

LISTA: Theoretical Linear Convergence, Practical Weights and Thresholds

Xiaohan Chen^{*}, Jialin Liu[†], Zhangyang Wang^{*}, and Wotao Yin[†]

^{*} TAMU, CSE

[†] UCLA, Math

NeurIPS'18

Overview

Recover **sparse** x^* from

$$b := Ax^* + \text{white noise}$$

Our methods improve on LISTA (Gregor&LeCun'10) and related work by

- learning fewer parameters (**faster training**)
- adding support detection (**faster recovery**)
- proving linear convergence and robustness (**theoretical guarantee**)

Review: ISTA and LISTA

ISTA (iterative soft thresholding)

$$x^{(k+1)} = \text{SoftThreshold}_\theta \left(x^{(k)} + \alpha A^T (b - Ax^{(k)}) \right).$$

α, θ are chosen by hand or cross validation.

Review: ISTA and LISTA

ISTA (iterative soft thresholding)

$$x^{(k+1)} = \text{SoftThreshold}_\theta \left(x^{(k)} + \alpha A^T (b - Ax^{(k)}) \right).$$

α, θ are chosen by hand or cross validation.

LISTA (Learned ISTA)

$$x^{(k+1)} = \text{SoftThreshold}_{\theta^k} \left(W_1^k b + W_2^k x^{(k)} \right).$$

θ^k, W_1^k, W_2^k are chosen by stochastic optimization

$$\underset{\{\theta^k, W_1^k, W_2^k\}}{\text{minimize}} \left\{ \mathbb{E}_{x^*, b} \|x^K(b) - x^*\|^2 \right\}$$

using synthesized (x^*, b) obeying $b = Ax^* + \text{white noise}$.

Review: ISTA and LISTA

ISTA (iterative soft thresholding)

$$x^{(k+1)} = \text{SoftThreshold}_{\theta} \left(x^{(k)} + \alpha A^T (b - Ax^{(k)}) \right).$$

α, θ are chosen by hand or cross validation.

LISTA (Learned ISTA)

$$x^{(k+1)} = \text{SoftThreshold}_{\theta^k} \left(W_1^k b + W_2^k x^{(k)} \right).$$

θ^k, W_1^k, W_2^k are chosen by stochastic optimization

$$\underset{\{\theta^k, W_1^k, W_2^k\}}{\text{minimize}} \left\{ \mathbb{E}_{x^*, b} \|x^K(b) - x^*\|^2 \right\}$$

using synthesized (x^*, b) obeying $b = Ax^* + \text{white noise}$.

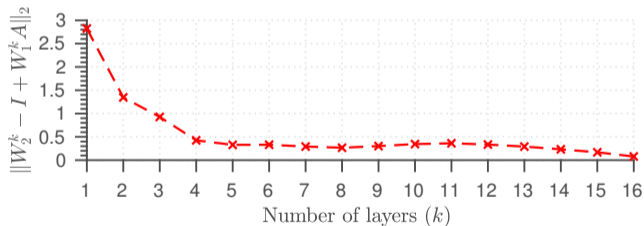
Compare: ISTA is slow, no training. LISTA is fast, difficult-to-train.

Proposed — coupled LISTA

LISTA-CP: couple W_1^k and W_2^k via

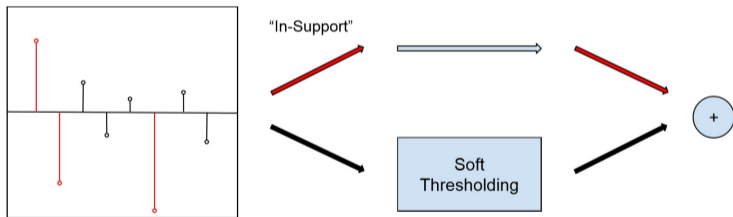
$$W_1^k A + W_2^k = I.$$

We show: $x^{(k)} \rightarrow x^*$ implies this relation to hold asymptotically.



LISTA-CPSS: support selection

- Only the large coordinates pass activations to the next iteration.
- Ideas from Linearized Bregman iteration (kicking)¹ and Fixed-Point Continuation method (FPC)².



¹Stanley Osher et al. '2011

²Elaine Hale, Wotao Yin, Yin Zhang '2008

Robust global linear convergence

Theorem

Fix A , sparsity level s , and noise level σ .

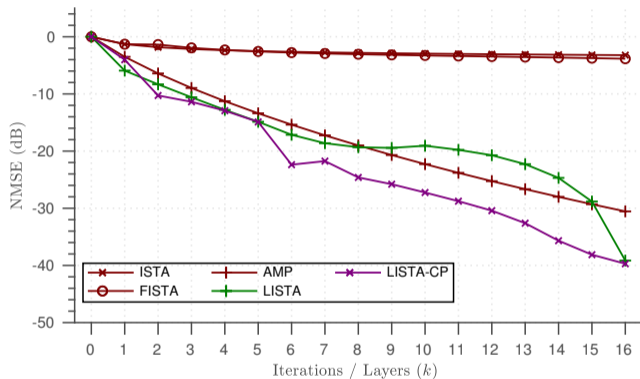
There exist $\{\theta^k, W_1^k\}$ such that LISTA-CP obeys

$$\|x^{(k)} - x^*\|_2 \leq sC_1 e^{-C_2 k} + C_3 \sigma, \quad k = 1, 2, \dots$$

where $C_1, C_2, C_3 > 0$ are constants.

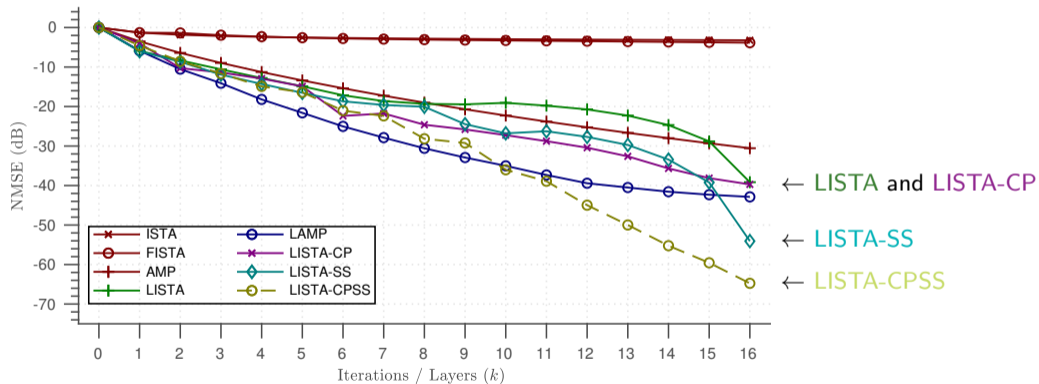
LISTA-CPSS improves the constants C_2, C_3 .

Weight coupling test



- CP can stabilize intermediate results.
- CP will not hurt final recovery performance.

Support selection test (no noise)



Thank you!

10:45 AM – 12:45 PM
Room 210 & 230 AB #163

Welcome to our poster for more details!