Statistical mechanics approach to 1-bit Compressed Sensing

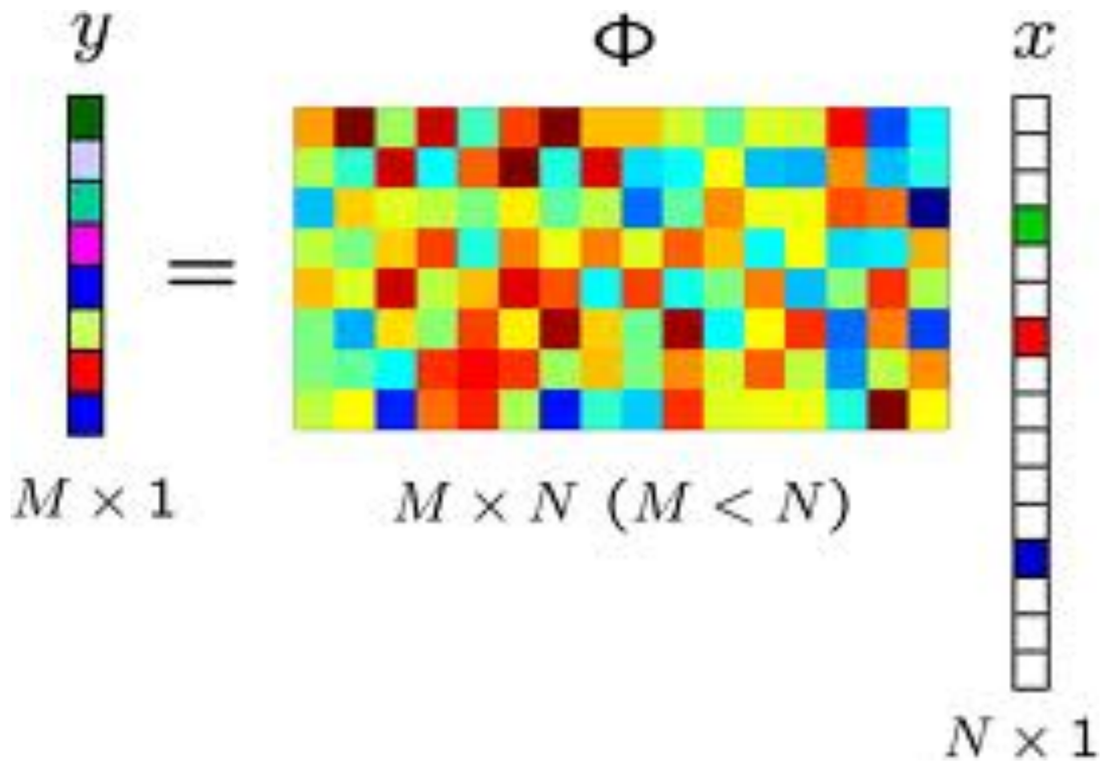
Yingying Xu
2012.Dec.7

Compressed Sensing

Measurement

Measurement matrix

Original signal



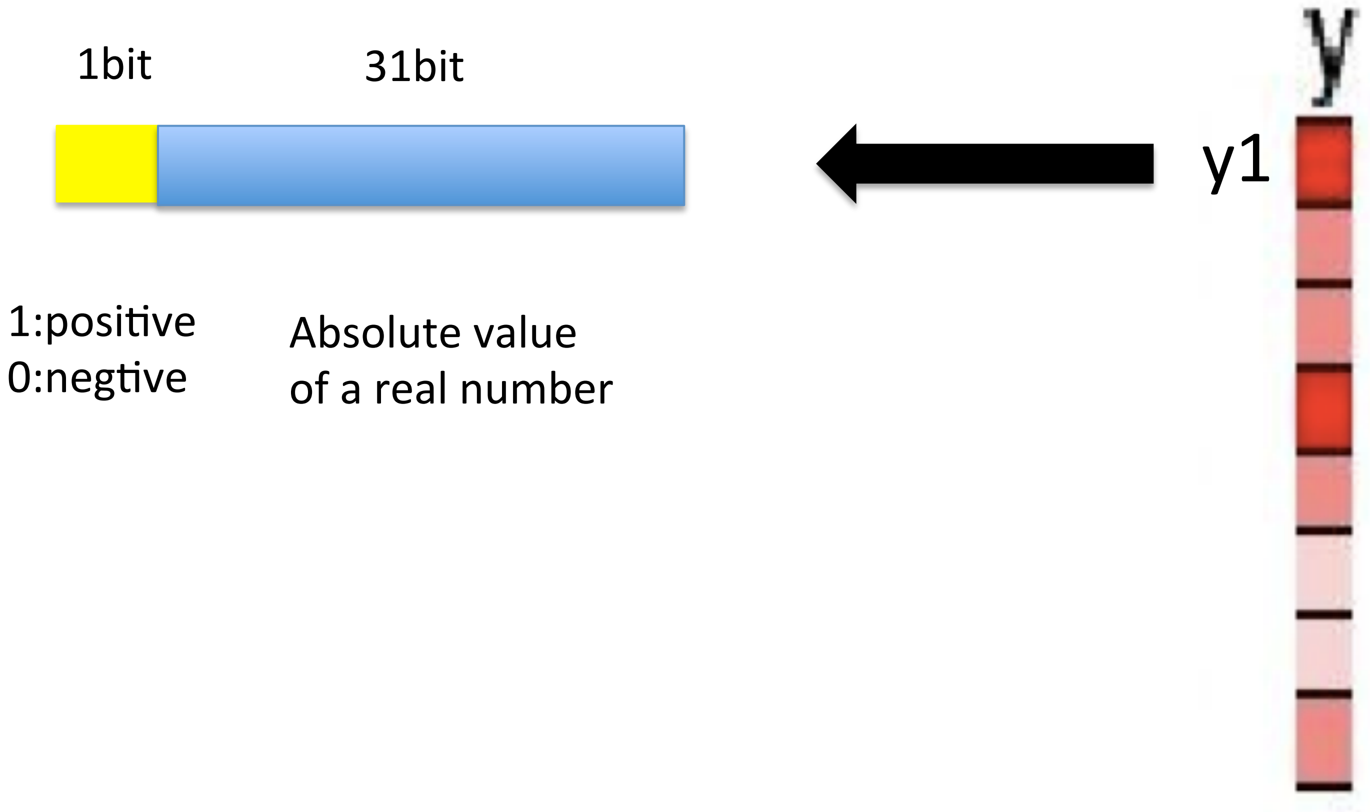
K<M<N

2k

1 bit Compressed sensing

$$\mathbf{y}=\mathrm{sign}(\Phi\mathbf{x})$$

$$\mathrm{sign}(x)=x/|x| \text{ for } x \neq 0$$



L1-norm minimization

Boufounos, P.T. and Baraniuk

$$\min_{x}\left\{\sum_{i=1}^{N}\left|x_i\right|_1\right\}\text{subj.to } \mathbf{y}=\mathrm{sign}(\Phi\mathbf{x}),\ \left|\mathbf{x}\right|^2=N \qquad (1)$$

L_1 norm   measurement   normalization

My approach:

Partition function

$$Z\left(\beta;\mathbf{A},\mathbf{x}^0\right)=\int dx\,\delta\left(\left|\mathbf{x}\right|^2-N\right)\Theta\left(\mathbf{y}\Phi\mathbf{x}\right)e^{-\beta\left|\mathbf{x}\right|}$$

Delta function   Step function

Performance assessment by the replica method

Free energy

$$\bar{f} = \lim_{\substack{N\to\infty \\ \beta\to\infty}} -\frac{1}{N\beta}\left[\ln Z\left(\beta;\Phi,\mathbf{x}^{\mathbf{0}}\right)\right]_{\Phi,\mathbf{x}^{\mathbf{0}}}$$

current problem: the partition function depends on the predetermined random variables Φ and x0, which requires us to assess the average of free energy density.

Replica method

$$\frac{1}{N}\left[\ln Z\left(\beta;\Phi,\mathbf{x}^{\mathbf{0}}\right)\right]_{\Phi,\mathbf{x}^{\mathbf{0}}} = \lim_{n\to 0}\frac{\partial}{\partial n}\frac{1}{N}\ln\left[Z^{n}\left(\beta;\Phi,\mathbf{x}^{\mathbf{0}}\right)\right]_{\Phi,\mathbf{x}^{\mathbf{0}}}$$

directly averaging the logarithm of the partition function is technically difficult. Therefore, we here resort to the replica method.

Obtain five equations

$$\frac{\partial \overline{f}}{\partial m} = 0 \Rightarrow \hat{m} = \frac{\alpha}{\pi\chi\rho}\sqrt{\rho - m^2}$$

$$\frac{\partial \overline{f}}{\partial \hat{m}} = 0 \Rightarrow m = \frac{2\rho\hat{m}}{\hat{Q}} H\left(\frac{1}{\sqrt{\hat{q} + \hat{m}^2}}\right)$$

$$\frac{\partial \overline{f}}{\partial \hat{q}} = 0 \Rightarrow \chi = \frac{2}{\hat{Q}}\left[(1-\rho)H\left(\frac{1}{\sqrt{\hat{q}}}\right) + \rho H\left(\frac{1}{\sqrt{\hat{q} + \hat{m}^2}}\right)\right]$$

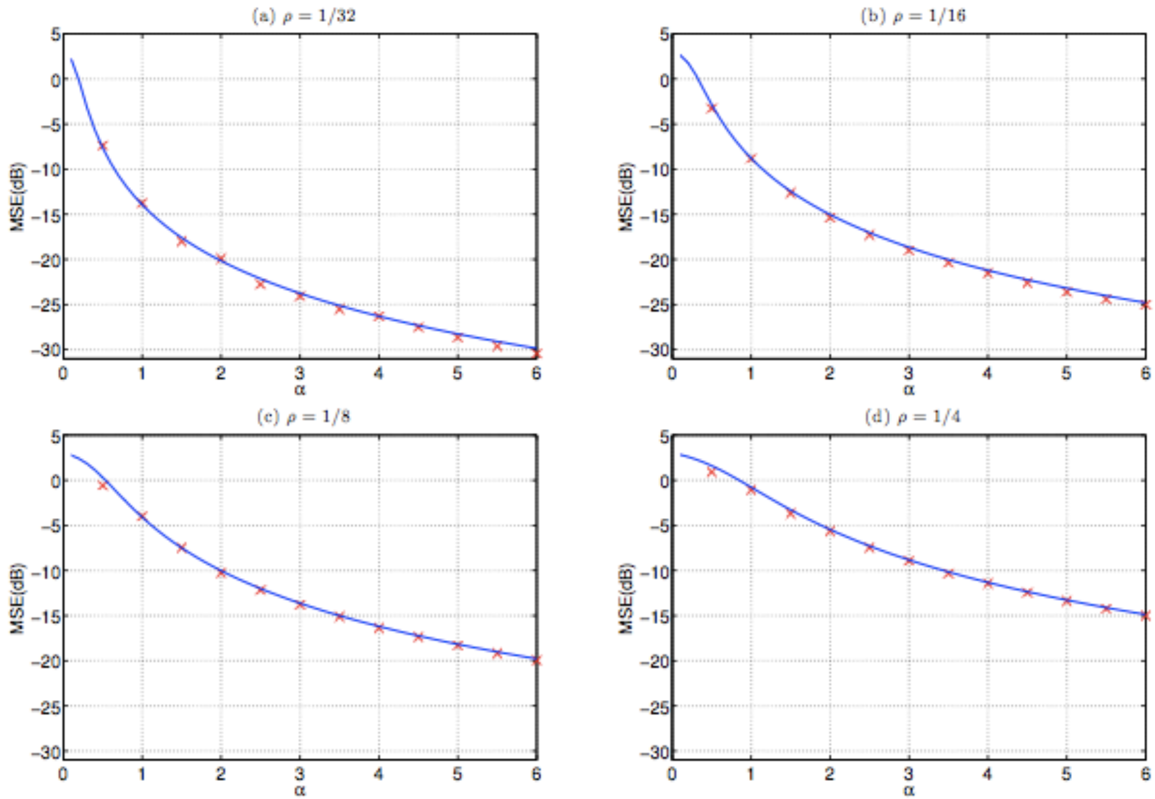
$$\frac{\partial \overline{f}}{\partial \chi} = 0 \Rightarrow \hat{q} = \frac{\alpha}{\pi\chi^2}\left(\arctan\left(\frac{\sqrt{\rho - m^2}}{m}\right) - \frac{m}{\rho}\sqrt{\rho - m^2}\right)$$

$$\frac{\partial \overline{f}}{\partial \hat{Q}} = 0 \Rightarrow \hat{Q}^2 = 2\left[(1-\rho)\left((\hat{q}+1)H\left(\frac{1}{\sqrt{\hat{q}}}\right) - \sqrt{\frac{\hat{q}}{2\pi}}e^{-\frac{1}{2\hat{q}}}\right) + \rho\left((\hat{q} + \hat{m}^2 + 1)H\left(\frac{1}{\sqrt{\hat{q} + \hat{m}^2}}\right) - \sqrt{\frac{\hat{q} + \hat{m}^2}{2\pi}}e^{-\frac{1}{2(\hat{q} + \hat{m}^2)}}\right)\right]$$

Results about MSE

$$\mathrm{MSE} = \left[\left| \frac{\hat{\boldsymbol{x}}}{|\hat{\boldsymbol{x}}|} - \frac{\boldsymbol{x}^0}{|\boldsymbol{x}^0|} \right|^2 \right]_{\boldsymbol{\Phi}, \boldsymbol{x}^0} = 2\left(1 - \frac{m}{\sqrt{\rho}}\right)$$

Mean squared error



A b c d corresponds to the cases of rho=0.03125, 0.0625, 0.125, and 0.25, respectively.

Curve: the theoretical prediction

Symbol ×: the experimental estimate obtained for algorithm RFPI from 1000 experiments

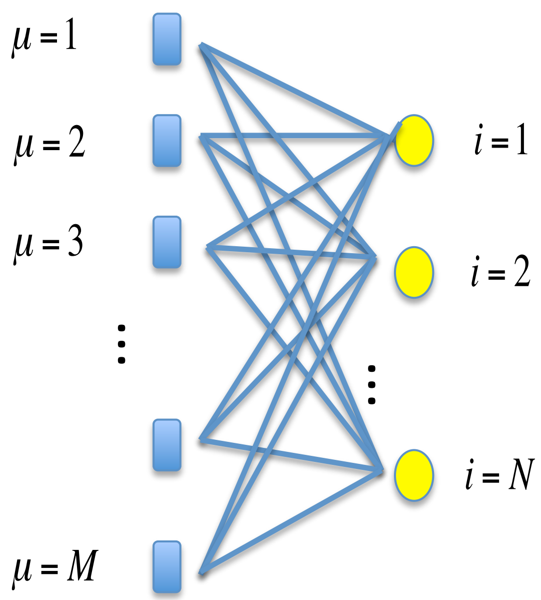
Developed an algorithm by Cavity method

$$\min_{x}\left\{\sum_{i=1}^{N}|x_i|_1\right\}\text{subj.to } \mathbf{y}=\text{sign}(\Phi\mathbf{x}),\ |\mathbf{x}|^2=N \qquad (1)$$

$$\min_{x,z>0}\max_{a,\Lambda}\left\{\sum_{i=1}^{N}|x_i|+\sum_{\mu=1}^{M}a_\mu\left(\sum_{i=1}^{N}\Phi_{\mu i}x_i-z_\mu\right)+\frac{\Lambda}{2}\left(\sum_{i=1}^{N}x_i^2-N\right)\right\}$$

Lagrange multipliers

Surplus variables



single body cost functions

$$L_i(x_i)=\frac{A_i}{2}x_i^2-H_ix_i+|x_i|,$$

$$L_\mu(a_\mu,z_\mu)=-\frac{B_\mu}{2}a_\mu^2+K_\mu a_\mu-z_\mu a_\mu,$$

where A_i, B_μ, H_i, K_μ are parameters

to be determined in a self-consistent manner.

Cavity method

μ-cavity systems: take out (a_μ, z_μ)

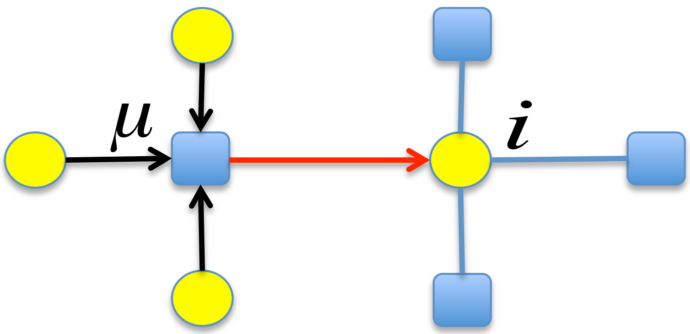
i-cavity systems: take out x_i

put (a_μ, z_μ) back into the μ-cavity system, remove x_i

$$L_{\mu\to i}(a_\mu, z_\mu) = \min_{\boldsymbol{x}\setminus x_i} \left\{ \sum_{j\neq i} \left(\Phi_{\mu j} a_\mu x_j + L_{j\to\mu}(x_j) \right) \right\} - z_\mu a_\mu$$

$$B_{\mu\to i} = \sum_{j\neq i} \frac{\Phi_{\mu j}^2}{A_{j\to\mu}} g''(H_{j\to\mu}),$$

$$K_{\mu\to i} = \sum_{j\neq i} \frac{\Phi_{\mu j}}{A_{j\to i}} g'(H_{j\to\mu}),$$



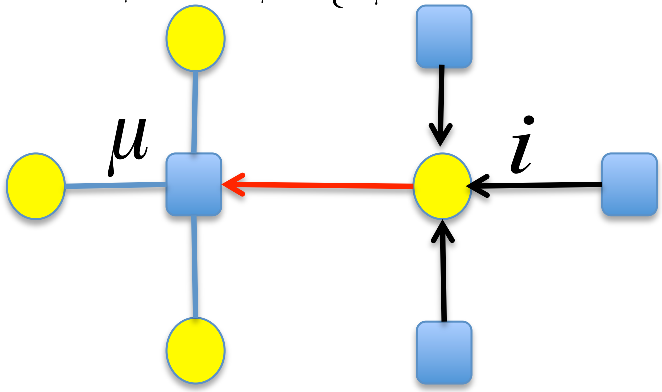
$$g(u) \equiv \frac{(|u|-1)^2}{2} \Theta(|u|-1)$$

put x_i back into the i-cavity system, remove (a_μ, z_μ)

$$L_{i\to\mu}(x_i) = \min_{z\setminus x_\mu > 0} \max_{a\setminus a_\mu} \left\{ \sum_{\nu\neq\mu} \left(\Phi_{\nu i} a_\nu x_i + L_{\nu\to i}(a_\nu, z_\nu) \right) \right\} + \frac{\Lambda}{2} x_i^2 + |x_i|$$

$$A_{i\to\mu} = \Lambda + \sum_{\nu\neq\mu} \frac{\Phi_{\nu i}^2}{B_{\nu\to i}} f''(K_{\nu\to i}),$$

$$H_{i\to\mu} = -\sum_{\nu\neq\mu} \frac{\Phi_{\nu i}}{B_{\nu\to i}} f'(K_{\nu\to i}).$$



$$f(u) \equiv \frac{u^2}{2} \Theta(-u)$$

Cavity method

i-cavity system → original system

$$B_{\mu\to i}=\sum_{j\neq i}\frac{\Phi_{\mu j}^2}{A_{j\to\mu}}g''\left(H_{j\to\mu}\right), \longrightarrow B_\mu=\sum_{i=1}^{N}\frac{\Phi_{\mu i}^2}{A_{i\to\mu}}g''\left(H_{i\to\mu}\right), \xrightarrow{N\to\infty} B=\frac{1}{NA}\sum_{i=1}^{N}g''\left(H_i\right),$$

$$K_{\mu\to i}=\sum_{j\neq i}\frac{\Phi_{\mu j}}{A_{j\to i}}g'\left(H_{j\to\mu}\right) \longrightarrow K_\mu=\sum_{i=1}^{N}\frac{\Phi_{\mu i}}{A_{i\to\mu}}g'\left(H_{i\to\mu}\right) \xrightarrow{\text{Taylor expansion}} K_\mu\approx\sum_{i=1}^{N}\Phi_{\mu i}\hat{x}_i-B\hat{a}_\mu.$$

$$\boxed{\hat{x}_i\equiv\frac{1}{A}g'\left(H_i\right)}$$

μ-cavity system → original system

$$A_{i\to\mu}=\Lambda+\sum_{\nu\neq\mu}\frac{\Phi_{\nu i}^2}{B_{\nu\to i}}f''\left(K_{\nu\to i}\right), \longrightarrow A_i=\Lambda+\sum_{\mu=1}^{M}\frac{\Phi_{\mu i}^2}{B_{\mu\to i}}f''\left(K_{\mu\to i}\right), \xrightarrow{N\to\infty} A=\Lambda+\frac{1}{NB}\sum_{\mu=1}^{M}f''\left(K_\mu\right),$$

$$H_{i\to\mu}=-\sum_{\nu\neq\mu}\frac{\Phi_{\nu i}}{B_{\nu\to i}}f'\left(K_{\nu\to i}\right). \longrightarrow H_i=-\sum_{\mu=1}^{M}\frac{\Phi_{\nu i}}{B_{\mu\to i}}f'\left(K_{\mu\to i}\right) \xrightarrow{\text{Taylor expansion}} \boxed{H_i\approx\sum_{\mu=1}^{M}\Phi_{\mu i}\hat{a}_\mu+\Gamma\hat{x}_i.}$$

$$f\left(u\right)\equiv\frac{u^2}{2}\Theta\left(-u\right)$$

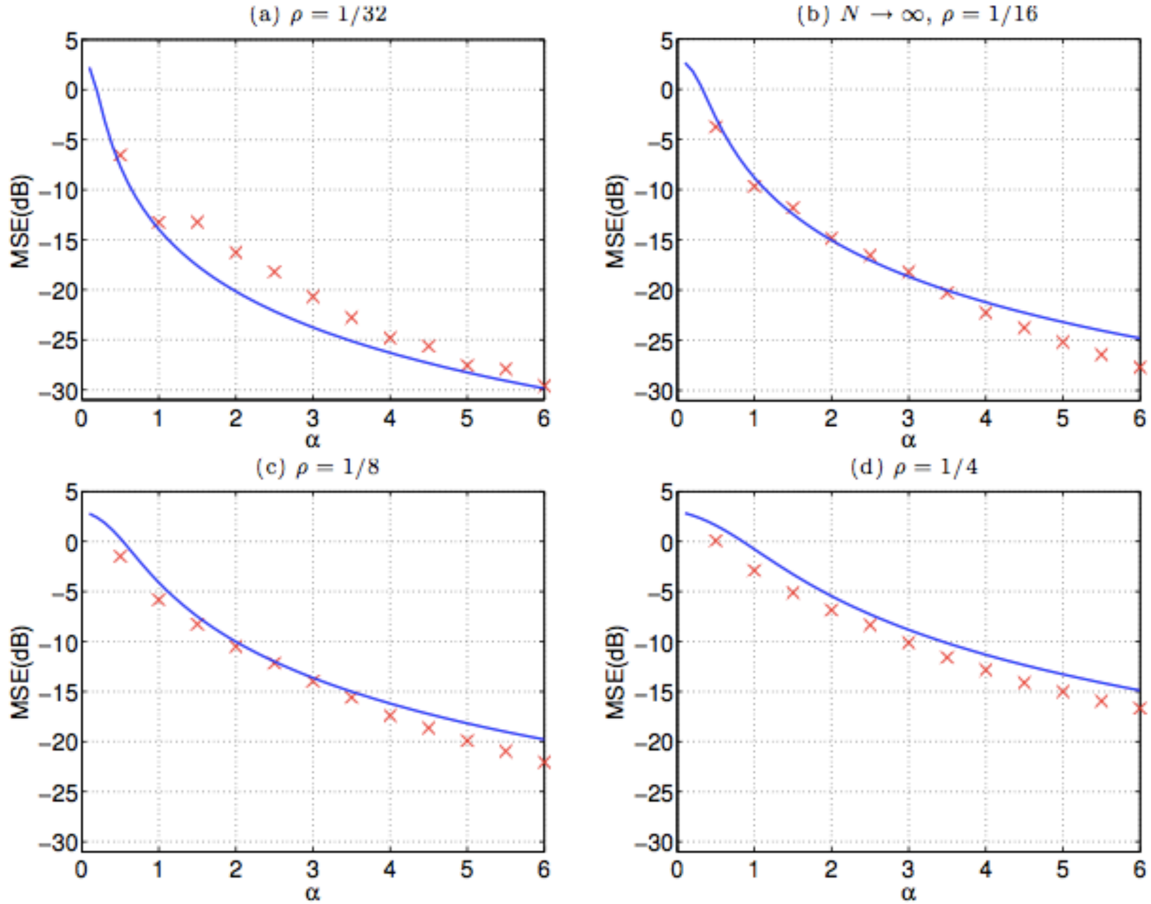
$$\Gamma\equiv\frac{1}{NB}\sum_{\mu=1}^{M}f''\left(K_\mu\right),\ \hat{a}_\mu\equiv-\frac{1}{B}f'\left(K_\mu\right)$$

Developed algorithm

Algorithm 2: CAVITY-INSPIRED SIGNAL RECOVERY$(\mathrm{B}, \mathbf{x}^*, \mathbf{H}^*)$

1) **Initialization** :
   X Seed : $\hat{\mathbf{x}}_0 \leftarrow \hat{\mathbf{x}}^*$
   H Seed : $\mathbf{H}_0 \leftarrow \mathbf{H}^*$
   Counter : $k \leftarrow 0$
2) **Counter Increase** :
   $k \leftarrow k+1$
3) **One-Sided Quadratic Gradient Descent** :
   $\mathbf{H}_k \leftarrow \mathbf{H}_{k-1} - \mathrm{B}^{-1}(\mathrm{Y}\Phi)^{\mathrm{T}} f'(\mathrm{Y}\Phi\hat{\mathbf{x}}_{k-1})$
4) **Assessment of Onsager Coefficient** :
   $\Gamma \leftarrow (\mathrm{NB})^{-1}\mathbf{1}^{\mathrm{T}} f''(\mathrm{Y}\Phi\hat{\mathbf{x}}_{k-1})$
5) **Self-feedback Cancellation** :
   $\tilde{\mathbf{H}}_k \leftarrow \mathbf{H}_k + \Gamma\hat{\mathbf{x}}_{k-1}$
6) **Shrinkage (l_1-Gradient Descent)** :
   $(\mathbf{u})_i \leftarrow \mathrm{sign}((\tilde{\mathbf{H}})_i)\max\{|(\tilde{\mathbf{H}})_i| - 1, 0\}$, for all i,
7) **Normalization** :
   $\hat{\mathbf{x}}_k \leftarrow \sqrt{N}\frac{\mathbf{u}}{||\mathbf{u}||_2}$
8) **Iteration** : Repeat from 2) until convergence.

Experiment result for algorithm CISR



Time cost:

CISR is 50 times faster!

	$K = 4$	$K = 8$	$K = 16$	$K = 32$
RFPI	25.7636(10.0799)s	27.8293(3.3566)s	33.3552(3.2914)s	35.4574(3.3869)s
CISR	0.0385(0.0583)s	0.0705(0.1058)s	0.0245(0.0346)s	0.0247(0.0207) s

Thank you!