A LEARNING APPROACH FOR WAVELET DESIGN

Dhruv Jawali[1], Abhishek Kumar[2], and Chandra Sekhar Seelamantula[2]
[1]National Mathematics Initiative, [2]Department of Electrical Engineering
Indian Institute of Science, Bangalore - 560012, India
Email: {dhruv13, kabhishek}@iisc.ac.in, chandra.sekhar@ieee.org

1. Motivation

- Multiresolution approximation (MRA) of signals, closely related to wavelets, finds many applications in signal analysis and image processing.
- A real-world application of wavelets involves choosing one among many families of analytically derived wavelets based on properties such as regularity, number of vanishing moments, compact support, symmetry, and ease of implementation.
- A learning framework for wavelet design allows one to pose the problem of searching for the 'ideal' wavelet for a given task as an optimization problem.

2. Wavelets and Filterbanks

- An orthonormal MRA has a scaling function $\phi(t)$, which satisfies the two-scale equation: $\hat{\phi}(2\omega) = \frac{1}{\sqrt{2}}\hat{\phi}(\omega)\hat{h}(\omega)$ [1], where the sequence $h[n]$ satisfies:

$$|\hat{h}(\omega)|^2 + |\hat{h}(\omega+\pi)|^2 = 2.$$

- Under certain conditions [2], it is possible to obtain the scaling $\phi(t)$ and wavelet $\psi(t)$ functions of an MRA from the filter $h[n]$ using:

$$\hat{\phi}(\omega) = \prod_{p=1}^{\infty} \frac{\hat{h}\left(e^{\mathrm{j}2^{-p}\omega}\right)}{\sqrt{2}}; \qquad \hat{\psi}(\omega) = \frac{1}{\sqrt{2}} e^{-\mathrm{j}\frac{\omega}{2}} \hat{h}^*\left(\frac{\omega}{2}+\pi\right)\hat{\phi}\left(\frac{\omega}{2}\right).$$

- The wavelet function $\psi(t)$ generates a basis of $L^2(\mathbb{R})$ through its shifts and scales.
- A wavelet $\psi(t)$ has p vanishing moments if, $\int_{-\infty}^{\infty} t^k \psi(t)dt = 0$ for $k = 0, \ldots, p-1$.
- Wavelets having p vanishing moments annihilate polynomials up to order $p-1$. Thus, more the number of vanishing moments, sparser are the representations of regular signals.
- A wavelet MRA can be viewed as analyzing using a perfect reconstruction filter bank (PRFB) [2], shown in Figure 1.
- Perfect reconstruction conditions:
  - PR1: No Distortion: $\hat{h}^*(e^{\mathrm{j}\omega})\hat{\tilde{h}}(e^{\mathrm{j}\omega}) + \hat{g}^*(e^{\mathrm{j}\omega})\hat{\tilde{g}}(e^{\mathrm{j}\omega}) = 2,$
  - PR2: Alias Cancellation: $\hat{h}^*(e^{\mathrm{j}(\omega+\pi)})\hat{\tilde{h}}(e^{\mathrm{j}\omega}) + \hat{g}^*(e^{\mathrm{j}(\omega+\pi)})\hat{\tilde{g}}(e^{\mathrm{j}\omega}) = 0.$
- A PRFB corresponds to a wavelet if, $\hat{g}(e^{\mathrm{j}\omega})|_{\omega=0} = 0 \Leftrightarrow \hat{h}(e^{\mathrm{j}\omega})|_{\omega=\pi} = 0.$

3. Our Contributions

- Viewing the PRFB (Figure 1) as a convolutional autoencoder (Figure 2), thus transforming the wavelet design problem into a learning problem.
- Filterbanks generating compactly supported orthogonal wavelets are learnt, with a given length of the filter L, and an arbitrary number of vanishing moments $p \leq \frac{L}{2}$.
- Vanishing moments are introduced in the learnt wavelet by restricting the filters in our model to have p roots at $\omega = \pi$ (Figure 3).
- PRFBs are learnt up to machine precision.

4. PRFBs as Convolutional Autoencoders

- We establish the similarity between PRFBs and convolutional autoencoders [3].
- Optimization problem: $h^*, \tilde{h}^*, g^*, \tilde{g}^* = \arg\min_{h,\tilde{h},g,\tilde{g}} \mathcal{L}(\mathbf{X}; h, \tilde{h}, g, \tilde{g}).$
- Cost function: $\mathcal{L}(\mathbf{X}; h, \tilde{h}, g, \tilde{g}) = \frac{1}{N}\sum_{i=1}^{N} \|\mathbf{x}_i - \tilde{\mathbf{x}}_i\|_2^2.$
- Dataset: $\mathbf{X} = \{\mathbf{x}_i \in \mathbb{R}^m\}_{1\leq i \leq N}$, $\mathbf{x}_i \in \mathcal{N}(0, I_{m\times m}).$
- Conjugate mirror filter (CMF) constraints: $\tilde{h}[n] = h[n], \tilde{g}[n] = g[n], g[n] = (-1)^{(1-n)}h[(2l+1)-n]$, where $l \in \mathbb{Z}$.

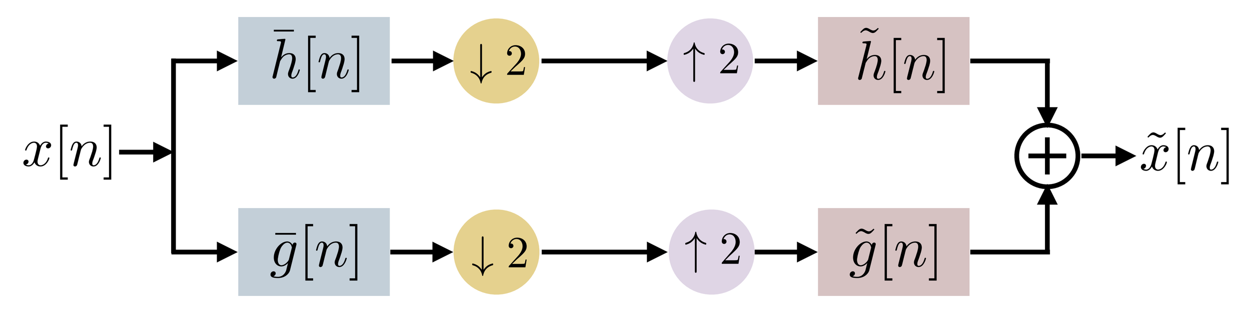


Figure 1. A perfect reconstruction filterbank (PRFB).

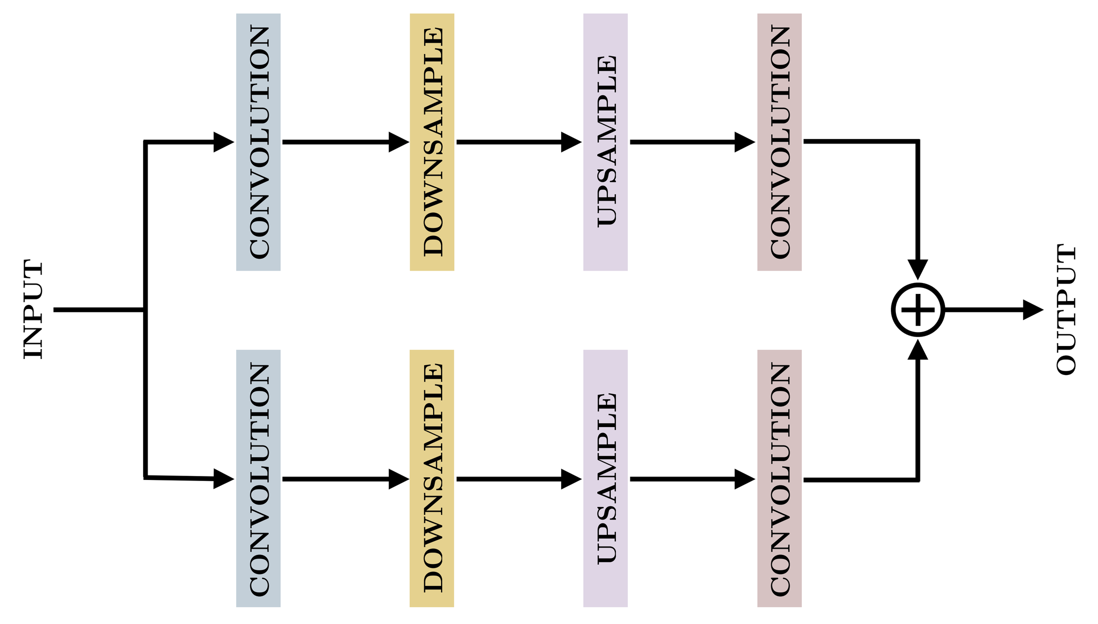


Figure 2. PRFB as a convolutional autoencoder.

- Incorporating vanishing moments:

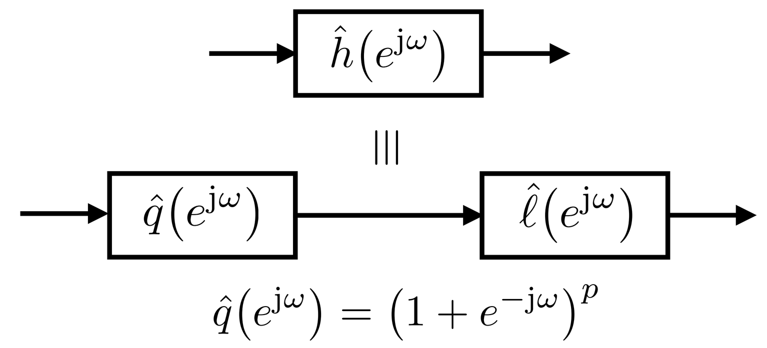


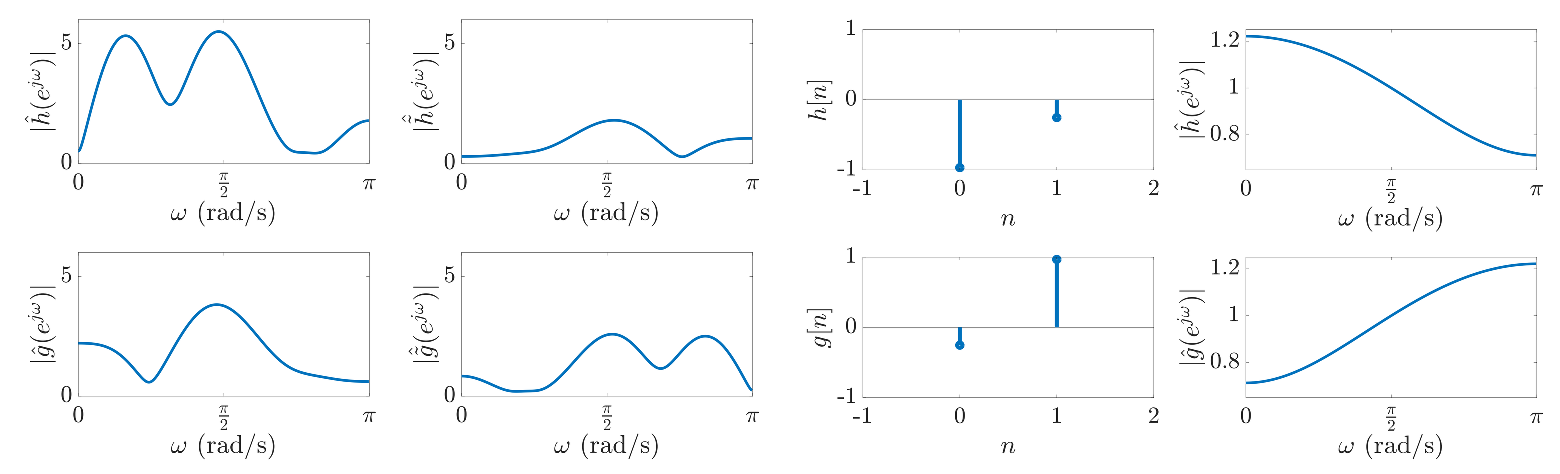
Figure 3. p vanishing moments can be imposed on the filter $h[n]$ by replacing it with a cascade of two filters $q[n]$ and $l[n]$ as shown. Only the coefficients of the filter $l[n]$ are learnt.

5. Implementation Details

- We used the Tensorflow Python library to implement the architecture.
- The Adam optimizer [4] is used to carry out gradient-descent optimization of the cost function.
- Convergence criteria:
  - The training loss goes below 10^{-28}; or
  - No change in the training loss for more than 100 iterations.
- Performance metric:
  - Signal-to-reconstruction-error ratio (SRER) $= 20\log_{10}\left\langle \frac{\|x\|_2}{\|x-\tilde{x}\|_2}\right\rangle$ dB.

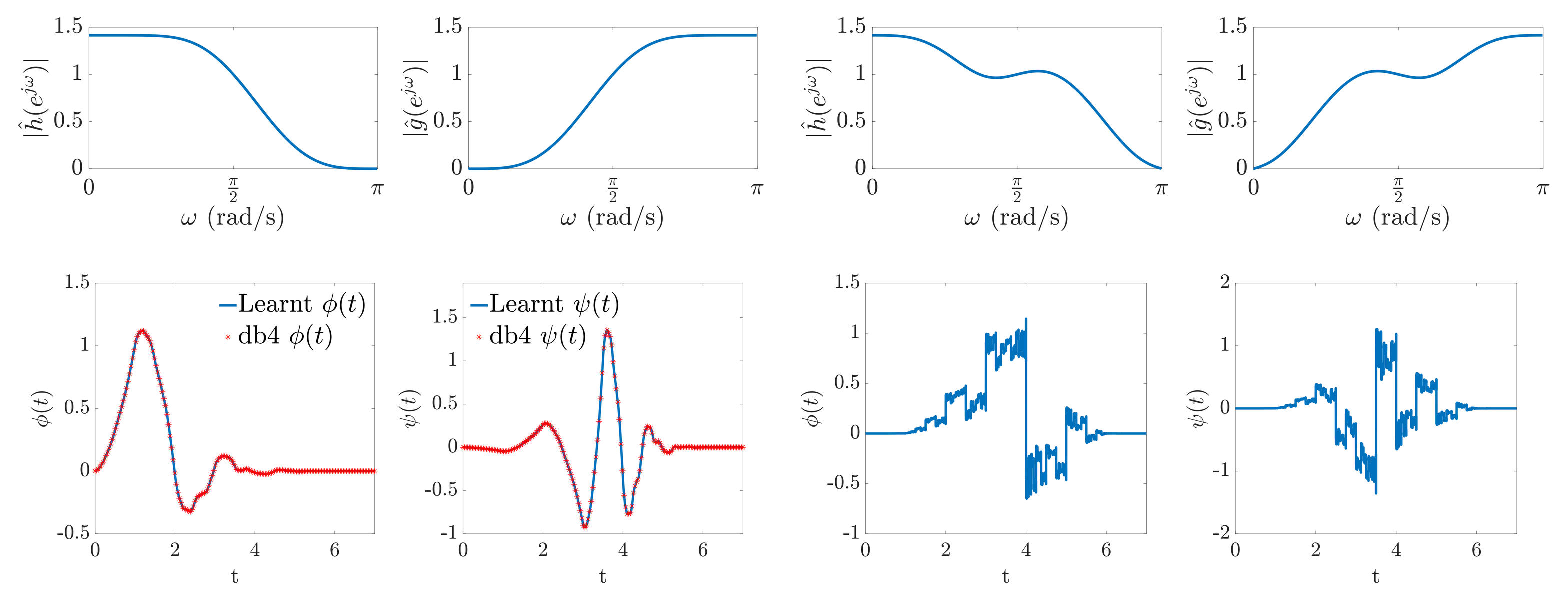
6. Results

Learnt PRFBs not resulting in a wavelet



- **Unconstrained:** Four filters of length 8 each are learnt.
  - None of the filters have zero magnitude at either $\omega = 0$ or $\omega = \pi$.
  - SRER = 204.83 dB.
- **CMF constraints:** One filter of length 2 is learnt.
  - $\hat{h}(\omega)$ takes nonzero value at $\omega = \pi$.
  - SRER = 211.40 dB.

Learnt PRFBs resulting in a wavelet



- **db4 wavelet [5]:** Length $L = 8$, $p = 4$ roots at $\omega = \pi$.
  - The learnt scaling and wavelet functions matches exactly with the db4 scaling and wavelet functions, respectively.
  - SRER = 236.58 dB.
- **A learnt wavelet that does not belong to the Daubechies family:** Length $L = 8$, $p = 4$ roots at $\omega = \pi$.
  - The proposed approach can also generate wavelets with an arbitrary number of vanishing moments.
  - SRER = 228.16 dB.

References

1. S. Mallat, "Multiresolution approximations and wavelet orthonormal bases of $L^2(\mathbb{R})$," Transactions of the American Mathematical Society, vol. 315, no. 1, pp. 69–87, 1989.
2. S. Mallat, A Wavelet Tour of Signal Processing: The Sparse Way. Academic press, 2008.
3. P. Baldi, "Autoencoders, unsupervised learning, and deep architectures," in Proceedings of ICML Workshop on Unsupervised and Transfer Learning, 2012, pp. 37–49.
4. D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980, 2014.
5. I. Daubechies, Ten lectures on wavelets. vol. 61, SIAM, 1992.

Acknowledgements: Indian Institute of Science, Pratiksha Trust, and IEEE SP Chapter Bangalore for travel fellowships.