# Consistent Basis Pursuit (CoBP) for Low-Complexity Signal Estimates in Quantized Compressed Sensing

Laurent Jacques

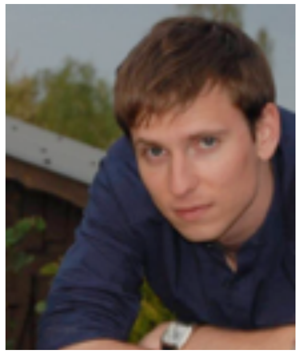
UCLouvain, Belgium

November 19th, 2015

Joint work with:

Amirafshar Moshtaghpour

Valerio Cambareri



Kevin Degraux

Christophe De Vleeschouwer

# Outline

1. Introduction to CS and QCS

2. Consistent Basis Pursuit for low-complexity signals

3. Quasi-isometric embeddings of low-complexity signals

4. Take-away messages & open questions

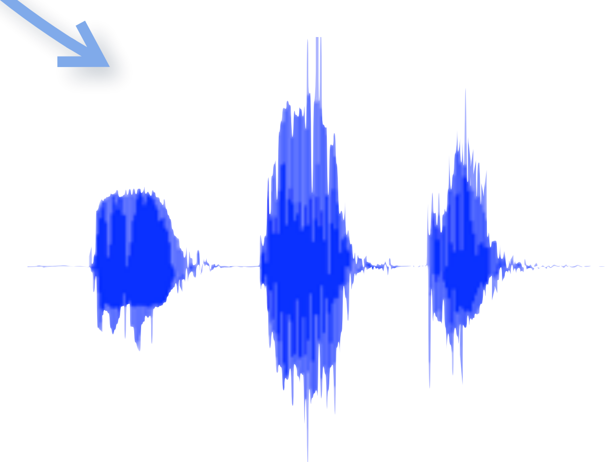
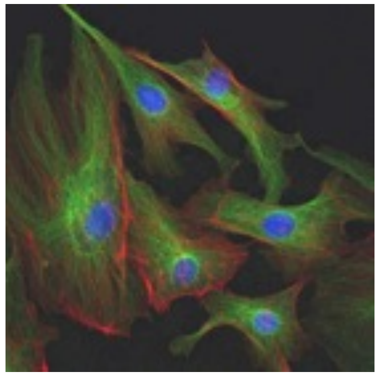
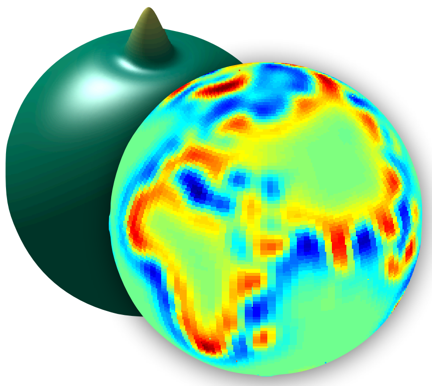
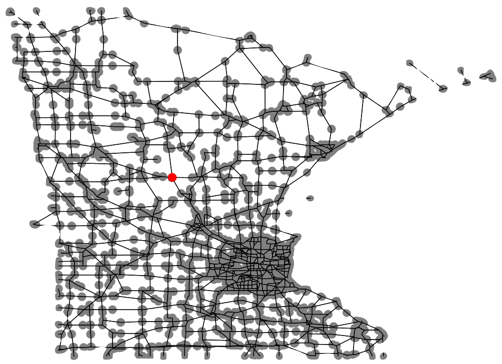
# 1. Introduction

# CS facts

# Compressed Sensing...

## *... in a nutshell:*

Generalize Dirac/Nyquist sampling:
1°) ask *few* (**linear**) *questions*
about your ***informative*** signal
2°) and recover it *differently* (**non-linearly**)"

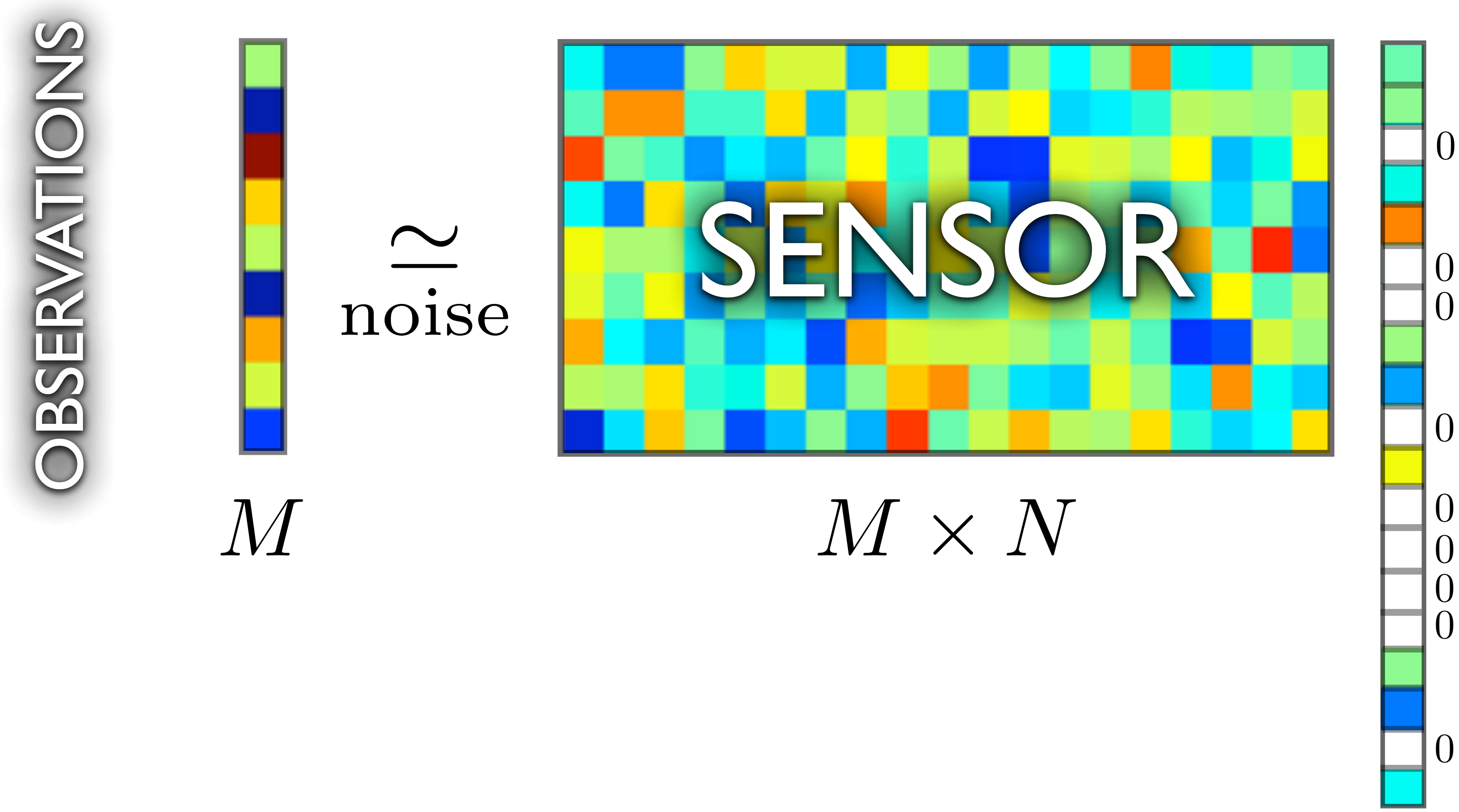





**e.g., sparse, structured, low-rank, ...**

# 1st, CS $\ni$ Generalized Linear Sensing!

$M$ questions | Sensing method | Signal

$$\boldsymbol{y} \underset{\text{noise}}{\simeq} \boldsymbol{\Phi}\, \boldsymbol{x}$$

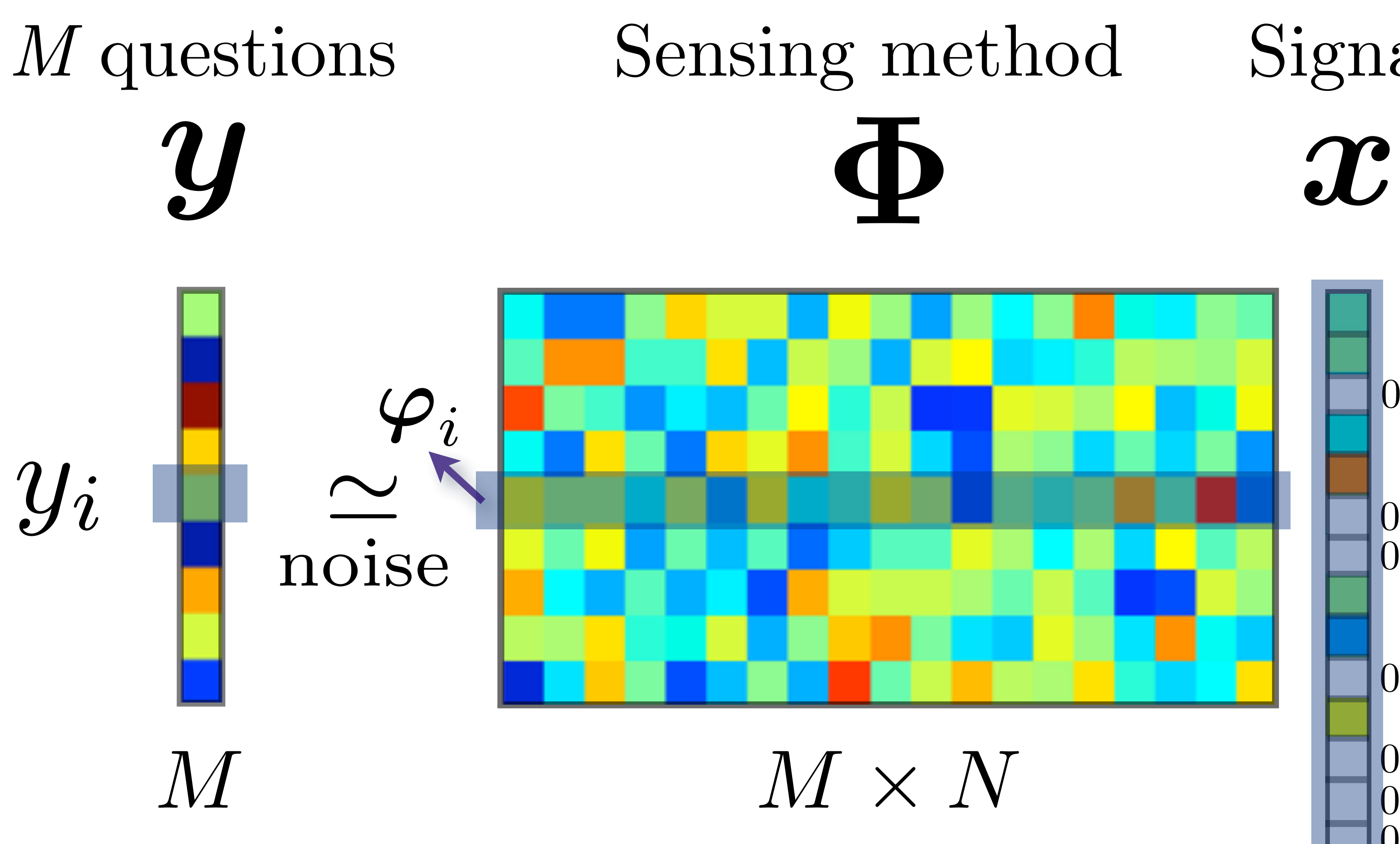



$M$    $M \times N$    $N$

Sparsity Prior ($\Psi = \mathrm{Id}$)

A signal in this discrete world

# 1st, CS $\ni$ Generalized Linear Sensing!




Sparsity Prior ($\Psi = \mathrm{Id}$)

A signal in this discrete world

**Generalized Linear Sensing!**

$$y_i \simeq \langle \boldsymbol{\varphi}_i, \boldsymbol{x} \rangle = \boldsymbol{\varphi}_i^T \boldsymbol{x}$$

$$1 \leq i \leq M$$

e.g., to be realized optically/analogically

# 2nd, CS $\ni$ Non-linear reconstruction!

Mmm, $M$ equations, $N$ unknowns?!

Ill-posed problem

You must regularize it!

(*intuition*: would you know the signal support,
much less unknowns)

# 2nd, CS $\ni$ Non-linear reconstruction!

<u>Possible reconstruction:</u> (others exist, e.g., greedy)

(Basis Pursuit DeNoise)

[Chen, Donoho, Saunders, 1998]

$$\boxed{\hat{\boldsymbol{x}} \;=\; \arg\min_{\boldsymbol{u}\in\mathbb{R}^N} \;\|\boldsymbol{u}\|_1 \text{ s.t. } \|\boldsymbol{y}-\boldsymbol{\Phi}\boldsymbol{u}\|\leq\epsilon}$$

Sparsity promotion: $\|\boldsymbol{u}\|_1 = \sum_j |u_j|$

Level of "noise"

Convexification of $\ell_0$-norm:
$\|\boldsymbol{u}\|_0 = |\mathrm{supp}\,\boldsymbol{u}| = |\{k : u_k \neq 0\}|$

# 2nd, CS $\ni$ Non-linear reconstruction!

BPDN instance optimality:

If $\frac{1}{\sqrt{M}}\boldsymbol{\Phi}$ respects the Restricted Isometry Property (RIP)

$$\boxed{(1-\rho)\|\boldsymbol{u}\|^2 \leq \tfrac{1}{M}\|\boldsymbol{\Phi u}\|^2 \leq (1+\rho)\|\boldsymbol{u}\|^2}$$

for all $\boldsymbol{u} \in \Sigma_{2K} := \{\boldsymbol{u} : \|\boldsymbol{u}\|_0 := |\operatorname{supp}\boldsymbol{u}| \leq 2K\}$

# 2nd, CS $\ni$ Non-linear reconstruction!

BPDN instance optimality:

If $\frac{1}{\sqrt{M}}\boldsymbol{\Phi}$ respects the Restricted Isometry Property (RIP)

$$\boxed{(1-\rho)\|\boldsymbol{u}\|^2 \leq \tfrac{1}{M}\|\boldsymbol{\Phi u}\|^2 \leq (1+\rho)\|\boldsymbol{u}\|^2}$$

for all $\boldsymbol{u} \in \Sigma_{2K} := \{\boldsymbol{u} : \|\boldsymbol{u}\|_0 := |\operatorname{supp}\boldsymbol{u}| \leq 2K\}$

Then, if $\rho < \sqrt{2}-1$ [Candès, 09],

(with $f \lesssim g \;\equiv\; \exists c > 0 : f \leqslant c\,g$)

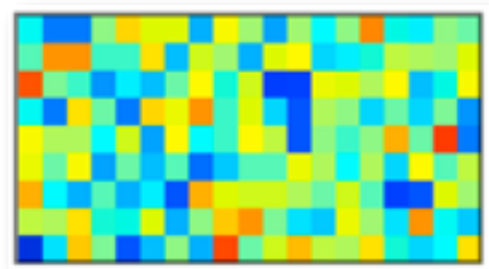
Robustness: *vs* sparse deviation + noise.

$$\|\boldsymbol{x} - \hat{\boldsymbol{x}}\| \;\lesssim\; \underbrace{\frac{1}{\sqrt{K}}\|\boldsymbol{x} - \boldsymbol{x}_K\|_1}_{e_0(K)} + \frac{\epsilon}{\sqrt{M}}$$

# 2nd, CS $\ni$ Non-linear reconstruction!

## Matrices with RIP?

$\mathbf{\Phi} \in \mathbb{R}^{M \times N}$, with $\Phi_{ij} \sim_{\rm iid} \mathcal{N}(0,1)$ and $M \gtrsim K \log N/K$.



## but also:

- Random sub-Gaussian ensembles (e.g., Bernoulli);
- random Fourier/Hadamard ensembles (structured sensing);
- random convolutions, spread-spectrum;
- ...

(see, *e.g.*, “A Mathematical Introduction to Compressive Sensing”, Rauhut, Foucart, Springer, 2013)

# Quantization context

## (Restricted to scalar quantization)

*Caveat* : **not** covered here:

- Sigma-Delta quantization for CS
  (see, *e.g.*, Kramer, Saab, Güntürk, Powell, Ward, ...)
- Vector quantization
  (see, *e.g.*, Goyal, Nguyen, Sun, ...)
- Universal quantization (periodic)
  (see, *e.g.*, Boufounos, Rane, ...)

# Compressive Sampling and Quantization

Compressed sensing theorist says:

*"Linearly sample a signal*

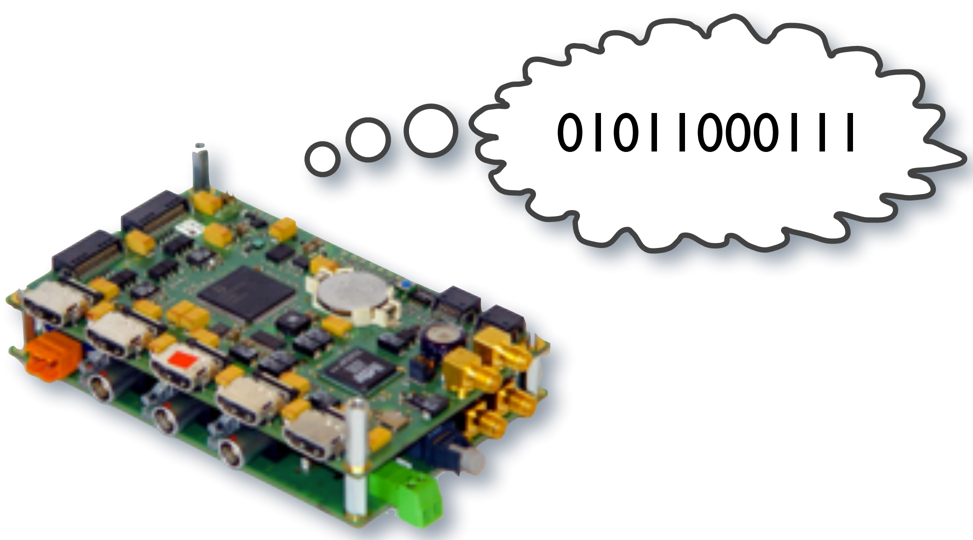
*at a rate function of*

*its intrinsic dimensionality"*


Information theorist and sensor designer say:

*"Okay, but I need to quantize/digitize my measurements!"*

*(e.g., in ADC)*




**Integration?**

**QCS theory?**

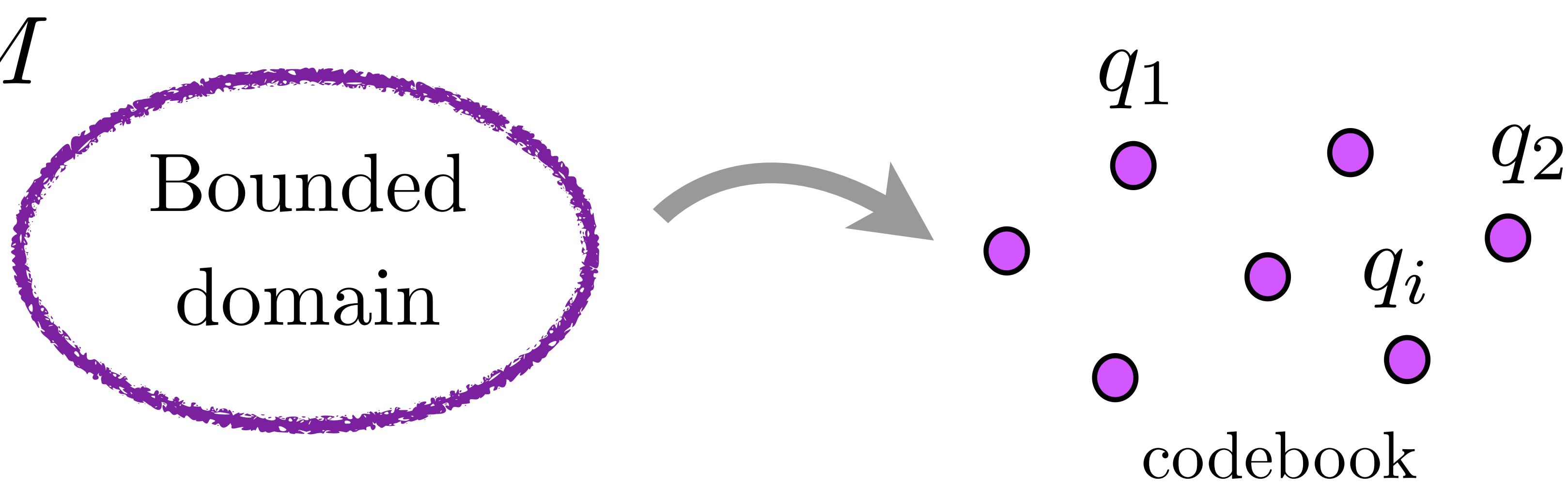
**Theoretical Bounds**

# What is quantization?

- <u>Generality</u>:

  Intuitively: "*Quantization maps a bounded continuous domain to a set of finite elements (or codebook)*"

$\mathbb{R}^M$




$$\mathcal{Q}[x] \in \{q_1, q_2, \cdots\}$$

- Oldest example: rounding off $\lfloor x \rfloor, \lceil x \rceil, \ldots \quad \mathbb{R} \to \mathbb{Z}$

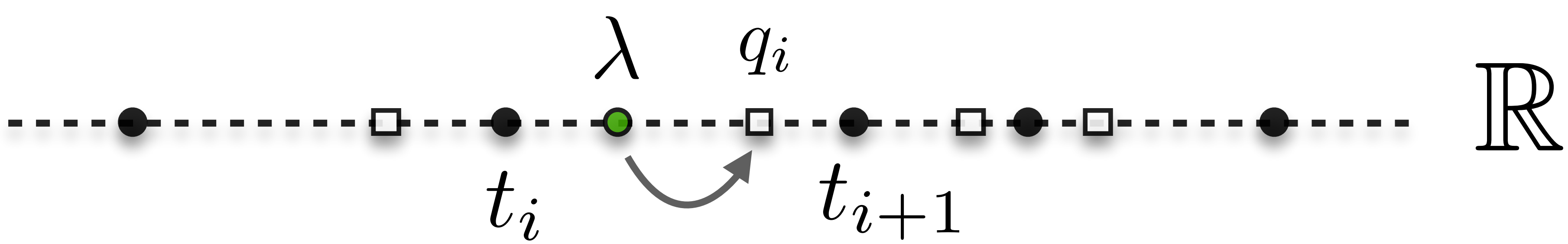
# Scalar quantization

Pulse Code Modulation - PCM
Memoryless Scalar Quantization - MSQ

Applied on each component of $M$- dimensional vectors:

$$\mathcal{Q}(\lambda) = q_i$$

□ : Level $\Omega = \{q_i\}$ (or codebook) ● : Thresholds $\mathcal{T} = \{t_i\}$



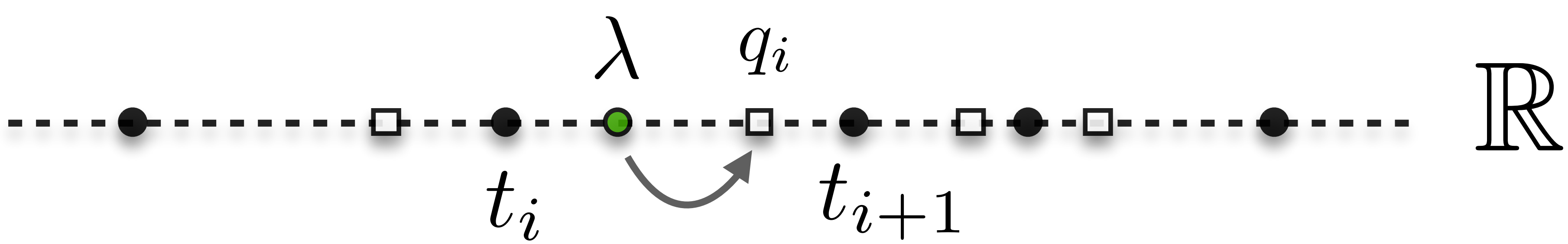
# Scalar quantization

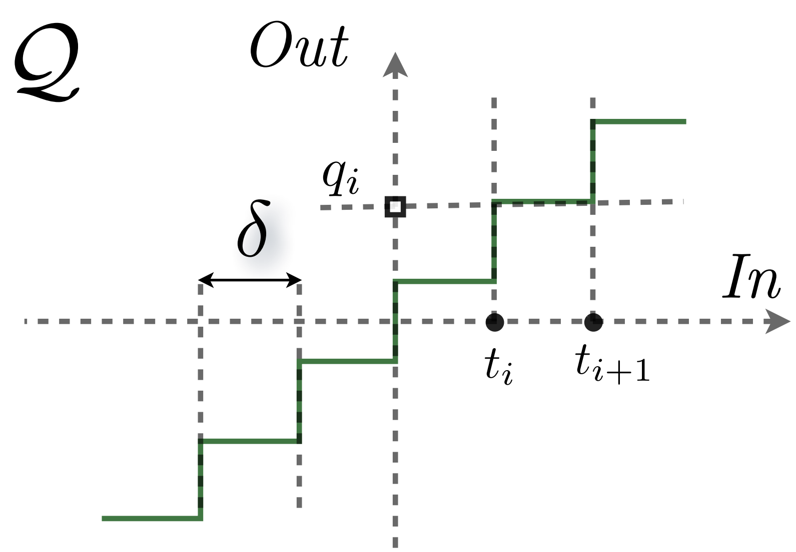
Pulse Code Modulation - PCM
Memoryless Scalar Quantization - MSQ

Applied on each component of $M$- dimensional vectors:

$$\mathcal{Q}(\lambda) = q_i$$

□ : Level $\Omega = \{q_i\}$ (or codebook) ● : Thresholds $\mathcal{T} = \{t_i\}$




Example: uniform, resolution $\delta > 0$

$$q_k = (k + 1/2)\delta$$

$$t_k = k\delta$$

$$\mathcal{Q}(t) = \delta(\lfloor \tfrac{t}{\delta} \rfloor + \tfrac{1}{2})$$

... with possible non-uniform adaption (Lloyd-Max)

# Quantizing Compressed Sensing?

With no additional noise:

e.g., basis pursuit, greedy methods, ...

$\mathbb{R}^M$ | $\boldsymbol{\Omega}$ codebook

$$\boldsymbol{x} \xrightarrow{\mathbb{R}^N} \boxed{\text{CS}} \xrightarrow{\boldsymbol{y} = \boldsymbol{\Phi x}} \boxed{\mathcal{Q}} \dashrightarrow{\boldsymbol{q} = \mathcal{Q}[\boldsymbol{y}]} \boxed{\text{Decoder}} \xrightarrow{\mathbb{R}^N} \hat{\boldsymbol{x}}$$

Finite codebook $\Rightarrow \hat{\boldsymbol{x}} \neq \boldsymbol{x}$

*i.e.*, impossibility to encode continuous domain in a finite number of elements.

# Quantizing Compressed Sensing?

With no additional noise:

e.g., basis pursuit, greedy methods, ...

$$\boldsymbol{x} \xrightarrow{\mathbb{R}^N} \boxed{\mathsf{CS}} \xrightarrow[\mathbb{R}^M]{\boldsymbol{y} = \boldsymbol{\Phi}\boldsymbol{x}} \underset{\text{codebook}}{\overset{\boldsymbol{\Omega}}{\boxed{\mathcal{Q}}}} \xrightarrow{\boldsymbol{q} = \mathcal{Q}[\boldsymbol{y}]} \boxed{\mathsf{Decoder}} \xrightarrow{\mathbb{R}^N} \hat{\boldsymbol{x}}$$

Finite codebook $\Rightarrow \hat{\boldsymbol{x}} \neq \boldsymbol{x}$

*i.e.*, impossibility to encode continuous domain in a finite number of elements.

Objective: Minimize $\|\hat{\boldsymbol{x}} - \boldsymbol{x}\|$

given a certain number of:

bits, measurements, or bits/meas.

Where to act?
Change CS, Q or decoder?
Some of them? all?

# Initial Approach for Quantized CS

# Former solution (Candès, Tao, ...)

1. (scalar) Quantization is like a noise

$$\boxed{\boldsymbol{q} \;=\; \mathcal{Q}\big[\boldsymbol{\Phi x}\big] = \boldsymbol{\Phi x} + \boldsymbol{n}}$$

quantization distortion

with $\mathcal{Q}(t) = \delta(\lfloor \frac{t}{\delta} \rfloor + \frac{1}{2})$ (componentwise)

$\longrightarrow$ Bounded:

$$\|\boldsymbol{n}\|_\infty = \|\mathcal{Q}(\boldsymbol{\Phi x}) - \boldsymbol{q}\|_\infty \leq \delta/2$$

# Former solution (Candès, Tao, ...)

1. (scalar) Quantization is like a noise

$$\boxed{\boldsymbol{q} \;=\; \mathcal{Q}\big[\boldsymbol{\Phi x}\big] = \boldsymbol{\Phi x} + \boldsymbol{n}}$$

2. CS is robust (e.g., with *basis pursuit denoise*)

$$\hat{\boldsymbol{x}} = \operatorname*{argmin}_{\boldsymbol{u}\in\mathbb{R}^N} \|\boldsymbol{u}\|_1 \text{ s.t. } \|\boldsymbol{\Phi u} - \boldsymbol{q}\| \leqslant \epsilon \quad \text{(BPDN)}$$

$\ell_2 - \ell_1$ instance optimality:

If $\|\boldsymbol{n}\| \leqslant \epsilon$ and $\frac{1}{\sqrt{M}}\boldsymbol{\Phi}$ is RIP$(\delta, 2K)$ with $\delta \leqslant \sqrt{2} - 1$, then

$$\|\hat{\boldsymbol{x}} - \boldsymbol{x}\| \;\lesssim\; \tfrac{\epsilon}{\sqrt{M}} + e_0(K),$$

with $e_0(K) = \|\boldsymbol{x} - \boldsymbol{x}_K\|_1/\sqrt{K}$.

# Former solution (Candès, Tao, ...)

1. (scalar) Quantization is like a noise

$\|\boldsymbol{\Phi u}-\boldsymbol{q}\|_\infty \le \delta/2$

$$\boxed{\boldsymbol{q} \;=\; \mathcal{Q}\big[\boldsymbol{\Phi x}\big] = \boldsymbol{\Phi x}+\boldsymbol{n}}$$

2. CS is robust (e.g., with *basis pursuit denoise*)

$$\hat{\boldsymbol{x}} = \operatorname*{argmin}_{\boldsymbol{u}\in\mathbb{R}^N} \|\boldsymbol{u}\|_1 \text{ s.t. } \|\boldsymbol{\Phi u}-\boldsymbol{q}\| \leqslant \epsilon \quad \text{(BPDN)}$$

$\ell_2-\ell_1$ instance optimality:

If $\|\boldsymbol{n}\| \leqslant \epsilon$ and $\frac{1}{\sqrt{M}}\boldsymbol{\Phi}$ is RIP$(\delta, 2K)$ with $\delta \leqslant \sqrt{2}-1$, then

$$\|\hat{\boldsymbol{x}}-\boldsymbol{x}\| \;\lesssim\; \delta + e_0(K),$$

with $e_0(K) = \|\boldsymbol{x}-\boldsymbol{x}_K\|_1/\sqrt{K}$.

Deterministic: $\epsilon^2 \le M\delta^2/4$
Stochastic: $\epsilon^2 \le M\delta^2/12 + c\sqrt{M}$ (w.h.p)

# Former solution (Candès, Tao, ...)

- In short:

$$\|\hat{\boldsymbol{x}} - \boldsymbol{x}\| \lesssim \delta + e_0(K),$$

But quantization error doesn't decay with $M$ !?

# Former solution (Candès, Tao, ...)

- In short:

$$\|\hat{\boldsymbol{x}} - \boldsymbol{x}\| \lesssim \delta + e_0(K),$$

But quantization error doesn't decay with $M$ !?

**Solution: *be consistent!***

Enforce $\mathcal{Q}[\boldsymbol{\Phi}\hat{\boldsymbol{x}}] = \mathcal{Q}[\boldsymbol{\Phi}\boldsymbol{x}]$!

"consistency condition"

# Consistent reconstructions in CS?

Issue: if $\hat{\boldsymbol{x}}$ solution of BPDN (adjusted to QCS)

*(i)* No Quantization Consistency (QC) !

$$\|\boldsymbol{\Phi}\hat{\boldsymbol{x}} - \mathcal{Q}[\boldsymbol{\Phi}\boldsymbol{x}]\| \leqslant \epsilon(\delta) \not\Rightarrow \underline{\mathcal{Q}[\boldsymbol{\Phi}\hat{\boldsymbol{x}}] = Q[\boldsymbol{\Phi}\boldsymbol{x}]}$$

(from BPDN constraint) $\Leftrightarrow \|\boldsymbol{\Phi}\hat{\boldsymbol{x}} - \mathcal{Q}[\boldsymbol{\Phi}\boldsymbol{x}]\|_\infty \leq \delta/2$

$\Rightarrow$ Sensing information is not fully exploited!

*(ii)* $\ell_2$ constraint in BPDN

$\approx$ Gaussian distribution (MAP - cond. log. lik.)

# But why looking for consistency?

First: Let $T$ the support of $\boldsymbol{x} \in \mathbb{R}^N$, $\boldsymbol{\Phi} \in \mathbb{R}^{M\times N}$, and $\boldsymbol{y} = \boldsymbol{\Phi}\boldsymbol{x}$.

**Proposition (Goyal, Vetterli, Thao, 98)** *If $T$ is known (with $|T| = K$), the best decoder* $\mathrm{Dec}()$ *provides a* $\hat{\boldsymbol{x}} = \mathrm{Dec}(\boldsymbol{y}, \boldsymbol{\Phi})$ *such that:*

$$\mathrm{RMSE} = (\mathbb{E}\|\boldsymbol{x} - \hat{\boldsymbol{x}}\|^2)^{1/2} \gtrsim (\tfrac{K}{M})\,\delta,$$

*where* $\mathbb{E}$ *is wrt a probability measure on* $\boldsymbol{x}_T$ *in a bounded set* $\mathcal{S} \subset \mathbb{R}^K$.

V. K Goyal, M. Vetterli, N. T. Thao, "Quantized Overcomplete Expansions in $\mathrm{R}^\mathrm{N}$: Analysis, Synthesis, and Algorithms" , IEEE Tran. IT, 44(1), 1998

# But why looking for consistency?

First: Let $T$ the support of $\boldsymbol{x} \in \mathbb{R}^N$, $\boldsymbol{\Phi} \in \mathbb{R}^{M\times N}$, and $\boldsymbol{y} = \boldsymbol{\Phi}\boldsymbol{x}$.

**Proposition (Goyal, Vetterli, Thao, 98)** *If $T$ is known (with $|T| = K$), the best decoder* Dec() *provides a* $\hat{\boldsymbol{x}} = \mathrm{Dec}(\boldsymbol{y}, \boldsymbol{\Phi})$ *such that:*

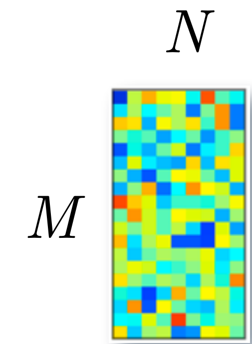
$$\mathrm{RMSE} = (\mathbb{E}\|\boldsymbol{x} - \hat{\boldsymbol{x}}\|^2)^{1/2} \gtrsim \big(\tfrac{K}{M}\big)\,\delta,$$

*where $\mathbb{E}$ is wrt a probability measure on $\boldsymbol{x}_T$ in a bounded set $\mathcal{S} \subset \mathbb{R}^K$.*

Bound achieved for $\boldsymbol{\Phi}_T = \mathrm{DFT} \in \mathbb{R}^{M\times K}$ and Dec() **consistent!**

V. K Goyal, M. Vetterli, N. T. Thao, "Quantized Overcomplete Expansions in $\mathrm{R}^N$: Analysis, Synthesis, and Algorithms" , IEEE Tran. IT, 44(1), 1998

# But why looking for consistency?




Second,

If $\mathbf{\Phi} \in \mathbb{R}^{M \times N}$ is a (random) frame in $\mathbb{R}^N$ $(M \geqslant N)$ and $\boldsymbol{y} = \mathbf{\Phi}\boldsymbol{x}$,

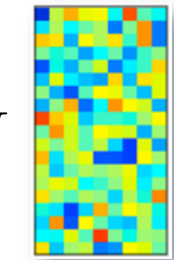
Then, for $\mathcal{Q}(\boldsymbol{y}) = \boldsymbol{y} + \boldsymbol{\xi}$ with $\xi_i \sim \mathcal{U}([-\frac{1}{2}\delta, \frac{1}{2}\delta])$,

and $\hat{\boldsymbol{x}}$ **consistent**, (i.e., $\mathcal{Q}(\mathbf{\Phi}\hat{\boldsymbol{x}}) = \mathcal{Q}(\mathbf{\Phi}\boldsymbol{x})$)

(achievable with *dithering* or under HRA)

This is equivalent to compressed sensing when the support of $\boldsymbol{x}$ is known.

# But why looking for consistency?




Second,

If $\boldsymbol{\Phi} \in \mathbb{R}^{M \times N}$ is a (random) frame in $\mathbb{R}^N$ $(M \geqslant N)$ and $\boldsymbol{y} = \boldsymbol{\Phi x}$,

Then, for $\mathcal{Q}(\boldsymbol{y}) = \boldsymbol{y} + \boldsymbol{\xi}$ with $\xi_i \sim \mathcal{U}([-\frac{1}{2}\delta, \frac{1}{2}\delta])$,

(achievable with *dithering* or under HRA)

and $\hat{\boldsymbol{x}}$ **consistent**,

(i.e., $\mathcal{Q}(\boldsymbol{\Phi}\hat{\boldsymbol{x}}) = \mathcal{Q}(\boldsymbol{\Phi x})$)

$$(\mathbb{E}_{\boldsymbol{\Phi},\boldsymbol{n}}\|\boldsymbol{x} - \hat{\boldsymbol{x}}\|^2)^{1/2} \lesssim (\tfrac{N}{M})\delta,$$

[Powell, Whitehouse, 2013]

(unit norm frame)

$$\|\boldsymbol{x} - \hat{\boldsymbol{x}}\| \lesssim (\tfrac{N}{M})\delta \cdot O(\log M, \log N, \log \eta),$$

[LJ 2014]

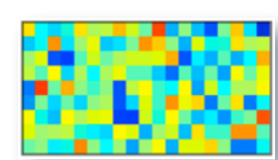
(Gaussian frame)

(uniformly)

with $\Pr \geqslant 1 - \eta$.

or $(\frac{K}{M})$ if $\boldsymbol{x}$ is $K$-sparse with Gaussian sensing matrix.

(with some logarithms)



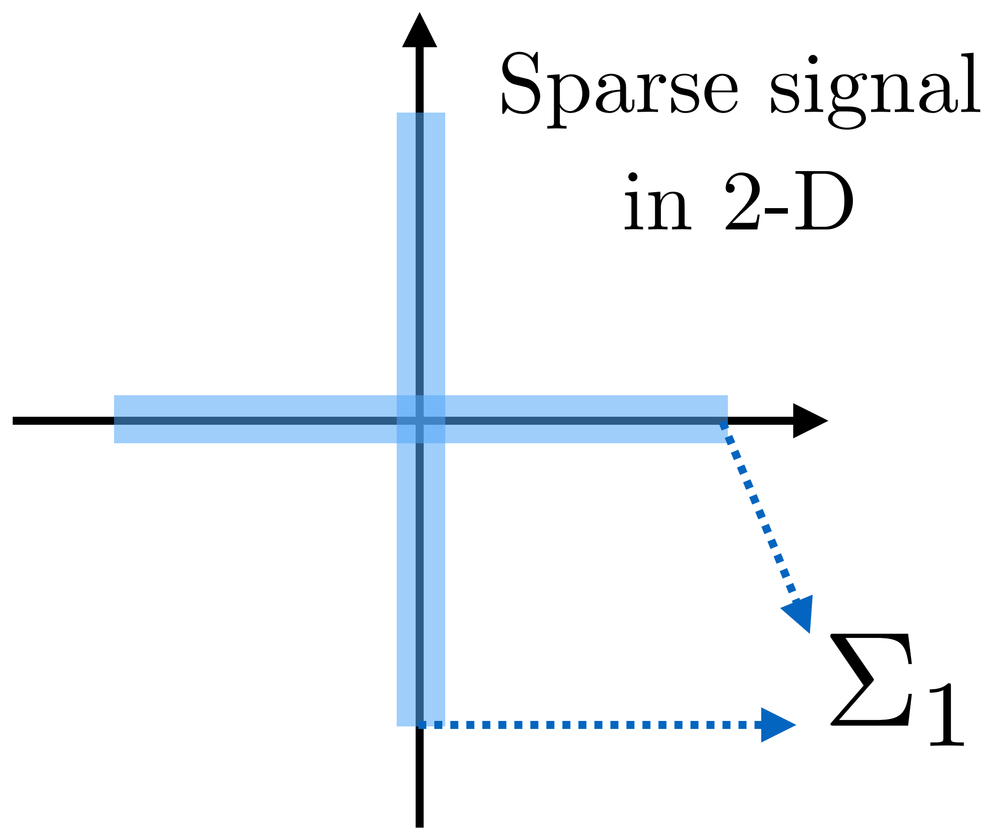
# 2. Consistent Basis Pursuit for low-complexity signals

# Low-complexity signal model

- Low complexity set $\boldsymbol{x}_0 \in \mathcal{K} \subset \mathbb{R}^N$ (more general than sparsity)
- Examples:

$$\mathcal{K} = \Sigma_K := \{\boldsymbol{u} \in \mathbb{R}^N : \|\boldsymbol{u}\|_0 := |\operatorname{supp} \boldsymbol{u}| \leqslant K\}$$

$$\mathcal{K} = \mathcal{C}_r := \{\boldsymbol{U} \in \mathbb{R}^{n \times n} \simeq \mathbb{R}^N : \operatorname{rank}(\boldsymbol{U}) \leqslant r\}$$




Matrix example:

$$\boldsymbol{U} = \boldsymbol{R}^T \boldsymbol{R} \in \mathcal{C}_r$$

with $\boldsymbol{R} \in \mathbb{R}^{r \times n}$, $(r \leq n)$

(example: hyperspectral imaging for linear unmixing)

# Low-complexity signal model

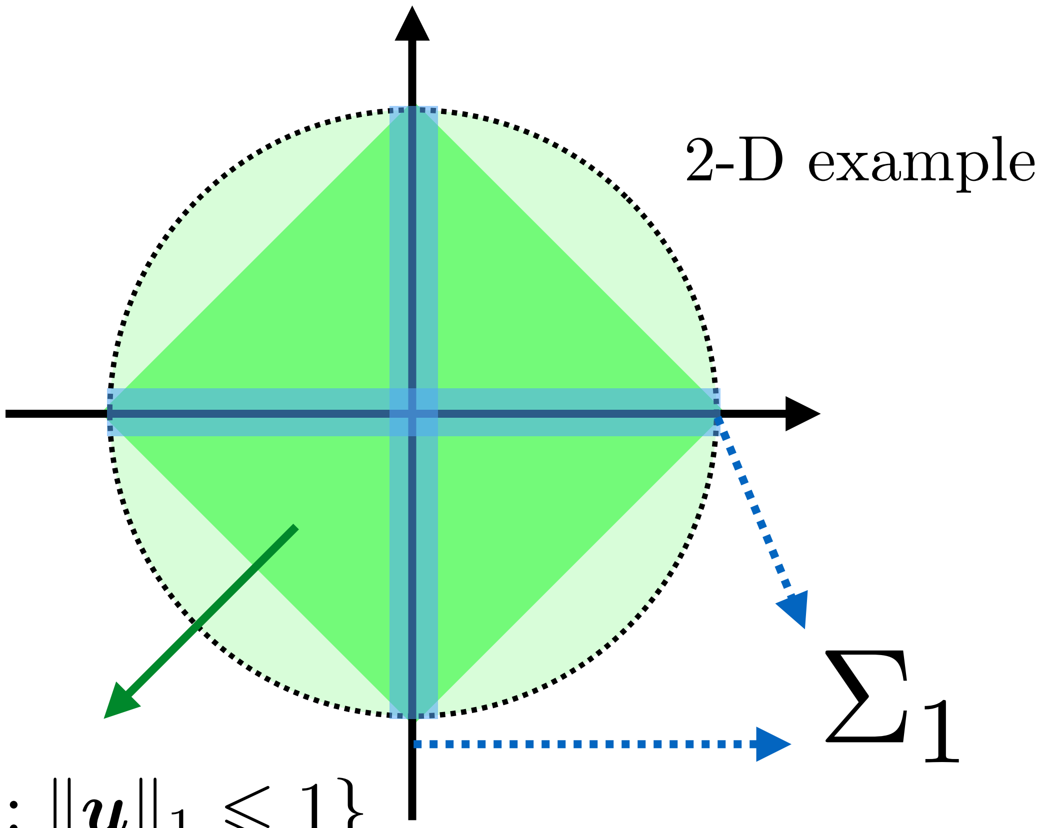
- Low complexity set $\boldsymbol{x}_0 \in \mathcal{K} \subset \mathbb{R}^N$
- Examples:

$$\mathcal{K} = \Sigma_K := \{\boldsymbol{u} \in \mathbb{R}^N : \|\boldsymbol{u}\|_0 := |\operatorname{supp} \boldsymbol{u}| \leqslant K\}$$

$$\mathcal{K} = \mathcal{C}_r := \{\boldsymbol{U} \in \mathbb{R}^{n\times n} \simeq \mathbb{R}^N : \operatorname{rank}(\boldsymbol{U}) \leqslant r\}$$

- Bounded convex hull:

$$\overline{\mathcal{K}} := \operatorname{conv}(\mathcal{K} \cap \mathbb{B}^N)$$




$\mathbb{B}_1^N = \{\boldsymbol{u} : \|\boldsymbol{u}\|_1 \leqslant 1\}$

# Low-complexity signal model

- Low complexity set $\boldsymbol{x}_0 \in \mathcal{K} \subset \mathbb{R}^N$
- Examples:

$$\mathcal{K} = \Sigma_K := \{\boldsymbol{u} \in \mathbb{R}^N : \|\boldsymbol{u}\|_0 := |\operatorname{supp} \boldsymbol{u}| \leqslant K\}$$

$$\mathcal{K} = \mathcal{C}_r := \{\boldsymbol{U} \in \mathbb{R}^{n \times n} \simeq \mathbb{R}^N : \operatorname{rank}(\boldsymbol{U}) \leqslant r\}$$

- Bounded convex hull:

$$\overline{\mathcal{K}} := \operatorname{conv}(\mathcal{K} \cap \mathbb{B}^N)$$

- Atomic norm: $\exists$ convex norm $\|\cdot\|_\sharp$ and a $s > 0$ s.t.

$$\overline{\mathcal{K}} \subset \overline{\mathcal{K}}_s := \{\boldsymbol{u} \in \mathbb{R}^N : \|\boldsymbol{u}\|_\sharp \leqslant s, \|\boldsymbol{u}\|_2 \leqslant 1\}$$

[Chandrasekaran 2012]

# Low-complexity signal model

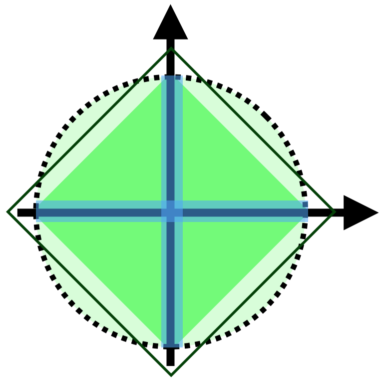
- Low complexity set $\boldsymbol{x}_0 \in \mathcal{K} \subset \mathbb{R}^N$
- Examples:

$$\mathcal{K} = \Sigma_K := \{\boldsymbol{u} \in \mathbb{R}^N : \|\boldsymbol{u}\|_0 := |\operatorname{supp} \boldsymbol{u}| \leqslant K\}$$

$$\mathcal{K} = \mathcal{C}_r := \{\boldsymbol{U} \in \mathbb{R}^{n \times n} \simeq \mathbb{R}^N : \operatorname{rank}(\boldsymbol{U}) \leqslant r\}$$

- Bounded convex hull:

$$\overline{\mathcal{K}} := \operatorname{conv}(\mathcal{K} \cap \mathbb{B}^N)$$



Example: for $\Sigma_K$

$\|\cdot\|_\sharp \leftrightarrow \|\cdot\|_1$

$s = \sqrt{K}$

- Atomic norm: $\exists$ convex norm $\|\cdot\|_\sharp$ and a $s > 0$ s.t.

$$\overline{\mathcal{K}} \subset \overline{\mathcal{K}}_s := \{\boldsymbol{u} \in \mathbb{R}^N : \|\boldsymbol{u}\|_\sharp \leqslant s, \|\boldsymbol{u}\|_2 \leqslant 1\}$$

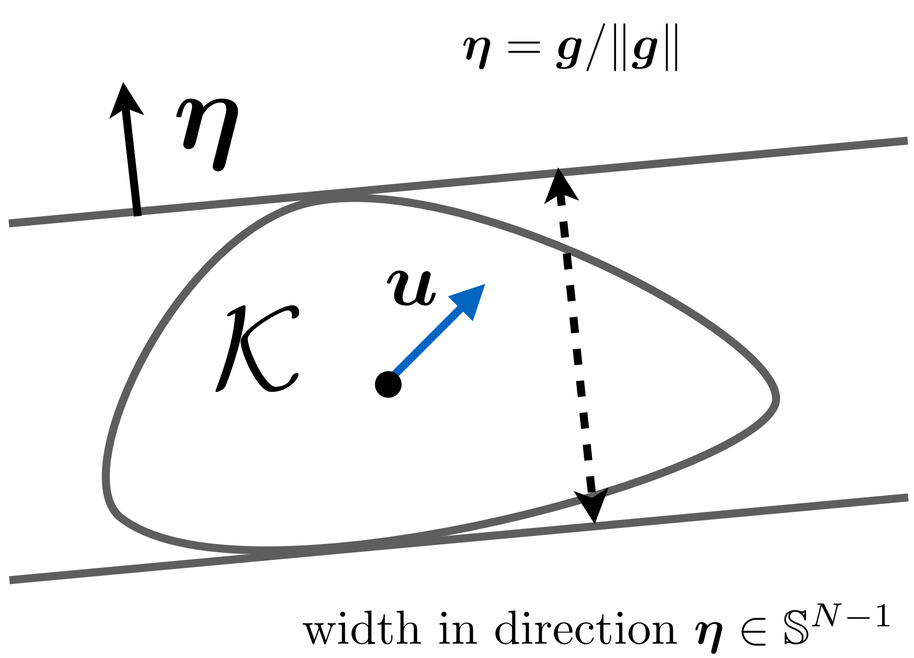
[Chandrasekaran 2012]

Additionally, this contains "*compressible*" signals under the initial low-complexity model!

# Low-complexity signal model

Measuring the "dimension" of $\mathcal{K}$ $\rightarrow$ Gaussian mean width:

$$w(\mathcal{K}) := \mathbb{E} \sup_{\boldsymbol{u} \in \mathcal{K}} |\langle \boldsymbol{g}, \boldsymbol{u} \rangle|, \ \text{ with } g_k \sim_{\text{iid}} \mathcal{N}(0,1)$$



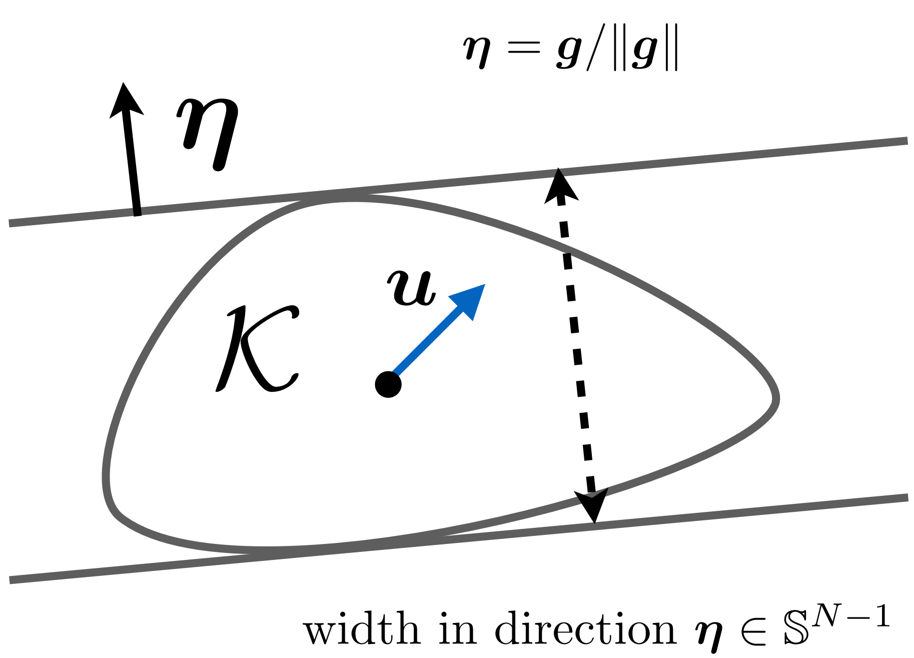

with $w(\mathcal{K}) \leqslant w(\mathcal{K}')$ if $\mathcal{K} \subset \mathcal{K}'$

[Plan, Vershynin, Chandrasekaran, ...]

# Low-complexity signal model

Measuring the "dimension" of $\mathcal{K}$ → Gaussian mean width:

$$w(\mathcal{K}) := \mathbb{E} \sup_{\boldsymbol{u} \in \mathcal{K}} |\langle \boldsymbol{g}, \boldsymbol{u} \rangle|, \ \text{with } g_k \sim_{\rm iid} \mathcal{N}(0,1)$$




with $w(\mathcal{K}) \leqslant w(\mathcal{K}')$ if $\mathcal{K} \subset \mathcal{K}'$

Examples:

$w^2(\mathcal{S}^{N-1}) \leqslant 4N$

$w^2(\mathcal{K}) \leqslant C \log |\mathcal{K}|$ (for finite sets)

$w^2(\mathcal{K} \cap \mathbb{B}^N) \leqslant L$ if subspace with $\dim \mathcal{K} = L$

$w^2(\Sigma_K \cap \mathbb{B}^N) \leqslant w^2(\overline{\Sigma}_K) \simeq K \log(2N/K)$

$w^2(\mathcal{C}_r) \leqslant w^2(\overline{\mathcal{C}}_r) \leqslant 4nr$

[Plan, Vershynin, Chandrasekaran, ...]

# QCS model

- As before: for Gaussian or sub-Gaussian $\mathbf{\Phi} \in \mathbb{R}^{M \times N}$

Uniform (bin $\delta > 0$) + dithering ($\xi_i \sim_{\text{iid}} \mathcal{U}([0, \delta])$)

$$\boxed{\boldsymbol{q} = \boldsymbol{A}(\boldsymbol{x}_0) := \mathcal{Q}(\mathbf{\Phi}\boldsymbol{x}_0 + \boldsymbol{\xi})}$$

# QCS model

- As before: for Gaussian or sub-Gaussian $\mathbf{\Phi} \in \mathbb{R}^{M \times N}$

  Uniform (bin $\delta > 0$) + dithering ($\xi_i \sim_{\rm iid} \mathcal{U}([0, \delta])$)

$$\boxed{\boldsymbol{q} = \boldsymbol{A}(\boldsymbol{x}_0) := \mathcal{Q}(\mathbf{\Phi}\boldsymbol{x}_0 + \boldsymbol{\xi})}$$

- Questions: knowing that $\boldsymbol{x}_0 \in \mathcal{K} \subset \mathbb{R}^N$,
  1. Theoretical bound on reconstruction error?

     $\Leftrightarrow$ proximity of consistent vectors

  2. Reconstruction algorithm?

     $\Leftrightarrow$ one solution: CoBP

# Proximity of consistent vectors [LJ 2015]

- Gaussian case: $\Phi_{ij} \sim_{\rm iid} \mathcal{N}(0,1)$

In all generality, provided

$$\boxed{M \gtrsim \frac{(1+\delta)^4}{\delta^2 \underline{\epsilon^4}} \underline{w^2(\mathcal{K})}}$$

Then, with $\Pr \geqslant 1 - 2\exp(-\frac{c\epsilon M}{1+\delta})$ and uniformly for all $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{K}$,

Proximity condition: $\boldsymbol{A}(\boldsymbol{x}_1) = \boldsymbol{A}(\boldsymbol{x}_2) \quad \Rightarrow \quad \|\boldsymbol{x}_1 - \boldsymbol{x}_2\| \leqslant \epsilon$

$\Rightarrow \epsilon = O(M^{-1/4})$ for consistent reconstruction $\in \mathcal{K}$!

# Proximity of consistent vectors [LJ 2015]

- Gaussian case: $\Phi_{ij} \sim_{\rm iid} \mathcal{N}(0,1)$

In all generality, provided

$$\boxed{M \gtrsim \frac{(1+\delta)^4}{\delta^2 \epsilon^4} w^2(\mathcal{K})}$$

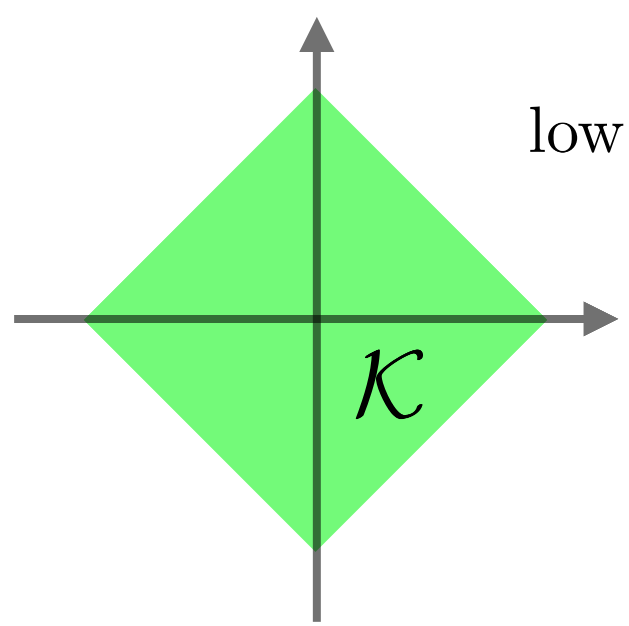
Then, with $\Pr \geqslant 1 - 2\exp(-\frac{c\epsilon M}{1+\delta})$ and uniformly for all $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{K}$,

**Proximity condition:** $\boldsymbol{A}(\boldsymbol{x}_1) = \boldsymbol{A}(\boldsymbol{x}_2) \quad \Rightarrow \quad \|\boldsymbol{x}_1 - \boldsymbol{x}_2\| \leqslant \epsilon$

$\Rightarrow \epsilon = O(M^{-1/4})$ for consistent reconstruction $\in \mathcal{K}$!

Interpretation:

low complexity set $\mathcal{K}$



# Proximity of consistent vectors [LJ 2015]

- Gaussian case: $\Phi_{ij} \sim_{\rm iid} \mathcal{N}(0,1)$

In all generality, provided

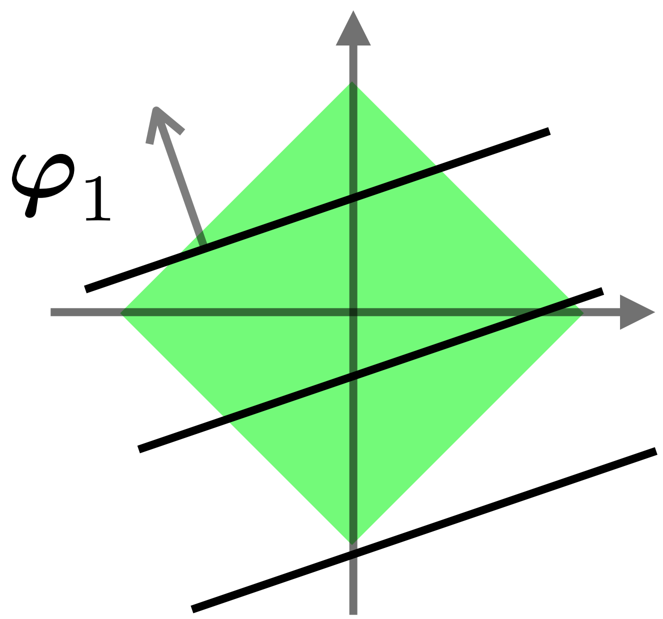
$$\boxed{M \gtrsim \frac{(1+\delta)^4}{\delta^2 \epsilon^4} w^2(\mathcal{K})}$$

Then, with $\Pr \geqslant 1 - 2\exp(-\frac{c\epsilon M}{1+\delta})$ and uniformly for all $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{K}$,

Proximity condition: $\boldsymbol{A}(\boldsymbol{x}_1) = \boldsymbol{A}(\boldsymbol{x}_2) \quad \Rightarrow \quad \|\boldsymbol{x}_1 - \boldsymbol{x}_2\| \leqslant \epsilon$

$\Rightarrow \epsilon = O(M^{-1/4})$ for consistent reconstruction $\in \mathcal{K}$!

Interpretation:



# Proximity of consistent vectors [LJ 2015]

- Gaussian case: $\Phi_{ij} \sim_{\text{iid}} \mathcal{N}(0,1)$

In all generality, provided

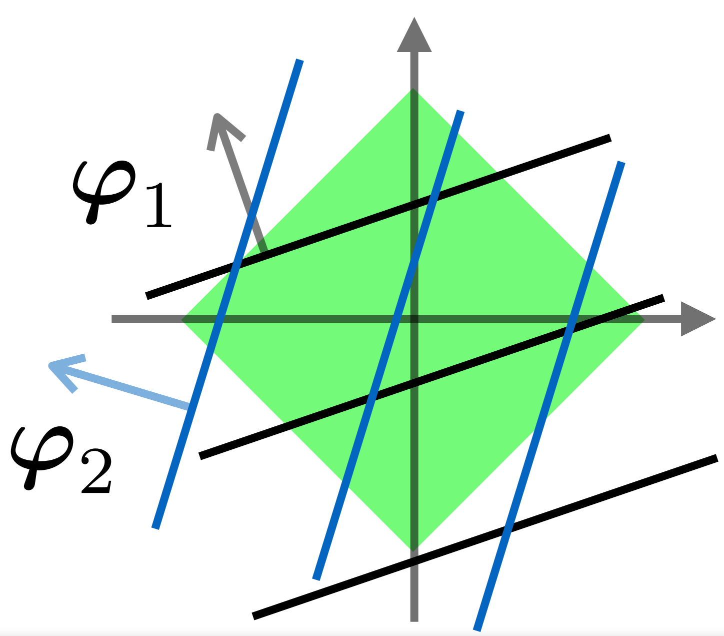
$$\boxed{M \gtrsim \frac{(1+\delta)^4}{\delta^2 \epsilon^4} w^2(\mathcal{K})}$$

Then, with $\Pr \geqslant 1 - 2\exp(-\frac{c\epsilon M}{1+\delta})$ and uniformly for all $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{K}$,

Proximity condition: $\quad \boldsymbol{A}(\boldsymbol{x}_1) = \boldsymbol{A}(\boldsymbol{x}_2) \quad \Rightarrow \quad \|\boldsymbol{x}_1 - \boldsymbol{x}_2\| \leqslant \epsilon$

$\Rightarrow \epsilon = O(M^{-1/4})$ for consistent reconstruction $\in \mathcal{K}$!

Interpretation:



# Proximity of consistent vectors [LJ 2015]

- Gaussian case: $\Phi_{ij} \sim_{\rm iid} \mathcal{N}(0,1)$

In all generality, provided

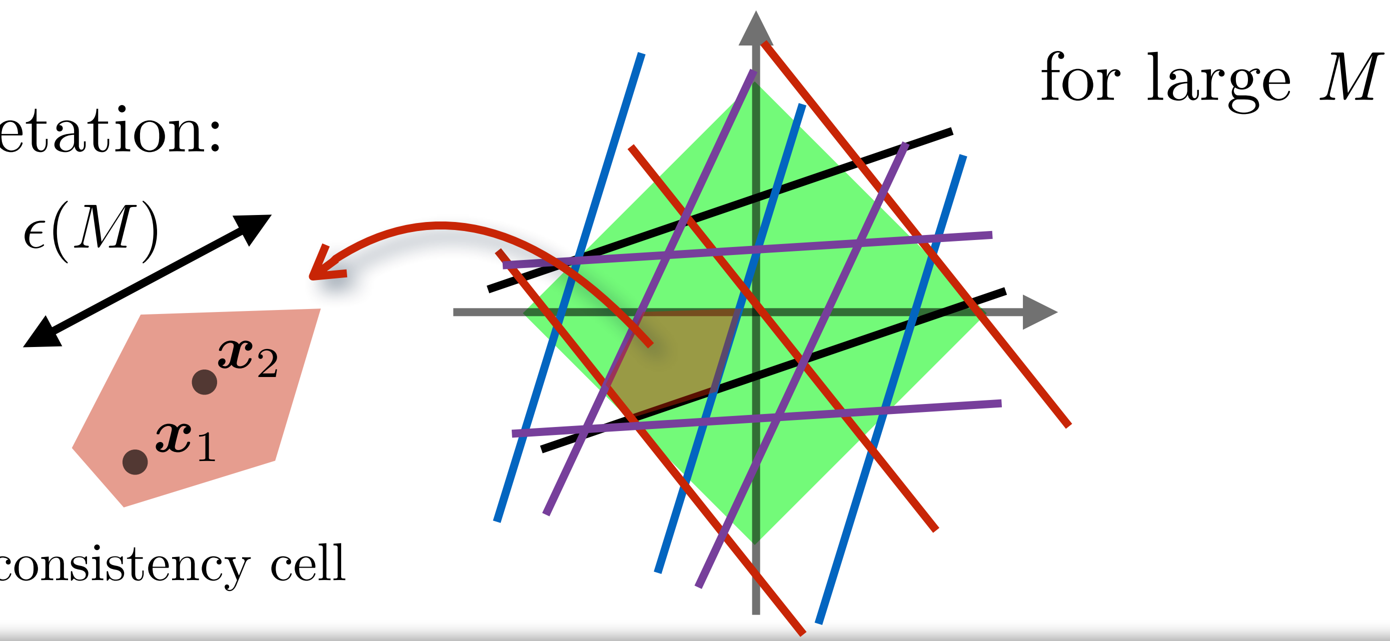
$$\boxed{M \gtrsim \frac{(1+\delta)^4}{\delta^2 \epsilon^4} w^2(\mathcal{K})}$$

Then, with $\Pr \geqslant 1 - 2\exp(-\frac{c\epsilon M}{1+\delta})$ and uniformly for all $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{K}$,

Proximity condition: $\boldsymbol{A}(\boldsymbol{x}_1) = \boldsymbol{A}(\boldsymbol{x}_2) \quad \Rightarrow \quad \|\boldsymbol{x}_1 - \boldsymbol{x}_2\| \leqslant \epsilon$

$\Rightarrow \epsilon = O(M^{-1/4})$ for consistent reconstruction $\in \mathcal{K}$!

Interpretation:



# Proximity of consistent vectors [LJ 2015]

- <u>Gaussian case</u>: $\Phi_{ij} \sim_{\text{iid}} \mathcal{N}(0,1)$

In all generality, provided

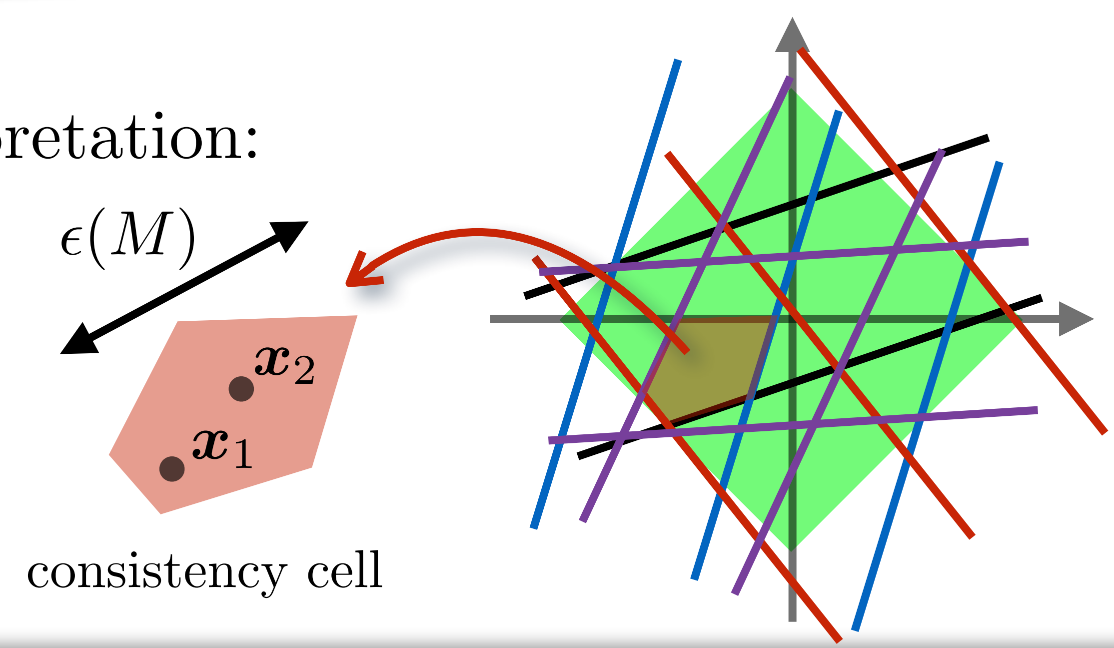
$$\boxed{M \gtrsim \frac{(1+\delta)^4}{\delta^2 \underline{\epsilon^4}} \underline{w^2(\mathcal{K})}}$$

Then, with $\Pr \geqslant 1 - 2\exp(-\frac{c\epsilon M}{1+\delta})$ and uniformly for all $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{K}$,

**<u>Proximity condition:</u>** $\boldsymbol{A}(\boldsymbol{x}_1) = \boldsymbol{A}(\boldsymbol{x}_2) \quad \Rightarrow \quad \|\boldsymbol{x}_1 - \boldsymbol{x}_2\| \leqslant \epsilon$

$\Rightarrow \epsilon = O(M^{-1/4})$ for consistent reconstruction $\in \mathcal{K}$!

Interpretation:




and don't forget

$\ell_1$-ball in high dimension

# Proximity of consistent vectors [LJ 2015]

- Gaussian case: $\Phi_{ij} \sim_{\rm iid} \mathcal{N}(0,1)$

In all generality, provided

$$\boxed{M \gtrsim \frac{(1+\delta)^4}{\delta^2 \epsilon^4} w^2(\mathcal{K})}$$

Then, with $\Pr \geqslant 1 - 2\exp(-\frac{c\epsilon M}{1+\delta})$ and uniformly for all $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{K}$,

Proximity condition: $\boldsymbol{A}(\boldsymbol{x}_1) = \boldsymbol{A}(\boldsymbol{x}_2) \quad \Rightarrow \quad \|\boldsymbol{x}_1 - \boldsymbol{x}_2\| \leqslant \epsilon$

$\Rightarrow \epsilon = O(M^{-1/4})$ for consistent reconstruction $\in \mathcal{K}$!

Special case:

For $\mathcal{K} = (\boldsymbol{\Psi}\Sigma_K) \cap \mathbb{B}^N$ and $\boldsymbol{\Psi}$ ONB

$$M \gtrsim \frac{2+\delta}{\epsilon} K \log\Big(\frac{N(2+\delta)^{3/2}}{K\delta\epsilon^{3/2}}\Big)$$

$$\Rightarrow \epsilon = O(M^{-1})!$$

# Proximity of consistent vectors [LJ 2015]

- <u>Sub-Gaussian case</u>: (e.g., iid Bernoulli or bounded $\Phi_{ij}$)

  Still ok but, for $K_0 \gtrsim \kappa(\boldsymbol{\Phi})$ and for all $\boldsymbol{x}_1, \boldsymbol{x}_2$ in

$$\reflectbox{$\Sigma$}_{K_0} := \{\boldsymbol{u} \in \mathbb{R}^N : K_0 \|\boldsymbol{u}\|_\infty^2 \leqslant \|\boldsymbol{u}\|_2^2\}.$$

  "*Amgis*" ;-)

  Then, with $\Pr \geqslant 1 - 2\exp(-\frac{c\epsilon M}{1+\delta})$, we have also

$$\boldsymbol{A}(\boldsymbol{x}_1) = \boldsymbol{A}(\boldsymbol{x}_2) \quad \Rightarrow \quad \|\boldsymbol{x}_1 - \boldsymbol{x}_2\| \leqslant \epsilon$$

# Proximity of consistent vectors [LJ 2015]

- <u>Sub-Gaussian case</u>: (e.g., iid Bernoulli or bounded $\Phi_{ij}$)

  Still ok but, for $K_0 \gtrsim \kappa(\mathbf{\Phi})$ and for all $\boldsymbol{x}_1, \boldsymbol{x}_2$ in

  $$\Game_{K_0} := \{\boldsymbol{u} \in \mathbb{R}^N : K_0 \|\boldsymbol{u}\|_\infty^2 \leqslant \|\boldsymbol{u}\|_2^2\}.$$

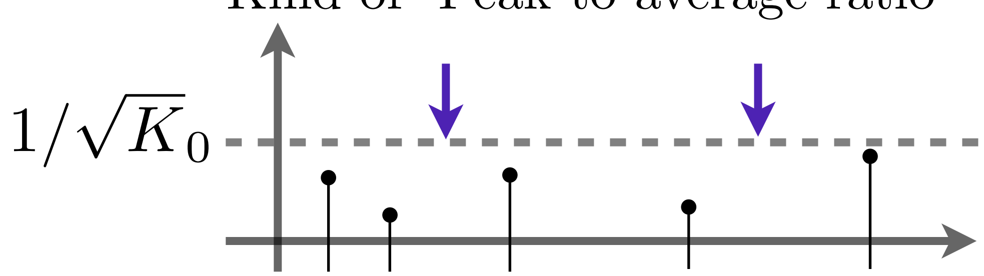
  "*Amgis*" ;-)

  Then, with $\Pr \geqslant 1 - 2\exp(-\frac{c\epsilon M}{1+\delta})$, we have also

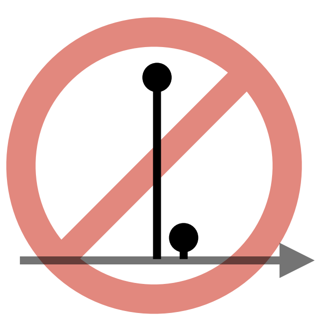
  $$\boldsymbol{A}(\boldsymbol{x}_1) = \boldsymbol{A}(\boldsymbol{x}_2) \quad \Rightarrow \quad \|\boldsymbol{x}_1 - \boldsymbol{x}_2\| \leqslant \epsilon$$

  $\Rightarrow$ Ok for "not too sparse" signals

Kind of "Peak to average ratio"




s.t. $\|\boldsymbol{u}\| = 1$



(*<u>remark</u>*: similar and earlier observations in 1-bit CS by Plan & Vershynin)

# Consistent Basis Pursuit (CoBP)

How to reconstruct
our low complexity signal?

# Consistent Basis Pursuit (CoBP)

- Gaussian case:

Actually, not a so new program, see e.g.,
[Milenkovitch, Dai, JL, Hammond, Fadili, ...]

$$\boldsymbol{x}^* := \operatorname*{argmin}_{\boldsymbol{u}\in\mathbb{R}^N} \|\boldsymbol{u}\|_{\underline{\sharp}} \ \text{s.t.}\ \boldsymbol{A}(\boldsymbol{u}) = \boldsymbol{A}(\boldsymbol{x}_0),\ \underline{\boldsymbol{u} \in \mathbb{B}^N}.$$

except for this

# Consistent Basis Pursuit (CoBP)

- <u>Gaussian case:</u>

$$\boldsymbol{x}^* := \operatorname*{argmin}_{\boldsymbol{u}\in\mathbb{R}^N} \|\boldsymbol{u}\|_\sharp \;\text{ s.t. }\; \boldsymbol{A}(\boldsymbol{u}) = \boldsymbol{A}(\boldsymbol{x}_0),\; \boldsymbol{u} \in \mathbb{B}^N.$$

If $\boldsymbol{A}$ respects "proximity condition" for all $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \overline{\mathcal{K}}_s \supset \mathcal{K}$, then, we have for all $\boldsymbol{x}_0 \in \overline{\mathcal{K}}_s$,

$$\|\boldsymbol{x}_0 - \boldsymbol{x}^*\|_2 \leqslant \epsilon.$$

# Consistent Basis Pursuit (CoBP)

- Gaussian case:

$$\boldsymbol{x}^* := \underset{\boldsymbol{u}\in\mathbb{R}^N}{\operatorname{argmin}} \|\boldsymbol{u}\|_\sharp \ \text{ s.t. } \ \boldsymbol{A}(\boldsymbol{u}) = \boldsymbol{A}(\boldsymbol{x}_0), \ \boldsymbol{u} \in \mathbb{B}^N.$$

If $\boldsymbol{A}$ respects "proximity condition" for all $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \overline{\mathcal{K}}_s \supset \mathcal{K}$, then, we have for all $\boldsymbol{x}_0 \in \overline{\mathcal{K}}_s$,

$$\|\boldsymbol{x}_0 - \boldsymbol{x}^*\|_2 \leqslant \epsilon.$$

$\Pr \geqslant 1 - 2\exp(-cM^{3/4}/\sqrt{\delta})$

**Corollary**: With high probability on Gaussian $\boldsymbol{\Phi}$ and uniform $\boldsymbol{\xi}$, $\forall \boldsymbol{x}_0 \in \overline{\mathcal{K}}_s$,

$$\|\boldsymbol{x}_0 - \boldsymbol{x}^*\|_2 = O\big(\tfrac{2+\delta}{\sqrt{\delta}}\,\big(\tfrac{w(\overline{\mathcal{K}}_s)^2}{M}\big)^{1/4}\big),$$

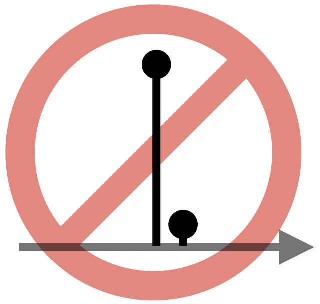
*i.e.*, $\|\boldsymbol{x}_0 - \boldsymbol{x}^*\|_2 = O\big(M^{-1/4}\big)$ if only $M$ varies.

# Consistent Basis Pursuit (CoBP)

- <u>sub-Gaussian case:</u>

$$\boldsymbol{x}^* := \operatorname*{argmin}_{\boldsymbol{u}\in\mathbb{R}^N} \|\boldsymbol{u}\|_\sharp \ \text{ s.t. } \begin{cases} \boldsymbol{A}(\boldsymbol{u}) = \boldsymbol{A}(\boldsymbol{x}_0), \\ \boldsymbol{u} \in \underline{\mathbb{B}^N \cap \lambda\mathbb{B}^N_\infty}, \end{cases}$$

to take care of signal "peaky-ness"



# Consistent Basis Pursuit (CoBP)

- sub-Gaussian case:

$$\boldsymbol{x}^* := \operatorname*{argmin}_{\boldsymbol{u}\in\mathbb{R}^N} \|\boldsymbol{u}\|_\sharp \ \text{ s.t. } \begin{cases} \boldsymbol{A}(\boldsymbol{u}) = \boldsymbol{A}(\boldsymbol{x}_0), \\ \boldsymbol{u} \in \mathbb{B}^N \cap \lambda\mathbb{B}^N_\infty, \end{cases}$$

**Proposition**: With high probability on sub-Gaussian $\boldsymbol{\Phi}$ and uniform $\boldsymbol{\xi}$, $\forall \boldsymbol{x}_0 \in \overline{\mathcal{K}}_s \cap \lambda\mathbb{B}^N_\infty$,

$$\|\boldsymbol{x}_0 - \boldsymbol{x}^*\|_2 = O\big(\tfrac{2+\delta}{\sqrt{\delta}}\,\big(\tfrac{w(\overline{\mathcal{K}}_s)^2}{M}\big)^{1/4} + \underline{\kappa(\boldsymbol{\Phi})\lambda}\big),$$

*i.e.*, $\|\boldsymbol{x}_0 - \boldsymbol{x}^*\|_2 = O\big(M^{-1/4} \underline{+ \lambda}\big)$ if only $M$ varies.

(possible) price to pay for sub-Gaussianity

# Consistent Basis Pursuit (CoBP)

- sub-Gaussian case:

$$\boldsymbol{x}^* := \operatorname*{argmin}_{\boldsymbol{u}\in\mathbb{R}^N} \|\boldsymbol{u}\|_\sharp \ \text{ s.t. } \begin{cases} \boldsymbol{A}(\boldsymbol{u}) = \boldsymbol{A}(\boldsymbol{x}_0), \\ \boldsymbol{u} \in \mathbb{B}^N \cap \lambda\mathbb{B}^N_\infty, \end{cases}$$

**Proposition**: With high probability on sub-Gaussian $\boldsymbol{\Phi}$ and uniform $\boldsymbol{\xi}$, $\forall \boldsymbol{x}_0 \in \overline{\mathcal{K}}_s \cap \lambda\mathbb{B}^N_\infty$,

$$\|\boldsymbol{x}_0 - \boldsymbol{x}^*\|_2 = O\big(\tfrac{2+\delta}{\sqrt{\delta}}\,\big(\tfrac{w(\overline{\mathcal{K}}_s)^2}{M}\big)^{1/4} + \underline{\kappa(\boldsymbol{\Phi})\lambda}\big),$$

*i.e.*, $\|\boldsymbol{x}_0 - \boldsymbol{x}^*\|_2 = O\big(M^{-1/4} \underline{+\,\lambda}\big)$ if only $M$ varies.

Is it bad? If you think so, then sample signals with

$$\boldsymbol{\Phi} \to \boldsymbol{\Phi}\boldsymbol{F}$$

with, *e.g.*, $\boldsymbol{F}$ = Fourier or Hadamard.

→ **Kind of "Spread the samples" trick ;-)**

# Experiments: implementation

- Solving CoBP: Convex Optimization

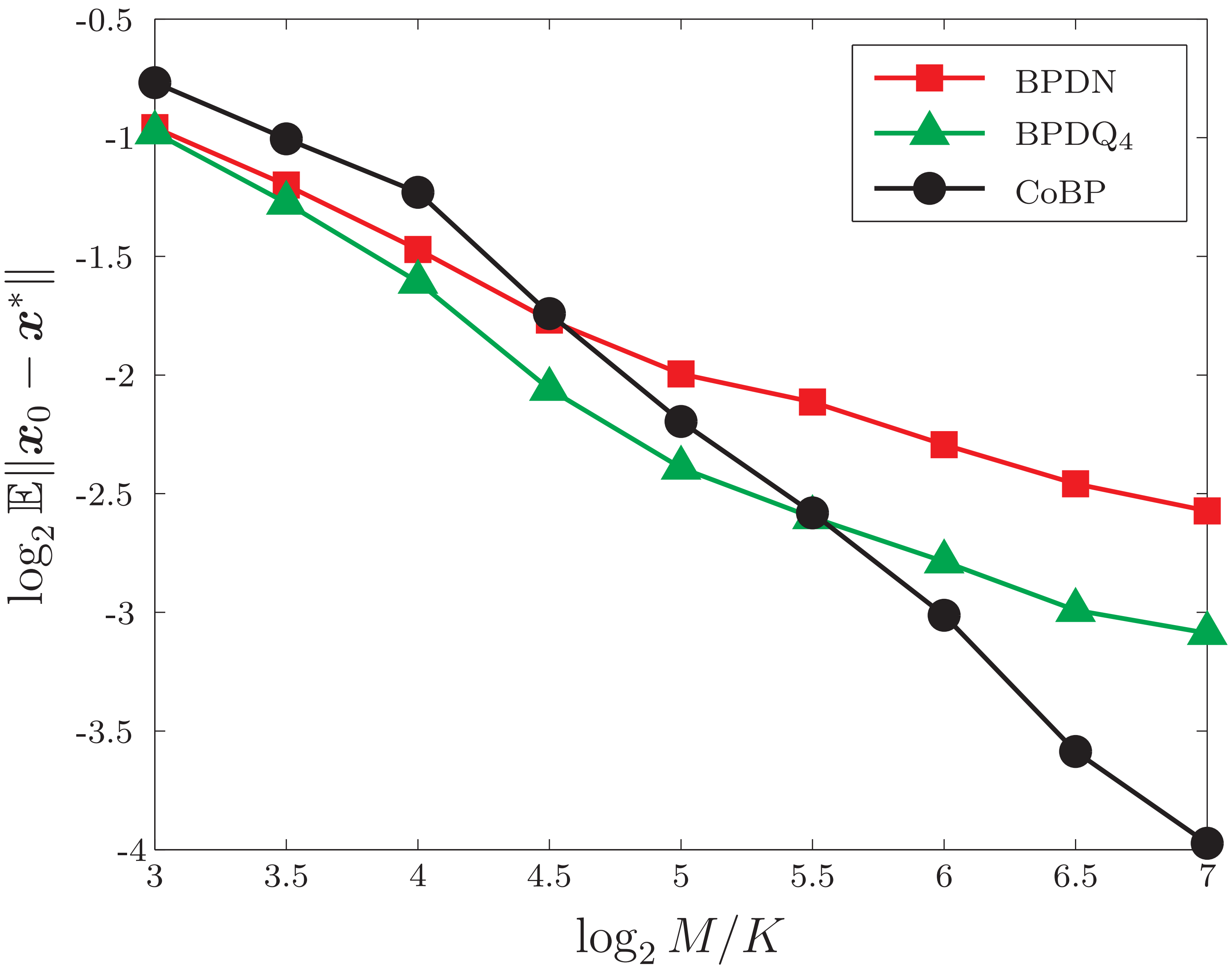
$$
\begin{aligned}
\boldsymbol{x}^* &:= \operatorname*{argmin}_{\boldsymbol{u}\in\mathbb{R}^N} \|\boldsymbol{u}\|_\sharp \ \text{ s.t. } \ \boldsymbol{A}(\boldsymbol{u}) = \boldsymbol{A}(\boldsymbol{x}_0), \ \boldsymbol{u} \in \mathbb{B}^N. \\
\Leftrightarrow &\quad \operatorname*{argmin}_{\boldsymbol{u}\in\mathbb{R}^N} \ \|\boldsymbol{u}\|_\sharp \ + \ \imath_{\text{consist.}}(\boldsymbol{u}) \ + \ \imath_{\mathbb{B}^N}(\boldsymbol{u})
\end{aligned}
$$

  - Many toolboxes available
  - We used a proximal algorithm, i.e., Parallel Proximal Algorithm (PPXA)
  - & the UNLocBoX toolbox (https://lts2.epfl.ch/unlocbox)

# Experiments: 1

$K$-sparse signals with Gaussian sensing

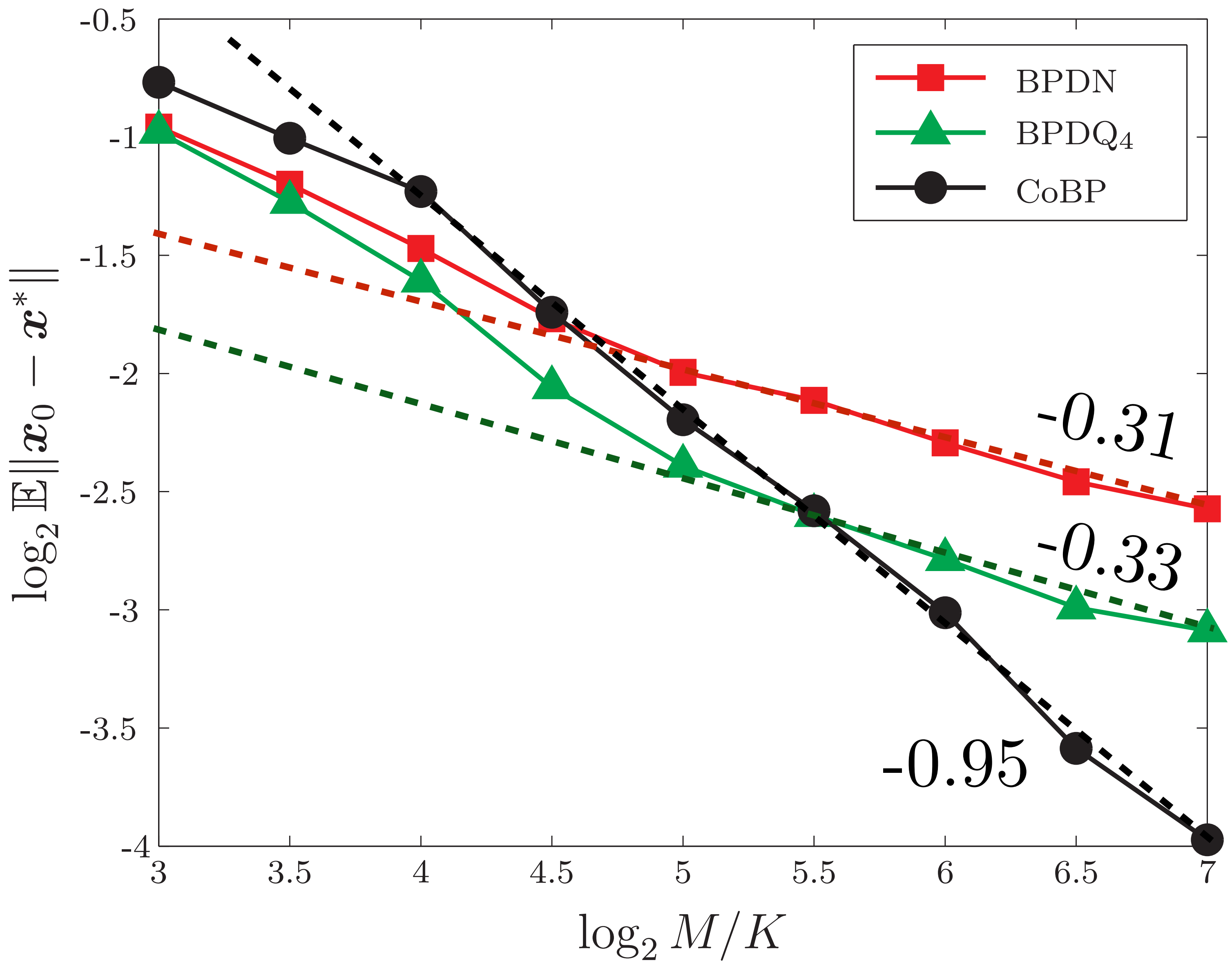
$N = 2048,\ K = 16,$
$B = 3,\ M/K \in [8, 128]$
20 trials per points



# Experiments: 1

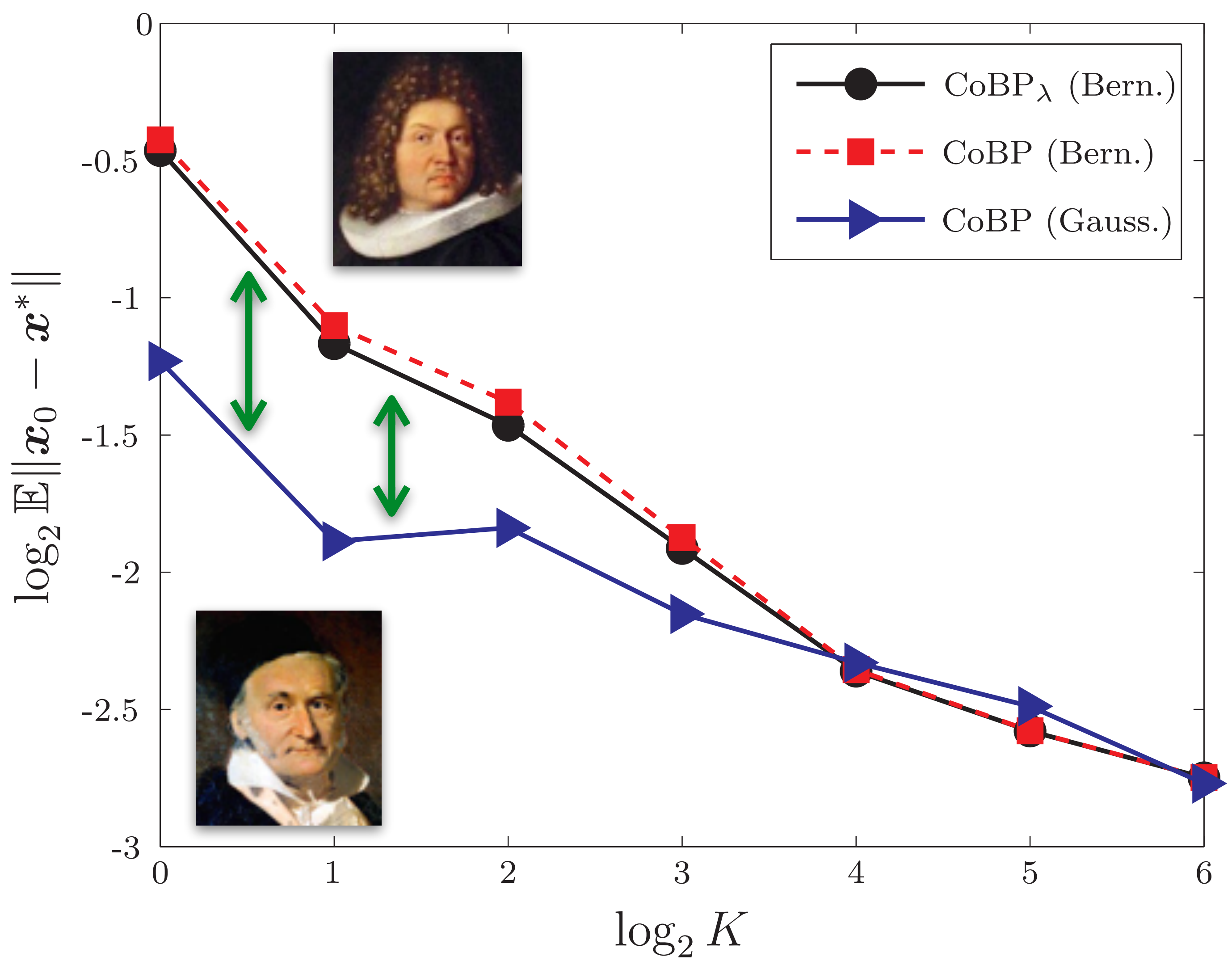
*K*-sparse signals with Gaussian sensing

$N = 2048,\ K = 16,$
$B = 3,\ M/K \in [8, 128]$
20 trials per points



# Experiments: 2

- Bernoulli vs Gauss(ian) sensing:




$N = 1024$, $K \in [1, 64]$,
$B = 4$, $M/K = 16$
20 trials per points

# Experiments: 3

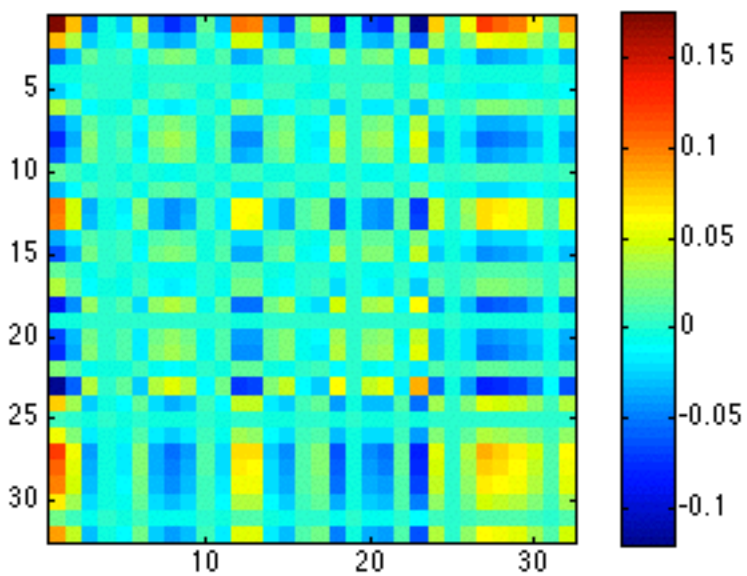
- Low-rank matrix and QCS (with $\boldsymbol{A}(\cdot) := \mathcal{Q}(\boldsymbol{\Phi}\,\cdot)$)

$$\boldsymbol{X}^* := \operatorname*{argmin}_{\boldsymbol{U}\in\mathbb{R}^{n\times n}} \|\boldsymbol{U}\|_* \ \text{ s.t. } \ \boldsymbol{A}(\mathrm{vec}(\boldsymbol{U})) = \boldsymbol{A}(\mathrm{vec}(\boldsymbol{X}_0)), \ \mathrm{vec}(\boldsymbol{U}) \in \mathbb{B}^N.$$

# Experiments: 3

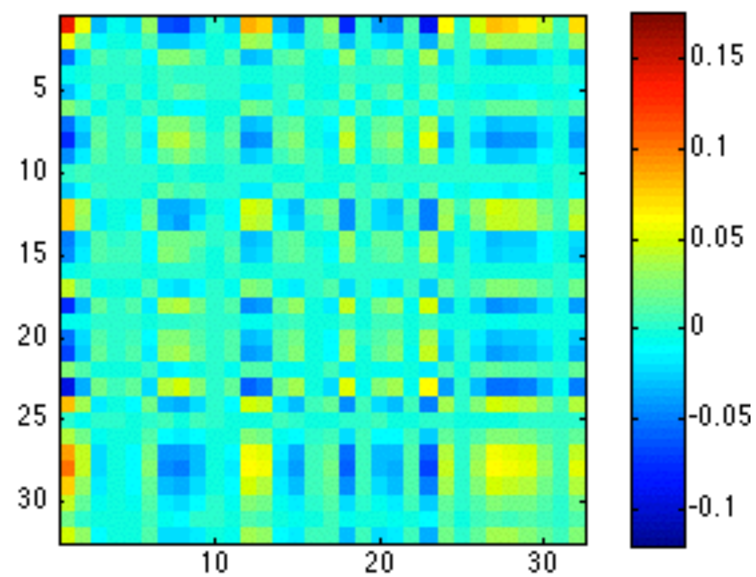
- Low-rank matrix and QCS (with $\boldsymbol{A}(\cdot) := \mathcal{Q}(\boldsymbol{\Phi}\,\cdot)$)

$$\boldsymbol{X}^* := \operatorname*{argmin}_{\boldsymbol{U}\in\mathbb{R}^{n\times n}} \|\boldsymbol{U}\|_* \ \text{ s.t. } \ \boldsymbol{A}(\mathrm{vec}(\boldsymbol{U})) = \boldsymbol{A}(\mathrm{vec}(\boldsymbol{X}_0)),\ \mathrm{vec}(\boldsymbol{U}) \in \mathbb{B}^N.$$



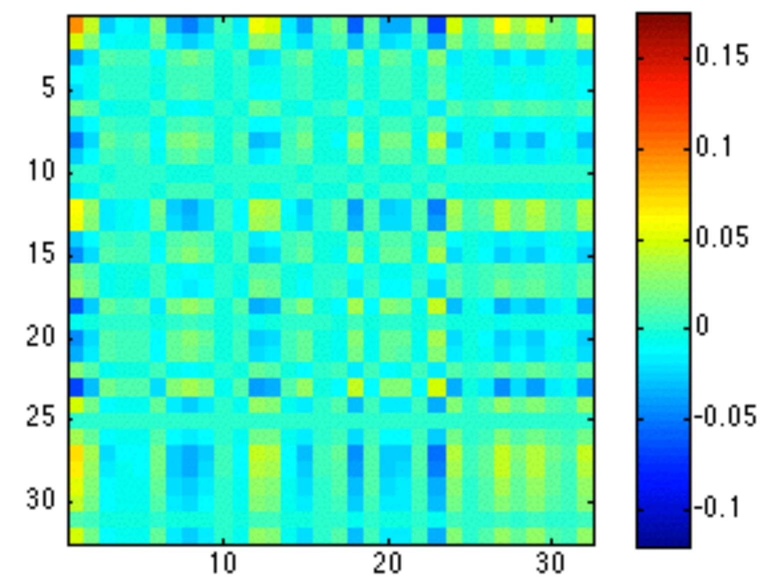

Original

$N = 1024 = n^2$ $(n = 32)$,
rank $r = 1$, $\boldsymbol{B = 2}$
Complexity $< P = 64$,
$M = 16P = N$




CoBP

SNR 11 dB




BPDN

SNR 6.9 dB

# 3. Quasi-isometric embeddings

## (or “how to quantize the RIP”)

# Quantizing the RIP?

Restricted isometry Property (RIP): (as an embedding preserving distances)

$$\boxed{(1-\rho)\|\boldsymbol{u}-\boldsymbol{v}\|^2 \leq \tfrac{1}{M}\|\boldsymbol{\Phi}\boldsymbol{u}-\boldsymbol{\Phi}\boldsymbol{v}\|^2 \leq (1+\rho)\|\boldsymbol{u}-\boldsymbol{v}\|^2}$$

Inserting quantization?

for all $\boldsymbol{u},\boldsymbol{v} \in \Sigma_K := \{\boldsymbol{u} : \|\boldsymbol{u}\|_0 := |\mathrm{supp}\,\boldsymbol{u}| \leq K\}$

# Quantizing the RIP?

Restricted isometry Property (RIP):

(as an embedding preserving distances)

$$\boxed{(1-\rho)\|\boldsymbol{u}-\boldsymbol{v}\|^2 \leq \tfrac{1}{M}\|\boldsymbol{\Phi u}-\boldsymbol{\Phi v}\|^2 \leq (1+\rho)\|\boldsymbol{u}-\boldsymbol{v}\|^2}$$

Inserting quantization?

for all $\boldsymbol{u},\boldsymbol{v} \in \Sigma_K := \{\boldsymbol{u} : \|\boldsymbol{u}\|_0 := |\operatorname{supp}\boldsymbol{u}| \leq K\}$

Why quantizing the RIP?

- since we can ;-)
- for future algorithm guarantees
- for nearest neighbors applications
- or "signal processing" in quantized CS domain

# Quantizing the RIP?

- Let's retake: for $\mathcal{Q}(\cdot) = \delta\lfloor \cdot/\delta \rfloor \in \delta\mathbb{Z}$

$$\boldsymbol{\psi}(\boldsymbol{u}) := \mathcal{Q}(\boldsymbol{\Phi}\boldsymbol{u} + \boldsymbol{\xi}), \text{ with } \boldsymbol{\psi}_j(\boldsymbol{u}) := \mathcal{Q}(\boldsymbol{\varphi}_j^T\boldsymbol{u} + \xi_j)$$

with $\boldsymbol{\Phi}_{ji} \sim_{\text{iid}} \mathcal{N}(0,1)$ and $\xi_j \sim_{\text{iid}} \mathcal{U}([0,\delta])$.

# Quantizing the RIP?

- Let's retake: for $\mathcal{Q}(\cdot) = \delta\lfloor \cdot/\delta \rfloor \in \delta\mathbb{Z}$

$$\boldsymbol{\psi}(\boldsymbol{u}) := \mathcal{Q}(\boldsymbol{\Phi}\boldsymbol{u} + \boldsymbol{\xi}), \text{ with } \boldsymbol{\psi}_j(\boldsymbol{u}) := \mathcal{Q}(\boldsymbol{\varphi}_j^T\boldsymbol{u} + \xi_j)$$

with $\boldsymbol{\Phi}_{ji} \sim_{\text{iid}} \mathcal{N}(0,1)$ and $\xi_j \sim_{\text{iid}} \mathcal{U}([0,\delta])$.

- Naive way: since $|a-b| - \delta \leq |\mathcal{Q}(a) - \mathcal{Q}(b)| \leq |a-b| + \delta, \quad \forall\, a,b \in \mathbb{R}$

$$(1-\rho)\,\|\boldsymbol{u}-\boldsymbol{v}\| - \delta \;\leq\; \tfrac{1}{\sqrt{M}}\,\|\boldsymbol{\psi}(\boldsymbol{u}) - \boldsymbol{\psi}(\boldsymbol{v})\| \;\leq\; (1+\rho)\,\|\boldsymbol{u}-\boldsymbol{v}\| + \delta,$$

whenever $\frac{1}{\sqrt{M}}\boldsymbol{\Phi}$ is RIP.

# Quantizing the RIP?

- Let's retake: for $\mathcal{Q}(\cdot) = \delta\lfloor \cdot/\delta \rfloor \in \delta\mathbb{Z}$

$$\boldsymbol{\psi}(\boldsymbol{u}) := \mathcal{Q}(\boldsymbol{\Phi u} + \boldsymbol{\xi}), \text{ with } \psi_j(\boldsymbol{u}) := \mathcal{Q}(\boldsymbol{\varphi}_j^T \boldsymbol{u} + \xi_j)$$

with $\boldsymbol{\Phi}_{ji} \sim_{\text{iid}} \mathcal{N}(0,1)$ and $\xi_j \sim_{\text{iid}} \mathcal{U}([0,\delta])$.

- Naive way: since $|a-b| - \delta \leq |\mathcal{Q}(a) - \mathcal{Q}(b)| \leq |a-b| + \delta, \quad \forall\, a,b \in \mathbb{R}$

$$(1-\rho)\,\|\boldsymbol{u}-\boldsymbol{v}\| - \delta \leq \tfrac{1}{\sqrt{M}}\,\|\boldsymbol{\psi}(\boldsymbol{u}) - \boldsymbol{\psi}(\boldsymbol{v})\| \leq (1+\rho)\,\|\boldsymbol{u}-\boldsymbol{v}\| + \delta,$$

**multiplicative error**

With $\rho = O(\sqrt{K/M})$.

(decaying, good!)

**additive error**

(constant, weird!?)

# Quantizing the RIP?

- Let's retake: for $\mathcal{Q}(\cdot) = \delta\lfloor \cdot/\delta \rfloor \in \delta\mathbb{Z}$

$$\boldsymbol{\psi}(\boldsymbol{u}) := \mathcal{Q}(\boldsymbol{\Phi u} + \boldsymbol{\xi}), \text{ with } \boldsymbol{\psi}_j(\boldsymbol{u}) := \mathcal{Q}(\boldsymbol{\varphi}_j^T \boldsymbol{u} + \xi_j)$$

with $\boldsymbol{\Phi}_{ji} \sim_{\text{iid}} \mathcal{N}(0,1)$ and $\xi_j \sim_{\text{iid}} \mathcal{U}([0,\delta])$.

- Solution:

1. Let's use another distance ($\ell_1$):

$$\tfrac{1}{M}\|\boldsymbol{\psi}(\boldsymbol{u}) - \boldsymbol{\psi}(\boldsymbol{v})\|_1 = \tfrac{1}{M}\sum_j |\boldsymbol{\psi}_j(\boldsymbol{u}) - \boldsymbol{\psi}_j(\boldsymbol{v})|$$

2. Let's study how it concentrates!

# Quantizing the RIP?

- Let's retake: for $\mathcal{Q}(\cdot) = \delta\lfloor \cdot/\delta \rfloor \in \delta\mathbb{Z}$

$$\boldsymbol{\psi}(\boldsymbol{u}) := \mathcal{Q}(\boldsymbol{\Phi u} + \boldsymbol{\xi}), \text{ with } \psi_j(\boldsymbol{u}) := \mathcal{Q}(\boldsymbol{\varphi}_j^T \boldsymbol{u} + \xi_j)$$

with $\boldsymbol{\Phi}_{ji} \sim_{\text{iid}} \mathcal{N}(0,1)$ and $\xi_j \sim_{\text{iid}} \mathcal{U}([0,\delta])$.

Quantized Gaussian Quasi-Isometric Embedding [LJ, 2015]

*Given an error* $0 < \epsilon < 1$, *and* $\mathcal{K} \subset \mathbb{R}^N$.
*If* $M$ *is such that*

$$M \gtrsim \epsilon^{-5} w(\mathcal{K})^2,$$

*then, for some* $c > 0$ *and for all* $\boldsymbol{u}, \boldsymbol{v} \in \mathcal{K}$, *and w.h.p., we have*

$$\Big(\sqrt{\tfrac{2}{\pi}} - \epsilon\Big)\,\|\boldsymbol{u} - \boldsymbol{v}\| - c\delta\epsilon \;\le\; \tfrac{1}{M}\,\|\boldsymbol{\psi}(\boldsymbol{u}) - \boldsymbol{\psi}(\boldsymbol{v})\|_1 \;\le\; \Big(\sqrt{\tfrac{2}{\pi}} + \epsilon\Big)\,\|\boldsymbol{u} - \boldsymbol{v}\| + c\delta\epsilon,$$

For $\mathcal{K} = \boldsymbol{A}\Sigma_K \cap \mathbb{B}^N$
and $\boldsymbol{A}$ ONB
$M \gtrsim \epsilon^{-2} K \log \frac{N}{K\delta\epsilon^{3/2}}$

all distortions decay with $M$!

# Quantizing the RIP?

- Let's retake: for $\mathcal{Q}(\cdot) = \delta\lfloor \cdot/\delta \rfloor \in \delta\mathbb{Z}$

$$\boldsymbol{\psi}(\boldsymbol{u}) := \mathcal{Q}(\boldsymbol{\Phi}\boldsymbol{u} + \boldsymbol{\xi}), \text{ with } \boldsymbol{\psi}_j(\boldsymbol{u}) := \mathcal{Q}(\boldsymbol{\varphi}_j^T \boldsymbol{u} + \xi_j)$$

with $\cancel{\boldsymbol{\Phi}_{ji} \sim_{\rm iid} \mathcal{N}(0,1)}$ and $\xi_j \sim_{\rm iid} \mathcal{U}([0,\delta])$.

OK for sub-Gaussian?
(e.g., Bernoulli)

# Quantizing the RIP?

- Let's retake: for $\mathcal{Q}(\cdot) = \delta\lfloor \cdot/\delta \rfloor \in \delta\mathbb{Z}$

$$\boldsymbol{\psi}(\boldsymbol{u}) := \mathcal{Q}(\boldsymbol{\Phi u} + \boldsymbol{\xi}),\ \text{with } \boldsymbol{\psi}_j(\boldsymbol{u}) := \mathcal{Q}(\boldsymbol{\varphi}_j^T \boldsymbol{u} + \xi_j)$$

with $\cancel{\boldsymbol{\Phi}_{ji} \sim_{\text{iid}} \mathcal{N}(0,1)}$ and $\xi_j \sim_{\text{iid}} \mathcal{U}([0,\delta])$.

Quantized **sub-**Gaussian Quasi-Isometric Embedding [LJ, 2015]

*Given an error* $0 < \epsilon < 1$, *and* $\mathcal{K} \subset \mathbb{R}^N$.
*If* $M$ *is such that*

$$M \gtrsim \epsilon^{-5} w(\mathcal{K})^2,$$

*then, w.h.p, for some* $c > 0$ *and for all* $\boldsymbol{u}, \boldsymbol{v} \in \mathcal{K}$ *with* $\boldsymbol{u} - \boldsymbol{v} \in C_{K_0}$, *we have*

$$\Big(\sqrt{\tfrac{2}{\pi}} - \epsilon - \tfrac{\kappa}{\sqrt{K_0}}\Big)\, \|\boldsymbol{u}-\boldsymbol{v}\| - c\delta\epsilon \ \le\ \tfrac{1}{M}\, \|\boldsymbol{\psi}(\boldsymbol{u}) - \boldsymbol{\psi}(\boldsymbol{v})\|_1 \ \le\ \Big(\sqrt{\tfrac{2}{\pi}} + \epsilon + \tfrac{\kappa}{\sqrt{K_0}}\Big)\, \|\boldsymbol{u}-\boldsymbol{v}\| + c\delta\epsilon.$$

For $\mathcal{K} = \boldsymbol{A}\Sigma_K \cap \mathbb{B}^N$
and $\boldsymbol{A}$ ONB
$M \gtrsim \epsilon^{-2} K \log \frac{N}{K\delta\epsilon^{3/2}}$

high $K_0$, less sparse but lower distortion!

But you can use the "spread the samples" trick!

# To conclude ...

# Take away messages

- Associating CS and Quantization
  provides many interesting questions:
  - geometrically (high dim. convex geom.)
  - numerically (not totally covered here)
  - with impacts in CS sensor design

# Take away messages

- Associating CS and Quantization provides many interesting questions:
  - geometrically (high dim. convex geom.)
  - numerically (not totally covered here)
  - with impacts in CS sensor design
- Beyond CS, quantifying random projections
  - leads to interesting embedding problems
  - possible impacts in dimensionality reductions

# Open questions


- CoBP robustness vs pre-quantisation noise?
  Do quasi-isometric embedding help?
- Quasi-isometric embedding for Hilbert spaces?
- Embeddings with other quantisation schemes?
  (link with machine learning?)
- Classification/clustering in the quantized domain?

# Thank you for the invitation!

A. Moshtaghpour, LJ, V. Cambareri, K. Degraux and C. De Vleeschouwer. "**Consistent Basis Pursuit for Signal and Matrix Estimates in Quantized Compressed Sensing",** IEEE Signal Processing Letters, vol. 23, no. 1, p. 25 - 29, january 2016.

P. T. Boufounos, LJ, F. Krahmer and R. Saab, "**Quantization and Compressive Sensing**", arXiv: 1405.1194, 2014 (to appear in Springer book "Compressed Sensing and Its Applications")

LJ, J. N. Laska, P. T. Boufounos, and R. G. Baraniuk, "**Robust 1-Bit Compressive Sensing via Binary Stable Embeddings of Sparse Vectors**", IEEE TIT, 59(4), pp. 2082-2102, 2013.

A. Powell, J. Whitehouse, "**Error bounds for consistent reconstruction: random polytopes and coverage processes**", to appear in FoCM, arXiv: 1405.7094

S. Dirksen, G. Lecué, H. Rauhut, "**On the gap between RIP-properties and sparse recovery conditions**" , arXiv: 1504.05073

LJ, "**Error Decay of (almost) Consistent Signal Estimations from Quantized Random Gaussian Projections**", submitted to TIT, arXiv: 1406.0022

LJ, "**A Quantized Johnson Lindenstrauss Lemma: The Finding of Buffon's Needle**", submitted to TIT, arXiv: 1309.1507

LJ, "**Small width, low distortions: quasi-isometric embeddings with quantized sub-Gaussian random projections"**, Submitted to TIT, arXiv: 1504.06170

+ references inside the presentation