Maximal Overlap Discrete Wavelet Transform

- abbreviation is MODWT (pronounced 'mod WT')
- transforms very similar to the MODWT have been studied in the literature under the following names:
  - undecimated DWT (or nondecimated DWT)
  - stationary DWT
  - translation invariant DWT
  - time invariant DWT
  - redundant DWT
- also related to notions of 'wavelet frames' and 'cycle spinning'
- basic idea: use values removed from DWT by downsampling

Quick Comparison of the MODWT to the DWT

- unlike the DWT, MODWT is not orthonormal (in fact MODWT is highly redundant)
- unlike the DWT, MODWT is defined naturally for all samples sizes (i.e., N need not be a multiple of a power of two)
- similar to the DWT, can form multiresolution analyses (MRAs) using MODWT, but with certain additional desirable features; e.g., unlike the DWT, MODWT-based MRA has details and smooths that shift along with $\mathbf{X}$ (if $\mathbf{X}$ has detail $\widetilde{\mathcal{D}}_j$, then $\mathcal{T}^m\mathbf{X}$ has detail $\mathcal{T}^m\widetilde{\mathcal{D}}_j$)
- similar to the DWT, an analysis of variance (ANOVA) can be based on MODWT wavelet coefficients
- unlike the DWT, MODWT discrete wavelet power spectrum same for $\mathbf{X}$ and its circular shifts $\mathcal{T}^m\mathbf{X}$

Definition of MODWT Wavelet & Scaling Filters: I

- recall that we can obtain DWT wavelet and scaling coefficients directly from $\mathbf{X}$ by filtering and downsampling:

$$\mathbf{X} \longrightarrow \boxed{H_j(\tfrac{k}{N})} \underset{\downarrow 2^j}{\longrightarrow} \mathbf{W}_j \text{ and } \mathbf{X} \longrightarrow \boxed{G_j(\tfrac{k}{N})} \underset{\downarrow 2^j}{\longrightarrow} \mathbf{V}_j$$

- transfer functions $H_j(\cdot)$ and $G_j(\cdot)$ are associated with impluse response sequences $\{h_{j,l}\}$ and $\{g_{j,l}\}$ via the usual relationships

$$\{h_{j,l}\} \longleftrightarrow H_j(\cdot) \text{ and } \{g_{j,l}\} \longleftrightarrow G_j(\cdot),$$

  and both filters have width $L_j = (2^j - 1)(L - 1) + 1$

- define MODWT filters $\{\tilde{h}_{j,l}\}$ and $\{\tilde{g}_{j,l}\}$ by renormalizing the DWT filters:

$$\tilde{h}_{j,l} = h_{j,l}/2^{j/2} \text{ and } \tilde{g}_{j,l} = g_{j,l}/2^{j/2}$$

Definition of MODWT Wavelet & Scaling Filters: II

- widths L_j of MODWT and DWT filters are the same
- whereas DWT filters have unit energy, MODWT filters satisfy

$$\sum_{l=0}^{L_j-1} \tilde{h}_{j,l}^2 = \sum_{l=0}^{L_j-1} \tilde{g}_{j,l}^2 = \frac{1}{2^j}$$

- let $\widetilde{H}_j(\cdot)$ and $\widetilde{G}_j(\cdot)$ be the corresponding transfer functions:

$$\widetilde{H}_j(f) = \frac{1}{2^{j/2}} H_j(f) \text{ and } \widetilde{G}_j(f) = \frac{1}{2^{j/2}} G_j(f)$$

so that

$$\{\tilde{h}_{j,l}\} \longleftrightarrow \widetilde{H}_j(\cdot) \text{ and } \{\tilde{g}_{j,l}\} \longleftrightarrow \widetilde{G}_j(\cdot)$$

Definition of MODWT Coefficients: I

- level j MODWT wavelet and scaling coefficients are *defined* to be output obtaining by filtering $\mathbf{X}$ with $\{\tilde{h}_{j,l}\}$ and $\{\tilde{g}_{j,l}\}$:

$$\mathbf{X} \longrightarrow \boxed{\widetilde{H}_j(\tfrac{k}{N})} \longrightarrow \widetilde{\mathbf{W}}_j \text{ and } \mathbf{X} \longrightarrow \boxed{\widetilde{G}_j(\tfrac{k}{N})} \longrightarrow \widetilde{\mathbf{V}}_j$$

- compare the above to its DWT equivalent:

$$\mathbf{X} \longrightarrow \boxed{H_j(\tfrac{k}{N})} \underset{\downarrow 2^j}{\longrightarrow} \mathbf{W}_j \text{ and } \mathbf{X} \longrightarrow \boxed{G_j(\tfrac{k}{N})} \underset{\downarrow 2^j}{\longrightarrow} \mathbf{V}_j$$

- DWT and MODWT have different normalizations for filters, and there is no downsampling by 2^j in the MODWT
- level J_0 MODWT consists of $J_0 + 1$ vectors, namely,

$$\widetilde{\mathbf{W}}_1, \widetilde{\mathbf{W}}_2, \ldots, \widetilde{\mathbf{W}}_{J_0} \text{ and } \widetilde{\mathbf{V}}_{J_0},$$

each of which has length N

Definition of MODWT Coefficients: II

- MODWT of level J_0 has $(J_0+1)N$ coefficients, whereas DWT has N coefficients for any given J_0
- whereas DWT of level J_0 requires N to be integer multiple of 2^{J_0}, MODWT of level J_0 is well-defined for *any* sample size N
- when N is divisible by 2^{J_0}, we can write

$$W_{j,t} = \sum_{l=0}^{L_j-1} h_{j,l} X_{2^j(t+1)-1-l \bmod N} \ \& \ \widetilde{W}_{j,t} = \sum_{l=0}^{L_j-1} \tilde{h}_{j,l} X_{t-l \bmod N},$$

and we have the relationship

$$W_{j,t} = 2^{j/2}\widetilde{W}_{j,2^j(t+1)-1} \ \&\text{, likewise, } \ V_{J_0,t} = 2^{J_0/2}\widetilde{V}_{J_0,2^{J_0}(t+1)-1}$$

(here $\widetilde{W}_{j,t}$ & $\widetilde{V}_{J_0,t}$ denote the tth elements of $\widetilde{\mathbf{W}}_j$ & $\widetilde{\mathbf{V}}_{J_0}$)

Properties of the MODWT

- as was true with the DWT, we can use the MODWT to obtain
  - a scale-based additive decomposition (MRA) and
  - a scale-based energy decomposition (ANOVA)
- in addition, the MODWT can be computed efficiently via a pyramid algorithm

MODWT Multiresolution Analysis: I

- starting from the definition

$$\widetilde{W}_{j,t} = \sum_{l=0}^{L_j-1} \tilde{h}_{j,l} X_{t-l \bmod N}, \text{ have } \widetilde{W}_{j,t} = \sum_{l=0}^{N-1} \tilde{h}^{\circ}_{j,l} X_{t-l \bmod N},$$

  where $\{\tilde{h}^{\circ}_{j,l}\}$ is $\{\tilde{h}_{j,l}\}$ periodized to length N

- can express the above in matrix notation as $\widetilde{\mathbf{W}}_j = \widetilde{\mathcal{W}}_j \mathbf{X}$, where $\widetilde{\mathcal{W}}_j$ is the $N \times N$ matrix given by

$$\begin{bmatrix} \tilde{h}^{\circ}_{j,0} & \tilde{h}^{\circ}_{j,N-1} & \tilde{h}^{\circ}_{j,N-2} & \tilde{h}^{\circ}_{j,N-3} & \cdots & \tilde{h}^{\circ}_{j,3} & \tilde{h}^{\circ}_{j,2} & \tilde{h}^{\circ}_{j,1} \\ \tilde{h}^{\circ}_{j,1} & \tilde{h}^{\circ}_{j,0} & \tilde{h}^{\circ}_{j,N-1} & \tilde{h}^{\circ}_{j,N-2} & \cdots & \tilde{h}^{\circ}_{j,4} & \tilde{h}^{\circ}_{j,3} & \tilde{h}^{\circ}_{j,2} \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ \tilde{h}^{\circ}_{j,N-2} & \tilde{h}^{\circ}_{j,N-3} & \tilde{h}^{\circ}_{j,N-4} & \tilde{h}^{\circ}_{j,N-5} & \cdots & \tilde{h}^{\circ}_{j,1} & \tilde{h}^{\circ}_{j,0} & \tilde{h}^{\circ}_{j,N-1} \\ \tilde{h}^{\circ}_{j,N-1} & \tilde{h}^{\circ}_{j,N-2} & \tilde{h}^{\circ}_{j,N-3} & \tilde{h}^{\circ}_{j,N-4} & \cdots & \tilde{h}^{\circ}_{j,2} & \tilde{h}^{\circ}_{j,1} & \tilde{h}^{\circ}_{j,0} \end{bmatrix}$$

MODWT Multiresolution Analysis: II

- recalling the DWT relationship $\mathcal{D}_j = \mathcal{W}_j^T \mathbf{W}_j$, define jth level MODWT detail as $\widetilde{\mathcal{D}}_j = \widetilde{\mathcal{W}}_j^T \widetilde{\mathbf{W}}_j$
- similar development leads to definition for jth level MODWT smooth as $\widetilde{\mathcal{S}}_j = \widetilde{\mathcal{V}}_j^T \widetilde{\mathbf{V}}_j$
- will now show that level J_0 MODWT-based MRA is given by

$$\mathbf{X} = \sum_{j=1}^{J_0} \widetilde{\mathcal{D}}_j + \widetilde{\mathcal{S}}_{J_0},$$

  which is analogous to the DWT-based MRA

MODWT Multiresolution Analysis: III

- since $\widetilde{\mathcal{D}}_j = \widetilde{\mathcal{W}}_j^T \widetilde{\mathbf{W}}_j$, let's look at $\widetilde{\mathcal{W}}_j^T$:

$$\begin{bmatrix} \tilde{h}^\circ_{j,0} & \tilde{h}^\circ_{j,1} & \tilde{h}^\circ_{j,2} & \tilde{h}^\circ_{j,3} & \cdots & \tilde{h}^\circ_{j,N-3} & \tilde{h}^\circ_{j,N-2} & \tilde{h}^\circ_{j,N-1} \\ \tilde{h}^\circ_{j,N-1} & \tilde{h}^\circ_{j,0} & \tilde{h}^\circ_{j,1} & \tilde{h}^\circ_{j,2} & \cdots & \tilde{h}^\circ_{j,N-4} & \tilde{h}^\circ_{j,N-3} & \tilde{h}^\circ_{j,N-2} \\ \tilde{h}^\circ_{j,N-2} & \tilde{h}^\circ_{j,N-1} & \tilde{h}^\circ_{j,0} & \tilde{h}^\circ_{j,1} & \cdots & \tilde{h}^\circ_{j,N-5} & \tilde{h}^\circ_{j,N-4} & \tilde{h}^\circ_{j,N-3} \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ \tilde{h}^\circ_{j,2} & \tilde{h}^\circ_{j,3} & \tilde{h}^\circ_{j,4} & \tilde{h}^\circ_{j,5} & \cdots & \tilde{h}^\circ_{j,N-1} & \tilde{h}^\circ_{j,0} & \tilde{h}^\circ_{j,1} \\ \tilde{h}^\circ_{j,1} & \tilde{h}^\circ_{j,2} & \tilde{h}^\circ_{j,3} & \tilde{h}^\circ_{j,4} & \cdots & \tilde{h}^\circ_{j,N-2} & \tilde{h}^\circ_{j,N-1} & \tilde{h}^\circ_{j,0} \end{bmatrix}$$

- since $\widetilde{\mathcal{V}}_j^T$ has a similar pattern, elements of $\widetilde{\mathcal{D}}_j$ & $\widetilde{\mathcal{S}}_j$ are thus

$$\widetilde{\mathcal{D}}_{j,t} = \sum_{l=0}^{N-1} \tilde{h}^\circ_{j,l} \widetilde{W}_{j,t+l \bmod N} \quad \& \quad \widetilde{\mathcal{S}}_{j,t} = \sum_{l=0}^{N-1} \tilde{g}^\circ_{j,l} \widetilde{V}_{j,t+l \bmod N}$$

MODWT Multiresolution Analysis: IV

- $\widetilde{\mathcal{D}}_j$ and $\widetilde{\mathcal{S}}_j$ both formed by cyclic cross-correlations, and hence
  - $\widetilde{\mathcal{D}}_j$ formed by filtering $\{\widetilde{W}_{j,t}\}$ with $\{\widetilde{H}_j^*(\frac{k}{N})\}$
  - $\widetilde{\mathcal{S}}_j$ formed by filtering $\{\widetilde{V}_{j,t}\}$ with $\{\widetilde{G}_j^*(\frac{k}{N})\}$
- in turn, $\{\widetilde{W}_{j,t}\}$ & $\{\widetilde{V}_{j,t}\}$ formed by filtering $\{X_t\} \longleftrightarrow \{\mathcal{X}_k\}$ with $\{\tilde{h}_{j,l}^{\circ}\} \longleftrightarrow \{\widetilde{H}_j(\frac{k}{N})\}$ & $\{\tilde{g}_{j,l}^{\circ}\} \longleftrightarrow \{\widetilde{G}_j(\frac{k}{N})\}$
- hence

$$\{\widetilde{\mathcal{D}}_{j,t}\} \longleftrightarrow \{\widetilde{H}_j^*(\tfrac{k}{N})\widetilde{H}_j(\tfrac{k}{N})\mathcal{X}_k\} = \{|\widetilde{H}_j(\tfrac{k}{N})|^2\mathcal{X}_k\}$$
$$\{\widetilde{\mathcal{S}}_{j,t}\} \longleftrightarrow \{\widetilde{G}_j^*(\tfrac{k}{N})\widetilde{G}_j(\tfrac{k}{N})\mathcal{X}_k\} = \{|\widetilde{G}_j(\tfrac{k}{N})|^2\mathcal{X}_k\}$$

MODWT Multiresolution Analysis: V

- since the DFT of a sum is the sum of the individual DFTs,

$$\{\widetilde{\mathcal{D}}_{j,t} + \widetilde{\mathcal{S}}_{j,t}\} \longleftrightarrow \{\left(|\widetilde{H}_j(\tfrac{k}{N})|^2 + |\widetilde{G}_j(\tfrac{k}{N})|^2\right) \mathcal{X}_k\}$$

- when $j \geq 2$, can reduce term in parentheses:

$$\begin{aligned}
|\widetilde{H}_j(\tfrac{k}{N})|^2 + |\widetilde{G}_j(\tfrac{k}{N})|^2 &= |\widetilde{H}(2^{j-1}\tfrac{k}{N})|^2 \prod_{l=0}^{j-2} |\widetilde{G}(2^l\tfrac{k}{N})|^2 + \prod_{l=0}^{j-1} |\widetilde{G}(2^l\tfrac{k}{N})|^2 \\
&= \left(|\widetilde{H}(2^{j-1}\tfrac{k}{N})|^2 + |\widetilde{G}(2^{j-1}\tfrac{k}{N})|^2\right) \prod_{l=0}^{j-2} |\widetilde{G}(2^l\tfrac{k}{N})|^2 \\
&= \tfrac{1}{2}\left(|H(2^{j-1}\tfrac{k}{N})|^2 + |G(2^{j-1}\tfrac{k}{N})|^2\right) |\widetilde{G}_{j-1}(\tfrac{k}{N})|^2 \\
&= |\widetilde{G}_{j-1}(\tfrac{k}{N})|^2
\end{aligned}$$

since $|H(f)|^2 + |G(f)|^2 = \mathcal{H}(f) + \mathcal{G}(f) = 2$

MODWT Multiresolution Analysis: VI

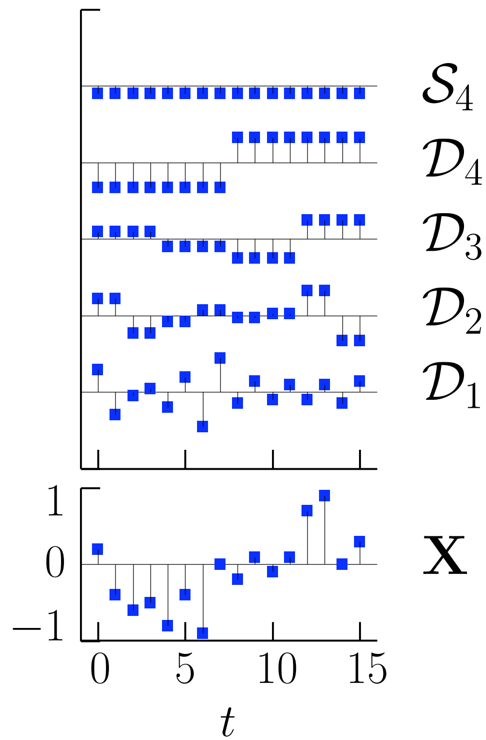
- implies $\{\widetilde{\mathcal{D}}_{j,t} + \widetilde{\mathcal{S}}_{j,t}\} \longleftrightarrow \{|\widetilde{G}_{j-1}(\frac{k}{N})|^2 \mathcal{X}_k\}$
- compare the above to $\{\widetilde{\mathcal{S}}_{j,t}\} \longleftrightarrow \{|\widetilde{G}_{j}(\frac{k}{N})|^2 \mathcal{X}_k\}$ and evoke the uniqueness of the DFT to get $\widetilde{\mathcal{S}}_{j-1} = \widetilde{\mathcal{D}}_j + \widetilde{\mathcal{S}}_j$ for $j \geq 2$
- hence $\widetilde{\mathcal{S}}_1 = \widetilde{\mathcal{D}}_2 + \widetilde{\mathcal{S}}_2 = \widetilde{\mathcal{D}}_2 + \widetilde{\mathcal{D}}_3 + \widetilde{\mathcal{S}}_3 = \cdots$, leading to

$$\widetilde{\mathcal{S}}_1 = \sum_{j=2}^{J_0} \widetilde{\mathcal{D}}_j + \widetilde{\mathcal{S}}_{J_0} \text{ and hence } \mathbf{X} = \sum_{j=1}^{J_0} \widetilde{\mathcal{D}}_j + \widetilde{\mathcal{S}}_{J_0}$$

if we use Exer. [172]: $\mathbf{X} = \widetilde{\mathcal{S}}_1 + \widetilde{\mathcal{D}}_1$ for all N & L

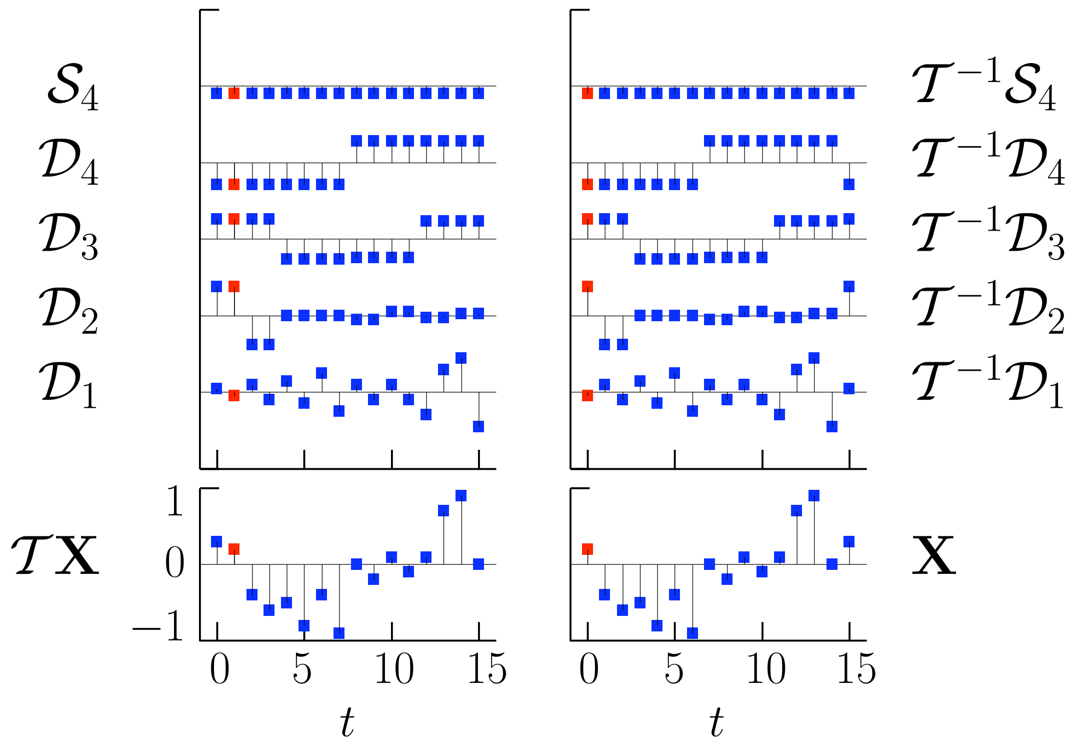
MODWT Multiresolution Analysis: VII

- if we form DWT-based MRAs for $\mathbf{X}$ and its circular shifts $\mathcal{T}^m\mathbf{X}$, $m = 1, \ldots, N-1$, we can obtain $\widetilde{\mathcal{D}}_j$ by appropriately averaging all N DWT-based details ('cycle spinning')



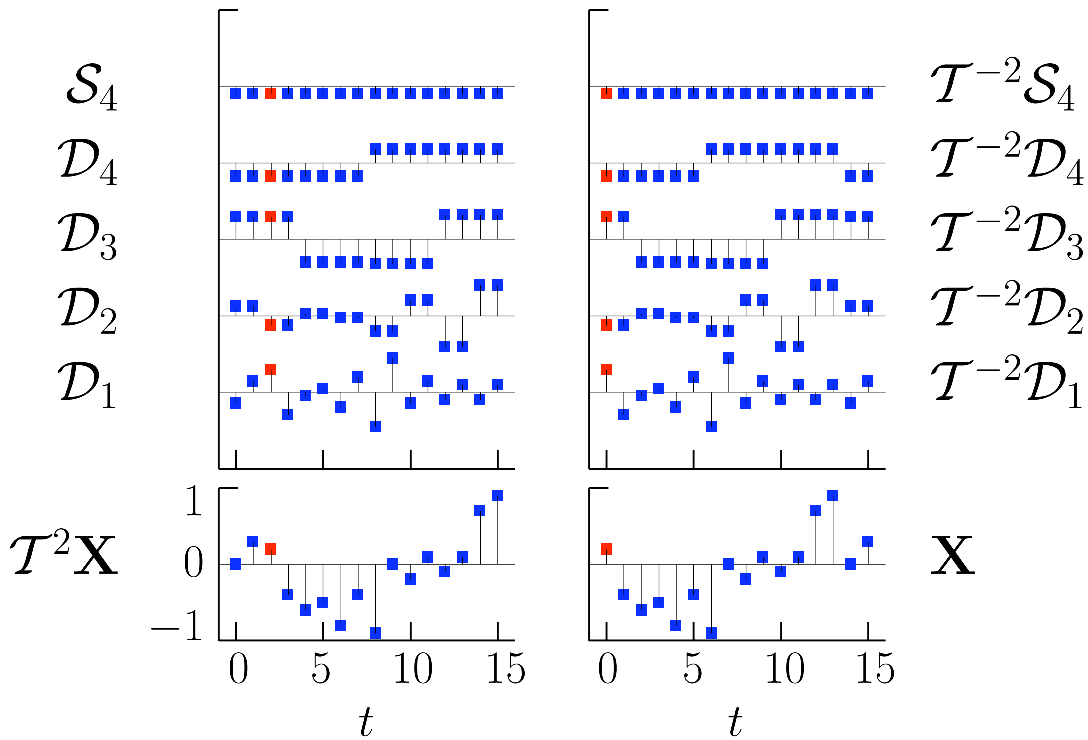
MODWT Multiresolution Analysis: VII

- if we form DWT-based MRAs for $\mathbf{X}$ and its circular shifts $\mathcal{T}^m\mathbf{X}$, $m = 1, \ldots, N-1$, we can obtain $\widetilde{\mathcal{D}}_j$ by appropriately averaging all N DWT-based details ('cycle spinning')



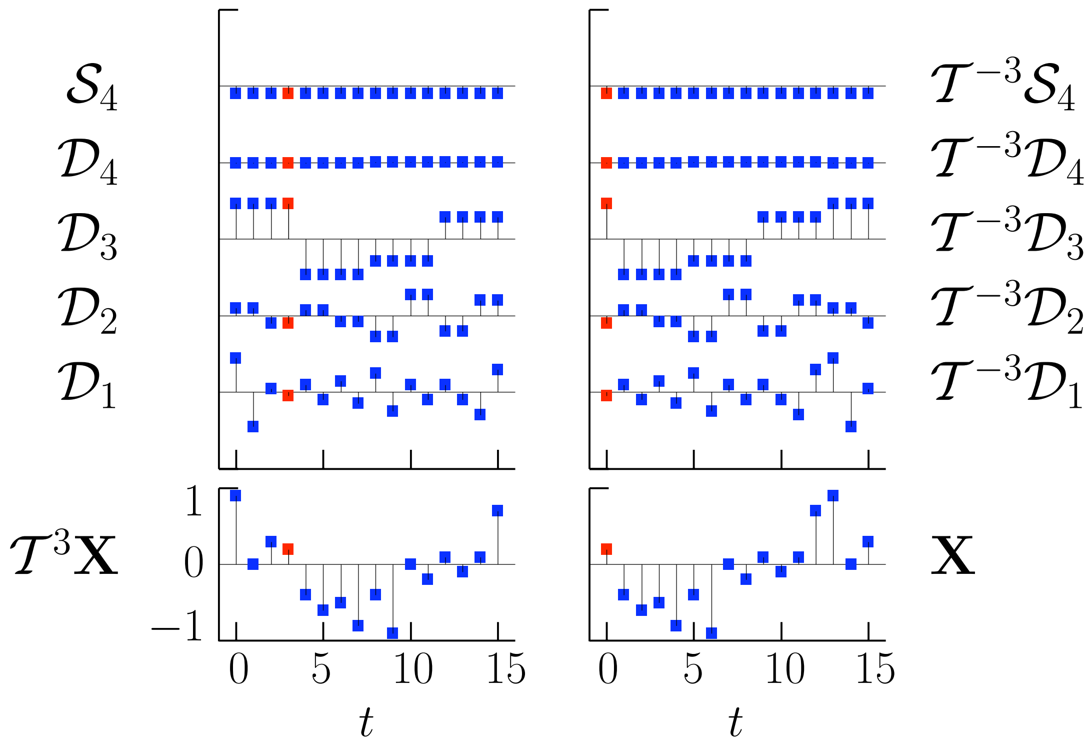
MODWT Multiresolution Analysis: VII

- if we form DWT-based MRAs for $\mathbf{X}$ and its circular shifts $\mathcal{T}^m\mathbf{X}$, $m = 1, \ldots, N-1$, we can obtain $\widetilde{\mathcal{D}}_j$ by appropriately averaging all N DWT-based details ('cycle spinning')



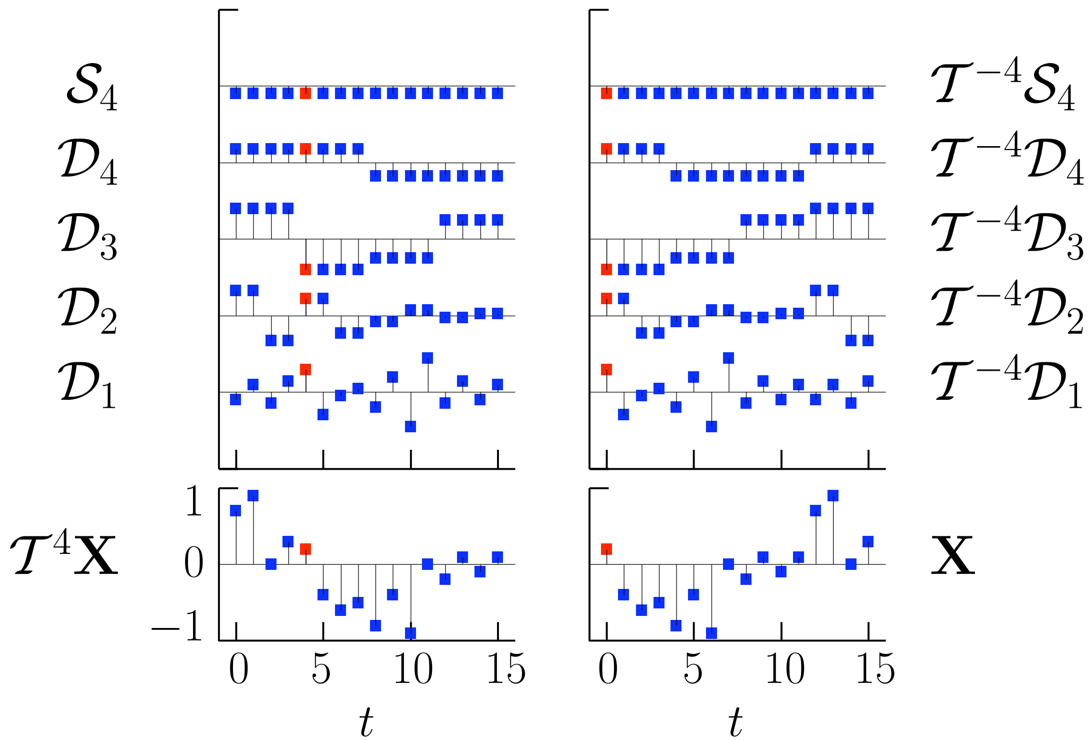
MODWT Multiresolution Analysis: VII

- if we form DWT-based MRAs for $\mathbf{X}$ and its circular shifts $\mathcal{T}^m\mathbf{X}$, $m = 1, \ldots, N-1$, we can obtain $\widetilde{\mathcal{D}}_j$ by appropriately averaging all N DWT-based details ('cycle spinning')



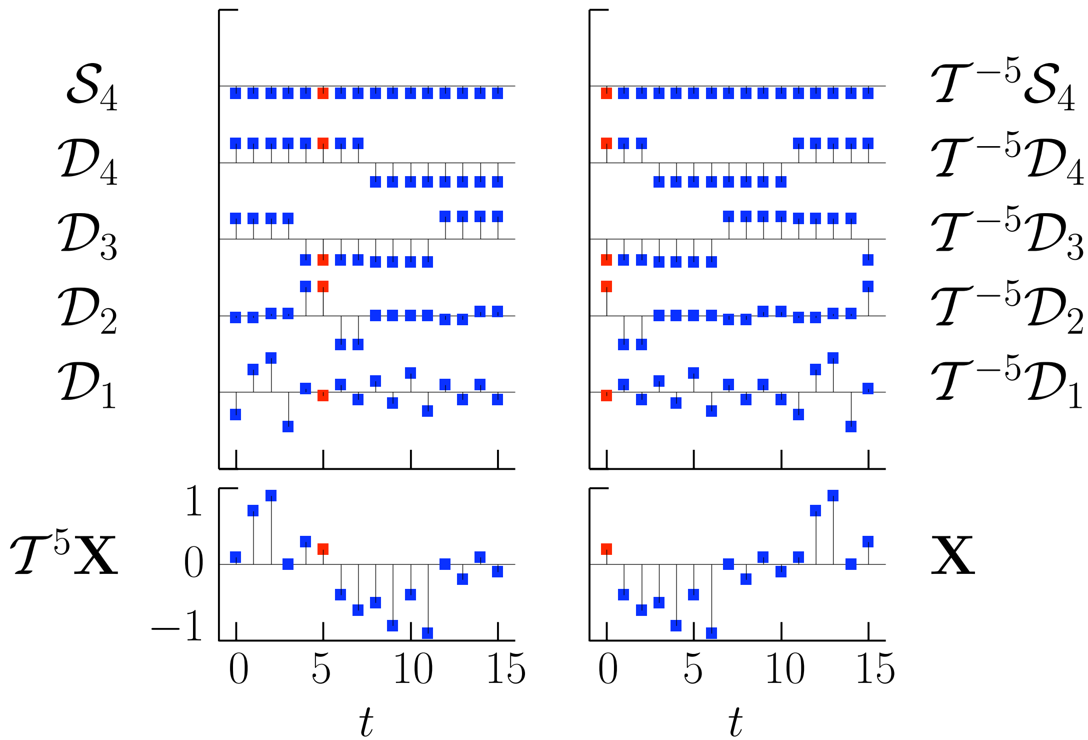
MODWT Multiresolution Analysis: VII

- if we form DWT-based MRAs for $\mathbf{X}$ and its circular shifts $\mathcal{T}^m\mathbf{X}$, $m = 1, \ldots, N-1$, we can obtain $\widetilde{\mathcal{D}}_j$ by appropriately averaging all N DWT-based details ('cycle spinning')



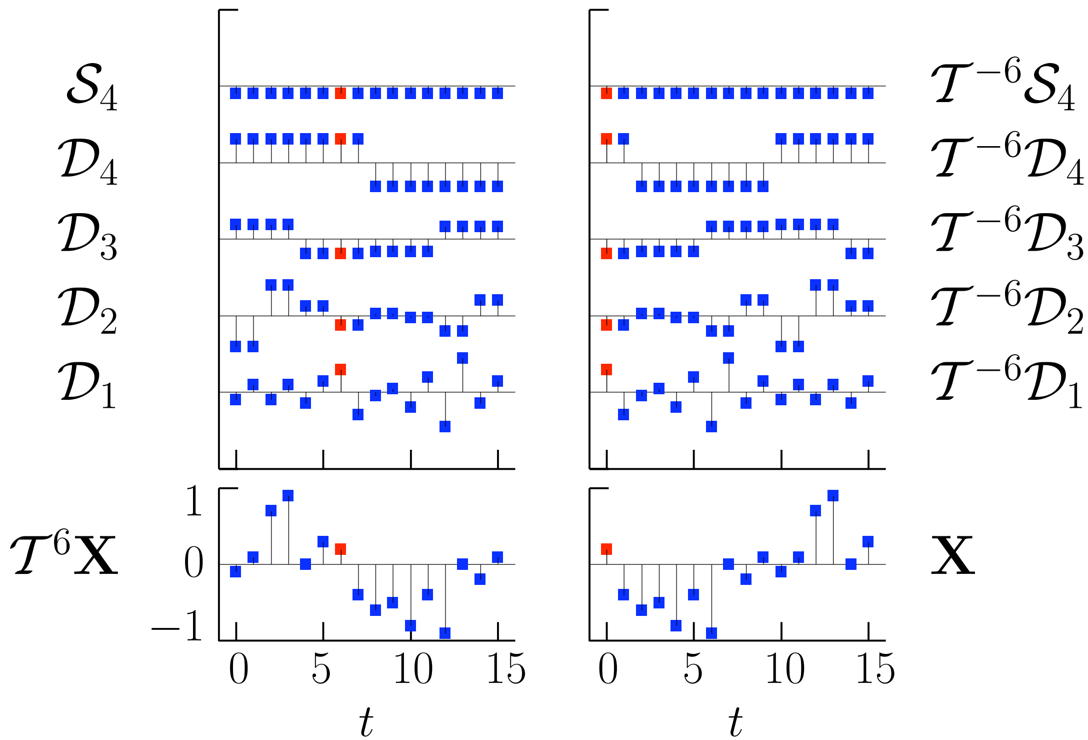
MODWT Multiresolution Analysis: VII

- if we form DWT-based MRAs for $\mathbf{X}$ and its circular shifts $\mathcal{T}^m\mathbf{X}$, $m = 1, \ldots, N-1$, we can obtain $\widetilde{\mathcal{D}}_j$ by appropriately averaging all N DWT-based details ('cycle spinning')



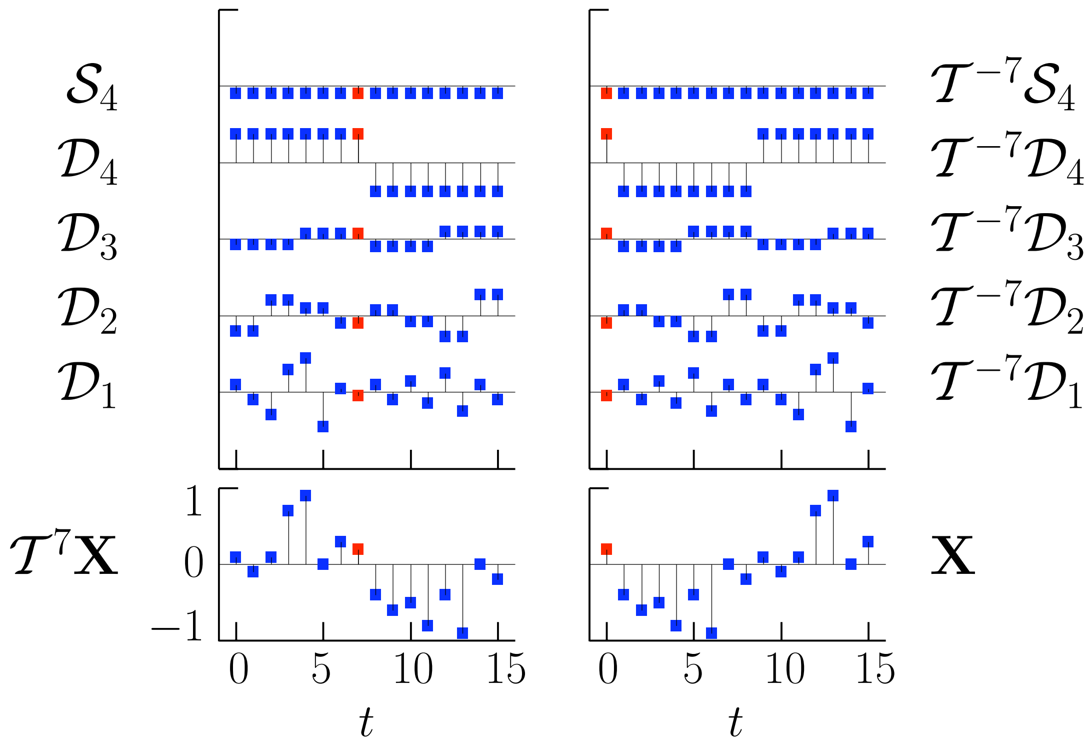
MODWT Multiresolution Analysis: VII

- if we form DWT-based MRAs for $\mathbf{X}$ and its circular shifts $\mathcal{T}^m\mathbf{X}$, $m = 1, \ldots, N-1$, we can obtain $\widetilde{\mathcal{D}}_j$ by appropriately averaging all N DWT-based details ('cycle spinning')



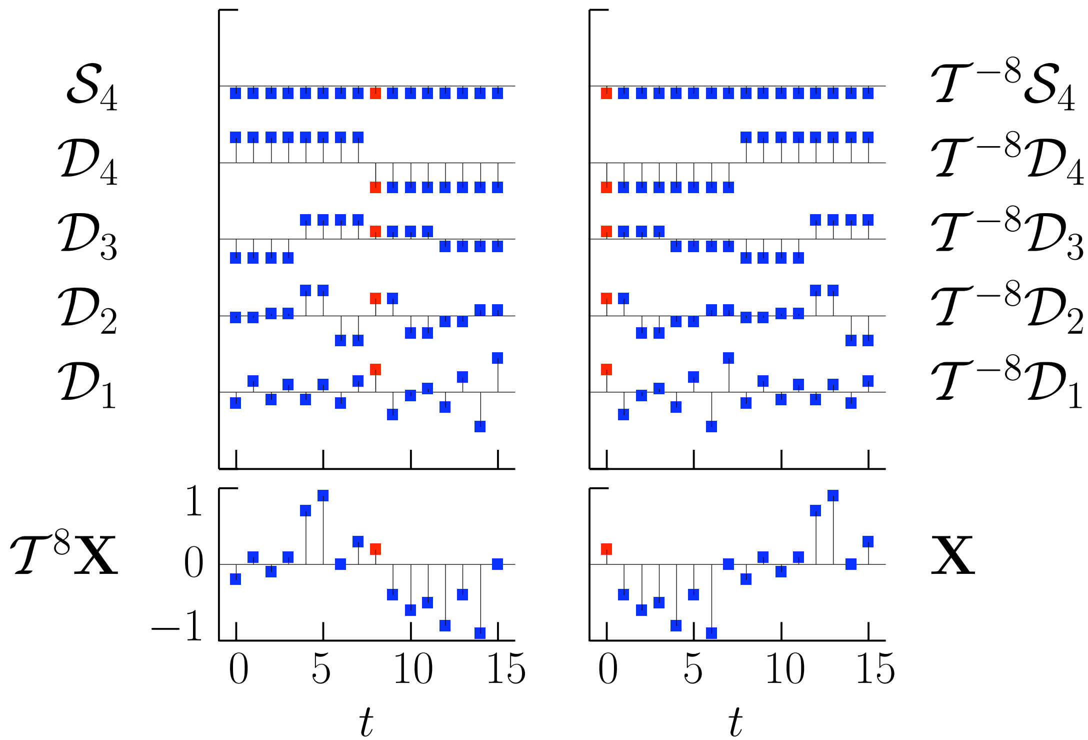
MODWT Multiresolution Analysis: VII

- if we form DWT-based MRAs for $\mathbf{X}$ and its circular shifts $\mathcal{T}^m\mathbf{X}$, $m = 1, \ldots, N-1$, we can obtain $\widetilde{\mathcal{D}}_j$ by appropriately averaging all N DWT-based details ('cycle spinning')



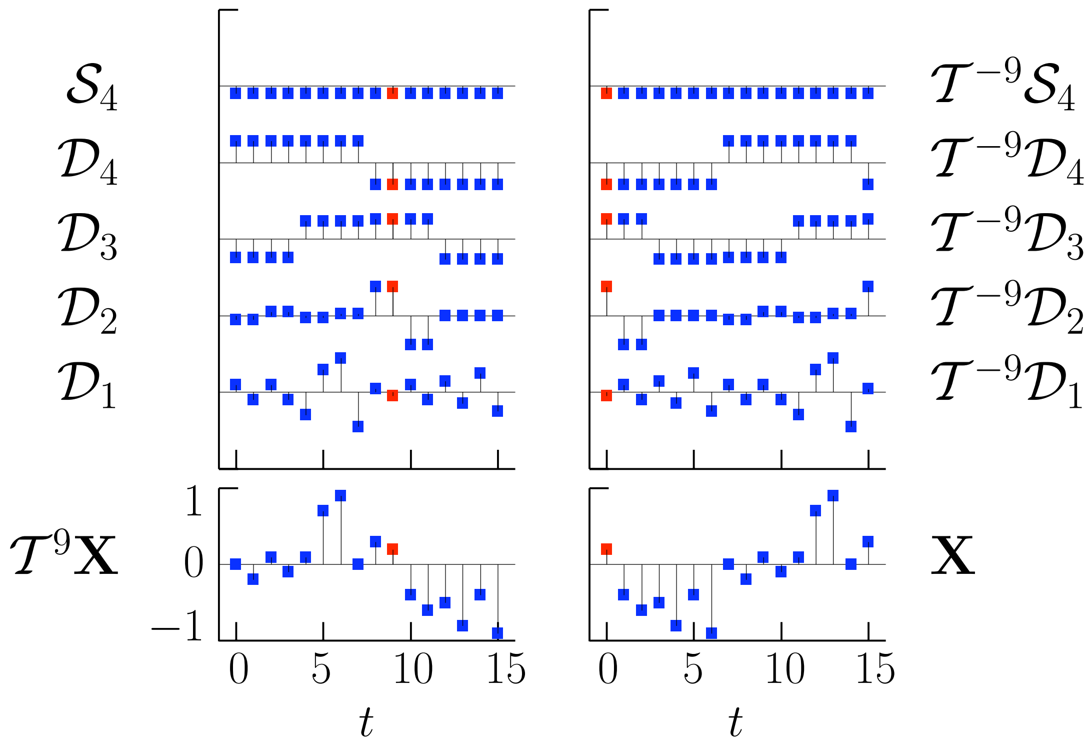
MODWT Multiresolution Analysis: VII

- if we form DWT-based MRAs for $\mathbf{X}$ and its circular shifts $\mathcal{T}^m\mathbf{X}$, $m = 1, \ldots, N-1$, we can obtain $\widetilde{\mathcal{D}}_j$ by appropriately averaging all N DWT-based details ('cycle spinning')



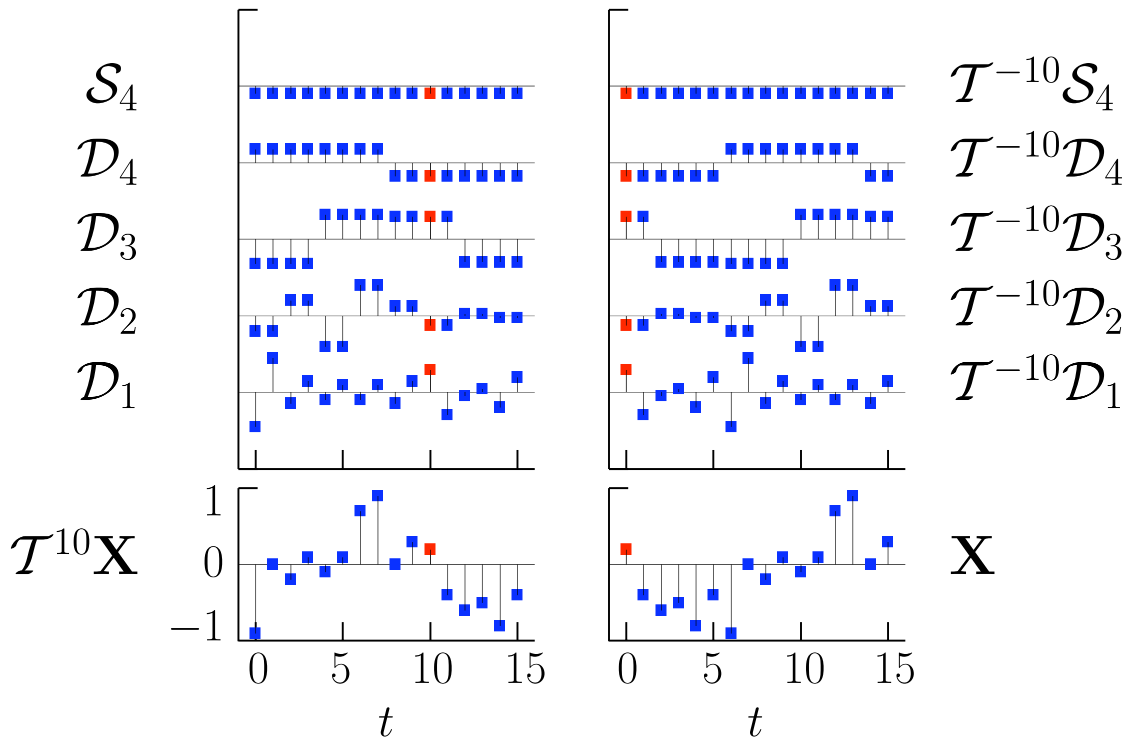
MODWT Multiresolution Analysis: VII

- if we form DWT-based MRAs for $\mathbf{X}$ and its circular shifts $\mathcal{T}^m\mathbf{X}$, $m = 1, \ldots, N-1$, we can obtain $\widetilde{\mathcal{D}}_j$ by appropriately averaging all N DWT-based details ('cycle spinning')



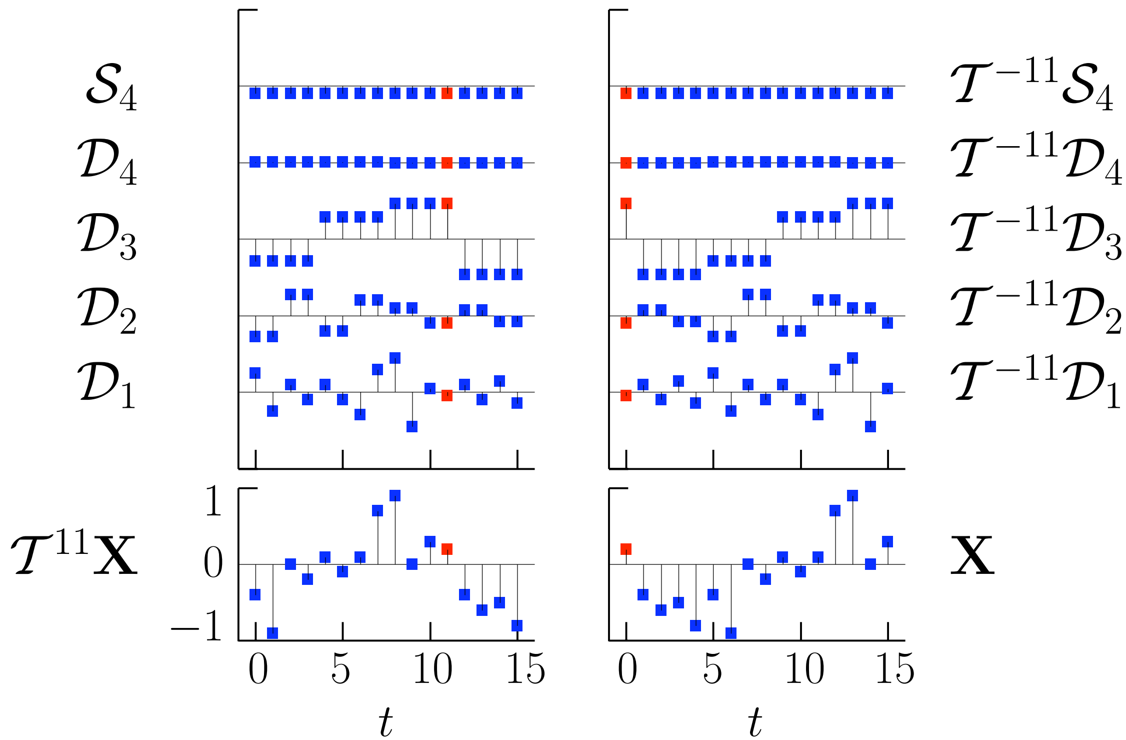
MODWT Multiresolution Analysis: VII

- if we form DWT-based MRAs for $\mathbf{X}$ and its circular shifts $\mathcal{T}^m\mathbf{X}$, $m = 1, \ldots, N-1$, we can obtain $\widetilde{\mathcal{D}}_j$ by appropriately averaging all N DWT-based details ('cycle spinning')



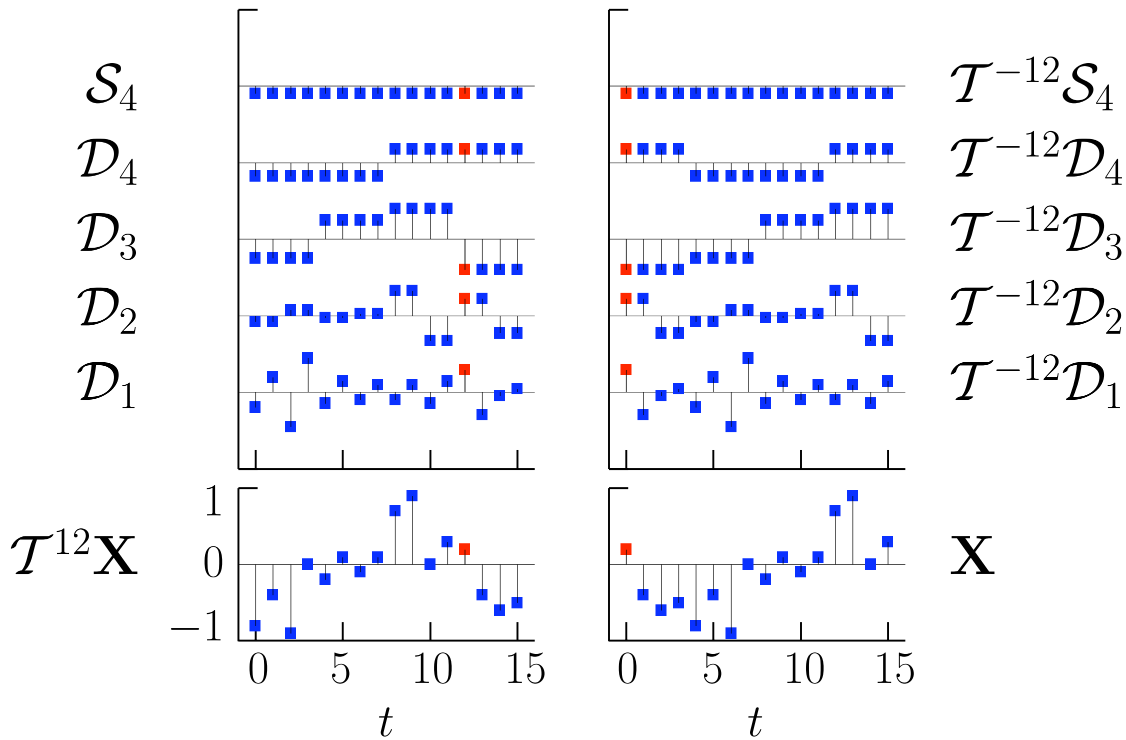
MODWT Multiresolution Analysis: VII

- if we form DWT-based MRAs for $\mathbf{X}$ and its circular shifts $\mathcal{T}^m\mathbf{X}$, $m = 1, \ldots, N-1$, we can obtain $\widetilde{\mathcal{D}}_j$ by appropriately averaging all N DWT-based details ('cycle spinning')



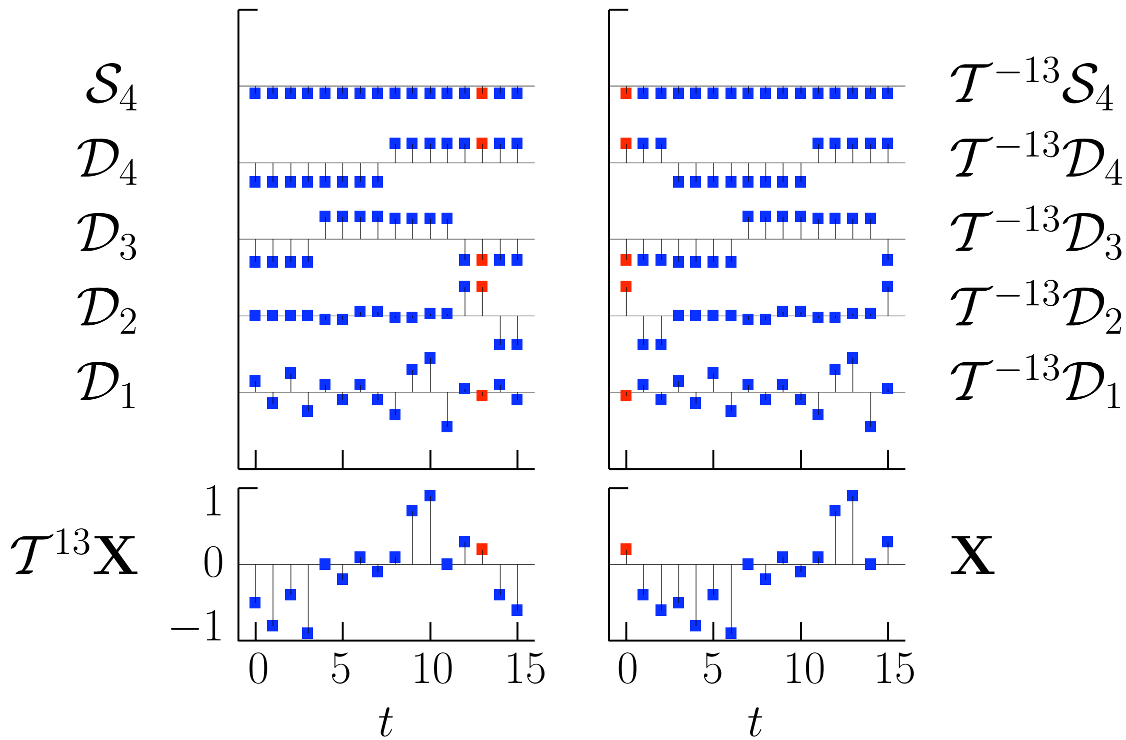
MODWT Multiresolution Analysis: VII

- if we form DWT-based MRAs for $\mathbf{X}$ and its circular shifts $\mathcal{T}^m\mathbf{X}$, $m = 1, \ldots, N-1$, we can obtain $\widetilde{\mathcal{D}}_j$ by appropriately averaging all N DWT-based details ('cycle spinning')



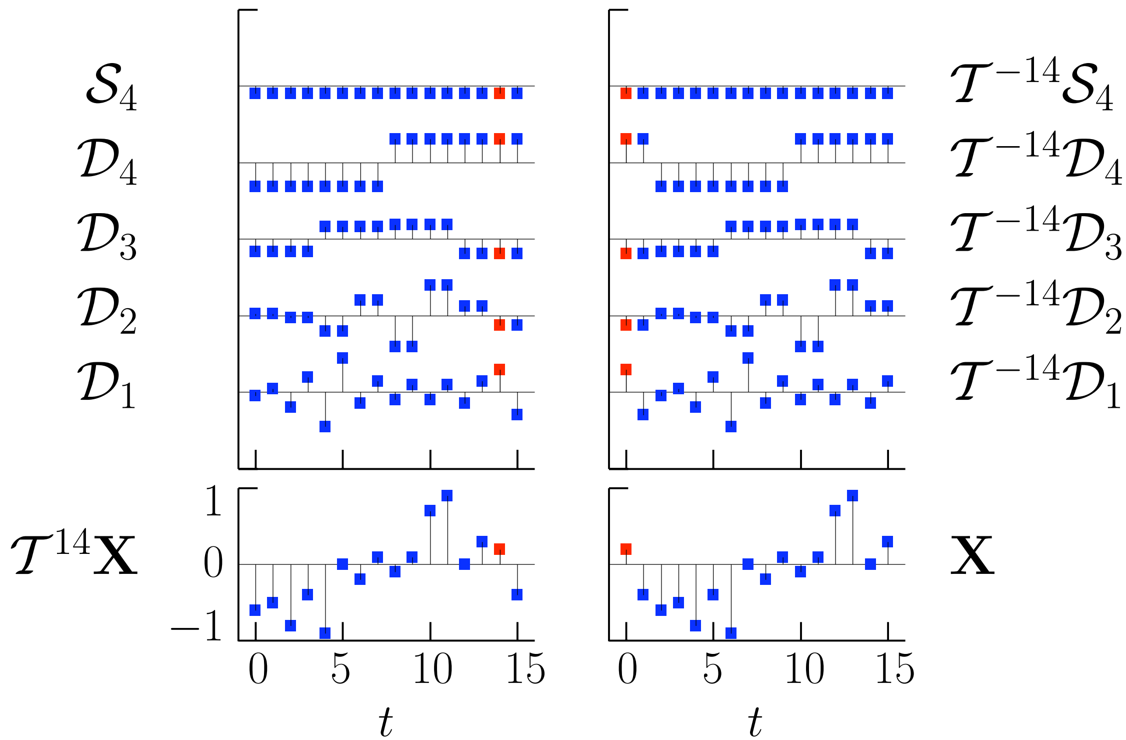
MODWT Multiresolution Analysis: VII

- if we form DWT-based MRAs for $\mathbf{X}$ and its circular shifts $\mathcal{T}^m\mathbf{X}$, $m = 1, \ldots, N-1$, we can obtain $\widetilde{\mathcal{D}}_j$ by appropriately averaging all N DWT-based details ('cycle spinning')



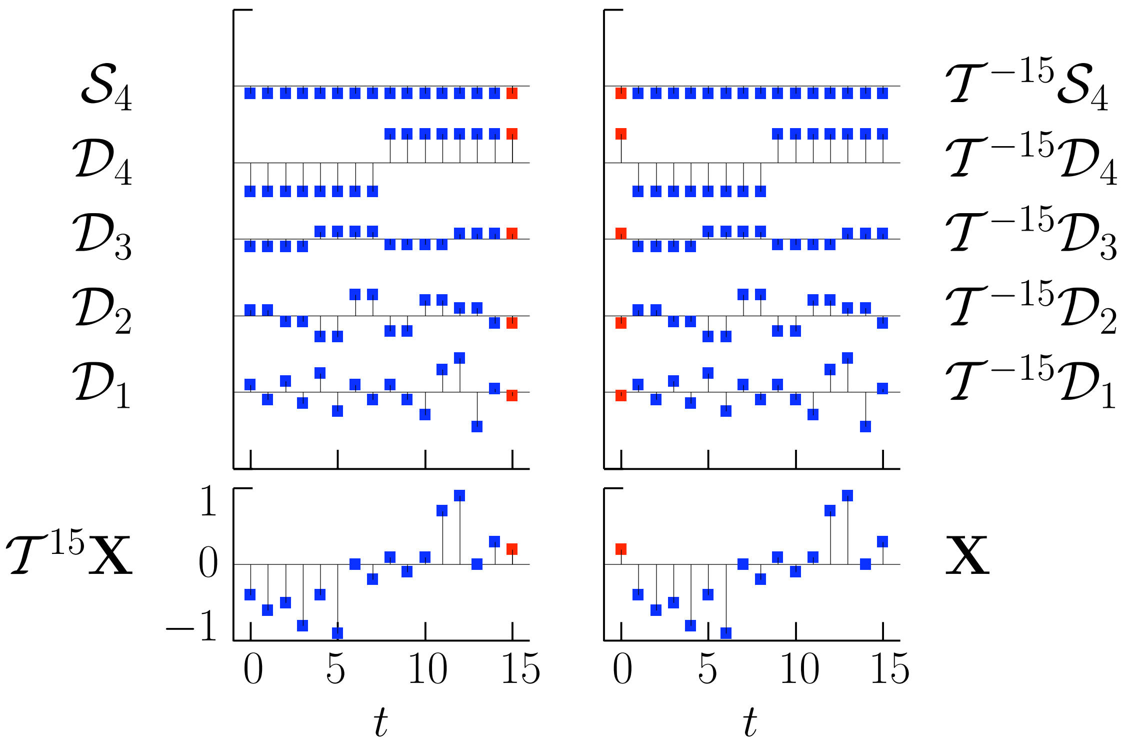
MODWT Multiresolution Analysis: VII

- if we form DWT-based MRAs for $\mathbf{X}$ and its circular shifts $\mathcal{T}^m\mathbf{X}$, $m = 1, \ldots, N-1$, we can obtain $\widetilde{\mathcal{D}}_j$ by appropriately averaging all N DWT-based details ('cycle spinning')



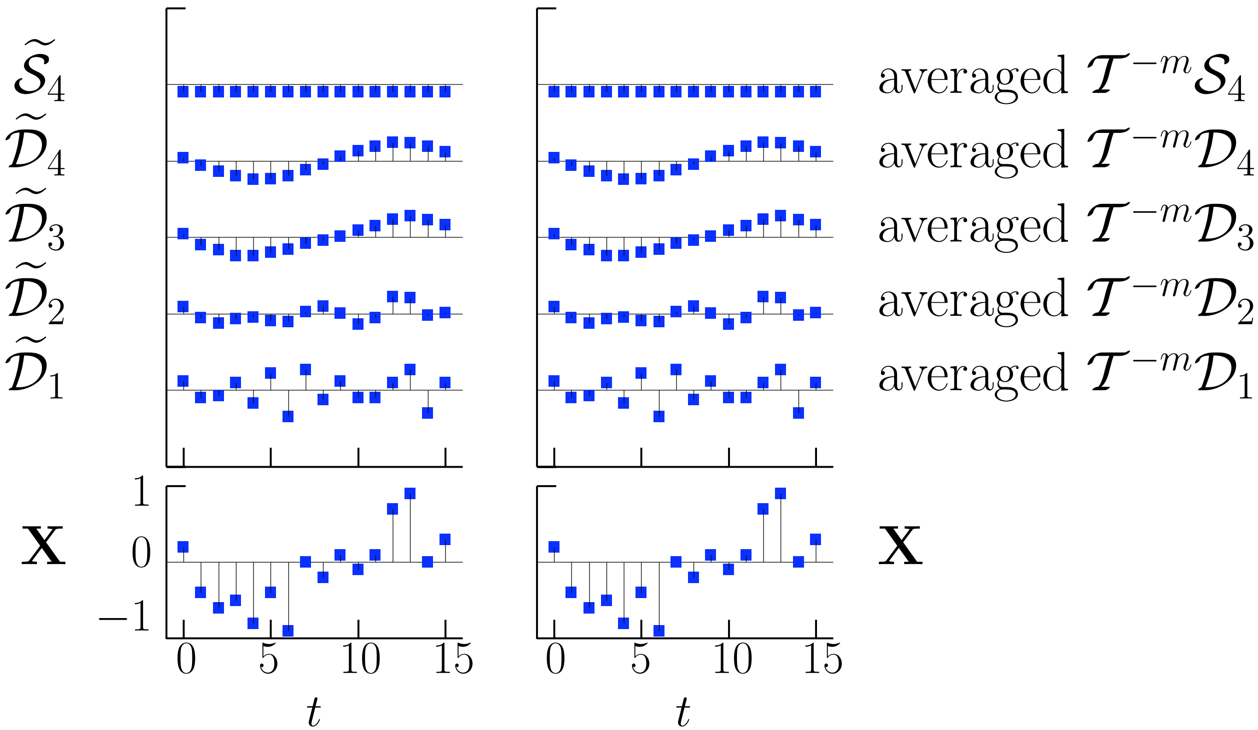
MODWT Multiresolution Analysis: VII

- if we form DWT-based MRAs for $\mathbf{X}$ and its circular shifts $\mathcal{T}^m\mathbf{X}$, $m = 1, \ldots, N-1$, we can obtain $\widetilde{\mathcal{D}}_j$ by appropriately averaging all N DWT-based details ('cycle spinning')



MODWT Multiresolution Analysis: VIII

- left-hand plots show $\widetilde{\mathcal{D}}_j$, while right-hand plots show average of $\mathcal{T}^{-m}\mathcal{D}_j$ in MRA for $\mathcal{T}^m\mathbf{X}$, $m = 0, 1, \ldots, 15$



MODWT Analysis of Variance: I

- for any $J_0 \geq 1$ & $N \geq 1$, will now show that

$$\|\mathbf{X}\|^2 = \sum_{j=1}^{J_0} \|\widetilde{\mathbf{W}}_j\|^2 + \|\widetilde{\mathbf{V}}_{J_0}\|^2,$$

leading to an analysis of the sample variance of $\mathbf{X}$:

$$\hat{\sigma}_X^2 = \frac{1}{N} \sum_{j=1}^{J_0} \|\widetilde{\mathbf{W}}_j\|^2 + \frac{1}{N}\|\widetilde{\mathbf{V}}_{J_0}\|^2 - \overline{X}^2,$$

which is analogous to the DWT-based analysis of variance

MODWT Analysis of Variance: II

- as before, let $\{\mathcal{X}_k\}$ be the DFT of $\{X_t\}$ so that

$$\{\widetilde{W}_{j,t}\} \longleftrightarrow \{\widetilde{H}_j(\tfrac{k}{N})\mathcal{X}_k\} \quad \& \quad \{\widetilde{V}_{j,t}\} \longleftrightarrow \{\widetilde{G}_j(\tfrac{k}{N})\mathcal{X}_k\}$$

- Parseval's theorem says:

$$\|\widetilde{\mathbf{W}}_j\|^2 = \frac{1}{N}\sum_{k=0}^{N-1} |\widetilde{H}_j(\tfrac{k}{N})|^2|\mathcal{X}_k|^2 \quad \& \quad \|\widetilde{\mathbf{V}}_j\|^2 = \frac{1}{N}\sum_{k=0}^{N-1} |\widetilde{G}_j(\tfrac{k}{N})|^2|\mathcal{X}_k|^2$$

- since $|\widetilde{H}_j(\frac{k}{N})|^2+|\widetilde{G}_j(\frac{k}{N})|^2 = |\widetilde{G}_{j-1}(\frac{k}{N})|^2$, $j \geq 2$, adding yields

$$\begin{aligned}
\|\widetilde{\mathbf{W}}_j\|^2 + \|\widetilde{\mathbf{V}}_j\|^2 &= \frac{1}{N}\sum_{k=0}^{N-1} \left(|\widetilde{H}_j(\tfrac{k}{N})|^2 + |\widetilde{G}_j(\tfrac{k}{N})|^2\right) |\mathcal{X}_k|^2 \\
&= \frac{1}{N}\sum_{k=0}^{N-1} |\widetilde{G}_{j-1}(\tfrac{k}{N})|^2|\mathcal{X}_k|^2 = \|\widetilde{\mathbf{V}}_{j-1}\|^2
\end{aligned}$$

MODWT Analysis of Variance: III

- using $\|\widetilde{\mathbf{V}}_{j-1}\|^2 = \|\widetilde{\mathbf{W}}_j\|^2 + \|\widetilde{\mathbf{V}}_j\|^2$ for $j = 2, 3, \ldots, J_0$ yields

$$\begin{aligned}
\|\widetilde{\mathbf{V}}_1\|^2 &= \|\widetilde{\mathbf{W}}_2\|^2 + \|\widetilde{\mathbf{V}}_2\|^2 \\
&= \|\widetilde{\mathbf{W}}_2\|^2 + \|\widetilde{\mathbf{W}}_3\|^2 + \|\widetilde{\mathbf{V}}_3\|^2 \\
&\vdots \\
&= \sum_{j=2}^{J_0} \|\widetilde{\mathbf{W}}_j\|^2 + \|\widetilde{\mathbf{V}}_{J_0}\|^2
\end{aligned}$$

- desired result

$$\|\mathbf{X}\|^2 = \sum_{j=1}^{J_0} \|\widetilde{\mathbf{W}}_j\|^2 + \|\widetilde{\mathbf{V}}_{J_0}\|^2,$$

now follows if we can show that $\|\mathbf{X}\|^2 = \|\widetilde{\mathbf{W}}_1\|^2 + \|\widetilde{\mathbf{V}}_1\|^2$, and this is the subject of Exer. [171a]

MODWT Pyramid Algorithm: I

- goal: compute $\widetilde{\mathbf{W}}_j$ & $\widetilde{\mathbf{V}}_j$ using $\widetilde{\mathbf{V}}_{j-1}$ rather than $\mathbf{X}$
- can obtain all 3 by filtering $\mathbf{X}$ directly:
  - to get $\widetilde{\mathbf{V}}_j$, use $\{\widetilde{G}_j(\frac{k}{N}) = \widetilde{G}_{j-1}(\frac{k}{N})\widetilde{G}(2^{j-1}\frac{k}{N})\}$
  - to get $\widetilde{\mathbf{W}}_j$, use $\{\widetilde{H}_j(\frac{k}{N}) = \widetilde{G}_{j-1}(\frac{k}{N})\widetilde{H}(2^{j-1}\frac{k}{N})\}$
  - to get $\widetilde{\mathbf{V}}_{j-1}$, use $\{\widetilde{G}_{j-1}(\frac{k}{N})\}$
- can get $\widetilde{\mathbf{V}}_j$ & $\widetilde{\mathbf{W}}_j$ using $\widetilde{G}(2^{j-1}\frac{k}{N})$ & $\widetilde{H}(2^{j-1}\frac{k}{N})$ on $\widetilde{\mathbf{V}}_{j-1}$
- Exer. [91]: if $\{\tilde{h}_l\} \longleftrightarrow \widetilde{H}(f)$, the inverse DFT of $\widetilde{H}(2^{j-1}f)$ is

$$\{\tilde{h}_0, \underbrace{0,\ldots,0,}_{2^{j-1}-1\text{ zeros}} \tilde{h}_1, \underbrace{0,\ldots,0,}_{2^{j-1}-1\text{ zeros}} \ldots, \tilde{h}_{L-2}, \underbrace{0,\ldots,0,}_{2^{j-1}-1\text{ zeros}} \tilde{h}_{L-1}\}$$

MODWT Pyramid Algorithm: II

- letting $\widetilde{V}_{0,t} \equiv X_t$, implies that, for all $j \geq 1$,

$$\widetilde{W}_{j,t} = \sum_{l=0}^{L-1} \tilde{h}_l \widetilde{V}_{j-1,t-2^{j-1}l \bmod N} \quad \& \quad \widetilde{V}_{j,t} = \sum_{l=0}^{L-1} \tilde{g}_l \widetilde{V}_{j-1,t-2^{j-1}l \bmod N}$$

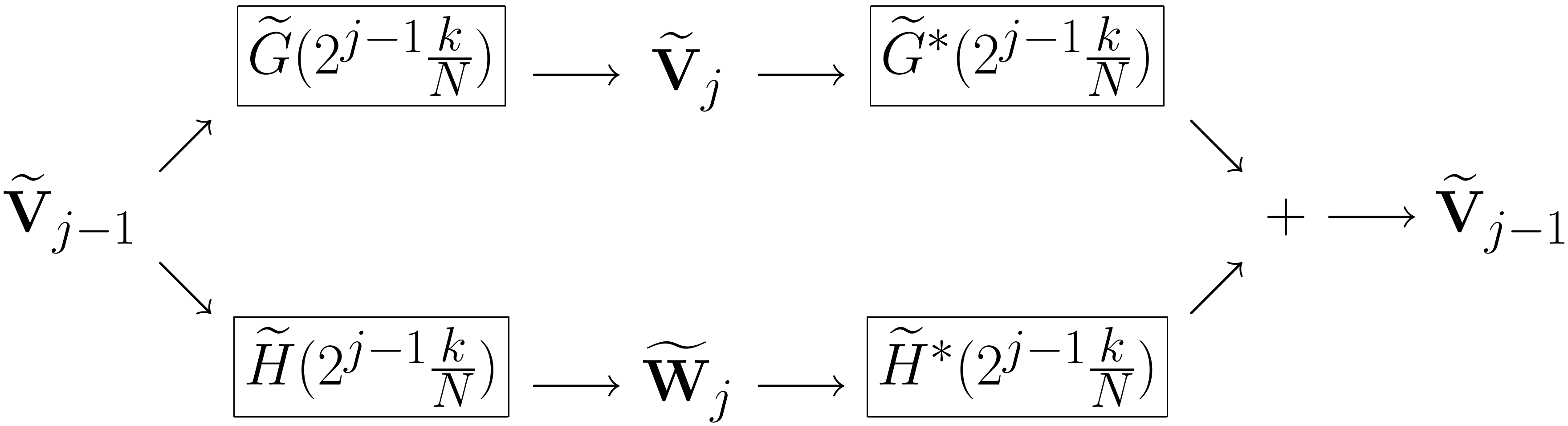
- algorithm requires $N \log_2(N)$ multiplications, which is the same as needed by fast Fourier transform algorithm
- inverse pyramid algorithm is given by

$$\widetilde{V}_{j-1,t} = \sum_{l=0}^{L-1} \tilde{h}_l \widetilde{W}_{j,t+2^{j-1}l \bmod N} + \sum_{l=0}^{L-1} \tilde{g}_l \widetilde{V}_{j,t+2^{j-1}l \bmod N}$$

(proof of this statement is the subject of Exer. [175])

MODWT Pyramid Algorithm: III

- pyramid algorithm summarized in following flow diagram:



- item [1] of Comments and Extensions to Sec. 5.5 has pseudo code for MODWT pyramid algorithm

MODWT Pyramid Algorithm: IV

- similar to DWT, can describe transform from $\widetilde{\mathbf{V}}_{j-1}$ to $\widetilde{\mathbf{W}}_j$ & $\widetilde{\mathbf{V}}_j$ as $\widetilde{\mathbf{W}}_j = \widetilde{\mathcal{B}}_j \widetilde{\mathbf{V}}_{j-1}$ & $\widetilde{\mathbf{V}}_j = \widetilde{\mathcal{A}}_j \widetilde{\mathbf{V}}_{j-1}$, where now $\widetilde{\mathcal{B}}_j$ & $\widetilde{\mathcal{A}}_j$ are $N \times N$ matrices
- rows of $\widetilde{\mathcal{B}}_j$ contain inverse DFT of $\{\widetilde{H}(2^{j-1}\frac{k}{N})\}$
- rows of $\widetilde{\mathcal{A}}_j$ contain inverse DFT of $\{\widetilde{G}(2^{j-1}\frac{k}{N})\}$
- example of $\widetilde{\mathcal{B}}_j$ with $j = 2$, $N = 12$ & $L = 4$:

$$\widetilde{\mathcal{B}}_2 \equiv \begin{bmatrix} \tilde{h}_0 & 0 & 0 & 0 & 0 & 0 & \tilde{h}_3 & 0 & \tilde{h}_2 & 0 & \tilde{h}_1 & 0 \\ 0 & \tilde{h}_0 & 0 & 0 & 0 & 0 & 0 & \tilde{h}_3 & 0 & \tilde{h}_2 & 0 & \tilde{h}_1 \\ \tilde{h}_1 & 0 & \tilde{h}_0 & 0 & 0 & 0 & 0 & 0 & \tilde{h}_3 & 0 & \tilde{h}_2 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \tilde{h}_3 & 0 & \tilde{h}_2 & 0 & \tilde{h}_1 & 0 & \tilde{h}_0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \tilde{h}_3 & 0 & \tilde{h}_2 & 0 & \tilde{h}_1 & 0 & \tilde{h}_0 \end{bmatrix}$$

MODWT Pyramid Algorithm: V

- Exer. [175]: like DWT, can express $\widetilde{\mathbf{W}}_j$ & $\widetilde{\mathbf{V}}_j \longrightarrow \widetilde{\mathbf{V}}_{j-1}$ as

$$\widetilde{\mathbf{V}}_{j-1} = \widetilde{\mathcal{B}}_j^T \widetilde{\mathbf{W}}_j + \widetilde{\mathcal{A}}_j^T \widetilde{\mathbf{V}}_j$$

- starting with $\widetilde{\mathbf{V}}_0 = \mathbf{X}$, J_0 recursions yield

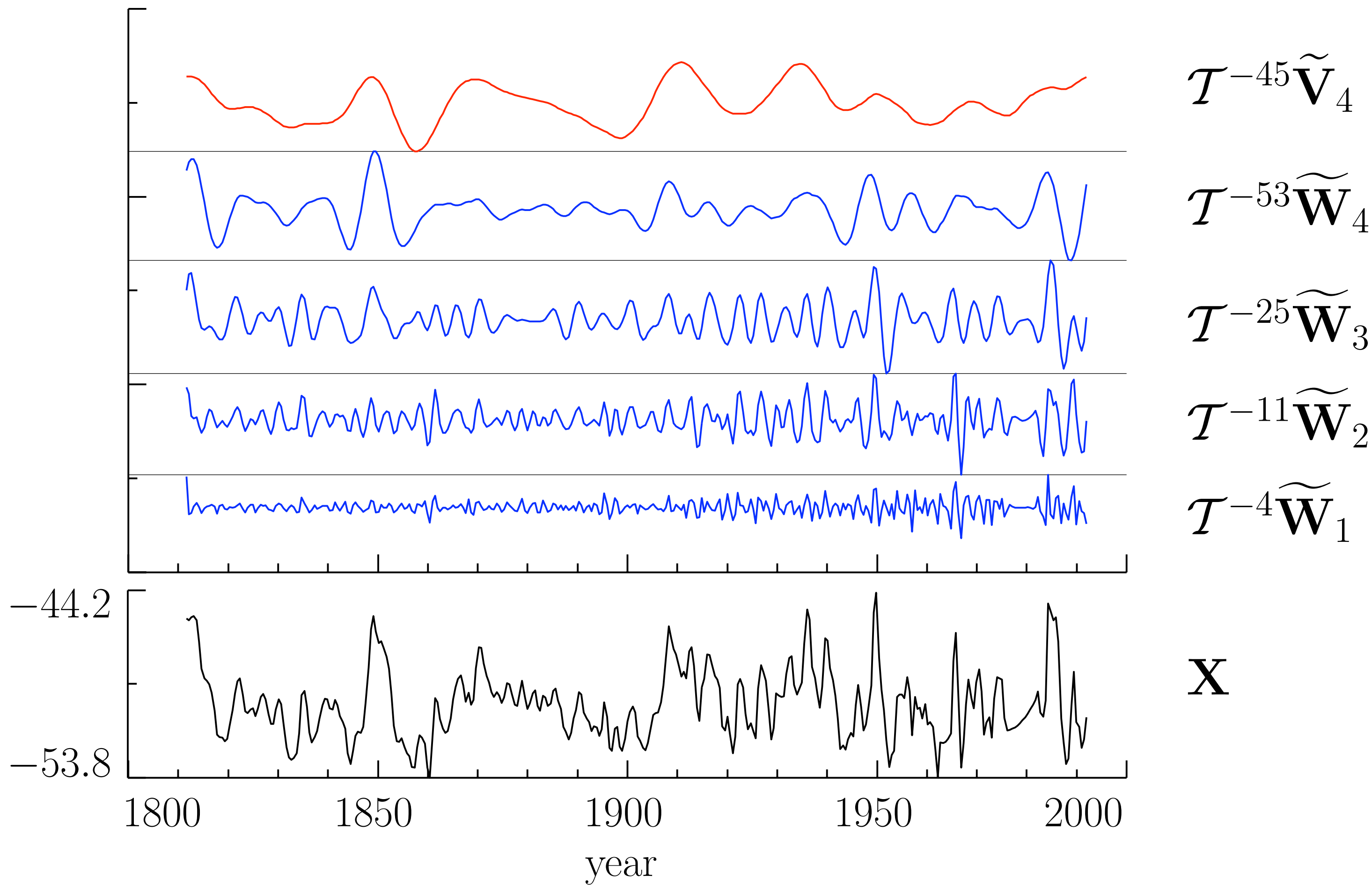
$$\begin{aligned}\mathbf{X} &= \underbrace{\widetilde{\mathcal{B}}_1^T \widetilde{\mathbf{W}}_1}_{\widetilde{\mathcal{D}}_1} + \underbrace{\widetilde{\mathcal{A}}_1^T \widetilde{\mathcal{B}}_2^T \widetilde{\mathbf{W}}_2}_{\widetilde{\mathcal{D}}_2} + \underbrace{\widetilde{\mathcal{A}}_1^T \widetilde{\mathcal{A}}_2^T \widetilde{\mathcal{B}}_3^T \widetilde{\mathbf{W}}_3}_{\widetilde{\mathcal{D}}_3} + \cdots \\ &\quad + \underbrace{\widetilde{\mathcal{A}}_1^T \cdots \widetilde{\mathcal{A}}_{J_0-1}^T \widetilde{\mathcal{B}}_{J_0}^T \widetilde{\mathbf{W}}_{J_0}}_{\widetilde{\mathcal{D}}_{J_0}} + \underbrace{\widetilde{\mathcal{A}}_1^T \cdots \widetilde{\mathcal{A}}_{J_0-1}^T \widetilde{\mathcal{A}}_{J_0}^T \widetilde{\mathbf{V}}_{J_0}}_{\widetilde{\mathcal{S}}_{J_0}}\end{aligned}$$

- since $\widetilde{\mathcal{D}}_j \equiv \widetilde{\mathcal{W}}_j^T \widetilde{\mathbf{W}}_j$ and $\widetilde{\mathcal{S}}_{J_0} \equiv \widetilde{\mathcal{V}}_{J_0}^T \widetilde{\mathbf{V}}_{J_0}$, we evidently have

$$\widetilde{\mathcal{W}}_j = \widetilde{\mathcal{B}}_j \widetilde{\mathcal{A}}_{j-1} \cdots \widetilde{\mathcal{A}}_1 \text{ and } \widetilde{\mathcal{V}}_{J_0} = \widetilde{\mathcal{A}}_{J_0} \widetilde{\mathcal{A}}_{J_0-1} \cdots \widetilde{\mathcal{A}}_1$$

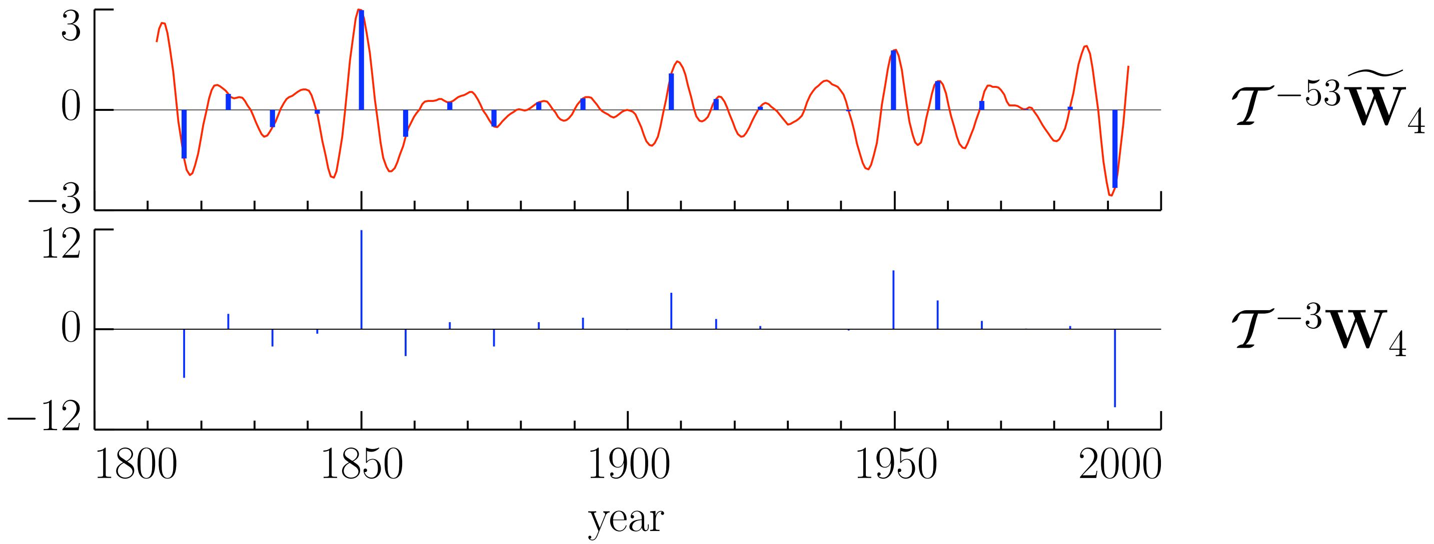
Example of $J_0 = 4$ LA(8) MODWT

- oxygen isotope records $\mathbf{X}$ from Antarctic ice core



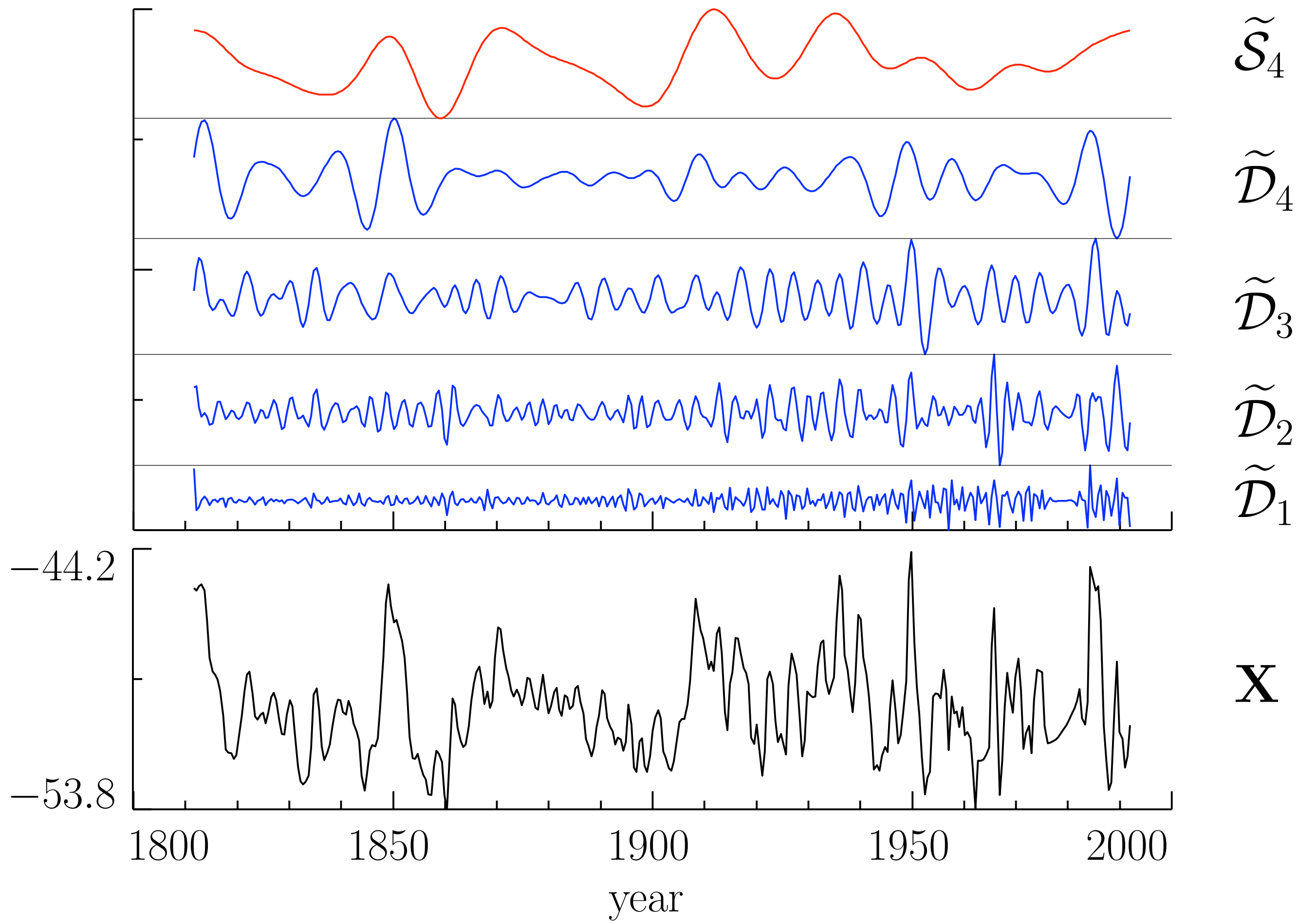
Relationship Between MODWT and DWT

- bottom plot shows $\mathbf{W}_4$ from DWT after circular shift $\mathcal{T}^{-3}$ to align coefficients properly in time (more about $\mathcal{T}$ later)
- top plot shows $\widetilde{\mathbf{W}}_4$ from MODWT and subsamples that, upon rescaling, yield $\mathbf{W}_4$ via $W_{4,t} = 4\widetilde{W}_{4,16(t+1)-1}$



Example of $J_0 = 4$ LA(8) MODWT MRA

- oxygen isotope records $\mathbf{X}$ from Antarctic ice core



Example of Variance Decomposition

- decomposition of sample variance from MODWT

$$\hat{\sigma}_X^2 \equiv \frac{1}{N}\sum_{t=0}^{N-1}\left(X_t - \overline{X}\right)^2 = \sum_{j=1}^{4}\frac{1}{N}\|\widetilde{\mathbf{W}}_j\|^2 + \frac{1}{N}\|\widetilde{\mathbf{V}}_4\|^2 - \overline{X}^2$$

- LA(8)-based example for oxygen isotope records
  - 0.5 year changes: $\frac{1}{N}\|\widetilde{\mathbf{W}}_1\|^2 \doteq 0.145\ (\doteq\ 4.5\%$ of $\hat{\sigma}_X^2)$
  - 1.0 years changes: $\frac{1}{N}\|\widetilde{\mathbf{W}}_2\|^2 \doteq 0.500\ (\doteq 15.6\%)$
  - 2.0 years changes: $\frac{1}{N}\|\widetilde{\mathbf{W}}_3\|^2 \doteq 0.751\ (\doteq 23.4\%)$
  - 4.0 years changes: $\frac{1}{N}\|\widetilde{\mathbf{W}}_4\|^2 \doteq 0.839\ (\doteq 26.2\%)$
  - 8.0 years averages: $\frac{1}{N}\|\widetilde{\mathbf{V}}_4\|^2 - \overline{X}^2 \doteq 0.969\ (\doteq 30.2\%)$
  - sample variance: $\hat{\sigma}_X^2 \doteq 3.204$

Summary of Key Points about the MODWT

- similar to the DWT, the MODWT offers
  - a scale-based multiresolution analysis
  - a scale-based analysis of the sample variance
  - a pyramid algorithm for computing the transform efficiently
- unlike the DWT, the MODWT is
  - defined for all sample sizes (no 'power of 2' restrictions)
  - unaffected by circular shifts to $\mathbf{X}$ in that coefficients, details and smooths shift along with $\mathbf{X}$ (example coming later)
  - highly redundant in that a level J_0 transform consists of $(J_0 + 1)N$ values rather than just N
- as we shall see, the MODWT can eliminate 'alignment' artifacts, but its redundancies are problematic for some uses