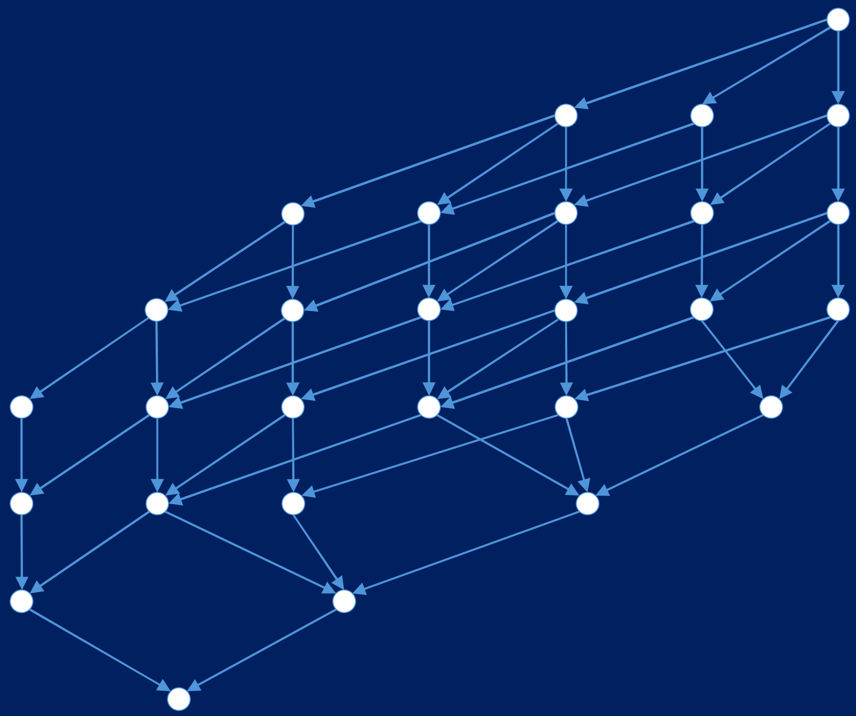
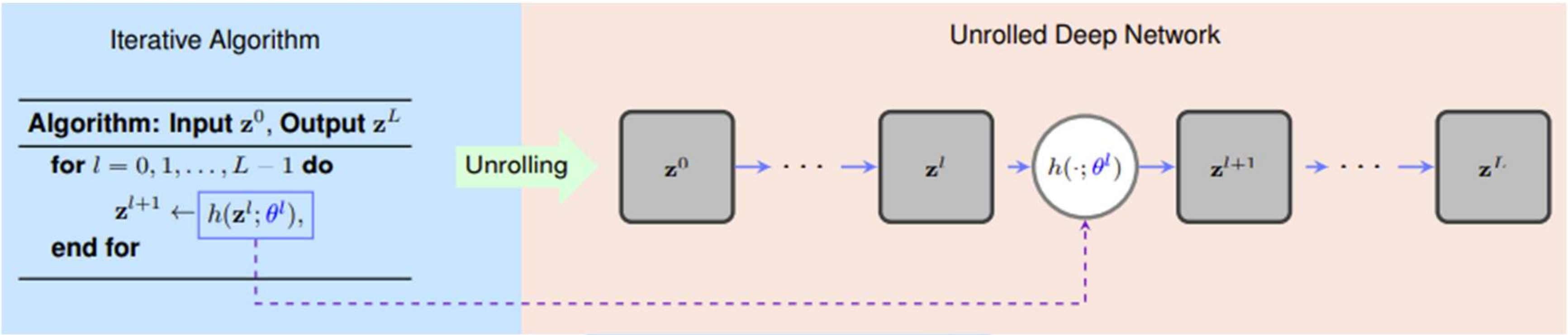
Graph Signal Processing

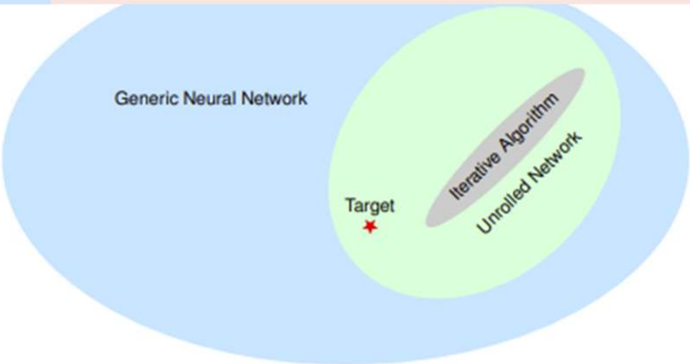
Bastian Seifert

Sommersemester 2025



Algorithm Unrolling Networks





Figures: Algorithm Unrolling: Interpretable, Efficient Deep Learning for Signal and Image Processing, Monga et al.

Network layers inherit interpretability from the iteration procedure

Literature

Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing 1406 2021
V Monga, Y Li, YC Eldar
IEEE Signal Processing Magazine 38 (2), 18-44

Graph unrolling networks: Interpretable neural networks for graph signal denoising 93 2021
S Chen, YC Eldar, L Zhao
IEEE Transactions on Signal Processing

Porting Signal Processing from Undirected to Directed Graphs: Case Study Signal Denoising with Unrolling Networks 1 2022
V Mihal, B Seifert, M Püschel
30th European Signal Processing Conference (EUSIPCO), 2076-2080

Sparsity

A signal representation is sparse if it has few non-zero components

$$\mathbf{x} = \sum_{k=0}^{K} a_k \mathbf{r}_k \text{ where } K \ll n$$

Sparsity can be measured by the ℓ_0 pseudo-norm

$$||\mathbf{x}||_0 = |\{x_i \mid x_i \neq 0\}|$$

which can be approximated by L1-norm

$$||\mathbf{x}||_1 = |x_1| + \ldots + |x_n|$$

Sparse coding

Given a set of signals

$$X = [\mathbf{x}_1, \ldots, \mathbf{x}_P]$$

we want to find a dictionary

$$D = [\mathbf{d}_1, \ldots, \mathbf{d}_K]$$

such that sparse signal representations

$$R = [\mathbf{r}_1, \ldots, \mathbf{r}_P]$$

are optimal approximations, i.e.,

$$||X - DR||_2^2$$

is minimized. So we have the optimization problem

$$\operatorname{argmin}_{D,R} \sum_{i=1}^{P} ||\mathbf{x}_i - D\mathbf{r}_i||_2^2 + \lambda ||\mathbf{r}_i||_0$$

The sparse coding problem can typically be either solved for D or R

Graph sparse coding

Assume the signal representations are the denoised signal, i.e., the signals are given as sparse signal propagated through a filter plus noise:

$$\mathbf{t}_i = \mathbf{x}_i + \mathbf{e}_i \text{ s.t. } \mathbf{x}_i = H\mathbf{s}_i$$

Then the graph sparse coding problem is the optimization problem

$$\min_{\mathbf{s}\in\mathbb{R}^n} \frac{1}{2}\|\mathbf{t}-\mathbf{x}\|_2^2 + \alpha\|\mathbf{s}\|_1, \quad \text{s.t. } \mathbf{x} = H\mathbf{s}$$

That is the dictionary is given by the graph filter.

Graph filter

Half-quadratic splitting

$$L(\mathbf{x}, \mathbf{s}, \mathbf{z}) = \frac{1}{2}\|\mathbf{t} - \mathbf{x}\|_2^2 + \alpha\|\mathbf{z}\|_1 + \frac{\mu_1}{2}\|\mathbf{x} - H\mathbf{s}\|_2^2 + \frac{\mu_2}{2}\|\mathbf{z} - \mathbf{s}\|_2^2$$

The variable z separates linear from non-linear part

Iterative algorithm

$$\mathbf{x} \leftarrow \frac{1}{1+\mu_1}(\mathbf{t} + \mu_1 H\mathbf{s}),$$
$$\mathbf{s} \leftarrow (\mu_1 H^T H + \mu_2 I)^{-1}(\mu_1 H^T \mathbf{x} + \mu_2 \mathbf{z}),$$
$$\mathbf{z} \leftarrow \text{argmin}_{\mathbf{z}}\|\mathbf{z}\|_1,$$

$\leftarrow$ softtresh$(z, \theta) = \text{sign}(z)(|z| - \theta)$

Goal: Learn step-sizes and filter from data

Unrolled Network

K_1 input signals $X = [\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(K_1)}] \in \mathbb{R}^{n \times K_1}$

K_2 output signals $Y = [\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(K_2)}]$

5-mode tensor

$$\mathbf{y}^{(k_2)} = \sum_{\ell=1}^{L} \sum_{k_1=1}^{K_1} (\mathbb{H}_{(\ell,k_1,k_2)} \odot A^{\ell})\mathbf{x}^{(k_1)}$$

$\mathbb{H} *_a X$ Trainable graph convolution

Matrix mit Dimension von A

Parametrized by learnable kernel function

$$\Psi_{i,j} = \psi_w(\mathbf{v}_j - \mathbf{v}_i) \in \mathbb{R}$$

based on truncated graph Fourier basis

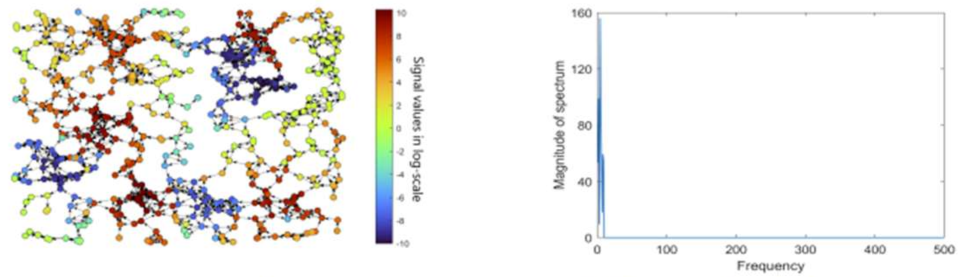
$$\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$$

Unrolled Network

$$\Sigma = [\mathbf{s}^{(1)}, \ldots, \mathbf{s}^{(k)}]$$

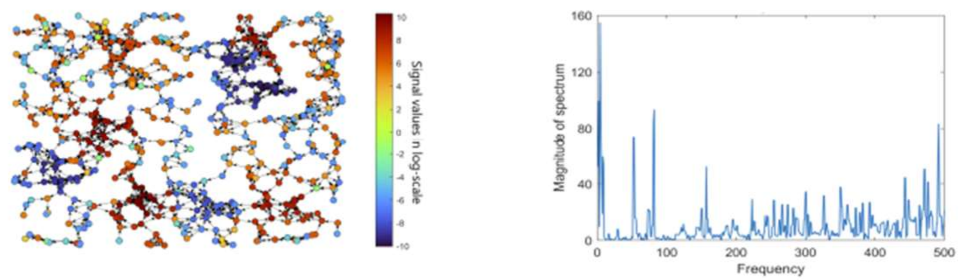
Use layers of the form

$$\begin{aligned}
X^{(b)} &\leftarrow \mathbb{A} *_a T + \mathbb{B} *_a \Sigma^{(b-1)}, \\
\Sigma^{(b)} &\leftarrow \mathbb{D} *_a X^{(b)} + \mathbb{E} *_a Z^{(b-1)}, \\
Z^{(b)} &\leftarrow \mathrm{thresh}_\alpha(\Sigma^{(b)}),
\end{aligned}$$



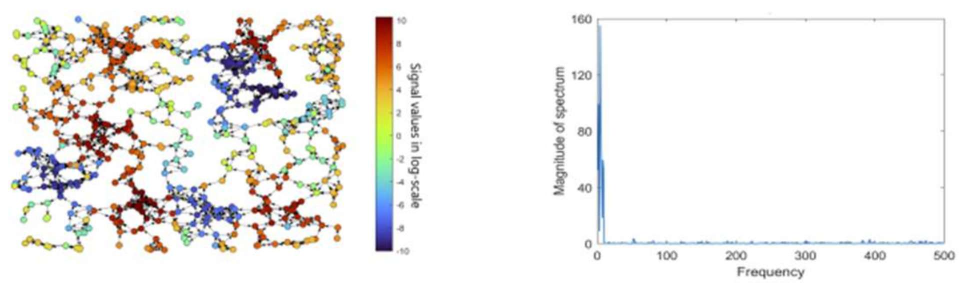
(a) Original signal (b) Spectrum (GenBC2) of the original signal

Fig. 2: Original signal (a) on graph, and (b) its spectrum.



(a) Noisy signal (b) Spectrum (GenBC2) of the noisy signal

Fig. 3: Noisy version of the signal in Figure 2 with an SNR ≈ 2.



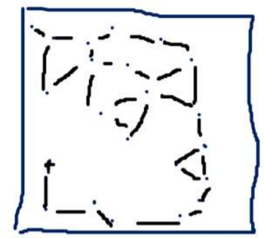
(a) Denoised signal (b) Spectrum (GenBC2) of the denoised signal

Other algorithms can be unrolled similar
(e.g., graph trend filtering)

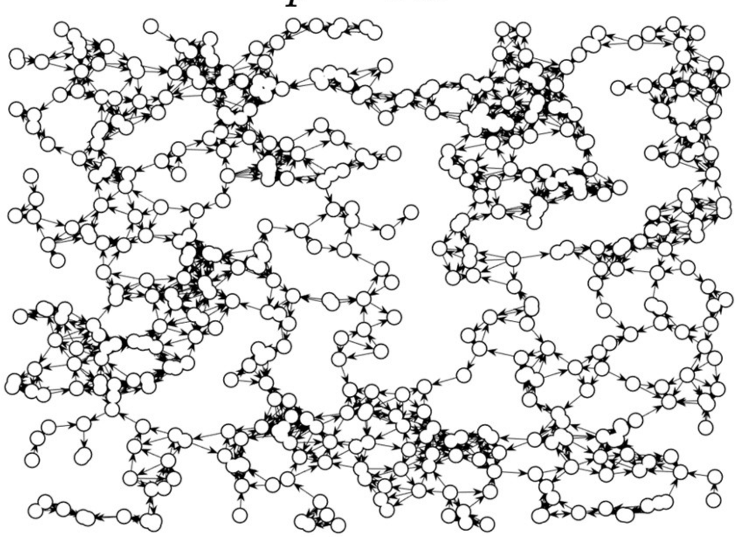
Experimental Setup: Digraphs Used

2 random geometric digraphs:

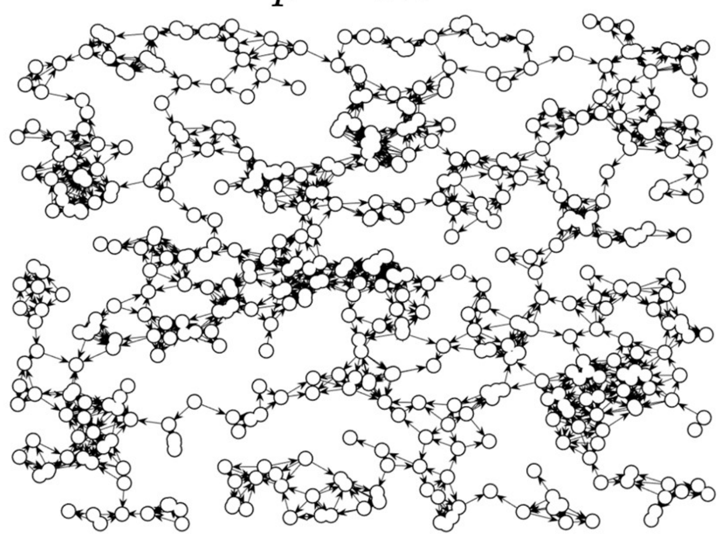
- 500 random points in $[0,1] \times [0,1]$
- connected if distance less than 0.065
- edge direction determined with probability $p\ (0.5, 0.9)$



$p = 0.5$



$p = 0.9$



Experimental Setup – digraph signals for denoising

For each basis:

- 6 sets of 100 low frequency signals
  (low frequency notion depends on basis!)
- one set for hyperparameter tuning, 5 sets for testing

Noise:

- additive Gaussian noise
- NMSE: $\frac{\|\mathbf{t}-\mathbf{x}\|_2^2}{\|\mathbf{x}\|_2^2} \approx 0.5$

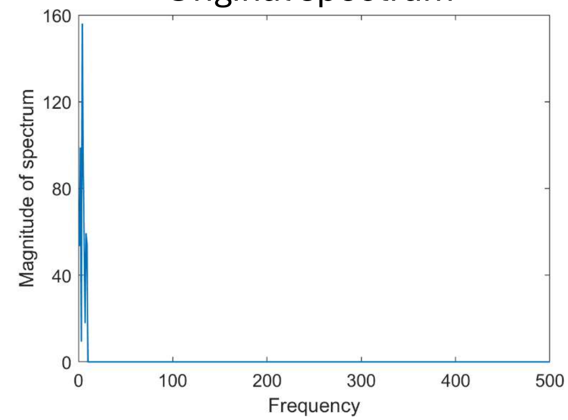
For each basis:

- apply 5 (4) methods

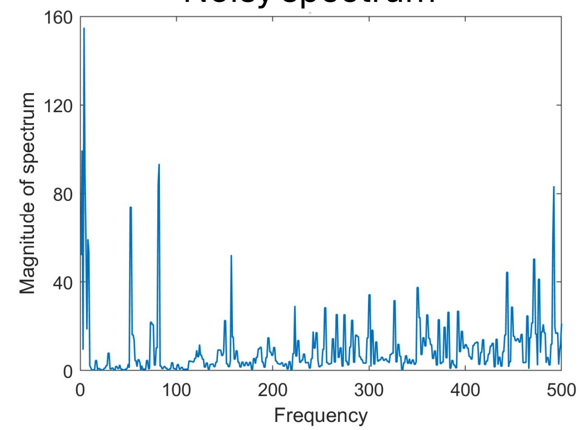
Total:

- 5 + 5 + 5 + 4 + 4 = 23 experiments

Original spectrum



Noisy spectrum



Results

ForgetDir never performs best

SelfLoops is close to AssocShift for GenBC

p = 0.5	ForgetDir	AdHoc	SelfLoops	UndShift	AssocShift
SprFreq	0.43 ± 0.01	0.46 ± 0.01	0.35 ± 0.01	0.19 ± 0.00	-
HermL	0.28 ± 0.01	0.31 ± 0.01	0.19 ± 0.00	0.07 ± 0.01	0.19 ± 0.04
StabApp	0.13 ± 0.00	0.08 ± 0.00	0.09 ± 0.00	0.10 ± 0.00	-
GenBC1	0.10 ± 0.00	0.01 ± 0.00	0.01 ± 0.00	0.03 ± 0.00	0.01 ± 0.00
GenBC2	0.10 ± 0.00	0.01 ± 0.00	0.01 ± 0.00	0.03 ± 0.00	0.01 ± 0.00

GenBC basis and shift work best

p = 0.9	ForgetDir	AdHoc	SelfLoops	UndShift	AssocShift
SprFreq	0.19 ± 0.01	0.19 ± 0.01	0.12 ± 0.00	0.06 ± 0.00	-
HermL	0.21 ± 0.01	0.28 ± 0.01	0.16 ± 0.01	0.07 ± 0.00	0.19 ± 0.01
StabApp	0.24 ± 0.00	0.10 ± 0.00	0.11 ± 0.00	0.11 ± 0.00	-
GenBC1	0.15 ± 0.01	0.03 ± 0.00	0.03 ± 0.01	0.06 ± 0.01	0.05 ± 0.00
GenBC2	0.15 ± 0.00	0.06 ± 0.00	0.06 ± 0.00	0.05 ± 0.00	0.01 ± 0.00

GenBC basis and shift work best