A new and improved recovery analysis for iterative hard thresholding algorithms in compressed sensing

Coralia Cartis (University of Edinburgh)

joint with

Andrew Thompson (University of Edinburgh)

NAIS Workshop on Advances in Large-Scale Optimization

Edinburgh, May 24, 2012

The compressed sensing formulation

Let $x \in \mathbb{R}^N$ be a given signal.

Suppose we obtain vector $b \in \mathbb{R}^n$ of noisy linear measurements

$$b = Ax + e,$$

where $A \in \mathbb{R}^{n \times N}$ is the measurement matrix and e is noise.

We assume

- $n < N \implies$ underdetermined system
- x sparse with $k < n$ non-zeros

Algorithms for sparse approximation

- The problem: Find (approximate) k-sparse x from an underdetermined system of linear equations.
- Frame as the nonconvex nonsmooth problem

$$\underset{y \in \mathbf{R}^N}{\text{minimize}} \quad \tfrac{1}{2}\|Ay - b\|_2^2 \quad \text{subject to} \quad \|y\|_0 \leq k$$

  - solve by gradient projection
  - when $Ax = b$, we seek the global solution

Typically, $\|y\|_0 \leq k$ is relaxed to $\|y\|_1 \leq \tau$

$\Longrightarrow$ convex problem

But here, we solve the original l_0-formulation.

Iterative Hard Thresholding (IHT) algorithm

$$\underset{y \in \mathbf{R}^N}{\text{minimize}} \quad \tfrac{1}{2}\|Ay - b\|_2^2 \quad \text{subject to} \quad \|y\|_0 \le k$$

Iterative Hard Thresholding (IHT) algorithm

$$\underset{y \in \mathbb{R}^N}{\text{minimize}} \quad \frac{1}{2}\|Ay - b\|_2^2 \quad \text{subject to} \quad \|y\|_0 \leq k$$

Iterative Hard Thresholding (IHT): [Blumensath and Davies, 2007]

Inputs: A, b, k and $\alpha \in (0, 1)$.

Initialize: $x^0 = 0$ and $m = 0$.

While some termination criterion is not satisfied, do:

$$x^{m+1} = H_k \left\{ x^m + \alpha A^T (b - Ax^m) \right\}^{(*)}$$

Output: $\hat{x} = x^m$. $\square$

$(*)$ where $H_k(x) : \mathbb{R}^N \to \mathbb{R}^N$ keeps the k largest entries of x.

State-of-the-art analyses

- **Restricted Isometry Property (RIP)**:

$$L_s = 1 - \min_{1 \le \|y\|_0 \le s} \frac{\|Ay\|_2^2}{\|y\|_2^2} \quad \text{and} \quad U_s = \max_{1 \le \|y\|_0 \le s} \frac{\|Ay\|_2^2}{\|y\|_2^2} - 1$$

State-of-the-art analyses

- **Restricted Isometry Property (RIP)**:

$$L_s = 1 - \min_{1 \leq \|y\|_0 \leq s} \frac{\|Ay\|_2^2}{\|y\|_2^2} \quad \text{and} \quad U_s = \max_{1 \leq \|y\|_0 \leq s} \frac{\|Ay\|_2^2}{\|y\|_2^2} - 1$$

- Prove that IHT moves closer to x in each iteration:

$$\|x^{m+1} - x\|_2 \leq \mu(L_{3k}, U_{3k})\|x^m - x\|_2 + \xi(U_{2k})\|e\|_2$$

State-of-the-art analyses

- **Restricted Isometry Property (RIP)**:

$$L_s = 1 - \min_{1 \le \|y\|_0 \le s} \frac{\|Ay\|_2^2}{\|y\|_2^2} \quad \text{and} \quad U_s = \max_{1 \le \|y\|_0 \le s} \frac{\|Ay\|_2^2}{\|y\|_2^2} - 1$$

- Prove that IHT moves closer to x in each iteration:

$$\|x^{m+1} - x\|_2 \le \mu(L_{3k}, U_{3k}) \|x^m - x\|_2 + \xi(U_{2k}) \|e\|_2$$

- $\Longrightarrow$ If $\mu(L_{3k}, U_{3k}) < 1$,

$$x^m \to x^* \quad \text{such that} \quad \|x^* - x\|_2 \le \frac{\xi(U_{2k})}{1 - \mu(L_{3k}, U_{3k})} \|e\|_2$$

[Blumensath & Davies (2007); Blanchard, CC, Tanner & AT (2010)]

The proportional-growth asymptotic framework

- Claim of compressed sensing: it is possible to sample proportional to the information content (sparsity): guaranteed recovery of x for $n \geq C \cdot k \ln\left(\frac{N}{k}\right)$.

The proportional-growth asymptotic framework

- Claim of compressed sensing: it is possible to sample proportional to the information content (sparsity): guaranteed recovery of x for $n \geq C \cdot k \ln\left(\frac{N}{k}\right)$.
- Proportional-growth asymptotic: for $(\delta, \rho) \in (0, 1]^2$, let $(k, n, N) \longrightarrow \infty$ such that

$$\lim_{n\to\infty} \frac{n}{N} = \delta \quad \text{and} \quad \lim_{n\to\infty} \frac{k}{n} = \rho.$$

The proportional-growth asymptotic framework

- Claim of compressed sensing: it is possible to sample proportional to the information content (sparsity): guaranteed recovery of x for $n \geq C \cdot k \ln\left(\frac{N}{k}\right)$.
- Proportional-growth asymptotic: for $(\delta, \rho) \in (0,1]^2$, let $(k, n, N) \longrightarrow \infty$ such that

$$\lim_{n\to\infty} \frac{n}{N} = \delta \quad \text{and} \quad \lim_{n\to\infty} \frac{k}{n} = \rho.$$

- Defines a phase space for asymptotic analysis.

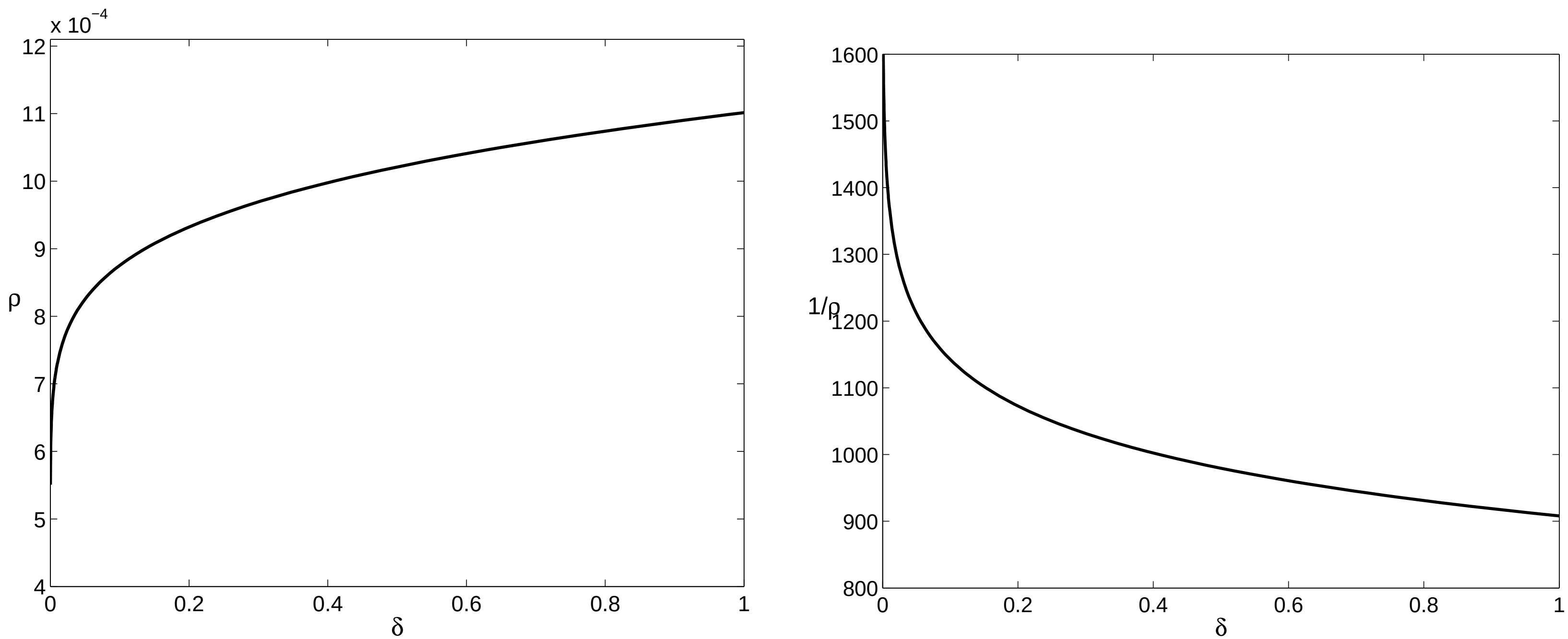
The proportional-growth asymptotic framework

- Claim of compressed sensing: it is possible to sample proportional to the information content (sparsity): guaranteed recovery of x for $n \geq C \cdot k \ln\left(\frac{N}{k}\right)$.
- Proportional-growth asymptotic: for $(\delta, \rho) \in (0,1]^2$, let $(k, n, N) \longrightarrow \infty$ such that

$$\lim_{n\to\infty} \frac{n}{N} = \delta \quad \text{and} \quad \lim_{n\to\infty} \frac{k}{n} = \rho.$$

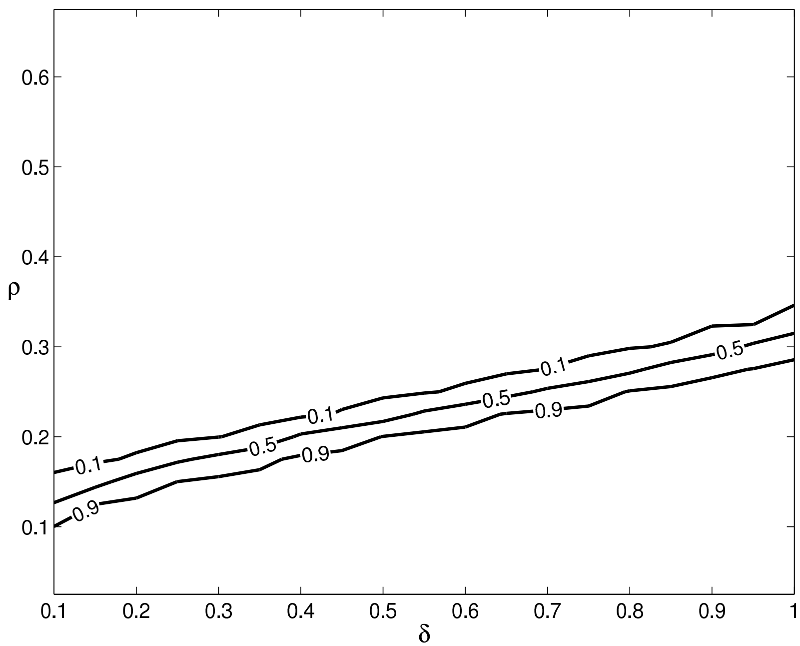
- Defines a phase space for asymptotic analysis.
- For example, RIP bounds for Gaussian matrices $L_k \longrightarrow \mathcal{L}(\delta, \rho)$ and $U_k \longrightarrow \mathcal{L}(\delta, \rho)$ [Bah and Tanner 2010]

RIP phase transition for IHT

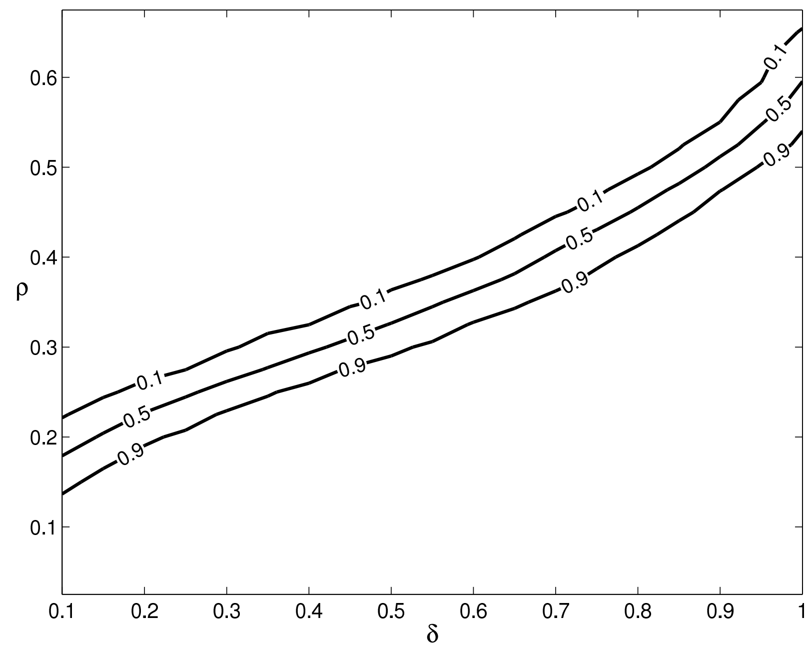


- Recovery guaranteed beneath the phase transition curve
- $n \geq 907k$ measurements needed to guarantee recovery

Empirical phase transitions for IHT



IHT



NIHT

- $\rho \sim 10^{-4}$ for RIP results
- Large gap between theory and average-case behaviour
- NIHT attains the same phase transition as for l_1-relaxation

Recovery guarantees: our approach

Aim: improve sparse-vector recovery guarantees.

Recovery guarantees: our approach

Aim: improve sparse-vector recovery guarantees.

Assumptions:

- **Noiseless case**: $\boldsymbol{e = 0} \implies \boldsymbol{b = Ax}$ and $\boldsymbol{x}$ is $\boldsymbol{k}$-sparse
- Any $\boldsymbol{2k}$ columns of $\boldsymbol{A}$ are linearly independent: $\boldsymbol{L_{2k} < 1}$.

Recovery guarantees: our approach

Aim: improve sparse-vector recovery guarantees.

Assumptions:

- **Noiseless case**: $e = 0 \implies b = Ax$ and x is k-sparse
- Any $2k$ columns of A are linearly independent: $L_{2k} < 1$.

Convergence condition: Suppose that $\alpha[1 + U_{2k}] < 1$.
Then the IHT iterates converge to a fixed point of IHT.

Recovery guarantees: our approach

Aim: improve sparse-vector recovery guarantees.

Assumptions:

- **Noiseless case**: $e = 0 \implies b = Ax$ and x is k-sparse
- Any $2k$ columns of A are linearly independent: $L_{2k} < 1$.

Convergence condition: Suppose that $\alpha[1 + U_{2k}] < 1$.
Then the IHT iterates converge to a fixed point of IHT.

Approach: derive conditions guaranteeing that

- IHT converges to some fixed point
- x is the only fixed point

$\implies$ IHT converges to x.

Fixed point analysis

A fixed point condition: [Blumensath & Davies]

Let $\bar{x}$ be k-sparse and supported on Γ. Then

$\bar{x}$ is a fixed point of IHT $\iff$ $A_\Gamma^T(b - A\bar{x}) = 0$ and

$$\min_{i \in \Gamma} |\bar{x}_i| \geq \alpha \max_{j \in \Gamma^C} \left| \left\{ A^T(b - A\bar{x}) \right\}_j \right|.$$

Fixed point analysis

A fixed point condition: [Blumensath & Davies]

Let $\bar{x}$ be k-sparse and supported on Γ. Then

$\bar{x}$ is a fixed point of IHT $\iff$ $A_\Gamma^T(b - A\bar{x}) = 0$ and

$$\min_{i\in\Gamma} |\bar{x}_i| \geq \alpha \max_{j\in\Gamma^C} |\left\{A^T(b - A\bar{x})\right\}_j|.$$

Thus $\bar{x}_\Gamma = A_\Gamma^\dagger b$ and so

- Any fixed point is a minimum-norm solution on some k-subspace.

Fixed point analysis

A fixed point condition: [Blumensath & Davies]

Let $\bar{x}$ be k-sparse and supported on Γ. Then

$\bar{x}$ is a fixed point of IHT $\iff$ $A_\Gamma^T(b - A\bar{x}) = 0$ and

$$\min_{i \in \Gamma} |\bar{x}_i| \geq \alpha \max_{j \in \Gamma^C} \left| \left\{ A^T(b - A\bar{x}) \right\}_j \right|.$$

Thus $\bar{x}_\Gamma = A_\Gamma^\dagger b$ and so

- Any fixed point is a minimum-norm solution on some k-subspace.
- But a minimum-norm solution is not necessarily a fixed point...

Single fixed point condition

Suppose

- $\bar{x}$ is a fixed point supported on Γ with $|\Gamma| = k$
- The original signal x is supported on Λ

Single fixed point condition

Suppose

- $\bar{x}$ is a fixed point supported on Γ with $|\Gamma| = k$
- The original signal x is supported on Λ

$$\min_{i\in\Gamma} |\bar{x}_i| \geq \alpha \max_{j\in\Gamma^C} |\left\{A^T(b - A\bar{x})\right\}_j|$$

$$\implies \|\bar{x}_{\Gamma\setminus\Lambda}\|_2 \geq \alpha \|\left\{A^T(b - A\bar{x})\right\}_{\Lambda\setminus\Gamma}\|_2$$

Single fixed point condition

Suppose

- $\bar{x}$ is a fixed point supported on Γ with $|\Gamma| = k$
- The original signal x is supported on Λ

$$\min_{i\in\Gamma} |\bar{x}_i| \geq \alpha \max_{j\in\Gamma^C} |\left\{A^T(b - A\bar{x})\right\}_j|$$

$$\implies \|\bar{x}_{\Gamma\setminus\Lambda}\|_2 \geq \alpha \|\left\{A^T(b - A\bar{x})\right\}_{\Lambda\setminus\Gamma}\|_2$$

$$\implies \|A_\Gamma^\dagger A_{\Lambda\setminus\Gamma} x_{\Lambda\setminus\Gamma}\|_2^2 \geq \alpha^2 \|A_{\Lambda\setminus\Gamma}^T (I - A_\Gamma A_\Gamma^\dagger) A_{\Lambda\setminus\Gamma} x_{\Lambda\setminus\Gamma}\|_2^2.$$

Single fixed point condition

Suppose

- $\bar{x}$ is a fixed point supported on Γ with $|\Gamma| = k$
- The original signal x is supported on Λ

$$\min_{i\in\Gamma} |\bar{x}_i| \geq \alpha \max_{j\in\Gamma^C} | \left\{A^T(b - A\bar{x})\right\}_j |$$

$$\implies \|\bar{x}_{\Gamma\setminus\Lambda}\|_2 \geq \alpha \| \left\{A^T(b - A\bar{x})\right\}_{\Lambda\setminus\Gamma} \|_2$$

$$\implies \|A_\Gamma^\dagger A_{\Lambda\setminus\Gamma} x_{\Lambda\setminus\Gamma}\|_2^2 \geq \alpha^2 \|A_{\Lambda\setminus\Gamma}^T (I - A_\Gamma A_\Gamma^\dagger) A_{\Lambda\setminus\Gamma} x_{\Lambda\setminus\Gamma}\|_2^2.$$

Theorem: Suppose

$$\|A_\Gamma^\dagger A_{\Lambda\setminus\Gamma} x_{\Lambda\setminus\Gamma}\|_2^2 < \alpha^2 \|A_{\Lambda\setminus\Gamma}^T (I - A_\Gamma A_\Gamma^\dagger) A_{\Lambda\setminus\Gamma} x_{\Lambda\setminus\Gamma}\|_2^2$$

for all $\Gamma \neq \Lambda$. Then x is the only fixed point of IHT.

Analysis for Gaussian matrices

Suppose $A \in \mathbb{R}^{n \times N}$ with entries distributed i.i.d. $N(0, 1/n)$ and suppose x is independent of A. Let Γ be an index set such that $|\Gamma| = k$ and $\Gamma \neq \Lambda$. Then

$$\frac{\|A_\Gamma^\dagger A_{\Lambda\setminus\Gamma} x_{\Lambda\setminus\Gamma}\|_2^2}{\|x_{\Lambda\setminus\Gamma}\|_2^2} = F_\Gamma, \quad \text{where} \quad F_\Gamma \sim \frac{k}{n-k+1} F(k, n-k+1);$$

$$\frac{\|A_{\Lambda\setminus\Gamma}^T (I - A_\Gamma A_\Gamma^\dagger) A_{\Lambda\setminus\Gamma} x_{\Lambda\setminus\Gamma}\|_2^2}{\|x_{\Lambda\setminus\Gamma}\|_2^2} \geq \left(\frac{n-k}{n}\right)^2 R_\Gamma^2,$$

$$\text{where} \quad R_\Gamma \sim \frac{1}{n-k} \chi_{n-k}^2.$$

Analysis for Gaussian matrices

Suppose $A \in \mathbb{R}^{n\times N}$ with entries distributed i.i.d. $N(0, 1/n)$ and suppose x is independent of A. Let Γ be an index set such that $|\Gamma| = k$ and $\Gamma \neq \Lambda$. Then

$$\frac{\|A_\Gamma^\dagger A_{\Lambda\setminus\Gamma} x_{\Lambda\setminus\Gamma}\|_2^2}{\|x_{\Lambda\setminus\Gamma}\|_2^2} = F_\Gamma, \quad \text{where} \quad F_\Gamma \sim \frac{k}{n-k+1} F(k, n-k+1);$$

$$\frac{\|A_{\Lambda\setminus\Gamma}^T (I - A_\Gamma A_\Gamma^\dagger) A_{\Lambda\setminus\Gamma} x_{\Lambda\setminus\Gamma}\|_2^2}{\|x_{\Lambda\setminus\Gamma}\|_2^2} \geq \left(\frac{n-k}{n}\right)^2 R_\Gamma^2,$$

$$\text{where} \quad R_\Gamma \sim \frac{1}{n-k} \chi_{n-k}^2.$$

Single FP condition: $F_\Gamma < \alpha^2 \left(\frac{n-k}{n}\right)^2 R_\Gamma^2$ for all $\Gamma \neq \Lambda$.

Asymptotic large deviations analysis

Recall the proportional-growth asymptotic:

$(k, n, N) \longrightarrow \infty$ such that

$$\lim_{n\to\infty} \frac{n}{N} = \delta \quad \text{and} \quad \lim_{n\to\infty} \frac{k}{n} = \rho.$$

Asymptotic large deviations analysis

Recall the proportional-growth asymptotic:

$(k, n, N) \longrightarrow \infty$ such that

$$\lim_{n\to\infty} \frac{n}{N} = \delta \quad \text{and} \quad \lim_{n\to\infty} \frac{k}{n} = \rho.$$

Upper tail bound for F-distribution:

Let $X_n^i \sim \frac{k}{n-k+1} F(k, n-k+1)$ for $i = 1, 2, \ldots, \binom{N}{k}$.

Then there exists a numerically computable function $\mathcal{IF}(\delta, \rho)$ such that for any $\epsilon > 0$,

$$\mathbb{P}\left\{\cap_i \left[X_n^i < \mathcal{IF}(\delta,\rho) + \epsilon\right]\right\} \longrightarrow 1 \ \text{ as } \ n \longrightarrow \infty.$$

Asymptotic large deviations analysis ...

Recall the proportional-growth asymptotic:

$(k, n, N) \longrightarrow \infty$ such that

$$\lim_{n\to\infty} \frac{n}{N} = \delta \quad \text{and} \quad \lim_{n\to\infty} \frac{k}{n} = \rho.$$

Asymptotic large deviations analysis ...

Recall the proportional-growth asymptotic:

$(k, n, N) \longrightarrow \infty$ such that

$$\lim_{n\to\infty} \frac{n}{N} = \delta \quad \text{and} \quad \lim_{n\to\infty} \frac{k}{n} = \rho.$$

Lower tail bound for normalized χ^2-distribution:

Let $X_n^i \sim \frac{1}{n-k}\chi^2_{n-k}$ for $i = 1, 2, \ldots, \begin{pmatrix} N \\ k \end{pmatrix}$.

Then there exists a numerically computable function $\mathcal{IL}(\delta, \rho)$ such that for any $\epsilon > 0$,

$$\mathbb{P}\left\{\cap_i \left[X_n^i > 1 - \mathcal{IL}(\delta, \rho) - \epsilon\right]\right\} \longrightarrow 1 \text{ as } n \longrightarrow \infty.$$

Comparison with RIP

For $A \in \mathbb{R}^{n\times N}$ Gaussian and $y \in \mathbb{R}^N$ k-sparse independent of A,

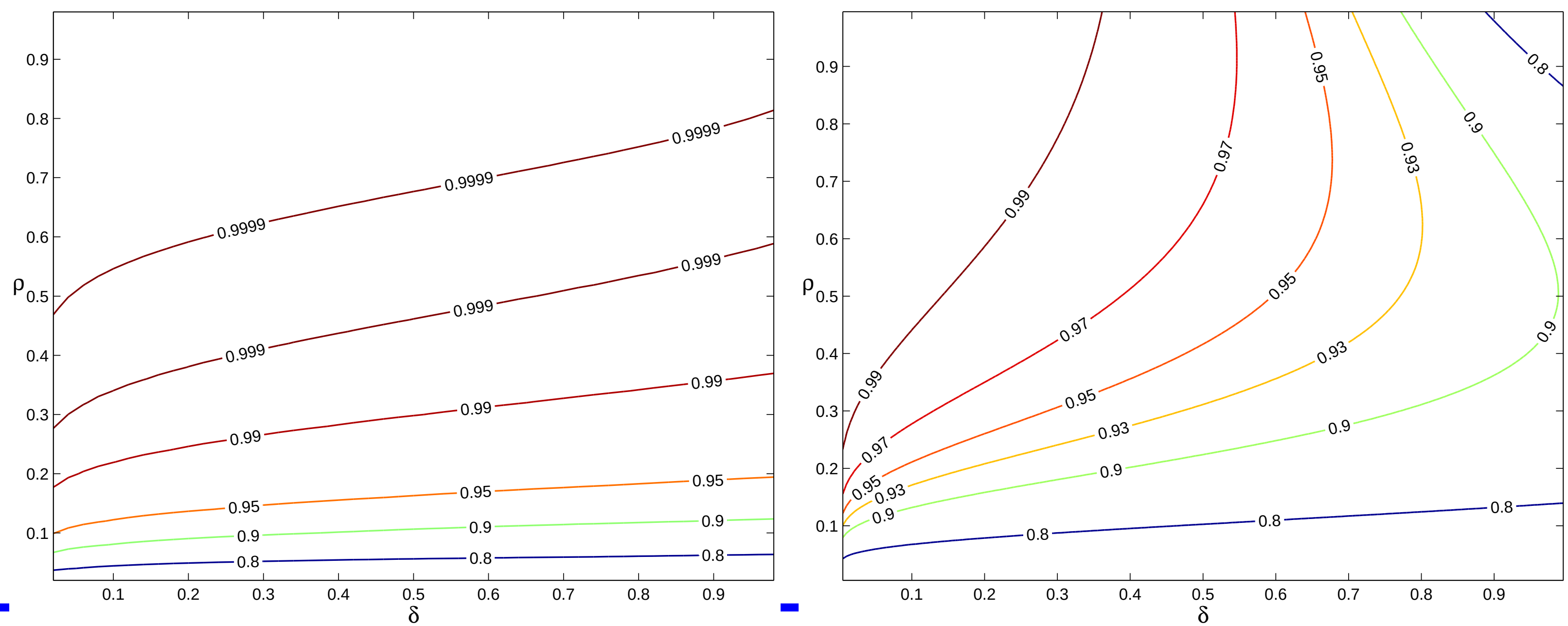
$$\frac{\|Ay\|_2^2}{\|y\|_2^2} \sim \frac{1}{n}\chi_n^2.$$

Comparison with RIP

For $A \in \mathbb{R}^{n \times N}$ Gaussian and $y \in \mathbb{R}^N$ k-sparse independent of A,

$$\frac{\|Ay\|_2^2}{\|y\|_2^2} \sim \frac{1}{n}\chi_n^2.$$

$$\mathcal{L}(\delta, \rho) \longrightarrow \mathcal{IL}(\delta, \rho)$$

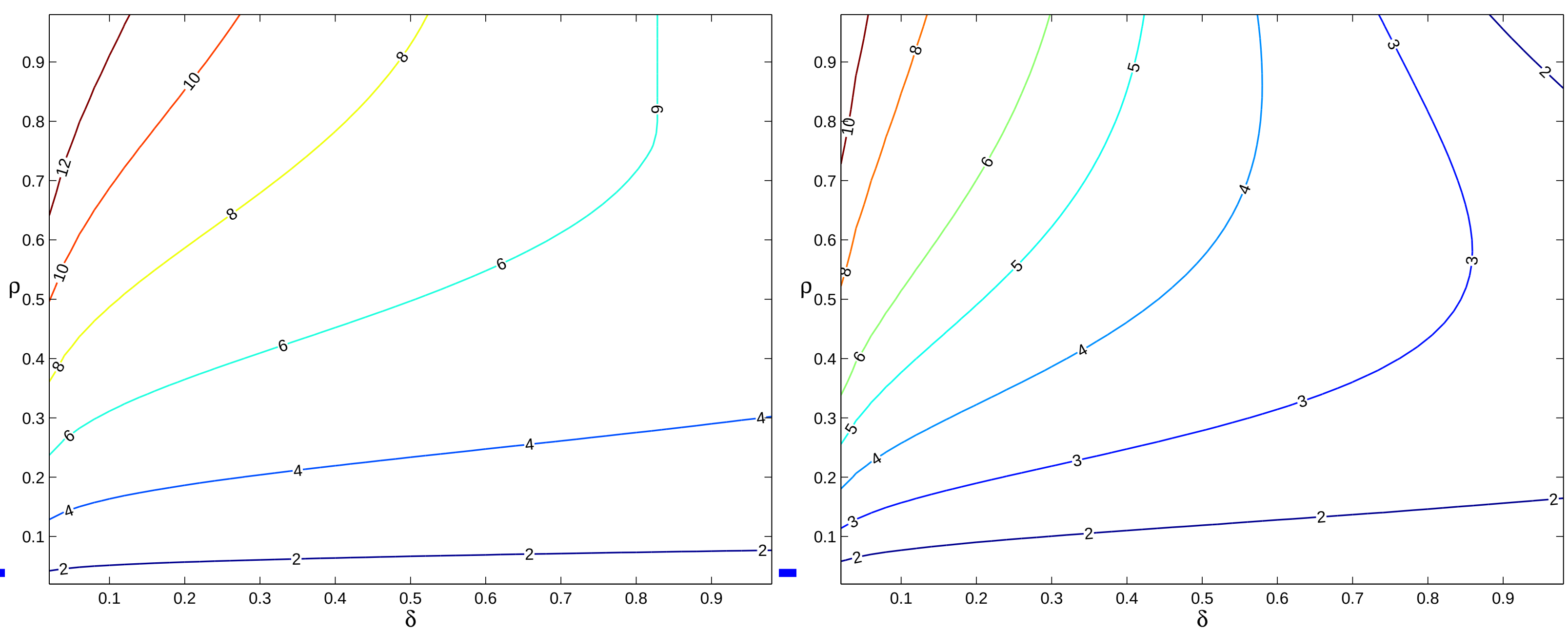


Comparison with RIP

For $A \in \mathbb{R}^{n \times N}$ Gaussian and $y \in \mathbb{R}^N$ k-sparse independent of A,

$$\frac{\|Ay\|_2^2}{\|y\|_2^2} \sim \frac{1}{n}\chi_n^2.$$

$$\mathcal{U}(\delta, \rho) \longrightarrow \mathcal{IU}(\delta, \rho)$$



Main recovery result for IHT

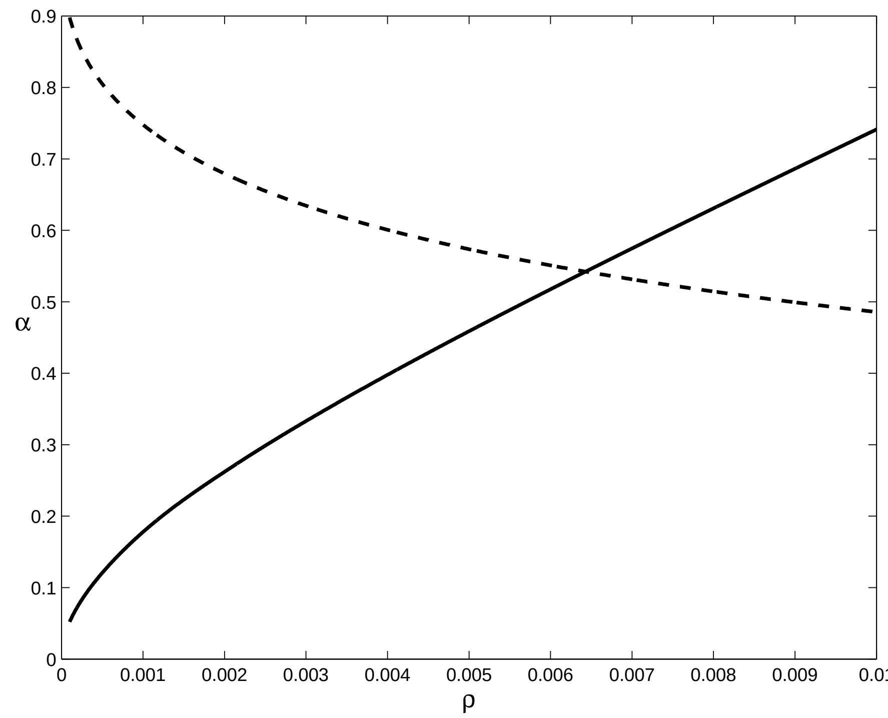
Single FP condition: $F_\Gamma < \alpha^2 \left(\frac{n-k}{n}\right)^2 R_\Gamma^2$ for all $\Gamma \neq \Lambda$

$$\overset{(k,n,N)\to\infty}{\Longrightarrow} \sqrt{\mathcal{IF}(\delta,\rho)} < \alpha(1-\rho)[1-\mathcal{IL}(\delta,\rho)].$$

Main recovery result for IHT

Single FP condition: $F_\Gamma < \alpha^2 \left(\frac{n-k}{n}\right)^2 R_\Gamma^2$ for all $\Gamma \neq \Lambda$

$\overset{(k,n,N)\to\infty}{\Longrightarrow} \quad \sqrt{\mathcal{IF}(\delta,\rho)} < \alpha(1-\rho)[1-\mathcal{IL}(\delta,\rho)]$.



Convergence condition:

$\alpha[1 + U_{2k}] < 1$

$\Downarrow$

$\alpha[1 + \mathcal{U}(\delta, 2\rho)] < 1$.

[Bah and Tanner, 2010]

Main recovery result for IHT

Single FP condition: $F_\Gamma < \alpha^2 \left(\frac{n-k}{n}\right)^2 R_\Gamma^2$ for all $\Gamma \neq \Lambda$

$\overset{(k,n,N)\to\infty}{\Longrightarrow} \quad \sqrt{\mathcal{IF}(\delta,\rho)} < \alpha(1-\rho)[1-\mathcal{IL}(\delta,\rho)]$.

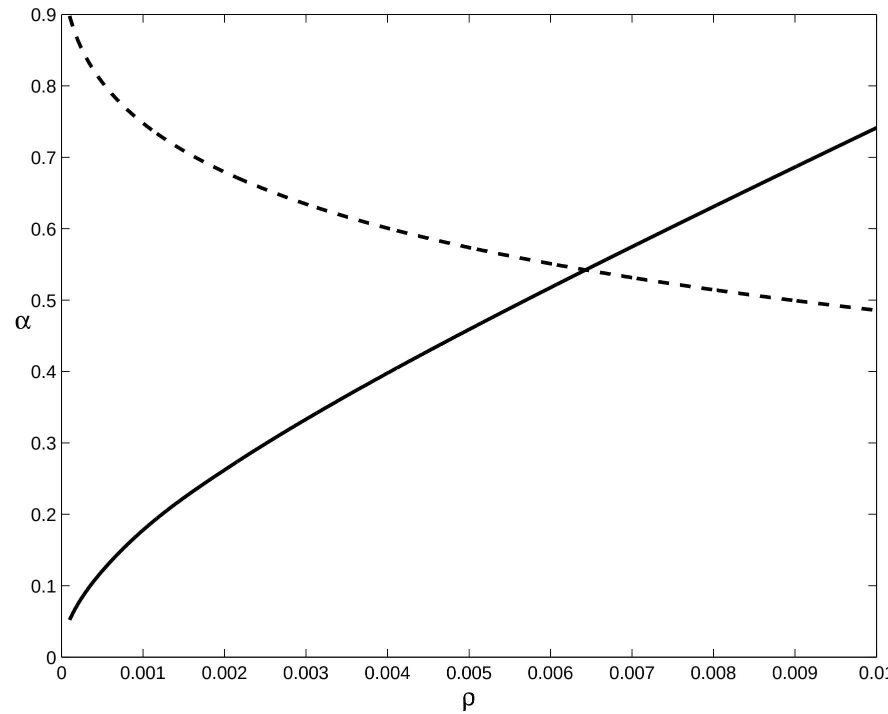
Convergence condition:

$\alpha[1+U_{2k}] < 1$

$\Downarrow$

$\alpha[1+\mathcal{U}(\delta,2\rho)] < 1$.

[Bah and Tanner, 2010]



$$\Longrightarrow \quad \frac{\sqrt{\mathcal{IF}(\delta,\rho)}}{(1-\rho)\,[1-\mathcal{IL}(\delta,\rho)]} < \alpha < \frac{1}{1+\mathcal{U}(\delta,2\rho)}$$

Main recovery result for IHT...

Theorem: Let $A \in \mathbb{R}^{n\times N}$ be a Gaussian matrix independent of x and consider the proportional growth asymptotic when $n/N \longrightarrow \delta$ and $k/n \longrightarrow \rho$ as $(k, n, N) \to \infty$. Define

$$\alpha^{min}(\delta,\rho) = \frac{\sqrt{\mathcal{IF}(\delta,\rho)}}{(1-\rho)\left[1-\mathcal{IL}(\delta,\rho)\right]} \quad \text{and} \quad \alpha^{max}(\delta,\rho) = \frac{1}{1+\mathcal{U}(\delta,2\rho)}.$$

If

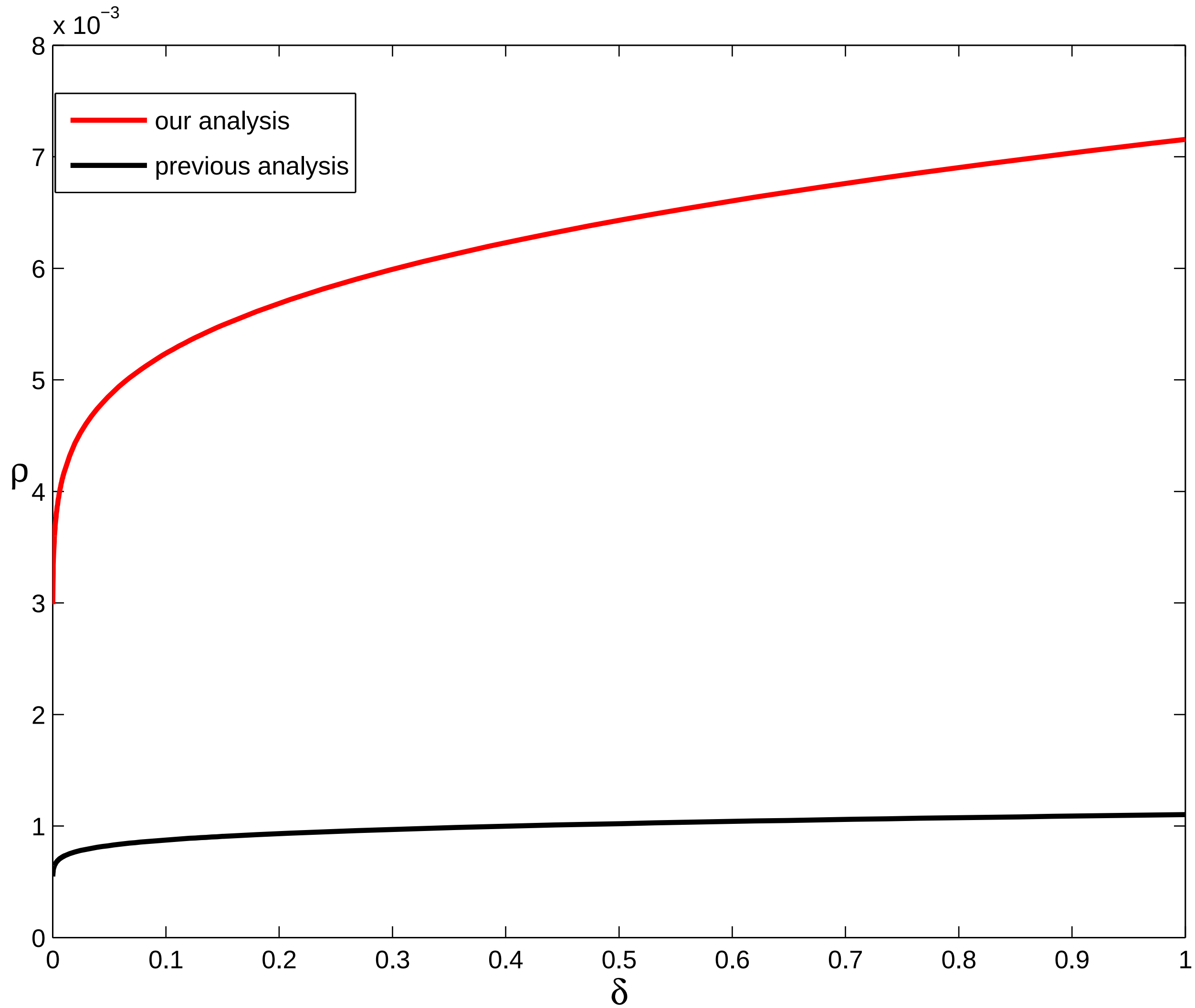
$$\alpha^{min}(\delta,\rho) < \alpha^{max}(\delta,\rho),$$

then IHT converges to x for any α satisfying

$$\alpha \in \left(\alpha^{min}(\delta,\rho), \alpha^{max}(\delta,\rho)\right),$$

with probability tending to 1 exponentially in n.

Phase transition for IHT



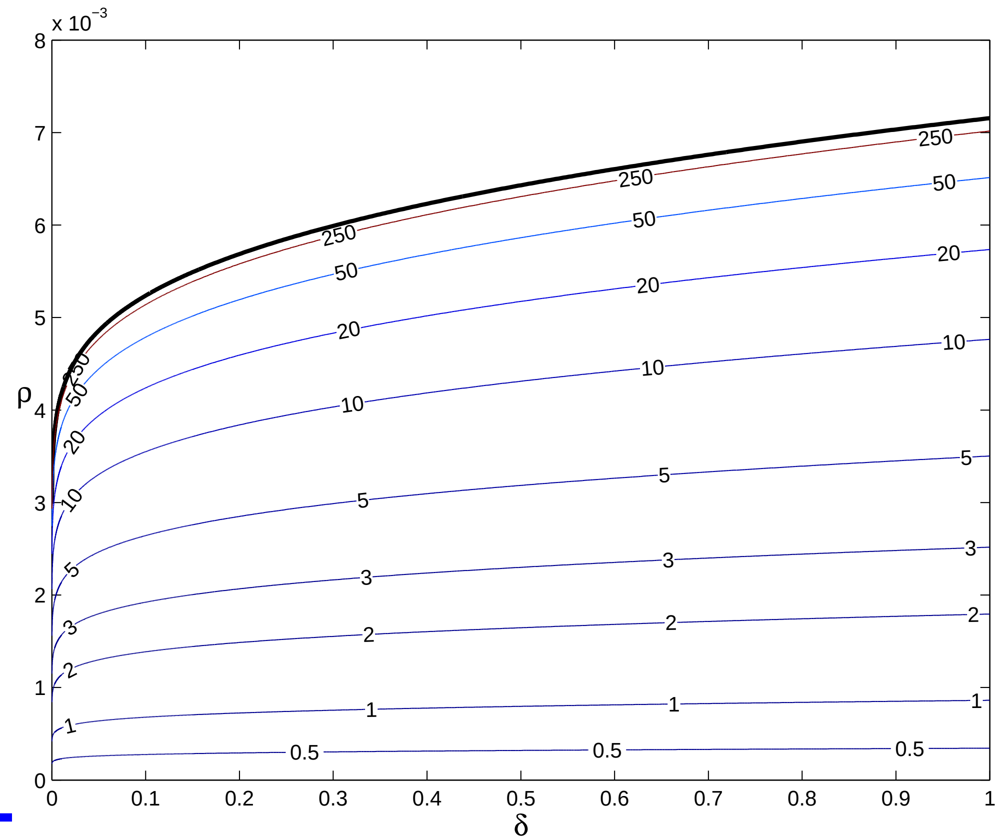
$\longrightarrow$ improvement by a factor of **7** on previous results.

Extension I: the noise case

Gaussian noise model: $b = Ax + e,\ \ e_i \sim N(0, \sigma^2/n)$.

We show that any fixed point $\bar{x}$ satisfies

$$\|\bar{x} - x\|_2 \leq \xi(\delta, \rho) \cdot \sigma.$$



Extension II: IHT variants

Normalised IHT (variable step-size)

- when $\Gamma^{m+1} = \Gamma^m$,

$$\alpha^m = \frac{\|A^T_{\Gamma^m}(b - Ax^m)\|_2^2}{\|A_{\Gamma^m}A^T_{\Gamma^m}(b - Ax^m)\|_2^2}$$

$\longrightarrow$ exact linesearch on the Γ^m face

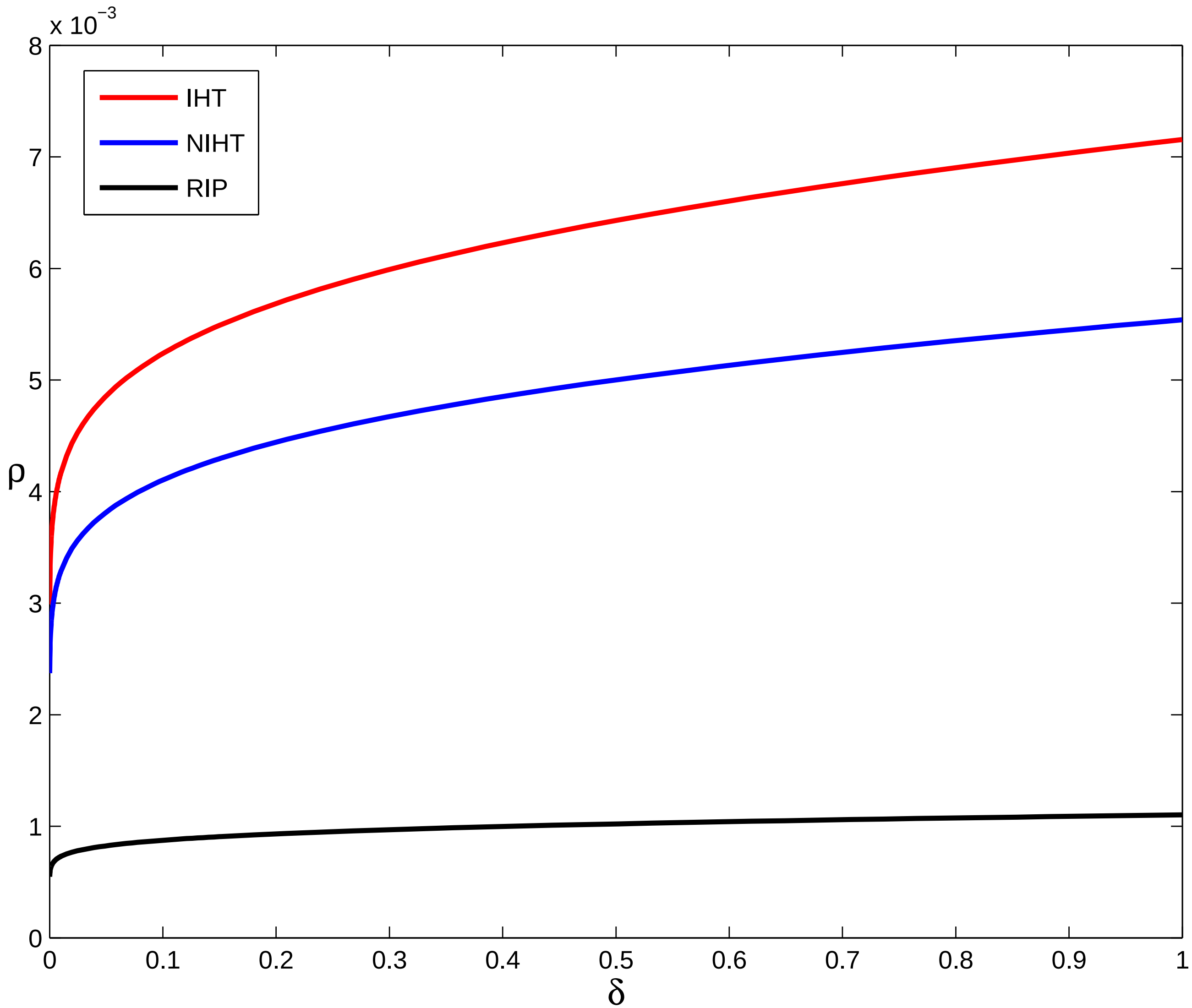
- otherwise employ a ‘sufficient decrease’ strategy.

Fixed points are not well-defined for NIHT

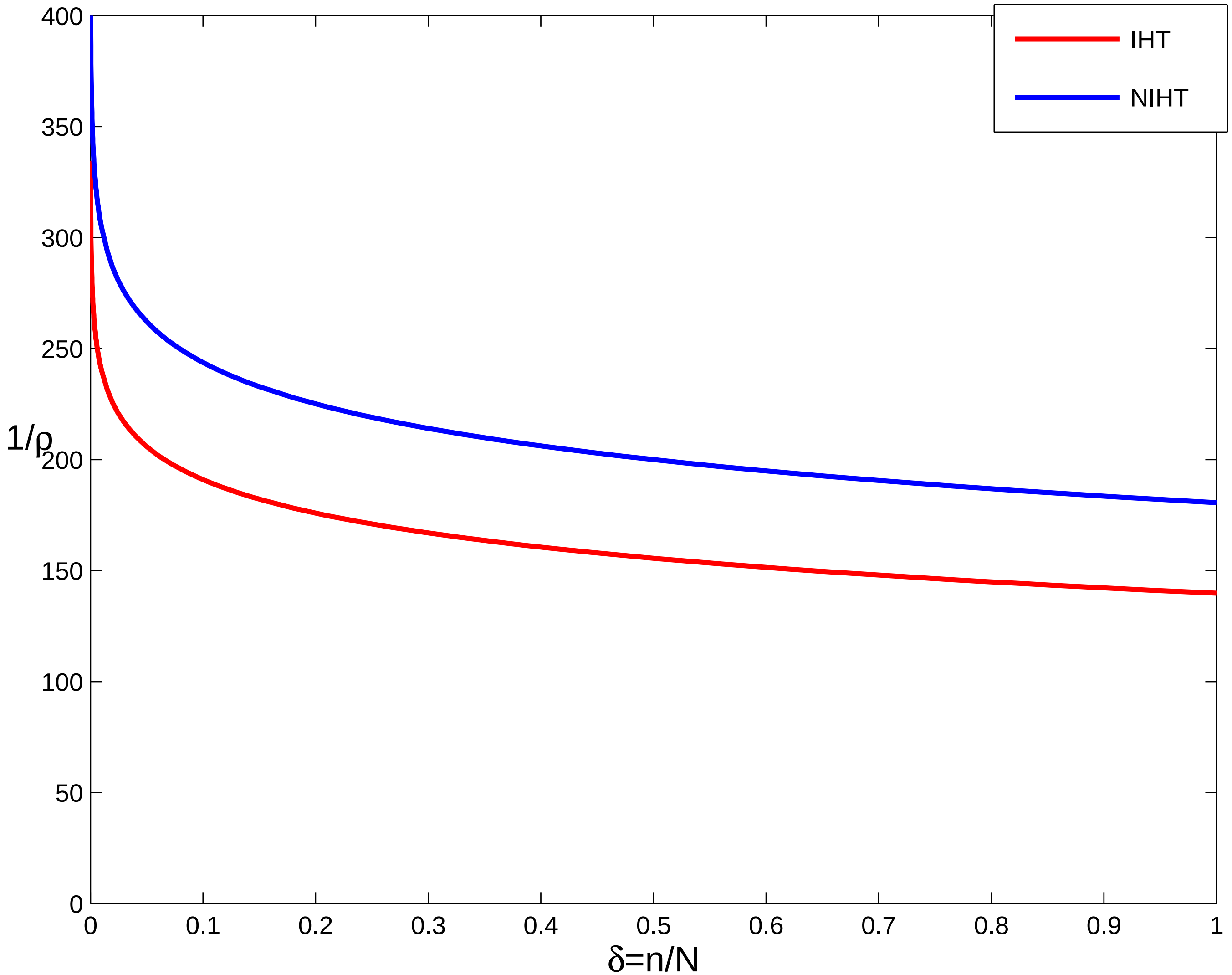
$\longrightarrow$ introduce concept of $\underline{\alpha}$-stable point.

A similar analysis gives an average phase transition for NIHT.

Recovery phase transitions



Inverse of the phase transitions



Summary and future work

- A new recovery analysis of IHT which considers its fixed points.
- An improved asymptotic recovery phase transition for Gaussian matrices.

Summary and future work

- A new recovery analysis of IHT which considers its fixed points.
- An improved asymptotic recovery phase transition for Gaussian matrices.

but...

- it still remains to fully close the gap between worst-case guarantees and average-case performance.

Summary and future work

- A new recovery analysis of IHT which considers its fixed points.
- An improved asymptotic recovery phase transition for Gaussian matrices.

but...

- it still remains to fully close the gap between worst-case guarantees and average-case performance.

Extension III. An even higher phase transition for wavelet trees, recovery if $n > 50k$ (binary).

References

- *A new and improved recovery analysis for iterative hard thresholding algorithms in compressed sensing*; CC, AT (in preparation)
- *Quantitative recovery conditions for iterative tree projection*; CC, AT (in preparation)
- *Phase transitions for greedy sparse approximation algorithms*; J.Blanchard, CC, J.Tanner, AT (Applied & Comp. Harm. Anal., 2011)
- *Improved bounds on restricted isometry constants for Gaussian matrices*; B.Bah, J.Tanner (SIAM J. on Matrix Analysis, 2010)
- *Iterative thresholding for sparse approximations*; T.Blumensath, M.Davies (J. of Fourier Anal. and Appl, 2008)
- *Normalised iterative hard thresholding; guaranteed stability and performance*; T.Blumensath, M.Davies (IEEE J. of Selected Topics in Sig. Proc, 2010)
- *Model-based compressive sensing*; R.Baraniuk, V.Cevher, M. Duarte, C. Hegde (IEEE Trans. in Inform. Theory, 2010)
- *Compressive single-pixel imaging*; A. Thompson (Technical report ERGO 11-006, 2011)