Living on the Edge

Phase Transitions in Random Convex Programs

Joel A. Tropp
Michael B. McCoy

Computing + Mathematical Sciences
California Institute of Technology

Joint with Dennis Amelunxen and Martin Lotz (Manchester),
including work of Samet Oymak and Babak Hassibi (Caltech)

Research supported in part by ONR, AFOSR, DARPA, and the Sloan Foundation

Phase Transitions

What is a Phase Transition?

Definition. A *phase transition* is a sharp change in the behavior of a computational problem as its parameters vary.

Example: Sparse linear inverse problem with random data

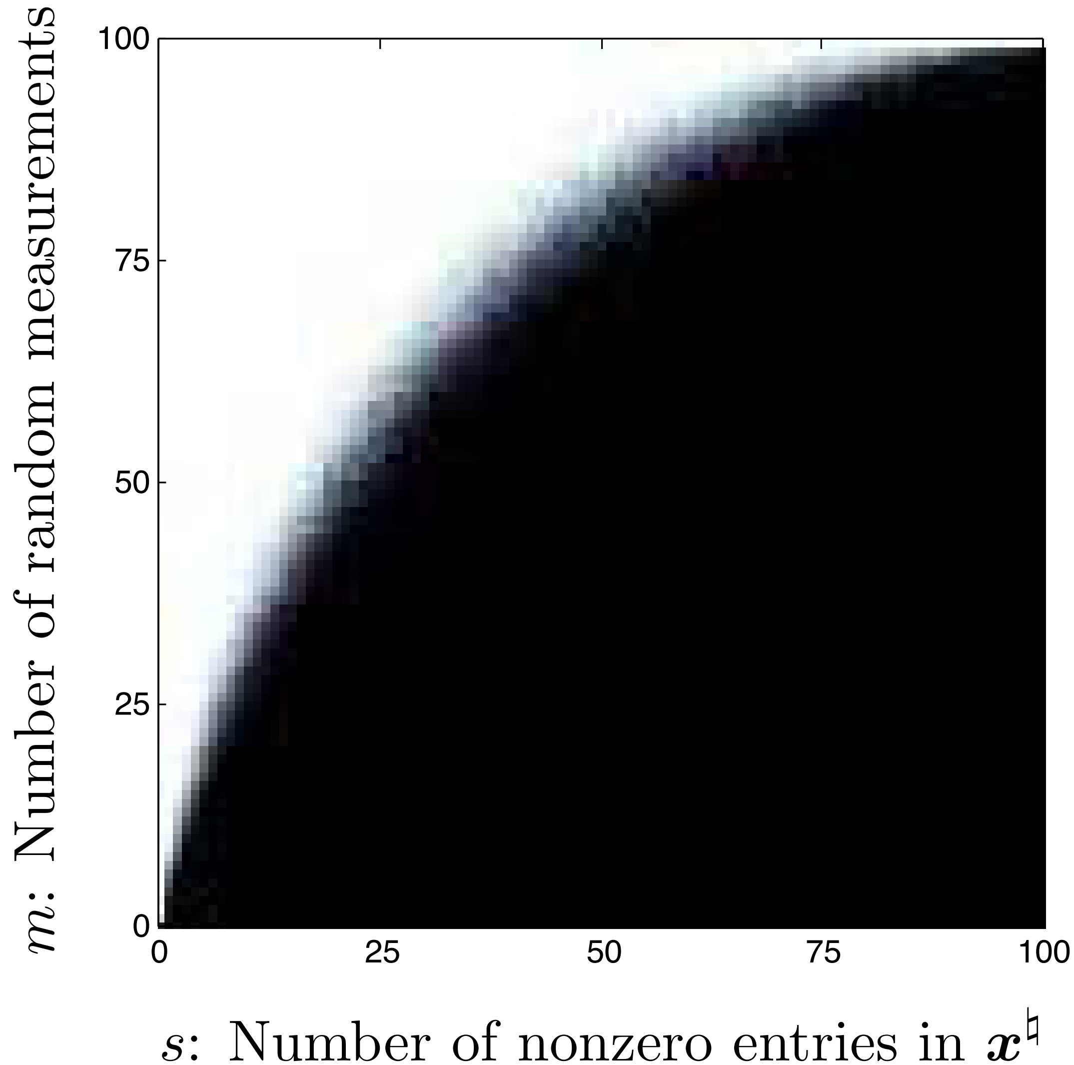
- Suppose $\boldsymbol{x}^\natural \in \mathbb{R}^d$ has s nonzero entries
- Acquire m random linear measurements of $\boldsymbol{x}^\natural$

$$z_i = \langle \boldsymbol{g}_i,\ \boldsymbol{x}^\natural \rangle \quad \text{for } i = 1, \dots, m$$

- Solve a convex optimization problem to reconstruct $\boldsymbol{x}^\natural$ from the data

$$\text{minimize} \quad \|\boldsymbol{x}\|_1 \quad \text{subject to} \quad \langle \boldsymbol{g}_i,\ \boldsymbol{x} \rangle = z_i \quad \text{for } i = 1, \dots, m$$

Example: Sparse Linear Inversion



Research Challenge...

Understand and predict phase transitions in random convex programs

Random Convex Programs

Examples...

- **Sensing.** Collect random measurements; reconstruct via optimization
- **Statistics.** Random data models; fit model via optimization
- **Coding.** Random channel models; decode via optimization

Motivations...

- **Average-case analysis.** Randomness describes "typical" behavior
- **Fundamental bounds.** Opportunities and limits for convex methods

Warmup:
Regularized Denoising

Setup for Regularized Denoising

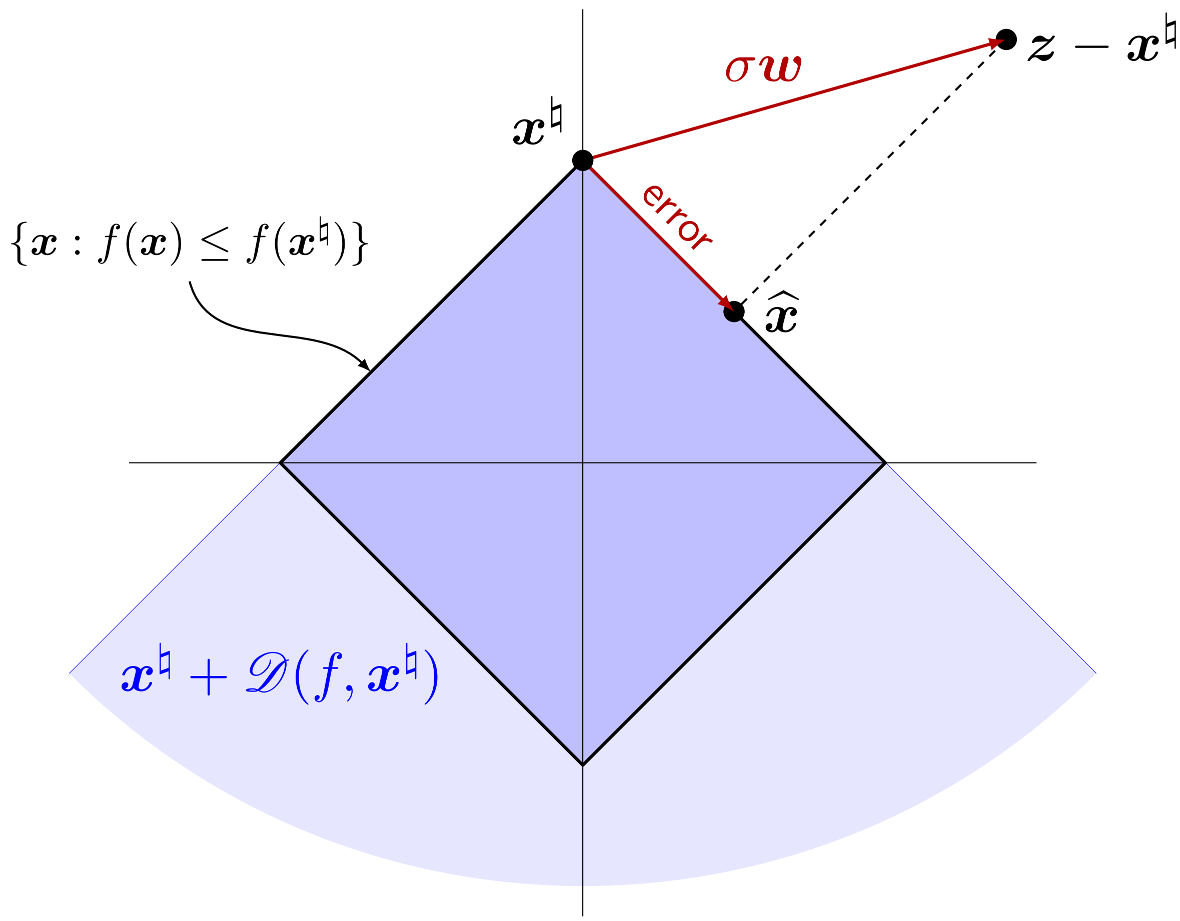
- Let $\boldsymbol{x}^\natural \in \mathbb{R}^d$ be "structured" but unknown
- Let $f : \mathbb{R}^d \to \mathbb{R}$ be a convex function that measures "structure"
- Observe $\boldsymbol{z} = \boldsymbol{x}^\natural + \sigma \boldsymbol{w}$ where $\boldsymbol{w} \sim \textsc{normal}(0, \mathbf{I})$
- Remove noise by solving the convex program*

$$\text{minimize} \quad \frac{1}{2}\|\boldsymbol{z} - \boldsymbol{x}\|_2^2 \quad \text{subject to} \quad f(\boldsymbol{x}) \le f(\boldsymbol{x}^\natural)$$

- **Hope:** The minimizer $\widehat{\boldsymbol{x}}$ approximates $\boldsymbol{x}^\natural$

*We assume the side information $f(\boldsymbol{x}^\natural)$ is available. This is equivalent** to knowing the optimal choice of Lagrange multiplier for the constraint.

Geometry of Denoising



The Risk of Regularized Denoising

Theorem 1. [Oymak & Hassibi 2013] **Assume**

- *We observe* $\boldsymbol{z} = \boldsymbol{x}^\natural + \sigma \boldsymbol{w}$ *where* $\boldsymbol{w}$ *is standard normal*

- *The vector* $\widehat{\boldsymbol{x}}$ *solves*

$$\text{minimize} \quad \frac{1}{2}\|\boldsymbol{z} - \boldsymbol{x}\|_2^2 \quad \text{subject to} \quad f(\boldsymbol{x}) \leq f(\boldsymbol{x}^\natural)$$

Then

$$\sup_{\sigma > 0} \frac{\mathbb{E}\,\|\widehat{\boldsymbol{x}} - \boldsymbol{x}^\natural\|^2}{\sigma^2} = \delta\big(\mathscr{D}(f, \boldsymbol{x}^\natural)\big)$$

where $\delta(\mathscr{D}(f, \boldsymbol{x}^\natural))$ *denotes the statistical dimension of the descent cone*

Statistical

Dimension

The Statistical Dimension

Definition. [Amelunxen, Lotz, McCoy, T 2013]

The *statistical dimension* of a closed, convex cone K is

$$\delta(K) := \mathbb{E}\left[\, \|\mathbf{\Pi}_K(\boldsymbol{g})\|_2^2 \,\right]$$

where

- $\mathbf{\Pi}_K$ is the Euclidean projection onto K
- $\boldsymbol{g}$ is a standard normal vector

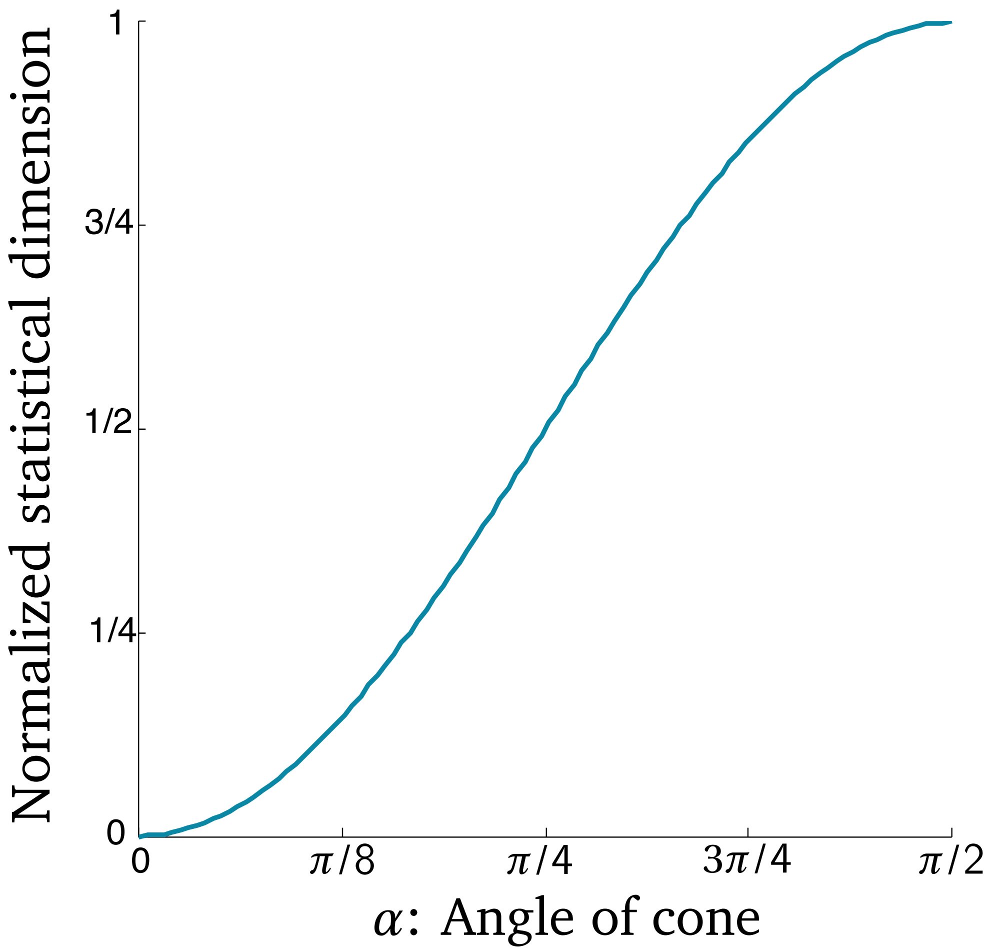
Intuition...

In stochastic geometry, a convex cone K with statistical dimension $\delta(K)$ behaves like a subspace with dimension $\lceil \delta(K) \rceil$

Basic Examples

Cone	Notation	Statistical Dimension
Subspace	L_j	j
Nonnegative orthant	$\mathbb{R}^d_+$	$\frac{1}{2}d$
Second-order cone	$\mathbb{L}^{d+1}$	$\frac{1}{2}(d+1)$
Real psd cone	$\mathbb{S}^d_+$	$\frac{1}{4}d(d-1)$
Complex psd cone	$\mathbb{H}^d_+$	$\frac{1}{2}d^2$

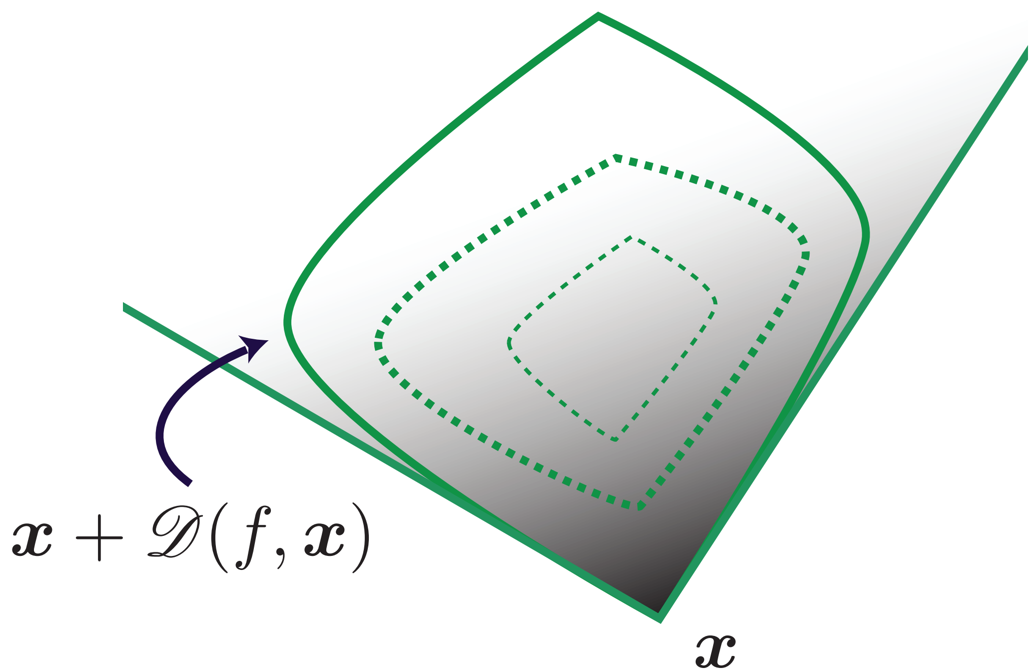
Circular Cones



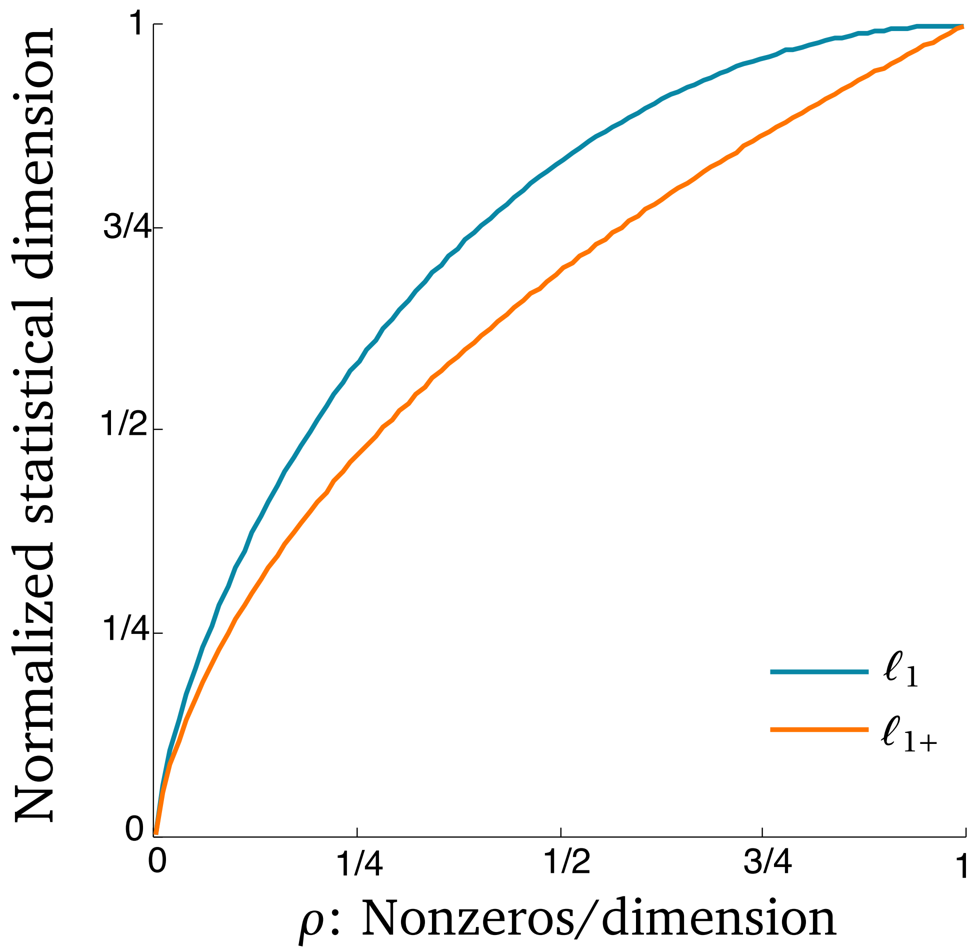
Descent Cones

Definition. The *descent cone* of a function f at a point $\boldsymbol{x}$ is

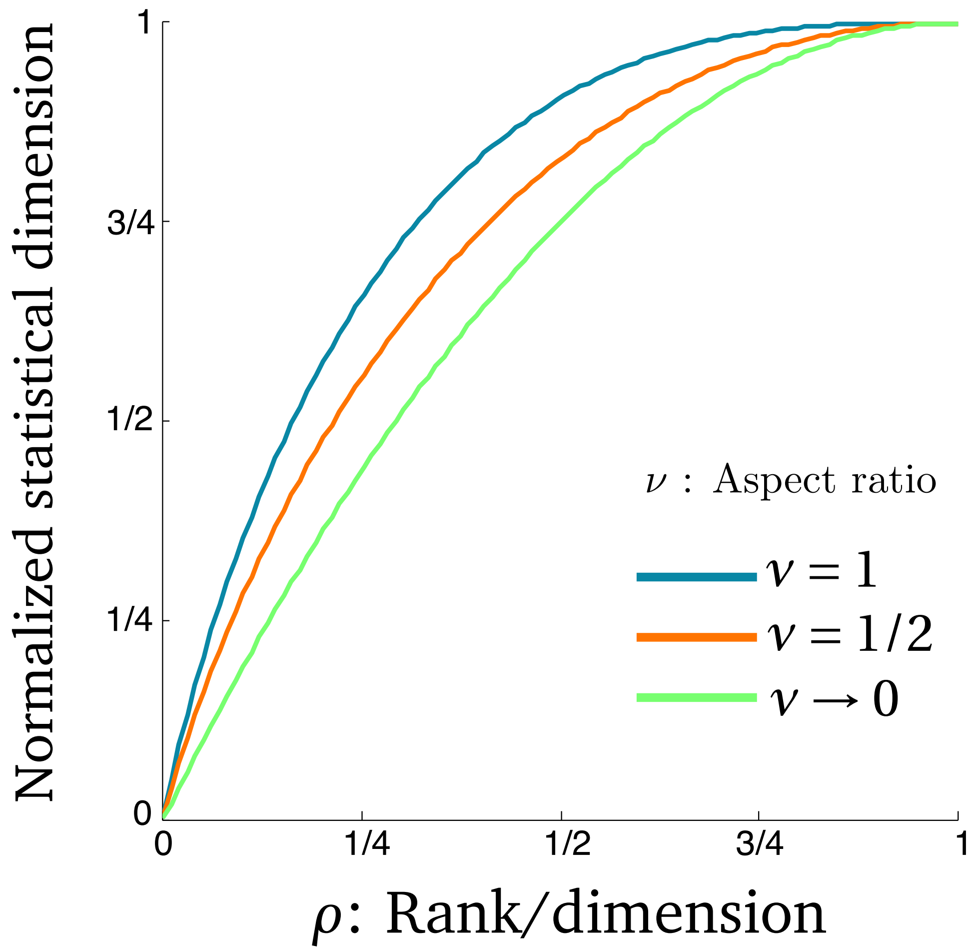
$$\mathscr{D}(f, \boldsymbol{x}) := \{\boldsymbol{h} : f(\boldsymbol{x} + \varepsilon\boldsymbol{h}) \le f(\boldsymbol{x}) \quad \texttt{for some } \varepsilon > 0\}$$



Descent Cone of ℓ_1 Norm at Sparse Vector



Descent Cone of S_1 Norm at Low-Rank Matrix



Statistical Dimension & Phase Transitions

- **Key Question:** When do two randomly oriented cones strike?

- **Intuition:** When do randomly oriented subspaces strike?

The Approximate Kinematic Formula

[Amelunxen, Lotz, McCoy, T 2013]

Let C and K be closed convex cones in $\mathbb{R}^d$

$$\delta(C) + \delta(K) \lesssim d \quad \Longrightarrow \quad \mathbb{P}\left\{C \cap \boldsymbol{Q}K = \{\boldsymbol{0}\}\right\} \approx 1$$

$$\delta(C) + \delta(K) \gtrsim d \quad \Longrightarrow \quad \mathbb{P}\left\{C \cap \boldsymbol{Q}K = \{\boldsymbol{0}\}\right\} \approx 0$$

where $\boldsymbol{Q}$ is a random orthogonal matrix

Regularized Linear Inverse Problems

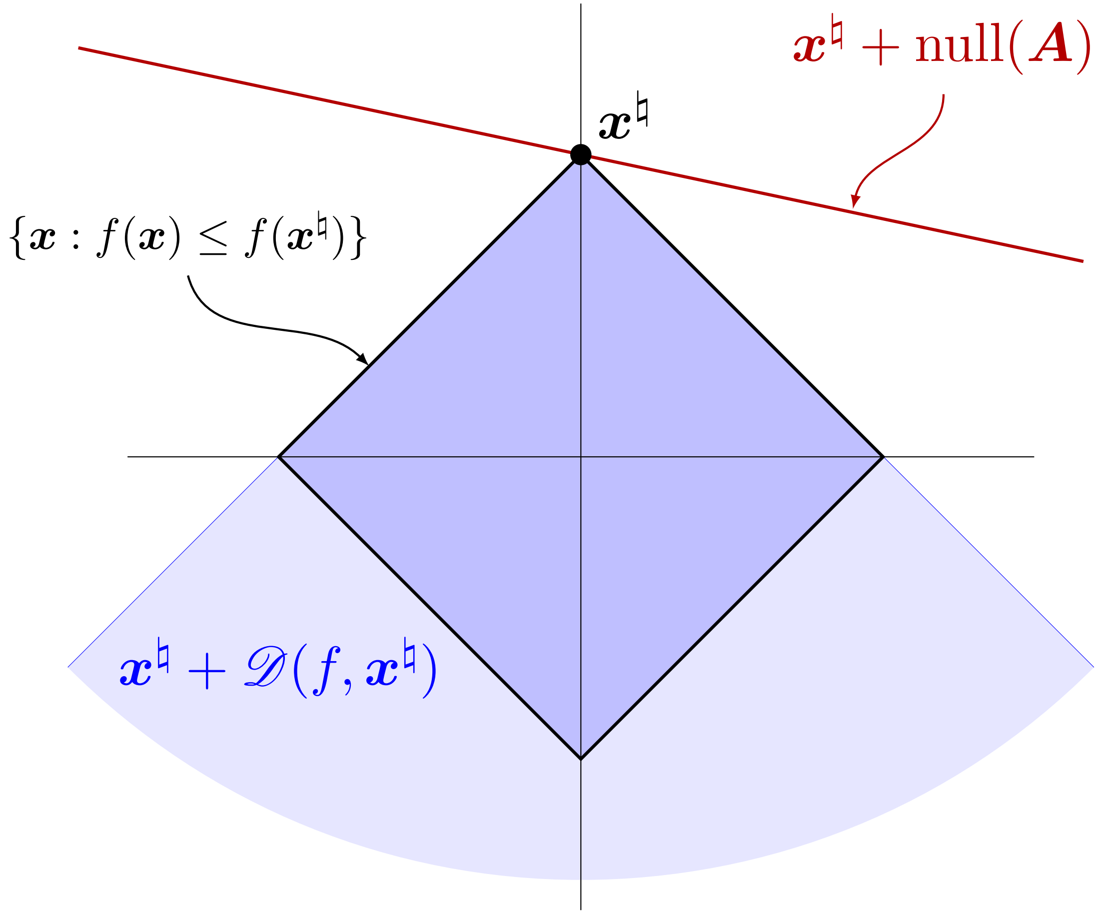
Setup for Linear Inverse Problems

- Let $\boldsymbol{x}^\natural \in \mathbb{R}^d$ be a structured, unknown vector
- Let $\boldsymbol{A} \in \mathbb{R}^{m \times d}$ be a measurement operator
- Observe $\boldsymbol{z} = \boldsymbol{A}\boldsymbol{x}^\natural$
- Find estimate $\widehat{\boldsymbol{x}}$ by solving convex program

$$\text{minimize} \quad f(\boldsymbol{x}) \quad \text{subject to} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{z}$$

- **Hope:** $\widehat{\boldsymbol{x}} = \boldsymbol{x}^\natural$

Geometry of Linear Inverse Problems



Linear Inverse Problems with Random Data

Theorem 2. [Amelunxen, Lotz, McCoy, T 2013] **Assume**

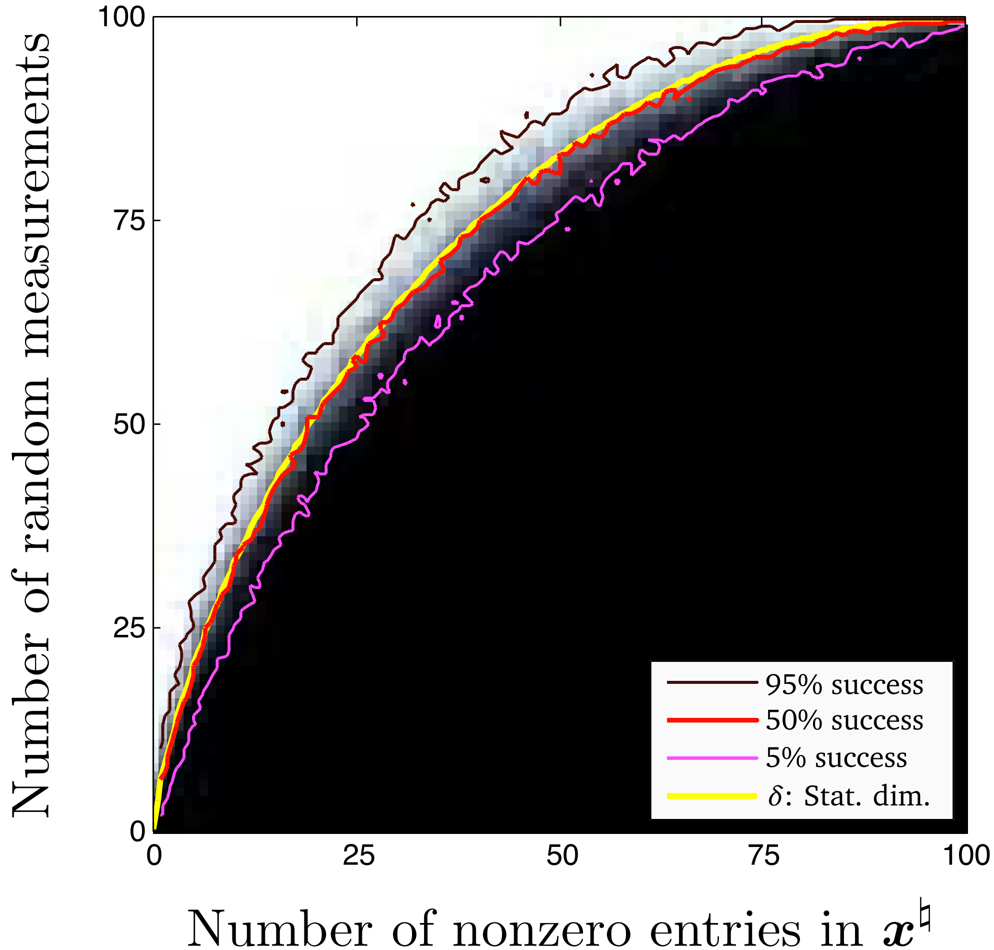
- *The vector* $\boldsymbol{x}^\natural \in \mathbb{R}^d$ *is unknown*
- *The observation* $\boldsymbol{z} = \boldsymbol{A}\boldsymbol{x}^\natural$ *where* $\boldsymbol{A} \in \mathbb{R}^{m \times d}$ *is standard normal*
- *The vector* $\widehat{\boldsymbol{x}}$ *solves*

$$\text{minimize} \quad f(\boldsymbol{x}) \quad \text{subject to} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{z}$$

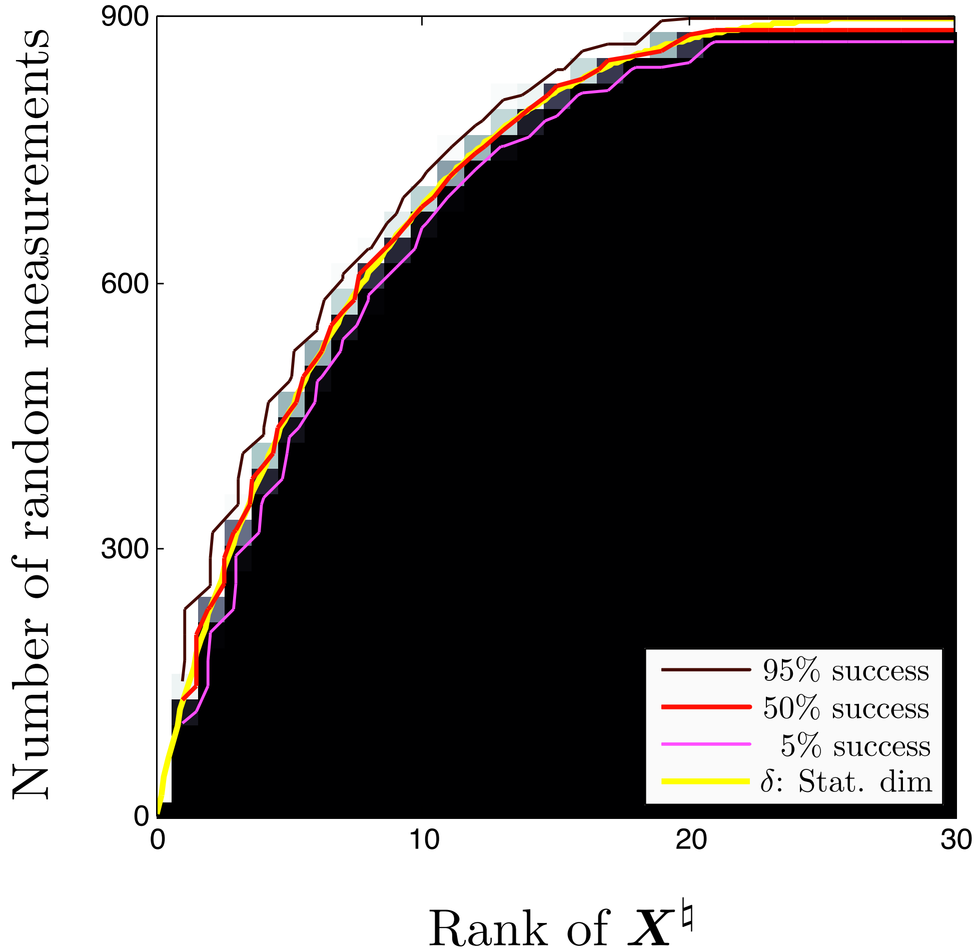
Then

$$m \gtrsim \delta\big(\mathscr{D}(f, \boldsymbol{x}^\natural)\big) \quad \Longrightarrow \quad \widehat{\boldsymbol{x}} = \boldsymbol{x}^\natural \quad \text{whp}$$
$$m \lesssim \delta\big(\mathscr{D}(f, \boldsymbol{x}^\natural)\big) \quad \Longrightarrow \quad \widehat{\boldsymbol{x}} \neq \boldsymbol{x}^\natural \quad \text{whp.}$$

Sparse Reconstruction via ℓ_1 Minimization



Low-Rank Recovery via S_1 Minimization



Demixing
Structured Signals

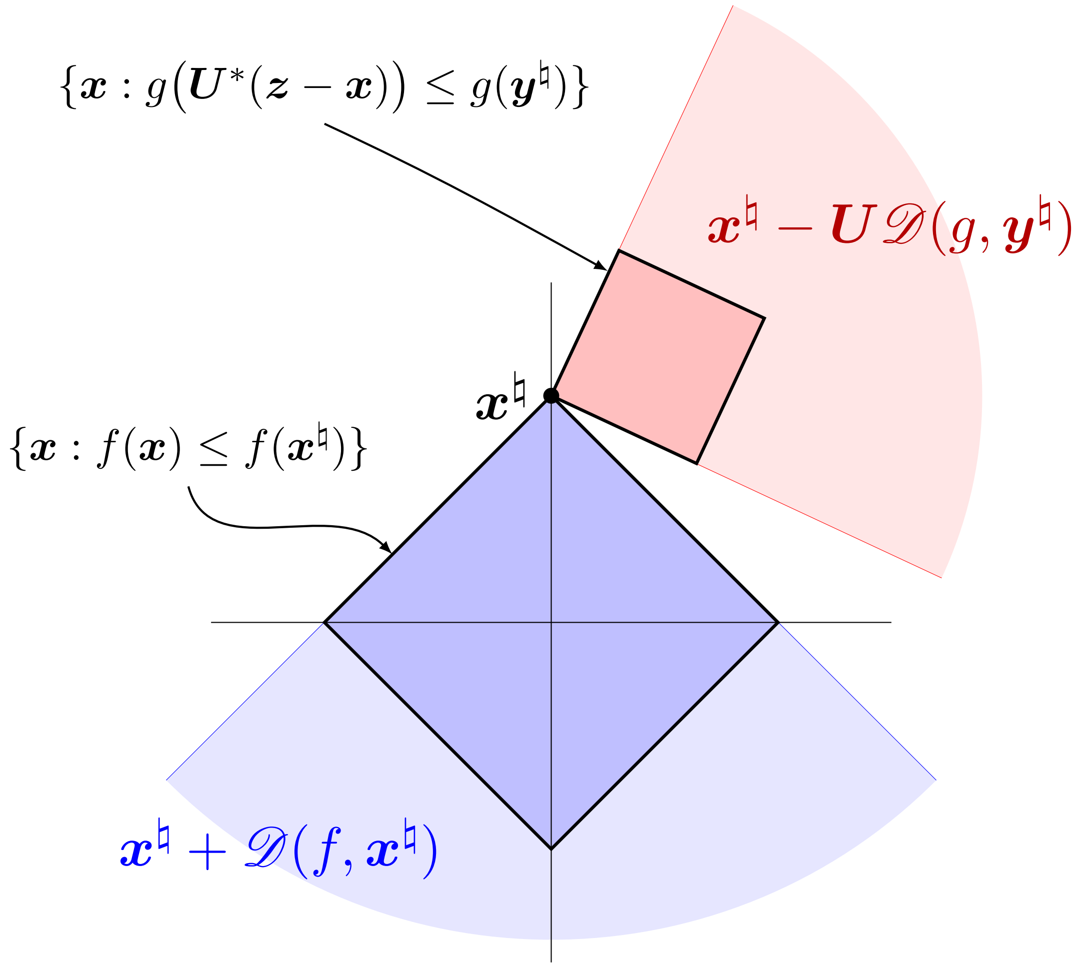
Setup for Demixing Problems

- Let $\boldsymbol{x}^\natural \in \mathbb{R}^d$ and $\boldsymbol{y}^\natural \in \mathbb{R}^d$ be structured, unknown vectors
- Let $\boldsymbol{U} \in \mathbb{R}^{d \times d}$ be a known orthogonal matrix
- Observe $\boldsymbol{z} = \boldsymbol{x}^\natural + \boldsymbol{U}\boldsymbol{y}^\natural$
- Reconstruct via convex program

$$\begin{aligned} \text{minimize} \quad f(\boldsymbol{x}) \quad \text{subject to} \quad & g(\boldsymbol{y}) \le g(\boldsymbol{y}^\natural) \\ & \boldsymbol{x} + \boldsymbol{U}\boldsymbol{y} = \boldsymbol{z} \end{aligned}$$

- **Hope:** $(\widehat{\boldsymbol{x}}, \widehat{\boldsymbol{y}}) = (\boldsymbol{x}^\natural, \boldsymbol{y}^\natural)$

Geometry of Demixing Problems



Demixing Problems with Random Incoherence

Theorem 3. [Amelunxen, Lotz, McCoy, T 2013] **Assume**

- *The vectors* $\boldsymbol{x}^\natural \in \mathbb{R}^d$ *and* $\boldsymbol{y}^\natural \in \mathbb{R}^d$ *are unknown*
- *The observation* $\boldsymbol{z} = \boldsymbol{x}^\natural + \boldsymbol{Q}\boldsymbol{y}^\natural$ *where* $\boldsymbol{Q}$ *is random orthogonal*
- *The pair* $(\widehat{\boldsymbol{x}}, \widehat{\boldsymbol{y}})$ *solves*

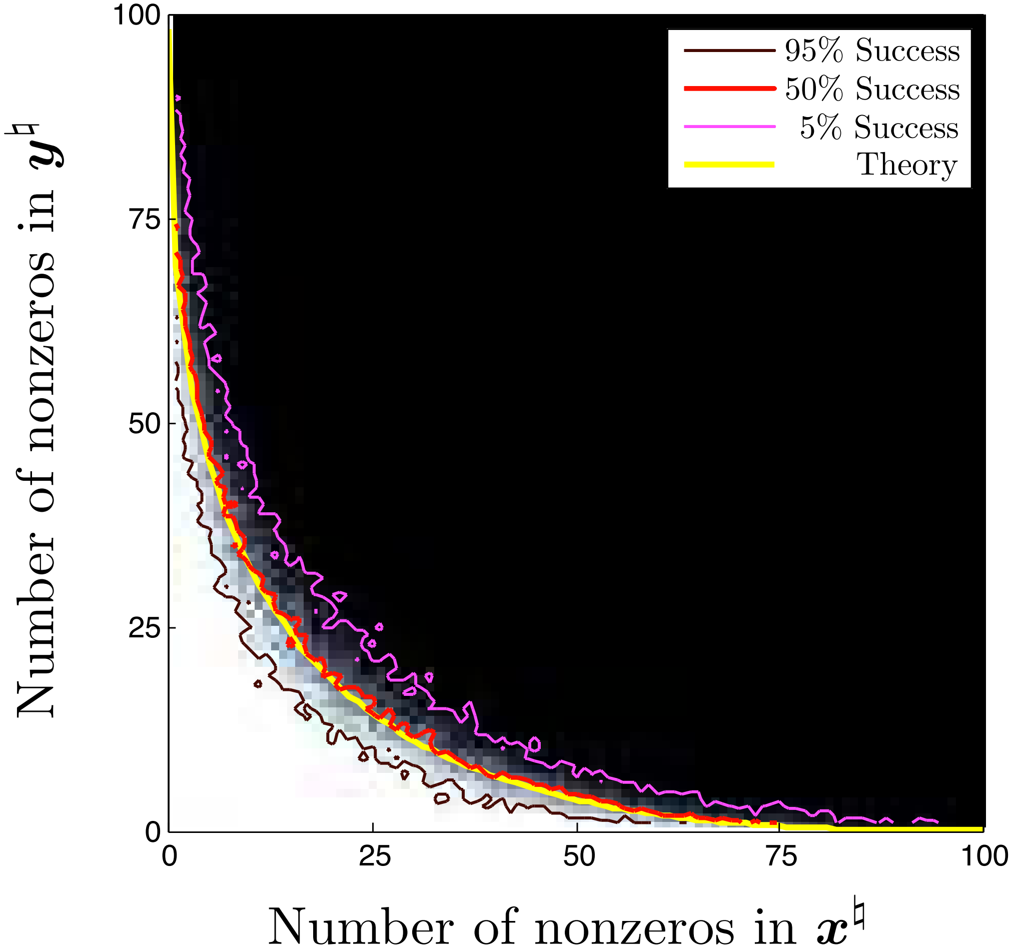
$$\text{minimize} \quad f(\boldsymbol{x}) \quad \text{subject to} \quad g(\boldsymbol{y}) \leq g(\boldsymbol{y}^\natural), \quad \boldsymbol{x} + \boldsymbol{Q}\boldsymbol{y} = \boldsymbol{z}$$

Then

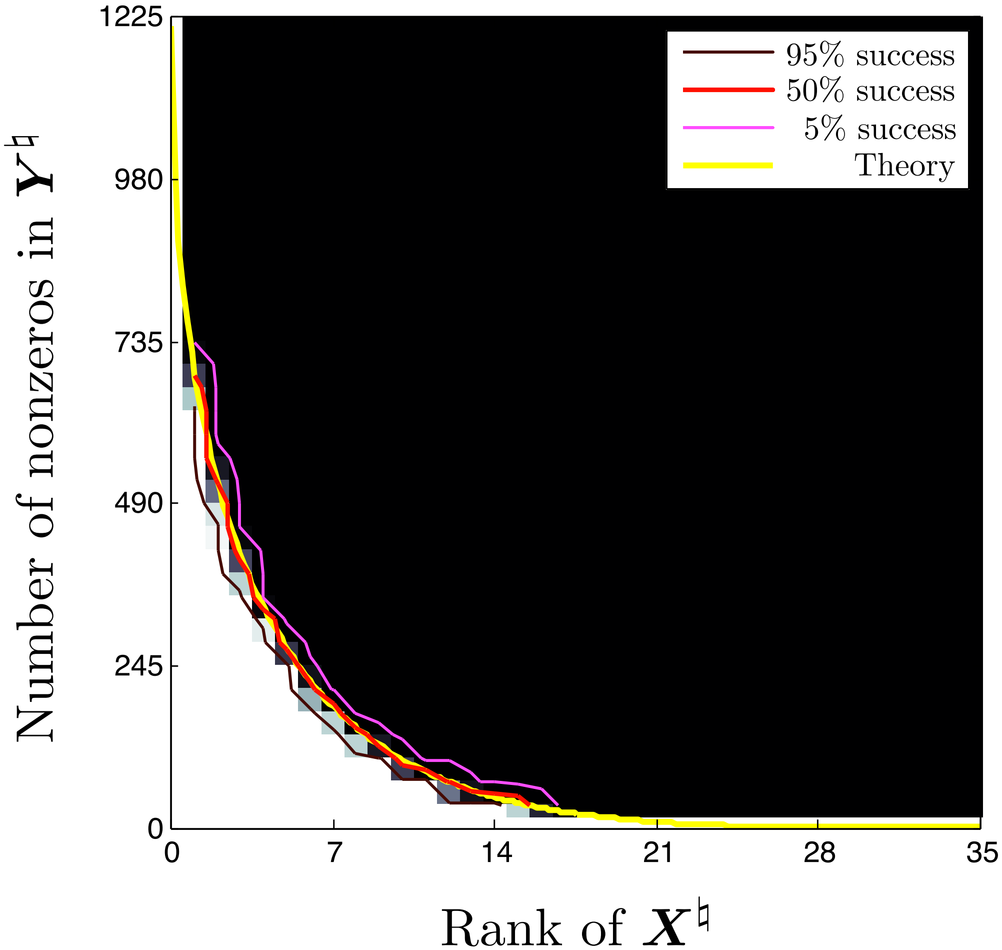
$$\delta\big(\mathscr{D}(f, \boldsymbol{x}^\natural)\big) + \delta\big(\mathscr{D}(g, \boldsymbol{y}^\natural)\big) \lesssim d \quad \Longrightarrow \quad (\widehat{\boldsymbol{x}}, \widehat{\boldsymbol{y}}) = (\boldsymbol{x}^\natural, \boldsymbol{y}^\natural) \quad \textit{whp}$$

$$\delta\big(\mathscr{D}(f, \boldsymbol{x}^\natural)\big) + \delta\big(\mathscr{D}(g, \boldsymbol{y}^\natural)\big) \gtrsim d \quad \Longrightarrow \quad (\widehat{\boldsymbol{x}}, \widehat{\boldsymbol{y}}) \neq (\boldsymbol{x}^\natural, \boldsymbol{y}^\natural) \quad \textit{whp}$$

Sparse + Sparse via $\ell_1 + \ell_1$ Minimization



Low-Rank + Sparse via $S_1 + \ell_1$ Minimization



Cone Programs with Random Constraints

Cone Program with Random Constraints

Theorem 4. [Amelunxen, Lotz, McCoy, T 2013] **Assume**

- *The cone* K *is proper*
- *The vectors* $\boldsymbol{u} \in \mathbb{R}^d$ *and* $\boldsymbol{b} \in \mathbb{R}^m$ *are standard normal*
- *The matrix* $\boldsymbol{A} \in \mathbb{R}^{m \times d}$ *is standard normal*

Consider the cone program

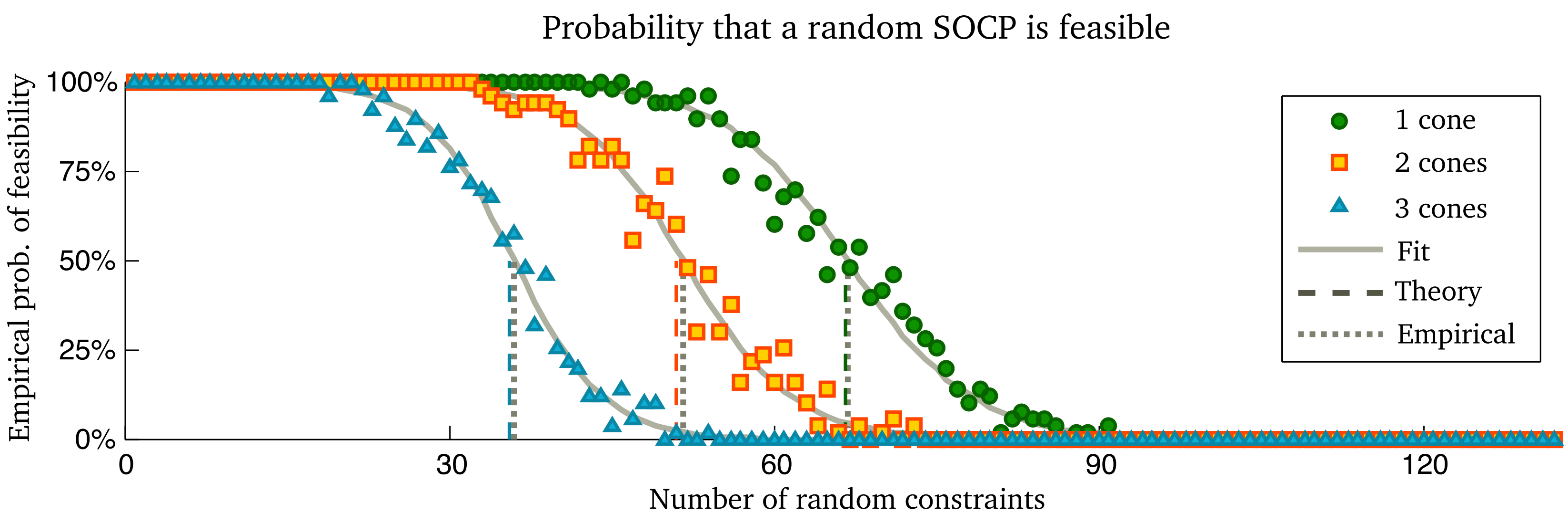
$$\text{minimize} \quad \langle \boldsymbol{u},\ \boldsymbol{x} \rangle \quad \text{subject to} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \quad \text{and} \quad \boldsymbol{x} \in K$$

Then

$$m \lesssim \delta(K) \quad \Longrightarrow \quad \text{the cone program is unbounded whp}$$

$$m \gtrsim \delta(K) \quad \Longrightarrow \quad \text{the cone program is infeasible whp}$$

Example: Some Random SOCPs



To learn more...

E-mail: `mccoy@cms.caltech.edu`
`jtropp@cms.caltech.edu`

Web: `http://users.cms.caltech.edu/~mccoy`
`http://users.cms.caltech.edu/~jtropp`

Papers:

- MT, "Sharp recovery bounds for convex deconvolution, with applications." arXiv cs.IT 1205.1580
- ALMT, "Living on the edge: A geometric theory of phase transitions in convex optimization." arXiv cs.IT 1303.6672
- Oymak & Hassibi, "Asymptotically exact denoising in relation to compressed sensing," arXiv cs.IT 1305.2714
- More to come!