Supplement to “Multiresolution analysis on the symmetric group”

Risi Kondor and Walter Dempsey
Department of Statistics and Department of Computer Science
The University of Chicago
{risi,wdempsey}@uchicago.edu

Appendix A: Requisite definitions from group theory and representation theory

Conjugation. We denote the complex conjugate of a number $z \in \mathbb{C}$ as z^* and the Hermitian conjugate of a matrix M as $M^{*\top}$.

Groups. A **group** is a set G endowed with an operation $G \times G \to G$ (usually denoted multiplicatively) obeying the following axioms:

G1. for any $x, y \in G,\ xy \in G$ (closure);
G2. for any $x, y, z \in G,\ x(yz) = (xy)z$ (associativity);
G3. there is a unique $e \in G$, called the **identity** of G, such that $ex = xe = x$ for any $x \in G$;
G4. for any $x \in G$, there is a corresponding element $x^{-1} \in G$ called the **inverse** of x, such that $x x^{-1} = x^{-1} x = e$.

We *do not* require that the group operation be commutative, i.e., in general, $xy \neq yx$. A subset H of G is called a **subgroup** of G, denoted $H < G$, if H itself forms a group with respect to the same operation as G, i.e., if for any $x, y \in H,\ xy \in H$. If $x \in G$ and $H < G$, then $xH = \{\, xh \mid h \in H \,\} \subset G$ is called a (left–)H–**coset**.

Representations. For the purposes of this paper a **representation** of G over $\mathbb{C}$ is a matrix-valued function $\rho\colon G \to \mathbb{C}^{d_\rho \times d_\rho}$ such that $\rho(x)\,\rho(y) = \rho(xy)$ for any $x, y \in G$. We call d_ρ the **order** or the **dimensionality** of ρ. Note that $\rho(e) = I$ for any representation. Two representations ρ_1 and ρ_2 of the same dimensionality d are said to be **equivalent** if for some invertible $T \in \mathbb{C}^{d\times d}$, $\rho_1(x) = T^{-1} \rho_2(x)\, T$ for any $x \in G$. A representation ρ is said to be **reducible** if it decomposes into a direct sum of smaller representations in the form

$$\rho(x) = T^{-1}\left(\rho_1(x) \oplus \rho_2(x)\right) T = T^{-1}\left(\begin{array}{c|c} \rho_1(x) & 0 \\ \hline 0 & \rho_2(x) \end{array}\right) T \qquad \forall\, x \in G$$

for some invertible $T \in \mathbb{C}^{d_\rho \times d_\rho}$.

Irreps. A maximal set of pairwise inequivalent, irreducible representations we call a system of **irreps**. It is possible to show that if $\mathcal{R}_1$ and $\mathcal{R}_2$ are two different systems of irreps of the same finite group G, then $\mathcal{R}_1$ and $\mathcal{R}_2$ have the same (finite) cardinality, and there is a bijection $\phi\colon \mathcal{R}_1 \to \mathcal{R}_2$ such that if $\phi(\rho_1) = \rho_2$, then ρ_1 and ρ_2 are equivalent. The theorem of total reducibility asserts that given a system of irreps $\{\rho_1, \rho_2, \ldots, \rho_k\}$, any representation ρ can be reduced into a direct sum of the ρ_i's in the sense that there is an invertible matrix T and sequence of mutiplicities $m_1, m_2, \ldots, m_k$ such that $\rho(x) = T^{-1}\left[\bigoplus_{i=1}^{k} \bigoplus_{j=1}^{m_k} \rho_i(x)\right] T$. For more information on representation theory, see, e.g. (J.–P. Serre: Linear Representations of Finite Groups, Springer–Verlag, 1977).

Unitarity irreps. A representation ρ is said to be **unitary** if $\rho(x)^{-1} = \rho(x^{-1}) = \rho(x)^\dagger$ for all $x \in G$. A finite group always has at least one system of irreps which is unitary. In particular, Young's orthogonal reprsentation (see below) is such a system. Throughout the paper by "irrep" we will tacitly always mean "unitary irrep".

Permutations and $\mathbb{S}_n$. A **permutation** of $\{1, \ldots, n\}$ is a bijective mapping $\sigma \colon \{1, \ldots, n\} \to \{1, \ldots, n\}$. The product of two permutations is defined by composition, $(\sigma_2\sigma_1)(i) = \sigma_2(\sigma_1(i))$ for all i. With respect to this operation the set of all $n!$ possible permutations of $\{1, \ldots, n\}$ form a group called the **symmetric group** of degree n, which we denote $\mathbb{S}_n$. For $m < n$ we identify $\mathbb{S}_m$ with the subgroup of permutations that only permute $\{1, \ldots, m\}$, i.e., $\mathbb{S}_m = \{\, \tau \in \mathbb{S}_n \mid \tau(m+1) = m+1,\ \tau(m+2) = m+2,\ \ldots,\ \tau(n) = n \,\}$.

Cycle notation. A **cycle** in a permutation $\sigma \in \mathbb{S}_n$ is a sequence $(c_1, c_2, \ldots, c_k)$ such that for $i = 1, 2, \ldots, k-1$, $\sigma(c_i) = c_{i+1}$ and $\sigma(c_k) = c_1$. Any permutation can be expressed as a product of disjoint cycles. Some special permutations that we are interested in are the **transpositions** (i, j), the **adjacent transpositions** $\tau_i = (i, i+1)$, and the **contiguous cycles** $[\![i, j]\!] = (i, i+1, \ldots, j)$.

Partitions and Young diagrams. A sequence of positive integers $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$ is said to be an **integer partition** of n (denoted $\lambda \vdash n$) if $\sum_{i=1}^k \lambda_i = n$ and $\lambda_i \geq \lambda_{i+1}$ for $i = 1, 2, \ldots, k-1$. The **Young diagram** of λ consists of $\lambda_1, \lambda_2, \ldots, \lambda_k$ boxes laid down in consecutive rows, as in

$$\begin{ytableau} ~ & ~ & ~ & ~ \\ ~ & ~ & ~ \\ ~ \end{ytableau}$$

for $\lambda = (4, 3, 1)$. We define $\Lambda_n = \{\, \lambda \mid \lambda \vdash n \,\}$.

Young tableaux. A Young diagram with numbers in its cells is called a **Young tableau**. A **standard Young tableau (SYT)** is a Young tableau in which each of the numbers $1, 2, \ldots, n$ is featured exactly once, and in such a way that in each row the numbers increase from left to right and in each column they increase from top to bottom. For example,

$$\begin{ytableau} 1 & 2 & 5 & 8 \\ 3 & 4 & 7 \\ 6 \end{ytableau} \tag{1}$$

is a standard tableau of shape $\lambda = (4, 3, 1)$. We define $\mathcal{T}_n$ to be the set of all standard tableaux of n boxes, $\mathcal{T}_\lambda \subseteq \mathcal{T}_n$ to be the set of all standard tableaux of shape $\lambda \vdash n$, and $\lambda(t)$ to be the shape of the tableau t.

Partial orders on partitons and SYT's. There is a natural partial order on partitions induced by the inclusion order of their respective diagrams. In particular, for $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$ and $\mu = (\mu_1, \mu_2, \ldots, \mu_\ell)$ we write $\lambda \geq \mu$ if $k \geq \ell$, and $\lambda_i \geq \mu_i$ for $i = 1, 2, \ldots, \ell$. For example,

$$\begin{ytableau} ~ & ~ & ~ & ~ \\ ~ & ~ & ~ \\ ~ \end{ytableau} \quad \geq \quad \begin{ytableau} ~ & ~ \\ ~ & ~ \\ ~ \end{ytableau} \qquad \text{and} \qquad \begin{ytableau} ~ & ~ & ~ & ~ \\ ~ & ~ & ~ \\ ~ \end{ytableau} \quad \geq \quad \begin{ytableau} ~ & ~ & ~ & ~ \end{ytableau}\ .$$

If $\lambda \geq \mu$, then we say that λ is a **descendant** of μ or that μ is an **ancestor** of λ. Similarly, for a pair of standard Young tableaux t and t' (or the corresponding Yamanouchi symbols) we write $t \geq t'$ (and say that t is a descendant of its ancestor t') if t' is a subtableau of t. For example,

$$\begin{ytableau} 1 & 2 & 5 & 8 \\ 3 & 4 & 7 \\ 6 \end{ytableau} \quad \geq \quad \begin{ytableau} 1 & 2 & 5 \\ 3 & 4 \\ 6 \end{ytableau}$$

Naturally, if t is of shape λ and t' is of shape μ, then $t \geq t'$ imples $\lambda \geq \mu$, but not the other way round.

Restriction and extension relations. For $n > m$ and $t \in \Lambda_n$ we define $\lambda \downarrow_m := \{\, \lambda' \in \Lambda_m \mid \lambda' \leq \lambda \,\}$ and for $\lambda' \in \Lambda_m$, we define $\lambda' \uparrow^n := \{\, \lambda \in \Lambda_n \mid \lambda \geq \lambda' \,\}$. Similarly, for $t \in \mathcal{T}_n$, we define $t \downarrow_m$ to be the *unique* $t' \in \mathcal{T}_m$ that is an ancestor of t, and we define $t' \uparrow^n = \{\, t \in \mathcal{T}_n \mid t \geq t' \,\}$. The $\uparrow$ and $\downarrow$ operators are also extended to sets of partitions/tableaux in the natural way.

Indexing the irreps. The significance of integer partitions and standard Young tableaux is that given a system of irreps $\mathcal{R}$ of $\mathbb{S}_n$, the former are in bijection with the individual irreps $\rho \in \mathcal{R}$, while the latter are in bijection with the rows and columns of the actual representation matrices. Exploiting these bijections, we label the irreps of $\mathbb{S}_n$ by the partitions $\lambda \in \Lambda_n$, and we label the individual rows and columns of $\rho_\lambda(\sigma)$ by the standard tableaux of shape λ. The dimensionality of each representation, $d_\lambda \equiv d_{\rho_\lambda}$ is determined by how many SYT there are of shape λ. A useful formula to answer this question is the so-called **hook rule**

$$d_\lambda = \frac{n!}{\prod_i \ell(i)}, \tag{2}$$

where i ranges over all the cells of λ, and $\ell(i)$ are the lengths of the corresponding "hooks", i.e., the number of cells to the right of cell i plus the number of cells below i plus one. For example, it is easy to check that $d_{(1)} = 1$, $d_{(n-1,1)} = n-1$, $d_{(n-2,2)} = n(n-3)/2$, and $d_{(n-2,1,1)} = (n-1)(n-2)/2$.

Young's Orthogonal Representation. The specific system of irreps that we use in this paper is called **Young's orthogonal representation**, or just **YOR**. A special feature of YOR is that its irreps are not only unitary, but also real-valued, hence the $\rho_\lambda(\sigma)$ matrices are orthogonal. YOR is defined by explicitly specifying the representation matrices corresponding to adjacent transpositions. For any standard tableau t, letting $\tau_i(t)$ be the tableau that we get from t by exchanging the numbers i and $i+1$ in its diagram, the rule defining $\rho_\lambda(\tau_i)$ in YOR is the following: if $\tau_i(t)$ is *not* a standard tableau, then the column of $\rho_\lambda(\tau_i)$ indexed by t is zero, except for the diagonal element $[\rho_\lambda(\tau_i)]_{t,t} = 1/d_t(i,i+1)$; if $\tau_i(t)$ is a standard tableau, then in addition to this diagonal element, we also have a single non-zero off-diagonal element, $[\rho_\lambda(\tau_i)]_{\tau_k(t),t} = (1 - 1/d_t(i,i+1)^2)^{1/2}$. All other matrix entries of $\rho_\lambda(\tau_i)$ are zero. In the above $d_t(i,i+1) = c_t(i+1) - c_t(i)$, where $c(j)$ is the column index minus the row index of the cell where j is located in t. Note that the trivial representation $\rho_{\text{triv}}(\sigma) \equiv 1$ is the irrep indexed by $\lambda = (n)$. Also note that for any $\lambda \vdash n$, each row/column of $\rho_\lambda(\tau_i)$ has at most two non-zero entries.

The Fourier transform. The **Fourier transform** of a function $f\colon \mathbb{S}_n \to \mathbb{C}$ is defined as the collection of matrices

$$\widehat{f}(\lambda) = \sum_{\sigma \in \mathbb{S}_n} f(\sigma)\, \rho_\lambda(\sigma) \qquad\qquad \lambda \vdash n. \tag{3}$$

As always, we assume that the ρ_λ representations are given in YOR. Since YOR is a system of real valued representations, if f is a real valued function, then the $\widehat{f}(\lambda)$ Fourier components are real valued matrices. Non-commutative Fourier transforms such as (25) enjoy many of the same properties as the usual Fourier transforms on the real line and the unit circle. In particular (25) is an invertible, unitary mapping $\mathbb{C}^{\mathbb{S}_n} \to \bigoplus_{\lambda \vdash n} \mathbb{C}^{d_\lambda \times d_\lambda}$. The **inverse Fourier transform** is

$$f(\sigma) = \frac{1}{n!} \sum_{\lambda \vdash n} d_\lambda \operatorname{tr}\big[\widehat{f}(\lambda)\, \rho_\lambda(\sigma)^{-1}\big].$$

$\mathbb{S}_n$–modules Given a vector space V, if there is a system of linear operators $\{T_\sigma\}_{\sigma \in \mathbb{S}_n}$ on V, such that $T_{\sigma_2} T_{\sigma_1} = T_{\sigma_2 \sigma_1}$ for all $\sigma_1, \sigma_2 \in \mathbb{S}_n$, then we say that V is an $\mathbb{S}_n$–module. If V is finite dimensional, then, fixing a basis, we immediately see the matrix representations of the T_σ operators form a representation of $\mathbb{S}_n$. Thus, any $\mathbb{S}_n$–module corresponds to a class of equivalent representations $\{\rho_\iota\}_\iota$ of $\mathbb{S}_n$, and V is irreducible if and only if the ρ_ι representations are irreducible.

Translations Letting T_τ be the translation operator mapping $f \mapsto f'$ where $f'(\sigma) := f(\tau^{-1}\sigma)$, it is easy to see that the Fourier transform satisfies the non-commutative translation theorem $\widehat{T_\tau f}(\lambda) = \rho_\lambda(\tau)\widehat{f}(\lambda)$. In particular, under translation, each column of each Fourier matrix is transformed independently, i.e., defining the functions $\psi_{t,t'}(\sigma) := [\rho_\lambda(\sigma)]_{t,t'}$, for fixed $t \in \mathcal{T}_n$, the space $M_t := \operatorname{span}\{\, \psi_{t',t} \mid t' \in \mathcal{T}_\lambda(t) \,\} \subseteq \mathbb{R}^{\mathbb{S}_n}$ will be an irreducible $\mathbb{S}_n$–module. Moreover, $\bigoplus_{t \in \mathcal{T}_n} M_t = \mathbb{R}^{\mathbb{S}_n}$.

Adapted representations In general, if $\rho\colon \mathbb{S}_n \to \mathbb{C}^{d \times d}$ is a representations of $\mathbb{S}_n$, then for $m < n$ we define the **restricted** representation $\rho\!\downarrow_m\colon \mathbb{S}_m \to \mathbb{C}^{d \times d}$ as simply $\rho\!\downarrow_m(\tau) = \rho(\tau)$ for all $\tau \in \mathbb{S}_m$.

In general, even if ρ is irreducible (over $\mathbb{S}_n$), $\rho{\downarrow}_m$ is usually not irreducible (over $\mathbb{S}_m$). However, by theorem of complete reducibility, it can be written as

$$\rho\downarrow_m(\tau) = T\Big[\bigoplus_{\lambda'\in A}\bigoplus_{i=1}^{k'_\lambda}\rho'_\lambda\Big]T^{-1} \qquad \tau\in\mathbb{S}_m \tag{4}$$

for some set $\{\rho'_\lambda\}_{\lambda'\in A}$ of irreps of $\mathbb{S}_m$, corresponding multiplicities $\{k'_\lambda\}_{\lambda'\in A}$, and basis transformation matrix T. In particular, relationships of this type hold for the $\rho=\rho_\lambda$ irreps of $\mathbb{S}_n$. If for a system of irreps of $\{\mathbb{S}_m\}_{m\in\mathbb{N}}$ all these relationships are such that each T matrix is just the identity, then we say that the system of representations is **adapted** to the chain of groups $\mathbb{S}_1<\mathbb{S}_2<\ldots<\mathbb{S}_n<\ldots$. By examining the definition of YOR, it is clear that it is a sytem of representations adapted to this chain. Moreover, due to the fact that for irreducible representations of the symmetric group each of the multiplicities in (26) is 1 (or 0), given the irreps of $\mathbb{S}_m$, the adapted irreps of $\mathbb{S}_{m+1}$ are uniquely determined. Thus (up to the permutation of rows and columns of the representation matrices), YOR is the *unique* adapted system of irreducible representations for $\mathbb{S}_n$. Adapted representations are sometimes also referred to as representations in a Gel'fand-Tsetlin basis.

Adapted modules By the close connection between modules and representations, a system of $\mathbb{S}_m$–modules form an adapted system of modules in the sense of Section 2.1 if and only if the corresponding representations are adapted to the chain $\mathbb{S}_1<\mathbb{S}_2<\ldots$. In particular the *unique* system of adapted $\mathbb{S}_m$ modules for $\mathbb{R}^{\mathbb{S}_n}$ is

$$M_t^{i_1\ldots i_k} := \text{span}\{\,\psi_{t,t'}^{i_1\ldots i_k}\mid t'\in\mathcal{T}_t\,\} \qquad k\in\{1,\ldots,n\} \qquad \{i_1\ldots i_k\}\subseteq\{1,\ldots,n\} \qquad t\in\mathcal{T}_n,$$

where

$$\psi_{t,t'}^{i_1\ldots i_k} := \rho_{\lambda(t)}(\mu_{i_1\ldots i_k}^{-1}\sigma),$$

and ρ_λ is in YOR. These modules correspond exactly to the columns of the Fourier transforms restricted to the various $\mu_{i_1\ldots i_k}\mathbb{S}_{n-k}$ cosets. Clausen's FFT [12] works by building up the Fourier transform from these smaller Fourier transforms.

Appendix B: Figures

$$\begin{tikzcd}[column sep=small]
 & & & \mathbb{S}_3 \arrow[dll] \arrow[d] \arrow[drr] & & & \\
 & \mathcal{S}_3 \arrow[dl] \arrow[dr] & & \mathcal{S}_2 \arrow[dl] \arrow[dr] & & \mathcal{S}_1 \arrow[dl] \arrow[dr] & \\
\mathcal{S}_{3,2}=\{[123]\} & & \mathcal{S}_{3,1}=\{[213]\} \quad \mathcal{S}_{2,3}=\{[132]\} & & \mathcal{S}_{2,1}=\{[312]\} \quad \mathcal{S}_{1,3}=\{[231]\} & & \mathcal{S}_{1,2}=\{[321]\}
\end{tikzcd}$$

Figure 1: The coset tree of $\mathbb{S}_3$. To denote the individual permutations at the leaves of this tree, we used one-line notation, i.e., $[a, b, c]$ is the permutation that maps 1 to a, 2 to b, etc..

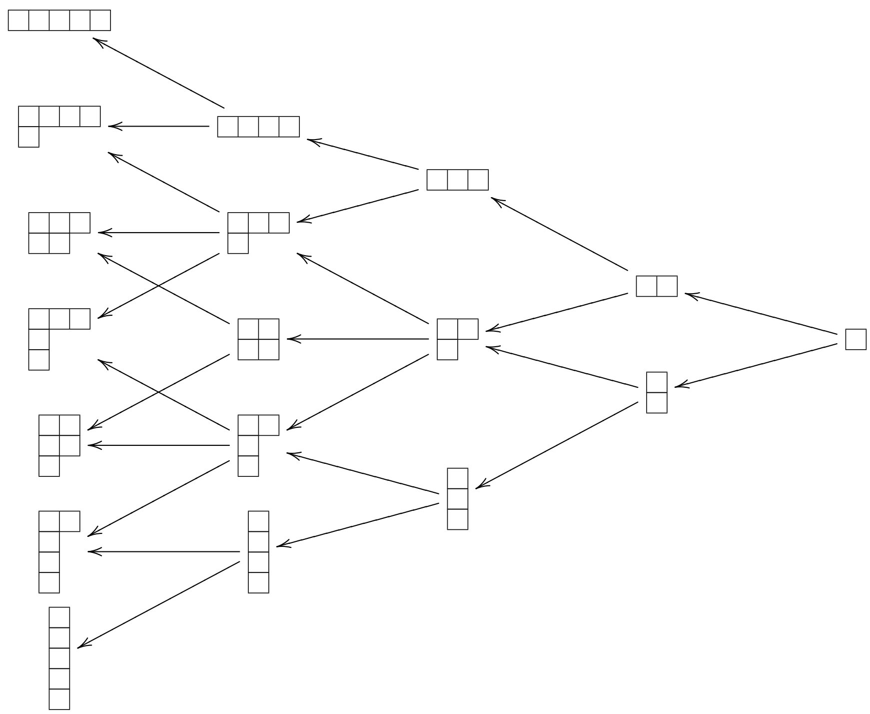

Figure 2: The diagram of the inclusion order $\leq$ between partitions (shown here up to $n = 5$) is also called the Young lattice or Bratelli diagram of $\mathbb{S}_n$. For any λ in the left hand column (corresponding to $k = 0$) of this diagram, $\lambda{\downarrow_{n-k}}$ is the set of its ancestors at level $n-k$, whereas $\lambda{\downarrow_{n-k}}{\uparrow^k}$ is the set of all partitions at level n that have at least one common ancestor with λ at level $n-k$. The $\leq$ order on standard Young tableaux induces a similar diagram, but that one is simpler because it is a tree.

Appendix C: Examples

Example 1 If $n = 5$ and we set $\nu_0 = \{\ytableaushort{12345}\}$, then $\nu_0 \downarrow_4 = \ytableaushort{1234}$, $\nu_0 \downarrow_3 = \ytableaushort{123}$, $\nu_0 \downarrow_2 = \ytableaushort{12}$, and $\nu_0 \downarrow_1 = \ytableaushort{1}$. Therefore, by Proposition 1,

$$\nu_1 = \ytableaushort{1234}\uparrow^5 = \left\{\ytableaushort{12345}, \ytableaushort{1234,5}\right\}$$

$$\nu_2 = \ytableaushort{123}\uparrow^5 = \left\{\ytableaushort{12345}, \ytableaushort{1234,5}, \ytableaushort{1235,4}, \ytableaushort{123,45}, \ytableaushort{123,4,5}\right\}$$

$$\nu_3 = \ytableaushort{12}\uparrow^5 = \left\{\ytableaushort{12345}, \ytableaushort{1234,5}, \ytableaushort{1235,4}, \ytableaushort{1245,3}, \ytableaushort{123,45}, \ytableaushort{124,35}, \ytableaushort{125,34}, \ytableaushort{123,4,5}, \ytableaushort{124,3,5}, \ytableaushort{125,3,4}, \ldots\right\}$$

$$\nu_4 = \ytableaushort{1}\uparrow^5 = \left\{\ytableaushort{12345}, \ytableaushort{1234,5}, \ytableaushort{1235,4}, \ytableaushort{1245,3}, \ytableaushort{1345,2}, \ytableaushort{123,45}, \ytableaushort{124,35}, \ytableaushort{125,34}, \ytableaushort{134,25}, \ytableaushort{135,24}, \ldots\right\}$$

and hence

$$\omega_0 = \left\{\ytableaushort{1234,5}\right\}$$

$$\omega_1 = \left\{\ytableaushort{1235,4}, \ytableaushort{123,45}, \ytableaushort{123,4,5}\right\}$$

$$\omega_2 = \left\{\ytableaushort{1245,3}, \ytableaushort{124,35}, \ytableaushort{125,34}, \ytableaushort{124,3,5}, \ytableaushort{125,3,4}, \ytableaushort{12,34,5}, \ytableaushort{12,35,4}, \ytableaushort{12,3,4,5}\right\}$$

$$\omega_3 = \left\{\ytableaushort{1345,2}, \ytableaushort{135,24}, \ytableaushort{134,2,5}, \ytableaushort{135,2,4}, \ytableaushort{145,2,3}, \ytableaushort{13,24,5}, \ytableaushort{13,25,4}, \ytableaushort{14,25,3}, \ytableaushort{13,2,4,5}, \ytableaushort{14,2,3,5}, \ytableaushort{15,2,3,4}, \ytableaushort{1,2,3,4,5}\right\}.$$

As explained in Section 3, this MRA is analogous to Haar wavelets on the real line. $\lrcorner$

Example 2 If $n = 5$ and $\nu_0 = \left\{\ytableaushort{135,24}\right\}$, then $\nu_0 \downarrow_4 = \left\{\ytableaushort{13,24}\right\}$, $\nu_0 \downarrow_3 = \left\{\ytableaushort{13,2}\right\}$, $\nu_0 \downarrow_2 = \left\{\ytableaushort{1,2}\right\}$, and $\nu_0 \downarrow_4 = \{\ytableaushort{1}\}$, therefore,

$$\omega_0 = \left\{\ytableaushort{13,24,5}\right\}$$

$$\omega_1 = \left\{\ytableaushort{1345,2}, \ytableaushort{134,25}, \ytableaushort{134,2,5}, \ytableaushort{135,2,4}, \ytableaushort{13,25,4}, \ytableaushort{13,2,4,5}\right\}$$

$$\omega_2 = \left\{\ytableaushort{14,25,3}, \ytableaushort{14,2,3,5}, \ytableaushort{15,2,3,4}, \ytableaushort{1,2,3,4,5}\right\}$$

$$\omega_3 = \left\{\ytableaushort{12345}, \ytableaushort{1234,5}, \ytableaushort{1235,4}, \ytableaushort{1254,3}, \ytableaushort{123,45}, \ytableaushort{124,35}, \ytableaushort{125,34}, \ytableaushort{123,4,5}, \ytableaushort{125,3,4}, \ytableaushort{12,34,5}, \ytableaushort{12,35,4}, \ytableaushort{12,3,4,5}\right\}.$$

$\lrcorner$

Example 3 If $n=5$ and we set $\overline{\nu}_0=\{(5)\}=\{\ydiagram{5}\}$, then $\overline{\nu}_1=\{\ydiagram{4}\}$, $\overline{\nu}_2=\{\ydiagram{3}\}$, $\overline{\nu}_3=\{\ydiagram{2}\}$, and $\overline{\nu}_4=\{\ydiagram{1}\}$. Therefore, by Proposition 2,

$$
\begin{aligned}
\overline{\nu}_1 &= \{\ydiagram{4}\}\!\uparrow^5= \left\{\ydiagram{5},\ydiagram{4,1}\right\}\\
\overline{\nu}_2 &= \{\ydiagram{3}\}\!\uparrow^5= \left\{\ydiagram{5},\ydiagram{4,1},\ydiagram{3,2},\ydiagram{3,1,1}\right\}\\
\overline{\nu}_3 &= \{\ydiagram{2}\}\!\uparrow^5= \left\{\ydiagram{5},\ydiagram{4,1},\ydiagram{3,2},\ydiagram{3,1,1},\ydiagram{2,2,1},\ydiagram{2,1,1,1}\right\}\\
\overline{\nu}_4 &= \{\ydiagram{1}\}\!\uparrow^5= \left\{\ydiagram{5},\ydiagram{4,1},\ydiagram{3,2},\ydiagram{3,1,1},\ydiagram{2,2,1},\ydiagram{2,1,1,1},\ydiagram{1,1,1,1,1}\right\}.
\end{aligned}
$$

and hence

$$
\overline{\omega}_0=\left\{\ydiagram{4,1}\right\},\quad \overline{\omega}_1=\left\{\ydiagram{3,2},\ydiagram{3,1,1}\right\},\quad \overline{\omega}_2=\left\{\ydiagram{2,2,1},\ydiagram{2,1,1,1}\right\},\quad \overline{\omega}_3=\left\{\ydiagram{1,1,1,1,1}\right\}.
$$

┘

Example 4 If $n=5$ and we set $\overline{\nu}_0=\left\{\ydiagram{5},\ydiagram{4,1}\right\}$, then $\overline{\nu}_0\!\downarrow_4=\left\{\ydiagram{4},\ydiagram{3,1}\right\}$, $\overline{\nu}_0\!\downarrow_3=\left\{\ydiagram{3},\ydiagram{2,1}\right\}$, $\overline{\nu}_0\!\downarrow_2=\left\{\ydiagram{2},\ydiagram{1,1}\right\}$, and $\overline{\nu}_0\!\downarrow_1=\{\ydiagram{1}\}$, leading to

$$
\begin{aligned}
\overline{\nu}_1 &= \left\{\ydiagram{4},\ydiagram{3,1}\right\}\!\uparrow^5= \left\{\ydiagram{5},\ydiagram{4,1},\ydiagram{3,2},\ydiagram{3,1,1}\right\}\\
\overline{\nu}_2 &= \left\{\ydiagram{3},\ydiagram{2,1}\right\}\!\uparrow^5= \left\{\ydiagram{5},\ydiagram{4,1},\ydiagram{3,2},\ydiagram{3,1,1},\ydiagram{2,2,1},\ydiagram{2,1,1,1}\right\}\\
\overline{\nu}_3 = \overline{\nu}_4 &= \left\{\ydiagram{2},\ydiagram{1,1}\right\}\!\uparrow^5= \left\{\ydiagram{5},\ydiagram{4,1},\ydiagram{3,2},\ydiagram{3,1,1},\ydiagram{2,2,1},\ydiagram{2,1,1,1}\right\}
\end{aligned}
$$

and hence

$$
\overline{\omega}_0=\left\{\ydiagram{3,2},\ydiagram{3,1,1}\right\},\quad \overline{\omega}_1=\left\{\ydiagram{2,2,1},\ydiagram{2,1,1,1}\right\},\quad \overline{\omega}_2=\left\{\ydiagram{1,1,1,1,1}\right\},\quad \overline{\omega}_3=\emptyset.
$$

┘

Example 5 For any n if we set $\nu_0=\{(n),(n-1,1)\}$, then $\nu_0\!\downarrow_{n-k}=\{(n-k),(n-k-1,1)\}$, and hence

$$
\begin{aligned}
\overline{\nu}_1 &= \{(n),(n-1,1),(n-2,2),(n-2,1,1)\},\\
\overline{\nu}_2 &= \{(n),(n-1,1),(n-2,2),(n-2,1,1),(n-3,3),(n-3,2,1),(n-3,1,1,1)\},
\end{aligned}
$$

and so on, leading to

$$
\begin{aligned}
\overline{\omega}_0 &= \{(n-2,2),(n-2,1,1)\}\\
\overline{\omega}_1 &= \{(n-3,3),(n-3,2,1),(n-3,1,1,1)\}\\
\overline{\omega}_2 &= \{(n-4,4),(n-4,3,1),(n-4,2,2),(n-4,2,1,1),(n-4,1,1,1,1)\},
\end{aligned}
$$

etc..

┘

Appendix D: Proofs and further results

Lemma 1 *If $V \subseteq \mathbb{R}^{\mathbb{S}_n}$ is a left $\mathbb{S}_n$–invariant space and $M \subseteq \mathbb{R}^{\mathbb{S}_n}$ is an irreducible $\mathbb{S}_{n-k}$–module with respect to the "internal" translation defined in (4) for some $i_1, \ldots, i_k$, then $M \cap V \neq \{0\}$ implies that $M \subseteq V$.*

Proof. Let f be any non-zero function in $M \cap V$ and consider $U := \mathrm{span}\{\, T_\tau^{i_1 \cdots i_k} f \mid \tau \in \mathbb{S}_{n-k} \,\}$. Clearly, U is an $\mathbb{S}_{n-k}$–module, so, since M is irreducible, we must have $U = M$. On the other hand, since any internal translation $T_\tau^{i_1 \cdots i_k}$ is equivalent to the global translation $T_{\mu_{i_1 \ldots i_k} \tau \mu_{i_1 \ldots i_k}}$, and V is left $\mathbb{S}_n$–invariant, each $T_\tau^{i_1 \cdots i_k} f$ must be in V, therefore $U \subseteq V$. ■

Lemma 2 *If $V \subseteq \mathbb{R}^{\mathbb{S}_n}$ is a left $\mathbb{S}_n$–invariant space and $\{M^{i_1 \cdots i_k}\}$ is a collection of $\mathbb{S}_{n-k}$–modules as in Section 2.1. Then if any one of the $\{M^{i_1 \cdots i_k}\}$ modules are subspaces of V, then they all are.*

Proof. Since, by definition, $M^{i'_1 \cdots i'_k} = T_{\mu_{i'_1 \ldots i'_k} \mu^{-1}_{i_1 \ldots i_k}} M^{i_1 \cdots i_k}$, and V is translation invariant, if $M^{i_1 \cdots i_k} \subseteq V$, then $M^{i'_1 \cdots i'_k} \subseteq V$ for any $\{i'_1, \ldots, i'_k\} \subseteq \{1, \ldots, n\}$. ■

Proof of Proposition 2. Since V_0 is both left- and right-invariant, it must be a sum of isotypics, so for some $\overline{\nu}_0 \subseteq \Lambda_n$, $V_0 = \bigoplus_{\lambda \in \overline{\nu}_0} U_\lambda$. Thus, by the same line of reasoning as in the proof of Proposition 1,

$$Q_k := \bigoplus_{t \in \xi_k} M_t \subseteq V_k \qquad \text{where} \qquad \xi_k = \{\, t' \in \mathcal{T}_\lambda \mid \lambda \in \nu_0 \,\} \downarrow_{n-k} \uparrow^n$$

It is easy to see that

$$\{\, t' \in \mathcal{T}_{\lambda'} \mid \lambda' \in \nu_0 \,\} \downarrow_{n-k} \uparrow^n = \{\, t \in \mathcal{T}_\lambda \mid \lambda \in \nu_0 \downarrow_{n-k} \uparrow^n \,\},$$

therefore,

$$Q_k = \bigoplus_{\lambda \in \nu_0 \downarrow_{n-k} \uparrow^n} \bigoplus_{t \in \mathcal{T}_\lambda} M_t = \bigoplus_{\lambda \in \nu_0 \downarrow_{n-k} \uparrow^n} U_\lambda.$$

However, this space being a sum of isotypics, it also satisfies the right-invariance condition of axiom Bi1, so, in fact, $Q_k = V_k$ and $\nu_k = \xi_k$. ■

Proof of Proposition 3. As customary in algebraic complexity theory, by a "scalar operation" we mean an operation of the type $y \leftarrow ax + b$, where $a, x, b \in \mathbb{C}$ are either constants or variables. Copying and moving data, as well as the bookkeeping operations involved in constructing Young tabeaux, controlling loops, etc. are assumed to be free. These details only make a difference of a constant prefactor, anyhow.

Thanks to the sparsity of YOR, if τ_j is the adjacent transposition $(j, j+1)$, multiplying a d_λ dimensional column vector $\boldsymbol{v}$ by $\rho_\lambda(\tau_j)$ takes only $2d_\lambda$ time. The contiguous cycle appearing in (22) can be written as $[\![i_{k+1}, n-k]\!] = \tau_{i_{k+1}} \tau_{i_{k+1}+1} \ldots \tau_{n-k-1}$, therefore,

$$\rho_\lambda([\![i_{k+1}, n-k]\!])\, \boldsymbol{v} = \rho_\lambda(\tau_{i_{k+1}}) \ldots \rho_\lambda(\tau_{n-k-1})\, \boldsymbol{v},$$

and thus, the entire operation in (22) can be performed in $2(n-k-i_{k+1})d_\lambda \leq 2(n-k-1)d_\lambda$ time. At a given level k an operation of this type must be performed for each non-zero branch of the coset tree at level $k+1$ (of which there are at most q) and for each

$$\lambda \in (\nu_0 \downarrow_{n-k}) \cup (\nu_0 \downarrow_{n-k-1} \uparrow^{n-k} \setminus \nu_0 \downarrow_{n-k}) = \nu_0 \downarrow_{n-k-1} \uparrow^{n-k}.$$

It is easy to see that

$$\sum_{\lambda \in \nu_0 \downarrow_{n-k-1} \uparrow^{n-k}} d_\lambda \;\leq \sum_{\lambda \in \nu_0 \downarrow_{n-1} \uparrow^n = \nu_1} d_\lambda,$$

therefore, setting $N = \sum_{t \in \nu_1} d_{\lambda(t)}$, the total complexity is bounded by

$$\sum_{k=0}^{n-2} 2qN(n-k-1) = n(n-1)pN < n^2 qN.$$

The $n^2 qM$ bound for Bi-CMRA is proven analogously. ■

Non-adapted modules

In the L-CMRA case if we do *not* assume that V_0 is a sum of adapted modules, the analysis becomes challening. While in this case there is no direct analog of Proposition 1, we can prove some weaker statements. Note that Bi-CMRA does not need a similar assumption, since by bi-invariance, V_0 must be a sum of isotypics, which can always be written as a sum of adapted modules.

In general, for two sets of (non-zero) adapted or non-adapted $\mathbb{S}_n$– resp. $\mathbb{S}_{n-k}$–modules $\{M_\alpha\}$ and $\{M^{i_1\cdots i_k}\}$ if

$$\bigoplus_{\alpha\in A} M_\alpha = \bigoplus_{i_1\ldots i_k} M^{i_1\cdots i_k},$$

then we write $\{M_\alpha\}_{\alpha\in A} = \{M^{i_1\cdots i_k}\}_{i_1\ldots i_k}\!\uparrow^n$ and say that $\{M_\alpha\}$ is induced by $\{M^{i_1\cdots i_k}\}$.

Proposition 1 *If $\{M_\alpha\}_{\alpha\in A} = \{M^{i_1\cdots i_k}\}_{i_1\ldots i_k}\!\uparrow^n$, and $M_\alpha \subseteq V_0$ for some $\alpha \in A$, then by axioms L1 and L2, $M^{i_1\cdots i_k} \subseteq V_k$ for all $\{i_1,\ldots,i_k\} \subseteq \{1,\ldots,n\}$.*

Proof. By the definition of induced modules

$$P_{i_1\ldots i_k} \bigoplus_{\alpha\in A} M_\alpha = M^{i_1\cdots i_k}.$$

Since $M_\alpha \neq \{0\}$, there must be at least some setting of the indices $i_1,\ldots,i_k$, for which $P_{i_1\ldots i_k} f \in M^{i_1\cdots i_k}$ for some non-zero $f \in M_\alpha$. Thus, $M^{i_1\cdots i_k} \cap V_k \neq \{0\}$. Since, by axiom L1, V_k is an $\mathbb{S}_n$–invariant space, Lemma 1 then implies that $M^{i_1\cdots i_k} \subseteq V_k$. Then, by Lemma 2, $M^{i'_1\cdots i'_k} \subseteq V_k$ for any $i'_1,\ldots,i'_k$. ■

Clearly, if none of the M_α modules are in V_0, then none of the $M^{i_1\cdots i_k}$ modules can be in V_k, either. However, it is not immediately clear that having a *single* $M_\alpha \subseteq V_0$ is sufficient to satisfy axiom LP3. The following result addresses this concern.

Proposition 2 *If $\{M_\alpha\}_{\alpha\in A} = \{M^{i_1\cdots i_k}\}_{i_1\ldots i_k}\!\uparrow^n$ and we fix some $\alpha \in A$, then for any $i_1,\ldots,i_k$ and any $g \in M^{i_1\cdots i_k}$ there is an $f \in M_\alpha$ such that $P_{i_1\ldots i_k} f = g$.*

Proof. As in the proof of Proposition 1, there must be some non-zero $f \in M_\alpha$ such that $g := P_{i_1\ldots i_k} f \in M^{i_1\cdots i_k}$ for some $i_1,\ldots,i_k$. Now if we define

$$U := \{\, T^{i_1\cdots i_k}_\tau g \mid \tau \in \mathbb{S}_{n-k} \,\} = \{\, P_{i_1\ldots i_k} T^{i_1\cdots i_k}_\tau f \mid \tau \in \mathbb{S}_{n-k} \,\},$$

then by the translation invariance of M_α, $T^{i_1\cdots i_k}_\tau f \in M_\alpha$ for any $\tau \in \mathbb{S}_{n-k}$, and therefore, $U \subseteq P_{i_1\ldots i_k} M_\alpha$. On the other hand, by the irreducibility of $M^{i_1\cdots i_k}$, we must have $U = M^{i_1\cdots i_k}$, just as in the proof of Lemma 1. Therefore, $P_{i_1\ldots i_k} M_\alpha = M^{i_1\cdots i_k}$.

To prove the statement for general $\{i_1,\ldots,i_k\} \subseteq 1,\ldots,n$ it is sufficient to consider that, by the transation invariance of M_α, each step of the above argument also holds for $f' := T_{\mu_{i'_1\ldots i'_k}\mu^{-1}_{i_1\ldots i_k}} f$. ■

Finally, we have the following result about which $\mathbb{S}_n$–modules are induced in V_k

Proposition 3 *If $\{M_\beta\}_{\beta\in B} = \{M^{i_1\cdots i_k}\}_{i_1\ldots i_k}\!\uparrow^n$ and for some $\{i_1\ldots i_k\} \subseteq \{1,\ldots,n\}$ we have $M^{i_1\cdots i_k} \subseteq V_k$, then $M_\beta \subseteq V_0$ for all $\beta \in B$.*

Proof. As in the proof of Propositions 1 and 2, since $M_\beta \neq \{0\}$, there must be some non-zero $f \in M_\beta$ such that $f \in M^{i'_1\cdots i'_k}$ for some $i'_1,\ldots,i'_k$. By Lemma 2, $M^{i'_1\cdots i'_k} \subseteq V_k$. Therefore, $f \in M_\beta \cap V_k$. Defining $U := \operatorname{span}\{\, T_\tau f \mid \tau \in \mathbb{S}_n \,\} \subseteq M_\beta$, since M_β is an $\mathbb{S}_n$–module, $U \subseteq M_\beta$, and since M_β is irreducible, this can only happen if $M_\beta = U$. On the other hand, V_k is also $\mathbb{S}_n$–invariant, so $U \subseteq V_k$. ■

Appendix E: Algorithms

function FastLCWT($n, \mu, \nu, flist$) {
$\omega \leftarrow \nu\downarrow_{n-1}\uparrow^{n} \setminus \nu$
$\{v_t \leftarrow 0_{d_{\lambda(t)}}\}_{t\in\nu}$ // Local Fourier coefficients
$\{w_t \leftarrow 0_{d_{\lambda(t)}}\}_{t\in\omega}$ // Local wavelet coefficients
$\{subflist_i \leftarrow \emptyset\}_{i=1}^{n}$ // List of function values to pass to each coset
$wlist \leftarrow \emptyset$ // List of computed wavelet coefficients

if $n = 1$ **then**
  $\{(\sigma, y)\} \leftarrow flist$
  $v_{t=\boxed{1}} = (y)$
  return $(\{v_t\}_{t\in\nu}, \emptyset)$
end if

for each $(\sigma, y) \in flist$ **do**
  $i \leftarrow \sigma(n)$
  $subflist_i \leftarrow subflist_i \cup \{(\llbracket i, n\rrbracket^{-1}\sigma, y)\}$
end for

for $i = 1$ **TO** n **do**
  if $subflist_i \neq \emptyset$ **then**
    $(\{v'_t\}_t, subwlist) \leftarrow$ FastLCWT$(n-1, \mu\llbracket i, n\rrbracket, \nu\downarrow_{n-1}, subflist_i)$
    for each $t \in \nu$ **do**
      $v_t \leftarrow v_t + \rho_{\lambda(t)}(\llbracket i, n\rrbracket)(v'_{t\downarrow_{n-1}}\uparrow^{n})$
    end for
    for each $t \in \omega$ **do**
      $w_t \leftarrow w_t + \rho_{\lambda(t)}(\llbracket i, n\rrbracket)(v'_{t\downarrow_{n-1}}\uparrow^{n})$
    end for
    $wlist \leftarrow wlist \cup subwlist$
  end if
end for

for each $t \in \omega$ **do**
  $wlist \leftarrow wlist \cup \{(\mu, t, w_t)\}$
end for

return $(\{v_t\}_{t\in\nu}, wlist)$
}

Algorithm 1: A more detailed description of the recursive procedure of Algorithm 1 to compute the L-CMRA transform of a sparse function f. When applied to the coset $\mu_{i_1\ldots i_k}\mathbb{S}_{n-k}$, this function is called with the arguments $m = n-k$, $\mu = \mu_{i_1\ldots i_k}$, $\nu = \nu_k$ and $flist$, which is a list of (σ, y) pairs, corresponding to each of the non-zero entries of f falling within $\mu_{i_1\ldots i_k}\mathbb{S}_{n-k}$. Thus, on the initial call, $m = n$, $\nu = e$, ,$\nu = \nu_0$, and $flist$ is a sparse representation of the entire function f. The function returns the scaling space coefficients $\{v_t\}_t\nu$, and the list $wlist$ of $(\mu_{i_1\ldots i_k}, t, w_t)$ triples, each corresponding to one of the $w_f(t; i_1 \ldots i_k)$ wavelet coefficients of (19).