ELECTRICAL ENGINEERING AND COMPUTER SCIENCE
UNIVERSITY of MICHIGAN

The accuracy of singular vectors of thresholded low-rank plus noise plus outlier matrices

Brian E. Moore and Raj Rao Nadakuditi

Model

- Signal-plus-noise-plus-outliers model:

$$\widetilde{X} = \theta u v^T + S + G \in \mathbb{R}^{m\times n}$$

  - $L = \theta u v^T$ is a low-rank matrix
  - S is a sparse outliers matrix

$$S_{ij} = \begin{cases} Q_{ij} & \text{with probability } p_s \\ 0 & \text{with probability } 1 - p_s, \end{cases}$$

  where Q_{ij} are drawn i.i.d. from an unknown distribution with

$$\mathbf{E}\, Q_{ij} = 0 \qquad \mathbf{E}\, Q_{ij}^2 = \sigma_q^2 \qquad \mathbf{E}\, Q_{ij}^4 < \infty$$

  - G is a noise matrix of i.i.d. elements

$$\mathbf{E}\, G_{ij} = 0 \qquad \mathbf{E}\, G_{ij}^2 = \sigma^2/n \qquad \mathbf{E}\, G_{ij}^4 < \infty$$

- Asymptotic regime: $m, n \to \infty$ such that $m/n \to c \in (0,1)$

Problem

- Given $\widetilde{X}$, estimate L as accurately as possible
- Applications:
  - Foreground/background separation
  - Dynamic MRI reconstruction
  - Many more...

Singular vector accuracy

Lemma (Singular vector accuracy):

Suppose $\widetilde{X}$ can be written as

$$\widetilde{X} = \alpha\left(\theta u v^T\right) + X + \Delta,$$

where X_{ij} are i.i.d. with

$$\mathbf{E}\, X_{ij} = 0 \qquad \mathbf{E}\, X_{ij}^2 = \sigma_x^2/n \qquad \mathbf{E}\, X_{ij}^4 < \infty,$$

and $\sigma_1(\Delta) \xrightarrow{a.s.} 0$. Then we have

$$|\langle v, \widetilde{v}_1\rangle|^2 \xrightarrow{a.s.} \begin{cases} 1 - \dfrac{c + \overline{\theta}^2}{\overline{\theta}^2\left(1+\overline{\theta}^2\right)} & \text{if } \overline{\theta} > c^{1/4} \\ 0 & \text{otherwise,} \end{cases}$$

with effective SNR

$$\overline{\theta} := \frac{\left(\lim_{n\to\infty} \alpha\right)\theta}{\sigma_x}$$

References

- F. Benaych-Georges and R. R. Nadakuditi. "The singular values and vectors of low rank perturbations of large rectangular random matrices." *Journal of Multivariate Analysis*, 111:120–135, 2012.
- B. E. Moore, R. R. Nadakuditi, and J. A. Fessler. "Improved robust PCA using low-rank denoising with optimal singular value shrinkage." *2014 IEEE workshop on statistical signal processing (SSP)*, 13–16, June 2014.

Oracle estimator

- Oracle estimator of L:

$$\widetilde{X}_{ij}^{\text{oracle}} = \begin{cases} \widetilde{X}_{ij} & \text{if } S_{ij} = 0 \\ 0 & \text{if } S_{ij} \neq 0 \end{cases}$$

- Replaces outlier-corrupted entries of $\widetilde{X}$ with zeros
- Result: $\widetilde{X}^{\text{oracle}}$ satisfies Lemma with

$$\overline{\theta}^{\text{oracle}} = \frac{\left(\lim_{n\to\infty}\sqrt{1-p_s}\right)\theta}{\sigma}$$

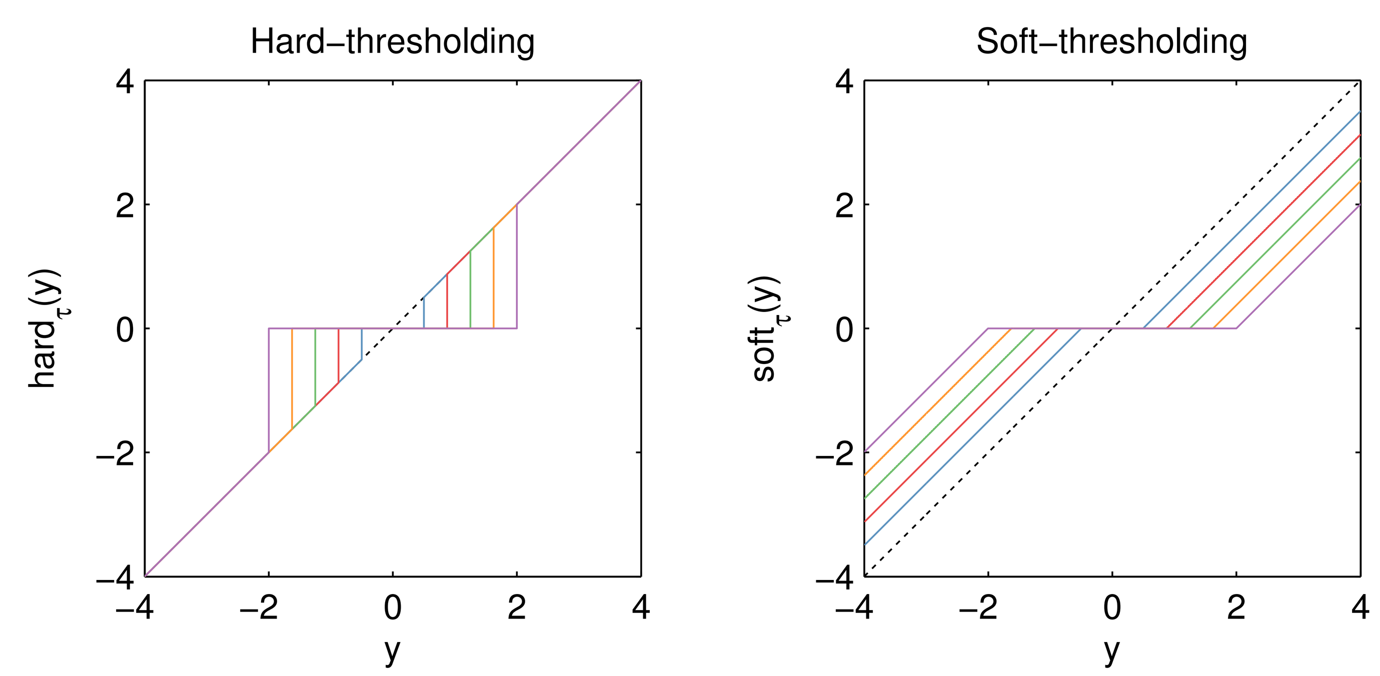
Thresholding-based estimators

- Given parameter $\tau > 0$, we consider two thresholding-based estimators of L:
  - Hard-thresholding

$$\widetilde{X}_{ij}^{\text{HT}} = \widetilde{X}_{ij} - \text{hard}_\tau(\widetilde{X}_{ij})$$

  - Soft-thresholding

$$\widetilde{X}_{ij}^{\text{ST}} = \widetilde{X}_{ij} - \text{soft}_\tau(\widetilde{X}_{ij})$$



Main result

Theorem (Sufficient conditions for equivalence of oracle and soft/hard-thresholding):

Suppose that $\tau \to 0$ such that $n\tau^2 \to \infty$ and $n\tau^3 \to 0$. Then $\widetilde{X}^{\text{HT}}$ and $\widetilde{X}^{\text{ST}}$ obey Lemma with effective SNR given by the following table:

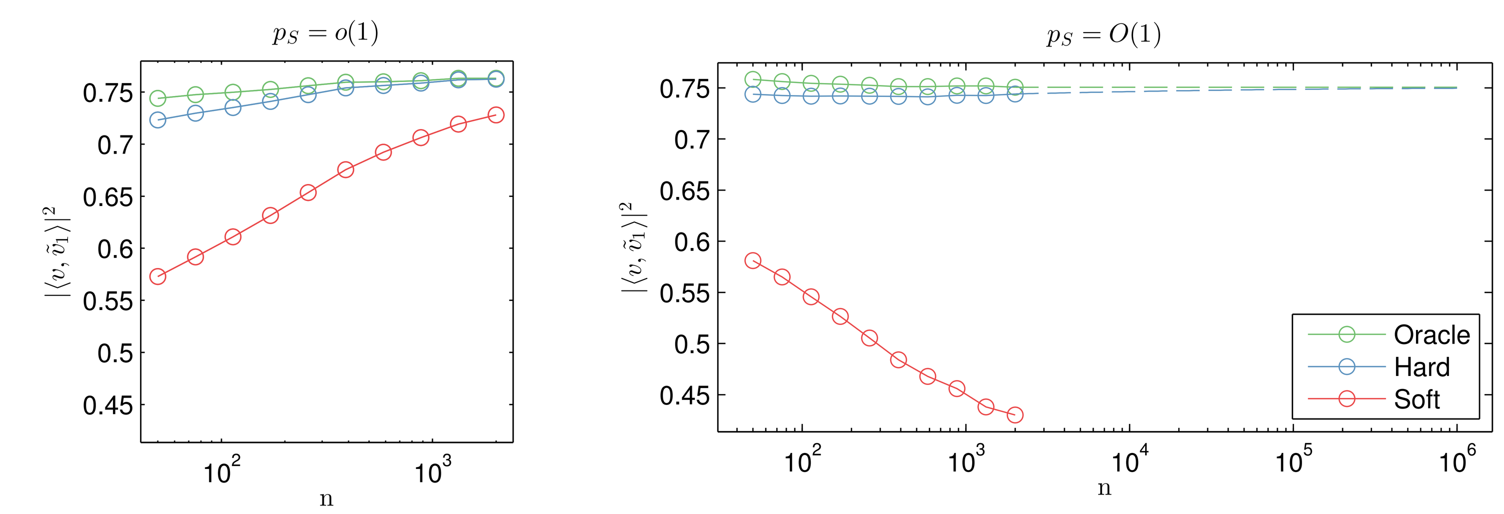
Table: Asymptotic accuracy of hard/soft-thresholding

	Sparse outliers $p_s = o(1)$	Dense outliers $p_s = O(1)$
Hard-thresholding	$\overline{\theta}^{\text{HT}} = \overline{\theta}^{\text{oracle}}$	$\overline{\theta}^{\text{HT}} = \overline{\theta}^{\text{oracle}}$
Soft-thresholding	$\overline{\theta}^{\text{ST}} = \overline{\theta}^{\text{oracle}}$	Theorem does not hold

Empirical validation of Theorem

- Gaussian noise and Laplacian outliers
- $\tau = O(\sqrt{\log(n)/n})$, which satisfies conditions of Theorem

Figure: Empirical singular vector accuracy



Application: Image denoising

- Rows of $\widetilde{X}$ contain corrupted instances of a vectorized image
  - Dense outliers: $p_s = O(1)$
  - Minimax threshold: $\tau = O(\sqrt{\log(n)/n})$
- *Only* hard-thresholding achieves oracle performance

Figure: First four observations

$\widetilde{X}(1,:)$ $\widetilde{X}(2,:)$ $\widetilde{X}(3,:)$ $\widetilde{X}(4,:)$

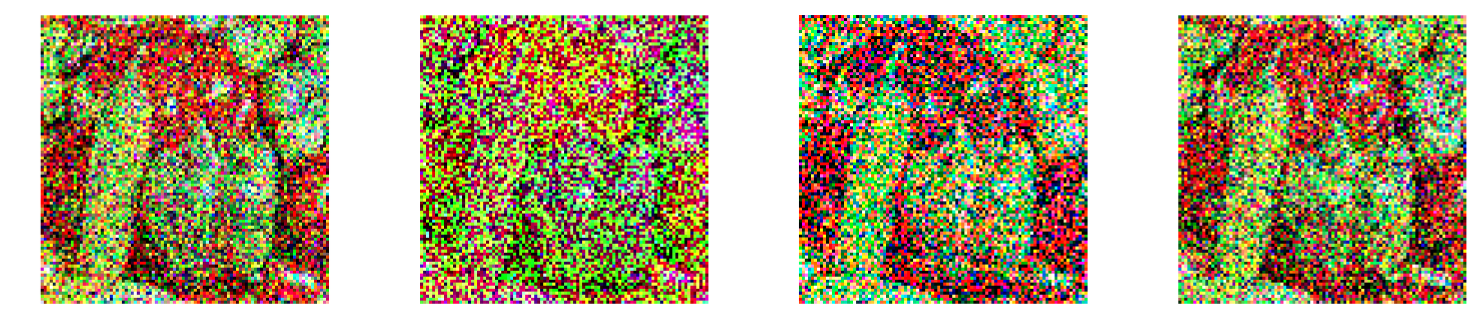
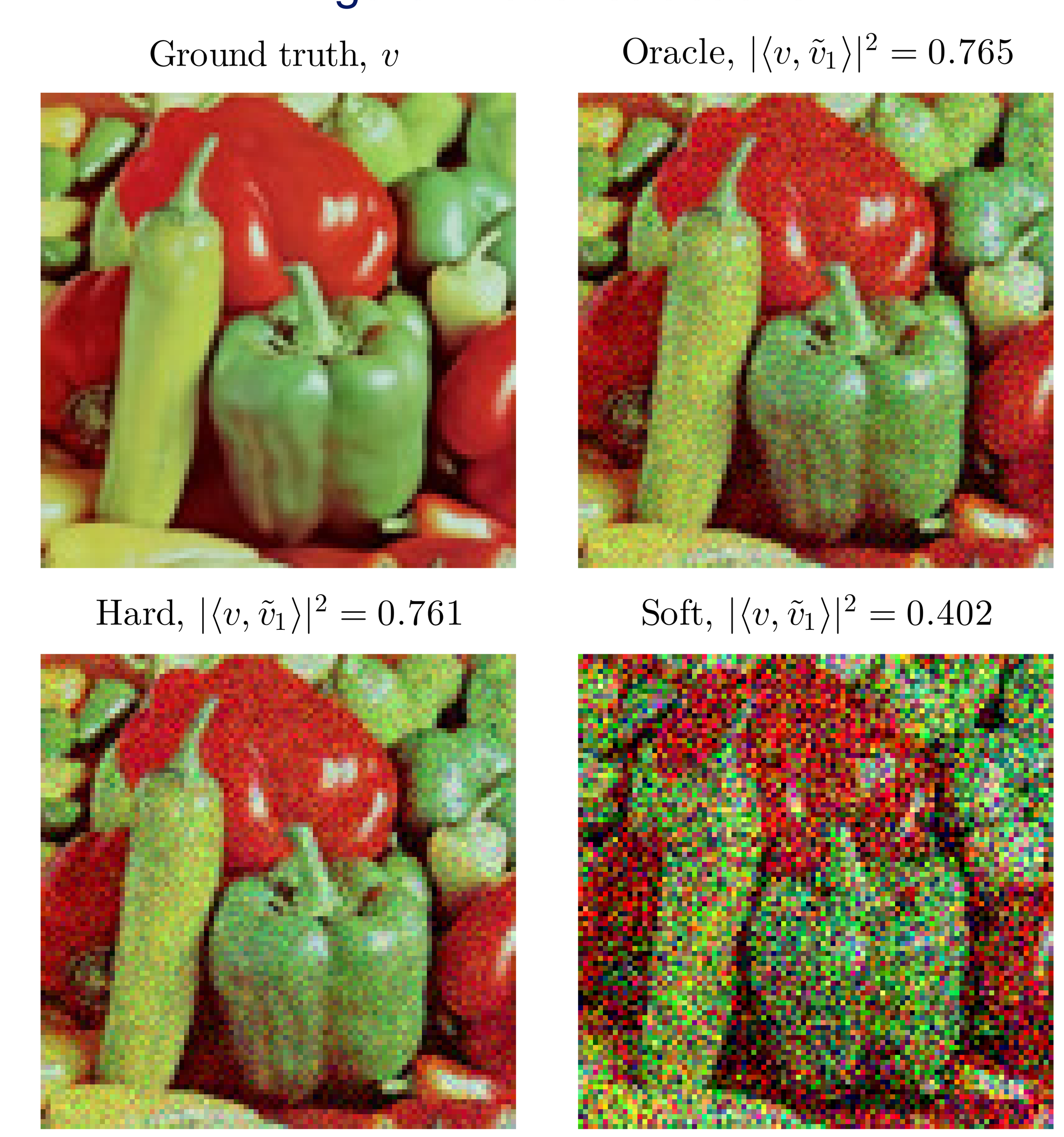


Figure: Reconstructions



Acknowledgements: Work supported by ARO MURI grant W911NF-11-1-0391.