ex Sparse normal means $y_i = \mu_i + \varepsilon_i$

$\varepsilon_i \sim N(0,1)$ and $MSE = \frac{1}{n}\sum_{i=1}^{n} (\hat{\mu}_i - \mu_i)^2$

(to translate to previous setting let new observation be $j \sim \text{unif}\{1,\dots,n\}$ and $y = \mu_j + \varepsilon$)

Hard thresholding: $\hat{\mu}_i = y_i \mathbb{1}\{|y_i| > \tau\}$

$$\mathbb{E}\, MSE = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\, (y_i - \mu_i)^2 \mathbb{1}\{|y_i| > \tau\} \to \varepsilon_i^2 \cdot \mathbb{1}\{|y_i| > \tau\}$$

$$+ \underbrace{\mathbb{E}\, \mu_i^2 \mathbb{1}\{|y_i| \le \tau\}}_{\mu_i^2 \mathbb{P}\{|y_i| \le \tau\}}$$

if $|\mu_i|$ is small then for τ large enough $\mathbb{P}\{|y_i| \le \tau\}$ is large (≈ 1) yet $\mu_i^2 \mathbb{P}\{|y_i| \le \tau\} \approx 0$.

if $|\mu_i|$ is large then τ ~~large~~ not too large $\Rightarrow \mathbb{P}\{|y_i| > \tau\} \approx 0$ and $\varepsilon_i^2 \mathbb{1}\{|y_i| > \tau\} \approx 0$.

Sparsity: suppose for $i \in S$, $|\mu_i| > \underline{\mu}$ and $i \notin S$, $\mu_i = 0$.

Set $\tau = \sqrt{2 \log(2n/\alpha)}$

fact: $\mathbb{P}\{|\varepsilon_i| > u\} \le 2e^{-u^2/2}$ for $N(0,1)$

so $\mathbb{P}\{|\varepsilon_i| > u \text{ for any } i = 1,\dots,n\} \le 2ne^{-u^2/2} = \alpha$