Information-Theoretic Limits of Group Testing: Phase Transitions, Noisy Tests, and Partial Recovery

Jonathan Scarlett
jonathan.scarlett@epfl.ch

Laboratory for Information and Inference Systems (LIONS)
École Polytechnique Fédérale de Lausanne (EPFL)
Switzerland

Bristol Statistics Seminar
[October, 2015]

Joint work with
Volkan Cevher @ LIONS

lions@epfl

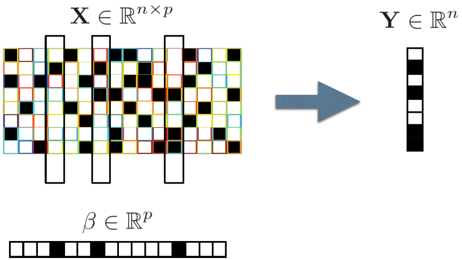
MARIE CURIE ACTIONS

FNSNF
FONDS NATIONAL SUISSE
SCHWEIZERISCHER NATIONALFONDS
FONDO NAZIONALE SVIZZERO
SWISS NATIONAL SCIENCE FOUNDATION

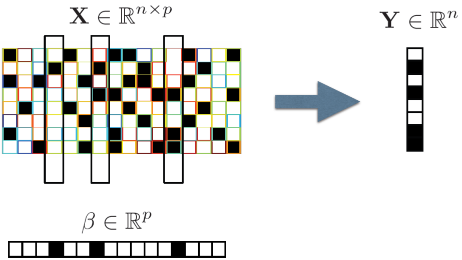
erc

EPFL

Group testing

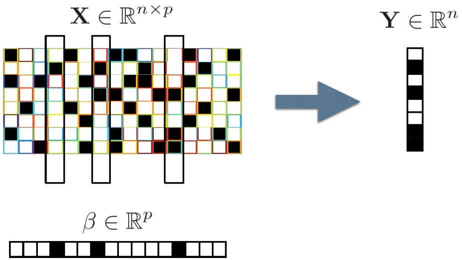


Group testing



- In this talk:
  - Defective set $S \sim \text{Uniform}\binom{p}{k}$
  - Measurement matrix $\mathbf{X} \sim \prod_{i=1}^{n} \prod_{j=1}^{p} P_X(x_{i,j})$ with $P_X \sim \text{Bernoulli}(\nu/k)$

Group testing



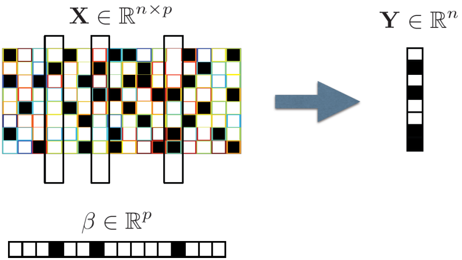
- In this talk:
  - Defective set $S \sim \text{Uniform}\binom{p}{k}$
  - Measurement matrix $\mathbf{X} \sim \prod_{i=1}^{n} \prod_{j=1}^{p} P_X(x_{i,j})$ with $P_X \sim \text{Bernoulli}(\nu/k)$
- Perfect recovery

$$P_\text{e} := \mathbb{P}[\hat{S} \neq S]$$

- Partial recovery

$$P_\text{e}(d_\text{max}) := \mathbb{P}\Big[|S \backslash \hat{S}| > d_\text{max} \cup |\hat{S} \backslash S| > d_\text{max}\Big]$$

Group testing



- In this talk:
  - Defective set $S \sim \text{Uniform}\binom{p}{k}$
  - Measurement matrix $\mathbf{X} \sim \prod_{i=1}^{n} \prod_{j=1}^{p} P_X(x_{i,j})$ with $P_X \sim \text{Bernoulli}(\nu/k)$
- Perfect recovery

$$P_{\text{e}} := \mathbb{P}[\hat{S} \neq S]$$

- Partial recovery

$$P_{\text{e}}(d_{\text{max}}) := \mathbb{P}\Big[|S \backslash \hat{S}| > d_{\text{max}} \cup |\hat{S} \backslash S| > d_{\text{max}}\Big]$$

- **Goal**: Conditions on n for $P_{\text{e}} \to 0$ or $P_{\text{e}}(d_{\text{max}}) \to 0$

Noiseless Group Testing (Exact Recovery)

- Observation model

$$Y = \mathbf{1}\left\{ \bigcup_{i \in S} \{X_i = 1\} \right\}$$

  - Bernoulli measurements, sparsity $k = O(p^{\theta})$

Noiseless Group Testing (Exact Recovery)

- Observation model

$$Y = \mathbf{1}\Big\{ \bigcup_{i \in S} \{X_i = 1\} \Big\}$$

  - Bernoulli measurements, sparsity $k = O(p^\theta)$

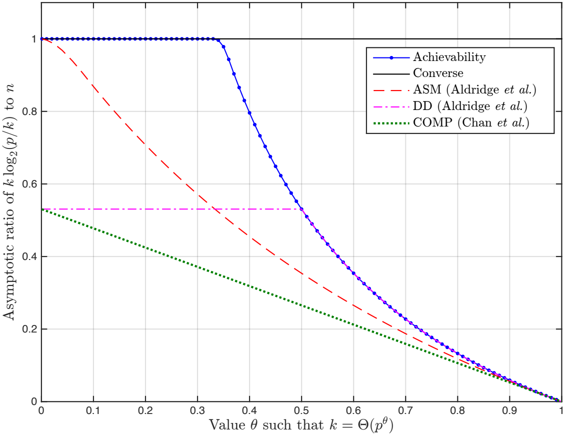
- Sufficient for $P_\mathrm{e} \to 0$:

$$n \geq \inf_{\nu > 0} \max\left\{ \frac{\theta}{e^{-\nu}\nu(1-\theta)}, \frac{1}{H_2(e^{-\nu})} \right\} \left(k \log \frac{p}{k} \right)(1+\eta)$$

- Necessary for $P_\mathrm{e} \not\to 1$:

$$n \geq \frac{k \log \frac{p}{k}}{\log 2}(1-\eta)$$

Noiseless Group Testing (Exact Recovery)



Key Implication: i.i.d. Bernoulli measurements are asymptotically as good as optimal adaptive measurements when $k = O(p^{1/3})$.

Noisy Group Testing (Exact Recovery)

- Observation model

$$Y = \mathbf{1}\left\{ \bigcup_{i \in S} \{X_i = 1\} \right\} \oplus Z$$

  where $Z \sim \mathrm{Bernoulli}(\rho)$
  - Bernoulli measurements, sparsity $k = O(p^{\theta})$

Noisy Group Testing (Exact Recovery)

- Observation model
$$Y = \mathbf{1}\bigg\{ \bigcup_{i \in S} \{X_i = 1\} \bigg\} \oplus Z$$
where $Z \sim \mathrm{Bernoulli}(\rho)$
  - Bernoulli measurements, sparsity $k = O(p^\theta)$
- Sufficient for $P_\mathrm{e} \to 0$:
$$n \ge \inf_{\delta_2 \in (0,1)} \max\bigg\{ \zeta(\rho, \delta_2, \theta), \frac{1}{\log 2 - H_2(\rho)} \bigg\} \bigg(k \log \frac{p}{k} \bigg) (1 + \eta)$$
where
$$\zeta(\rho, \delta_2, \theta) := \frac{2}{\log 2} \max\bigg\{ \frac{2(1 + \frac{1}{3}\delta_2(1 - 2\rho)) \frac{\theta}{1-\theta}}{\delta_2^2 (1 - 2\rho)^2}, \frac{\frac{1+2\theta}{1-\theta}}{(1 - 2\rho) \log \frac{1-\rho}{\rho} (1 - \delta_2)} \bigg\}.$$

Noisy Group Testing (Exact Recovery)

- Observation model

$$Y = \mathbf{1}\left\{ \bigcup_{i \in S} \{X_i = 1\} \right\} \oplus Z$$

  where $Z \sim \mathrm{Bernoulli}(\rho)$
  - Bernoulli measurements, sparsity $k = O(p^\theta)$
- Sufficient for $P_\mathrm{e} \to 0$:

$$n \ge \inf_{\delta_2 \in (0,1)} \max\left\{ \zeta(\rho, \delta_2, \theta), \frac{1}{\log 2 - H_2(\rho)} \right\} \left(k \log \frac{p}{k} \right) (1 + \eta)$$

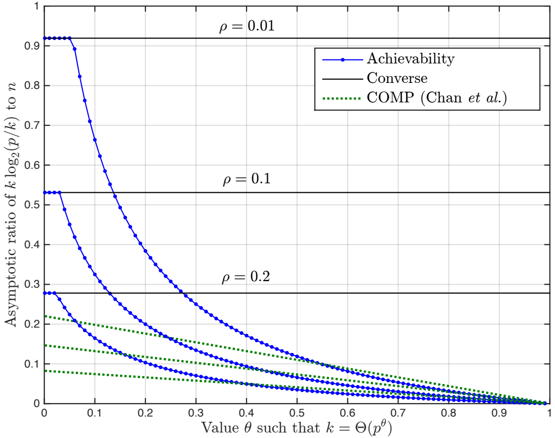
  where

$$\zeta(\rho, \delta_2, \theta) := \frac{2}{\log 2} \max\left\{ \frac{2(1 + \frac{1}{3}\delta_2(1 - 2\rho)) \frac{\theta}{1-\theta}}{\delta_2^2 (1 - 2\rho)^2}, \frac{\frac{1+2\theta}{1-\theta}}{(1 - 2\rho) \log \frac{1-\rho}{\rho} (1 - \delta_2)} \right\}.$$

- Necessary for $P_\mathrm{e} \not\to 1$:

$$n \le \frac{k \log \frac{p}{k}}{\log 2 - H_2(\rho)} (1 - \eta)$$

Noisy Group Testing (Exact Recovery)



Key Implication: i.i.d. Bernoulli measurements are asymptotically as good as optimal adaptive measurements when $k = O(p^\theta)$ for sufficiently small θ.

Partial Recovery

- Recovery criterion

$$P_e(d_{\max}) := \mathbb{P}\Big[|S \backslash \hat{S}| > d_{\max} \cup |\hat{S} \backslash S| > d_{\max}\Big]$$

where $d_{\max} = \lfloor \alpha^* k \rfloor$ for some $\alpha^* \in (0, 1)$

Partial Recovery

- Recovery criterion

$$P_{\mathrm{e}}(d_{\max}) := \mathbb{P}\Big[|S\backslash\hat{S}| > d_{\max} \cup |\hat{S}\backslash S| > d_{\max}\Big]$$

  where $d_{\max} = \lfloor \alpha^* k \rfloor$ for some $\alpha^* \in (0,1)$

- Sufficient for $P_{\mathrm{e}}(d_{\max}) \to 0$:

$$n \geq \frac{k \log \frac{p}{k}}{\log 2 - H_2(\rho)}(1+\eta)$$

- Necessary for $P_{\mathrm{e}}(d_{\max}) \not\to 1$:

$$n \leq \frac{(1-\alpha^*) k \log \frac{p}{k}}{\log 2 - H_2(\rho)}(1-\eta)$$

Partial Recovery

- Recovery criterion

$$P_e(d_{\max}) := \mathbb{P}\Big[|S\backslash\hat{S}| > d_{\max} \cup |\hat{S}\backslash S| > d_{\max}\Big]$$

  where $d_{\max} = \lfloor \alpha^* k \rfloor$ for some $\alpha^* \in (0,1)$
- Sufficient for $P_e(d_{\max}) \to 0$:

$$n \geq \frac{k \log \frac{p}{k}}{\log 2 - H_2(\rho)}(1+\eta)$$

- Necessary for $P_e(d_{\max}) \not\to 1$:

$$n \leq \frac{(1-\alpha^*) k \log \frac{p}{k}}{\log 2 - H_2(\rho)}(1-\eta)$$

- **For small θ, the reduction is at most a factor $1-\alpha^*$ asymptotically**

Channel coding

e.g. see [Wainwright, 2009], [Atia and Saligrama, 2012], [Aksoylar et al., 2013]

Thresholding Techniques for Channel Coding

- Mutual information

$$I(X;Y) := \sum_{x,y} P_{XY}(x,y) \log \frac{P_{Y|X}(y|x)}{P_Y(y)}$$

Thresholding Techniques for Channel Coding

- Mutual information

$$I(X;Y) := \sum_{x,y} P_{XY}(x,y) \log \frac{P_{Y|X}(y|x)}{P_Y(y)}$$

- Information density

$$\imath(x;y) := \log \frac{P_{Y|X}(y|x)}{P_Y(y)}$$

Thresholding Techniques for Channel Coding

- Mutual information

$$I(X;Y) := \sum_{x,y} P_{XY}(x,y) \log \frac{P_{Y|X}(y|x)}{P_Y(y)}$$

- Information density

$$\imath(x;y) := \log \frac{P_{Y|X}(y|x)}{P_Y(y)}$$

- Non-asymptotic channel coding bounds [Han, 2003]

$$P_{\rm e} \le \mathbb{P}\Big[\imath^n(\mathbf{X};\mathbf{Y}) \le nR - \log\delta\Big] + \delta$$
$$P_{\rm e} \ge \mathbb{P}\Big[\imath^n(\mathbf{X};\mathbf{Y}) \le nR + \log\delta\Big] - \delta$$

Non-Asymptotic Bounds

- Information density

$$\imath(x_{s_{\mathrm{dif}}}; y | x_{s_{\mathrm{eq}}}) := \log \frac{P_{Y|X_{s_{\mathrm{dif}}} X_{s_{\mathrm{eq}}}}(y | x_{s_{\mathrm{dif}}}, x_{s_{\mathrm{eq}}})}{P_{Y|X_{s_{\mathrm{eq}}}}(y | x_{s_{\mathrm{eq}}})}$$

  where $(s_{\mathrm{dif}}, s_{\mathrm{eq}})$ is a partition of s

Non-Asymptotic Bounds

- Information density

$$\imath(x_{s_{\text{dif}}}; y | x_{s_{\text{eq}}}) := \log \frac{P_{Y|X_{s_{\text{dif}}} X_{s_{\text{eq}}}}(y | x_{s_{\text{dif}}}, x_{s_{\text{eq}}})}{P_{Y|X_{s_{\text{eq}}}}(y | x_{s_{\text{eq}}})}$$

  where $(s_{\text{dif}}, s_{\text{eq}})$ is a partition of s

- Achievability

$$P_{\text{e}} \leq \mathbb{P}\bigg[\bigcup_{s_{\text{dif}}, s_{\text{eq}}} \bigg\{ \imath^n(\mathbf{X}_{s_{\text{dif}}}; \mathbf{Y} | \mathbf{X}_{s_{\text{eq}}}) \leq \log \binom{p-k}{|s_{\text{dif}}|} + \Delta \bigg\} \bigg] + \delta_1$$

Non-Asymptotic Bounds

- Information density

$$\imath(x_{s_{\mathrm{dif}}}; y | x_{s_{\mathrm{eq}}}) := \log \frac{P_{Y|X_{s_{\mathrm{dif}}} X_{s_{\mathrm{eq}}}}(y | x_{s_{\mathrm{dif}}}, x_{s_{\mathrm{eq}}})}{P_{Y|X_{s_{\mathrm{eq}}}}(y | x_{s_{\mathrm{eq}}})}$$

  where $(s_{\mathrm{dif}}, s_{\mathrm{eq}})$ is a partition of s

- Achievability

$$P_{\mathrm{e}} \le \mathbb{P}\Bigg[\bigcup_{s_{\mathrm{dif}}, s_{\mathrm{eq}}} \bigg\{ \imath^n(\mathbf{X}_{s_{\mathrm{dif}}}; \mathbf{Y} | \mathbf{X}_{s_{\mathrm{eq}}}) \le \log \binom{p-k}{|s_{\mathrm{dif}}|} + \Delta \bigg\} \Bigg] + \delta_1$$

- Converse

$$P_{\mathrm{e}} \ge \mathbb{P}\Bigg[\imath^n(\mathbf{X}_{s_{\mathrm{dif}}}; \mathbf{Y} | \mathbf{X}_{s_{\mathrm{eq}}}) \le \log \binom{p-k+|s_{\mathrm{dif}}|}{|s_{\mathrm{dif}}|} + \log \delta_1 \Bigg] - \delta_1$$

Achievability Bound

- Decoder that searches for the unique set $s \in \mathcal{S}$ such that

$$\imath^n(\mathbf{x}_{s_{\mathrm{dif}}}; \mathbf{y} | \mathbf{x}_{s_{\mathrm{eq}}}) > \gamma_{|s_{\mathrm{dif}}|}$$

  for all $(s_{\mathrm{dif}}, s_{\mathrm{eq}})$ of s with $s_{\mathrm{dif}} \neq \emptyset$.

Achievability Bound

- Decoder that searches for the unique set $s \in \mathcal{S}$ such that

$$\imath^n(\mathbf{x}_{s_{\mathrm{dif}}}; \mathbf{y}|\mathbf{x}_{s_{\mathrm{eq}}}) > \gamma_{|s_{\mathrm{dif}}|}$$

  for all $(s_{\mathrm{dif}}, s_{\mathrm{eq}})$ of s with $s_{\mathrm{dif}} \neq \emptyset$.

- Initial bound:

$$P_{\mathrm{e}} \le \mathbb{P}\Bigg[\bigcup_{(s_{\mathrm{dif}}, s_{\mathrm{eq}})} \Big\{\imath^n(\mathbf{X}_{s_{\mathrm{dif}}}; \mathbf{Y}|\mathbf{X}_{s_{\mathrm{eq}}}) \le \gamma_{|s_{\mathrm{dif}}|}\Big\}\Bigg] + \sum_{\bar{s} \in \mathcal{S} \setminus \{s\}} \mathbb{P}\Big[\imath^n(\mathbf{X}_{\bar{s} \setminus s}; \mathbf{Y}|\mathbf{X}_{\bar{s} \cap s}) > \gamma_{|s_{\mathrm{dif}}|}\Big],$$

Achievability Bound

- Decoder that searches for the unique set $s \in \mathcal{S}$ such that

$$\imath^n(\mathbf{x}_{s_{\mathrm{dif}}};\mathbf{y}|\mathbf{x}_{s_{\mathrm{eq}}}) > \gamma_{|s_{\mathrm{dif}}|}$$

  for all $(s_{\mathrm{dif}}, s_{\mathrm{eq}})$ of s with $s_{\mathrm{dif}} \neq \emptyset$.
- Analysis of second term (with $\ell := |\bar{s} \backslash s|$):

$$\mathbb{P}\Big[\imath^n(\mathbf{X}_{\bar{s}\backslash s};\mathbf{Y}|\mathbf{X}_{\bar{s}\cap s}) > \gamma_\ell\Big]$$

$$= \sum_{\mathbf{x}_{\bar{s}\cap s},\mathbf{x}_{\bar{s}\backslash s},\mathbf{y}} P_X^{n\times(k-\ell)}(\mathbf{x}_{\bar{s}\cap s}) P_X^{n\times\ell}(\mathbf{x}_{\bar{s}\backslash s}) P^n_{Y|X_{s_{\mathrm{eq}}}}(\mathbf{y}|\mathbf{x}_{\bar{s}\cap s})$$

$$\times \mathbb{1}\left\{ \log \frac{P^n_{Y|X_{s_{\mathrm{dif}}}X_{s_{\mathrm{eq}}}}(\mathbf{y}|\mathbf{x}_{\bar{s}\backslash s}, \mathbf{x}_{\bar{s}\cap s})}{P^n_{Y|X_{s_{\mathrm{eq}}}}(\mathbf{y}|\mathbf{x}_{\bar{s}\cap s})} > \gamma_\ell \right\}$$

Achievability Bound

- Decoder that searches for the unique set $s \in \mathcal{S}$ such that

$$\imath^n(\mathbf{x}_{s_{\mathrm{dif}}}; \mathbf{y} | \mathbf{x}_{s_{\mathrm{eq}}}) > \gamma_{|s_{\mathrm{dif}}|}$$

  for all $(s_{\mathrm{dif}}, s_{\mathrm{eq}})$ of s with $s_{\mathrm{dif}} \neq \emptyset$.
- Analysis of second term (with $\ell := |\bar{s} \backslash s|$):

$$\mathbb{P}\Big[\imath^n(\mathbf{X}_{\bar{s}\backslash s}; \mathbf{Y} | \mathbf{X}_{\bar{s}\cap s}) > \gamma_\ell\Big]$$
$$= \sum_{\mathbf{x}_{\bar{s}\cap s}, \mathbf{x}_{\bar{s}\backslash s}, \mathbf{y}} P_X^{n\times(k-\ell)}(\mathbf{x}_{\bar{s}\cap s}) P_X^{n\times\ell}(\mathbf{x}_{\bar{s}\backslash s}) P^n_{Y|X_{s_{\mathrm{eq}}}}(\mathbf{y}|\mathbf{x}_{\bar{s}\cap s})$$
$$\times \mathbb{1}\left\{ \log \frac{P^n_{Y|X_{s_{\mathrm{dif}}}X_{s_{\mathrm{eq}}}}(\mathbf{y}|\mathbf{x}_{\bar{s}\backslash s}, \mathbf{x}_{\bar{s}\cap s})}{P^n_{Y|X_{s_{\mathrm{eq}}}}(\mathbf{y}|\mathbf{x}_{\bar{s}\cap s})} > \gamma_\ell \right\}$$
$$\leq \sum_{\mathbf{x}_{\bar{s}\cap s}, \mathbf{x}_{\bar{s}\backslash s}, \mathbf{y}} P_X^{n\times(k-\ell)}(\mathbf{x}_{\bar{s}\cap s}) P_X^{n\times\ell}(\mathbf{x}_{\bar{s}\backslash s}) P^n_{Y|X_{s_{\mathrm{dif}}}X_{s_{\mathrm{eq}}}}(\mathbf{y}|\mathbf{x}_{\bar{s}\backslash s}, \mathbf{x}_{\bar{s}\cap s}) e^{-\gamma_\ell}$$

Achievability Bound

- Decoder that searches for the unique set $s \in \mathcal{S}$ such that

$$\imath^n(\mathbf{x}_{s_{\mathrm{dif}}};\mathbf{y}|\mathbf{x}_{s_{\mathrm{eq}}}) > \gamma_{|s_{\mathrm{dif}}|}$$

for all $(s_{\mathrm{dif}}, s_{\mathrm{eq}})$ of s with $s_{\mathrm{dif}} \neq \emptyset$.

- Analysis of second term (with $\ell := |\bar{s}\backslash s|$):

$$\begin{aligned}
&\mathbb{P}\Big[\imath^n(\mathbf{X}_{\bar{s}\backslash s};\mathbf{Y}|\mathbf{X}_{\bar{s}\cap s}) > \gamma_\ell\Big] \\
&\quad= \sum_{\mathbf{x}_{\bar{s}\cap s},\mathbf{x}_{\bar{s}\backslash s},\mathbf{y}} P_X^{n\times(k-\ell)}(\mathbf{x}_{\bar{s}\cap s}) P_X^{n\times\ell}(\mathbf{x}_{\bar{s}\backslash s}) P^n_{Y|X_{s_{\mathrm{eq}}}}(\mathbf{y}|\mathbf{x}_{\bar{s}\cap s}) \\
&\qquad\qquad \times \mathbb{1}\left\{\log\frac{P^n_{Y|X_{s_{\mathrm{dif}}}X_{s_{\mathrm{eq}}}}(\mathbf{y}|\mathbf{x}_{\bar{s}\backslash s},\mathbf{x}_{\bar{s}\cap s})}{P^n_{Y|X_{s_{\mathrm{eq}}}}(\mathbf{y}|\mathbf{x}_{\bar{s}\cap s})} > \gamma_\ell\right\} \\
&\quad\leq \sum_{\mathbf{x}_{\bar{s}\cap s},\mathbf{x}_{\bar{s}\backslash s},\mathbf{y}} P_X^{n\times(k-\ell)}(\mathbf{x}_{\bar{s}\cap s}) P_X^{n\times\ell}(\mathbf{x}_{\bar{s}\backslash s}) P^n_{Y|X_{s_{\mathrm{dif}}}X_{s_{\mathrm{eq}}}}(\mathbf{y}|\mathbf{x}_{\bar{s}\backslash s},\mathbf{x}_{\bar{s}\cap s}) e^{-\gamma_\ell} \\
&\quad= e^{-\gamma_\ell}
\end{aligned}$$

Achievability Bound

- Decoder that searches for the unique set $s \in \mathcal{S}$ such that

$$\imath^n(\mathbf{x}_{s_{\mathrm{dif}}}; \mathbf{y} | \mathbf{x}_{s_{\mathrm{eq}}}) > \gamma_{|s_{\mathrm{dif}}|}$$

  for all $(s_{\mathrm{dif}}, s_{\mathrm{eq}})$ of s with $s_{\mathrm{dif}} \neq \emptyset$.
- After some re-arrangements and choosing $\{\gamma_\ell\}$ appropriately,

$$P_{\mathrm{e}} \leq \mathbb{P}\Bigg[\bigcup_{s_{\mathrm{dif}}, s_{\mathrm{eq}}} \bigg\{ \imath^n(\mathbf{X}_{s_{\mathrm{dif}}}; \mathbf{Y} | \mathbf{X}_{s_{\mathrm{eq}}}) \leq \log \binom{p-k}{|s_{\mathrm{dif}}|} + \Delta \bigg\} \Bigg] + \delta_1$$

Converse Bound

- Consider a genie that reveals part of the defective set, $S_{\mathrm{eq}} \subseteq S$, to the decoder. The decoder is left to estimate $S_{\mathrm{dif}} := S \backslash S_{\mathrm{eq}}$.
- Starting point:

$$P_{\mathrm{e}}(s_{\mathrm{eq}}) \geq \mathbb{P}[\mathcal{A}(s_{\mathrm{eq}})] - \mathbb{P}[\mathcal{A}(s_{\mathrm{eq}}) \cap \text{no error}],$$

where

$$\mathcal{A}(s_{\mathrm{eq}}) = \Big\{ \imath^n(\mathbf{X}_{S_{\mathrm{dif}}}; \mathbf{Y} | \mathbf{X}_{s_{\mathrm{eq}}}) \leq \gamma \Big\}.$$

Converse Bound

- Consider a genie that reveals part of the defective set, $S_{\text{eq}} \subseteq S$, to the decoder. The decoder is left to estimate $S_{\text{dif}} := S \backslash S_{\text{eq}}$.
- Analysis of second term:

$$\mathbb{P}[\mathcal{A}(s_{\text{eq}}) \cap \text{no error}]$$
$$= \sum_{s_{\text{dif}} \in \mathcal{S}_{\text{dif}}(s_{\text{eq}})} \frac{1}{\binom{p-k+\ell}{\ell}} \sum_{(\mathbf{x},\mathbf{y}) \in \mathcal{D}(s_{\text{dif}}|s_{\text{eq}})} P_X^{n \times p}(\mathbf{x})$$
$$\times P^n_{Y|X_{s_{\text{dif}}} X_{s_{\text{eq}}}}(\mathbf{y}|\mathbf{x}_{s_{\text{dif}}}, \mathbf{x}_{s_{\text{eq}}}) \mathbb{1}\left\{ \log \frac{P^n_{Y|X_{s_{\text{dif}}} X_{s_{\text{eq}}}}(\mathbf{y}|\mathbf{x}_{s_{\text{dif}}}, \mathbf{x}_{s_{\text{eq}}})}{P^n_{Y|X_{s_{\text{eq}}}}(\mathbf{y}|\mathbf{x}_{s_{\text{eq}}})} \le \gamma \right\}$$

Converse Bound

- Consider a genie that reveals part of the defective set, $S_{\text{eq}} \subseteq S$, to the decoder. The decoder is left to estimate $S_{\text{dif}} := S \backslash S_{\text{eq}}$.
- Analysis of second term:

$$\mathbb{P}[\mathcal{A}(s_{\text{eq}}) \cap \text{no error}]$$

$$= \sum_{s_{\text{dif}} \in \mathcal{S}_{\text{dif}}(s_{\text{eq}})} \frac{1}{\binom{p-k+\ell}{\ell}} \sum_{(\mathbf{x},\mathbf{y}) \in \mathcal{D}(s_{\text{dif}}|s_{\text{eq}})} P_X^{n \times p}(\mathbf{x})$$

$$\times P^n_{Y|X_{s_{\text{dif}}} X_{s_{\text{eq}}}}(\mathbf{y}|\mathbf{x}_{s_{\text{dif}}}, \mathbf{x}_{s_{\text{eq}}}) \mathbb{1}\left\{ \log \frac{P^n_{Y|X_{s_{\text{dif}}} X_{s_{\text{eq}}}}(\mathbf{y}|\mathbf{x}_{s_{\text{dif}}}, \mathbf{x}_{s_{\text{eq}}})}{P^n_{Y|X_{s_{\text{eq}}}}(\mathbf{y}|\mathbf{x}_{s_{\text{eq}}})} \leq \gamma \right\}$$

$$\leq \frac{1}{\binom{p-k+\ell}{\ell}} \sum_{s_{\text{dif}} \in \mathcal{S}_{\text{dif}}(s_{\text{eq}})} \sum_{(\mathbf{x},\mathbf{y}) \in \mathcal{D}(s_{\text{dif}}|s_{\text{eq}})} P_X^{n \times p}(\mathbf{x}) P^n_{Y|X_{s_{\text{eq}}}}(\mathbf{y}|\mathbf{x}_{s_{\text{eq}}}) e^{\gamma}$$

Converse Bound

- Consider a genie that reveals part of the defective set, $S_{\text{eq}} \subseteq S$, to the decoder. The decoder is left to estimate $S_{\text{dif}} := S \backslash S_{\text{eq}}$.
- Analysis of second term:

$$\mathbb{P}[\mathcal{A}(s_{\text{eq}}) \cap \text{no error}]$$

$$= \sum_{s_{\text{dif}} \in \mathcal{S}_{\text{dif}}(s_{\text{eq}})} \frac{1}{\binom{p-k+\ell}{\ell}} \sum_{(\mathbf{x},\mathbf{y}) \in \mathcal{D}(s_{\text{dif}}|s_{\text{eq}})} P_X^{n\times p}(\mathbf{x})$$

$$\times P^n_{Y|X_{s_{\text{dif}}}X_{s_{\text{eq}}}}(\mathbf{y}|\mathbf{x}_{s_{\text{dif}}}, \mathbf{x}_{s_{\text{eq}}}) \mathbb{1}\left\{ \log \frac{P^n_{Y|X_{s_{\text{dif}}}X_{s_{\text{eq}}}}(\mathbf{y}|\mathbf{x}_{s_{\text{dif}}}, \mathbf{x}_{s_{\text{eq}}})}{P^n_{Y|X_{s_{\text{eq}}}}(\mathbf{y}|\mathbf{x}_{s_{\text{eq}}})} \le \gamma \right\}$$

$$\le \frac{1}{\binom{p-k+\ell}{\ell}} \sum_{s_{\text{dif}} \in \mathcal{S}_{\text{dif}}(s_{\text{eq}})} \sum_{(\mathbf{x},\mathbf{y}) \in \mathcal{D}(s_{\text{dif}}|s_{\text{eq}})} P_X^{n\times p}(\mathbf{x}) P^n_{Y|X_{s_{\text{eq}}}}(\mathbf{y}|\mathbf{x}_{s_{\text{eq}}}) e^{\gamma}$$

$$= \frac{e^{\gamma}}{\binom{p-k+\ell}{\ell}}.$$

Converse Bound

- Consider a genie that reveals part of the defective set, $S_{\rm eq} \subseteq S$, to the decoder. The decoder is left to estimate $S_{\rm dif} := S \backslash S_{\rm eq}$.
- After some re-arrangements and choosing $\{\gamma_\ell\}$ appropriately,

$$P_{\rm e} \ge \mathbb{P}\bigg[\imath^n(\mathbf{X}_{s_{\rm dif}}; \mathbf{Y} | \mathbf{X}_{s_{\rm eq}}) \le \log \binom{p-k+|s_{\rm dif}|}{|s_{\rm dif}|} + \log \delta_1\bigg] - \delta_1$$

Non-Asymptotic Bounds

- Information density

$$\imath(x_{s_{\mathrm{dif}}}; y | x_{s_{\mathrm{eq}}}) := \log \frac{P_{Y|X_{s_{\mathrm{dif}}} X_{s_{\mathrm{eq}}}}(y | x_{s_{\mathrm{dif}}}, x_{s_{\mathrm{eq}}})}{P_{Y|X_{s_{\mathrm{eq}}}}(y | x_{s_{\mathrm{eq}}})}$$

  where $(s_{\mathrm{dif}}, s_{\mathrm{eq}})$ is a partition of s

- Achievability

$$P_{\mathrm{e}} \leq \mathbb{P}\Bigg[\bigcup_{s_{\mathrm{dif}}, s_{\mathrm{eq}}} \bigg\{ \imath^n(\mathbf{X}_{s_{\mathrm{dif}}}; \mathbf{Y} | \mathbf{X}_{s_{\mathrm{eq}}}) \leq \log \binom{p-k}{|s_{\mathrm{dif}}|} + \Delta \bigg\} \Bigg] + \delta_1$$

- Converse

$$P_{\mathrm{e}} \geq \mathbb{P}\Bigg[\imath^n(\mathbf{X}_{s_{\mathrm{dif}}}; \mathbf{Y} | \mathbf{X}_{s_{\mathrm{eq}}}) \leq \log \binom{p-k+|s_{\mathrm{dif}}|}{|s_{\mathrm{dif}}|} + \log \delta_1 \Bigg] - \delta_1$$

Application Techniques

- General steps for application
  1. Bound the tail probabilities $\mathbb{P}\Big[\imath^n(\mathbf{X}_{s_{\mathrm{dif}}};\mathbf{Y}|\mathbf{X}_{s_{\mathrm{eq}}}) \le \mathbb{E}[\imath^n] \pm n\delta\Big]$

Application Techniques

- General steps for application
  1. Bound the tail probabilities $\mathbb{P}\Big[\imath^n(\mathbf{X}_{s_{\mathrm{dif}}};\mathbf{Y}|\mathbf{X}_{s_{\mathrm{eq}}}) \le \mathbb{E}[\imath^n] \pm n\delta\Big]$
  2. Control and simplify the remainder terms

Application Techniques

- General steps for application
  1. Bound the tail probabilities $\mathbb{P}\Big[\imath^n(\mathbf{X}_{s_{\mathrm{dif}}}; \mathbf{Y}|\mathbf{X}_{s_{\mathrm{eq}}}) \le \mathbb{E}[\imath^n] \pm n\delta\Big]$
  2. Control and simplify the remainder terms
- General form of corollaries: $P_{\mathrm{e}} \to 0$ if

$$n \ge \max_{(s_{\mathrm{dif}}, s_{\mathrm{eq}})} \frac{\log \binom{p-k}{|s_{\mathrm{dif}}|}}{I(X_{s_{\mathrm{dif}}}; Y|X_{s_{\mathrm{eq}}})}(1+\eta)$$

and $P_{\mathrm{e}} \to 1$ if

$$n \le \max_{(s_{\mathrm{dif}}, s_{\mathrm{eq}})} \frac{\log \binom{p-k+|s_{\mathrm{dif}}|}{|s_{\mathrm{dif}}|}}{I(X_{s_{\mathrm{dif}}}; Y|X_{s_{\mathrm{eq}}})}(1-\eta)$$

Concentration Inequalities

- Noiseless and noisy cases *(This one alone is enough for partial recovery!)*:

$$\mathbb{P}\left[\left|\imath^n(\mathbf{X}_{s_{\mathrm{dif}}};\mathbf{Y}|\mathbf{X}_{s_{\mathrm{eq}}}) - nI(\ell)\right| \geq n\delta\right] \leq 2\exp\left(-\frac{\delta^2 n}{4(8+\delta)}\right)$$

  - Proved using Bernstein's inequality
  - The only property used is that $|\mathcal{Y}| = 2$

Concentration Inequalities

- Noiseless and noisy cases *(This one alone is enough for partial recovery!)*:

$$\mathbb{P}\left[\left|\imath^n(\mathbf{X}_{s_{\mathrm{dif}}};\mathbf{Y}|\mathbf{X}_{s_{\mathrm{eq}}}) - nI(\ell)\right| \ge n\delta\right] \le 2\exp\left(-\frac{\delta^2 n}{4(8+\delta)}\right)$$

  - Proved using Bernstein's inequality
  - The only property used is that $|\mathcal{Y}| = 2$

- Noiseless case:

$$\mathbb{P}\left[\imath^n \le nI(\ell)(1-\delta_2)\right] \le \exp\left(-n\frac{\ell}{k}e^{-\nu}\nu\Big((1-\delta_2)\log(1-\delta_2)+\delta_2\Big)(1-\epsilon)\right)$$

  - Proved by writing $\imath^n(\mathbf{X}_{s_{\mathrm{dif}}};\mathbf{Y}|\mathbf{X}_{s_{\mathrm{eq}}})$ in terms of Binomial random variables, then applying Binomial tail bounds

Concentration Inequalities

- Noiseless and noisy cases *(This one alone is enough for partial recovery!)*:

$$\mathbb{P}\left[\left|\imath^n(\mathbf{X}_{s_{\mathrm{dif}}};\mathbf{Y}|\mathbf{X}_{s_{\mathrm{eq}}}) - nI(\ell)\right| \ge n\delta\right] \le 2\exp\left(-\frac{\delta^2 n}{4(8+\delta)}\right)$$

  - Proved using Bernstein's inequality
  - The only property used is that $|\mathcal{Y}| = 2$

- Noiseless case:

$$\mathbb{P}\left[\imath^n \le nI(\ell)(1-\delta_2)\right] \le \exp\left(-n\frac{\ell}{k}e^{-\nu}\nu\Big((1-\delta_2)\log(1-\delta_2)+\delta_2\Big)(1-\epsilon)\right)$$

  - Proved by writing $\imath^n(\mathbf{X}_{s_{\mathrm{dif}}};\mathbf{Y}|\mathbf{X}_{s_{\mathrm{eq}}})$ in terms of Binomial random variables, then applying Binomial tail bounds

- Noisy case with crossover probability ρ:

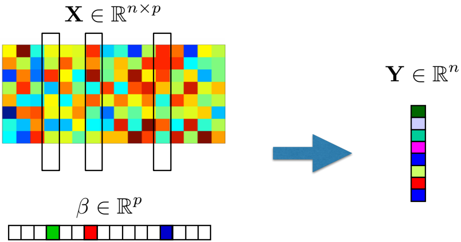
$$\mathbb{P}\left[\imath^n \le nI(\ell)(1-\delta_2)\right] \le \exp\left(-n\frac{\ell}{k}e^{-\nu}\nu\left(\frac{\delta_2^2(1-2\rho)^2}{2(1+\frac{1}{3}\delta_2(1-2\rho))}\right)(1-\epsilon)\right)$$

  - Proved using Bennet's inequality – may be somewhat crude

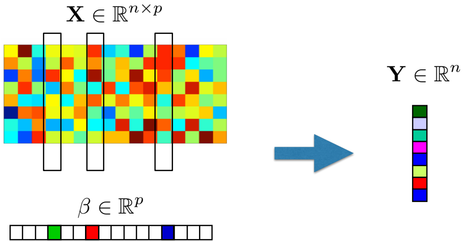
Recap of Results

- **Noiseless case (exact recovery):**
  - Sufficient for $P_{\mathrm{e}} \to 0$:
$$n \geq \inf_{\nu > 0} \max \left\{ \frac{\theta}{e^{-\nu}\nu(1-\theta)}, \frac{1}{H_2(e^{-\nu})} \right\} \left(k \log \frac{p}{k} \right) (1+\eta)$$
  - Necessary for $P_{\mathrm{e}} \not\to 1$:
$$n \geq \frac{k \log \frac{p}{k}}{\log 2} (1-\eta)$$
- **Noisy case (exact recovery):**
  - Sufficient for $P_{\mathrm{e}} \to 0$:
$$n \geq \inf_{\delta_2 \in (0,1)} \max \left\{ \zeta(\rho, \delta_2, \theta), \frac{1}{\log 2 - H_2(\rho)} \right\} \left(k \log \frac{p}{k} \right) (1+\eta).$$
  - Necessary for $P_{\mathrm{e}} \not\to 1$:
$$n \leq \frac{k \log \frac{p}{k}}{\log 2 - H_2(\rho)} (1-\eta)$$
- **Noiseless and noisy cases (partial recovery):**
  - Sufficient for $P_{\mathrm{e}}(d_{\max}) \to 0$:
$$n \geq \frac{k \log \frac{p}{k}}{\log 2 - H_2(\rho)} (1+\eta)$$
  - Necessary for $P_{\mathrm{e}}(d_{\max}) \not\to 1$:
$$n \leq \frac{(1-\alpha^*) k \log \frac{p}{k}}{\log 2 - H_2(\rho)} (1-\eta)$$

More General Suport Recovery

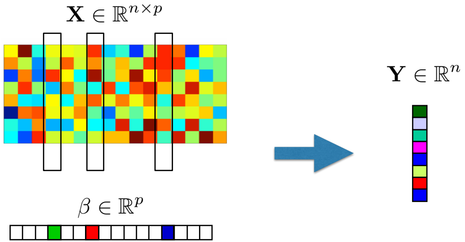


More General Suport Recovery



- General probabilistic models
  - Support $S \sim \mathrm{Uniform}\binom{p}{k}$
  - Measurement matrix $\mathbf{X} \sim \prod_{i=1}^{n} \prod_{j=1}^{p} P_X(x_{i,j})$
  - Non-zero entries $\beta_S \sim P_{\beta_S}$
  - Observations $(\mathbf{Y}|\mathbf{X}, \beta) \sim P_{\mathbf{Y}|\mathbf{X}_S \beta_S}$

More General Suport Recovery



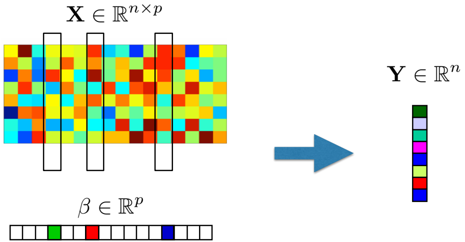
- General probabilistic models
  - Support $S \sim \mathrm{Uniform}\binom{p}{k}$
  - Measurement matrix $\mathbf{X} \sim \prod_{i=1}^{n} \prod_{j=1}^{p} P_X(x_{i,j})$
  - Non-zero entries $\beta_S \sim P_{\beta_S}$
  - Observations $(\mathbf{Y}|\mathbf{X}, \beta) \sim P_{\mathbf{Y}|\mathbf{X}_S \beta_S}$
- Support recovery

$$P_{\mathrm{e}} := \mathbb{P}[\hat{S} \neq S]$$

- Partial recovery

$$P_{\mathrm{e}}(d_{\max}) := \mathbb{P}\Big[|S \backslash \hat{S}| > d_{\max} \cup |\hat{S} \backslash S| > d_{\max}\Big]$$

More General Suport Recovery



- General probabilistic models
  - Support $S \sim \mathrm{Uniform}\binom{p}{k}$
  - Measurement matrix $\mathbf{X} \sim \prod_{i=1}^{n} \prod_{j=1}^{p} P_X(x_{i,j})$
  - Non-zero entries $\beta_S \sim P_{\beta_S}$
  - Observations $(\mathbf{Y}|\mathbf{X}, \beta) \sim P_{\mathbf{Y}|\mathbf{X}_S \beta_S}$
- Support recovery

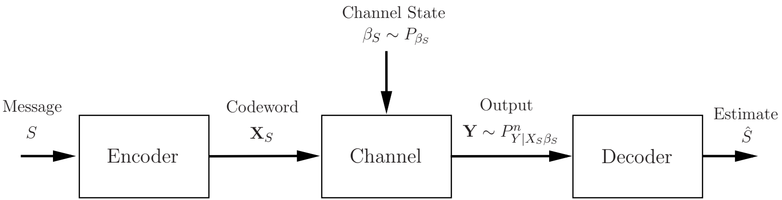
$$P_{\mathrm{e}} := \mathbb{P}[\hat{S} \neq S]$$

- Partial recovery

$$P_{\mathrm{e}}(d_{\max}) := \mathbb{P}\Big[|S \backslash \hat{S}| > d_{\max} \cup |\hat{S} \backslash S| > d_{\max}\Big]$$

- **Goal**: Conditions on n for $P_{\mathrm{e}} \to 0$ or $P_{\mathrm{e}}(d_{\max}) \to 0$

Channel coding



e.g. see [Wainwright, 2009], [Atia and Saligrama, 2012], [Aksoylar et al., 2013]

More General Suport Recovery

- Steps for applying generalized non-asymptotic bounds:
  1. Construct a "typical set" $\mathcal{T}_\beta$ of non-zero entries β_S

More General Suport Recovery

- Steps for applying generalized non-asymptotic bounds:
  1. Construct a "typical set" $\mathcal{T}_\beta$ of non-zero entries β_S
  2. Bound the tail probabilities $\mathbb{P}\Big[\imath^n(\mathbf{X}_{s_{\mathrm{dif}}};\mathbf{Y}|\mathbf{X}_{s_{\mathrm{eq}}},\beta_s) \le \mathbb{E}[\imath^n] \pm n\delta \,\Big|\, \beta_s = b_s\Big]$

More General Suport Recovery

- Steps for applying generalized non-asymptotic bounds:
  1. Construct a "typical set" $\mathcal{T}_\beta$ of non-zero entries β_S
  2. Bound the tail probabilities $\mathbb{P}\Big[\imath^n(\mathbf{X}_{s_{\mathrm{dif}}};\mathbf{Y}|\mathbf{X}_{s_{\mathrm{eq}}},\beta_s) \le \mathbb{E}[\imath^n] \pm n\delta \,\Big|\, \beta_s = b_s\Big]$
  3. Control and simplify the remainder terms

More General Suport Recovery

- Steps for applying generalized non-asymptotic bounds:
  1. Construct a "typical set" $\mathcal{T}_\beta$ of non-zero entries β_S
  2. Bound the tail probabilities $\mathbb{P}\Big[\imath^n(\mathbf{X}_{s_{\mathrm{dif}}};\mathbf{Y}|\mathbf{X}_{s_{\mathrm{eq}}},\beta_s) \le \mathbb{E}[\imath^n] \pm n\delta \,\Big|\, \beta_s = b_s\Big]$
  3. Control and simplify the remainder terms
- General form of corollaries: $P_{\mathrm{e}} \to 0$ if

$$n \ge \max_{(s_{\mathrm{dif}},s_{\mathrm{eq}})} \frac{\log \binom{p-k}{|s_{\mathrm{dif}}|}}{I(X_{s_{\mathrm{dif}}};Y|X_{s_{\mathrm{eq}}},\beta_s=b_s)}(1+\eta) \quad \text{for all } b_s \in \mathcal{T}_\beta$$

and $P_{\mathrm{e}} \to 1$ if

$$n \le \max_{(s_{\mathrm{dif}},s_{\mathrm{eq}})} \frac{\log \binom{p-k+|s_{\mathrm{dif}}|}{|s_{\mathrm{dif}}|}}{I(X_{s_{\mathrm{dif}}};Y|X_{s_{\mathrm{eq}}},\beta_s=b_s)}(1-\eta) \quad \text{for all } b_s \in \mathcal{T}_\beta$$

Exact Recovery for Linear and 1-bit Models

Observation models

$$Y = \langle X, \beta \rangle + Z$$

$$Y = \text{sign}\Big(\langle X, \beta \rangle + Z \Big)$$

Exact Recovery for Linear and 1-bit Models

Observation models

$$Y = \langle X, \beta \rangle + Z$$

$$Y = \mathrm{sign}\Big(\langle X, \beta \rangle + Z\Big)$$

- Case 1:
  - Gaussian X, discrete β_S, sparsity $k = \Theta(1)$, low SNR
  - Necessary and sufficient conditions:

Linear	1-bit
$\max_{(s_{\mathrm{dif}}, s_{\mathrm{eq}})} \frac{\lvert s_{\mathrm{dif}} \rvert \log p}{\frac{1}{2\sigma^2} \sum_{i \in s_{\mathrm{dif}}} b_i^2}$	$\max_{(s_{\mathrm{dif}}, s_{\mathrm{eq}})} \frac{\lvert s_{\mathrm{dif}} \rvert \log p}{\frac{1}{\pi\sigma^2} \sum_{i \in s_{\mathrm{dif}}} b_i^2}$

  - Only factor $\frac{\pi}{2}$ difference

Exact Recovery for Linear and 1-bit Models

Observation models

$$Y = \langle X, \beta \rangle + Z$$

$$Y = \text{sign}\Big(\langle X, \beta \rangle + Z\Big)$$

- Case 1:
  - Gaussian X, discrete β_S, sparsity $k = \Theta(1)$, low SNR
  - Necessary and sufficient conditions:

Linear	1-bit
$\max_{(s_{\text{dif}}, s_{\text{eq}})} \frac{\lvert s_{\text{dif}} \rvert \log p}{\frac{1}{2\sigma^2} \sum_{i \in s_{\text{dif}}} b_i^2}$	$\max_{(s_{\text{dif}}, s_{\text{eq}})} \frac{\lvert s_{\text{dif}} \rvert \log p}{\frac{1}{\pi\sigma^2} \sum_{i \in s_{\text{dif}}} b_i^2}$

  - Only factor $\frac{\pi}{2}$ difference
- Case 2:
  - Gaussian X, fixed β_S, sparsity $k = \Theta(p)$, moderate SNR
  - Conditions:

Linear	1-bit
$\Theta(p)$ sufficient	$\Omega\Big(p\sqrt{\log p}\Big)$ necessary

Partial Recovery for Linear and 1-bit Models

- Partial recovery

$$P_e(d_{\max}) := \mathbb{P}\Big[|S \backslash \hat{S}| > d_{\max} \cup |\hat{S} \backslash S| > d_{\max}\Big]$$

  - Linear and 1-bit models
  - Gaussian X, Gaussian β_S, sparsity $k = o(p)$, allowed errors $d_{\max} = \lfloor \alpha^* k \rfloor$

Partial Recovery for Linear and 1-bit Models

- Partial recovery

$$P_{\mathrm{e}}(d_{\max}) := \mathbb{P}\Big[|S \backslash \hat{S}| > d_{\max} \cup |\hat{S} \backslash S| > d_{\max}\Big]$$

  - Linear and 1-bit models
  - Gaussian X, Gaussian β_S, sparsity $k = o(p)$, allowed errors $d_{\max} = \lfloor \alpha^* k \rfloor$

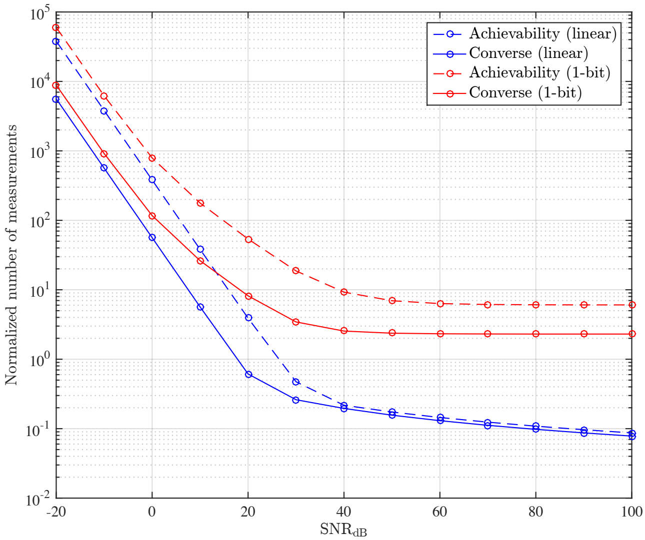
- Sufficient for $P_{\mathrm{e}} \to 0$:

$$n \geq \max_{\alpha \in [\alpha^*, 1]} \frac{\alpha k \log \frac{p}{k}}{f(\alpha)} (1 + \eta)$$

- Necessary for $P_{\mathrm{e}} \not\to 1$:

$$n \geq \max_{\alpha \in [\alpha^*, 1]} \frac{(\alpha - \alpha^*) k \log \frac{p}{k}}{f(\alpha)} (1 - \eta)$$

Partial Recovery



Conclusion

- Contributions:
  - New information-theoretic limits for group testing and other support recovery problems
  - Exact thresholds (phase transitions), or near-exact

Conclusion

- Contributions:
  - New information-theoretic limits for group testing and other support recovery problems
  - Exact thresholds (phase transitions), or near-exact

- Implications for group testing:
  - Optimality of i.i.d. Bernoulli measurements in many cases
  - Limited gain by moving to partial recovery

Conclusion

- Contributions:
  - New information-theoretic limits for group testing and other support recovery problems
  - Exact thresholds (phase transitions), or near-exact
- Implications for group testing:
  - Optimality of i.i.d. Bernoulli measurements in many cases
  - Limited gain by moving to partial recovery
- Future work:
  - Closing the remaining gaps (better concentration inequalities?)
  - Non-i.i.d. measurements
  - Applications to other non-linear models
  - "Information-spectrum" type analysis of other statistical problems
  - . . .

Conclusion

- Contributions:
  - New information-theoretic limits for group testing and other support recovery problems
  - Exact thresholds (phase transitions), or near-exact
- Implications for group testing:
  - Optimality of i.i.d. Bernoulli measurements in many cases
  - Limited gain by moving to partial recovery
- Future work:
  - Closing the remaining gaps (better concentration inequalities?)
  - Non-i.i.d. measurements
  - Applications to other non-linear models
  - "Information-spectrum" type analysis of other statistical problems
  - . . .

Further Details

- Further details (group testing):

  http://infoscience.epfl.ch/record/206886

  (accepted to 2016 SODA conference)

- Further details (general models):

  http://arxiv.org/abs/1501.07440

  (submitted to IEEE Transactions on Information Theory)

References I

M. Wainwright, "Information-theoretic limits on sparsity recovery in the high-dimensional and noisy setting," *IEEE Trans. Inf. Theory*, vol. 55, no. 12, pp. 5728–5741, Dec. 2009.

G. Atia and V. Saligrama, "Boolean compressed sensing and noisy group testing," *IEEE Trans. Inf. Theory*, vol. 58, no. 3, pp. 1880–1901, March 2012.

C. Aksoylar, G. Atia, and V. Saligrama, "Sparse signal processing with linear and non-linear observations: A unified Shannon theoretic approach," April 2013, http://arxiv.org/abs/1304.0682.

T. S. Han, *Information-Spectrum Methods in Information Theory*. Springer, 2003.