A Novel Iterative Thresholding Algorithm for Arctangent Regularization Problem

Zihao He[†], Qianyu Shu[‡], Jinming Wen[†*], and Hing Cheung So[§]

Jinan University[†], Sichuan Normal University[‡], Pazhou Lab[*], City University of Hong Kong[§]

Introduction

- The **proximity operator of an arctangent penalty** is derived, expressed using hyperbolic functions of sine and cosine.
- An arctangent regularization iterative thresholding (ARIT) algorithm is proposed, which offers closed-form solutions for subproblems associated with the arctangent penalty.
- Experimental results demonstrate that the ARIT algorithm achieves better performance than several existing iterative thresholding algorithms in terms of the **probability of successful recovery, phase transition and running time**.

Compressed Sensing and Sparse Recovery

Compressed sensing (CS) [1] is a sampling technique that allows an s-sparse signal $\boldsymbol{x} \in \mathbb{R}^N$ to be stably recovered from a much smaller number of measurements than that required by the Nyquist-Shannon sampling theory. The primary objective of CS is to **recover $\boldsymbol{x}$ from a low-dimensional measurements vector $\boldsymbol{b}$**:

$$\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{v},$$

where $\boldsymbol{A} \in \mathbb{R}^{M\times N}$ $(M \ll N)$ is the measurement matrix and $\boldsymbol{v} \in \mathbb{R}^M$ is a noise vector.

Arctangent Penalty and Regularization Problem

The arctangent penalty is expressed as [2]:

$$\mathcal{R}_c(x) := \arctan(c|x|),$$

where $c > 0$ is a constant. The arctangent regularization problem is defined as:

$$\min_{\boldsymbol{x}\in\mathbb{R}^N} \Big\{ \underbrace{\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2 + \lambda \sum_{i=1}^{N} \arctan(c|x_i|)}_{\mathcal{F}_{\lambda,c}(\boldsymbol{x})} \Big\}.$$

As the above minimization problem is a non-convex and non-smooth optimization problem which is hard to solve directly, we apply the **majorization-minimization (MM)** method to solve it. A surrogate function is constructed:

$$\mathcal{G}_{\lambda,c,\eta,\boldsymbol{z}}(\boldsymbol{x}) = \mathcal{F}_{\lambda,c}(\boldsymbol{x}) + (\boldsymbol{x}-\boldsymbol{z})^T\left(\eta\boldsymbol{I} - \boldsymbol{A}^T\boldsymbol{A}\right)(\boldsymbol{x}-\boldsymbol{z}),$$

where $\eta \geq \|\boldsymbol{A}\|_2^2$ and $\boldsymbol{z}$ is a certain vector. Minimizing $\mathcal{G}_{\lambda,c,\eta,\boldsymbol{z}}(\boldsymbol{x})$ is equivalent to minimizing

$$\mathcal{Q}_{\lambda,c,\eta,\boldsymbol{z}}(\boldsymbol{x}) = \|\boldsymbol{x} - \mathcal{T}(\boldsymbol{z})\|_2^2 + \frac{\lambda}{\eta}\sum_{i=1}^{N}\arctan(c|x_i|),$$

where

$$\mathcal{T}(\boldsymbol{z}) = \boldsymbol{z} + \frac{1}{\eta}\boldsymbol{A}^T(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{z}).$$

Closed-form Thresholding Operator

For given $\lambda, \eta, c \in \mathbb{R}^+$ and $u \in \mathbb{R}$, denote

$$g_{u,\lambda,\eta,c}(x) := (x-u)^2 + \frac{\lambda}{\eta}\arctan(c|x|), \quad (1)$$

$$p(u) := \frac{1}{3c^2} - \frac{u^2}{9}, \quad q(u) := \frac{\lambda}{4\eta c} - \frac{u}{3c^2} - \frac{u^3}{27}.$$

If λ satisfies $0 < \lambda < \frac{16\sqrt{3}\eta}{9c^2}$, then **the global minimizer of (1) is given by**:

$$x = h(u) = \begin{cases} \mathrm{sign}(u)\bar{h}(|u|), & |u| > \frac{\lambda c}{2\eta} \\ 0, & |u| \leq \frac{\lambda c}{2\eta} \end{cases}, \quad (2)$$

where

$$\bar{h}(|u|) = \begin{cases} -2r\cosh(\frac{\vartheta}{3}) + \frac{|u|}{3}, & p(|u|) < 0 \\ (-2q(|u|))^{\frac{1}{3}} + \frac{|u|}{3}, & p(|u|) = 0 \\ -2r\sinh(\frac{\vartheta}{3}) + \frac{|u|}{3}, & p(|u|) > 0 \end{cases}$$

with $r = \mathrm{sign}(q(|u|))\sqrt{|p(|u|)|}$ and

$$\vartheta = \begin{cases} \mathrm{arcosh}\left(\frac{q(|u|)}{r^3}\right), & p(|u|) < 0 \\ \mathrm{arsinh}\left(\frac{q(|u|)}{r^3}\right), & p(|u|) > 0 \end{cases}.$$

The ARIT algorithm

Input: $\boldsymbol{b}$, $\boldsymbol{A}$, constants $\eta \geq \|\boldsymbol{A}\|_2^2$, $c > 0$, $0 < k < \frac{16\sqrt{3}}{9}$, and $\varepsilon > 0$.

Initialize: $n = 0$, $\boldsymbol{x}^{[0]} = \boldsymbol{0}$, $0 < \lambda^{[0]} \leq \frac{k\eta}{c^2}$.

until the stopping rule is met:

1: $\mathcal{T}\left(\boldsymbol{x}^{[n]}\right) = \boldsymbol{x}^{[n]} + \frac{1}{\eta}\boldsymbol{A}^T\left(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}^{[n]}\right)$;

2: **if** $\left\|\boldsymbol{A}\boldsymbol{x}^{[n]} - \boldsymbol{b}\right\|_2 \geq \varepsilon$ **then**

$$\lambda^{[n+1]} = \min\left\{\lambda^{[n]}, \frac{2\eta\left|\mathcal{T}\left(\boldsymbol{x}^{[n]}\right)\right|_{[s+1]}}{c}, \iota\right\};$$

3: **else**

$$\lambda^{[n+1]} = \lambda^{[n]};$$

4: **end if**

5: $\boldsymbol{x}^{[n+1]} = \mathcal{H}_{\lambda^{[n+1]},\eta,c}\left(\mathcal{T}\left(\boldsymbol{x}^{[n]}\right)\right)$;

Note that $\mathcal{H}_{\lambda,\eta,c}(\boldsymbol{u}) = [h(u_1),\ldots,h(u_N)]^T$;

6: $n = n + 1$;

Output: vector $\boldsymbol{x}^{[n]}$.

Experimental Results

We compare the ARIT algorithm with the IHT [3], half [4], log-sum [5], AIT-soft [6], and AIT-SCAD [6] algorithms in the following aspects.

Probability of Successful Recovery

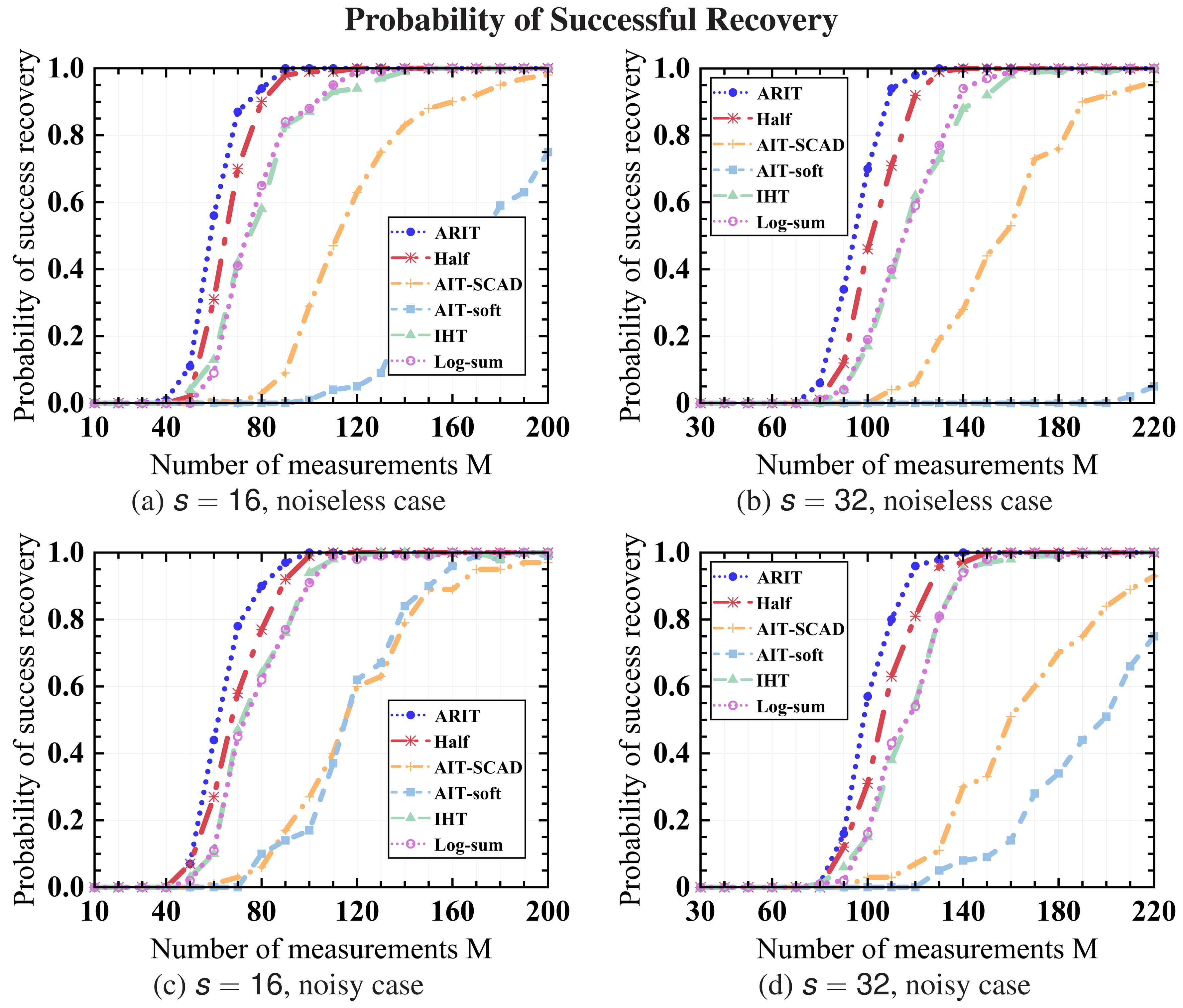

Figure: Probability of successful recovery for recovering s-sparse signals in noiseless and noisy cases with $N = 256$. **We consider $\boldsymbol{x}$ to be successfully recovered if** $\frac{\|\boldsymbol{x}-\hat{\boldsymbol{x}}\|_2}{\|\boldsymbol{x}\|_2} < 0.001 + \frac{2\|\boldsymbol{v}\|_2}{\|\boldsymbol{A}\boldsymbol{x}\|_2}$ [7], where $\boldsymbol{v}$ is the noise vector, and $\boldsymbol{x}$ and $\hat{\boldsymbol{x}}$ are the true signal and recovered signal, respectively.

Phase Transition

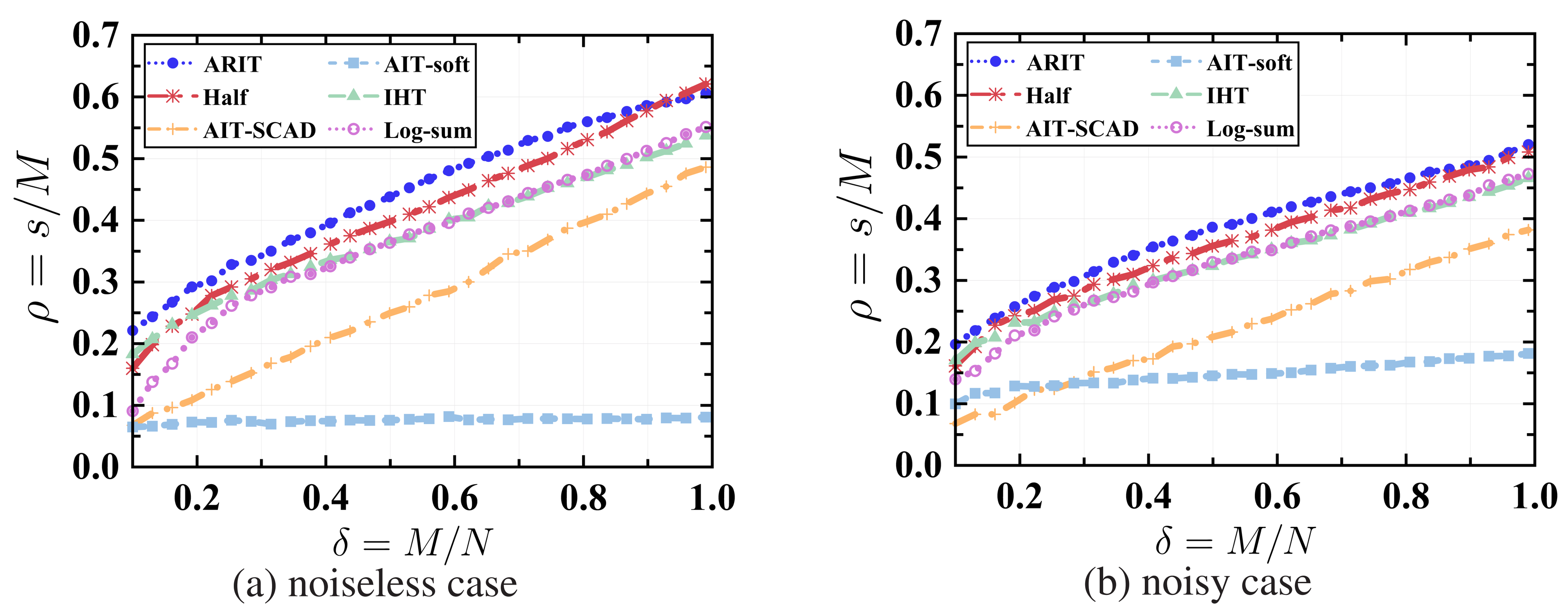

Figure: Phase transition curves for 50% **successful recovery rate** with $N = 1024$.

Running Time

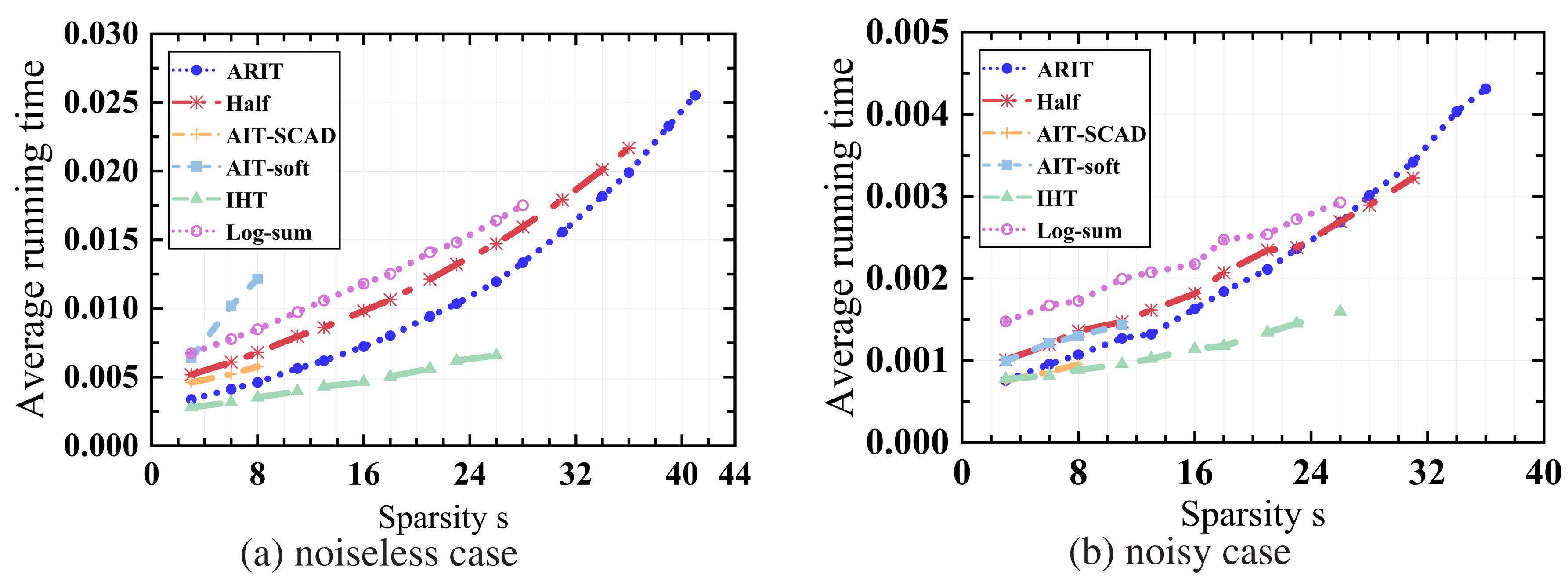

Figure: Average running time required for a successful recovery with $M = 128$ and $N = 256$. When an algorithm **successfully recovers $\boldsymbol{x}$ with a probability over 90%**, we record its average running time for performing a successful recovery.

References

[1] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
[2] E. J. Candès, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted ℓ_1 minimization," *J. Fourier Anal. Appl.*, vol. 14, pp. 877–905, 2008.
[3] R. Garg and R. Khandekar, "Gradient descent with sparsification: An iterative algorithm for sparse recovery with restricted isometry property," in *Proc. 26th Annu. Int. Conf. Mach. Learn.*, 2009, pp. 337–344.
[4] Z. Xu, X. Chang, F. Xu, and H. Zhang, "$L_{1/2}$ regularization: A thresholding representation theory and a fast solver," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 23, no. 7, pp. 1013–1027, 2012.
[5] X. Zhou, X. Liu, G. Zhang, L. Jia, X. Wang, and Z. Zhao, "An iterative threshold algorithm of log-sum regularization for sparse problem," *IEEE Trans. Circuits Syst. Video Technol.*, pp. 1–1, 2023.
[6] Y. Wang, J. Zeng, Z. Peng, X. Chang, and Z. Xu, "Linear convergence of adaptively iterative thresholding algorithms for compressed sensing," *IEEE Trans. Signal Process.*, vol. 63, no. 11, pp. 2957–2971, 2015.
[7] J. D. Blanchard, J. Tanner and K. Wei, "Conjugate gradient iterative hard thresholding: Observed noise stability for compressed sensing," *IEEE Trans. Signal Process.*, vol. 63, no. 2, pp. 528–537, 2015.