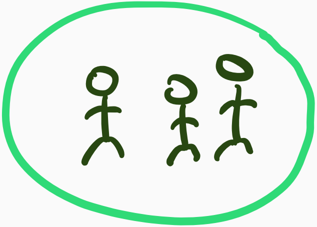
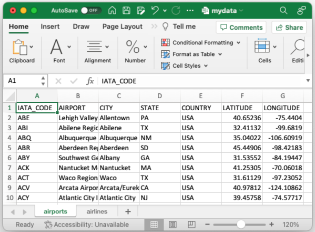
Statistical-Computational Tradeoffs in Mixed Sparse Linear Regression

Gabriel Arpino, Ramji Venkataramanan

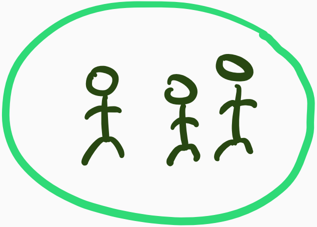
October 3rd, 2023

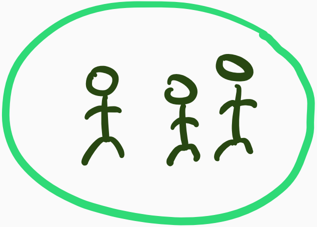
UNIVERSITY OF CAMBRIDGE



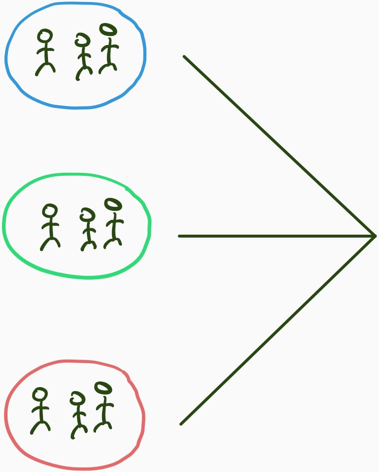


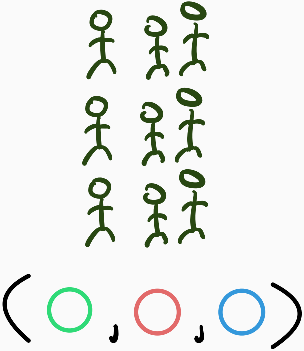
IATA_CODE	AIRPORT	CITY	STATE	COUNTRY	LATITUDE	LONGITUDE
ABE	Lehigh Valley	Allentown	PA	USA	40.65236	-75.4404
ABI	Abilene Regic	Abilene	TX	USA	32.41132	-99.6819
ABQ	Albuquerque	Albuquerque	NM	USA	35.04022	-106.60919
ABR	Aberdeen Re	Aberdeen	SD	USA	45.44906	-98.42183
ABY	Southwest Ge	Albany	GA	USA	31.53552	-84.19447
ACK	Nantucket M	Nantucket	MA	USA	41.25305	-70.06018
ACT	Waco Region	Waco	TX	USA	31.61129	-97.23052
ACV	Arcata Airpor	Arcata/Eurek	CA	USA	40.97812	-124.10862
ACY	Atlantic City I	Atlantic City	NJ	USA	39.45758	-74.57717



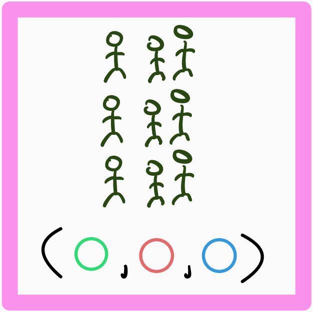


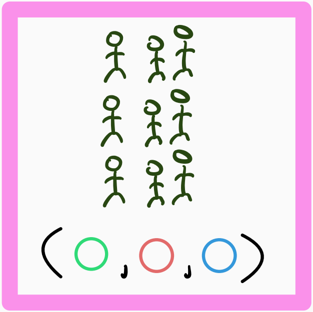
High-Dimensionality



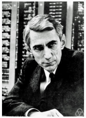
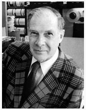


Heterogeneity

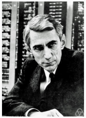
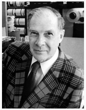


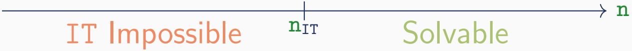


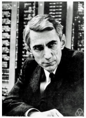
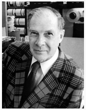
Q: minimal information and computation to recover?

n

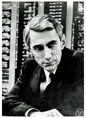
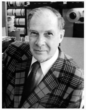

n
IT Impossible
n_{IT}
Solvable

$\texttt{n}$

IT Impossible $\texttt{n}_{\texttt{IT}}$ Solvable

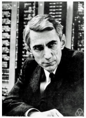
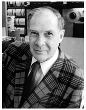
Ex: Communications [S98], Graphical Models [SW09], Regression [MM18], ...

IT Impossible n_{IT} Solvable n

Ex: Communications [S98], Graphical Models [SW09], Regression [MM18], . . .

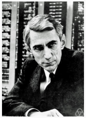
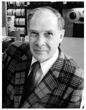
IT Impossible n_{IT} n_{alg} poly-time n

IT Impossible n_{IT} Solvable n

Ex: Communications [S98], Graphical Models [SW09], Regression [MM18], . . .

IT Impossible n_{IT} Computational Barrier n_{alg} poly-time n

IT Impossible n_{IT} Solvable n

Ex: Communications [S98], Graphical Models [SW09], Regression [MM18], . . .

IT Impossible n_{IT} Computational Barrier n_{alg} poly-time n

Ex: Sparse PCA [BR13], Group Testing [CGHWZ22], Sparse Linear Regression [GZ22], . . .

Computational Barriers

Failure of efficient families: MCMC, AMP, Stat Query [BMNW22, GJ19, BBHLS21, ...]

Computational Barriers

Failure of efficient families: MCMC, AMP, Stat Query [BMNW22, GJ19, BBHLS21, ...]

Average Case Reductions [BB20, BR13, ...]

Computational Barriers

Failure of efficient families: MCMC, AMP, Stat Query [BMNW22, GJ19, BBHLS21, . . .]

Average Case Reductions [BB20, BR13, . . .]

Statistical Physics [DKMZ11, GMZ22, . . .]

Computational Barriers

Failure of efficient families: MCMC, AMP, Stat Query [BMNW22, GJ19, BBHLS21, . . .]

Average Case Reductions [BB20, BR13, . . .]

Statistical Physics [DKMZ11, GMZ22, . . .]

Low-Degree Conjecture [H18]:

Computational Barriers

Failure of efficient families: MCMC, AMP, Stat Query [BMNW22, GJ19, BBHLS21, . . .]

Average Case Reductions [BB20, BR13, . . .]

Statistical Physics [DKMZ11, GMZ22, . . .]

Low-Degree Conjecture [H18]:

polynomials of degree $\log(\text{dimension})$ w.r.t input

Computational Barriers

Failure of efficient families: MCMC, AMP, Stat Query [BMNW22, GJ19, BBHLS21, . . .]

Average Case Reductions [BB20, BR13, . . .]

Statistical Physics [DKMZ11, GMZ22, . . .]

Low-Degree Conjecture [H18]:

$$\text{polynomials of degree} \log(\text{dimension}) \text{ w.r.t input} \iff \text{poly-time algorithms}$$

Mixed Sparse Linear Regression (MSLR)

Bad News

Good News

Proof

Mixed Linear Regression

$$y_i = \begin{cases} \langle \boldsymbol{x}_i, \boxed{\boldsymbol{\beta}_1} \rangle + w_i, & \text{with indep. prob. } \phi \\ \langle \boldsymbol{x}_i, \boxed{\boldsymbol{\beta}_2} \rangle + w_i, & \text{with indep. prob. } 1 - \phi \end{cases}$$

Mixed Linear Regression

$$y_i = \begin{cases} \langle \boldsymbol{x}_i, \boxed{\boldsymbol{\beta}_1} \rangle + w_i, & \text{with indep. prob. } \phi \\ \langle \boldsymbol{x}_i, \boxed{\boldsymbol{\beta}_2} \rangle + w_i, & \text{with indep. prob. } 1-\phi \end{cases}$$

◦ $\mathtt{n}$ observations $(y_i, \boldsymbol{x}_i)_{i\in[\mathtt{n}]}$

Mixed Linear Regression

$$y_i = \begin{cases} \langle \boldsymbol{x}_i, \boldsymbol{\beta}_1 \rangle + w_i, & \text{with indep. prob. } \phi \\ \langle \boldsymbol{x}_i, \boldsymbol{\beta}_2 \rangle + w_i, & \text{with indep. prob. } 1-\phi \end{cases}$$

- $\mathtt{n}$ observations $(y_i, \boldsymbol{x}_i)_{i \in [\mathtt{n}]}$
- $\boldsymbol{x}_i \in \mathbb{R}^p$ covariates

Mixed Linear Regression

$$y_i = \begin{cases} \langle \boldsymbol{x}_i, \boxed{\boldsymbol{\beta}_1} \rangle + w_i, & \text{with indep. prob. } \phi \\ \langle \boldsymbol{x}_i, \boxed{\boldsymbol{\beta}_2} \rangle + w_i, & \text{with indep. prob. } 1-\phi \end{cases}$$

- $\mathtt{n}$ observations $(y_i, \boldsymbol{x}_i)_{i\in[\mathtt{n}]}$
- $\boldsymbol{x}_i \in \mathbb{R}^p$ covariates
- $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2 \in \mathbb{R}^p$

Mixed Linear Regression

$$y_i = \begin{cases} \langle \boldsymbol{x}_i, \boldsymbol{\beta}_1 \rangle + w_i, & \text{with indep. prob. } \phi \\ \langle \boldsymbol{x}_i, \boldsymbol{\beta}_2 \rangle + w_i, & \text{with indep. prob. } 1 - \phi \end{cases}$$

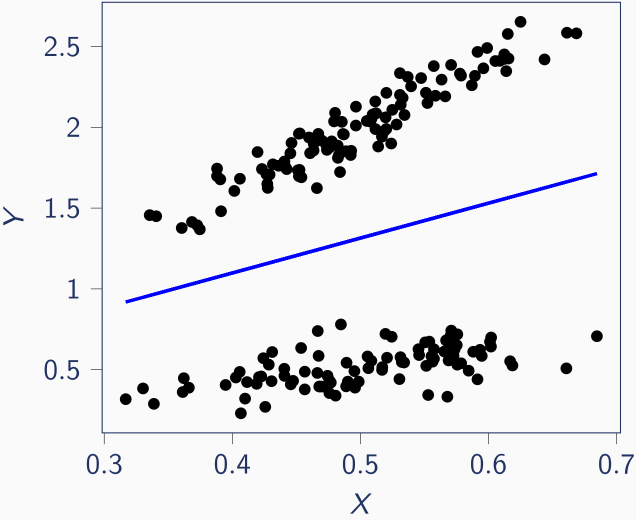
- $\mathtt{n}$ observations $(y_i, \boldsymbol{x}_i)_{i \in [\mathtt{n}]}$
- $\boldsymbol{x}_i \in \mathbb{R}^p$ covariates
- $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2 \in \mathbb{R}^p$
- $w_i \in \mathbb{R}$ noise

Mixed Linear Regression

$$y_i = \begin{cases} \langle \boldsymbol{x}_i, \boxed{\boldsymbol{\beta}_1} \rangle + w_i, & \text{with indep. prob. } \phi \\ \langle \boldsymbol{x}_i, \boxed{\boldsymbol{\beta}_2} \rangle + w_i, & \text{with indep. prob. } 1-\phi \end{cases}$$

- $\mathtt{n}$ observations $(y_i, \boldsymbol{x}_i)_{i \in [\mathtt{n}]}$
- $\boldsymbol{x}_i \in \mathbb{R}^p$ covariates
- $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2 \in \mathbb{R}^p$
- $w_i \in \mathbb{R}$ noise

Goal: Recover $\boxed{\boldsymbol{\beta}_1}$, $\boxed{\boldsymbol{\beta}_2}$ w.h.p as $p \to \infty$, given *only* $(y_i, \boldsymbol{x}_i)_{i \in [\mathtt{n}]}$



Applications: *Health Care* [IHHM'22], *Music Perception* [VT'02], *Market Segmentation* [WK'00], ...

Model Assumptions

$$y_i = \begin{cases} \langle \boldsymbol{x}_i, \boldsymbol{\beta}_1 \rangle + w_i, & \text{with indep. prob. } \phi \\ \langle \boldsymbol{x}_i, \boldsymbol{\beta}_2 \rangle + w_i, & \text{with indep. prob. } 1 - \phi \end{cases}$$

Model Assumptions

$$y_i = \begin{cases} \langle \boldsymbol{x}_i, \boxed{\boldsymbol{\beta}_1} \rangle + w_i, & \text{with indep. prob. } \phi \\ \langle \boldsymbol{x}_i, \boxed{\boldsymbol{\beta}_2} \rangle + w_i, & \text{with indep. prob. } 1 - \phi \end{cases}$$

- $\texttt{n} << p$

Model Assumptions

$$y_i = \begin{cases} \langle \boldsymbol{x}_i, \boxed{\boldsymbol{\beta}_1} \rangle + w_i, & \text{with indep. prob. } \phi \\ \langle \boldsymbol{x}_i, \boxed{\boldsymbol{\beta}_2} \rangle + w_i, & \text{with indep. prob. } 1 - \phi \end{cases}$$

- $\mathtt{n} << p$
- $\boldsymbol{x}_i \overset{i.i.d}{\sim} \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_p)$

Model Assumptions

$$y_i = \begin{cases} \langle \boldsymbol{x}_i, \boxed{\boldsymbol{\beta}_1} \rangle + w_i, & \text{with indep. prob. } \phi \\ \langle \boldsymbol{x}_i, \boxed{\boldsymbol{\beta}_2} \rangle + w_i, & \text{with indep. prob. } 1 - \phi \end{cases}$$

- $\mathtt{n} << p$
- $\boldsymbol{x}_i \overset{i.i.d}{\sim} \mathcal{N}(\mathbf{0}, \boldsymbol{I}_p)$
- $w_i \overset{i.i.d}{\sim} \mathcal{N}(0, \sigma^2)$

Model Assumptions

$$y_i = \begin{cases} \langle \boldsymbol{x}_i, \boldsymbol{\beta}_1 \rangle + w_i, & \text{with indep. prob. } \phi \\ \langle \boldsymbol{x}_i, \boldsymbol{\beta}_2 \rangle + w_i, & \text{with indep. prob. } 1 - \phi \end{cases}$$

- $\mathtt{n} << p$
- $\boldsymbol{x}_i \overset{i.i.d}{\sim} \mathcal{N}(\mathbf{0}, \boldsymbol{I}_p)$
- $w_i \overset{i.i.d}{\sim} \mathcal{N}(0, \sigma^2)$
- $\|\boldsymbol{\beta}_1\|_0 = \|\boldsymbol{\beta}_2\|_0 = k << p$

Model Assumptions

$$y_i = \begin{cases} \langle \boldsymbol{x}_i, \boldsymbol{\beta}_1 \rangle + w_i, & \text{with indep. prob. } \phi \\ \langle \boldsymbol{x}_i, \boldsymbol{\beta}_2 \rangle + w_i, & \text{with indep. prob. } 1 - \phi \end{cases}$$

- $\mathtt{n} << p$
- $\boldsymbol{x}_i \overset{i.i.d}{\sim} \mathcal{N}(\mathbf{0}, \boldsymbol{I}_p)$
- $w_i \overset{i.i.d}{\sim} \mathcal{N}(0, \sigma^2)$
- $\|\boldsymbol{\beta}_1\|_0 = \|\boldsymbol{\beta}_2\|_0 = k << p$
- $\|\boldsymbol{\beta}_1\|_2^2 = \|\boldsymbol{\beta}_2\|_2^2 = \Theta(k)$

Model Assumptions

$$y_i = \begin{cases} \langle \boldsymbol{x}_i, \boxed{\boldsymbol{\beta}_1} \rangle + w_i, & \text{with indep. prob. } \phi \\ \langle \boldsymbol{x}_i, \boxed{\boldsymbol{\beta}_2} \rangle + w_i, & \text{with indep. prob. } 1 - \phi \end{cases}$$

- $\mathtt{n} << p$
- $\boldsymbol{x}_i \overset{i.i.d}{\sim} \mathcal{N}(\mathbf{0}, \boldsymbol{I}_p)$
- $w_i \overset{i.i.d}{\sim} \mathcal{N}(0, \sigma^2)$
- $\|\boldsymbol{\beta}_1\|_0 = \|\boldsymbol{\beta}_2\|_0 = k << p$
- $\|\boldsymbol{\beta}_1\|_2^2 = \|\boldsymbol{\beta}_2\|_2^2 = \Theta(k)$
- $\mathtt{SNR} = \frac{\|\boldsymbol{\beta}_1\|_2^2}{\sigma^2}$

Literature

$$y_i = \begin{cases} \langle \boldsymbol{x}_i, \boldsymbol{\beta}_1 \rangle + w_i, & \text{with indep. prob. } \phi \\ \langle \boldsymbol{x}_i, \boldsymbol{\beta}_2 \rangle + w_i, & \text{with indep. prob. } 1 - \phi \end{cases}$$

Literature

$$y_i = \begin{cases} \langle \boldsymbol{x}_i, \boldsymbol{\beta}_1 \rangle + w_i, & \text{with indep. prob. } \phi \\ \langle \boldsymbol{x}_i, \boldsymbol{\beta}_2 \rangle + w_i, & \text{with indep. prob. } 1 - \phi \end{cases}$$

Max. likelihood is non-convex, NP-Hard [YCS'14]

Literature

$$y_i = \begin{cases} \langle \boldsymbol{x}_i, \boldsymbol{\beta}_1 \rangle + w_i, & \text{with indep. prob. } \phi \\ \langle \boldsymbol{x}_i, \boldsymbol{\beta}_2 \rangle + w_i, & \text{with indep. prob. } 1 - \phi \end{cases}$$

Max. likelihood is non-convex, NP-Hard [YCS'14]

AM (noiseless), AMP (proportional regime), ... [YCS'14], [TV'23]

Literature

$$y_i = \begin{cases} \langle \boldsymbol{x}_i, \boldsymbol{\beta}_1 \rangle + w_i, & \text{with indep. prob. } \phi \\ \langle \boldsymbol{x}_i, \boldsymbol{\beta}_2 \rangle + w_i, & \text{with indep. prob. } 1-\phi \end{cases}$$

Max. likelihood is non-convex, NP-Hard [YCS'14]

AM (noiseless), AMP (proportional regime), ... [YCS'14], [TV'23]

Best: ℓ_1-penalization + EM needs $\mathtt{n} = \tilde{\Omega}(k^2)$ [SBG'10], [MP'22]

Literature

$$y_i = \begin{cases} \langle \boldsymbol{x}_i, \boldsymbol{\beta}_1 \rangle + w_i, & \text{with indep. prob. } \phi \\ \langle \boldsymbol{x}_i, \boldsymbol{\beta}_2 \rangle + w_i, & \text{with indep. prob. } 1 - \phi \end{cases}$$

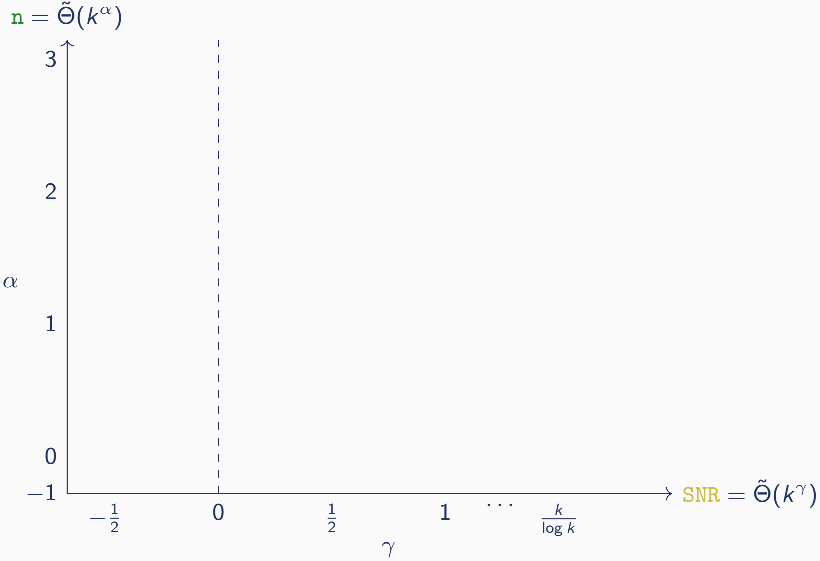
Max. likelihood is non-convex, NP-Hard [YCS'14]

AM (noiseless), AMP (proportional regime), . . . [YCS'14], [TV'23]

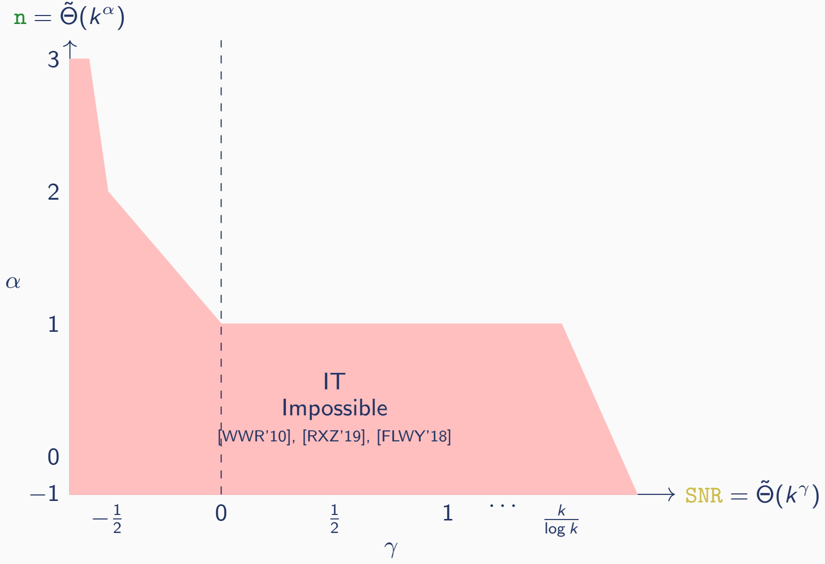
Best: ℓ_1-penalization + EM needs $\mathtt{n} = \tilde{\Omega}(k^2)$ [SBG'10], [MP'22]

Linear Regression needs just $\mathtt{n} = \tilde{\Omega}(k)$ [RXZ'19], [GZ'22]

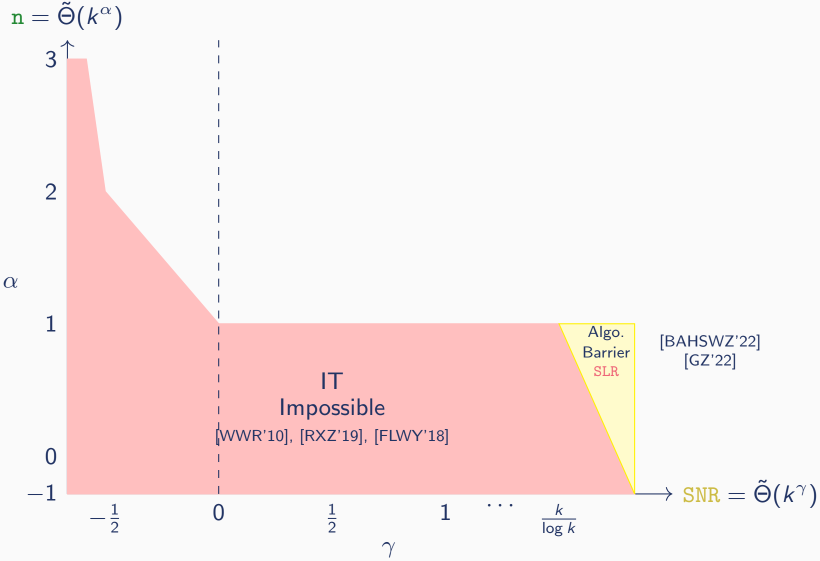
Mixed Sparse Linear Regression



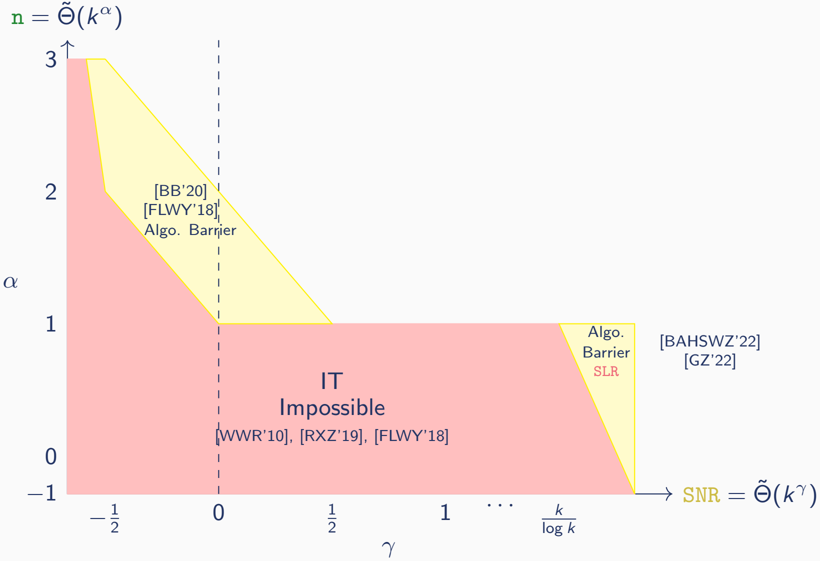
Mixed Sparse Linear Regression



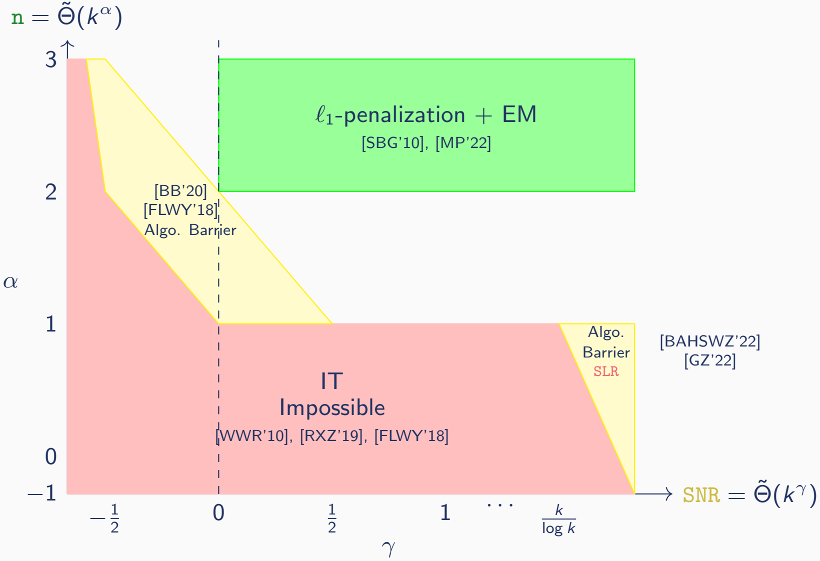
Mixed Sparse Linear Regression



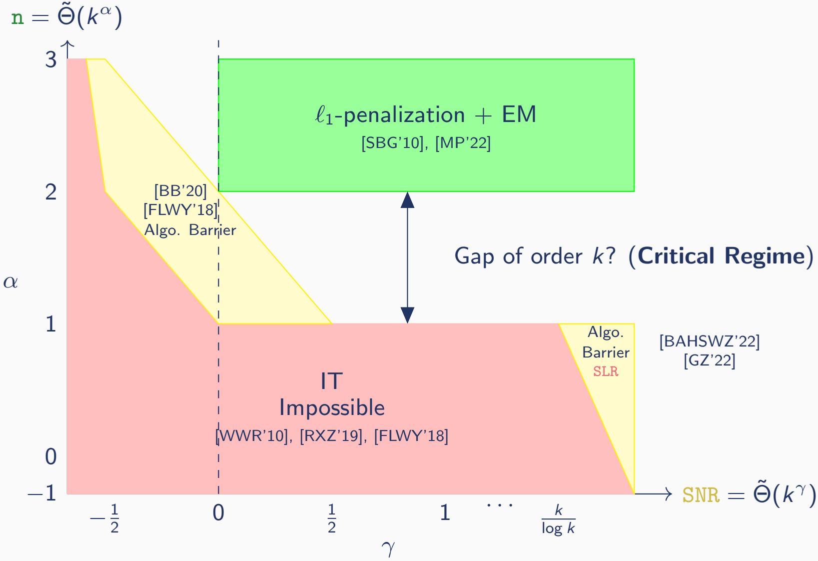
Mixed Sparse Linear Regression



Mixed Sparse Linear Regression



Mixed Sparse Linear Regression



Extreme Cases

$$y_i = \begin{cases} \langle \boldsymbol{x}_i, \boldsymbol{\beta}_1 \rangle + w_i, & \text{with prob. } \phi \\ \langle \boldsymbol{x}_i, \boldsymbol{\beta}_2 \rangle + w_i, & \text{with prob. } 1 - \phi \end{cases}$$

Extreme Cases

$$y_i = \begin{cases} \langle \boldsymbol{x}_i, \boldsymbol{\beta}_1 \rangle + w_i, & \text{with prob. } \phi \\ \langle \boldsymbol{x}_i, \boldsymbol{\beta}_2 \rangle + w_i, & \text{with prob. } 1 - \phi \end{cases}$$

- $\phi = 1$: **Linear Regression**

$$y_i = \langle \boldsymbol{x}_i, \boldsymbol{\beta} \rangle + w_i$$

Extreme Cases

$$y_i = \begin{cases} \langle \boldsymbol{x}_i, \boldsymbol{\beta}_1 \rangle + w_i, & \text{with prob. } \phi \\ \langle \boldsymbol{x}_i, \boldsymbol{\beta}_2 \rangle + w_i, & \text{with prob. } 1-\phi \end{cases}$$

- $\phi = 1$: **Linear Regression**

$$y_i = \langle \boldsymbol{x}_i, \boldsymbol{\beta} \rangle + w_i$$

- $\boldsymbol{\beta}_1 = -\boldsymbol{\beta}_2$, $\phi = 1/2$: **Symmetric Mixed Linear Regression**

$$y_i = \begin{cases} \langle \boldsymbol{x}_i, \boldsymbol{\beta} \rangle + w_i, & \text{with prob. } 1/2 \\ -\langle \boldsymbol{x}_i, \boldsymbol{\beta} \rangle + w_i, & \text{with prob. } 1/2 \end{cases}$$

Main Results: Bad News

Theorem. For $k = o(\sqrt{p})$ and $k \ll \mathtt{n} \ll k^2$, analytic polynomials of degree less than $\frac{k^2}{\mathtt{n}}$ cannot solve a hypothesis testing variant of Symmetric Mixed Linear Regression.

Main Results: Bad News

Theorem. For $k = o(\sqrt{p})$ and $k \ll \mathtt{n} \ll k^2$, analytic polynomials of degree less than $\frac{k^2}{\mathtt{n}}$ cannot solve a hypothesis testing variant of Symmetric Mixed Linear Regression.

$$\Downarrow$$

Low-Degree Conjecture [H18] $+$ Reduction to Recovery

$$\Downarrow$$

Main Results: Bad News

Theorem. For $k = o(\sqrt{p})$ and $k \ll \mathtt{n} \ll k^2$, analytic polynomials of degree less than $\frac{k^2}{\mathtt{n}}$ cannot solve a hypothesis testing variant of Symmetric Mixed Linear Regression.

$$\Downarrow$$

$$\text{Low-Degree Conjecture [H18]} + \text{Reduction to Recovery}$$

$$\Downarrow$$

Corollary. For $k = o(\sqrt{p})$ and $k \ll \mathtt{n} \ll k^2$, any randomized algorithm requires runtime $\exp\left(\tilde{\Omega}\left(\frac{k^2}{\mathtt{n}}\right)\right)$ to exactly recover β_1, β_2 in Symmetric Mixed Linear Regression.

Main Results: Bad News

Theorem. For $k = o(\sqrt{p})$ and $k \ll \mathtt{n} \ll k^2$, analytic polynomials of degree less than $\frac{k^2}{\mathtt{n}}$ cannot solve a hypothesis testing variant of Symmetric Mixed Linear Regression.

$$\Downarrow$$

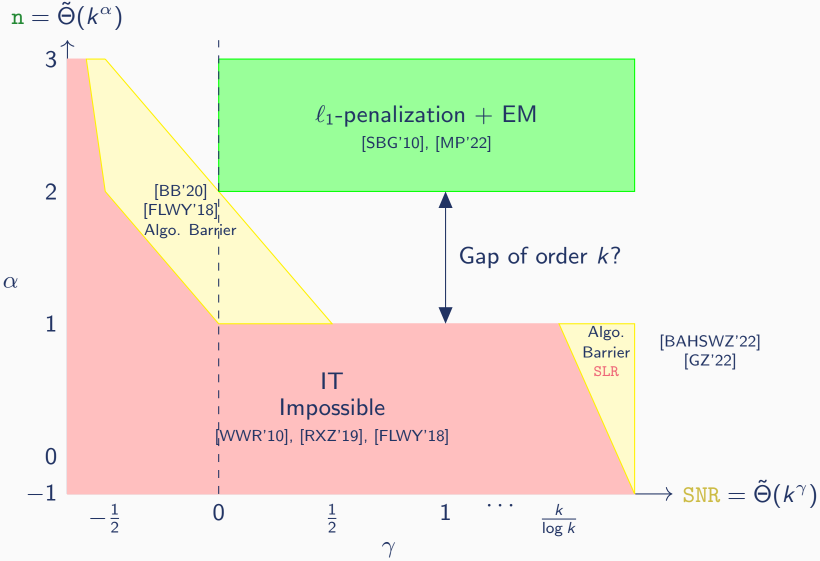
$$\text{Low-Degree Conjecture [H18]} + \text{Reduction to Recovery}$$

$$\Downarrow$$

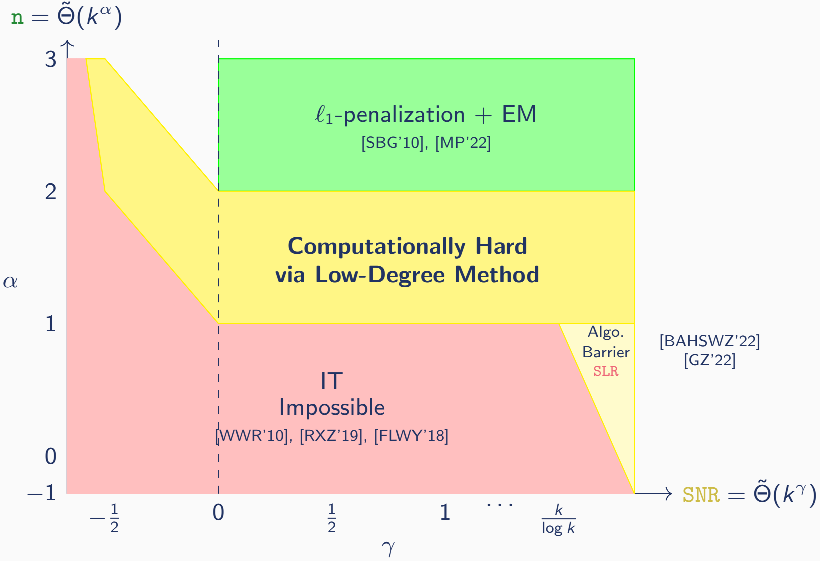
Corollary. For $k = o(\sqrt{p})$ and $k \ll \mathtt{n} \ll k^2$, any randomized algorithm requires runtime $\exp\left(\tilde{\Omega}\left(\frac{k^2}{\mathtt{n}}\right)\right)$ to exactly recover β_1, β_2 in Symmetric Mixed Linear Regression.

This runtime is *super-polynomial* in the input!

Mixed Sparse Linear Regression



Mixed Sparse Linear Regression



Smooth Tradeoff

Corollary. For $k = o(\sqrt{p})$ and $k \ll \mathtt{n} \ll k^2$,

any randomized algorithm requires runtime $\exp\left(\tilde{\Omega}\left(\frac{k^2}{\mathtt{n}}\right)\right)$ to exactly recover $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$ in Mixed Sparse Linear Regression.

Smooth Tradeoff

Corollary. For $k = o(\sqrt{p})$ and $k \ll \mathtt{n} \ll k^2$,

any randomized algorithm requires runtime $\exp\left(\tilde{\Omega}\left(\frac{k^2}{\mathtt{n}}\right)\right)$ to exactly recover $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$ in Mixed Sparse Linear Regression.

$\uparrow$ **Sample size** $\mathtt{n}$ $\overset{\text{Low-Degree Conjecture [H'18]}}{\Longleftrightarrow}$ $\downarrow$ **Runtime** $\exp\left(\frac{k^2}{\mathtt{n}}\right)$

Smooth Tradeoff

Corollary. For $k = o(\sqrt{p})$ and $k \ll \mathtt{n} \ll k^2$,

any randomized algorithm requires runtime $\exp\left(\tilde{\Omega}\left(\frac{k^2}{\mathtt{n}}\right)\right)$ to exactly recover $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$ in Mixed Sparse Linear Regression.

$\uparrow$ **Sample size** $\mathtt{n}$ $\xLeftrightarrow{\text{Low-Degree Conjecture [H'18]}}$ $\downarrow$ **Runtime** $\exp\left(\frac{k^2}{\mathtt{n}}\right)$

Reminiscent of Sparse PCA [DKWB22]:

Smooth Tradeoff

Corollary. For $k = o(\sqrt{p})$ and $k \ll \mathtt{n} \ll k^2$,

any randomized algorithm requires runtime $\exp\left(\tilde{\Omega}\left(\frac{k^2}{\mathtt{n}}\right)\right)$ to exactly recover $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$ in Mixed Sparse Linear Regression.

$\uparrow$ **Sample size** $\mathtt{n}$ $\xLeftrightarrow{\text{Low-Degree Conjecture [H'18]}}$ $\downarrow$ **Runtime** $\exp\left(\frac{k^2}{\mathtt{n}}\right)$

Reminiscent of Sparse PCA [DKWB22]:

Recover k-sparse $\boldsymbol{x}$ from $y_i \overset{i.i.d}{\sim} \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_p + \beta \boldsymbol{x}\boldsymbol{x}^T)$, $\quad \frac{\mathtt{n}}{p} = \Theta(1)$

Smooth Tradeoff

Corollary. For $k = o(\sqrt{p})$ and $k \ll \mathtt{n} \ll k^2$, any randomized algorithm requires runtime $\exp\left(\tilde{\Omega}\left(\frac{k^2}{\mathtt{n}}\right)\right)$ to exactly recover $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$ in Mixed Sparse Linear Regression.

$$\uparrow \textbf{Sample size } \mathtt{n} \quad \xLeftrightarrow{\text{Low-Degree Conjecture [H'18]}} \quad \downarrow \textbf{Runtime } \exp\left(\tfrac{k^2}{\mathtt{n}}\right)$$

Reminiscent of Sparse PCA [DKWB22]:

Recover k-sparse $\boldsymbol{x}$ from $y_i \overset{i.i.d}{\sim} \mathcal{N}(\mathbf{0}, \boldsymbol{I}_p + \beta \boldsymbol{x}\boldsymbol{x}^T)$, $\quad \frac{\mathtt{n}}{p} = \Theta(1)$

$$\uparrow \textbf{Sample size } \mathtt{n} \quad \xLeftrightarrow{\text{Low-Degree Conjecture [H'18]}} \quad \downarrow \textbf{Runtime } \exp\left(\tfrac{k^2}{\mathtt{n}}\right)$$

Implications

Symmetric Mixed Linear Regression

$$y_i = \begin{cases} \langle \boldsymbol{x}_i, \boldsymbol{\beta} \rangle + w_i, & \text{with prob. } 1/2 \\ -\langle \boldsymbol{x}_i, \boldsymbol{\beta} \rangle + w_i, & \text{with prob. } 1/2 \end{cases}$$

Implications

Symmetric Mixed Linear Regression

$$y_i = \begin{cases} \langle \boldsymbol{x}_i, \boldsymbol{\beta} \rangle + w_i, & \text{with prob. } 1/2 \\ -\langle \boldsymbol{x}_i, \boldsymbol{\beta} \rangle + w_i, & \text{with prob. } 1/2 \end{cases}$$

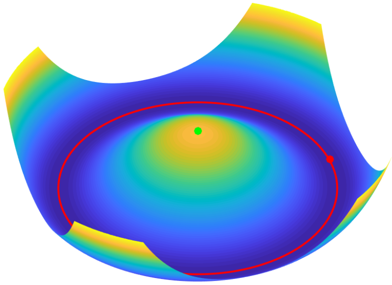
$$\mathit{Symmetry} : \boldsymbol{\beta} \mapsto \pm\boldsymbol{\beta}$$

Implications

Symmetric Mixed Linear Regression

$$y_i = \begin{cases} \langle \boldsymbol{x}_i, \boldsymbol{\beta} \rangle + w_i, & \text{with prob. } 1/2 \\ -\langle \boldsymbol{x}_i, \boldsymbol{\beta} \rangle + w_i, & \text{with prob. } 1/2 \end{cases}$$

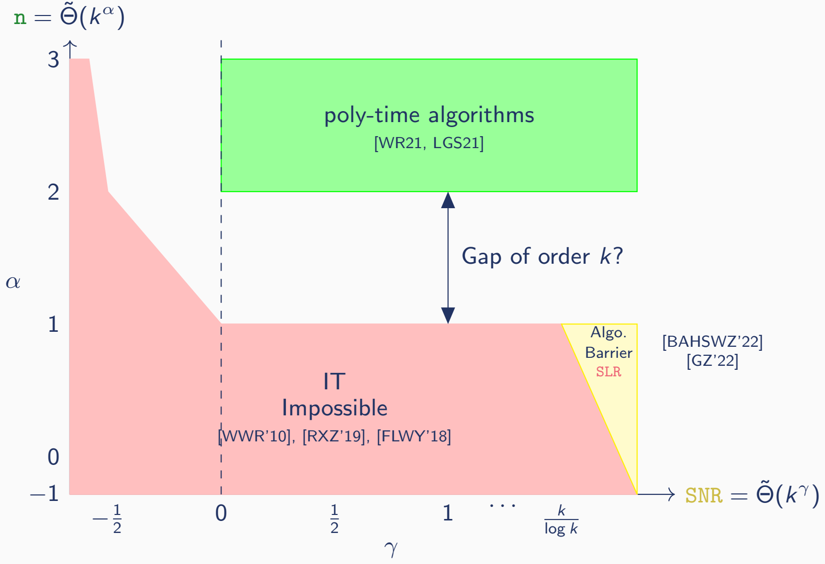
$$\mathit{Symmetry} : \boldsymbol{\beta} \mapsto \pm \boldsymbol{\beta}$$



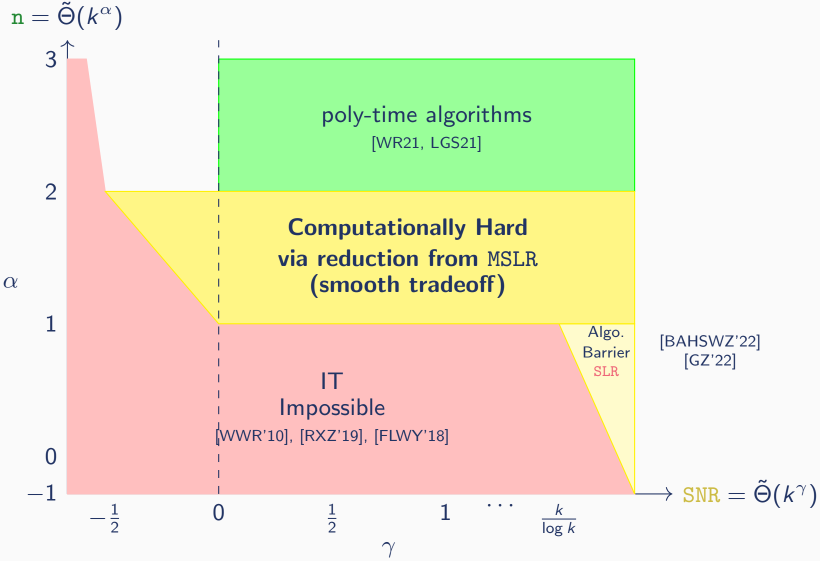
Sparse Phase Retrieval

$$y_i = |\langle \boldsymbol{x}_i, \boldsymbol{\beta} \rangle| + w_i$$

Sparse Phase Retrieval



Sparse Phase Retrieval

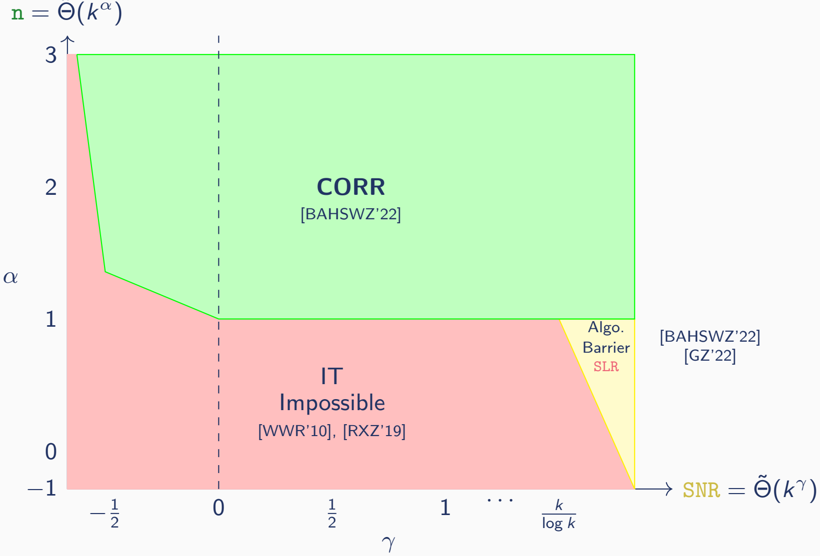


Main Results: Good News

Non-Symmetric Mixed Linear Regression

$$\boldsymbol{\beta}_1 \neq -\boldsymbol{\beta}_2 \text{ or } \phi \neq \frac{1}{2}$$

Non-Symmetric Mixed Linear Regression



Non-Symmetric Mixed Linear Regression

$\boldsymbol{\beta}_1 \neq -\boldsymbol{\beta}_2$ or $\phi \neq \frac{1}{2}$

More Good News

Non-Symmetric Mixed Linear Regression

$\boldsymbol{\beta}_1 \neq -\boldsymbol{\beta}_2$ or $\phi \neq \frac{1}{2}$

Theorem. Provided $\mathtt{n} = \tilde{\Omega}(k)$, the **CORR** algorithm recovers the joint support of $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$ w.h.p.

More Good News

Non-Symmetric Mixed Linear Regression

$\boldsymbol{\beta}_1 \neq -\boldsymbol{\beta}_2$ or $\phi \neq \frac{1}{2}$

Theorem. Provided $\mathtt{n} = \tilde{\Omega}(k)$, the **CORR** algorithm recovers the joint support of $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$ w.h.p.

Translates to full recovery when combined with dense algorithms [YCS14, CYC14]

Proof Idea

The Low-Degree Method [HS17]

Hypothesis Test between two distributions with $o(1)$ error:

- Null model (pure noise) $\boldsymbol{z} \sim \mathbb{Q}_p$
- Planted model (with "signal") $\boldsymbol{z} \sim \mathbb{P}_p$

The Low-Degree Method [HS17]

Hypothesis Test between two distributions with $o(1)$ error:

- Null model (pure noise) $\boldsymbol{z} \sim \mathbb{Q}_p$
- Planted model (with "signal") $\boldsymbol{z} \sim \mathbb{P}_p$

$$L(\boldsymbol{z}) := \frac{d\mathbb{P}}{d\mathbb{Q}}(\boldsymbol{z})$$

The Low-Degree Method [HS17]

Hypothesis Test between two distributions with $o(1)$ error:

- Null model (pure noise) $\boldsymbol{z} \sim \mathbb{Q}_p$
- Planted model (with "signal") $\boldsymbol{z} \sim \mathbb{P}_p$

$$L(\boldsymbol{z}) := \frac{d\mathbb{P}}{d\mathbb{Q}}(\boldsymbol{z}), \quad \langle f, g \rangle := \mathbb{E}_{\boldsymbol{z}\sim\mathbb{Q}} f(\boldsymbol{z}) g(\boldsymbol{z}), \quad \|f\|^2 := \langle f, f \rangle$$

The Low-Degree Method [HS17]

Hypothesis Test between two distributions with $o(1)$ error:

- Null model (pure noise) $\boldsymbol{z} \sim \mathbb{Q}_p$
- Planted model (with "signal") $\boldsymbol{z} \sim \mathbb{P}_p$

$$L(\boldsymbol{z}) := \frac{d\mathbb{P}}{d\mathbb{Q}}(\boldsymbol{z}), \quad \langle f, g \rangle := \mathbb{E}_{\boldsymbol{z} \sim \mathbb{Q}} f(\boldsymbol{z}) g(\boldsymbol{z}), \quad \|f\|^2 := \langle f, f \rangle$$

Lemma. [MRZ15] If $\|L\| = O(1)$ as $p \to \infty$, then impossible to distinguish $\mathbb{P}$ from $\mathbb{Q}$ with $o(1)$ error.

The Low-Degree Method [HS17]

Hypothesis Test between two distributions with $o(1)$ error:

- Null model (pure noise) $\boldsymbol{z} \sim \mathbb{Q}_p$
- Planted model (with "signal") $\boldsymbol{z} \sim \mathbb{P}_p$

$$L(\boldsymbol{z}) := \frac{d\mathbb{P}}{d\mathbb{Q}}(\boldsymbol{z}), \quad \langle f, g \rangle := \mathbb{E}_{\boldsymbol{z} \sim \mathbb{Q}} f(\boldsymbol{z}) g(\boldsymbol{z}), \quad \|f\|^2 := \langle f, f \rangle$$

Lemma. [MRZ15] If $\|L\| = O(1)$ as $p \to \infty$, then impossible to distinguish $\mathbb{P}$ from $\mathbb{Q}$ with $o(1)$ error.

Low-Deg. Conjecture. [H18, CGHWZ22] If $\|L^{\leq D}\| = O(1)$ for some $D = \omega(\log p)$, then any algorithm distinguishing $\mathbb{P}$ from $\mathbb{Q}$ with $o(1)$ error requires runtime $\exp(\tilde{\Omega}(D))$.

$$L(z) := \frac{d\mathbb{P}}{d\mathbb{Q}}(z), \quad \langle f, g \rangle := \mathbb{E}_{z \sim \mathbb{Q}} f(z) g(z), \quad \|f\|^2 := \langle f, f \rangle$$

$$L(\boldsymbol{z}) := \frac{d\mathbb{P}}{d\mathbb{Q}}(\boldsymbol{z}), \quad \langle f, g \rangle := \mathbb{E}_{\boldsymbol{z}\sim\mathbb{Q}} f(\boldsymbol{z}) g(\boldsymbol{z}), \quad \|f\|^2 := \langle f, f \rangle$$

Proposition. [KWB19] The unique solution f^* to the optimization problem

$$\max_{f \text{ deg } D} \frac{\mathbb{E}_{\boldsymbol{z}\sim\mathbb{P}} f(\boldsymbol{z})}{\mathbb{E}_{\boldsymbol{z}\sim\mathbb{Q}}[f(\boldsymbol{z})^2]}$$

is the normalized LDLR $f^* = L^{\leq D}/\|L^{\leq D}\|$. The value is $\|L^{\leq D}\|$.

$$L(\boldsymbol{z}) := \frac{d\mathbb{P}}{d\mathbb{Q}}(\boldsymbol{z}), \quad \langle f, g \rangle := \mathbb{E}_{\boldsymbol{z}\sim\mathbb{Q}} f(\boldsymbol{z}) g(\boldsymbol{z}), \quad \|f\|^2 := \langle f, f \rangle$$

Proposition. [KWB19] The unique solution f^* to the optimization problem

$$\max_{f \text{ deg } D} \frac{\mathbb{E}_{\boldsymbol{z}\sim\mathbb{P}} f(\boldsymbol{z})}{\mathbb{E}_{\boldsymbol{z}\sim\mathbb{Q}}[f(\boldsymbol{z})^2]} = \max_{f \text{ deg } D} \frac{\langle L, f \rangle}{\|f\|^2}$$

is the normalized LDLR $f^* = L^{\leq D}/\|L^{\leq D}\|$. The value is $\|L^{\leq D}\|$.

Detection Variant

$$y_i = \begin{cases} \langle \boldsymbol{x}_i, \boldsymbol{\beta}_1 \rangle + w_i, & \text{with indep. prob. } \phi \\ \langle \boldsymbol{x}_i, \boldsymbol{\beta}_2 \rangle + w_i, & \text{with indep. prob. } 1 - \phi \end{cases}$$

Detection Variant

$$y_i = \begin{cases} \langle \boldsymbol{x}_i, \boldsymbol{\beta}_1 \rangle + w_i, & \text{with indep. prob. } \phi \\ \langle \boldsymbol{x}_i, \boldsymbol{\beta}_2 \rangle + w_i, & \text{with indep. prob. } 1-\phi \end{cases}$$

Under $\mathbb{P}$, observe $\boldsymbol{x}_i \overset{i.i.d}{\sim} \mathcal{N}(\mathbf{0}, \boldsymbol{I}_p)$, and y_i as above.

Detection Variant

$$y_i = \begin{cases} \langle \boldsymbol{x}_i, \boldsymbol{\beta}_1 \rangle + w_i, & \text{with indep. prob. } \phi \\ \langle \boldsymbol{x}_i, \boldsymbol{\beta}_2 \rangle + w_i, & \text{with indep. prob. } 1 - \phi \end{cases}$$

Under $\mathbb{P}$, observe $\boldsymbol{x}_i \overset{i.i.d}{\sim} \mathcal{N}(\mathbf{0}, \boldsymbol{I}_p)$, and y_i as above.

Under $\mathbb{Q}$, observe $(y_i, \boldsymbol{x}_i) \overset{i.i.d}{\sim} \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} (\|\boldsymbol{\beta}_1\|_2^2 + \sigma^2) & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I}_p \end{bmatrix}\right)$.

Detection Variant

$$y_i = \begin{cases} \langle \boldsymbol{x}_i, \boldsymbol{\beta}_1 \rangle + w_i, & \text{with indep. prob. } \phi \\ \langle \boldsymbol{x}_i, \boldsymbol{\beta}_2 \rangle + w_i, & \text{with indep. prob. } 1-\phi \end{cases}$$

Under $\mathbb{P}$, observe $\boldsymbol{x}_i \overset{i.i.d}{\sim} \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_p)$, and y_i as above.

Under $\mathbb{Q}$, observe $(y_i, \boldsymbol{x}_i) \overset{i.i.d}{\sim} \mathcal{N}\left(\boldsymbol{0}, \begin{bmatrix} (\|\boldsymbol{\beta}_1\|_2^2 + \sigma^2) & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_p \end{bmatrix}\right).$

Goal: Test whether $(y_i, \boldsymbol{x}_i)_{i \in [\mathtt{n}]}$ came from $\mathbb{P}$ or from $\mathbb{Q}$, with o(1) error as $p \to \infty$.

Proof Notation

$$\boldsymbol{y} = \boldsymbol{X}\beta_1 \odot \boldsymbol{z} + \boldsymbol{X}\beta_2 \odot (1 - \boldsymbol{z}) + \boldsymbol{w}, \quad z_i \overset{i.i.d}{\sim} \text{Bernoulli}(\phi)$$

Proof Notation

$$\boldsymbol{y} = \boldsymbol{X}\beta_1 \odot \boldsymbol{z} + \boldsymbol{X}\beta_2 \odot (1 - \boldsymbol{z}) + \boldsymbol{w}, \quad z_i \overset{i.i.d}{\sim} \text{Bernoulli}(\phi)$$

$$\boldsymbol{X}_j : \text{column}, \quad \boldsymbol{x}_i : \text{row}$$

$$\boldsymbol{\alpha} \in \mathbb{R}^{\mathtt{n} \times (p+1)} : |\boldsymbol{\alpha}| = \sum_{i,j} \alpha_{i,j}, \boldsymbol{\alpha}! = \prod_{i,j} \alpha_{i,j}!$$

Goal: compute $\|L^{\leq D}\|^2$

Goal: compute $\|L^{\leq D}\|^2$

Input: $(\boldsymbol{y}, \boldsymbol{X}) \in \mathbb{R}^{\mathtt{n} \times (p+1)}$

Goal: compute $\|L^{\leq D}\|^2$

Input: $(\boldsymbol{y}, \boldsymbol{X}) \in \mathbb{R}^{\mathtt{n}\times(p+1)}$

orthonormal basis for poly. deg D w.r.t $\mathbb{Q}$

Goal: compute $\|L^{\leq D}\|^2$

Input: $(\boldsymbol{y}, \boldsymbol{X}) \in \mathbb{R}^{\mathtt{n}\times(p+1)}$

orthonormal basis for poly. deg D w.r.t $\mathbb{Q}$

$$\frac{1}{\sqrt{\boldsymbol{\alpha}!}} H_{\boldsymbol{\alpha}}(\boldsymbol{u}) = \frac{1}{\sqrt{\boldsymbol{\alpha}!}} \prod_{i=1}^{\mathtt{n}} \prod_{j=1}^{p+1} H_{\alpha_{i,j}}(u_{i,j}), \quad \boldsymbol{\alpha}, \boldsymbol{u} \in \mathbb{R}^{\mathtt{n}\times(p+1)}$$

$\|L^{\leq D}\|^2$

$$\|L^{\le D}\|^2 = \sum_{0 \le |\boldsymbol{\alpha}| \le D} \frac{1}{\boldsymbol{\alpha}!} \langle L, H_{\boldsymbol{\alpha}} \rangle^2$$

$$\|L^{\leq D}\|^2 = \sum_{0 \leq |\boldsymbol{\alpha}| \leq D} \frac{1}{\boldsymbol{\alpha}!} \langle L, H_{\boldsymbol{\alpha}} \rangle^2$$

$$= \sum_{0 \leq |\boldsymbol{\alpha}| \leq D} \frac{1}{\boldsymbol{\alpha}!} \mathbb{E}_{\mathbb{Q}}[L(\boldsymbol{y}, \boldsymbol{X}) H_{\boldsymbol{\alpha}}(\boldsymbol{X}, \boldsymbol{y})]^2$$

$$
\begin{aligned}
\|L^{\leq D}\|^2 &= \sum_{0 \leq |\boldsymbol{\alpha}| \leq D} \frac{1}{\boldsymbol{\alpha}!} \langle L, H_{\boldsymbol{\alpha}} \rangle^2 \\
&= \sum_{0 \leq |\boldsymbol{\alpha}| \leq D} \frac{1}{\boldsymbol{\alpha}!} \mathbb{E}_{\mathbb{Q}}[L(\boldsymbol{y}, \boldsymbol{X}) H_{\boldsymbol{\alpha}}(\boldsymbol{X}, \boldsymbol{y})]^2 \\
&= \sum_{0 \leq |\boldsymbol{\alpha}| \leq D} \frac{1}{\boldsymbol{\alpha}!} \mathbb{E}_{\mathbb{P}}[H_{\boldsymbol{\alpha}}(\boldsymbol{X}, \boldsymbol{y})]^2
\end{aligned}
$$

$$
\begin{aligned}
\|L^{\leq D}\|^2 &= \sum_{0\leq|\boldsymbol{\alpha}|\leq D} \frac{1}{\boldsymbol{\alpha}!}\langle L, H_{\boldsymbol{\alpha}}\rangle^2 \\
&= \sum_{0\leq|\boldsymbol{\alpha}|\leq D} \frac{1}{\boldsymbol{\alpha}!}\mathbb{E}_{\mathbb{Q}}[L(\boldsymbol{y},\boldsymbol{X})H_{\boldsymbol{\alpha}}(\boldsymbol{X},\boldsymbol{y})]^2 \\
&= \sum_{0\leq|\boldsymbol{\alpha}|\leq D} \frac{1}{\boldsymbol{\alpha}!}\mathbb{E}_{\mathbb{P}}[H_{\boldsymbol{\alpha}}(\boldsymbol{X},\boldsymbol{y})]^2 \\
&= \sum_{0\leq|\boldsymbol{\alpha}|\leq D} \frac{1}{\boldsymbol{\alpha}!}
\end{aligned}
$$

$$
\mathbb{E}_{\mathbb{P}}\left[\prod_{i=1}^{\mathtt{n}} \underbrace{\left(\prod_{j=1}^{p} H_{\alpha_{i,j}}(X_{i,j})\right)}_{\text{GIP}} H_{\alpha_{i,p+1}}\left(\underbrace{\frac{(\boldsymbol{X}\beta_1\odot\boldsymbol{z}+\boldsymbol{X}\beta_2\odot(1-\boldsymbol{z})+\boldsymbol{w})}{\sqrt{\|\beta\|_2^2+\sigma^2}}}_{\mathcal{N}(0,1)}\right)\right]^2
$$

$$
\begin{aligned}
&\|L^{\leq D}\|^2 \\
&= \sum_{0 \leq |\boldsymbol{\alpha}| \leq D} \frac{1}{\boldsymbol{\alpha}!} \\
&\mathbb{E}_{\mathbb{P}}\left[\underbrace{\prod_{i=1}^{\mathrm{n}}\left(\prod_{j=1}^{p} H_{\alpha_{i,j}}(X_{i,j})\right)}_{\text{GIP}} H_{\alpha_{i,p+1}}\left(\underbrace{\frac{(\boldsymbol{X}\beta_1 \odot \boldsymbol{z} + \boldsymbol{X}\beta_2 \odot (1-\boldsymbol{z}) + \boldsymbol{w})}{\sqrt{\|\beta\|_2^2 + \sigma^2}}}_{\mathcal{N}(0,1)}\right)\right]^2
\end{aligned}
$$

$$
\begin{aligned}
&\|L^{\leq D}\|^2 \\
&= \sum_{0 \leq |\boldsymbol{\alpha}| \leq D} \frac{1}{\boldsymbol{\alpha}!} \\
&\mathbb{E}_{\mathbb{P}} \left[\prod_{i=1}^{\mathtt{n}} \underbrace{\left(\prod_{j=1}^{p} H_{\alpha_{i,j}}(X_{i,j}) \right)}_{\text{GIP}} H_{\alpha_{i,p+1}} \left(\underbrace{\frac{(\boldsymbol{X}\boldsymbol{\beta}_1 \odot \boldsymbol{z} + \boldsymbol{X}\boldsymbol{\beta}_2 \odot (1 - \boldsymbol{z}) + \boldsymbol{w})}{\sqrt{\|\boldsymbol{\beta}\|_2^2 + \sigma^2}}}_{\mathcal{N}(0,1)} \right) \right]^2 \\
&\approx \sum_{0 \leq |\boldsymbol{\alpha}| \leq D} \frac{1}{\boldsymbol{\alpha}!} \mathbb{E}_{\mathbb{P}} \left[\prod_{(i,j)=(1,1)}^{(\mathtt{n},p+1)} (\boldsymbol{\beta}_{1,j} z_i + \boldsymbol{\beta}_{2,j}(1 - z_i))^{\boldsymbol{\alpha}_{i,j}} \cdot H_{\tilde{\alpha}_{i,p+1}}(\mathcal{N}(0,1)) \right]^2
\end{aligned}
$$

$$
\begin{aligned}
&\|L^{\le D}\|^2 \\
&= \sum_{0\le|\boldsymbol{\alpha}|\le D} \frac{1}{\boldsymbol{\alpha}!} \\
&\mathbb{E}_{\mathbb{P}}\left[\underbrace{\prod_{i=1}^{\mathtt{n}}\left(\prod_{j=1}^{p} H_{\alpha_{i,j}}(X_{i,j})\right)}_{\text{GIP}} H_{\alpha_{i,p+1}}\left(\underbrace{\frac{(\boldsymbol{X}\boldsymbol{\beta}_1\odot\boldsymbol{z}+\boldsymbol{X}\boldsymbol{\beta}_2\odot(1-\boldsymbol{z})+\boldsymbol{w})}{\sqrt{\|\boldsymbol{\beta}\|_2^2+\sigma^2}}}_{\mathcal{N}(0,1)}\right)\right]^2 \\
&\approx \sum_{0\le|\boldsymbol{\alpha}|\le D} \frac{1}{\boldsymbol{\alpha}!}\mathbb{E}_{\mathbb{P}}\left[\prod_{(i,j)=(1,1)}^{(\mathtt{n},p+1)}(\boldsymbol{\beta}_{1,j}z_i+\boldsymbol{\beta}_{2,j}(1-z_i))^{\boldsymbol{\alpha}_{i,j}}\cdot H_{\tilde{\alpha}_{i,p+1}}(\mathcal{N}(0,1))\right]^2 \\
&\approx \sum_{0\le|\boldsymbol{\alpha}|\le D} \frac{1}{\boldsymbol{\alpha}!}\mathbb{E}_{\mathbb{P}}\left[\prod_{(i,j)=(1,1)}^{(\mathtt{n},p+1)}(\boldsymbol{\beta}_{1,j}z_i+\boldsymbol{\beta}_{2,j}(1-z_i))^{\boldsymbol{\alpha}_{i,j}}\right]^2
\end{aligned}
$$

$$\|L^{\leq D}\|^2$$

$$= \sum_{0\leq|\boldsymbol{\alpha}|\leq D} \frac{1}{\boldsymbol{\alpha}!}$$

$$\mathbb{E}_{\mathbb{P}}\left[\prod_{i=1}^{\mathtt{n}} \underbrace{\left(\prod_{j=1}^{p} H_{\alpha_{i,j}}(X_{i,j})\right)}_{\text{GIP}} H_{\alpha_{i,p+1}}\left(\underbrace{\frac{(\boldsymbol{X}\boldsymbol{\beta}_1 \odot \boldsymbol{z} + \boldsymbol{X}\boldsymbol{\beta}_2 \odot (1-\boldsymbol{z}) + \boldsymbol{w})}{\sqrt{\|\boldsymbol{\beta}\|_2^2 + \sigma^2}}}_{\mathcal{N}(0,1)}\right)\right]^2$$

$$\approx \sum_{0\leq|\boldsymbol{\alpha}|\leq D} \frac{1}{\boldsymbol{\alpha}!}\mathbb{E}_{\mathbb{P}}\left[\prod_{(i,j)=(1,1)}^{(\mathtt{n},p+1)} (\boldsymbol{\beta}_{1,j} z_i + \boldsymbol{\beta}_{2,j}(1-z_i))^{\boldsymbol{\alpha}_{i,j}} \cdot H_{\tilde{\alpha}_{i,p+1}}(\mathcal{N}(0,1))\right]^2$$

$$\approx \sum_{0\leq|\boldsymbol{\alpha}|\leq D} \frac{1}{\boldsymbol{\alpha}!}\mathbb{E}_{\mathbb{P}}\left[\prod_{(i,j)=(1,1)}^{(\mathtt{n},p+1)} (\boldsymbol{\beta}_{1,j} z_i + \boldsymbol{\beta}_{2,j}(1-z_i))^{\boldsymbol{\alpha}_{i,j}}\right]^2$$

$$\overset{\text{symm.}}{\approx}$$

$$
\begin{aligned}
&\|L^{\leq D}\|^2 \\
&= \sum_{0\leq|\boldsymbol{\alpha}|\leq D} \frac{1}{\boldsymbol{\alpha}!} \\
&\mathbb{E}_{\mathbb{P}}\left[\prod_{i=1}^{\mathtt{n}} \underbrace{\left(\prod_{j=1}^{p} H_{\alpha_{i,j}}(X_{i,j})\right)}_{\text{GIP}} H_{\alpha_{i,p+1}}\left(\underbrace{\frac{(\boldsymbol{X}\boldsymbol{\beta}_1 \odot \boldsymbol{z} + \boldsymbol{X}\boldsymbol{\beta}_2 \odot (1-\boldsymbol{z}) + \boldsymbol{w})}{\sqrt{\|\boldsymbol{\beta}\|_2^2 + \sigma^2}}}_{\mathcal{N}(0,1)}\right)\right]^2 \\
&\approx \sum_{0\leq|\boldsymbol{\alpha}|\leq D} \frac{1}{\boldsymbol{\alpha}!} \mathbb{E}_{\mathbb{P}}\left[\prod_{(i,j)=(1,1)}^{(\mathtt{n},p+1)} (\boldsymbol{\beta}_{1,j} z_i + \boldsymbol{\beta}_{2,j}(1-z_i))^{\boldsymbol{\alpha}_{i,j}} \cdot H_{\tilde{\alpha}_{i,p+1}}(\mathcal{N}(0,1))\right]^2 \\
&\approx \sum_{0\leq|\boldsymbol{\alpha}|\leq D} \frac{1}{\boldsymbol{\alpha}!} \mathbb{E}_{\mathbb{P}}\left[\prod_{(i,j)=(1,1)}^{(\mathtt{n},p+1)} (\boldsymbol{\beta}_{1,j} z_i + \boldsymbol{\beta}_{2,j}(1-z_i))^{\boldsymbol{\alpha}_{i,j}}\right]^2 \\
&\overset{\text{symm.}}{\approx} \sum_{d}^{D} \mathbb{E}_{\mathbb{P}}[\langle \boldsymbol{\beta}_1^{(1)}, \boldsymbol{\beta}_1^{(2)} \rangle^{2d}]
\end{aligned}
$$

$$
\begin{aligned}
&\|L^{\leq D}\|^2 \\
&= \sum_{0\leq|\boldsymbol{\alpha}|\leq D} \frac{1}{\boldsymbol{\alpha}!} \\
&\mathbb{E}_{\mathbb{P}}\left[\prod_{i=1}^{\mathtt{n}} \underbrace{\left(\prod_{j=1}^{p} H_{\alpha_{i,j}}(X_{i,j})\right)}_{\text{GIP}} H_{\alpha_{i,p+1}}\left(\underbrace{\frac{(\boldsymbol{X}\boldsymbol{\beta}_1\odot\boldsymbol{z}+\boldsymbol{X}\boldsymbol{\beta}_2\odot(1-\boldsymbol{z})+\boldsymbol{w})}{\sqrt{\|\boldsymbol{\beta}\|_2^2+\sigma^2}}}_{\mathcal{N}(0,1)}\right)\right]^2 \\
&\approx \sum_{0\leq|\boldsymbol{\alpha}|\leq D} \frac{1}{\boldsymbol{\alpha}!}\mathbb{E}_{\mathbb{P}}\left[\prod_{(i,j)=(1,1)}^{(\mathtt{n},p+1)} (\boldsymbol{\beta}_{1,j}z_i+\boldsymbol{\beta}_{2,j}(1-z_i))^{\boldsymbol{\alpha}_{i,j}}\cdot H_{\tilde{\alpha}_{i,p+1}}(\mathcal{N}(0,1))\right]^2 \\
&\approx \sum_{0\leq|\boldsymbol{\alpha}|\leq D} \frac{1}{\boldsymbol{\alpha}!}\mathbb{E}_{\mathbb{P}}\left[\prod_{(i,j)=(1,1)}^{(\mathtt{n},p+1)} (\boldsymbol{\beta}_{1,j}z_i+\boldsymbol{\beta}_{2,j}(1-z_i))^{\boldsymbol{\alpha}_{i,j}}\right]^2 \\
&\overset{\text{symm.}}{\approx} \sum_{d}^{D}\mathbb{E}_{\mathbb{P}}[\langle\boldsymbol{\beta}_1^{(1)},\boldsymbol{\beta}_1^{(2)}\rangle^{2d}] = O(1) \text{ for } D\leq\frac{k^2}{\mathtt{n}}.
\end{aligned}
$$

Proof Idea: Good News

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta}_1 \odot \boldsymbol{z} + \boldsymbol{X}\boldsymbol{\beta}_2 \odot (1 - \boldsymbol{z}), \quad z_i \overset{i.i.d}{\sim} \text{Bernoulli}(\phi)$$

Proof Idea: Good News

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta}_1 \odot \boldsymbol{z} + \boldsymbol{X}\boldsymbol{\beta}_2 \odot (1 - \boldsymbol{z}), \quad z_i \overset{i.i.d}{\sim} \text{Bernoulli}(\phi)$$

$$\mathcal{S}_1 := \{j \in [p] : \boldsymbol{\beta}_{1,j} > 0\}$$

Proof Idea: Good News

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta}_1 \odot \boldsymbol{z} + \boldsymbol{X}\boldsymbol{\beta}_2 \odot (1 - \boldsymbol{z}), \quad z_i \overset{i.i.d}{\sim} \text{Bernoulli}(\phi)$$

$$\mathcal{S}_1 := \{j \in [p] : \boldsymbol{\beta}_{1,j} > 0\}$$

[BAHSWZ22]:

$$\textbf{CORR}(\boldsymbol{X}, \boldsymbol{y}) := \left\{ j \in [p] : \left| \frac{\langle \boldsymbol{X}_j, \boldsymbol{y} \rangle}{\|\boldsymbol{y}\|_2} \right| \geq \sqrt{2(1 + \epsilon/2) \log 2p} \right\}$$

Proof Idea: Good News

$$\boldsymbol{y} = \boldsymbol{X}\beta_1 \odot \boldsymbol{z} + \boldsymbol{X}\beta_2 \odot (1 - \boldsymbol{z}), \quad z_i \overset{i.i.d}{\sim} \text{Bernoulli}(\phi)$$

$$\mathcal{S}_1 := \{j \in [p] : \beta_{1,j} > 0\}$$

[BAHSWZ22]:

$$\textbf{CORR}(\boldsymbol{X}, \boldsymbol{y}) := \left\{ j \in [p] : \left| \frac{\langle \boldsymbol{X}_j, \boldsymbol{y} \rangle}{\|\boldsymbol{y}\|_2} \right| \geq \sqrt{2(1 + \epsilon/2) \log 2p} \right\}$$

$$u_j := \frac{\langle \boldsymbol{X}_j, \boldsymbol{y} \rangle}{\|\boldsymbol{y}\|_2}$$

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta}_1 \odot \boldsymbol{z} + \boldsymbol{X}\boldsymbol{\beta}_2 \odot (1 - \boldsymbol{z}), \quad \boldsymbol{z} \overset{i.i.d}{\sim} \text{Bernoulli}(\phi)$$

$$\textbf{CORR}(\boldsymbol{X}, \boldsymbol{y}) := \left\{ j \in [p] : \left| \frac{\langle \boldsymbol{X}_j, \boldsymbol{y} \rangle}{\|\boldsymbol{y}\|_2} \right| \geq \sqrt{2(1 + \epsilon/2) \log 2p} \right\}$$

$$u_j := \frac{\langle \boldsymbol{X}_j, \boldsymbol{y} \rangle}{\|\boldsymbol{y}\|_2}$$

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta}_1 \odot \boldsymbol{z} + \boldsymbol{X}\boldsymbol{\beta}_2 \odot (1 - \boldsymbol{z}), \quad \boldsymbol{z} \overset{i.i.d}{\sim} \text{Bernoulli}(\phi)$$

$$\textbf{CORR}(\boldsymbol{X}, \boldsymbol{y}) := \left\{ j \in [p] : \left| \frac{\langle \boldsymbol{X}_j, \boldsymbol{y} \rangle}{\|\boldsymbol{y}\|_2} \right| \geq \sqrt{2(1 + \epsilon/2)\log 2p} \right\}$$

$$u_j := \frac{\langle \boldsymbol{X}_j, \boldsymbol{y} \rangle}{\|\boldsymbol{y}\|_2}$$

$$j \in (\mathcal{S}_1 \cup \mathcal{S}_2)^{\complement} : u_j \overset{i.i.d}{\sim} \mathcal{N}(0, 1)$$

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta}_1 \odot \boldsymbol{z} + \boldsymbol{X}\boldsymbol{\beta}_2 \odot (1 - \boldsymbol{z}), \quad \boldsymbol{z} \overset{i.i.d}{\sim} \text{Bernoulli}(\phi)$$

$$\textbf{CORR}(\boldsymbol{X}, \boldsymbol{y}) := \left\{ j \in [p] : \left| \frac{\langle \boldsymbol{X}_j, \boldsymbol{y} \rangle}{\|\boldsymbol{y}\|_2} \right| \geq \sqrt{2(1 + \epsilon/2) \log 2p} \right\}$$

$$u_j := \frac{\langle \boldsymbol{X}_j, \boldsymbol{y} \rangle}{\|\boldsymbol{y}\|_2}$$

$$j \in (\mathcal{S}_1 \cup \mathcal{S}_2)^{\complement} : u_j \overset{i.i.d}{\sim} \mathcal{N}(0, 1)$$

$$\max_{j \in [p]} u_j < \sqrt{2(1 + \epsilon/2) \log 2p}$$

$$\boldsymbol{y} = \boldsymbol{X}\beta_1 \odot \boldsymbol{z} + \boldsymbol{X}\beta_2 \odot (1 - \boldsymbol{z}), \quad \boldsymbol{z} \overset{i.i.d}{\sim} \text{Bernoulli}(\phi)$$

$$\textbf{CORR}(\boldsymbol{X}, \boldsymbol{y}) := \left\{ j \in [p] : \left| \frac{\langle \boldsymbol{X}_j, \boldsymbol{y} \rangle}{\|\boldsymbol{y}\|_2} \right| \geq \sqrt{2(1 + \epsilon/2) \log 2p} \right\}$$

$$u_j := \frac{\langle \boldsymbol{X}_j, \boldsymbol{y} \rangle}{\|\boldsymbol{y}\|_2}$$

$j \in \mathcal{S}_1 \cap \mathcal{S}_2$:

$$\mathbb{E}\left[\boldsymbol{X}_{j,i} | y_i, \beta_1, \boldsymbol{z}_i = 1\right] = \frac{y_i \cdot \boldsymbol{\beta}_{1,j}}{\|\boldsymbol{\beta}_1\|_2^2 + \sigma^2}$$

$$\boldsymbol{y} = \boldsymbol{X}\beta_1 \odot \boldsymbol{z} + \boldsymbol{X}\beta_2 \odot (1 - \boldsymbol{z}), \quad \boldsymbol{z} \overset{i.i.d}{\sim} \text{Bernoulli}(\phi)$$

$$\textbf{CORR}(\boldsymbol{X}, \boldsymbol{y}) := \left\{ j \in [p] : \left| \frac{\langle \boldsymbol{X}_j, \boldsymbol{y} \rangle}{\|\boldsymbol{y}\|_2} \right| \geq \sqrt{2(1 + \epsilon/2) \log 2p} \right\}$$

$$u_j := \frac{\langle \boldsymbol{X}_j, \boldsymbol{y} \rangle}{\|\boldsymbol{y}\|_2}$$

$j \in \mathcal{S}_1 \cap \mathcal{S}_2$:

$$\mathbb{E}\left[\boldsymbol{X}_{j,i} | y_i, \boldsymbol{\beta}_1, \boldsymbol{z}_i = 1\right] = \frac{y_i \cdot \boldsymbol{\beta}_{1,j}}{\|\boldsymbol{\beta}_1\|_2^2 + \sigma^2}$$

$$\mathbb{E}\left[u_j \mid \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2\right] = \frac{\|\boldsymbol{y}_{\{\boldsymbol{z}=1\}}\|_2^2 \boldsymbol{\beta}_{1,j} + \|\boldsymbol{y}_{\{\boldsymbol{z}=0\}}\|_2^2 \boldsymbol{\beta}_{2,j}}{\|\boldsymbol{y}\|_2 (\|\boldsymbol{\beta}\|_2^2 + \sigma^2)}$$

$$\mathbb{E}\left[u_j \mid \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2\right] = \frac{\|\boldsymbol{y}_{\{\boldsymbol{z}=1\}}\|_2^2 \boldsymbol{\beta}_{1,j} + \|\boldsymbol{y}_{\{\boldsymbol{z}=0\}}\|_2^2 \boldsymbol{\beta}_{2,j}}{\|\boldsymbol{y}\|_2 (\|\boldsymbol{\beta}\|_2^2 + \sigma^2)}$$

$$\mathbb{E}\left[u_j \mid \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2\right] = \frac{\|\boldsymbol{y}_{\{\boldsymbol{z}=1\}}\|_2^2 \boldsymbol{\beta}_{1,j} + \|\boldsymbol{y}_{\{\boldsymbol{z}=0\}}\|_2^2 \boldsymbol{\beta}_{2,j}}{\|\boldsymbol{y}\|_2 (\|\boldsymbol{\beta}\|_2^2 + \sigma^2)}$$

$$\|\boldsymbol{y}\|_2 \approx \sqrt{\mathtt{n}(\|\boldsymbol{\beta}\|_2^2 + \sigma^2)}$$

$$\mathbb{E}\left[u_j \mid \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2\right] = \frac{\|\boldsymbol{y}_{\{z=1\}}\|_2^2 \boldsymbol{\beta}_{1,j} + \|\boldsymbol{y}_{\{z=0\}}\|_2^2 \boldsymbol{\beta}_{2,j}}{\|\boldsymbol{y}\|_2 (\|\boldsymbol{\beta}\|_2^2 + \sigma^2)}$$

$$\|\boldsymbol{y}\|_2 \approx \sqrt{\mathtt{n}(\|\boldsymbol{\beta}\|_2^2 + \sigma^2)}$$

$$u_j \approx \frac{\phi \mathtt{n}(\|\boldsymbol{\beta}\|_2^2 + \sigma^2)\boldsymbol{\beta}_{1,j} + (1-\phi)\mathtt{n}(\|\boldsymbol{\beta}\|_2^2 + \sigma^2)\boldsymbol{\beta}_{2,j}}{\sqrt{\mathtt{n}(\|\boldsymbol{\beta}\|_2^2 + \sigma^2)}(\|\boldsymbol{\beta}\|_2^2 + \sigma^2)}$$

$$\mathbb{E}\left[u_j \mid \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2\right] = \frac{\|\boldsymbol{y}_{\{z=1\}}\|_2^2 \boldsymbol{\beta}_{1,j} + \|\boldsymbol{y}_{\{z=0\}}\|_2^2 \boldsymbol{\beta}_{2,j}}{\|\boldsymbol{y}\|_2 (\|\boldsymbol{\beta}\|_2^2 + \sigma^2)}$$

$$\|\boldsymbol{y}\|_2 \approx \sqrt{\mathtt{n}(\|\boldsymbol{\beta}\|_2^2 + \sigma^2)}$$

$$\begin{aligned} u_j &\approx \frac{\phi \mathtt{n}(\|\boldsymbol{\beta}\|_2^2 + \sigma^2)\boldsymbol{\beta}_{1,j} + (1-\phi)\mathtt{n}(\|\boldsymbol{\beta}\|_2^2 + \sigma^2)\boldsymbol{\beta}_{2,j}}{\sqrt{\mathtt{n}(\|\boldsymbol{\beta}\|_2^2 + \sigma^2)}(\|\boldsymbol{\beta}\|_2^2 + \sigma^2)} \\ &= \sqrt{\frac{\mathtt{n}}{\|\boldsymbol{\beta}\|_2^2 + \sigma^2}}(\phi \boldsymbol{\beta}_{1,j} + (1-\phi)\boldsymbol{\beta}_{2,j}) \end{aligned}$$

$$\mathbb{E}\left[u_j \mid \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2\right] = \frac{\|\boldsymbol{y}_{\{z=1\}}\|_2^2 \boldsymbol{\beta}_{1,j} + \|\boldsymbol{y}_{\{z=0\}}\|_2^2 \boldsymbol{\beta}_{2,j}}{\|\boldsymbol{y}\|_2(\|\boldsymbol{\beta}\|_2^2 + \sigma^2)}$$

$$\|\boldsymbol{y}\|_2 \approx \sqrt{\mathtt{n}(\|\boldsymbol{\beta}\|_2^2 + \sigma^2)}$$

$$\begin{aligned}
u_j &\approx \frac{\phi \mathtt{n}(\|\boldsymbol{\beta}\|_2^2 + \sigma^2)\boldsymbol{\beta}_{1,j} + (1-\phi)\mathtt{n}(\|\boldsymbol{\beta}\|_2^2 + \sigma^2)\boldsymbol{\beta}_{2,j}}{\sqrt{\mathtt{n}(\|\boldsymbol{\beta}\|_2^2 + \sigma^2)}(\|\boldsymbol{\beta}\|_2^2 + \sigma^2)} \\
&= \sqrt{\frac{\mathtt{n}}{\|\boldsymbol{\beta}\|_2^2 + \sigma^2}}(\phi \boldsymbol{\beta}_{1,j} + (1-\phi)\boldsymbol{\beta}_{2,j}) \\
&\overset{\mathtt{n}=\tilde{\Omega}(k)}{>} \sqrt{2(1+\epsilon/2)\log 2p}.
\end{aligned}$$

Summary

- Narrow symmetric case: $\mathtt{n} = \tilde{\Omega}(k^2)$
- Broad non-symmetric case: $\mathtt{n} = \tilde{\Omega}(k)$

Future

Future

- Sub-exp algorithms for `MSLR`, Sparse Phase Retrieval

Future

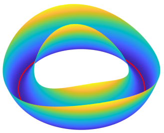
- Sub-exp algorithms for `MSLR`, Sparse Phase Retrieval
- Adversarial instances of "robust" inference are mixtures, ex: Robust Mean Est., Robust Lin Reg [DK23, BB20]

Future

- Sub-exp algorithms for `MSLR`, Sparse Phase Retrieval
- Adversarial instances of “robust” inference are mixtures, ex: Robust Mean Est., Robust Lin Reg [DK23, BB20]

Eigenspace Computation

Compute the principal subspace of a symmetric matrix.



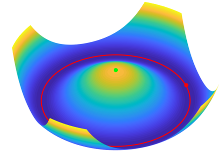
$\min_{\boldsymbol{X}^*\boldsymbol{X}=\boldsymbol{I}} -\frac{1}{2}\operatorname{trace}[\boldsymbol{X}^*\boldsymbol{A}\boldsymbol{X}].$

Symmetry: $\boldsymbol{X} \mapsto \boldsymbol{X}\boldsymbol{R}$

$\mathbb{G} = \mathsf{O}(r)$

Generalized Phase Retrieval

Recover a complex vector $\boldsymbol{x}_o$ from magnitude measurements $\boldsymbol{y} = |\boldsymbol{A}\boldsymbol{x}_o|$.



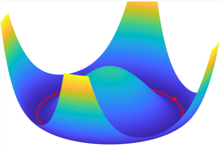
$\min_{\boldsymbol{x}} \frac{1}{2}\|\boldsymbol{y}^2 - |\boldsymbol{A}\boldsymbol{x}|^2\|_2^2.$

Symmetry: $\boldsymbol{x} \mapsto \boldsymbol{x}e^{i\phi}$

$\mathbb{G} = \mathbb{S}^1 \cong \mathsf{O}(2)$

Matrix Recovery

Recover a low-rank matrix $\boldsymbol{X} = \boldsymbol{U}\boldsymbol{V}^*$ from incomplete/corrupted observations



$\min_{\boldsymbol{U},\boldsymbol{V}} \mathcal{L}(\boldsymbol{Y} - \mathcal{A}[\boldsymbol{U}\boldsymbol{V}^*]) + \rho(\boldsymbol{U},\boldsymbol{V}).$

Symmetry: $(\boldsymbol{U},\boldsymbol{V}) \mapsto (\boldsymbol{U}\boldsymbol{\Gamma}, \boldsymbol{V}\boldsymbol{\Gamma}^{-*})$

$\mathbb{G} = \mathsf{GL}(r)$ or $\mathbb{G} = \mathsf{O}(r)$

[WM22]

Appendix: The Low-Degree Method [HS17]

Goal: *Hypothesis Test* between two distributions with $o(1)$ error:

- Null model (pure noise) $\boldsymbol{z} \sim \mathbb{Q}_p$
- Planted model (with "signal") $\boldsymbol{z} \sim \mathbb{P}_p$

Compute

$$\max_{f \deg D} \frac{\mathbb{E}_{\boldsymbol{z}\sim\mathbb{P}}[f(\boldsymbol{z})]}{\sqrt{\mathbb{E}_{\boldsymbol{z}\sim\mathbb{Q}}[f(\boldsymbol{z})^2]}} \qquad \frac{\text{mean in } \mathbb{P}}{\text{fluctuations in } \mathbb{Q}}$$

Appendix: The Low-Degree Method [HS17]

Goal: *Hypothesis Test* between two distributions with $o(1)$ error:

- Null model (pure noise) $\boldsymbol{z} \sim \mathbb{Q}_p$
- Planted model (with "signal") $\boldsymbol{z} \sim \mathbb{P}_p$

Can we find a low-degree polynomial $f(\boldsymbol{z})$ that is big for $\boldsymbol{z} \sim \mathbb{P}$ and small for $\boldsymbol{z} \sim \mathbb{Q}$?

Compute

$$\max_{f \deg D} \frac{\mathbb{E}_{\boldsymbol{z}\sim\mathbb{P}}[f(\boldsymbol{z})]}{\sqrt{\mathbb{E}_{\boldsymbol{z}\sim\mathbb{Q}}[f(\boldsymbol{z})^2]}} \qquad \frac{\text{mean in } \mathbb{P}}{\text{fluctuations in } \mathbb{Q}}$$

Appendix: The Low-Degree Method [HS17]

$$\max_{f \deg D} \frac{\mathbb{E}_{\boldsymbol{z} \sim \mathbb{P}}[f(\boldsymbol{z})]}{\sqrt{\mathbb{E}_{\boldsymbol{z} \sim \mathbb{Q}}[f(\boldsymbol{z})^2]}}$$

Appendix: The Low-Degree Method [HS17]

$$\max_{f \deg D} \frac{\mathbb{E}_{\boldsymbol{z} \sim \mathbb{P}}[f(\boldsymbol{z})]}{\sqrt{\mathbb{E}_{\boldsymbol{z} \sim \mathbb{Q}}[f(\boldsymbol{z})^2]}}$$

$$= \max_{f \deg D} \frac{\mathbb{E}_{\boldsymbol{z} \sim \mathbb{Q}}[L f(\boldsymbol{z})]}{\sqrt{\mathbb{E}_{\boldsymbol{z} \sim \mathbb{Q}}[f(\boldsymbol{z})^2]}}$$

$$L = \frac{d\mathbb{P}}{d\mathbb{Q}}$$

Appendix: The Low-Degree Method [HS17]

$$\max_{f \deg D} \frac{\mathbb{E}_{\boldsymbol{z}\sim\mathbb{P}}[f(\boldsymbol{z})]}{\sqrt{\mathbb{E}_{\boldsymbol{z}\sim\mathbb{Q}}[f(\boldsymbol{z})^2]}}$$

$$= \max_{f \deg D} \frac{\mathbb{E}_{\boldsymbol{z}\sim\mathbb{Q}}[Lf(\boldsymbol{z})]}{\sqrt{\mathbb{E}_{\boldsymbol{z}\sim\mathbb{Q}}[f(\boldsymbol{z})^2]}}$$

$$= \max_{f \deg D} \frac{\langle L, f\rangle}{\|f\|}$$

$$L = \frac{d\mathbb{P}}{d\mathbb{Q}}$$

$$\langle f, g\rangle := \mathbb{E}_{\boldsymbol{z}\sim\mathbb{Q}}[f(\boldsymbol{z})g(\boldsymbol{z})]$$

Appendix: The Low-Degree Method [HS17]

$$\max_{f \deg D} \frac{\mathbb{E}_{\boldsymbol{z} \sim \mathbb{P}}[f(\boldsymbol{z})]}{\sqrt{\mathbb{E}_{\boldsymbol{z} \sim \mathbb{Q}}[f(\boldsymbol{z})^2]}}$$

$$= \max_{f \deg D} \frac{\mathbb{E}_{\boldsymbol{z} \sim \mathbb{Q}}[Lf(\boldsymbol{z})]}{\sqrt{\mathbb{E}_{\boldsymbol{z} \sim \mathbb{Q}}[f(\boldsymbol{z})^2]}} \qquad L = \frac{d\mathbb{P}}{d\mathbb{Q}}$$

$$= \max_{f \deg D} \frac{\langle L, f \rangle}{\|f\|} \qquad \langle f, g \rangle := \mathbb{E}_{\boldsymbol{z} \sim \mathbb{Q}}[f(\boldsymbol{z})g(\boldsymbol{z})]$$

$$= \|L^{\leq D}\|$$

Appendix: The Low-Degree Conjecture

Informal. [H18, CGHWZ22] Assume "sufficiently nice" distributions $\mathbb{P}$, $\mathbb{Q}$. If $\|L^{\leq D}\| = O(1)$ for some $D = \omega(\log p)$, then algorithms require runtime $\exp(\tilde{\Omega}(D))$ to distinguish $\mathbb{P}$ from $\mathbb{Q}$.

Symmetric Mixed Linear Detection

$$y_i = \begin{cases} \langle \boldsymbol{x}_i, \boldsymbol{\beta} \rangle + w_i, & \text{with indep. prob. } 1/2 \\ -\langle \boldsymbol{x}_i, \boldsymbol{\beta} \rangle + w_i, & \text{with indep. prob. } 1/2 \end{cases}$$

Under $\mathbb{P}$, observe $\boldsymbol{x}_i \overset{i.i.d}{\sim} \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_p)$, and y_i as above.

Under $\mathbb{Q}$, observe $(y_i, \boldsymbol{x}_i) \overset{i.i.d}{\sim} \mathcal{N}\left(\boldsymbol{0}, \begin{bmatrix} (\|\boldsymbol{\beta}_1\|_2^2 + \sigma^2) & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_p \end{bmatrix}\right)$.

Goal: Test whether $(y_i, \boldsymbol{x}_i)_{i \in [n]}$ came from $\mathbb{P}$ or from $\mathbb{Q}$, with vanishing error probability as $p \to \infty$.

More Bad News

Symmetric Mixed Linear Regression: $\mathtt{SNR} = \infty$

$$y_i = \pm\langle \boldsymbol{x}_i, \boldsymbol{\beta} \rangle$$

More Bad News

Symmetric Mixed Linear Regression: $\texttt{SNR} = \infty$

$$y_i = \pm\langle \boldsymbol{x}_i, \boldsymbol{\beta}\rangle \overset{|\cdot|}{\Longleftarrow} |y_i| = |\langle \boldsymbol{x}_i, \boldsymbol{\beta}\rangle|$$

More Bad News

Symmetric Mixed Linear Regression: $\texttt{SNR} = \infty$

$$y_i = \pm\langle \boldsymbol{x}_i, \boldsymbol{\beta}\rangle \overset{|\cdot|}{\Longleftarrow} |y_i| = |\langle \boldsymbol{x}_i, \boldsymbol{\beta}\rangle|$$

Corollary. (Conditional on LDC) For $k = o(\sqrt{p})$ and $k \lesssim \texttt{n} \lesssim k^2$, runtime $\exp\left(\frac{k^2}{\texttt{n}}\right)$ is required to solve a detection variant of Sparse Phase Retrieval.

More Bad News

Symmetric Mixed Linear Regression: $\mathtt{SNR} = \infty$

$$y_i = \pm\langle \boldsymbol{x}_i, \boldsymbol{\beta} \rangle \xLeftarrow{|\cdot|} |y_i| = |\langle \boldsymbol{x}_i, \boldsymbol{\beta} \rangle|$$

Corollary. (Conditional on LDC) For $k = o(\sqrt{p})$ and $k \lesssim \mathtt{n} \lesssim k^2$, runtime $\exp\left(\frac{k^2}{\mathtt{n}}\right)$ is required to solve a detection variant of Sparse Phase Retrieval.

$\implies$ Provides rigorous evidence for computational hardness of Sparse Phase Retrieval when $n = \tilde{o}(k^2)$!