DIMACS Series in Discrete Mathematics
and Theoretical Computer Science
Volume 00, 0000

Efficient Computation of Isotypic Projections for the Symmetric Group

PERSI DIACONIS AND DANIEL ROCKMORE

March 1, 1992

ABSTRACT. Spectral analysis on the symmetric group S_n calls for computing projections of functions defined on S_n and its homogeneous spaces, onto invariant subspaces. In particular, for the analysis of partially ranked data, the appropriate homogeneous spaces are given as quotients by Young subgroups. Here the naive character theoretic approach to computing projections requires $O(n \cdot n!)$ operations.

In this paper two types of polynomial time algorithms (quadratic in the size of the homogeneous space) are presented for partially ranked data. The first approach makes use of a more careful organization of the character theoretic computation and is applicable to arbitrary finite groups and their homogeneous spaces. The second approach makes use of the techniques of the combinatorial Radon transform.

1. Introduction

Let G be a finite group acting transitively on a set X. Often X is called a **homogeneous space** for G. Let $L(X)$ denote the vector space of complex-valued functions on X. Then $L(X)$ naturally admits a (permutation) representation ρ of G defined by

$$(\rho(s)(f))(x) = f(s^{-1}x) \tag{1.1}$$

for all $s \in G$ and all $x \in X$. As a representation space for G, $L(X)$ has a natural decomposition as a direct sum of isotypic subspaces

1991 *Mathematics Subject Classification.* Primary 20C40, 65U05; Secondary 05-04, 05E10.

The first author was supported in part by NSF Grant DMS 86-00235. The second author was supported in part by an NSF Mathematical Sciences Postdoctoral Fellowship. He also thanks Laci Babai and the Department of Computer Science at the University of Chicago for their hospitality while much of this was written.

This paper is in final form and no version will be submitted elsewhere.

$$L(X) = \oplus_{i=1}^{m} V_i.$$

The general problem may be stated as given a vector $f \in L(X)$ how may we quickly compute the projection of f onto any one (or some subset) of the V_i?

We were led to considering this calculation in the course of analyzing data on certain homogeneous spaces for the symmetric group. To explain, let S_n denote the symmetric group on n letters. Let λ be a partition of n. Thus,

$$\lambda = (\lambda_1, ..., \lambda_k)$$

where $\lambda_i \geq \lambda_{i+1} > 0$ and $\lambda_1 + ... + \lambda_k = n$. The corresponding Young subgroup $S_\lambda \leq S_n$ permutes the first λ_1 coordinates among themselves, the next λ_2 coordinates among themselves, and so on. Let X^λ denote the quotient space S_n/S_λ. The associated permutation representation of S_n on $M^\lambda = L(X^\lambda) = \{f : X^\lambda \longrightarrow \mathbb{R}\}$ is then of dimension $n!/\prod_i \lambda_i!$.

As motivation, consider the California Lotto game. This involves a choice of a random unordered 6-element subset of $\{1, 2, \ldots, 49\}$. Gamblers try to guess at the 6-set. The collection of 6-sets can be represented as the set of cosets, $S_{49}/S_6 \times S_{43}$. Indeed, S_{49} acts transitively on the 6-sets and the subgroup fixing $\{1, 2, 3, 4, 5, 6\}$ is $S_6 \times S_{43}$. Here, $\mid X^{(43,6)} \mid = 13,983,816$. About 10^6 people play the game every week. This leads to a function $f : X^{(43,6)} \longrightarrow \mathbb{R}$, $f(s)$ being the number of people choosing the subset s. The state takes about half of the total bet and splits the rest among the people choosing the winning 6-set.

Gamblers are interested in the function $f(s)$ - if one can pick a rarely chosen 6-set, one avoids having to split the prize. Chernoff [**8**] shows how this can actually lead to bets with positive expectation for the Massachusetts Lottery. The state is also interested in $f(s)$ - if there are no winners, the lottery is "rolled over" and betting activity increases. To make accurate fiscal projections, the number of roll-overs must be predicted and the data in $f(s)$ is clearly relevant.

One analysis of $f(s)$ begins by looking at the popularity of individual numbers. This is measured by averages like

$$\overline{f}(i) = \sum_{s \ni i} f(s).$$

Following this one might look at the popularity of pairs, triples and so on. The functions on 6-sets decompose naturally into distinct invariant irreducible subspaces (denoted $S^{(49-j,j)}$),

$$M^{(43,6)} = S^{49} \oplus S^{(48,1)} \oplus \ ... \ \oplus S^{(43,6)}.$$

The invariant subspace $S^{(49-j,j)}$ naturally measures the "pure" contribution of the popularity of the various j-sets. The analysis described above amounts to computing the projection of f onto these invariant subspaces. This approach is called **spectral analysis** by Diaconis ([**10**],[**11**]) who develops the subject in some detail.

Spectral analysis can be applied to partially ranked data of "shape λ". This involves n items (e.g. a list of 40 movies). Respondents are asked to choose from a list of n items their favorite λ_1 items (but not rank within) then their next favorite λ_2 items, and so on. For example, ranking the top 10 movies (in order) out of 40 leads to data of the shape (30,1,1,...,1) (10 ones). Sample surveys and the "q-sort" data collected by psychologists leads to other shapes. In each case spectral analysis offers a systematic method of analysis ([**10**], [**5**]).

Projections onto isotypic subspaces is a mainstay of the practical implementation of the spectral analysis approach to data. This generalizes to some degree the classical sums of squares decomposition of the standard ANOVA approach [**13**]. Computation of the isotypic projections is followed by computation of inner products of the projections with a sequence of easily interpretable vectors. Computation of the inner products is well understood. This paper focusses on the computation of the isotypic projections.

In this paper two algorithms are given for computation of the isotypic projections. The first method employs a more careful organization of the character theoretic approach and in fact, applies in the general case of decomposing the permutation representation of an arbitrary finite group G and homogeneous space X. This is developed fully, with special attention paid to the case in which $L(X)$ is multiplicity-free - as is the case for k-sets of an n-set. The results are developed as a series of reductions, each yielding greater computational improvements. Efficiency however, is gained at the cost of stronger assumptions on explicit knowledge of the representation theory of the associated symmetry group. Each subsection here contains an idea that is worth isolating. The strongest general result is contained in Theorem 2.4 with a slightly stronger result (Theorem 2.6) holding for distance transitive graphs.

The second algorithm uses ideas related to the **combinatorial Radon transform.** Use of an appropriate Radon transform (and its inverse) yield the decomposition of $M^{(n-k,k)}$. The work in this section is rather preliminary. Possible generalizations of this approach are discussed.

2. A character theoretic approach

2.1. Generalities. Let G be a finite group acting transitively on the finite set $X = \{x_1, ..., x_n\}$. Then X is called a **homogeneous space** for G. For any $x \in X$ let $Stab(x)$ denote the subgroup of elements of G which fix x. Let $L(X)$ denote the vector space of complex-valued functions on X. Then $L(X)$ naturally defines a representation ρ of G. This is precisely the permutation representation of G on the quotient space $G/Stab(x_i)$ for any i. (Equivalently this is the representation obtained by inducing the trivial representation from $Stab(x_i)$ to G.) As a vector space this is naturally equivalent to the complex vector space generated by the points $\{x_1, \ldots, x_n\}$. This associates a function $f \in L(X)$ to the point $f(x_1)x_1 + \ldots + f(x_n)x_n$.

Following Serre [**22**], the **isotypic decomposition** of $L(X)$ may be explained as follows. Let $\rho_1, \ldots, \rho_h$ be a complete set of inequivalent irreducible representations of G and let χ_i denote the character of ρ_i. As a representation space for G, $L(X)$ has a direct sum decomposition into irreducible G-invariant subspaces. That is,

$$L(X) = \oplus_{i=1}^{m} U_i. \tag{2.1}$$

where

$$\rho(s)U_i = U_i \qquad (U_i \neq 0)$$

and no nontrivial subspace of U_i has this property. For each i, $1 \leq i \leq m$, $\rho(s)$ restricted to U_i gives an irreducible representation of G. For each j, $1 \leq j \leq h$, let V_j be the subspace of $L(X)$ given by the direct sum over all U_i which define representations equivalent to ρ_j. Thus, the decomposition (2.1) may be rewritten as

$$L(X) = \oplus_{i=1}^{h} V_i$$

where V_i is the i^{th} **isotypic** subspace of ρ. (Note that some of these V_i may be 0.)

2.2. Computing projections via characters. The following classical theorem (see e.g. [22], Theorem 8) defines the projection matrices.

THEOREM 2.1. *With the above notation the projection matrix p_i of V onto V_i along $\oplus_{j \neq i} V_j$ is given by the formula*

$$p_i = \frac{dim(\rho_i)}{\mid G \mid} \sum_{s \in G} \chi_i(s)^* \rho(s).$$

Theorem 2.1 is completely general. In our particular case, where ρ is the permutation representation on $G/Stab(x_1)$ more can be said.

THEOREM 2.2. *Let X be a homogeneous space for G and let ρ be the associated permutation representation of G in $L(X)$. Let $\chi_1, \ldots, \chi_h$ be the irreducible characters of G and $\rho_1, \ldots, \rho_h$ be a corresponding set of irreducible representations.*

If $f = f^{(1)} + \cdots + f^{(h)}$ is the isotypic deomposition of $f \in L(X)$ (some $f^{(i)}$ may be 0) then for all $i \leq h$ and all $x \in X$:

$$f^{(i)}(x) = \frac{dim(\rho_i)}{\mid G \mid} \sum_{s \in G} \chi_i(s) f(sx). \tag{2.2}$$

PROOF. The definition of the permutation representation (1.1) and the definition of the projection p_i (Theorem 2.1) immediately yield the sequence of equalities,

$$\begin{aligned} f^{(i)}(x) = p_i f(x) &= \frac{dim(\rho_i)}{\mid G \mid} \sum_{s \in G} \chi_i^*(s)(\rho(s)f)(x) \\ &= \frac{dim(\rho_i)}{\mid G \mid} \sum_{s \in G} \chi_i^*(s) f(s^{-1}x). \\ &= \frac{dim(\rho_i)}{\mid G \mid} \sum_{s \in G} \chi_i(s^{-1}) f(s^{-1}x), \end{aligned}$$

with the last equation following from the fact that $\chi_i^*(s) = \chi_i(s^{-1})$. $\square$

Theorem 2.2 gives the following simple, albeit improvable, upper bound on the time needed to compute one projection.

COROLLARY 2.1. *The projections of f onto any isotypic subspace may be computed in at most $\mid X \mid\mid G \mid$ operations.*

Remark: The computational model used here counts a single complex multiplication and addition as one "operation". As such, this does not take into account the inner group multiplication (i.e. sx for $s \in G$, $x \in X$) needed to compute the point for which the function value is retrieved.

As a first step in reducing the number of calculations, consider the summation appearing in (2.2),

$$\sum_{s \in G} \chi_i(s) f(sx). \tag{2.3}$$

The idea is to decompose (2.3) into subsums which will occur repeatedly as the values $f^{(i)}(x)$ are computed over all $x \in X$. These subsums need be computed once and then stored to be recalled when needed. To do this, a closer investigation of the structure of the quotient space $G/Stab(x_i)$ is useful.

Fix $\{ s_1, ..., s_n \} \subseteq G$ such that

$$(i) \;\; s_1 = identity$$

$$(ii) \;\; s_i x_1 = x_i.$$

Also, let $H = Stab(x_1)$.

PROPOSITION 2.1. *With all the notation as above,*
(i) $Stab(x_i) = s_i H s_i^{-1}$.
(ii) Coset representatives for $G/Stab(x_i)$ are given by the set

$$\{s_i s_j s_i^{-1} \mid 1 \leq j \leq n\}.$$

Let

$$\chi_i^{(j)} = \sum_{s\in H} \chi_i(s_j s). \tag{2.4}$$

As will emerge, these $\chi_i^{(j)}$ are "generalized spherical functions". The following proposition shows that to compute projections, only certain sums involving the $\chi_i^{(j)}$ are needed.

PROPOSITION 2.2. *With notation as in Theorem 2.2 and (2.4),*

$$f^{(i)}(x_k) \;=\; \frac{dim(\rho_i)}{\mid G\mid} \sum_{j=1}^{n} f(s_k x_j)\chi_i^{(j)}. \tag{2.5}$$

PROOF. Consider first the computation of $f^{(i)}(x_1)$. Ignoring the outer scalar factor, the summation admits the following decomposition.

$$\begin{aligned}
\sum_{s\in G} \chi_i(s) f(sx_1) &= \sum_{j=1}^{n}\sum_{s\in H} f(s_j s x_1)\chi_i(s_j s)\\
&= \sum_{j=1}^{n} f(s_j x_1) \sum_{s\in H} \chi_i(s_j s)\\
&= \sum_{j=1}^{n} f(x_j)\chi_i^{(j)}
\end{aligned}$$

The subsums $\{\chi_i^{(j)}\}$ in fact determine the i^{th} projection. To calculate $f^{(i)}(x_k)$ where $k \neq 1$, (2.2) gives

$$f^{(i)}(x_k) = \frac{dim(\rho_i)}{\mid G\mid} \sum_{s\in G} f(sx_k)\chi_i(s).$$

Again, considering only the summation,

$$\begin{aligned}
\sum_{s\in G} f(sx_k)\chi_i(s) &= \sum_{j=1}^{n} \sum_{t\in Stab(x_k)} f(s_k s_j s_k^{-1} t x_k)\chi_i(s_k s_j s_k^{-1} t)\\
&= \sum_{j=1}^{n}\sum_{s\in H} f(s_k s_j s_k^{-1} s_k s s_k^{-1} x_k)\chi_i(s_k s_j s_k^{-1} s_k s s_k^{-1})\\
&= \sum_{j=1}^{n}\sum_{s\in H} f(s_k s_j s x_1)\chi_i(s_k s_j s s_k^{-1})\\
&= \sum_{j=1}^{n} f(s_k x_j) \sum_{s\in H} \chi_i(s_k s_j s s_k^{-1})\\
&= \sum_{j=1}^{n} f(s_k x_j) \sum_{s\in H} \chi_i(s_j s)
\end{aligned}$$

where the first equality follows from Proposition 2 and the last line follows from the fact that χ_i is a class function. So, finally

$$f^{(i)}(x_k) = \frac{dim(\rho_i)}{\mid G \mid} \sum_{j=1}^{n} f(s_k x_j)\chi_i^{(j)}.$$

□

Thus, the following initial reduction is obtained.

LEMMA 2.1. *For any fixed i, the projections onto the isotypic V_i may be computed in at most $O(\mid G \mid + \mid X \mid^2)$ operations.*

PROOF. We need at most $O(\mid G \mid)$ operations to calculate the subsums $\chi_i^{(j)}$ for all j. Then, for each x_k, $\mid X \mid$ additional multiplications to calculate $f^{(i)}(x_k)$. □

Note: The subsums $\chi_i^{(j)}$ need only be computed once - ever - and then stored. In situations in which X remains constant and f changes frequently and must be analyzed repeatedly, the above idea may prove useful. Of course, this does have the drawback that if $\mid G \mid$ is large (e.g. 49!) then these startup costs may prove prohibitive.

To reduce this $O(\mid G \mid)$ term, the subsums $\chi_i^{(j)}$ must be analyzed more closely. Since χ_i is a class function for G, for any t in the normalizer of H in G (denoted $N_G(H)$),

$$\begin{aligned} \chi_i^{(j)} &= \sum_{s \in H} \chi_i(t s_j s t^{-1}) \\ &= \sum_{s \in H} \chi_i(t s_j t^{-1} t s t^{-1}) \\ &= \sum_{s \in H} \chi_i(t s_j t^{-1} s). \end{aligned}$$

Thus,

LEMMA 2.2. *If s_j and s_k are conjugate by elements of the normalizer of H in G then*

$$\chi_i^{(j)} = \chi_i^{(k)}.$$

Pause for a moment to consider this reduction for a particular class of examples of interest, the homogeneous spaces for S_n given by its action on the collections of k-sets of an n-set where $k \leq \frac{n}{2}$. The canonical n-set is $\{1, ..., n\}$ and S_n acts in the usual way on the collection of subsets of size k. Let this homogeneous space be denoted as $X^{(n-k,k)}$. This may be realized by the obvious action of S_n on the space of binary n-tuples with k ones. (Here the k-set $\{i_1, ..., i_k\}$ with $i_j < i_{j+1}$ is represented by the n-tuple with 1's in positions i_j and 0's elsewhere.) The stabilizer of the point $\{1, ..., n-k\}$ is the subgroup $S_{n-k} \times S_k$ where S_{n-k} is the subgroup acting only on the elements $\{1, ..., n-k\}$

and S_k the subgroup which permutes $\{n-k+1, ..., n\}$. Let $M^{(n-k,k)}$ denote the the vector space $L(X^{(n-k,k)})$. The isotypic decomposition of these spaces is well-known. In section 4 this will be given in more detail, but for now it will suffice to state without proof,

THEOREM 2.3. *The isotypic decomposition of $M^{(n-k,k)}$ where $k \leq \frac{n}{2}$ is given by*

$$M^{(n-k,k)} = S^{(n)} \oplus S^{(n-1,1)} \oplus ... \oplus S^{(n-k,k)}$$

where $S^{(n-j,j)}$ denotes an irreducible representation of S_n of dimension $\binom{n}{j} - \binom{n}{j-1}$.

To compute the projections onto the $S^{(n-j,j)}$, the coset representatives of $S_n/S_{n-k} \times S_k$ must first be determined.

PROPOSITION 2.3. *Coset representatives for $S_n/S_{n-k} \times S_k$ are given by the identity and all products of transpositions of the form*

$$(i_1 j_1)...(i_r j_r)$$

where $1 \leq r \leq k$ and both

$$n-k+1 \leq j_1 < ... < j_r \leq n$$

and

$$1 \leq i_1 < ... < i_r \leq n-k.$$

PROOF. Consider the action of S_n on the binary n-tuples with k ones and in particular consider the action on the vector

$$(\underbrace{0, 0, ..., 0}_{n-k\ 0's}, \underbrace{1, 1, ..., 1}_{k\ 1's}).$$

By inspection, the above permutations yield all other such n-tuples. $\square$

It is easy to see that for $k < n/2$, the normalizer of $S_{n-k} \times S_k$ in S_n is just $S_{n-k} \times S_k$. For $k = n/2$ it is isomorphic to a wreath product, $S_2\ wr\ S_{n/2}$, containing $S_{n/2} \times S_{n/2}$. The following lemma determines the number of orbits for the coset representatives under the action of $S_{n-k} \times S_k$ by conjugation.

LEMMA 2.3. *The coset representatives given for $S_n/S_{n-k} \times S_k$ in Lemma 2.2 are divided into $k+1$ orbits with respect to the action of $S_{n-k} \times S_k$ by conjugation.*

PROOF. By inspection one can see that under the action of $S_{n-k} \times S_k$ the coset representatives, as given in proposition 3, break up into equivalence classes determined by the number of transpositions needed to write them. These coset representatives can be taken to be the set

$$\{1, (1,n), (1,n-1)(2,n), ..., (1,n-k+1)\cdots(n-k,n)\}.$$

□

Examples: (These have been grouped by conjugacy class with respect to the action of the stabilizing subgroup.)

(1) Coset Representatives for $S_6/S_5 \times S_1$:

$$\{1\},\ \{(16), (26), (36), (46), (56)\}$$

(2) Coset Representatives for $S_6/S_4 \times S_2$:

$$\{1\}, \{(16), (26), (36), (46), (15), (25), (35), (45)\},$$
$$\{(15)(26), (15)(36), (15)(46), (25)(36), (25)(46), (35)(46)\}$$

(3) Coset Representatives for $S_6/S_3 \times S_3$:

$$\{1\}, \{(16), (26), (36), (15), (25), (35), (14), (24), (34)\},$$
$$\{(14)(25), (14)(35), (14)(26), (14)(36),$$
$$(24)(35), (24)(36), (15)(26), (15)(36), (25)(36)\},$$
$$\{(14)(25)(36)\}$$

This reduction gives the following improvements. These will be further refined later.

COROLLARY 2.2. *For fixed $j \le k$ at most*

$$O((k+1)(n-k)!k! + \binom{n}{k}^2)$$

operations are need to calculate the projection of any $f \in M^{(n-k,k)}$ onto $S^{(n-j,j)}$.

PROOF. At most $(n-k)!k!$ operations are needed to calculate each of the $k+1$ different $\chi_i^{(j)}$. Then for each of the $\binom{n}{k}$ points x at most another $\binom{n}{k}$ will be required to calculate $f^{(i)}(x)$. □

COROLLARY 2.3. *At most*

$$O(k((k+1)(n-k)!k! + \binom{n}{k}^2))$$

operations are needed to calculate the projections onto every $S^{(n-j,j)}$.

The reductions in this section have in fact been implemented and used. See [**5**] for such an example.

2.3. Computing projections via matrix representations. As noted previously, assuming that this is a computation which is to be performed frequently, the character subsum calculation can be viewed as a "start-up" cost. The actual "interesting" computation takes only $O(\binom{n}{k}^2)$ operations. This compares very favorably with the naive bound of $(n!)^2$.

However, the associated start-up cost for the character theory approach of section 2.2 may still prove to be prohibitive. For instance, in the example cited in the introduction, in which we wished to compute projections from $M^{(43,6)}$, this would require performing at least 43! operations. Clearly further reduction is necessary. This may be achieved if certain explicit matrix representations are well-understood. This is now explained.

The sum $\chi_i^{(j)}$ is a sum of traces of a certain collection of matrices. Let ρ_i be an explicit matrix representation with character χ_i. Then,

$$\begin{aligned} \chi_i^{(j)} &= \sum_{s\in H} \chi_i(s_j s) \\ &= \sum_{s\in H} trace(\rho_i(s_j s)) \\ &= trace(\rho_i(s_j) \sum_{s\in H} \rho_i(s)). \end{aligned}$$

Consider the inner sum $\sum_{s\in H} \rho_i(s)$. As this is the sum of the values of a matrix representation over an entire subgroup (the representation ρ_i restricted to the subgroup H) this is "known". To explain, recall that for any irreducible unitary representation η of a group K,

$$\sum_{t\in K} \eta(t) = \begin{cases} \mid K \mid & \text{if } \eta = 1_K \\ d_\eta \times d_\eta \quad 0 \text{ matrix otherwise} & \end{cases} \tag{2.6}$$

where 1_K denotes the trivial representation of K ([**22**], section 2.2). Let $I_H(\rho_i)$ be the number of times that the trivial representation occurs in ρ_i restricted to H. Then recall,

$$I_H(\rho_i) = \langle \rho_i \downarrow H, 1_H \rangle = \langle \rho_i, 1_H \uparrow G \rangle. \tag{2.7}$$

In (2.7) restriction of a representation is denoted by $\downarrow$, while $\uparrow$ denotes induction. The notation $\langle \eta, \psi \rangle$ denotes the usual inner product of two representations. Then (2.7) follows from Frobenius reciprocity ([**22**], Theorem 13). Note that $1_H \uparrow G$ is just the representation ρ.

Say that a matrix representation ψ of G is H-**adapted** if for any $t \in H$,

$$\psi(t) = \begin{pmatrix} \eta_1(t) & 0 & \cdots & 0 \\ 0 & \eta_2(t) & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \eta_l(t) \end{pmatrix} \tag{2.8}$$

where each η_i is an irreducible unitary representation of H.

Then application of (2.6) shows that

$$\chi_i^{(j)} = trace(\rho_i(s_j) \begin{pmatrix} \mid H \mid & 0 & \dots & 0 \\ 0 & \mid H \mid & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix})$$

where the inner matrix has $I_H(\rho_i)$ occurrences of $\mid H \mid$ on the diagonal, 0 elsewhere. Thus,

$$\chi_i^{(j)} = \mid H \mid \sum_{k=1}^{I_H(\rho_i)} (\rho_i(s_j))_{kk}. \tag{2.9}$$

Finally, note that if s_l and s_j are in the same $H\backslash G/H$ double coset, then assuming that the representation ρ_i is H-adapted,

$$(\rho_i(s_j))_{k,k} = (\rho_i(s_l))_{k,k}.$$

Hence, a further reduction in the start-up cost is obtained.

THEOREM 2.4. *Let ρ_i be an irreducible representation of G corresponding to the isotypic subspace V_i such that ρ_i is H-adapted. Then assuming the precomputation of the set of matrix elements*

$$\{(\rho_i(s_j))_{k,k} \mid 1 \leq j \leq m, 1 \leq k \leq I_H(\rho_i)\},$$

the projection onto V_i may be computed in at most

$$O(\mid X \mid^2 + \mid H\backslash G/H \mid \cdot I_H(\rho_i))$$

operations.

Remark: Theorem 4 depends upon the knowledge of explicit matrix representations for those representations occuring in the permutation representation on $L(X)$. Furthermore, assuming that the representations are known, they must be given in bases that are H-adapted. Assuming the existence of some realization of the representations as matrix representations the latter problem of constructing an appropriate change of basis has been solved (cf. [**4**], section 2 or [**21**], Appendix).

For the symmetric group S_n, Young's orthogonal form provides explicit formulas for a complete set of unitary irreducible matrix representations (cf. [**18**], section 25) evaluated at the set of pairwise adjacent transpositions. Additionally, for the permutation representations arising from quotients by Young subgroups, the decomposition into irreducibles is given by Young's Rule (cf. [**18**], section 14).

For an arbitrary group, explicit matrix representations may not be known. Babai and Ronyai [**2**] treat the problem of constructing a complete set by decomposing the group algebra.

It is worth pointing out that more generally, it may also prove fruitful to view the problem of computing projections in the context of matrix algebras. Babai, Friedl and Stricker [**1**], extending work of Friedl and Ronyai [**16**] show that deterministic polynomial time (in the dimension) algorithms exist for decomposing semisimple matrix algebras over number fields (closed under conjugate-transpose) into their simple components. There has been no investigation of the practicality of these algorithms.

2.4. The multiplicity-free case. A particular case of interest, which includes $M^{(n-k,k)}$, is that in which $L(X)$ is **multiplicity-free**. In this case $I_H(\eta)$ is 1 or 0 for each representation η of G. So,

$$L(X) \;=\; V_1 \oplus \ldots \oplus V_l$$

where the V_i are not only isotypics, but also inequivalent irreducible subspaces. Recall that (G,H) is said to be a **Gelfand pair** if the permutation representation of G associated to G/H is multiplicity-free. Assume that ρ_i is an irreducible representation which occurs in $L(X)$ in this case and furthermore assume that it is given as an H-adapted matrix representation of G. Then for any $t \in H$,

$$\rho_i(t) \;=\; \begin{pmatrix} 1 & 0 \\ 0 & B(t) \end{pmatrix}$$

where $B(t)$ is block diagonal

$$B(t) = \begin{pmatrix} \eta_1(t) & 0 & \cdots & 0 \\ 0 & \eta_2(t) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \eta_m(t) \end{pmatrix}$$

and the η_j are non-trivial unitary matrix representations of H (depending on ρ_i). Then for all $s \in G$ the H-bi-invariant function that picks out the 1,1 entry of the matrix $\rho_i(s)$ is called the i^{th} **spherical function** of the Gelfand pair (G,H). Let these be denoted σ_i, then

$$\chi_i^{(j)} = \mid H \mid \sigma_i(s_j). \tag{2.10}$$

Thus, computation of the projection onto the i^{th} isotypic (denoted $f^{(i)}$) is given as a **spherical transform**. That is, recall

$$f^{(i)}(x_k) = \frac{\chi_i(1)}{|G|} \cdot \sum_{j=1}^{n} f(s_k x_j)\chi_i^{(j)} \tag{2.11}$$

$$= \frac{\chi_i(1)}{|G|} \cdot |H| \cdot \sum_{j=1}^{n} f(s_k x_j)\sigma_i(s_j) \tag{2.12}$$

where (2.12) follows from (2.10). To reduce this computation further we take advantage of the H-bi-invariance of the spherical functions. For this, consider the action of the subgroup H on X. By applying Frobenius reciprocity (or equivalently, using "Mackey theory"), it can be shown that the number of orbits is equal to the number of nontrivial isotypic subspaces. Thus the partitioning into orbits can be written as

$$X = \Omega_1 \coprod \cdots \coprod \Omega_l$$

so that l is also the number of H, H double cosets.

This partitioning allows for a rewriting of (2.12) by

$$f^{(i)}(x_k) = \frac{\chi_i(1)}{|G|} \cdot |H| \cdot \sum_{r=1}^{l} \sum_{x \in \Omega_r} f(s_k x)\sigma_i(\Omega_r) \tag{2.13}$$

$$= \frac{\chi_i(1)}{|G|} \cdot |H| \cdot \sum_{r=1}^{l} \sigma_i(\Omega_r) \sum_{x \in \Omega_r} f(s_k x) \tag{2.14}$$

$$= \frac{\chi_i(1)}{|G|} \cdot |H| \cdot \sum_{r=1}^{l} \sigma_i(\Omega_r) f_k(\Omega_r) \tag{2.15}$$

where $f_k(\Omega_r) = \sum_{x \in \Omega_r} f(s_k x)$. Finally, write

$$\widehat{f_k}(i) = |H| \cdot \sum_{r=1}^{l} f_k(\Omega_r)\sigma_i(\Omega_r) \tag{2.16}$$

and call (2.16) the i^{th} **spherical transform** of f_k.

Note that at most $|X|$ additions are needed to compute any given f_k, so at most $|X|^2$ are required to compute all f_k. At most $|H \backslash G / H|$ operations are required to compute any single spherical transform, so at most $|H \backslash G / H| \cdot |X|$ operations are needed for each of the spherical transforms necessary for a single projection. This is summarized by the Theorem 2.5 below.

THEOREM 2.5. *Let X be a homogeneous space for G that yields a multiplicity-free representation of G. Then, assuming that the spherical functions for the associated Gelfand pair are known, the projection onto any one of the isotypics may be calculated as a sequence of spherical transforms (eq. 2.16). This requires at most $|X|^2$ additions, followed by at most an additional $|H \backslash G / H| \cdot |X|$*

operations. To compute the projections onto every isotypic would require an additional $\mid H\backslash G/H \mid^2 \cdot \mid X \mid$ *operations.*

COROLLARY 2.4. *All isotypic projections for* $M^{(n-k,k)}$ *may be computed using* $\binom{n}{k}^2$ *additions and* $\binom{n}{k}(k+1)^2$ *operations.*

2.5. Distance transitive graphs. As a final reduction in this section, the interesting case of multiplicty-free representations arising from distance-transitive graphs is considered. In this case, the set X is to be considered as an undirected graph with the usual metric given by edge-distance between points. It is assumed that G acts as isometries on the graph, so that $d(sx, sy) = (x, y)$ for all $s \in G$ and $x, y \in X$. The graph is said to be **distance transitive** if given any two pairs of points (x, y), (x', y') such that $d(x, y) = d(x', y')$, there exists $s \in G$ such that $(sx, sy) = (x', y')$.

Let Ω_k denote the set of points of distance k from the basepoint x_1. Then the algebra of functions constant on each Ω_k is isomorphic to the algebra of H-bi-invariant functions on G ($H = Stab(x_1)$), which has basis equal to the set of associated spherical functions. See Brouwer, Cohen and Neumaier [**7**] or Stanton [**23**] for excellent surveys of the subject of distance transitive graphs and their spherical functions as well as for pointers to other relevant literature.

For the spherical functions which arise from distance transitive graphs, spherical transforms may be computed more efficiently than by direct computation. Recent work of Driscoll, Healy and Rockmore shows that for such graphs, the **discrete spherical transform** defined as the set of spherical transforms $\{\widehat{f}(i)\}_i$, can be computed in at most $O(n \log^2 n)$ operations versus n^2 operations required for direct computation. Their algorithm makes use of the fact that for distance transitive graphs, the spherical functions satisfy a three-term recurrence. See [**15**] for details.

Such graphs may be also studied from the point of view of association schemes. Bannai and Ito [**3**] give a full treatment of the subject.

The set of k-sets of an n-set naturally form a graph where two k-sets are joined if and only if they differ by a single element. This graph is in fact distance transitive and is sometimes referred to as the Johnson scheme in the association scheme literature. The associated spherical functions can be written as Hahn polynomials so that the isotypic projections can be computed as Hahn polynomial transforms. In [**15**] this example is worked out in full detail (section 3.3). Using this result gives a final reduction for computing projections for $M^{(n-k,k)}$.

THEOREM 2.6. *All isotypic projections for* $M^{(n-k,k)}$ *may be computed using at most* $\binom{n}{k}^2$ *additions followed by* $O(\binom{n}{k} k \log^2 k)$ *operations.*

3. Radon Transform Approach

In the particular case of k-sets of an n-set, special techniques are available. We believe that these techniques extend to other standard examples (see section

3.3).

3.1. k-sets of an n-set. For $k \leq n/2$, let X denote the $\binom{n}{k}$ k-element subsets of $\{1, \ldots, n\}$. Abbreviate $L(X) = M^{(n-k,k)}$ as M^k in this section. For $0 < j \leq k \leq n/2$, define the **Radon transform** $R : M^j \longrightarrow M^k$ by summing over subsets:

$$Rf(t) = \sum_{s \subset t} f(s)$$

for $\mid s \mid = j$ and $\mid t \mid = k$.

This type of combinatorial Radon transform was introduced by Bolker [**6**]. Kung [**20**] gives an extensive survey. Diaconis and Graham [**12**] discuss algorithmic aspects.

Each of these authors prove that this particular version of the transform is one-to-one when $0 < j \leq k \leq n/2$. Graham, Li and Li [**17**] give an explicit left inverse. To describe this, choose a basis of delta functions for M^k:

$$\delta_t(t') = \begin{cases} 1 & \text{if } t = t', \\ 0 & \text{otherwise.} \end{cases}$$

Choose a similar basis for M^j. Then the Radon transform can be regarded as an $\binom{n}{k}$ by $\binom{n}{j}$ matrix, R, with entries

$$R_{ts} = \begin{cases} 1 & \text{if } s \subset t, \\ 0 & \text{otherwise} \end{cases} \tag{3.1}$$

for all $\mid s \mid = j$ and $\mid t \mid = k$.

The left inverse R^- is an $\binom{n}{j}$ by $\binom{n}{k}$ matrix satisfying $R^- R = I$, the $\binom{n}{j}$ by $\binom{n}{j}$ identity matrix. Graham, Li and Li [**17**] give R^- explicitly as

$$R^-_{s,t} = \frac{(-1)^{k-j}(k-j)}{(-1)^{|t-s|} \mid t-s \mid} \frac{1}{\binom{n-j}{|t-s|}} \tag{3.2}$$

with $\mid t - s \mid$ denoting the cardinality of the set of elements in t but not in s.

It follows from the definitions that both R and R^- are S_n-homomorphisms. To state the first lemma, regard M^j and M^k as inner product spaces by defining

$$\langle \delta_s \mid \delta_{s'} \rangle = \begin{cases} 1 & \text{if } s = s', \\ 0 & \text{otherwise} \end{cases}$$

LEMMA 3.1. *For R and R^- defined above let $\pi = RR^-$. Then π is a projection with range isomorphic to M^j.*

PROOF. Clearly $\pi^2 = \pi$. By multiplying out, the general entry of π can be written down explicitly. If t and t' are k-sets, the t, t' entry of π only depends on the cardinality of $\mid t \cap t' \mid = z$ (say). So $0 \leq z \leq k$ and

$$\pi_{tt'} = \frac{k-j}{n-j}\sum_{u=0}^{j}(-1)^{j-u}\frac{\binom{z}{u}\binom{k-z}{j-u}}{\binom{n-z-1}{n-u-1}}. \tag{3.3}$$

This clearly shows that $\pi_{tt'} = \pi_{t't}$. So, π is a projection. Since π commutes with the action of S_n, the range of π is an invariant subspace. Let δ_s be a basis for M^j as s runs through j-sets. The $\bar{\delta_t} = R\delta_t$ are linearly independent and in the range of π which is thus isomorphic to M^j. $\square$

The decomposition of M^k into irreducibles is well-known to be multiplicity-free. Using the notation of James [**18**] this decomposition is given by

$$M^k \cong S^{(n)} \oplus S^{(n-1,1)} \oplus S^{(n-2,2)} \oplus \cdots \oplus S^{(n-k,k)} \tag{3.4}$$

$$\cong M^{k-1} \oplus S^{(n-k,k)}. \tag{3.5}$$

See Stanton [**23**] for a direct proof. Bolker [**6**] gives a proof using Radon transform arguments. Using induction Lemma 1 yields

Corollary 3.1. *Let $\pi^{(j)}$ be the projection from M^k described in Lemma 1. Then a projection onto the irreducible subspace $S^{(n-j,j)} \subset M^k$ can be computed as $\pi^{(j)} - \pi^{(j-1)}$, $1 \le j \le k$, with $\pi^{(0)}$ equal to the projection onto $S^{(n)}$ given by $1/\binom{n}{k} \cdot J$. (J denotes the $\binom{n}{k}$ by $\binom{n}{k}$ matrix of all ones.)*

The rows and columns of the projections matrices $\pi^{(j)}$ are indexed by the k-sets, with any given entry only depending on the cardinality of the intersection of the associated k-sets (see eqn. (3.3)). Computation of this intersection requires at most on the order of $k \log k$ operations. Consequently, computation of this matrix will require at most $O(\binom{n}{k}^2 k \log k)$ operations. Thus to obtain all projection matrices requires at most $O(\binom{n}{k}^2 k^2 \log k)$ operations. To compare, the startup cost required of the spherical function approach is $\binom{n}{k}^2$ additions. Either approach is feasible.

Having built the projection matrices, the isotypic projections may be computed as a sequence of k matrix-vector multiplications.

Theorem 3.1. *To compute the projections of an arbitrary element of M^k onto each of the $k+1$ isotypic subspaces via Radon transform methods requires a startup cost of at most $O(k^2 \log k\binom{n}{k}^2)$ operations and then an additional $k\binom{n}{k}^2$ operations.*

3.2. Further directions. The techniques of section 3.1 seem to present many possiblities for further investigation. We close here by indicating a few of these.

1. $GL_n(\mathbb{F}_q)/GL_{n-k}(\mathbb{F}_q) \times GL_k(\mathbb{F}_q)$. The quotient $S_n/S_{n-k} \times S_k$ has a "q-analogue" given by $GL_n(\mathbb{F}_q)/GL_{n-k}(\mathbb{F}_q) \times GL_k(\mathbb{F}_q)$. The associated permutation representation of $GL_n(\mathbb{F}_q)$ is given by the action of $GL_n(\mathbb{F}_q)$ on k-dimensional subspaces of $\mathbb{F}_q^n$. As for k-sets of an n-set, this is a multiplicity-free representation (and even a distance transitive graph with associated spherical functions given by the q-Hahn polynomials). The techniques of the Radon transform are again available. If $X^k(q)$ denotes the set of k-dimensional subspaces of F_q^n, then for $j < k \leq n/2$ define $R : L(X^j(q)) \longrightarrow L(X^k(q))$ by

$$Rf(y) = \sum_{x \subset y} f(x).$$

2. General Young subgroups and parabolic subgroups. Let $\lambda \vdash n$ and let S_λ denote the corresponding Young subgroup of S_n and G_λ the corresponding parabolic subgroup of $GL_n(\mathbb{F}_q)$. Work of James suggests that the techniques of the Radon transform could be applicable in this situation for computing projections onto not only isotypic subspaces, but even onto a complete set of irreducible subspaces. For S_n see ([**18**], section 17) and for $GL_n(\mathbb{F}_q)$ see ([**19**], section 15).

3. Connections between Radon transform and character theory. There must be a clear connection between the material of section 2 and the that of section 3.1. This has not yet been investigated.

4. We would be remiss if we did not point out the closely related area of noncommutative fast Fourier transforms. See Clausen [**9**] and Diaconis and Rockmore [**14**] and the references cited there. These ideas speed computation of the discrete Fourier transform by successive restriction through a tower of subgroups. Such restrictions have not been invoked in any of the above approaches. This idea may be of use for the computation considered here and provide even further speedups.

References

1. L. Babai, K. Friedl, and M. Stricker, *Decomposition of *-closed algebras in polynomial time*, Technical Report, University of Chicago, Department of Computer Science, 1992.
2. L. Babai and L. Ronyai, *Computing irreducible representations of finite groups*, Math. Comp. **55** (1990), 705–722.
3. E. Bannai and T. Ito, *Algebraic Combinatorics I. Association Schemes.* Benjamin/Cummings Pub. Co., Inc., Menlo Park, CA 1984.
4. U. Baum, *Existence and efficient construction of fast Fourier transforms on supersolvable groups*, Ph.D. Thesis, University of Bonn, 1991.
5. L. Beckett and P. Diaconis, *Spectral analysis for discrete longitudinal data*, Technical Report, Department of Mathematics, Harvard University, 1991.
6. E. Bolker, *The finite Radon transform*, Contemporary Math **63** (1987), 27–50.
7. A. E. Brouwer, A. M. Cohen and A. C. Neumaier, *Distance Regular Graphs*, Springer-Verlag, Berlin, 1990.
8. H. Chernoff, *An analysis of the Massachusetts numbers game*, Math. Intelligencer, **3** (1981), 166–172.
9. M. Clausen, *Fast generalized Fourier transforms*, Theoret. Comp. Sci. **67** (1989), 55–63.
10. P. Diaconis, *Spectral analysis for ranked data*, Ann. of Stat., **17** (1989), 349-379.

11. P. Diaconis, *Group Representations in Probability and Statistics*, Inst. of Math. Stat., Hayward, CA, 1989.
12. P. Diaconis and R. Graham, *The Radon transform on* $\mathbb{Z}_2^k$, Pacific J. of Math. **118** (1985), 323–345.
13. P. Diaconis and D. Rockmore, *Representation theoy and ANOVA for designed experiments*, Technical Report, Harvard University, Department of Mathematics, 1991.
14. P. Diaconis and D. Rockmore, *Efficient computation of the Fourier transform on finite groups*, Jour. of AMS **3** (1990), 297–332.
15. J. Driscoll, D. Healy, and D. Rockmore, *Fast spherical transforms on distance transitive graphs*, Technical Report, Dartmouth College, Department of Mathematics and Computer Science, 1991.
16. K. Friedl and L. Ronyai, *Polynomial time solutions of some problems in computational algebra*, Proc. 17th ACM STOC, 1983, 153–162.
17. R. L. Graham, R. Li, and W. Li, *On the structure of t-designs*, SIAM J. Alg. and Disc. Meth. **1** (1980), 8–14.
18. G. D. James, *The Representation Theory of the Symmetric Group*, Lecture Notes in Math., vol. 682, Springer-Verlag, Berlin, 1978.
19. G. D. James, *Representations of General Linear Groups*, London Math. Soc. Lecture Notes, vol. 94, Cambridge Univ. Press, Great Britain, 1984.
20. J. P. Kung, *Radon transforms in combinatorics and lattice theory*, Contemporary Math. **57** (1986), 36–74.
21. D. Rockmore, *Fast Fourier analysis for finite groups*, Ph.D. Thesis, Harvard University, 1989.
22. J. P. Serre, *Linear Representations of Finite Groups*, Springer-Verlag, NY , 1986.
23. D. Stanton, *Orthogonal polynomials and Chevalley groups*, Special Functions: Group Theoretical Aspects and Applications. (ed. R. Askey, et. al.), 1984, 87–128.

Dept. of Mathematics, Harvard University, Cambridge, Massachusetts 02138

Dept. of Mathematics and Computer Science, Dartmouth College, Hanover, New Hampshire 03755