The engine behind efficient signal processing is a good signal model that identifies and exploits the underlying data structure. The intrinsic geometric structure of the data can be represented using a weighted graph.

**Our Goal:** incorporate the graph structure into the representative model – while being data adaptive and keeping it scalable for high dimensional data.

---

**Graph-Haar Wavelet Construction**

- Recursively partition the graph:
  - Spectral bisection using the Fiedler vector

  - \( \Omega_1 = \{ i | \phi^T \psi[i] \geq 0 \} \)
  - \( \Omega_2 = \{ i | \phi^T \psi[i] < 0 \} \)

- Construct an orthogonal graph-Haar wavelet basis, relying on the obtained partition

\[
\phi_i = \begin{cases} 
\frac{1}{\sqrt{N}} & i \in \Omega_1 \\
\frac{1}{\sqrt{N}} \sqrt{\frac{\phi^T \psi[i]}{\phi^T \psi[\Omega_1]}} & i \in \Omega_2 \\
0 & \text{else}
\end{cases}
\]

---

**Graph Sparse-Dictionary Learning (GSDL)**

Introducing double sparsity [1]: learning sparse combinations of the graph-Haar wavelet functions:

\[
\begin{align*}
\text{arg min}_{A, X} & \quad \|Y - \Phi AX\|_F^2 \\
\text{s.t.} & \quad \|x_i\|_0 \leq T \quad \forall i \\
& \quad \|A_i\|_0 \leq P \quad \forall j, \quad \|\Phi A_j\|_2 = 1 \quad \forall j
\end{align*}
\]

- \( Y = \) data matrix, \( \Phi = \) base dictionary,
- \( A = \) overcomplete sparse matrix,
- \( X = \) sparse representation matrix

Fusion of the analytic and trainable paradigms:

- Multi-scale nature
- Topology aware
- Data adaptive
- Efficient implementation and deployment (orthogonal + sparse)
- Less parameters: applicable for higher dimensions

---

Let’s Make it even more adaptive...

**Top-Used Atom Visualization**

Do the learned functions possess the desired multi-scale appearance? Compare with K-SVD [2], the graph Polynomial dictionary [3] and DGRDL [4].

---

**References**


---

**Performance in Signal Processing: Synthetic Data**

For piecewise-smooth signals on a medium graph, GSDL outperforms K-SVD [2], the Polynomial dictionary [3], DGRDL [4] and a direct use of \( \Phi \).

![Performance Graph](image-url)

For high dimensional data, most other methods become infeasible and could not be compared, yet GSDL remains efficient!

---

**Performance in Signal Processing: Flickr Data**

Analyzing the daily number of distinct Flickr users uploading photos near Trafalgar Square, London.

![Performance Graph](image-url)