the eigenvalues of V_γ^* as an orthogonal complement of this span.

Interesting new theory is needed here, as the multiplicity of V_ν^* in $\mathbb{R}[\Pi_\gamma]$ equals the Kostka number, and therefore we need to lift the eigenvectors of shape ν in $K_{\gamma,\nu}$ linearly independent ways to $\mathbb{R}[\Pi_\gamma]$. In Prop. 8 below, we show that lifting and projecting with set partitions constructed from "column-strict" tableaux is the same as Young's rule for decomposing $\mathbb{R}[\Pi_\gamma]$ into irreducible modules with multiplicity $K_{\gamma,\nu}$. Therefore, using mappings that come from different column-strict tableaux give us linearly independent images in $\mathbb{R}[\Pi_\gamma]$ as needed.

The Kostka number $K_{\nu,\gamma}$ equals the number of column-strict tableaux of shape ν and *content* γ; that is, a tableau constructed by filling the boxes of ν with γ_1 ones, γ_2 twos, and so on, such that the rows weakly increase and the columns strictly increase. For example, the $K_{\gamma,\nu} = 3$ column-strict tableaux T_1, T_2, T_3 of shape $\nu = [5, 4]$ and content $\gamma = [4, 2, 2, 1]$ are shown in Fig. 30. Note that from this definition, $K_{\gamma,\nu} = 0$ if $\nu \not\trianglerighteq \gamma$. If T is a column-strict tableaux of shape ν and content γ, then define $\xi_T \in \Pi_\gamma$ to be the set partition with j in row r if the jth box of T_i, read in reading order, contains r. By reading order, we mean left-to-right across the rows from top to bottom. The set partitions $\{\xi_{T_1}, \xi_{T_2}, \xi_{T_3}\}$ are shown in Fig. 30. For example, ξ_{T_2} has 5 in row 3 because the 5th box of T_2 contains 3.

Proposition 8 *If $\pi_1 \in \Pi_\nu$ is the reading-order set partition and $\{\xi_i\}_{i=1}^{K_{\gamma,\nu}} \subseteq \Pi_\gamma$ are the set partitions corresponding to the column-strict tableaux $\{T_i\}_{i=1}^{K_{\gamma,\nu}}$ of shape ν and content γ, then the matrices $\mathbf{T}_{\xi_i,\pi_1} = \mathbf{B}_{\xi_i}^\top \mathbf{B}_{\pi_1}$, $1 \leq i \leq K_{\gamma,\nu}$, give $\mathbb{S}_n$-module isomorphisms between $V_\nu^* \subseteq \mathbb{R}[\Pi_\nu]$ and $K_{\gamma,\nu}$ linearly independent copies $\{V_{\nu,i}^*\}_{i=1}^{K_{\gamma,\nu}}$ of V_ν^* in $\mathbb{R}[\Pi_\gamma]$.*

Proof The transformation $\mathbf{T}_{\xi_i,\pi_1}$ is a right $\mathbb{S}_n$-module homomorphism, since both $\mathbf{B}_{\xi_i}, \mathbf{B}_{\pi_1}$ are. We will show that $\mathbf{T}_{\xi_i,\pi_1} = c_i \Theta_{T_i}$, for a nonzero scalar c_i, where Θ_{T_i} is the $\mathbb{S}_n$-module isomophism given by Young's rule defined in [83, Def. 2.9.3]. These transformations Θ_{T_i} are known to give $K_{\gamma,\nu}$ linearly independent isomorphisms from $V_\nu^* \in \mathbb{R}[\Pi_\nu]$ to $V_{\nu,i}^* \in \mathbb{R}[\Pi_\gamma]$ by [83, Thm. 2.11.2].

Since $\mathbf{T}_{\xi_i,\pi_1}$ is a module homomorphism, and $\mathbb{S}_n$ acts transitively on Π_ν, it is sufficient to show that they are equal at $\mathbf{e}_{\pi_1}$. Let $\mathbb{S}_{\pi_1} = \{\sigma \in \mathbb{S}_n \mid \sigma(\pi_1) = \pi_1\}$ be the stabilizer subgroup of π_1. Then

$$\mathbf{B}_{\xi_i}^\top \mathbf{B}_{\pi_1} \mathbf{e}_{\pi_1} = \sum_{\sigma \in \mathbb{S}_{\pi_1}} \mathbf{B}_{\xi_i}^\top \mathbf{e}_\sigma = \sum_{\sigma \in \mathbb{S}_{\pi_1}} \sum_{\mu \in \Pi_\gamma, \sigma(\mu) = \xi_i} \mathbf{e}_\mu = \sum_{\sigma \in \mathbb{S}_{\pi_1}} \mathbf{e}_{\sigma^{-1}(\xi_i)} = \sum_{\sigma \in \mathbb{S}_{\pi_1}} \mathbf{e}_{\sigma(\xi_i)}.$$

Thus the coefficient of $\mathbf{e}_\mu$ in $\mathbf{B}_{\xi_i}^\top \mathbf{B}_{\pi_1} \mathbf{e}_\pi$ is $c_{\mu,\xi_i,\pi_1} = \#\{\sigma \in \mathbb{S}_n \mid \sigma(\xi_i) = \mu, \sigma(\pi_1) = \pi_1\}$. This set is a coset of the stabilizer of (ξ_i, π_1) under the action of $\mathbb{S}_n$ on $\Pi_\gamma \times \Pi_\nu$, so its cardinality is constant for all μ, which we denote by c_{ξ_i} (equal to the size of this stabilizer). By comparing with the action of Θ_{T_i} in [83, Def. 2.9.3] (identifying ordered set partitions here with "tabloids" there), we have that $\mathbf{T}_{\xi_i,\pi_1} = c_\xi \Theta_{T_i}$, where ξ_i is the set partition filled according to T_i. □

		$V^*_{\square\square\square\square\square\square}$	$V^*_{\square\square\square\square}$	$V^*_{\square\square\square}$	$V^*_{\square\square\square}$	$V^*_{\square\square}$	$V^*_{\square\square}$
		1	5	9	10	5	16
$\mathbb{R}[\Pi_{\square\square\square\square\square\square}]$	1	1					
$\mathbb{R}[\Pi_{\square\square\square\square}]$	6	1	1				
$\mathbb{R}[\Pi_{\square\square\square}]$	15	1	1	1			
$\mathbb{R}[\Pi_{\square\square\square}]$	30	1	2	1	1		
$\mathbb{R}[\Pi_{\square\square}]$	20	1	1	1	0	1	
$\mathbb{R}[\Pi_{\square\square}]$	60	1	2	2	1	1	1

Fig. 29 The (upper-left quadrant of the) table of Kostka numbers $K_{\gamma,\nu}$ giving the multiplicity of the irreducible module V_ν^* as a component of the module $\mathbb{R}[\Pi_\gamma]$. The dimension d_ν of V_ν^* is given below it in red and the dimension m_γ of $\mathbb{R}[\Pi_\gamma]$ is given to its right in blue. The dimensions sum as $m_\gamma = \sum_{\nu \vdash n} K_{\gamma,\mu} d_\nu$; for example, $60 = 1 \cdot 1 + 2 \cdot 5 + 2 \cdot 9 + 1 \cdot 10 + 1 \cdot 5 + 1 \cdot 16$.

$$\nu = \;\;\boxed{\;\;\;\;} \qquad \pi_1 = \begin{array}{|c|c|c|c|c|}\hline 1&2&3&4&5\\\hline 6&7&8&9\\\cline{1-4}\end{array} \qquad T_i: \; \begin{array}{|c|c|c|c|c|}\hl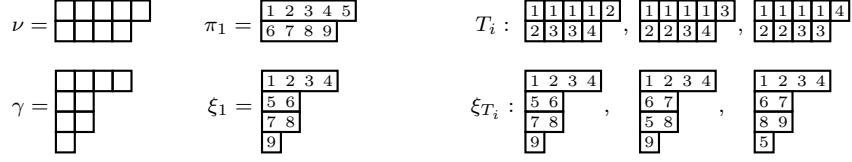ine 1&1&1&1&2\\\hline 2&3&3&4\\\cline{1-4}\end{array}, \; \begin{array}{|c|c|c|c|c|}\hline 1&1&1&1&3\\\hline 2&2&3&4\\\cline{1-4}\end{array}, \; \begin{array}{|c|c|c|c|c|}\hline 1&1&1&1&4\\\hline 2&2&3&3\\\cline{1-4}\end{array}$$

$$\gamma = \;\;\boxed{\;\;\;\;} \qquad \xi_1 = \begin{array}{|c|c|c|c|}\hline 1&2&3&4\\\hline 5&6\\\cline{1-2}7&8\\\cline{1-2}9\\\cline{1-1}\end{array} \qquad \xi_{T_i}: \; \begin{array}{|c|c|c|c|}\hline 1&2&3&4\\\hline 5&6\\\cline{1-2}7&8\\\cline{1-2}9\\\cline{1-1}\end{array}, \; \begin{array}{|c|c|c|c|}\hline 1&2&3&4\\\hline 6&7\\\cline{1-2}5&8\\\cline{1-2}9\\\cline{1-1}\end{array}, \; \begin{array}{|c|c|c|c|}\hline 1&2&3&4\\\hline 6&7\\\cline{1-2}8&9\\\cline{1-2}5\\\cline{1-1}\end{array}$$

Fig. 30 When $\nu = [5,4]$ and $\gamma = [4,2,2,1]$, the Kostka number is $K_{\gamma,\nu} = 3$. The three column-strict tableaux T_1, T_2, T_3 of shape ν and content γ are shown here. They each have $\gamma_1 = 4$ ones, $\gamma_2 = 2$ twos, $\gamma_3 = 2$ threes, and $\gamma_4 = 1$ four. The set partitions $\pi_1 \in \Pi_\nu$ and $\xi_1 \in \Pi_\gamma$ are in reading order, and $\xi_{T_1}, \xi_{T_2}, \xi_{T_3} \subseteq \Pi_\gamma$ are the set-partitions with T_i fillings; that is ξ_{T_i} has j is row r if the jth box of T_i (read in reading order) contains r.

To summarize, from the structure of the Schreier graphs, we know that the graph Laplacian of the Schreier graph $\mathbb{P}_\gamma$ has the following spectral decomposition:

$$\boldsymbol{\mathcal{L}}_{\mathbb{P}_\gamma} = \sum_{\lambda \in \Lambda_\gamma} \sum_{k=1}^{\kappa_{\gamma,\lambda}} \lambda \mathbf{v}_{\gamma,\lambda,k} \mathbf{v}_{\gamma,\lambda,k}^\top + \sum_{\nu \rhd \gamma} \sum_{i=1}^{K_{\gamma,\nu}} \sum_{\lambda \in \Lambda_\nu} \sum_{k=1}^{\kappa_{\nu,\lambda}} \lambda (\mathbf{T}_{\xi_i,\pi_1}^{\nu,\gamma} \mathbf{v}_{\nu,\lambda,k})(\mathbf{T}_{\xi_i,\pi_1}^{\nu,\gamma} \mathbf{v}_{\nu,\lambda,k})^\top.$$

Thus, when the Laplacian eigenvectors of the Schreier graphs that precede $\mathbb{P}_\gamma$ in dominance order have already been computed, to find an orthonormal eigenbasis for V_γ^*, it suffices to diagonalize the rank d_γ matrix

$$\boldsymbol{\mathcal{L}}_{\mathbb{P}_\gamma} - \sum_{\nu \rhd \gamma} \sum_{i=1}^{K_{\gamma,\nu}} \sum_{\lambda \in \Lambda_\nu} \sum_{k=1}^{\kappa_{\nu,\lambda}} \lambda (\mathbf{T}_{\xi_i,\pi_1}^{\nu,\gamma} \mathbf{v}_{\nu,\lambda,k})(\mathbf{T}_{\xi_i,\pi_1}^{\nu,\gamma} \mathbf{v}_{\nu,\lambda,k})^\top, \tag{25}$$

as opposed to the rank m_γ matrix $\boldsymbol{\mathcal{L}}_{\mathbb{P}_\gamma}$. The complexity of forming and computing the eigendecomposition of the matrix in (25) is $\mathcal{O}(m_\gamma d_\gamma^2 + m_\gamma^2 (\sum_{\nu \rhd \gamma} d_\nu K_{\gamma,\nu}))$.

6.3 Efficient Computation of the Analysis Coefficients

As detailed in (23), the analysis coefficients associated with shape γ can be computed either by lifting each eigenvector $\mathbf{v}_{\gamma,\lambda,k}$ up to the permutahedron in z_γ different ways and taking the inner product of each with the signal $\mathbf{f}$, or by projecting the signal down to the Schreier graph in z_γ different ways and taking the inner product between each of the projections and each eigenvector $\mathbf{v}_{\gamma,\lambda,k}$. Since the characteristic matrix used to lift or project between the Schreier $\mathbb{P}_\gamma$ and the permutahedron $\mathbb{P}_n$ is a sparse matrix with $n!$ entries equal to 1 (one per row) and the remainder equal to 0, the respective complexities of these two approaches are $\mathcal{O}(n!d_\gamma z_\gamma)$ and $\mathcal{O}(z_\gamma(n! + d_\gamma m_\gamma))$. However, the more problematic issue with both of these approaches is the memory required to store all z_γ characteristic matrices $\mathbf{B}_{\bar{\pi}}$ associated with each shape γ, which is $4n!z_\gamma$ bytes (e.g., storing these $\mathbf{B}_{\bar{\pi}}$ matrices for all $\bar{\pi} \in \bar{\Pi}_{[4,3,2,1]}$ alone would require approximately 383GB).

To reduce the required memory, we use the following method which only requires storing the single characteristic matrix $\mathbf{B}_\gamma = \mathbf{B}_{\pi_1}$ associated with the reading-ordered set partition π_1 for each shape γ; these matrices are computed dynamically, as detailed in Sec. 6.1. For each $\bar{\pi} \in \bar{\Pi}_\gamma$, we let $\sigma \in \mathbb{S}_n$ be a permutation such that $\sigma(\pi_1) = \bar{\pi}$. Then we have

$$\langle \mathbf{f}, \boldsymbol{\varphi}_{\gamma,\lambda,k,\bar{\pi}} \rangle = \bar{c}_\gamma \langle \mathbf{f}, \mathbf{B}_{\bar{\pi}} \mathbf{v}_{\gamma,\lambda,k} \rangle = \bar{c}_\gamma \langle \mathbf{f}, \mathbf{B}_{\sigma(\pi_1)} \mathbf{v}_{\gamma,\lambda,k} \rangle = \bar{c}_\gamma \langle \mathbf{f}, \rho_L(\sigma) \mathbf{B}_{\pi_1} \mathbf{v}_{\gamma,\lambda,k} \rangle = \bar{c}_\gamma \langle \mathbf{B}_{\pi_1}^\top \rho_L(\sigma^{-1}) \mathbf{f}, \mathbf{v}_{\gamma,\lambda,k} \rangle,$$

where the third equality follows from Prop. 4 and the fourth equality follows from the orthogonality of $\rho_L(\sigma)$. Therefore, to compute the analysis coefficient $\langle \mathbf{f}, \boldsymbol{\varphi}_{\gamma,\lambda,k} \rangle$ we can compute instead $\langle \mathbf{B}_{\pi_1}^\top \rho_L(\sigma^{-1}) \mathbf{f}, \mathbf{v}_{\gamma,\lambda,k} \rangle$, which amounts to reordering the data vector $\mathbf{f}$ by σ^{-1}, projecting it down to the Schreier graph using a single matrix $\mathbf{B}_{\pi_1}$, and then taking the inner product with $\mathbf{v}_{\gamma,\lambda,k}$.

The permutation σ^{-1} is recorded as the product of a minimal sequence of adjacent transpositions, and the reorderings are computed by sequentially applying these transpositions. Thus, the third and final phase of the setup portion consists of computing σ^{-1} for each shape γ and every $\bar{\pi} \in \bar{\Pi}_\gamma$ by constructing a *path* in

$\mathbb{P}_\gamma$ from the reading-order set partition $\pi_1 \in \bar{\Pi}_\gamma$ to $\bar{\pi} \in \bar{\Pi}_\gamma$, which is a sequence $(\pi_1, \pi_2, \ldots, \pi_r)$ such that $\pi_{i+1} = (j_i, j_i + 1)(\pi_i)$ for $i = 1, \ldots, r-1$ and $\pi_r = \bar{\pi}$. Thus, paths give the sequence of adjacent transpositions that transform π_1 into $\bar{\pi}$. A path is *minimal* if there is no path in $\mathbb{P}_\gamma$ from π_1 to $\bar{\pi}$ with fewer edges (equivalently, no way to transform π_1 to $\bar{\pi}$ with fewer swaps). Minimal paths can be constructed via a breadth-first search [10]. One example of a minimal path constructed in this manner is shown in Fig. 27. Moreover, all minimal paths to $\bar{\pi} \in \bar{\Pi}_\gamma$ live entirely in $\bar{\Pi}_\gamma$. To see this, we use the fact that the length of a minimal path to $\pi \in \Pi_\gamma$ equals the number of inversions in π, where an inversion is a pair (i, j) with $i < j$ and j in a higher row than i in π (see [3] for a proof). Thus, inversion numbers go up at each step in a minimal path from π_1 to π. A minimal path never leaves $\bar{\Pi}_\gamma$ and returns to $\bar{\Pi}_\gamma$, for that would require the number of inversions to go down. So for each $\bar{\pi} \in \bar{\Pi}_\gamma$, we pre-compute and save the (length $n!$) permutation that rearranges $\mathbf{f}$ into $\rho_L(\sigma^{-1})\mathbf{f}$. These permutations are constructed by first dynamically making the permutation corresponding to each of the $n-1$ adjacent swaps. Then we perform a tree traversal on the Schreier graph and multiply adjacent swaps to get the permutation corresponding to each Schreier vertex in $\bar{\Pi}_\gamma$. For each Schreier graph $\mathbb{P}_\gamma$, the complexity of computing these permutations is $\mathcal{O}(z_\gamma n!)$.

Finally, in the analysis portion of the code (the only part that is dependent on the data), for each $\gamma \vdash n$ and each $\bar{\pi} \in \bar{\Pi}_\gamma$, the vector $\mathbf{f}$ is reordered into $\rho_L(\sigma^{-1})\mathbf{f}$ by the stored permutation and projected by $\mathbf{B}_{\pi_1}^\top$ to the Schreier graph $\mathbb{P}_\gamma$, where its inner product with each of the Laplacian eigenvectors $\{\mathbf{v}_{\gamma,\lambda,k}\}$ is computed and multiplied by the constant $\bar{c}_\gamma$. This portion has complexity $\mathcal{O}(z_\gamma(2n! + m_\gamma d_\gamma))$ for each shape γ.

Remark 6 The implementation as described thus far requires us to store z_γ permutation vectors of length $n!$ for each shape, which can lead to memory issues as n grows (e.g., for $n = 10$, the permutation vectors associated with the liftings in $\bar{\bar{\Pi}}_{[4,3,2,1]}$ require approximatley 170 GB of memory). Thus, for larger n, we also implement an alternative version of the code that is more memory efficient. In this second version, we save only the permutations corresponding to the adjacent transpositions and the lists of adjacent transpositions leading from reading-ordered set partitions to the other vertices of the Schreier graphs. During the analysis phase, we perform a tree traversal of each Schreier graph, and at each step, the data vector is permuted by the adjacent transposition corresponding to the edge in the tree. The rearranged vectors are then projected down to the corresponding Schreier graph, where the inner products with the Laplacian eigenvectors of $\mathbb{P}_\gamma$ are computed. While this variant is more efficient from a memory standpoint, the downside is that the work of computing the permutations that rearrange $\mathbf{f}$ into each $\rho_L(\sigma^{-1})\mathbf{f}$ from phase 3 of the setup portion now needs to be done for each new data vector $\mathbf{f}$.

6.4 Subsampling of the Dictionary Atoms

The total number of dictionary atoms in $\mathcal{D}$ from Thm. 1 is $\sum_{\gamma \vdash n} d_\gamma m_\gamma$. We briefly mention three ways to reduce the number of atoms in order to improve the computational efficiency of the applying the analysis operator.

First, we always use the less redundant dictionary $\bar{\mathcal{D}}$ from Thm. 2, which provides exactly the same information but avoids identical atoms in order to reduce the total number of atoms to $\sum_{\gamma \vdash n} d_\gamma z_\gamma$. However, this quantity still grows faster than $n!$, meaning that the redundancy of the dictionary $\bar{\mathcal{D}}$ also increases as n increases.

Second, in many applications, the most relevant information lies in the isotypic components associated with the first handful of symmetry types and/or the Laplacian eigenspaces of the permutahedron associated with the lowest eigenvalues. In this case, we do not need to compute all of the analysis coefficients, which significantly reduces the overall complexity, as the computational bottlenecks lie in the symmetry types that are later in the dominance order. If the energy decomposition onto each isotypic component is still desired, we can leverage the bipartite nature of the permutahedron by computing the inner products between the atoms associated with the transpose shape with the element-wise product of the signal and the sign vector of the permutahedron; for example,

$$\sum_{z=1}^{15} \left| \left\langle \mathbf{f}, \varphi_{\boxplus,7.1774,\bar{\pi}_z} \right\rangle \right|^2 = \sum_{z=1}^{10} \left| \left\langle \bar{\mathbf{f}}, \varphi_{\boxplus,0.8226,\bar{\pi}_z} \right\rangle \right|^2,$$

where $\bar{f}(\sigma) = \text{sign}(\sigma)f(\sigma)$. We denote by $\mathcal{F}_n$ the set of integer partitions of n that cannot be written as the transpose of a symmetry type that precedes it in lexicographic order (e.g., when $n = 10$, $\mathcal{F}_n$ contains 22 of the 42 symmetry types, which are shown below in Fig. 33).

Third, and most efficaciously, we can simply compute transform coefficients for atoms associated with the first k symmetry types, which again is often where the most interesting information resides. As an example, if $n = 10$, $|\mathcal{D}| = \sum_{\gamma \vdash n} d_\gamma m_\gamma = 419,571,370$ (redundancy factor of 115.6); $|\bar{\mathcal{D}}| = \sum_{\gamma \vdash n} d_\gamma z_\gamma = 44,711,456$ (10.7% of $|\mathcal{D}|$, redundancy factor of 12.3); $\sum_{\gamma \in \mathcal{F}_n} d_\gamma z_\gamma = 18,004,348$ (40.3% of $|\bar{\mathcal{D}}|$, redundancy factor of 5.0); there are $\sum_{\gamma \in \mathcal{F}_n^8} d_\gamma z_\gamma = 98,866$ atoms generated from the first $k = 8$ shapes (0.55% of $\sum_{\gamma \in \mathcal{F}_n} d_\gamma z_\gamma$, redundancy factor of 0.03); and, finally, there are $\sum_{\gamma \in \mathcal{F}_n^8} \min\{2, d_\gamma\} z_\gamma = 1,821$ atoms generated from the first $k = 8$ shapes and up to two eigenvalues from each shape (1.8% of the previous quantity, redundancy factor of 0.0005). These values are shown in Fig. 31 for a wider range of n. Importantly, if we only plan to use atoms generated from the first k shapes in the analysis, we only need to compute the corresponding Schreier adjacency matrices, lifting matrices, paths, and eigendecompositions described in Secs. 6.1 and 6.2 for these shapes. We demonstrate the resulting computational savings in Sec. 6.5. Applications that could benefit from such a reduced transform with the top k shapes include lossy compression and machine learning problems, for which the reduced transform coefficients can serve as low-dimensional feature vectors for the high-dimensional ranked data.

n	$\sum_{\gamma \in \mathcal{F}_n} z_\gamma$	$\sum_{\gamma \in \mathcal{F}_n} d_\gamma z_\gamma$	$\sum_{\gamma \in \mathcal{F}_n} d_\gamma z_\gamma / n!$	$\sum_{\gamma \in \mathcal{F}_n^8} d_\gamma z_\gamma$	$\sum_{\gamma \in \mathcal{F}_n^8} \min\{2, d_\gamma\} z_\gamma$
3	4	7	1.167	7	7
4	8	19	0.792	19	15
5	26	131	1.092	131	51
6	107	1326	1.842	1326	213
7	295	6987	1.386	6987	589
8	1570	96895	2.403	15975	759
9	5507	843313	2.324	42541	1255
10	34427	18004348	4.962	98866	1821
11	139877	181831409	4.555	211751	2553
12	823242	3657722234	7.636	424832	3479
15	$1.6839 \cdot 10^8$	$2.9336 \cdot 10^{13}$	22.434	2562211	7731
20	$2.2896 \cdot 10^{12}$	$2.9314 \cdot 10^{20}$	120.490	26769956	21891
25	$8.0413 \cdot 10^{16}$	$1.7136 \cdot 10^{28}$	1104.810	168121101	49551
30	$5.0450 \cdot 10^{21}$	$2.7008 \cdot 10^{36}$	10182.000	758346521	97211

Fig. 31 Number of dictionary atoms for different values of n, the number of candidates. The third column ($\sum_{\gamma \in \mathcal{F}_n} d_\gamma z_\gamma$) gives the number of atoms in the proposed dictionary $\bar{\mathcal{D}}$ associated with the shapes in $\mathcal{F}_n$, and the fourth column is the resulting redundancy of the transform (number of atoms divided by the length of the data vector $\mathbf{f}$). The fifth and sixth columns contain the number of atoms when the dictionary is further restricted to those atoms associated with the top $k = 8$ shapes in $\mathcal{F}_n^8$ and either all eigenvalues in those shapes or at most two eigenvalues in each shape, respectively.

6.5 Computational Summary

To summarize, when we perform the transform using all atoms associated with the shapes in $\mathcal{F}_n$, phase 1 of the setup portion is negligible when compared to phases 2 and 3 of the setup and the computation of the analysis coefficients. Fig. 32 shows the times required for our MATLAB implementations to perform the proposed transform over all shapes in $\mathcal{F}_n$, on a 2.3 GHz MacBook Pro laptop with 32GB of RAM. The code is not yet optimized in the sense that we have not precompiled any C/C++ subroutines into MEX functions. As an example, for $n = 9$ candidates (362,880 vertices or possible rankings), the main implementation of the code performs the offline computations in the three phases of the setup portion in approximately 43 seconds, and then computes the analysis coefficients for each data vector $\mathbf{f}$ in approximately 26 seconds.

When we only compute the analysis coefficients associated with the first k symmetry types ($\mathcal{F}_n^k$), the two main computational bottlenecks are computing the permutations that rearrange $\mathbf{f}$ into $\rho_L(\sigma^{-1})\mathbf{f}$ and computing the analysis coefficients, each of which has complexity of $\mathcal{O}(z_\gamma n!)$ for shape γ. The overall complexity is therefore equal to $\mathcal{O}(n^{\bar{m}_k} n!)$, where $\bar{m}_k$ is defined to be the smallest value of m that results in $\sum_{i=1}^m p(i) \geq k$, where $p(i)$ is the partition number of integer i. For $k = 5, 8, 15, 30$, $\bar{m}_k$ is equal to 3,4,5,6, respectively. So, for the example above with the top $k = 8$ shapes, which includes all shapes in lexicographic order up to $[n - 4, 4]$ (i.e., where most of the useful information resides), the complexity is $\mathcal{O}(n^4 n!) = \mathcal{O}(n! \log^4(n!))$. With $n = 10$ and $k = 8$,
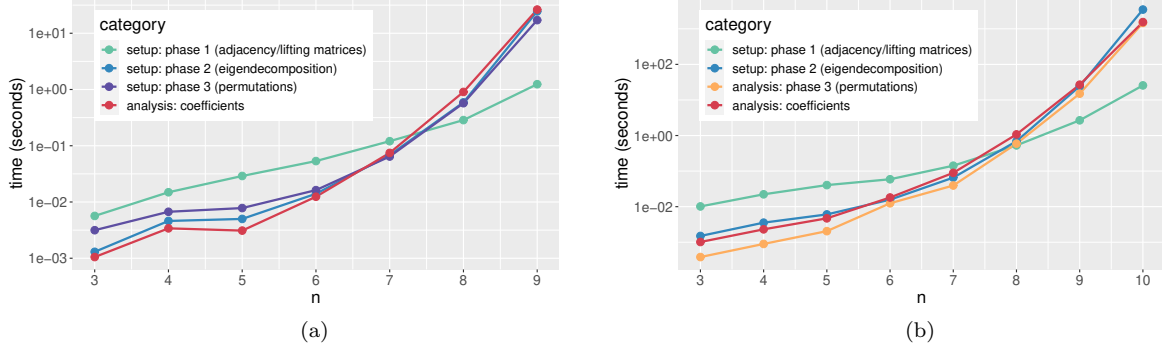
Fig. 32 Computation times for the setup and analysis portions in our (a) base implementation, and (b) more memory efficient implementation that moves some of the phase 3 setup computation into the analysis portion of the code.

which includes the computation of all of the largest magnitude coefficients in Fig. 19 for the sushi preference data, both the setup and analysis portions run in under one minute.

Finally, to synthesize a signal from the analysis coefficients (i.e., perform the inverse transform), we need to lift each Schreier Laplacian eigenvector back to the permutahedron one time, reorder these vectors in different ways, and then take a linear combination of the reordered vectors:

$$\mathbf{f}_{\mathrm{rec}} = \sum_{\gamma \vdash n} \sum_{\lambda \in \Lambda_\gamma} \sum_{k=1}^{\kappa_{\gamma,\lambda}} \sum_{\bar{\pi} \in \bar{\Pi}_\gamma} \alpha_{\gamma,\lambda,k,\bar{\pi}} \boldsymbol{\varphi}_{\gamma,\lambda,k,\bar{\pi}} = \sum_{\gamma \vdash n} \bar{c}_\gamma \sum_{\lambda \in \Lambda_\gamma} \sum_{k=1}^{\kappa_{\gamma,\lambda}} \sum_{\bar{\pi} \in \bar{\Pi}_\gamma} \alpha_{\gamma,\lambda,k,\bar{\pi}} \rho_L(\sigma) \mathbf{B}_{\pi_1} \mathbf{v}_{\gamma,\lambda,k},$$

where $\alpha_{\gamma,\lambda,k,\bar{\pi}} = \langle \mathbf{f}, \boldsymbol{\varphi}_{\gamma,\lambda,k,\bar{\pi}} \rangle$.

7 Related Work, Revisited

7.1 Bases, Frames, and Interpretability

As discussed in Sec. 3.1, Diaconis [29, p. 955] remarks that there is not a natural choice of basis of the irreducible components $V_{\gamma,i}$ with interpretable basis elements, and circumvents this issue via Mallows' method of projecting an overcomplete spanning set of interpretable functions (two of which are shown in Fig. 7) onto the isotypic component W_γ. Interestingly, these projections also form a tight frame for W_γ.

Proposition 9 *For each $\pi, \xi \in \Pi_\gamma$, define $\boldsymbol{\delta}_{\pi,\xi} \in \mathbb{R}[\mathbb{S}_n]$ by*

$$\boldsymbol{\delta}_{\pi,\xi}(\sigma) := \begin{cases} 1, & \textit{if } \sigma(\xi) = \pi \\ 0, & \textit{otherwise} \end{cases}, \sigma \in \mathbb{S}_n;$$

that is, $\boldsymbol{\delta}_{\pi,\xi}(\sigma) = 1$ if and only if σ places the candidates in each block of π in the ranking slots of the corresponding block of ξ. Let $(\boldsymbol{\delta}_{\pi,\xi})_\gamma$ be the orthogonal projection of $\boldsymbol{\delta}_{\pi,\xi}$ onto W_γ. Then the collection of vectors $\{(\boldsymbol{\delta}_{\pi,\xi})_\gamma\}_{\pi,\xi \in \Pi_\gamma}$ is a tight frame for W_γ.

Proof For $\tau \in \mathbb{S}_n$ we have $\tau \boldsymbol{\delta}_{\pi,\xi} = \boldsymbol{\delta}_{\tau(\pi),\xi}$ and $\boldsymbol{\delta}_{\pi,\xi} \tau = \boldsymbol{\delta}_{\pi,\tau^{-1}(\xi)}$, since $\sigma(\xi) = \pi$ if and only if $\tau\sigma(\xi) = \tau(\pi)$ if and only if $\sigma\tau(\tau^{-1}\xi) = \pi$. Furthermore, projection onto the isotypic component W_γ is in the center of the group algebra $\mathbb{R}[\mathbb{S}_n]$ (see (18) and [97, (13,19)]), so it commutes with both the left and right action of the group, giving $\tau(\boldsymbol{\delta}_{\pi,\xi})_\gamma = (\tau \boldsymbol{\delta}_{\pi,\xi})_\gamma = (\boldsymbol{\delta}_{\tau(\pi),\xi})_\gamma$ and $(\boldsymbol{\delta}_{\pi,\xi})_\gamma \tau = (\boldsymbol{\delta}_{\pi,\xi} \tau)_\gamma = (\boldsymbol{\delta}_{\pi,\tau^{-1}(\xi)})_\gamma$. For any fixed $\pi, \xi \in \Pi_\gamma$, the irreducible left $\mathbb{S}_n$-module generated by $(\boldsymbol{\delta}_{\pi,\xi})_\gamma$ is given by

$$\mathbb{R}[\mathbb{S}_n](\boldsymbol{\delta}_{\pi,\xi})_\gamma = \mathbb{R}\text{-span}\{\tau(\boldsymbol{\delta}_{\pi,\xi})_\gamma \mid \tau \in \mathbb{S}_n\} = \mathbb{R}\text{-span}\{(\boldsymbol{\delta}_{\tau(\pi),\xi})_\gamma \mid \tau \in \mathbb{S}_n\} = \{(\boldsymbol{\delta}_{\pi,\xi})_\gamma \mid \pi \in \Pi_\gamma\},$$

where the last equality follows from the fact that $\mathbb{S}_n$ acts transitively on Π_γ. Since this module is irreducible, it follows from [97, Thm 10.5], that $\{(\boldsymbol{\delta}_{\tau(\pi),\xi})_\gamma \mid \tau \in \mathbb{S}_n\}$ is a tight group frame for its span, and the span is isomorphic to V_γ. The symmetric argument with $\mathbb{S}_n$ acting on the right produces a copy of the right irreducible $\mathbb{S}_n$-module V_γ^* and proves that $\{(\boldsymbol{\delta}_{\pi,\xi})_\gamma \mid \xi \in \Pi_\gamma\}$ is a tight frame for its span. The isotypic component is isomorphic to a tensor product $W_\gamma \cong V_\gamma \otimes V_\gamma^*$ of irreducible left and right $\mathbb{S}_n$ modules, respectively. Therefore, by [97, Cor. 5.1], $\{(\boldsymbol{\delta}_{\pi,\xi})_\gamma \mid \pi, \xi \in \Pi_\gamma\}$ is a tight frame for W_γ.   □

On the other hand, we can distinguish a subset of our frame vectors that yield a basis for each $V_{\gamma,i}$. An ordered set partition $\pi \in \Pi_\gamma$ is *standard* if when the rows are put in increasing order from left to right, the columns also are in increasing order from top to bottom. For example, the five ordered set partitions in the bottom right corner of the Schreier graph in Fig. 12 are standard. Standard ordered set partitions are in bijection with standard tableaux and $d_\gamma = \dim(V_\gamma)$ equals the number of standard ordered set partitions [83]. Let Π_γ^{std} denote the set of standard ordered set partitions. The atoms coming from liftings by standard set partitions provide the desired basis.

Proposition 10 *For $\gamma \vdash n$, $\lambda \in \Lambda_\gamma$, and $1 \le k \le \kappa_{\gamma,\lambda}$, the set $\{\boldsymbol{\varphi}_{\gamma,\lambda,k,\pi} \mid \pi \in \Pi_\gamma^{std}\}$ is a basis for the irreducible $\mathbb{S}_n$ module $\mathbb{R}[\mathbb{S}_n]\boldsymbol{\varphi}_{\gamma,\lambda,k,\pi_0} \cong V_\gamma$ (for any $\pi_0 \in \Pi_\gamma$).*

Proof The module $V_\gamma^* \subseteq \mathbb{R}[\Pi_\gamma]$ is spanned by the polytabloids $\{q_\pi \mid \pi \in \Pi_\gamma\}$ with symmetric group action $q_\pi \sigma = q_{\sigma^{-1}(\pi)}$ (see (21)) and has a basis consisting of standard polytabloids $\{q_\pi \mid \pi \in \Pi_\gamma^{\text{std}}\}$ by [83, Thm. 2.5.2]. We show that this property transfers to $\mathbb{R}[\mathbb{S}_n]\boldsymbol{\varphi}_{\gamma,\lambda,k,\pi_0} = \mathbb{R}[\mathbb{S}_n](\mathbf{B}_{\pi_0}\mathbf{v}_{\gamma,\lambda,k})$. For any $0 \ne \mathbf{v} \in V_\gamma^*$, the map $\mathbb{R}[\mathbb{S}_n](\mathbf{B}_{\pi_0}\mathbf{v}) \to V_\gamma^*$ given by sending $\sigma\mathbf{B}_{\pi_0}\mathbf{v}$ to $\mathbf{v}\sigma^{-1}$ is an isomorphism (the representing matrices are transposes of one another to account for the change from left to right modules). Moreover, for $0 \ne \mathbf{v}, \mathbf{w} \in V_\gamma^* \subseteq \mathbb{R}[\Pi_\gamma]$, the action of the symmetric group on $(\mathbb{R}[\mathbb{S}_n]\mathbf{B}_{\pi_0})\mathbf{v}$ is identical to the action on $(\mathbb{R}[\mathbb{S}_n]\mathbf{B}_{\pi_0})\mathbf{w}$. To see this, observe that $\sigma\mathbf{B}_{\pi_0}\mathbf{v} = \mathbf{B}_{\sigma(\pi_0)}\mathbf{v}$ and $\sigma\mathbf{B}_{\pi_0}\mathbf{w} = \mathbf{B}_{\sigma(\pi_0)}\mathbf{w}$. Moreover, if $\sum_{\sigma \in \mathbb{S}_n} c_\sigma \sigma\mathbf{B}_{\pi_0}\mathbf{v} = \sum_{\sigma \in \mathbb{S}_n} c_\sigma \mathbf{B}_{\sigma(\pi_0)}\mathbf{v} = 0$ is a dependence relation, then there is $x \in \mathbb{R}[\mathbb{S}_n]$ so that $\mathbf{w} = \mathbf{v}x$ so $\sum_{\sigma \in \mathbb{S}_n} c_\sigma \mathbf{B}_{\sigma(\pi_0)}\mathbf{w} = \sum_{\sigma \in \mathbb{S}_n} c_\sigma \sigma\mathbf{B}_{\pi_0}(\mathbf{v}x) = (\sum_{\sigma \in \mathbb{S}_n} c_\sigma \sigma\mathbf{B}_\pi\mathbf{v})x = 0$, since the left and right group actions commute. The converse argument holds, so dependence relations on $\{\mathbf{B}_\pi\mathbf{v}\}_{\pi \in \Pi_\gamma}$ and $\{\mathbf{B}_\pi\mathbf{w}\}_{\pi \in \Pi_\gamma}$ are the same. If $\mathbf{v} = q_{\pi_0}$, then $\mathbf{B}_{\sigma(\pi_0)}q_{\pi_0} = \sigma\mathbf{B}_{\pi_0}q_{\pi_0}$ which maps to $q_{\pi_0}\sigma^{-1} = q_{\sigma(\pi_0)}$ under the isomorphism, so for each $\pi \in \Pi_\gamma$, $\mathbf{B}_\pi q_{\pi_0}$ maps to q_π. It follows that $\{\mathbf{B}_\pi q_{\pi_0} \mid \pi \in \Pi_\gamma^{\text{std}}\}$ is a basis for $\mathbb{R}[\mathbb{S}_n](\mathbf{B}_{\pi_0}q_{\pi_0})$. The result holds for $\mathbb{R}[\mathbb{S}_n](\mathbf{B}_{\pi_0}\mathbf{v}_{\gamma,\lambda,k})$ by the fact that the dependence relations are the same for $\{\mathbf{B}_\pi q_{\pi_0}\}_{\pi \in \Pi_\gamma}$ and $\{\mathbf{B}_\pi\mathbf{v}_{\gamma,\lambda,k}\}_{\pi \in \Pi_\gamma}$.   □

To recap, there are two important differences between the data analysis methods we propose here and those presented by Diaconis in [29]. First, we directly construct each spanning vector as an interpretable function in $V_{\gamma,i}$, whereas the Mallows vectors are constructed as interpretable functions but then projected, during which some of the interpretability may be lost. Second, and probably more importantly, the atoms of the form $\boldsymbol{\varphi}_{\gamma,\lambda,k,\bar{\pi}}$ that we use reside in a single Laplacian eigenspace of the permutahedron, whereas the atoms of the form $(\boldsymbol{\delta}_{\pi,\xi})_\gamma$ from Prop. 9 (and shown in the bottom of Fig. 7) reside in multiple Laplacian eigenspaces of the permutahedron, making the corresponding analysis coefficients less interpretable from a smoothness perspective. Our proposed frame is therefore a more refined representation that enables us to capture both smoothness and symmetry information about the ranking data.

7.2 Smoothness, Notions of Frequency, and Symmetry Types

We mentioned in Sec. 3.2 that the symmetry types carry a notion of frequency: the Laplacian eigenvectors of the Cayley graph induced by the generating set of all transpositions, Γ_n, that reside in isotypic components that occur earlier in the dominance ordering are smoother functions with respect to the graph structure. This notion of frequency and relationship between dominance ordering and smoothness does not carry over directly to the setting of the permutahedron, as all but the first and last isotypic components contain Laplacian eigenvectors of $\mathbb{P}_n$ associated with different eigenvalues. However, we believe there is still an interesting relationship between

the dominance ordering of the isotypic components and the smoothness with respect to the permutahedron of the smoothest eigenvectors within each isotypic component.

Conjecture 1 For each $\gamma \vdash n$, define $\tilde{\lambda}_\gamma := \min_{\lambda \in \Lambda_\gamma}\{\lambda\}$. If $\nu \rhd \gamma$, then $\tilde{\lambda}_\nu < \tilde{\lambda}_\gamma$.

We have numerically verified this conjecture for $n \leq 10$. Fig. 33 shows an example of Conj. 1 with $n = 10$.

7.3 Computational Complexity and Efficient Algorithms

A naive implementation of the Fourier transform on $\mathbb{S}_n$ requires $\mathcal{O}((n!)^2)$ time. The FFT of Clausen and Baum [19] reduces the time to $\mathcal{O}(n^3 n!)$ and Maslen [68] improves this to $\mathcal{O}(n^2 n!)$. It is conceivable that we could leverage some of the ideas from these works to further improve the complexity of applying our proposed transform. However, these methods also run into space constraints as they save time by storing many vectors of length $n!$ in their implementation. For example, most transforms on $\mathbb{S}_n$ are built recursively from $\mathbb{S}_{n-1}$ and the sequences of adjacent swaps $(i, i+1)(i+1, i+2)\cdots(n-1, n)$ for each $1 \leq i \leq n-1$ (coset representatives of $\mathbb{S}_{n-1}$ in $\mathbb{S}_n$). For these reasons in part, [53, 76] implement algorithms to compute the FFT on individual isotypic components.

Another possible computational improvement we have not yet explored is to approximately compute the transform coefficients, avoiding the eigendecompositions of the Schreier Laplacians. This is the approach taken in most scalable graph signal processing algorithms [43, 89].
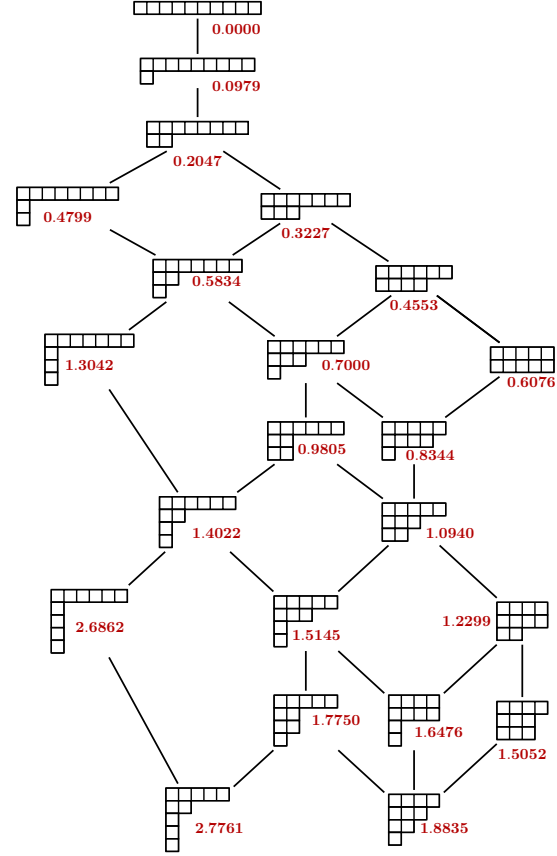

Fig. 33 Dominance order poset with each of the 22 shapes in $\mathcal{F}_{10}$ labeled by the smallest eigenvalue $\tilde{\lambda}_\gamma$ associated with an eigenvector in the isotypic component W_γ. Conj. 1 also holds for the 20 conjugate shapes, which are not shown.

7.4 Parametric Distance-Based Models

Of the probability models for ranked data, the most closely related to our framework are distance-based and multistage models that use the Kendall distance metric ([67, Ch. 6], [2, Ch. 8.3], and [100, Sec. 4.3] all contain overviews of distance-based models). For example, Mallows' ϕ-model [65], [28, Ch. 6] assumes there is a single modal ranking σ_0 and the probability of a voter voting for another ranking σ is proportional to $e^{-\tau d_K(\sigma, \sigma_0)}$, where τ is a single dispersion parameter and the Kendall distance $d_K(\sigma, \sigma_0)$ is equal to the number of hops separating σ and σ_0 on the permutahedron. Generalizations of this model include mixture models with multiple modes and dispersion rates to represent heterogeneous voting populations [74], generalized Mallows models that allow for different dispersion parameters at each position of the permutation [37], and mixtures of generalized Mallows models [72]. A review and comparison of these models is in [13] and software implementations are detailed in [46, 77].

From a graph signal processing perspective, fitting these models is closely related to blind deconvolution on graphs [85, 47, 79]; e.g.,

$$\min_{\mathbf{x}, h}\{\|\mathbf{f} - h(\boldsymbol{\mathcal{L}})\mathbf{x}\|_2\} \quad \text{s.t. } \|\mathbf{x}\|_0 \leq S, \tag{26}$$

where $\mathbf{x}$ is a sparse input signal whose support corresponds to the S modes used in the mixture of Mallows models, h is a diffusion filter which may be a polynomial of small degree or have a parametrized form such as $h(\lambda_\ell) = \alpha e^{-\tau \lambda_\ell}$, and $\mathcal{L}$ is the graph Laplacian of either the (unweighted) permutahedron or a weighted variant that places different edge weights on edges corresponding to swaps of candidates in different adjacent positions in the rankings.

While our proposed approach models the length $n!$ data vector as the linear combination of a much higher number of atoms, $\mathbf{f} = \sum_{\phi \in \bar{\mathcal{D}}} \langle \mathbf{f}, \phi \rangle \phi$, the approach can also be used to generate compressed representations with atoms of the same form. This is because ranked data found in applications is typically smooth with respect to the permutahedron, and therefore the magnitudes of most of the coefficients $\langle \mathbf{f}, \phi \rangle$ are small (see, e.g., Fig. 17). Thus, sparse linear combinations of these atoms can yield effective approximations to high-dimensional ranked data; i.e., $\mathbf{f} \approx \sum_{\phi \in \bar{\mathcal{D}}'} \alpha_\phi \phi$, where $\bar{\mathcal{D}}'$ is a subset of the tight frame dictionary $\bar{\mathcal{D}}$ with cardinality $|\bar{\mathcal{D}}'| \ll n!$.

8 Extensions and Future Directions

We conclude with a brief mention of three lines of future work.

8.1 Closed Form Computation of the Schreier Eigenvalues and Eigenvectors

Availability of closed-form formulas for the Laplacian eigenvalues and eigenvectors of the Schreier graphs could eliminate the need to perform the computations mentioned above altogether. Closed forms are known for Cayley graphs generated by all transpositions [31] and for Cayley graphs generated by transposing i with n for each i (the star graph) [36], but no general, closed formula is known in the case of adjacent transpositions (the permutahedron). Partial results are known [5,33,38] when γ is a "hook shape," i.e., $\gamma = [n-k, 1, \ldots, 1]$ for $0 \le k \le n-1$.

Lemma 3 ([5,38]) *Define the k-fold Cartesian product graph*

$$Q_{n,k} := Q_n^{\times k} = \underbrace{Q_n \times Q_n \times \cdots \times Q_n}_{k \, copies}, \tag{27}$$

which is the k-dimensional cube graph of side length n. For each subset $I = \{i_1, \ldots, i_k\} \subseteq \{1, \ldots, n-1\}$ of size k, $\lambda_I := \sum_{j=1}^{k} \lambda_{i_j}^{Q_n}$ is a Laplacian eigenvalue of $Q_{n,k}$, and the k-fold wedge product

$$\mathbf{w}_{\gamma, \lambda_I} := \mathbf{v}_{i_1}^{Q_n} \wedge \mathbf{v}_{i_2}^{Q_n} \wedge \cdots \wedge \mathbf{v}_{i_k}^{Q_n} = \frac{1}{\sqrt{k!}} \sum_{\sigma \in \mathbb{S}_k} \mathsf{sign}(\sigma) \mathbf{v}_{i_{\sigma(1)}}^{Q_n} \otimes \mathbf{v}_{i_{\sigma(2)}}^{Q_n} \otimes \cdots \otimes \mathbf{v}_{i_{\sigma(k)}}^{Q_n} \tag{28}$$

is a Laplacian eigenvector of $Q_{n,k}$ associated with λ_I.

As shown in Fig. 34, the vertices of $Q_{n,k}$ can be labeled by the set of k-tuples from the alphabet $\{1, \ldots, n\}$:

$$V(Q_{n,k}) = \{(a_1, \ldots, a_k) \mid 1 \le a_j \le n\},$$

and the vertices of the Schreier graph $\mathbb{P}_{[n-k,1,\ldots,1]}$ can be labeled by the k-permutations of $\{1, \ldots, n\}$:

$$V(\mathbb{P}_{[n-k,1,\ldots,1]}) = \{(a_1, \ldots, a_k) \mid 1 \le a_j \le n, a_i \ne a_j \text{ if } i \ne j\}.$$

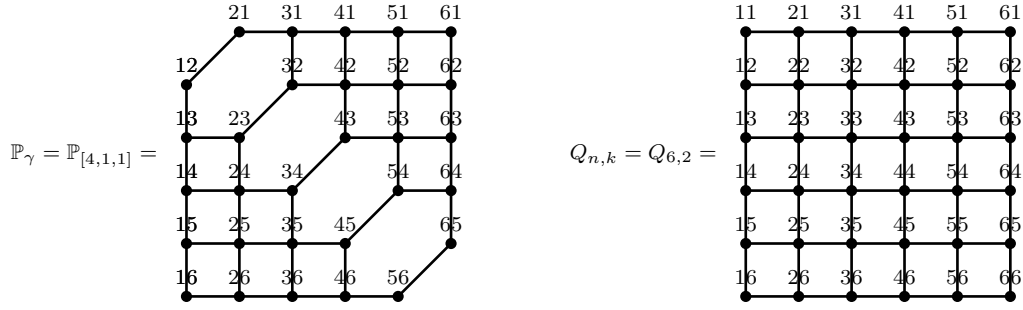Thus $V(\mathbb{P}_{[n-k,1,\ldots,1]}) \subseteq V(Q_{n,k})$.

Fig. 34 Embedding of the Schreier graph $\mathbb{P}_\gamma$ in the k-dimensional cube graph $Q_{n,k}$.

Proposition 11 ([5,38]) *Let γ be a "hook shape"; i.e., $\gamma = [n-k, 1^k] = [n-k, 1, \ldots, 1]$ for $0 \leq k \leq n-1$. For each subset $I = \{i_1, \ldots, i_k\} \subseteq \{1, \ldots, n-1\}$ of size k, $\lambda_I := \sum_{j=1}^{k} \lambda_{i_j}^{Q_n}$ is a Laplacian eigenvalue of $\mathbb{P}_\gamma$ with the associated eigenvector given by $\mathbf{v}_{\gamma,\lambda_I} := \mathbf{w}_{\gamma,\lambda_I}|_{V(\mathbb{P}_\gamma)}$, the restriction of $\mathbf{w}_{\gamma,\lambda_I}$, viewed as a function on the vertices of $Q_{n,k}$, to the vertices of $\mathbb{P}_\gamma$. Thus, for any hook shape $\gamma = [n-k, 1^k]$, the complexity of computing each eigenvector of $\mathbb{P}_\gamma$ is $\mathcal{O}(k!n^k)$, and the complexity of computing all eigenvectors of $\mathbb{P}_\gamma$ is $\mathcal{O}\left(\frac{n!k!n^k}{(n-k)!}\right)$, which is bounded by $\mathcal{O}(k!n^{2k})$.*

Remark 7 Earlier, Edelman and White [33] studied the integer eigenvalues of the permutahedron, found the integer eigenvalues from the hook shapes, and also conjectured that the only integer eigenvalues that appear in non-hook shapes also appear in the hook shapes for the same value of n.

Remark 8 The permutahedron $\mathbb{P}_n$ is the same as the Full-Flag Johnson graph $FJ(n,1)$ studied in [25], in which Dai uses the recursive structure of Full-Flag Johnson graphs to compute a subset of the adjacency spectrum of $\mathbb{P}_n$, namely the eigenvalues of the matrix M_n in [25, Lem. 10]. That matrix M_n is actually equal to the adjacency matrix of the Schreier graph $\mathbb{P}_{[n,n-1]}$, which is a path graph with additional self loops that make it regular. Since, as mentioned in Sec. 5.2.2, the Laplacian eigenvalues of $\mathbb{P}_{[n,n-1]}$ are known in closed form, an alternative closed form of the eigenvalues of Dai's M_n matrix is $\{n - 3 + 2\cos(\frac{\pi\ell}{n})\}_{\ell=0,1,\ldots,n-1}$, and we do indeed know which subset of the spectrum of $\mathbb{P}_n$ is attained from this matrix.

The Schreier graph $\mathbb{P}_{[n-2,2]}$ (shown, e.g., in Fig. 12) closely resembles the quartered Aztec diamond, the adjacency spectrum of which is studied and specified by Ciucu in [18, Eq. (2.4)]. However, the quartered Aztec diamond graph studied by Ciucu does not include the self loops that are present in $\mathbb{P}_{[n-2,2]}$, and is therefore not regular and the Laplacian eigenvalues do not follow immediately. We have not yet found a way to successfully adapt the approach of [18] to write down a closed-form formula for the Laplacian eigenvalues of $\mathbb{P}_{[n-2,2]}$.

8.2 Partial and Incomplete Rankings

Our approach to construct tight frames for signals on the permutahedron can be extended to (i) *partial rankings*, in which voters are allowed to include ties for one or more candidates [24, 29, 94, 6, 95, 52, 55], and (ii) *incomplete rankings*, in which voters may only rank a subset of the candidates [52, 64, 20, 90, 91]. A common example of partial rankings is when voters list their top k candidates, and the remaining $n-k$ are implicitly considered to be tied. An example of incomplete rankings is balanced incomplete block design, where each voter is assigned k out of the n objects to rank, in such a way that each object is judged by the same number of voters and each pair of objects is presented to the same number of voters [67, Ch. 11], [32, 21, 7]. This is used, for example, if there are too many objects for an individual to effectively compare due to time constraints or cognitive limitations. Our frame vectors may yield a naturally interpretable spanning set for the possible vote tallies in this experimental design. Both of these extensions merit further investigation given the myriad applications in which partial and incomplete rankings appear.

8.3 Tight Frames for Analyzing Data on Other Cayley Graphs, Groups, and Combinatorial Structures

The frame construction presented here works when the permutahedron is replaced by the Cayley graph $\Gamma(\mathbb{S}_n, S)$ corresponding to any generating set $S \subseteq \mathbb{S}_n$. For example, if S is the set of all transpositions, then the Cayley graph is the graph Γ_n discussed in Sec. 3.2, and if $S = \{(i, n) \mid 1 \leq i < n\}$, then the Cayley graph is the star graph studied in [36]. In any of these cases, our methods yield a tight spectral frame with respect to the graph, and signals can be projected down to and interpreted on the corresponding Schreier graphs.

Furthermore, these methods extend to any finite group. The equitable partition $\sim_\pi$ (see Def. 2) on $\mathbb{S}_n$ induced by the ordered set partition $\pi \in \Pi_\gamma$ has equivalence classes equal to the right cosets of the stabilizer subgroup $\mathbb{S}_\pi = \{\sigma \in \mathbb{S}_n \mid \sigma(\pi) = \pi\}$, and the Schreier graph $\mathbb{P}_\gamma$ is isomorphic to the Cayley graph determined by $\mathbb{S}_n$ acting on these cosets. When $\mathbb{S}_n$ is replaced by any finite group $\mathbb{G}$ and $\mathbb{S}_\mu$ is replaced by a subgroup $\mathbb{H} \leq \mathbb{G}$, one obtains tight spectral frames (with the same geometric and energy-preserving properties) for studying data on $\mathbb{G}$ with respect to any Cayley graph. These methods naturally extend to groups that generalize the symmetric group such as Weyl groups, complex reflection groups, and finite general linear groups. For example, when applied to the hyperoctahedral group (the Weyl group of type B), these methods provide a setting to analyze data on signed permutations. One can further extend these methods to analyze data on matchings, subsets, and set partitions by using the representation theory of the corresponding semisimple algebras with "group-like" structure, the Brauer, rook monoid, and partition algebras. Fourier analysis methods on these algebras are initiated in the recent work [69].

References

1. Akritidis, L., Katsaros, D., Bozanis, P.: Effective rank aggregation for metasearching. J. Syst. Softw **84**(1), 130–143 (2011)
2. Alvo, M., Yu, P.L.H.: Statistical Methods for Ranking Data. Springer (2014)
3. Armon, S., Halverson, T.: Transition matrices between Young's natural and seminormal representations. Electronic J. Combin. **28**(3) (2021)
4. Arrow, K.J.: A difficulty in the concept of social welfare. J. Political Econ **58**(4), 328–346 (1950)
5. Bacher, R.: Valeur propre minimale du laplacien de Coxeter pour le groupe symétrique. J. Algebra **167**(2), 460–472 (1994)
6. Baggerly, K.A.: Visual estimation of structure in ranked data. Ph.D. thesis, Rice University (1995)
7. Bailey, R., Diaconis, P., Rockmore, D.N., Rowley, C.: A spectral analysis approach for experimental designs. In: Excursions in Harmonic Analysis, Vol. 4, pp. 367–395. Springer (2015)
8. Basha, T., Moses, Y., Avidan, S.: Photo sequencing. In: Proc. ECCV, pp. 654–667. Springer (2012)
9. Bennett, P.N., Chickering, D.M., Mityagin, A.: Learning consensus opinion: Mining data from a labeling game. In: Proc. ACM WWW, pp. 121–130 (2009)
10. Bondy, J.A., Murty, U.S.: Graph Theory. Springer (2000)
11. Borda, J.d.: Mémoire sur les élections au scrutin. Histoire de l'Academie Royale des Sciences pour 1781 (Paris, 1784) (1784)
12. Breitling, R., Armengaud, P., Amtmann, A., Herzyk, P.: Rank products: A simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. FEBS letters **573**(1-3), 83–92 (2004)
13. Ceberio, J., Irurozki, E., Mendiburu, A., Lozano, J.A.: A review of distances for the Mallows and Generalized Mallows estimation of distribution algorithms. Comput. Optim. Appl. **62**(2), 545–564 (2015)
14. Chamberlin, J.R., Cohen, J.L., Coombs, C.H.: Social choice observed: Five presidential elections of the American Psychological Association. J. Politics **46**(2), 479–502 (1984)
15. Chen, W.: How to order sushi. PhD diss., Harvard University (2014)
16. Chen, X., Bennett, P.N., Collins-Thompson, K., Horvitz, E.: Pairwise ranking aggregation in a crowdsourced setting. In: Proc. ACM WSDM, pp. 193–202 (2013)
17. Christensen, O.: Frames and Bases. Birkhauser (2008)
18. Ciucu, M.: Symmetry classes of spanning trees of Aztec diamonds and perfect matchings of odd squares with a unit hole. J. Algebr. Comb. **27**(4), 493–538 (2008)
19. Clausen, M., Baum, U.: Fast Fourier transforms for symmetric groups: Theory and implementation. Math. Comput. **61**(204), 833–847 (1993)
20. Clémençon, S., Jakubowicz, J., Sibony, E.: Multiresolution analysis of incomplete rankings. arXiv preprint arXiv:1403.1994 (2014)
21. Cochran, W.G., Cox, G.: Experimental Designs. Wiley (1957)
22. Cohen, A., Mallows, C.: Analysis of ranking data. Bell Laboratories Memorandum (1980)
23. Condorcet, N.d.: Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix. L'Imprimerie Royale (1785)
24. Critchlow, D.E.: Metric methods for analyzing partially ranked data, vol. 34. Springer-Verlag (1985)
25. Dai, I.: Diameter bounds and recursive properties of Full-Flag Johnson graphs. Discrete Math. **341**(7), 1932–1944 (2018)

26. DeConde, R.P., Hawley, S., Falcon, S., Clegg, N., Knudsen, B., Etzioni, R.: Combining results of microarray experiments: A rank aggregation approach. Stat. Appl. Genet. Mol. Biol. **5**(1) (2006)
27. Devlin, S., Uminsky, D.: Identifying group contributions in NBA lineups with spectral analysis. J. Sports Anal. **6**(3), 215–234 (2020)
28. Diaconis, P.: Group Representations in Probability and Statistics, *Institute of Mathematical Statistics Lecture Notes— Monograph Series*, vol. 11. Institute of Mathematical Statistics, Hayward, CA (1988)
29. Diaconis, P.: A generalization of spectral analysis with application to ranked data. Ann. Statist. **17**(3), 949–979 (1989)
30. Diaconis, P., Rockmore, D.: Efficient computation of isotypic projections for the symmetric group. DIMACS Ser. Discrete Math. Theoret. Comput. Sci **11**, 87–104 (1993)
31. Diaconis, P., Shahshahani, M.: Generating a random permutation with random transpositions. Z. Wahrsch. Verw. Gebiete **57**(2), 159–179 (1981)
32. Durbin, J.: Incomplete blocks in ranking experiments. Br. J. Stat. Psychol. **4**(2), 85–90 (1951)
33. Edelman, P., White, D.: Codes, transforms and the spectrum of the symmetric group. Pacific J. Math. **143**(1), 47–67 (1990)
34. Etingof, P.I., Golberg, O., Hensel, S., Liu, T., Schwendner, A., Vaintrob, D., Yudovina, E.: Introduction to representation theory, vol. 59. American Mathematical Soc. (2011)
35. Where is ranked choice voting used? `https://www.fairvote.org/rcv#where_is_ranked_choice_voting_used`
36. Flatto, L., Odlyzko, A.M., Wales, D.B.: Random shuffles and group representations. Ann. Probab. **13**(1), 154–178 (1985)
37. Fligner, M.A., Verducci, J.S.: Distance based ranking models. J. R. Stat. Soc. Series B Stat. Methodol. **48**(3), 359–369 (1986)
38. Friedman, J.: On Cayley graphs on the symmetric group generated by transpositions. Combinatorica **20**(4), 505–519 (2000)
39. Gabriel, K.R.: Biplots. Wiley StatsRef: Statistics Reference Online (2014)
40. Ghandehari, M., Guillot, D., Hollingsworth, K.: A non-commutative viewpoint on graph signal processing. In: Proc. Int. Conf. Samp. Theory and Appl. (2019)
41. Godsil, C., Royle, G.F.: Algebraic Graph Theory, vol. 207. Springer Science & Business Media (2013)
42. Grossman, J., Minton, G.: Inversions in ranking data. Discrete Math. **309**(20), 6149–6151 (2009)
43. Hammond, D.K., Vandergheynst, P., Gribonval, R.: Wavelets on graphs via spectral graph theory. Appl. Comput. Harmon. Anal. **30**(2), 129–150 (2011)
44. Huang, J., Guestrin, C., Guibas, L.: Fourier theoretic probabilistic inference over permutations. J. Mach. Learn. Res. **10**, 997–1070 (2009)
45. Huang, J., Guestrin, C., Guibas, L.J.: Efficient inference for distributions on permutations. In: Adv. Neural Inf. Process. Syst., pp. 697–704 (2008)
46. Irurozki, E., Calvo, B., Lozano, J.A.: PerMallows: An R package for Mallows and generalized Mallows models. J. Stat. Softw. **71**(1), 1–30 (2016)
47. Iwata, K., Yamada, K., Tanaka, Y.: Graph blind deconvolution with sparseness constraint. arXiv preprint arXiv:2010.14002 (2020)
48. Jiang, X., Sun, J., Guibas, L.: A Fourier-theoretic approach for inferring symmetries. Comput. Geom. **47**(2), 164–174 (2014)
49. Kakarala, R.: A signal processing approach to Fourier analysis of ranking data: The importance of phase. IEEE Trans. Signal Process. **59**(4), 1518–1527 (2011)
50. Kamishima, T.: Nantonac collaborative filtering: Recommendation based on order responses. In: Proc. ACM Int. Conf. Knowl. Discovery Data Mining, pp. 583–588 (2003)
51. Kamishima, T., Akaho, S.: Efficient clustering for orders. In: D. Zighed, S. Tsumoto, Z. Ras, H. Hacid (eds.) Mining Complex Data, pp. 261–279. Springer (2009)
52. Kidwell, P., Lebanon, G., Cleveland, W.: Visualizing incomplete and partially ranked data. IEEE Trans. Vis. Comput. Graphics **14**(6), 1356–1363 (2008)
53. Kondor, R.: $S_n$ob: A C++ toolkit for fast Fourier transforms on the symmetric group (2006)
54. Kondor, R.: A Fourier space algorithm for solving quadratic assignment problems. In: Proc. ACM-SIAM SODA, pp. 1017–1028 (2010)
55. Kondor, R., Barbosa, M.S.: Ranking with kernels in Fourier space. In: Proc. COLT, pp. 451–463 (2010)
56. Kondor, R., Borgwardt, K.M.: The skew spectrum of graphs. In: Proc. Int. Conf. Mach. Learn., pp. 496–503 (2008)
57. Kondor, R., Dempsey, W.: Multiresolution analysis on the symmetric group. Proc. Neural Inf. Process. Syst. pp. 1637–1645 (2012)
58. Kondor, R., Howard, A., Jebara, T.: Multi-object tracking with representations of the symmetric group. In: Artif. Intell. and Statist., pp. 211–218 (2007)
59. Kovačević, J., Chebira, A.: Life beyond bases: The advent of frames (part I). IEEE Signal Process. Mag. **24**(4), 86–104 (2007)
60. Kovačević, J., Chebira, A.: Life beyond bases: The advent of frames (part II). IEEE Signal Process. Mag. **24**(5), 115–125 (2007)
61. Li, X., Wang, X., Xiao, G.: A comparative study of rank aggregation methods for partial and top ranked lists in genomic applications. Brief. Bioinform. **20**(1), 178–189 (2017)
62. Lin, S.: Rank aggregation methods. Wiley Interdiscip. Rev. Comput. Stat. **2**(5), 555–570 (2010)
63. Liu, Q., Crispino, M., Scheel, I., Vitelli, V., Frigessi, A.: Model-based learning from preference data. Annu. Rev. Stat. Appl. **6**, 329–354 (2019)
64. Malandro, M.E.: Inverse semigroup spectral analysis for partially ranked data. Appl. Comput. Harmon. Anal. **35**(1), 16–38 (2013)
65. Mallows, C.L.: Non-null ranking models. I. Biometrika **44**(1/2), 114–130 (1957)
66. Marden, J.I.: Use of nested orthogonal contrasts in analyzing rank data. J. Am. Stat. Assoc. **87**(418), 307–318 (1992)

67. Marden, J.I.: Analyzing and modeling rank data. Chapman and Hall/CRC (1995)
68. Maslen, D.: The efficient computation of Fourier transforms on the symmetric group. Math. Comp. **67**(223), 1121–1147 (1998)
69. Maslen, D., Rockmore, D., Wolff, S.: The efficient computation of Fourier transforms on semisimple algebras. J. Fourier Anal. Appl. **24**, 1377–1400 (2018)
70. McCullagh, P.: Permutations and regression models. In: Probability Models and Statistical Analyses for Ranking Data, pp. 196–215. Springer (1993)
71. McCullagh, P., Ye, J.: Matched pairs and ranked data. In: Probability Models and Statistical Analyses for Ranking Data, pp. 299–306. Springer (1993)
72. Meilă, M., Chen, H.: Dirichlet process mixtures of generalized Mallows models. In: Proc. Uncertainty in Artif. Intell., pp. 358–367 (2010)
73. 2017 Minneapolis election results. `http://vote.minneapolismn.gov/results/2017/`
74. Murphy, T.B., Martin, D.: Mixtures of distance-based models for ranking data. Comput. Stat. Data Anal. **41**(3-4), 645–655 (2003)
75. Ortega, A., Frossard, P., Kovačević, J., Moura, J.M., Vandergheynst, P.: Graph signal processing: Overview, challenges, and applications. Proc. IEEE **106**(5), 808–828 (2018)
76. Plumb, G., Pachauri, D., Kondor, R., Singh, V.: S_nFFT: A Julia toolkit for Fourier analysis of functions over permutations. J. Mach. Learn. Res. **16**(1), 3469–3473 (2015)
77. Qian, Z., Philip, L.: Weighted distance-based models for ranking data using the R package rankdist. J. Stat. Softw. **90**(1), 1–31 (2019)
78. Raman, K., Joachims, T.: Bayesian ordinal peer grading. In: Proc. ACM Conf. Learning@ Scale, pp. 149–156 (2015)
79. Ramírez, D., Marques, A.G., Segarra, S.: Graph-signal reconstruction and blind deconvolution for structured inputs. Signal Processing **188**, 108180 (2021)
80. Rockmore, D.: Some applications of generalized FFTs. In: Proc. DIMACS Wkshp. Groups and Comput., pp. 329–369 (1997)
81. Rockmore, D., Kostelec, P., Hordijk, W., Stadler, P.F.: Fast Fourier transform for fitness landscapes. Appl. Comput. Harmon. Anal. **12**(1), 57–76 (2002)
82. Rubinstein, R., Bruckstein, A.M., Elad, M.: Dictionaries for sparse representation modeling. Proc. IEEE **98**(6), 1045–1057 (2010)
83. Sagan, B.E.: The Symmetric Group: Representations, Combinatorial Algorithms, and Symmetric Functions, vol. 203. Springer Science & Business Media (2013)
84. Schreier, O.: Die untergruppen der freien gruppen. In: Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg, vol. 5, pp. 161–183. Springer (1927)
85. Segarra, S., Mateos, G., Marques, A.G., Ribeiro, A.: Blind identification of graph filters. IEEE Trans. Signal Process. **65**(5), 1146–1159 (2016)
86. Shani, G., Gunawardana, A.: Evaluating recommendation systems. In: Recommender Systems Handbook, pp. 257–297. Springer (2011)
87. Shuman, D.: Localized spectral graph filter frames: A unifying framework, survey of design considerations, and numerical comparison. IEEE Signal Process. Mag. **37**(6), 43–63 (2020)
88. Shuman, D. I, Narang, S.K., Frossard, P., Ortega, A., Vandergheynst, P.: The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains. IEEE Signal Process. Mag. **30**(3), 83–98 (2013)
89. Shuman, D. I, Vandergheynst, P., Kressner, D., Frossard, P.: Distributed signal processing via Chebyshev polynomial approximation. IEEE Trans. Signal Inf. Process. Netw. **4**(4), 736–751 (2018)
90. Sibony, E.: Multiresolution analysis of ranking data. Ph.D. thesis, Télécom ParisTech (2016)
91. Sibony, E., Clémençon, S., Jakubowicz, J.: A multiresolution analysis framework for the statistical analysis of incomplete rankings. arXiv preprint arXiv:1601.00399 (2016)
92. Stoyanovich, J., Jacob, M., Gong, X.: Analyzing crowd rankings. In: Proc. ACM WebDB, pp. 41–47 (2015)
93. Strang, G.: The discrete cosine transform. SIAM Review **41**(1), 135–147 (1999)
94. Thompson, G.: Generalized permutation polytopes and exploratory graphical methods for ranked data. Ann. Statist. **21**(3), 1401–1430 (1993)
95. Ukkonen, A.: Visualizing sets of partial rankings. In: Proc. Int. Symp. Intell. Data Anal., pp. 240–251 (2007)
96. Uminsky, D., Banuelos, M., González-Albino, L., Garza, R., Nwakanma, S.A.: Detecting higher order genomic variant interactions with spectral analysis. In: Proc. Eur. Signal Process. Conf., pp. 1–5 (2019)
97. Waldron, S.F.: An Introduction to Finite Tight Frames. Springer (2018)
98. Wang, J., Srebro, N., Evans, J.: Active collaborative permutation learning. In: Proc. ACM SIGKDD, pp. 502–511 (2014)
99. Wong, H.S., Chin, T.J., Yu, J., Suter, D.: Mode seeking over permutations for rapid geometric model fitting. Pattern Recogn. **46**(1), 257–271 (2013)
100. Yu, P.L., Gu, J., Xu, H.: Analysis of ranking data. Wiley Interdiscip. Rev. Comput. Stat. **11**(6), e1483 (2019)
101. Zbikowski, C.: How the 2017 Ward 3 election in Minneapolis foreshadows our local political future. Streets.mn (2019). `https://streets.mn/2019/05/20/how-the-2017-ward-3-election-in-minneapolis-foreshadows-our-local-political-future/`