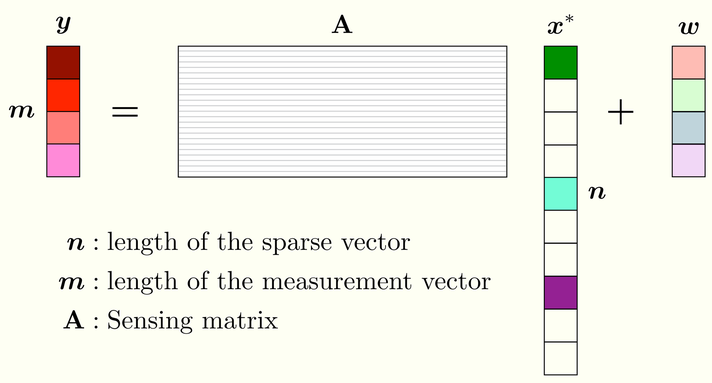
Some Intriguing Observations on the Learnt Matrices of the Deep Unfolded Networks

Kartheek Kumar Reddy Nareddy †§
Inbasekaran Perumal ‡
Chandra Sekhar Seelamantula †

† Department of Electrical Engineering Indian Institute of Science Bengaluru
§ Institute of Data Science, German Aerospace Center (DLR), 07745 Jena, Germany
‡ Department of Electronics and Communication Engineering National Institute of Technology Karnataka Mangaluru

1. Sparse recovery challenge

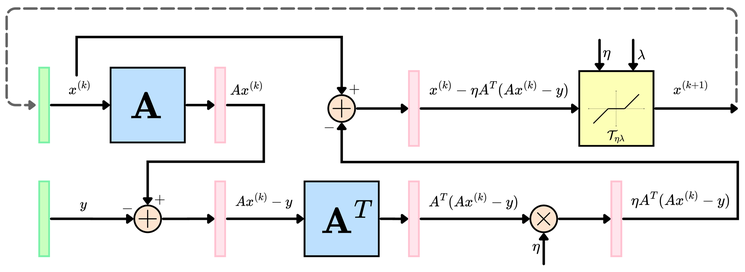
$$\boldsymbol{y} = \mathbf{A}\boldsymbol{x}^* + \boldsymbol{w}$$



$\boldsymbol{n}$: length of the sparse vector
$\boldsymbol{m}$: length of the measurement vector
$\mathbf{A}$: Sensing matrix

$$\underset{\boldsymbol{x}\in\mathbb{R}^n}{\text{minimize}} \ \frac{1}{2}\|\mathbf{A}\boldsymbol{x}-\boldsymbol{y}\|_2^2 + \lambda\|\boldsymbol{x}\|_1$$

2. ISTA vs LISTA



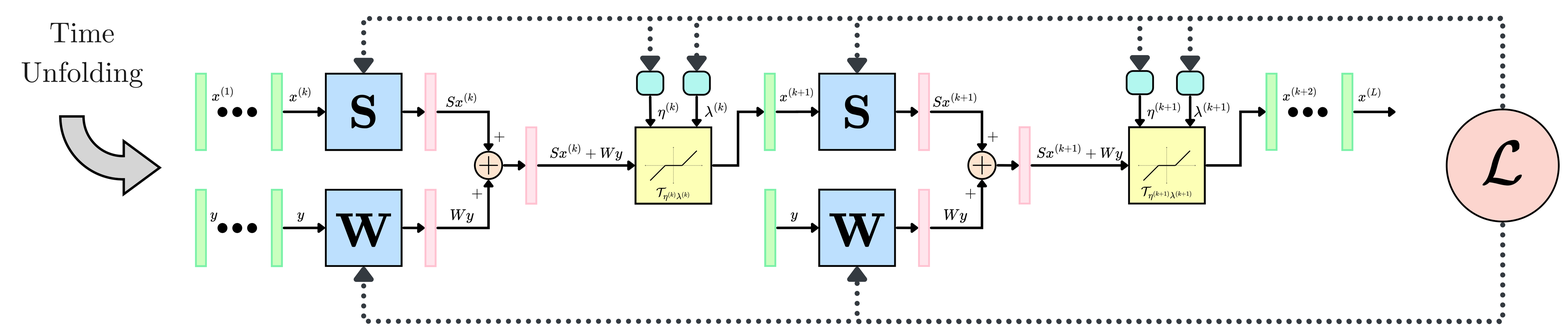
ISTA

$$z^{(k+1)} = x^{(k)} - \eta A^T(Ax^{(k)} - y)$$
$$x^{(k+1)} = T_{\eta\lambda}(z^{(k+1)})$$
$$T_{\eta\lambda}(x) = sgn(x)\cdot\max(|x| - \eta\lambda, 0)$$

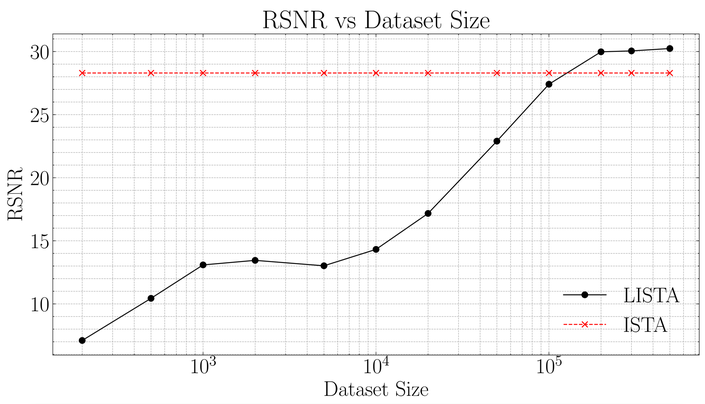
LISTA

$$z^{(k+1)} = (I - \eta A^T A)x^{(k)} + \eta A^T y$$
$$z^{(k+1)} = Sx^{(k)} + Wy$$
$$x^{(k+1)} = T_{\eta^{(k)}\lambda^{(k)}}(z^{(k+1)})$$

Time Unfolding

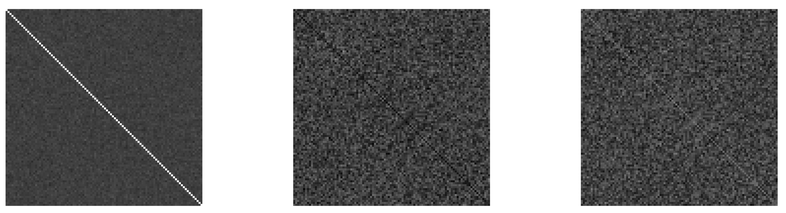


3. Effect of Training Dataset



RSNR improves with larger LISTA datasets (10^5+).

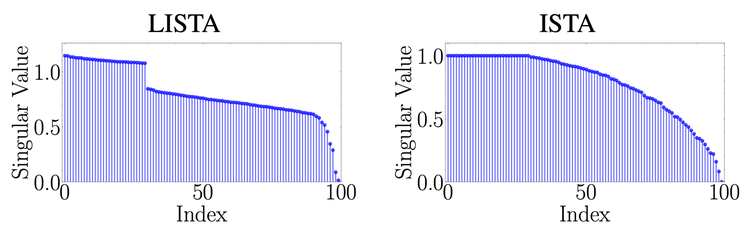
4. S Matrix Visualization



ISTA | Sign Inverted LISTA | LISTA

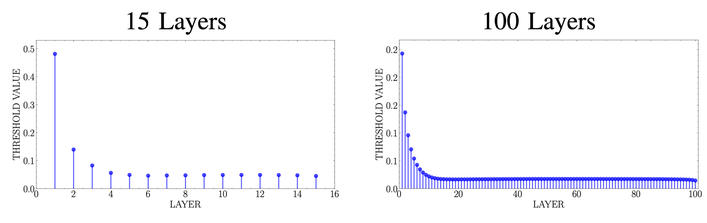
ISTA shows diagonal dominance from its S = I - ηA^{T}A structure while LISTA exhibits negative diagonal dominance.

5. SVD of S Matrix



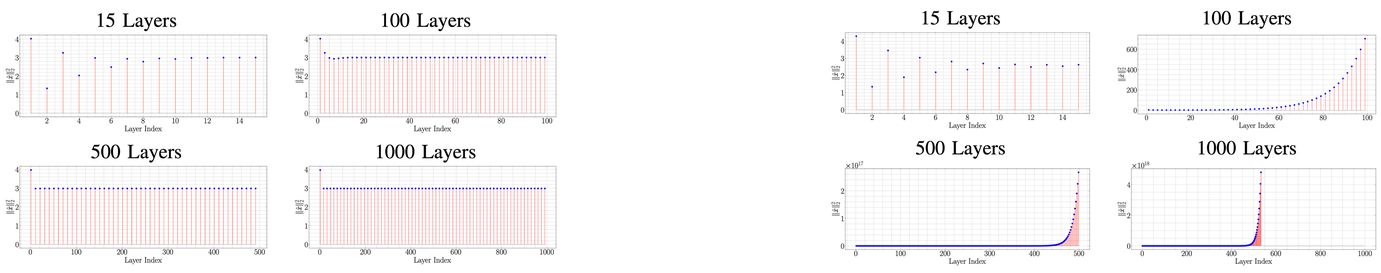
In ISTA, the singular values of the S matrix are bounded by one, while in LISTA, they exceed one, indicating potential instability.

6. Learned Threshold Values



As network depth increases, LISTA's learned thresholds decrease, indicating saturation and an adaptive strategy to balance stability and signal amplification.

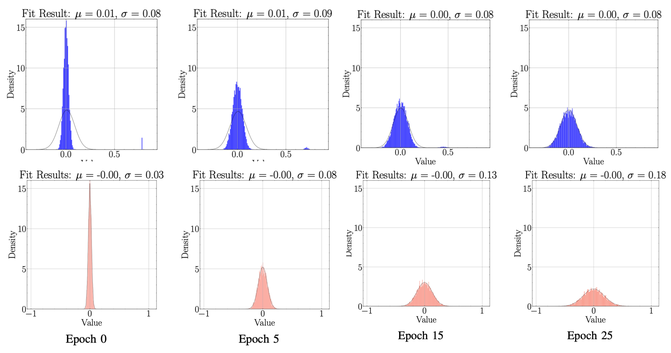
7. Impact of the Soft Threshold Function on Stability



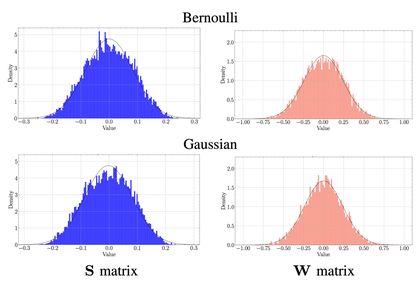
The soft threshold function keeps the squared ℓ2 norm of the reconstructed signal bounded, ensuring stability across different network depths.

Removing the soft threshold function causes the squared ℓ2 norm to grow exponentially, leading to instability as the number of layers increases

8. Evolution of Matrices

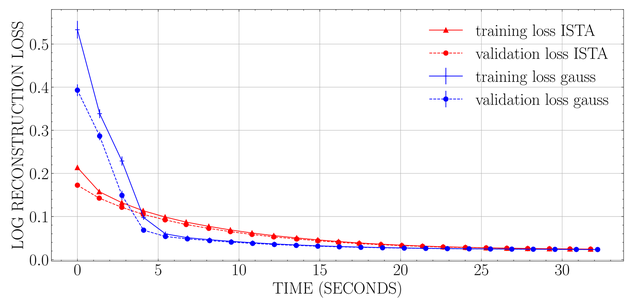


Evolution of the S and W matrices over training epochs. The weights are nearly Gaussian distributed.



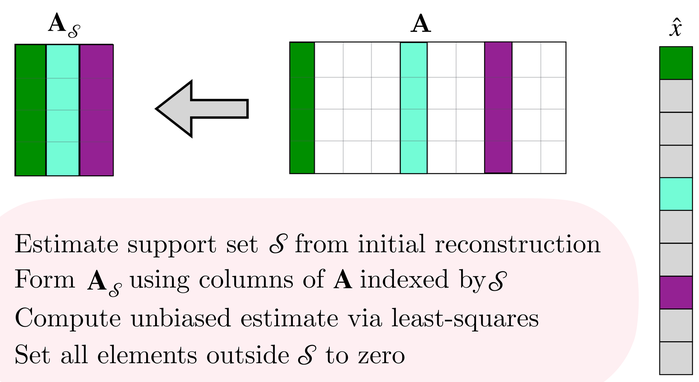
S and W matrices exhibit Gaussian like pattern independent of sensing matrix type.

10. Effect of Initialization

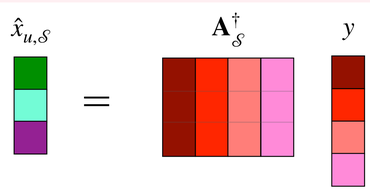


The network weights initialized to Gaussian distribution converged faster compared to ISTA-based initialization.

11. Unbiasing Technique



Estimate support set $\mathcal{S}$ from initial reconstruction
Form $\mathbf{A}_{\mathcal{S}}$ using columns of $\mathbf{A}$ indexed by $\mathcal{S}$
Compute unbiased estimate via least-squares
Set all elements outside $\mathcal{S}$ to zero



Technique	250 dB	50 dB	40 dB	30 dB
ISTA	44.5 ± 4.3	43.5 ± 4.3	37.8 ± 2.7	28.2 ± 2.5
LISTA	30.4 ± 2.3	30.5 ± 2.3	30.1 ± 2.3	26.4 ± 3.0
U-ISTA	238.3 ± 55.5	54.3 ± 3.8	41.1 ± 6.3	32.2 ± 3.7
U-LISTA	115.7 ± 52.3	51.4 ± 7.7	44.0 ± 4.2	32.6 ± 4.1

Prefix U- denotes unbiased variant.