Appendix

6.1 Evaluating Performance

To evaluate the clustering performance we used the *Adjusted Rand Index* (ARI) as a measure of clustering accuracy [25], defined as:

$$ARI \equiv \frac{\sum_{i<j}\binom{n_{ij}}{2} - \frac{1}{\binom{n}{2}}[\sum_i \binom{a_i}{2}][\sum_j \binom{b_j}{2}]}{\frac{1}{2}[\sum_i \binom{a_i}{2} + \sum_j \binom{b_j}{2}] - \frac{1}{\binom{n}{2}}[\sum_i \binom{a_i}{2}][\sum_j \binom{b_j}{2}]} \tag{13}$$

where a_i denotes the number of signals whose true cluster is i, , b_j the number of signals assigned to cluster j and n_{ij} the number of signals whose true cluster is i and are assigned to cluster j, with $n = \sum_i a_i = \sum_j b_j = \sum_{i,j} n_{ij}$.

6.2 Additional Figures for Simulation

The coefficients of the clustering in simulations are shown in Figure 6.

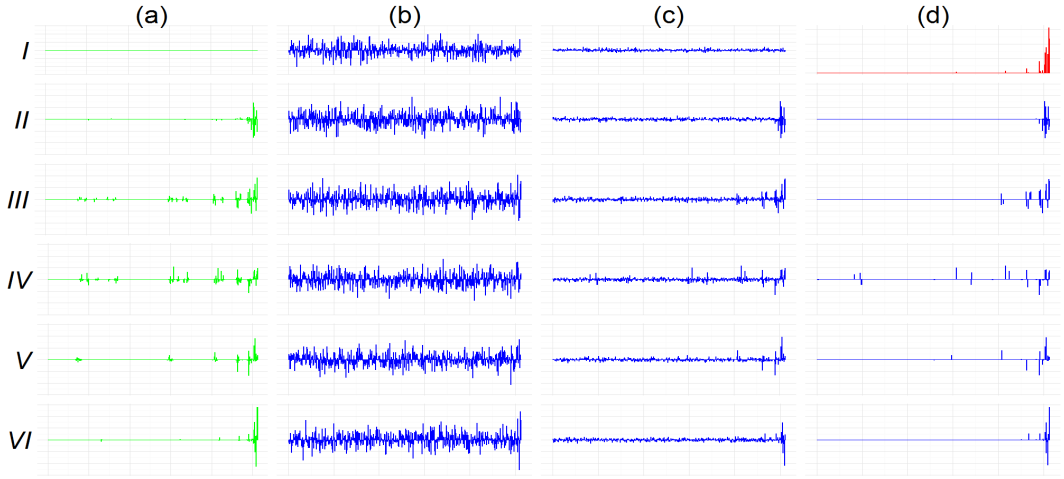

Figure 6: **Simulated cluster data coefficients**. Wavelet coefficients for six clusters from the simulation described in main text (see Figure 1): I. Flat curve, II. Heavisine, III. Blocks, IV. Bumps, V. Doppler, VI. Piecewise polynomial. **(a)** Coefficients for true cluster centers.**(b)** Coefficients for one individual curve from each cluster. SNR is too low to allow individual curve smoothing to reliably estimate cluster centers. **(c)** Coefficients of cluster centers returned by SPARCWave resemble the original clusters coefficients. **(d)** Coefficients after Wavelets smoothing of returned cluster centers improves estimation of original cluster centers for cluster center from $II. - VI.$ (except for the trivial flat curve), shown in blue. The w_i coefficients fitted in SPARCWave, shown in red. Although a few wavelet coefficients at fine resolutions are large for some of the cluster centers, the informative coefficients for clustering are all at the coarsest levels.

6.3 Additional Details and Figures for Real Data

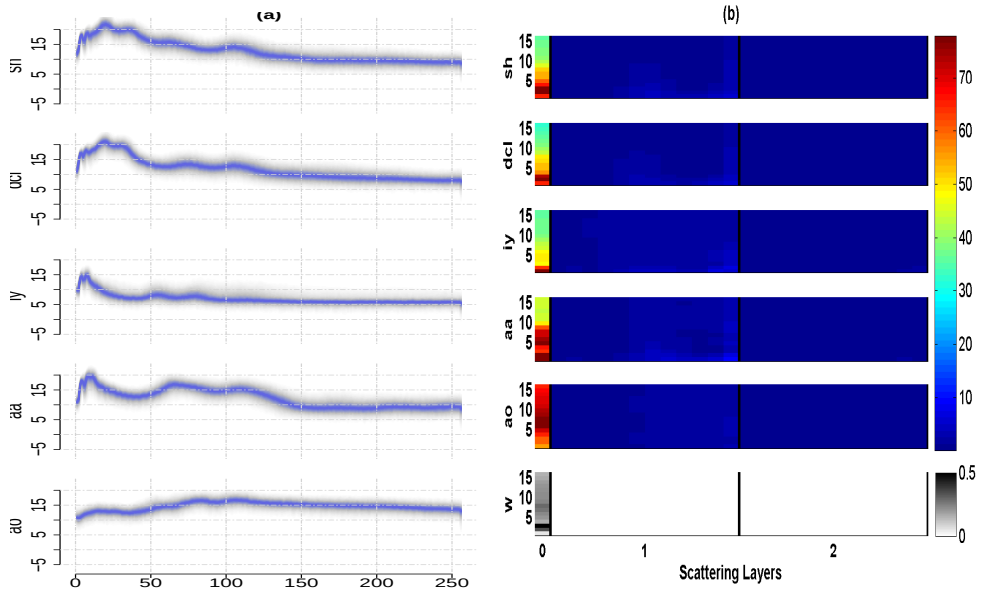

Figure 7: **Clustering results for the Phoneme dataset**. **(a)** Fanplots as in Figure 5. **(b)** Heat maps for different clusters and for weighs as in Figure 5. Most of the clustering information is in the 0-th layer.

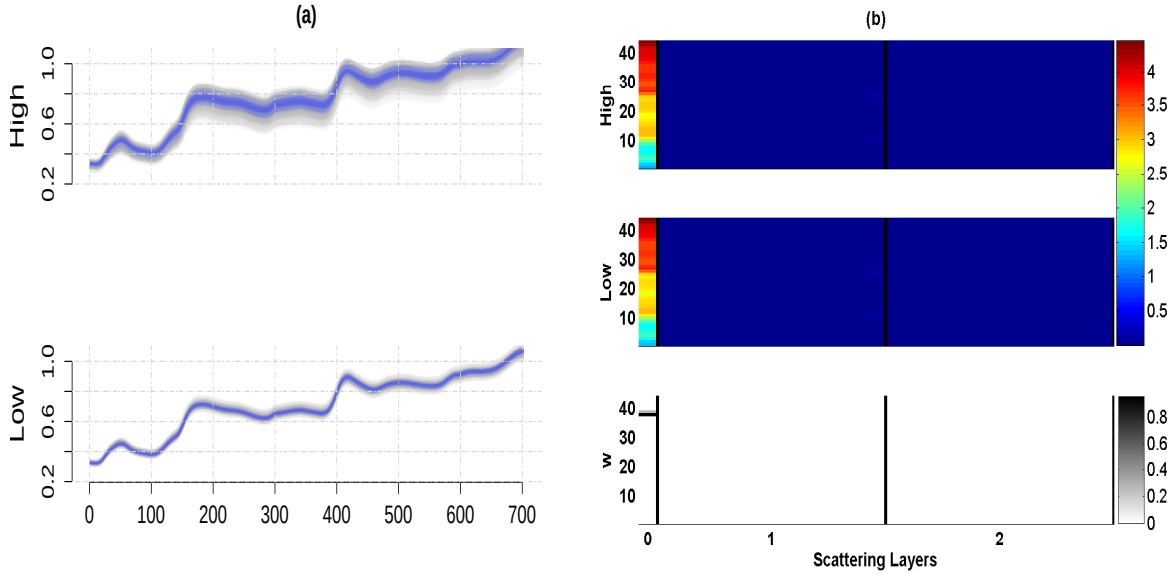

Figure 8: **Clustering results for the Wheat dataset**. **(a)** Fanplots as in Figure 5. **(b)** Heat maps for different clusters and for weighs as in Figure 5. Most of the clustering information is in the 0-th layer.

Table 2: Parameters for running Scattering transform for each of the three real datasets. We used $L = 2$ in all datasets. The different sizes T of the original vectors lead to different sizes for the scattering feature vectors $\mathbf{v}$. In the *scatnet* software implementation, a few of the scattering coefficients $v^{(l)}_{j_1,..,j_l}(k\Delta_t)$ are filtered. The choice of the frequency resolutions a_1 and a_2 were different for different datasets - for audio datasets (such as Phenome), it is known that higher frequency is required to represent the information in a signal [1], and we used $a_1 = 2^{1/8}$. The SPARCScatter method selects only a few of the scattering coefficients at the top two layers.

Parameter	**Growth**	**Phoneme**	**Wheat**
n	92	4507	100
T	30	256	701
K	2	5	2
L	2	2	2
$\|\mathbf{v}\|$	42	400	924
$\|\|\mathbf{w}\|\|_0$	10	16	2
Δ_t	32	32	32
(a_1, a_2)	$(2^{1/2}, 2)$	$(2^{1/8}, 2)$	$(2^{1/8}, 2)$