

Power Series based Hard Thresholding Algorithms for Sparse Signal Recovery

Mrityunjoy Chakraborty



Department of Electronics and Electrical Communication Engineering (E&ECE)
Indian Institute of Technology, Kharagpur
Kharagpur, West Bengal, India

"Power Series Based Hard Thresholding Algorithms for Sparse Signal Recovery," K. A. Bapat and M. Chakraborty, IEEE Signal Processing Letters, vol. 32, pp. 2474-2478, June 2025; also EUSIPCO-2026.

- 1 Overview of Sparse recovery
- 2 Power Series based Unified framework

Linear System

- $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.
- Linear system $\mathbf{y} = \mathbf{A}\mathbf{x}$, $\mathbf{y} \in \mathbb{F}^m$, $\mathbf{A} \in \mathbb{F}^{m \times N}$, $m < N$.

$$\underbrace{\begin{bmatrix} \mathbf{y} \end{bmatrix}}_{m \times 1} = \underbrace{\begin{bmatrix} \mathbf{A} \end{bmatrix}}_{m \times N} \underbrace{\begin{bmatrix} \mathbf{x} \end{bmatrix}}_{N \times 1} \quad (1)$$

- Non-trivial null space : One solution exists $\implies$ Infinite solutions exists
- Unique Recovery not possible in general
- What if $\mathbf{x}$ is sparse (say K -sparse)
- Unique solution under some conditions on m .

Sparse solutions of Linear system

- Ideally, solve the ℓ_0 -minimization problem:

$$(P_0) \quad \min \|\mathbf{z}\|_0 \quad \text{subject to} \quad \mathbf{y} = \mathbf{A}\mathbf{z}, \quad (2)$$

where $\|\cdot\|_0$ counts the nonzero elements.

- Known to be NP-hard in general.
- Seminal works by Candés, Romberg, Tao and Donoho established that solving

$$(P_1) : \quad \min \|\mathbf{z}\|_1 \quad \text{subject to} \quad \mathbf{y} = \mathbf{A}\mathbf{z}, \quad (3)$$

guarantees exact solution as that of (P_0) .

- (P_1) is a convex problem $\rightarrow$ LP formulation.
- Computationally expensive $\mathcal{O}(m^2 N^{\frac{3}{2}})$ or $\mathcal{O}(N^3)$, depending on the specific implementation.
- Led to development of Greedy and Thresholding based methods. Complexity $\mathcal{O}(mNK)$.
- Fast and capable of yielding near optimal solutions.
- How to design sensing matrix $\mathbf{A}$?
- If $\mathbf{y}$, $\mathbf{A}$ are given, how to recover K -sparse $\mathbf{x}$.

Restricted Isometry Property

The K -th order RIC δ_K of a matrix $\mathbf{A} \in \mathbb{C}^{m \times N}$ is the smallest number $\delta \geq 0$ such that

$$(1 - \delta)\|\mathbf{x}\|_2^2 \leq \|\mathbf{Ax}\|_2^2 \leq (1 + \delta)\|\mathbf{x}\|_2^2, \quad (4)$$

holds for all K -sparse vectors $\mathbf{x} \in \mathbb{C}^N$.

- Usually we say $\mathbf{A}$ satisfies RIP (of order K) if $\delta_K < 1$.
- Random matrices: i.i.d (sub)-Gaussian satisfy RIP with high probability.

$$m \geq CK \log(N/K), \quad (5)$$

- For *large dimensions* $C \approx 2$.
- If $\mathbf{x} = \mathbf{\Psi z}$, $\mathbf{\Psi}$ unitary, and $\mathbf{A}$ drawn randomly as i.i.d (sub-Gaussian), $\mathbf{A}\mathbf{\Psi}$ satisfies RIP w.h.p.
- The above definition implies that eigenvalues of the matrix $\mathbf{A}_S^H \mathbf{A}_S \in [1 - \delta_s, 1 + \delta_s]$, where $S \subseteq [N]$ is any indexing set with cardinality $|S| \leq s$.

- For a positive integer L , $[L]$ denotes the set $\{1, \dots, L\}$.
- For a set $S \subseteq [N]$, $\bar{S} := [N] \setminus S$ and $|S|$ denote the complement and the cardinality of S respectively.
- Support of a vector $\mathbf{u}$, denoted by $\text{supp}(\mathbf{u})$, is the index set of nonzero entries of $\mathbf{u}$.
- For a matrix $\mathbf{U}$, $\|\mathbf{U}\|_{2 \rightarrow 2}$ denotes its spectral norm.
- For a set $\Omega \subseteq [N]$, $(\mathbf{u})_{\Omega} \in \mathbb{C}^N$ is a vector having entries $[(\mathbf{u})_{\Omega}]_i = u_i$ if $i \in \Omega$, and other entries set to zeros.
- $\mathbf{u}_{\Omega} \in \mathbb{C}^{|\Omega|}$ is the restriction of $\mathbf{u}$ to set Ω .
- $\mathbf{A}_{\Omega}$: a submatrix of $\mathbf{A}$ having columns indexed by Ω .
- For two sets A and B , $A \Delta B := (A \setminus B) \cup (B \setminus A)$.

Greedy and Thresholding based Methods

We consider the problem :

$$\min \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 \quad \text{subject to} \quad \|\mathbf{x}\|_0 \leq K. \quad (6)$$

- IHT :

- a) Gradient descent update + hard thresholding $\rightarrow$

$$H_K[\mathbf{x}^n + \mathbf{A}^H(\mathbf{y} - \mathbf{A}\mathbf{x}^n)],$$

- where H_K : K -level hard thresholding operator.

- c) Needs more iterations for convergence (even after correct support)

- HTP :

- a) IHT + LS of size $K \rightarrow$ if IHT produces

$$\mathbf{u}^{n+1} = H_K[\mathbf{x}^n + \mathbf{A}^H(\mathbf{y} - \mathbf{A}\mathbf{x}^n)]$$

- then in HTP,

$$\mathbf{x}^{n+1} = \arg \min_{\mathbf{u} \in \mathbb{C}^N} \|\mathbf{y} - \mathbf{A}\mathbf{u}\|_2^2 \quad \text{subject to} \quad \text{supp}(\mathbf{u}) \subseteq \text{supp}(\mathbf{u}^{n+1})$$

- b) Faster Convergence.

- Newton step based IHT, HTP
- Follows the well-known Newton update in optimization

$$\mathbf{x}^{n+1} = \mathbf{x}^n - \mu(\nabla^2 f(\mathbf{x}^n))^{-1} \cdot \nabla f(\mathbf{x}^n), \quad (7)$$

- Very fast convergence compared to standard counterparts.
- However, it has poorer theoretical guarantees - requires more measurements for successful recovery.
- Typical iteration of NSIHT:

$$\mathbf{x}^{n+1} = H_K[\mathbf{x}^n + \mu(\mathbf{A}^T \mathbf{A} + \epsilon \mathbf{I})^{-1} \mathbf{A}^T (\mathbf{y} - \mathbf{A} \mathbf{x}^n)] \quad (8)$$

- Pseudo-inverse based HT algorithm (PHT):

$$\mathbf{x}^{n+1} = H_K[\mathbf{x}^n + \mu \mathbf{A}^\dagger (\mathbf{y} - \mathbf{A} \mathbf{x}^n)] \quad (9)$$

Proposed Power Series based HT algorithms

- We generalize the cost function to $\|\mathbf{Z}(\mathbf{y} - \mathbf{A}\mathbf{u})\|_2^2$, where $\mathbf{Z}^H \mathbf{Z} = q(\mathbf{A}\mathbf{A}^H)$, with $q(t)$ being a power series in t with coefficients $\alpha_i \in \mathbb{R}$, given by $q(t) = \sum_{i=c}^{\infty} \alpha_i t^i$ where c is an integer.
- Given $\lambda_i \geq 0, i = 1, \dots, m$ be the eigenvalues of the positive semidefinite matrix $\mathbf{A}\mathbf{A}^H$, we assume, $q(t)$ is convergent for $t = \lambda_i, i = 1, \dots, m$ and $q(\lambda_i) > 0$.
- Since, for each eigenvalue λ_i of $\mathbf{A}\mathbf{A}^H$, $q(\lambda_i)$ is an eigenvalue of $q(\mathbf{A}\mathbf{A}^H)$, we have $q(\mathbf{A}\mathbf{A}^H)$ positive definite.
- We define $\beta_{\min} := \min_{\lambda_i} q(\lambda_i)$ and $\beta_{\max} := \max_{\lambda_i} q(\lambda_i), i \in [m]$.
- IHT and HTP based updates performed using the gradient $-(\mathbf{A}^H q(\mathbf{A}\mathbf{A}^H)(\mathbf{y} - \mathbf{A}\mathbf{x}^n))$.

Table: POWER SERIES BASED HARD THRESHOLDING ALGORITHMS

- **Input:** $\mathbf{A}$, $\mathbf{y}$, K , $q(t)$.
- **Initialization:** K -sparse $\mathbf{x}^0$ (typically $\mathbf{x}^0 = \mathbf{0}$).
- **while** (stopping criteria not met)
 - $\mathbf{z}^{n+1} = \mathbf{x}^n + \mathbf{A}^H q(\mathbf{A}\mathbf{A}^H)(\mathbf{y} - \mathbf{A}\mathbf{x}^n)$,
 - $\mathbf{x}^{n+1} = \begin{cases} H_K[\mathbf{z}^{n+1}] & \text{// PSIHT,} \\ \arg \min_{\mathbf{u} \in \mathbb{C}^N} \|\mathbf{y} - \mathbf{A}\mathbf{u}\|_2^2 \text{ subject to } \text{supp}(\mathbf{u}) \subseteq \text{supp}(H_K[\mathbf{z}^{n+1}]). & \text{// PSHTP.} \end{cases}$
- end while**
- **Output:** K -sparse vector $\mathbf{x}^*$, $\hat{S} = \text{supp}(\mathbf{x}^*)$.

- Special cases
 - a) PSIHT, $q(t) = \mu t^0 = \mu \implies$ IHT algorithm.
 - b) PSHTP $q(t) = \mu t^0 = \mu \implies$ HTP algorithm.
 - c) PSIHT $q(t) = \mu(1 + \frac{t}{\epsilon})^{-1} \implies$ NSIHT algorithm.
 - d) PSHTP $q(t) = \mu(1 + \frac{t}{\epsilon})^{-1} \implies$ NSHTP algorithm.
 - e) PSIHT $q(t) = \mu t^{-1} \implies$ PHT algorithm.
- Opens up the possibilities of designing new algorithms with improved characteristics by suitable choice of the power series.

Lemma 1

Let $\mathbf{z} \in \mathbb{C}^N$ be a given vector and $\mathbf{x} \in \mathbb{C}^N$ be any K -sparse vector, then

$$\|(\mathbf{x} - H_K(\mathbf{z}))_{\overline{T}}\|_2 \leq \sqrt{2}\|(\mathbf{x} - \mathbf{z})_{S \Delta T}\|_2, \quad (10)$$

$$\|\mathbf{x} - H_K(\mathbf{z})\|_2 \leq \eta\|(\mathbf{x} - \mathbf{z})_{S \cup T}\|_2, \quad (11)$$

where $S = \text{supp}(\mathbf{x})$, $T = \text{supp}(H_K(\mathbf{z}))$ and $\eta = (\sqrt{5} + 1)/2$.

J. Shen and P. Li, "A tight bound of hard thresholding", *Journal of Machine Learning Research*, vol. 18, no 208, pp. 1–42, 2018.

Lemma 2

Let $\mathbf{G} \in \mathbb{C}^{m \times m}$ be a Hermitian matrix and $\mathbf{W} \in \mathbb{C}^{m \times r}$ be any matrix. Then following relations hold:

$$\lambda_{\max}(\mathbf{W}^H \mathbf{G} \mathbf{W}) \leq \lambda_{\max}(\mathbf{G}) \cdot \lambda_{\max}(\mathbf{W}^H \mathbf{W}), \quad (12)$$

$$\text{and } \lambda_{\min}(\mathbf{W}^H \mathbf{G} \mathbf{W}) \geq \lambda_{\min}(\mathbf{G}) \cdot \lambda_{\min}(\mathbf{W}^H \mathbf{W}). \quad (13)$$

Lemma 3

Let $\mathbf{A} \in \mathbb{C}^{m \times N}$, $m < N$ and $\mathbf{e} \in \mathbb{C}^m$ be a vector. If $S \subseteq N$ be an indexing set with cardinality $|S| \leq s$, then following is satisfied:

$$\|(\mathbf{A}^H q(\mathbf{A} \mathbf{A}^H) \mathbf{e})_S\|_2 \leq \beta_{\max} \sqrt{1 + \delta_s} \|\mathbf{e}\|_2. \quad (14)$$

Lemma 4

Let $\mathbf{A} \in \mathbb{C}^{m \times N}$ and $\mathbf{u}, \mathbf{v} \in \mathbb{C}^N$ be any two vectors. If $S := \text{supp}(\mathbf{u}) \cup \text{supp}(\mathbf{v})$, and $T \subseteq [N]$ are two indexing sets, then

$$|\langle \mathbf{u}, (\mathbf{I} - \mathbf{B})\mathbf{v} \rangle| \leq \rho_s \|\mathbf{u}\|_2 \|\mathbf{v}\|_2, \quad (15)$$

$$\|[(\mathbf{I} - \mathbf{B})\mathbf{v}]_T\|_2 \leq \rho_{s'} \|\mathbf{v}\|_2, \quad (16)$$

where $\mathbf{B} = \mathbf{A}^H q(\mathbf{A}\mathbf{A}^H)\mathbf{A}$, $s' = |T \cup \text{supp}(\mathbf{v})|$, $s = |S|$, and for any positive integer k ,

$$\rho_k := \max\{1 - (1 - \delta_k)\beta_{\min}, (1 + \delta_k)\beta_{\max} - 1\}. \quad (17)$$

Proof: We have, $\langle \mathbf{u}, (\mathbf{I} - \mathbf{B})\mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle - \langle \mathbf{A}\mathbf{u}, q(\mathbf{A}\mathbf{A}^H)\mathbf{A}\mathbf{v} \rangle$

$$= \langle \mathbf{u}_S, \mathbf{v}_S \rangle - \langle \mathbf{A}_S \mathbf{u}_S, q(\mathbf{A}\mathbf{A}^H)\mathbf{A}_S \mathbf{v}_S \rangle = \langle \mathbf{u}_S, \mathbf{C}\mathbf{v}_S \rangle,$$

where $\mathbf{C} = \mathbf{I} - \mathbf{A}_S^H q(\mathbf{A}\mathbf{A}^H)\mathbf{A}_S$. Applying Cauchy-Schwarz inequality,

$$|\langle \mathbf{u}, (\mathbf{I} - \mathbf{B})\mathbf{v} \rangle| \leq \|\mathbf{u}_S\|_2 \|\mathbf{C}\mathbf{v}_S\|_2 \leq \|\mathbf{u}_S\|_2 \|\mathbf{C}\|_{2 \rightarrow 2} \|\mathbf{v}_S\|_2.$$

Since $\mathbf{C}$ is a Hermitian matrix, one has

$$\|\mathbf{C}\|_{2 \rightarrow 2} = \max\{|\lambda_{\max}(\mathbf{C})|, |\lambda_{\min}(\mathbf{C})|\}. \quad (18)$$

Using (13), $\lambda_{\max}(\mathbf{C}) = 1 - \lambda_{\min}(\mathbf{A}_S^H q(\mathbf{A}\mathbf{A}^H)\mathbf{A}_S) \leq 1 - \lambda_{\min}(\mathbf{A}_S^H \mathbf{A}_S) \cdot \lambda_{\min}(q(\mathbf{A}\mathbf{A}^H)) \leq 1 - (1 - \delta_s)\beta_{\min}$, where the last step is a consequence of Definition of RIC and the fact that $\beta_{\min} > 0$.

Similarly, it follows that $\lambda_{\min}(\mathbf{C}) \geq 1 - (1 + \delta_s)\beta_{\max}$. However, since $\lambda_{\min}(\mathbf{C}) \leq \lambda_{\max}(\mathbf{C})$, from (18) we have

$$\|\mathbf{C}\|_{2 \rightarrow 2} \leq \max\{1 - (1 - \delta_s)\beta_{\min}, (1 + \delta_s)\beta_{\max} - 1\}.$$

Using the above results along with (18), while noting that $\|\mathbf{v}_S\|_2 = \|\mathbf{v}\|_2$ and $\|\mathbf{u}_S\|_2 = \|\mathbf{u}\|_2$ proof of (15) is complete. If $\mathbf{v}' = [(\mathbf{I} - \mathbf{B})\mathbf{v}]_T$, then $\|\mathbf{v}'\|_2^2 = \langle \mathbf{v}', [(\mathbf{I} - \mathbf{B})\mathbf{v}] \rangle \leq \rho_{s'} \|\mathbf{v}'\|_2 \|\mathbf{v}\|_2$. Simplifying, proof of (16) is complete.

Convergence Conditions :

- We show that both the proposed algorithms satisfy,

$$\|\mathbf{x} - \mathbf{x}^{n+1}\|_2 \leq \psi \|\mathbf{x} - \mathbf{x}^n\|_2 + \phi \|\mathbf{e}\|_2, \quad (19)$$

where ψ, ϕ depend only on $\beta_{\min}, \beta_{\max}$, and the RICs of $\mathbf{A}$ of various orders.

- Proceeding backwards recursively, (19) can be written as $\|\mathbf{x} - \mathbf{x}^{n+1}\|_2 \leq \psi^{n+1} \|\mathbf{x} - \mathbf{x}^0\|_2 + \frac{1-\psi^{n+1}}{1-\psi} \phi \|\mathbf{e}\|_2$. For convergence, it is necessary and sufficient to have $\psi < 1$, for which $\lim_{n \rightarrow \infty} \|\mathbf{x} - \mathbf{x}^{n+1}\|_2 \leq \frac{1}{1-\psi} \phi \|\mathbf{e}\|_2$.

Theorem 2.1 (PSIHT)

If the matrix $\mathbf{A}$ satisfies $\delta_{3K} < 1/\eta \simeq 0.618$ and $q(t)$ is chosen such that $\beta_{\min}$ and $\beta_{\max}$ satisfy

$$\frac{1 - \frac{1}{\eta}}{1 - \delta_{3K}} < \beta_{\min} \leq \beta_{\max} < \frac{1 + \frac{1}{\eta}}{1 + \delta_{3K}}, \quad (20)$$

then the estimate produced by the PSIHT algorithm $\{\mathbf{x}^n\}_{n \geq 1}$, satisfies (19) with $\psi = \eta\rho_{3K} < 1$ and $\phi = \eta\beta_{\max}\sqrt{1 + \delta_{2K}}$.

Proof: We have, $\mathbf{x}^{n+1} = H_K[\mathbf{z}^{n+1}]$, where $\mathbf{z}^{n+1} = \mathbf{x}^n + \mathbf{A}^H q(\mathbf{A}\mathbf{A}^H)(\mathbf{y} - \mathbf{A}\mathbf{x}^n)$. Using Lemma-1 and defining $\mathbf{B} := \mathbf{A}^H q(\mathbf{A}\mathbf{A}^H)\mathbf{A}$, $T := S \cup S^{n+1}$, $S = \text{supp}(\mathbf{x})$, we have,

$$\|\mathbf{x} - \mathbf{x}^{n+1}\|_2 \leq \eta \|(\mathbf{z}^{n+1} - \mathbf{x})_T\|_2, \quad (21)$$

$$\begin{aligned}
 &= \eta \|(\mathbf{x}^n + \mathbf{A}^H q(\mathbf{A}\mathbf{A}^H)(\mathbf{y} - \mathbf{A}\mathbf{x}^n) - \mathbf{x})_T\|_2 \\
 &\leq \eta \|((\mathbf{I} - \mathbf{B})(\mathbf{x} - \mathbf{x}^n))_T\|_2 + \eta \|(\mathbf{A}^H q(\mathbf{A}\mathbf{A}^H)\mathbf{e})_T\|_2 \\
 &\leq \eta \rho_{3K} \|(\mathbf{x} - \mathbf{x}^n)\|_2 + \eta \beta_{\max} \sqrt{1 + \delta_{2K}} \|\mathbf{e}\|_2,
 \end{aligned}$$

which results in

$$\|\mathbf{x} - \mathbf{x}^{n+1}\|_2 \leq \psi \|(\mathbf{x} - \mathbf{x}^n)\|_2 + \phi \|\mathbf{e}\|_2. \quad (22)$$

To ensure that $\psi < 1$, one requires

- $\eta(1 - (1 - \delta_{3K})\beta_{\min}) < 1$, or equivalently, $\beta_{\min} > (1 - \frac{1}{\eta})/(1 - \delta_{3K})$,
- $\eta(1 - (1 + \delta_{3K})\beta_{\max}) > -1$, or, equivalently, $\beta_{\max} < (1 + \frac{1}{\eta})/(1 + \delta_{3K})$.

Together with the fact that $\beta_{\max} \geq \beta_{\min}$, these lead to (20). Note that for (20) to hold, we require, $(1 + \frac{1}{\eta})/(1 + \delta_{3K}) > (1 - \frac{1}{\eta})/(1 - \delta_{3K})$, which is satisfied if $\delta_{3K} < 1/\eta$.

This completes the proof. □

Theorem 2.2 (PSHTP)

If the sensing matrix satisfies $\delta_{3K} < \frac{1}{\sqrt{3}}$ and $q(t)$ is chosen to satisfy

$$\frac{1 - \sqrt{\frac{1 - \delta_{3K}^2}{2}}}{1 - \delta_{3K}} < \beta_{\min} \leq \beta_{\max} < \frac{1 + \sqrt{\frac{1 - \delta_{3K}^2}{2}}}{1 + \delta_{3K}}, \quad (23)$$

then the estimate produced by PSHTP algorithm $\{\mathbf{x}^n\}_{n \geq 1}$ satisfies (19) with the coefficients ψ and ϕ given by

$$\psi = \frac{\sqrt{2}\rho_{3K}}{\sqrt{1 - \delta_{2K}^2}} < 1, \quad (24)$$

$$\phi = \frac{\sqrt{2}\beta_{\max}\sqrt{1 + \delta_{2K}}}{\sqrt{1 - \delta_{2K}^2}} + \frac{\sqrt{1 + \delta_K}}{1 - \delta_{2K}}. \quad (25)$$

- While the existing algorithms like IHT, HTP, NSIHT, NSHTP, PHT etc are special cases of the proposed power series based approach, **this opens up possibilities of designing new algorithms with improved performance by suitable choice of $q(t)$.**

- Improved guarantees:

a) For both NSIHT and NSHTP, $\psi_{\text{new}} < \psi_{\text{old}} \implies$ Faster convergence.

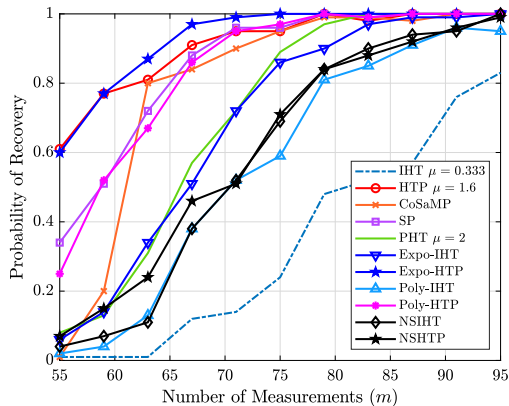
Algorithm	Previous condition	Improved condition
b) NSIHT	$\delta_{3K} < 1/\sqrt{3} \approx 0.577$	$\delta_{3K} < 1/\eta \approx 0.618$
NSHTP	$\delta_{3K} < 1/2$	$\delta_{3K} < 1/\sqrt{3}$

c) Relaxed parameter ranges.

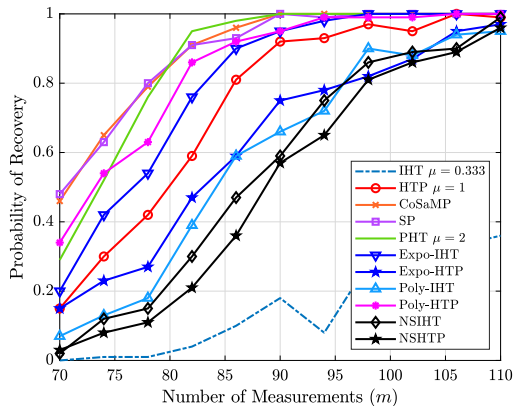
- A practical choice to choose $q(t)$ is to take it to be monotonic in t , as, in such case, $\beta_{\min} := \min_{\lambda_i} q(\lambda_i) = q(\lambda_{\min})$ and $\beta_{\max} := \max_{\lambda_i} q(\lambda_i) = q(\lambda_{\max})$, where $\lambda_{\min}$ and $\lambda_{\max}$ are the minimum and maximum eigenvalues of $\mathbf{A}\mathbf{A}^H$ respectively.

Simulation Results

- The following choices of $q(t)$ were considered:
 - a) Exponential: $q(t) = \mu \exp(-\lambda t)$ (Expo-IHT, Expo-HTP).
 - b) Unit degree polynomial: $q(t) = \mu(1 - t/\epsilon)$ (Poly-IHT, Poly-HTP)
- Support of the target K -sparse signal $\mathbf{x}$ generated from $[N]$ following a uniform distribution without repetition and the nonzero entries of $\mathbf{x}$ are taken as i.i.d according to:
 - a) Normal distribution (zero mean, unit variance), and
 - b) Bernoulli ($x_i = \pm 1$, equi-probable).
- Maximum of 100 iterations permitted for all algorithms
- Stopping criterion : $\text{MSD}(n) := \frac{\|\mathbf{x}^n - \mathbf{x}\|_2^2}{\|\mathbf{x}\|_2^2} < 10^{-4}$
- Ensemble averaging over 100 independent realizations.
- Entries of $\mathbf{A}$: i.i.d. Gaussian with zero mean and variance $\frac{1}{m}$



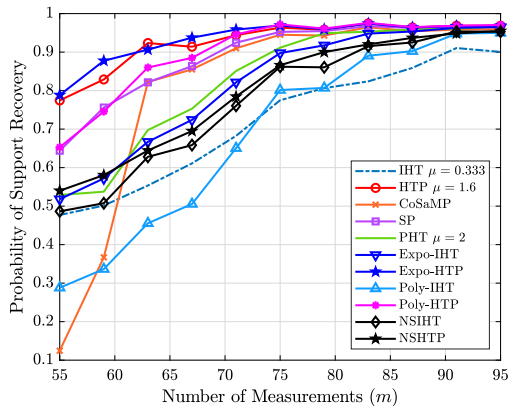
(a) Gaussian distribution, Noiseless measurements



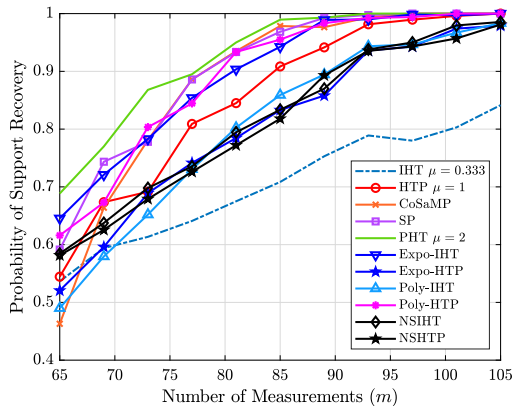
(b) Bernoulli distribution, Noiseless measurements.

Figure: Performance of various algorithms for $N = 200$, $K = 20$.

- Next, we consider a noisy scenario, where the measurements are given by $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{e}$, with the noise vector $\mathbf{e}$ consisting of zero mean, i.i.d Gaussian entries with variance $\sigma^2 = 10^{-4}$.



(a) Gaussian distribution, noisy measurements



(b) Bernoulli distribution, noisy measurements.

Figure: Performance of various algorithms for $N = 200$, $K = 20$.

Thank you. Questions...