What can we learn from the Compressed Sensing theory?

Laurent Jacques (ICTEAM/ELEN, UCL)

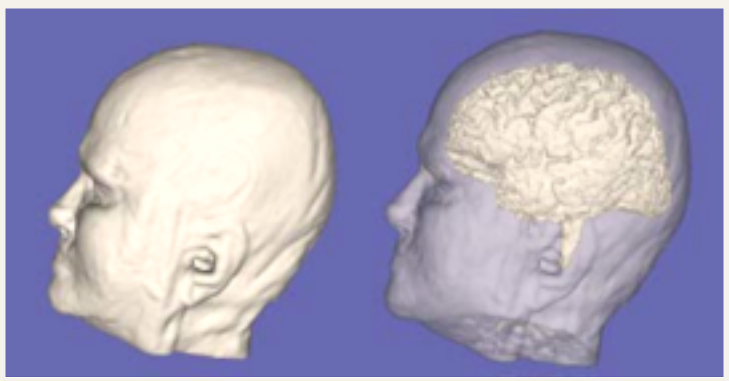
March 22, 2012.

Ecole Centrale de Lille, France.

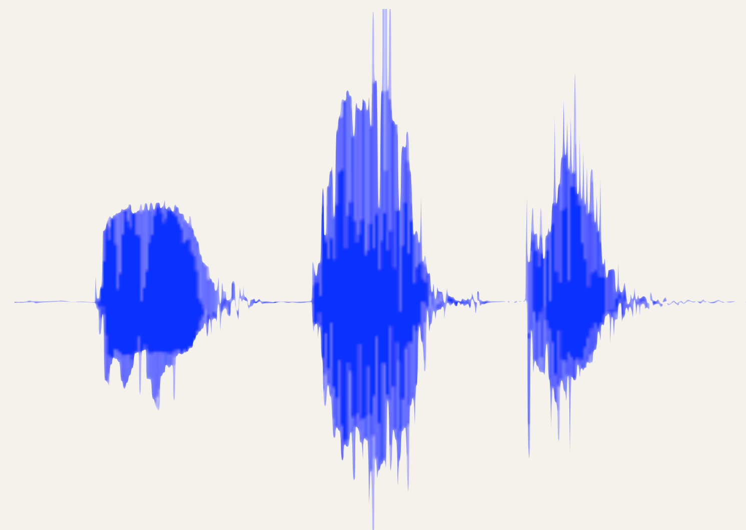
I. Sparsity is everywhere ...

Sparse Signal Representation (1/2)

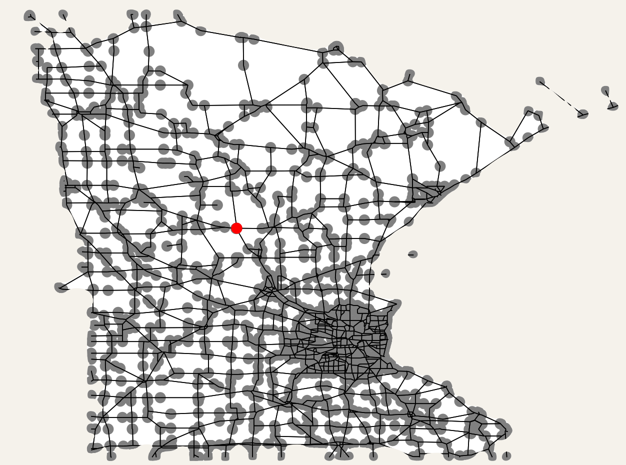
Informative signals are composed of structures ...



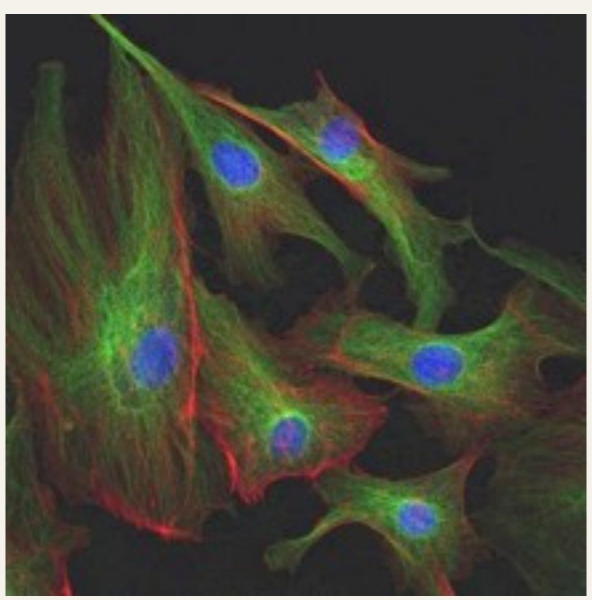
3-D data



Speech signal



Data on Graph



Biology

Video

Astronomy

Sparse Signal Representation (2/2)

- Signal $x \in \mathbb{R}^N$ (e.g. N = pixel number, voxels, graph nodes, ...)
- There exists a "sparsity" basis (e.g. wavelets, Fourier, ...)

$$\Psi = (\Psi_1, \cdots, \Psi_D) \in \mathbb{R}^{N \times D}$$

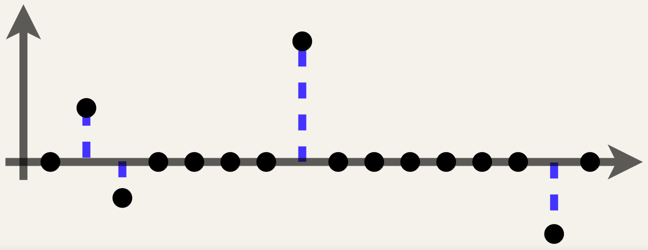
where x has a linear representation

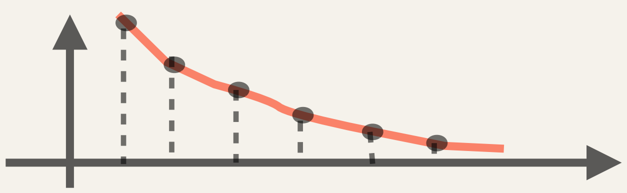
$$x = \sum_{j=1}^{D} \alpha_j \Psi_j = \Psi\alpha$$

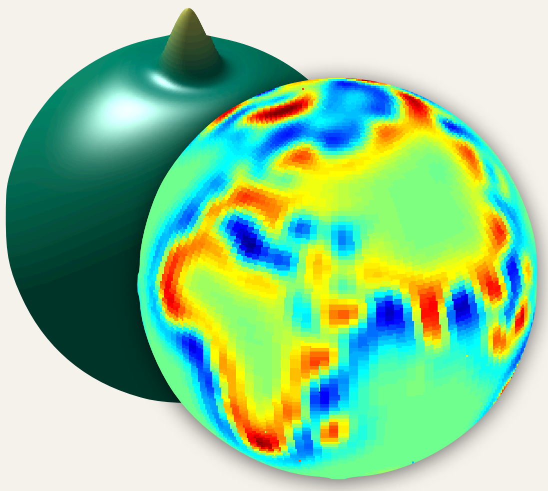
$\alpha \in \mathbb{R}^D$ is the *coefficient vector.*

- <u>**Sparsity**</u> :

$$\#\{i : \alpha_i \neq 0\} \ll N \qquad \|\alpha - \alpha_K\| \ll \|\alpha\|$$

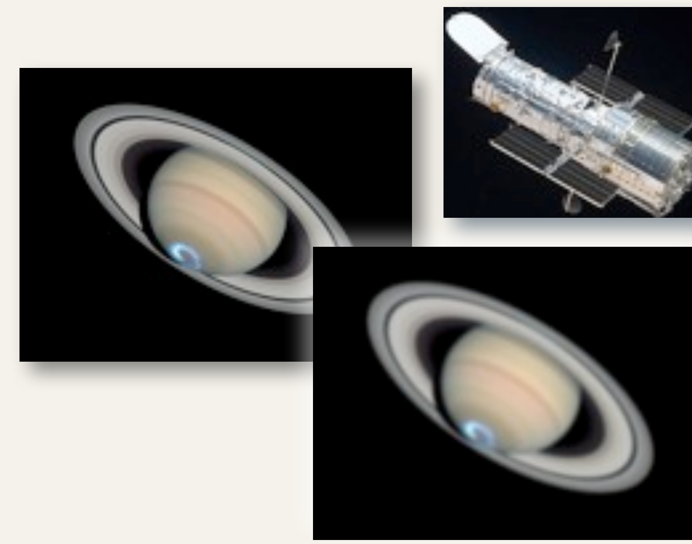






General Sparsity Application

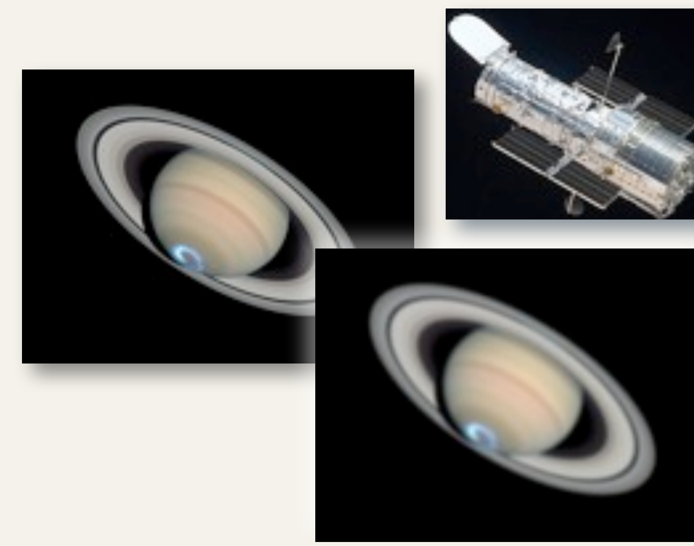
1. Data Compression/Transmission (by definition);
2. Data restoration :
   * Denoising,
   * Debluring,
   * Inpainting, ...



3. Simplified model and interpretation (e.g. in ML)

General Sparsity Application

1. <u>Data Compression/Transmission</u> (by definition);
2. <u>Data restoration</u> :
   - ∗ Denoising,
   - ∗ Debluring,
   - ∗ Inpainting, ...



3. <u>Simplified model and interpretation</u> (*e.g.*, in ML)

More generally,

For regularizing (stabilizing) inverse problems

<u>Impact on data sampling philosophy !</u>

(applications: optics, biomedical, astronomy, ...)

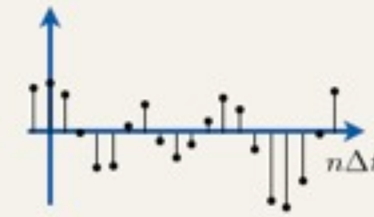
II. Adjusting sampling to sparsity

Sampling with Sparsity (1/2)

- Paradigm shift:

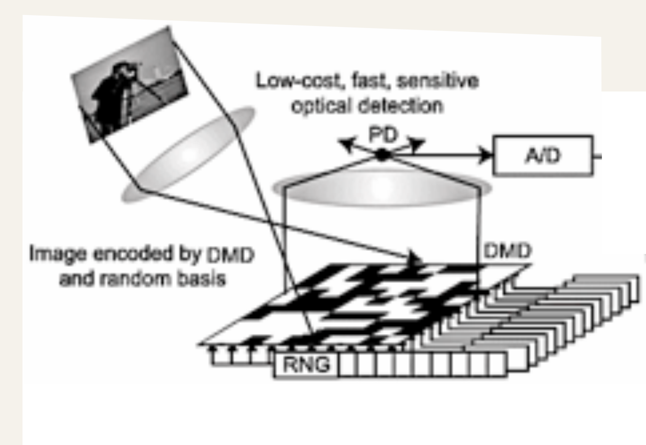
  "Computer readable" sensing

  \+ prior information



World → Signal → Sensing Device → Sensing →  → Signal → Human

Optimized setup: sampling rate $\propto$ information



- Examples:

  Radio-Interferometry, Compressed Sensing, MRI, Deflectometry, Seismology, ...

Sampling with Sparsity (2/2)

but ... non-linear reconstruction schemes!

Regularized inverse problems:

Reconstruct $x \in \mathbb{R}^N$ from $y = \text{Sensing}(x) \in \mathbb{R}^M$ given a sparse model on x.

Examples: Tomography, frequency/partial observations, ...

$$x^* = \operatorname*{argmin}_{u \in \mathbb{R}^N} \mathcal{S}(u) \text{ s.t. } \text{Sensing}(u) \approx \text{Sensing}(x)$$

Sparsity metric:

e.g., small $\mathcal{S}(\alpha) = \|\alpha\|_1$ if $u = \Psi\alpha$,
small Total Variation $\mathcal{S}(u) = \|\nabla u\|$

Noise: Gaussian, Poisson, ...

III. Compressed Sensing

Compressed Sensing

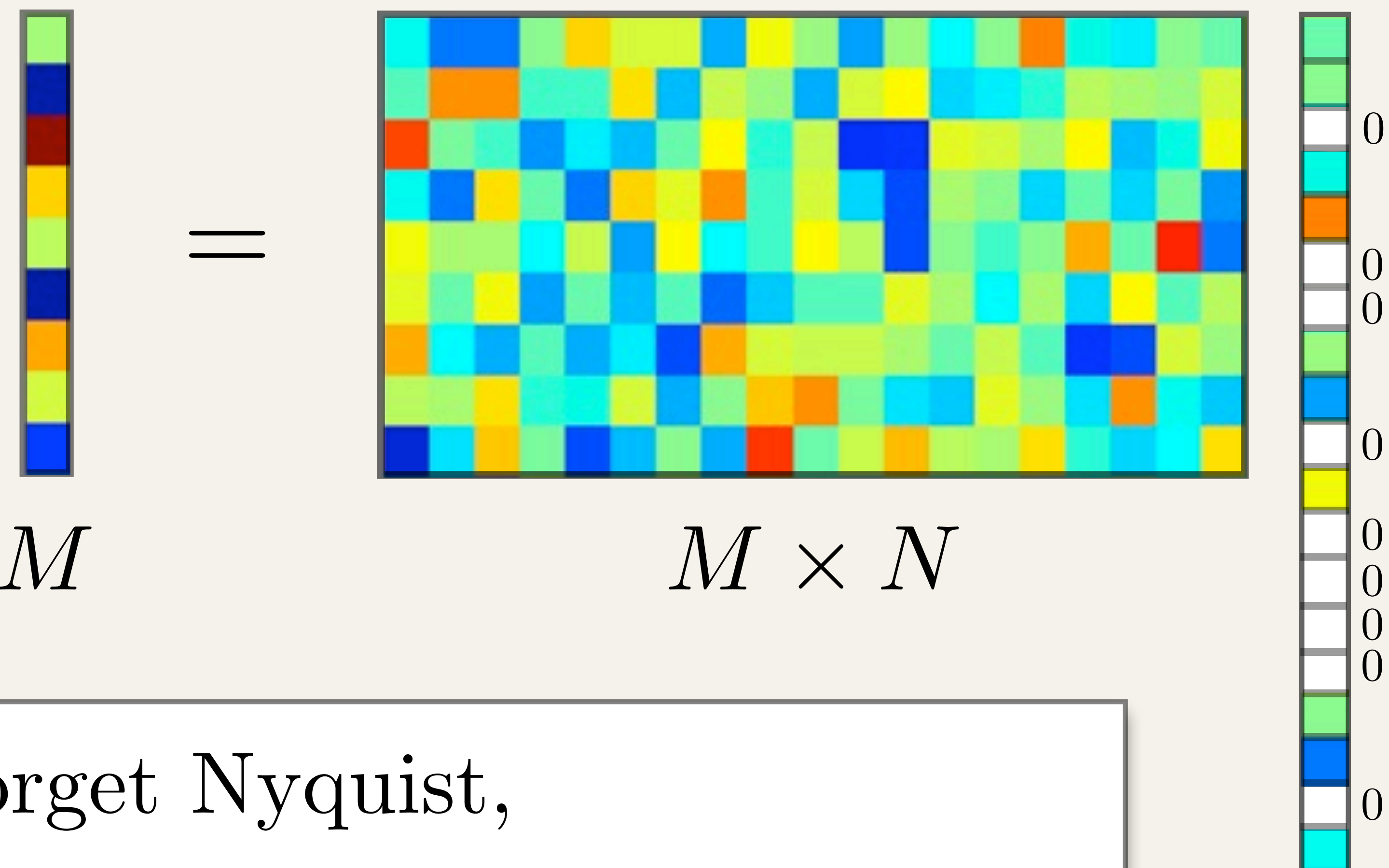
CS in a nutshell:

"Forget" Dirac, forget Nyquist,
ask *few* (linear) *questions*
about your informative (sparse) signal,
and recover it *differently* (non-linearly)"

Compressed Sensing

M questions — Sensing method — Signal

$\boldsymbol{y}$ — $\boldsymbol{\Phi}$ — $\boldsymbol{x}$



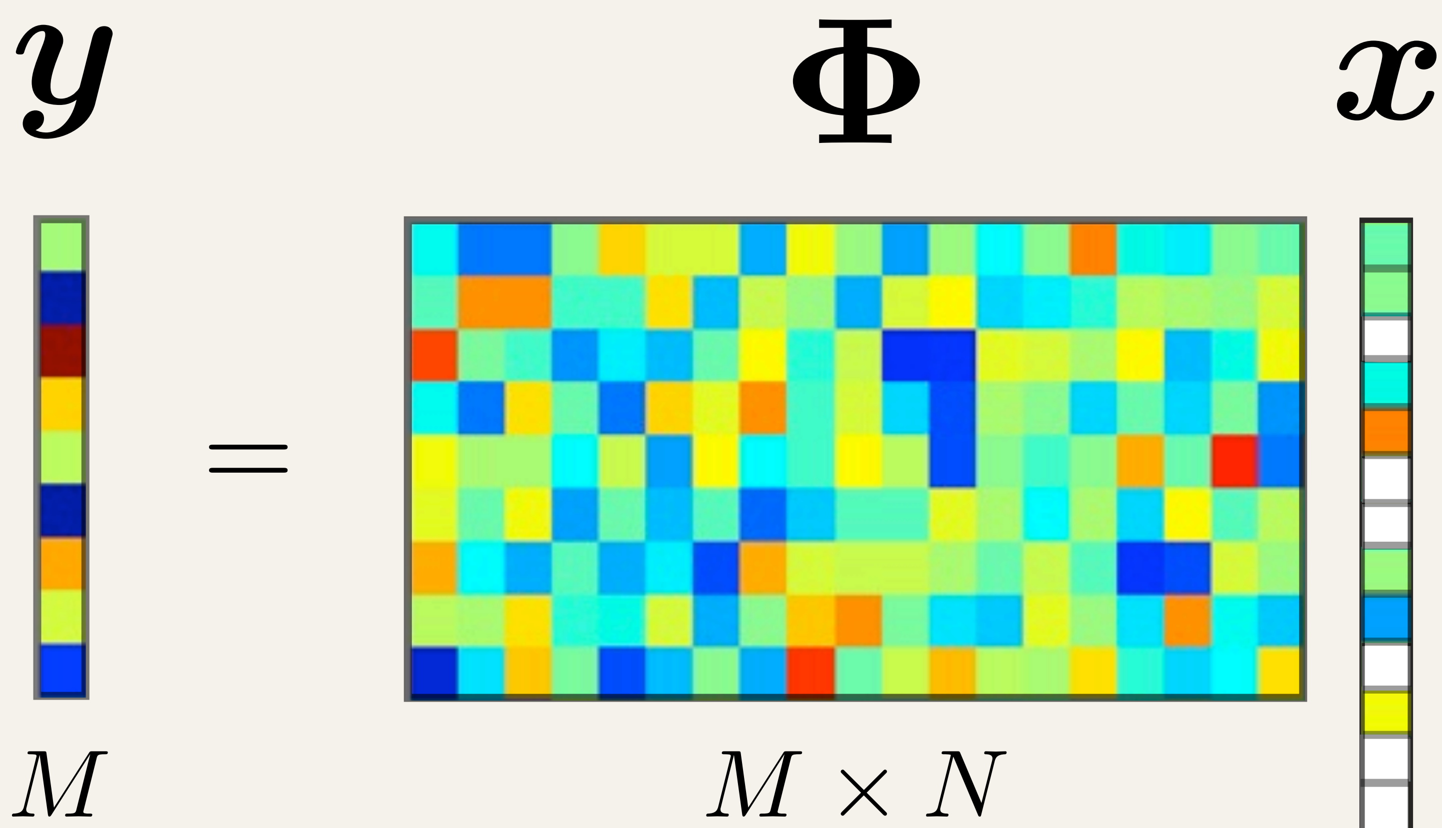
Sparsity Prior ($\Psi = \mathrm{Id}$)

M — $M \times N$ — N

CS in a nutshell:

"Forget" Dirac, forget Nyquist,
ask *few* (linear) *questions*
about your informative (sparse) signal,
and recover it *differently* (non-linearly)"

Compressed Sensing

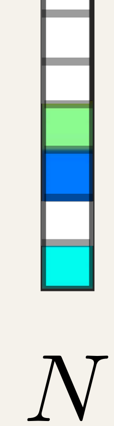


Non-linear reconstruction:

If $\boldsymbol{x}$ is K-sparse and if $\boldsymbol{\Phi}$ well "conditioned"

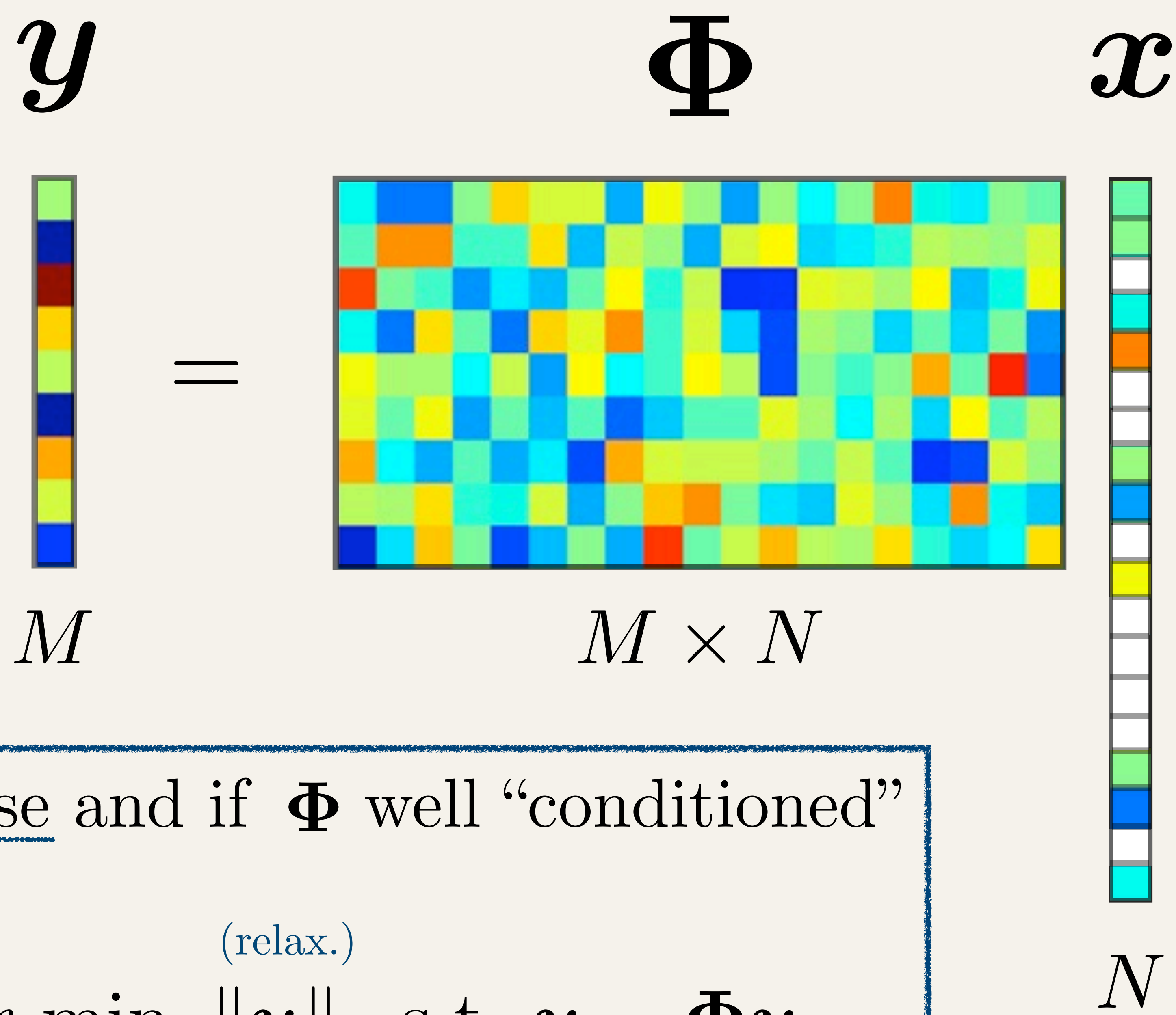
then:

$$\boldsymbol{x}^* = \underset{\boldsymbol{u} \in \mathbb{R}^N}{\arg\min} \ \|\boldsymbol{u}\|_0 \text{ s.t. } \boldsymbol{y} = \boldsymbol{\Phi}\boldsymbol{u}$$



$\|\boldsymbol{u}\|_0 = \#\{j : u_j \neq 0\}$

Compressed Sensing



If $\boldsymbol{x}$ is K-sparse and if $\boldsymbol{\Phi}$ well "conditioned"

then:

(relax.)

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{u}\in\mathbb{R}^N} \|\boldsymbol{u}\|_{\cancel{2}\;1} \text{ s.t. } \boldsymbol{y} = \boldsymbol{\Phi}\boldsymbol{u}$$

$\|\boldsymbol{u}\|_1 = \sum_j |u_j|$

(Basis Pursuit) [Chen, Donoho, Saunders, 1998]

Compressed Sensing

$\exists\ \delta \in (0,1)$ Restricted Isometry Property

$$\sqrt{1-\delta}\,\|\boldsymbol{v}\|_2 \ \leqslant\ \|\boldsymbol{\Phi v}\|_2 \ \leqslant\ \sqrt{1+\delta}\,\|\boldsymbol{v}\|_2$$

for all $2K$ sparse signals $\boldsymbol{v}$.

any subset of $2K$ columns
is an *isometry*

If $\boldsymbol{x}$ is K-sparse and if $\boldsymbol{\Phi}$ well "conditioned"
then:

(relax.)

$$\boldsymbol{x}^* = \ \arg\min_{\boldsymbol{u}\in\mathbb{R}^N} \ \|\boldsymbol{u}\|_{\cancel{0}\,1} \ \text{s.t.}\ \boldsymbol{y} = \boldsymbol{\Phi u}$$

if $\delta < \sqrt{2}-1$ [Candes 08]

$\|\boldsymbol{u}\|_1 = \sum_j |u_j|$

(Basis Pursuit) [Chen, Donoho, Saunders, 1998]

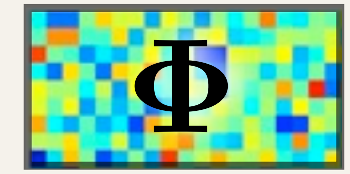
Compressed Sensing

$\exists\ \delta \in (0,1)$ Restricted Isometry Property

$$\sqrt{1-\delta}\,\|\boldsymbol{v}\|_2 \ \leqslant\ \|\boldsymbol{\Phi v}\|_2 \ \leqslant\ \sqrt{1+\delta}\,\|\boldsymbol{v}\|_2$$

for all $2K$ sparse signals $\boldsymbol{v}$.

any subset of $2K$ columns is an *isometry*

Examples: 

- Gaussian
- Bernoulli
- Random Fourier
- ....

$$M = O(K \ln N/K) \ll N$$

$$\boldsymbol{\Phi} \in \mathbb{R}^{M\times N},\ \ \Phi_{ij} \sim_{\rm iid} \mathcal{N}(0,1)$$

If $\boldsymbol{x}$ is K-sparse and if $\boldsymbol{\Phi}$ well "conditioned"

then:

(relax.)

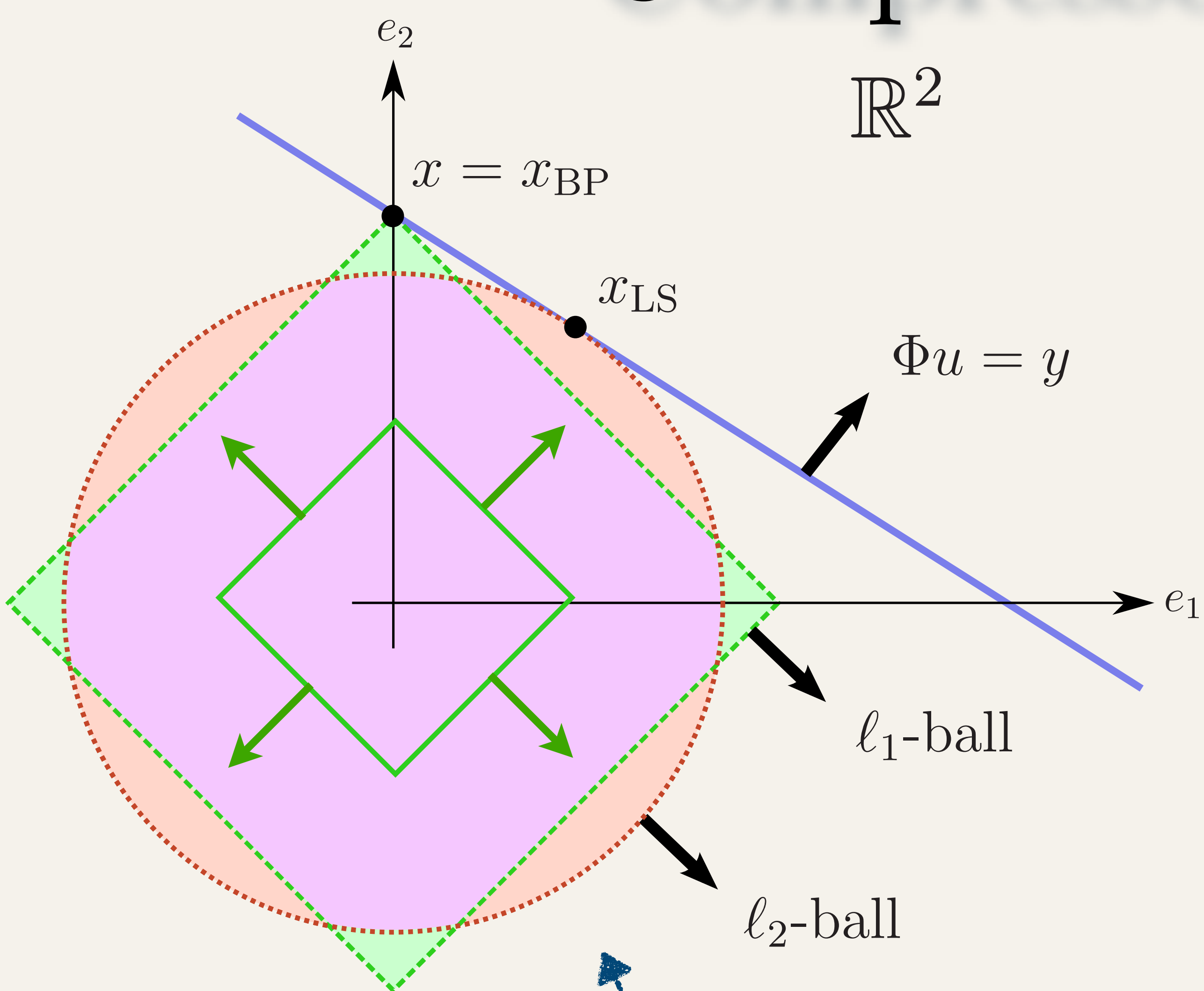
$$\boldsymbol{x}^* = \ \underset{\boldsymbol{u}\in\mathbb{R}^N}{\arg\min}\ \|\boldsymbol{u}\|_{\underset{1}{\cancel{0}}} \text{ s.t. } \boldsymbol{y} = \boldsymbol{\Phi u}$$

if $\delta < \sqrt{2}-1$ [Candes 08]

$\|\boldsymbol{u}\|_1 = \sum_j |u_j|$

(Basis Pursuit) [Chen, Donoho, Saunders, 1998]

Compressed Sensing



If $\boldsymbol{x}$ is K-sparse and if $\boldsymbol{\Phi}$ well "conditioned"

then:

(relax.)

$$\boldsymbol{x}^* = \underset{\boldsymbol{u}\in\mathbb{R}^N}{\arg\min}\ \|\boldsymbol{u}\|_{\cancel{2}\,\mathbf{1}} \ \text{s.t.}\ \boldsymbol{y} = \boldsymbol{\Phi}\boldsymbol{u}$$

$\|\boldsymbol{u}\|_1 = \sum_j |u_j|$

(Basis Pursuit) [Chen, Donoho, Saunders, 1998]

Solvers:

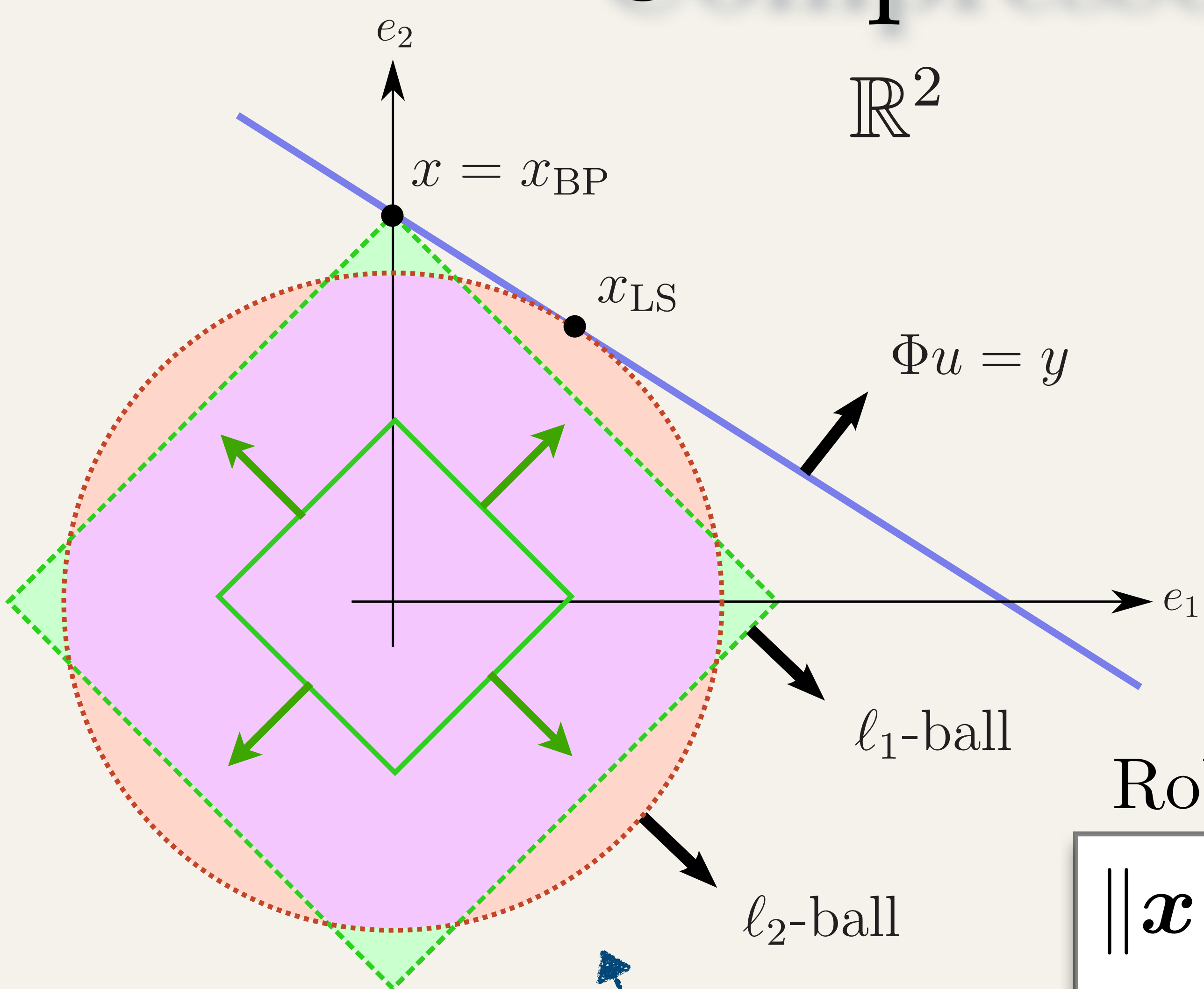
Linear Programming,

Interior Point Method,

Proximal Methods,

... **Tons** of toolboxes ...

Compressed Sensing



Robustness: *vs* sparse deviation + noise.

$$\|\boldsymbol{x} - \boldsymbol{x}^*\| \leqslant C \frac{1}{\sqrt{K}} \|\boldsymbol{x} - \boldsymbol{x}_K\|_1 + D\epsilon$$

If $\boldsymbol{x}$ is K-sparse and if $\boldsymbol{\Phi}$ well "conditioned" then:

(relax.)

$\|\boldsymbol{y} - \boldsymbol{\Phi u}\| \leqslant \epsilon$

$$\boldsymbol{x}^* = \underset{\boldsymbol{u} \in \mathbb{R}^N}{\arg\min} \; \|\boldsymbol{u}\|_{\cancel{0}\,1} \text{ s.t. } \boldsymbol{y} = \boldsymbol{\Phi u}$$

$\|\boldsymbol{u}\|_1 = \sum_j |u_j|$

(Basis Pursuit) [Chen, Donoho, Saunders, 1998]

Solvers:

Linear Programming,
Interior Point Method,
Proximal Methods,
... **Tons** of toolboxes ...

IV. CS and Quantization

Joint work with:

D. Hammond (U. Oregon, USA)

J. Fadili (ENSICaen, France)

Quantization of CS Measurements

* Turning Measurements into bits → scalar quantization

$$\boxed{y \;=\; \mathcal{Q}\big[\Phi x\big] \;\in \Omega^M,}$$

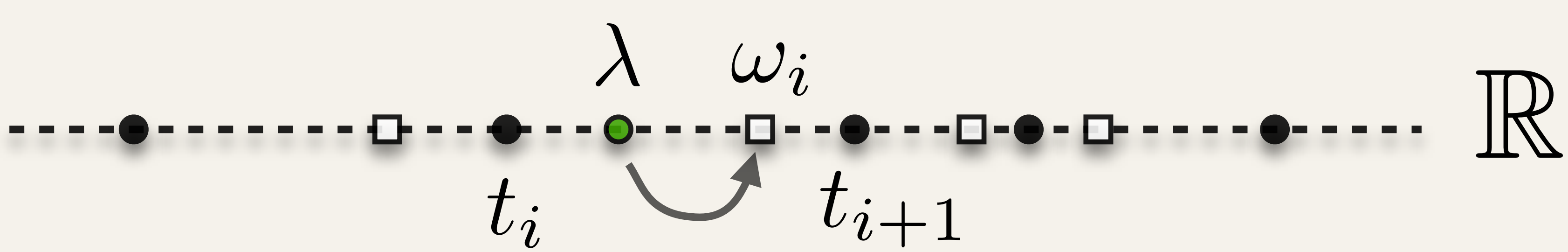
with:

$$\Omega = \{\omega_i \in \mathbb{R} : 1 \leqslant i \leqslant B\}, \;(B = 2^R) \qquad \text{(levels)}$$

$$T = \{t_i \in \overline{\mathbb{R}} : 1 \leqslant i \leqslant B+1, t_i \leqslant t_{i+1}\} \qquad \text{(thresholds)}$$

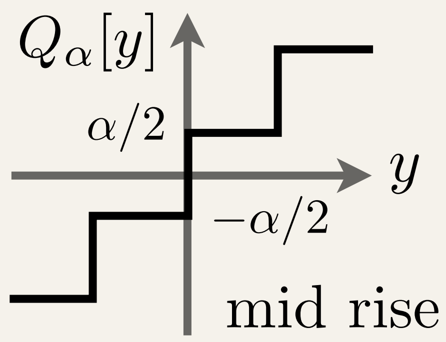
$$\forall \lambda \in \mathbb{R}, \qquad \mathcal{Q}[\lambda] = \omega_i \;\Leftrightarrow\; \lambda \in \mathcal{R}_i \triangleq [t_i, t_{i+1}),$$

$$\forall u \in \mathbb{R}^M, \quad (\mathcal{Q}[u])_j = \mathcal{Q}[u_j]$$



(e.g. Lloyd-Max)

Quantization and CS: the former association



- Most of the time... Uniform Quantization

$$(Q_\alpha[y])_i \quad = \quad \alpha\lfloor (y)_i/\alpha \rfloor + \alpha/2.$$

with $t_i = i\alpha,\ t_{i+1} - t_i = \alpha,\ \omega_i = (t_i + t_{i+1})/2 = (i + 1/2)\alpha$

- Reconstruction: since $\|y - \Phi x\|_\infty \leqslant \frac{\alpha}{2}$
  - quantization ≈ uniform noise: $u_i \; = \; y_i - (\Phi x)_i \sim_{\mathrm{iid}} \mathcal{U}([-\frac{\alpha}{2}, \frac{\alpha}{2}])$
  - BPDN with $\epsilon_2^2(\alpha) \; = \; E[\|u\|_2^2] \; + \; \kappa\,\mathrm{Var}^{\frac{1}{2}}[\|u\|_2^2]$
    $\; = \; M\frac{\alpha^2}{12} \; + \; \kappa\, M^{\frac{1}{2}}\frac{\alpha^2}{6\sqrt{5}}$

If RIP,

$$\boxed{\begin{gathered}\boldsymbol{x}^* = \operatorname*{arg\,min}_{\boldsymbol{u}\in\mathbb{R}^N} \|\boldsymbol{u}\|_1 \text{ s.t. } \|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{u}\|_2 \leqslant \epsilon_2 \\ \|\boldsymbol{x} - \boldsymbol{x}^*\| = O(\alpha)\end{gathered}}$$

Quantization and CS: the former association

* Most of the time... Uniform Quantization

$$(Q_\alpha[y])_i \quad = \quad \alpha\lfloor (y)_i/\alpha \rfloor + \alpha/2.$$

with $t_i = i\alpha,\ t_{i+1} - t_i = \alpha,\ \omega_i = (t_i + t_{i+1})/2 = (i + 1/2)\alpha$

* Reconstruction: since $\|y - \Phi x\|_\infty \leqslant \frac{\alpha}{2}$
  * quantization ≈ uniform noise: $u_i \ = \ y_i - (\Phi x)_i \sim_{\rm iid} \mathcal{U}([-\frac{\alpha}{2}, \frac{\alpha}{2}])$
  * BPDN with $\epsilon_2^2(\alpha) \ = \ E[\|u\|_2^2] \ + \ \kappa\, \mathrm{Var}^{\frac{1}{2}}[\|u\|_2^2]$
    $= \ M\frac{\alpha^2}{12} \ + \ \kappa\, M^{\frac{1}{2}}\frac{\alpha^2}{6\sqrt{5}}$

But, not adapted: if x^* solution of BPDN,

* $Q_\alpha\big[\Phi x^*\big] \ \neq y$, i.e no Quantization Consistency (QC)
* ℓ_2 constraint ≈ Gaussian distribution (MAP - cond. log. lik.)

Uniform CS Dequantizers

New optimization schemes (towards the constraint $\|y - \Phi v\|_\infty$):

$$\boxed{\Delta_{1,p}(y,\epsilon) \;=\; \arg\min_{v\in\mathbb{R}^N} \|v\|_1 \text{ s.t. } \|y-\Phi v\|_p \;\leqslant \epsilon, \quad (\mathbf{BPDQ}_p)}$$

for "Basis Pursuit DeQuantizer" of *moment* $p \geqslant 1$

BPDQ_p adapted to GGD noise (of shape p): $n_i \sim \mathrm{pdf} \propto \exp -|t/b|^p \;(b>0)$

If Φ is RIP_p of order K, *i.e.*,

$\exists\; \mu_p > 0,\; \delta \in (0,1)$,

$$\sqrt{1-\delta}\,\|v\|_2 \;\leqslant\; \tfrac{1}{\mu_p}\,\|\Phi v\|_p \;\leqslant\; \sqrt{1+\delta}\,\|v\|_2,$$

for all K sparse signals v.

Then, reconstruction error $\underline{O(\alpha/\sqrt{p+1})}$.

Uniform CS Dequantizers

New optimization schemes (towards the constraint $\|y - \Phi v\|_\infty$):

$$\boxed{\Delta_{1,p}(y,\epsilon) \;=\; \mathop{\arg\min}_{v\in\mathbb{R}^N} \|v\|_1 \text{ s.t. } \|y - \Phi v\|_p \;\leqslant \epsilon, \quad (\mathbf{BPDQ}_p)}$$

for "Basis Pursuit DeQuantizer" of *moment* $p \geqslant 1$

BPDQ_p adapted to GGD noise (of shape p): $n_i \sim \mathrm{pdf} \propto \exp -|t/b|^p \;\; (b > 0)$

If Φ is RIP_p of order K, *i.e.*,

$$\boxed{\begin{array}{l} \exists\ \mu_p > 0,\ \ \delta \in (0,1), \\ \quad \sqrt{1-\delta}\,\|v\|_2 \;\leqslant\; \frac{1}{\mu_p}\,\|\Phi v\|_p \;\leqslant\; \sqrt{1+\delta}\,\|v\|_2, \\ \text{for all } K \text{ sparse signals } v. \end{array}}$$

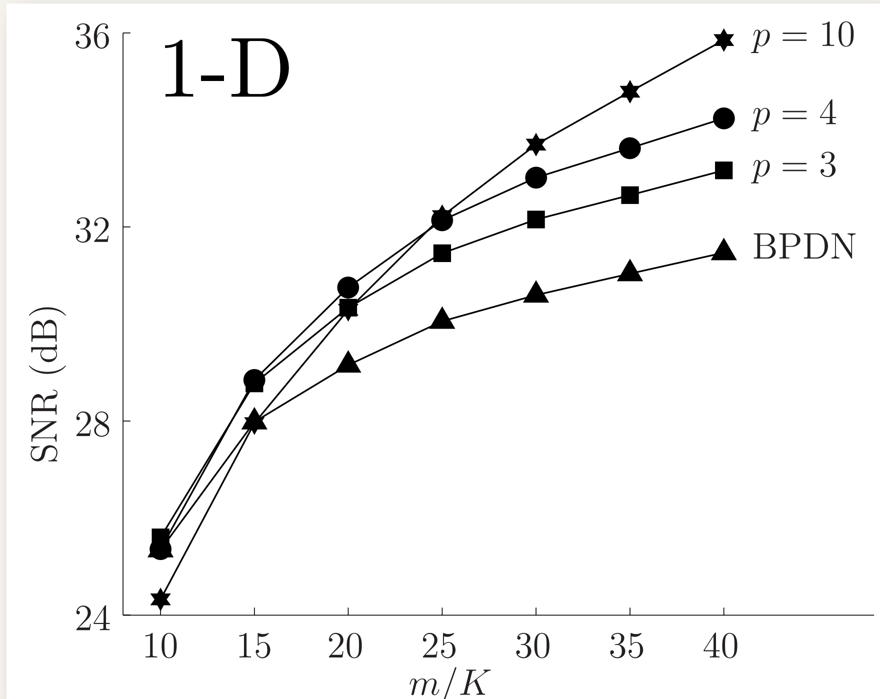
Then, reconstruction error $O(\alpha/\sqrt{p+1})$.

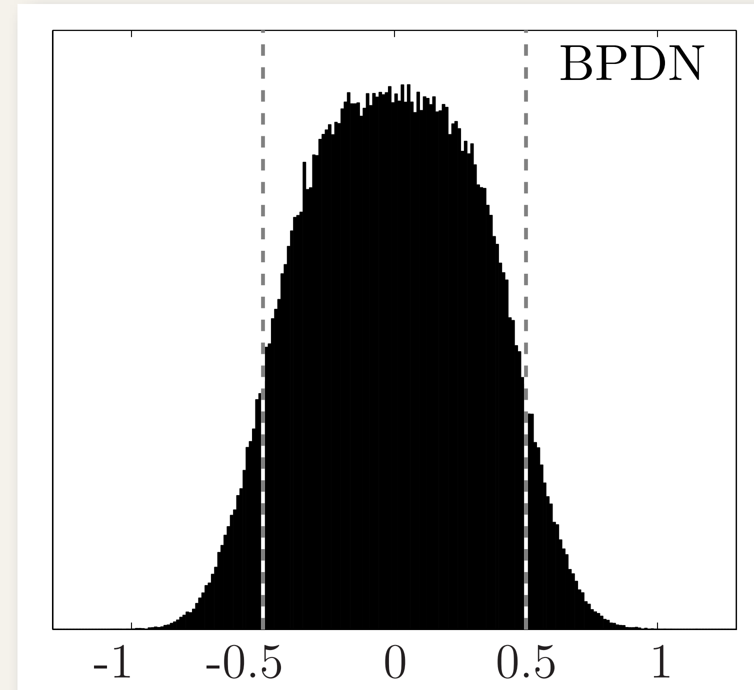
But no free lunch: $M = O\big(\,(K \log N/K\,)^{p/2}\,\big)$

$\Longrightarrow$ Another reading: limited range of valid p for a given M (and K)!

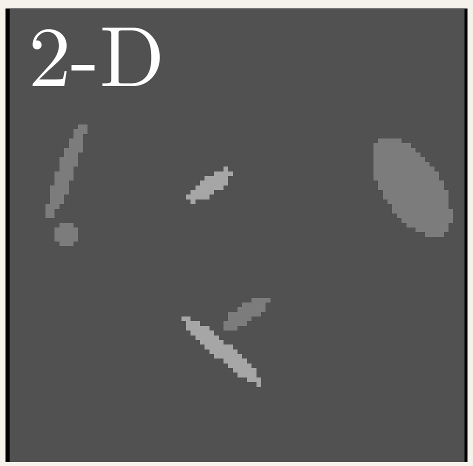
Numerical illustration

Histograms of
$\alpha^{-1}(y - \Phi x_p^*)_i$





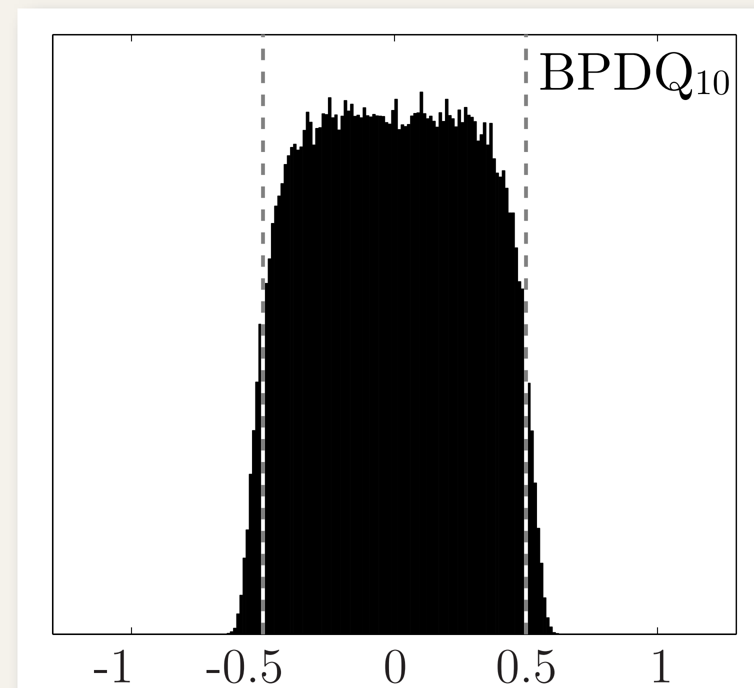
Original



$N/M = 8$
12 bins/meas.
Rand. Fourier

BPDN-TV
SNR: 8.96 dB

BPDQ_{10}-TV
SNR: 12.03 dB



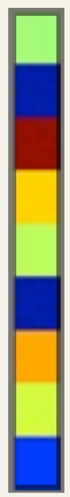
V. 1-bit CS

Joint work with:

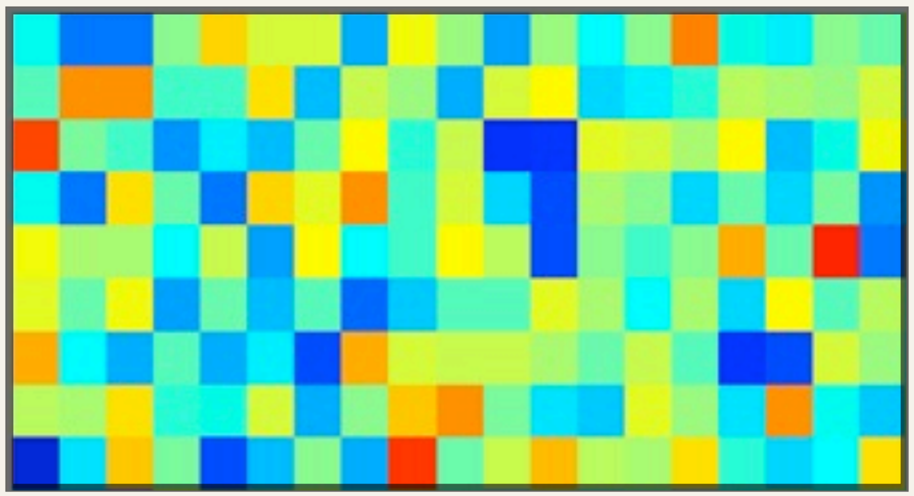
J. Laska (Rice U., USA)
P. Boufounos (MERL, USA)
R. Baraniuk (Rice U., USA)

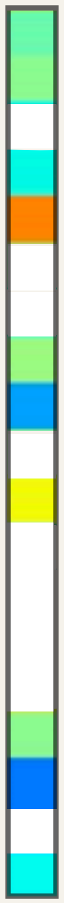
Compressed Sensing

$\boldsymbol{y}$ $\qquad\qquad\qquad$ $\boldsymbol{\Phi}$ $\qquad$ $\boldsymbol{x}$



$=$

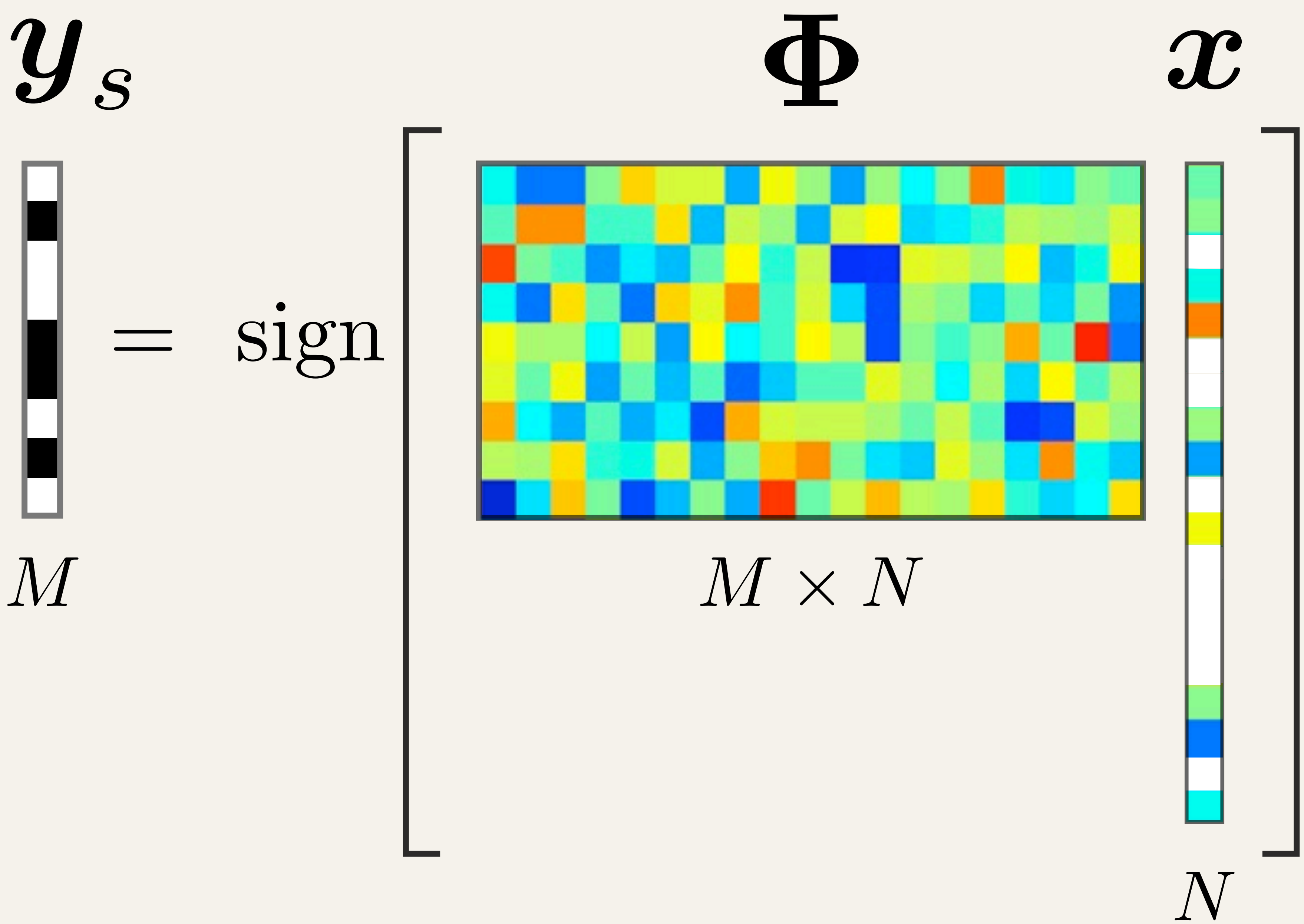




M $\qquad\qquad\qquad$ $M \times N$

N

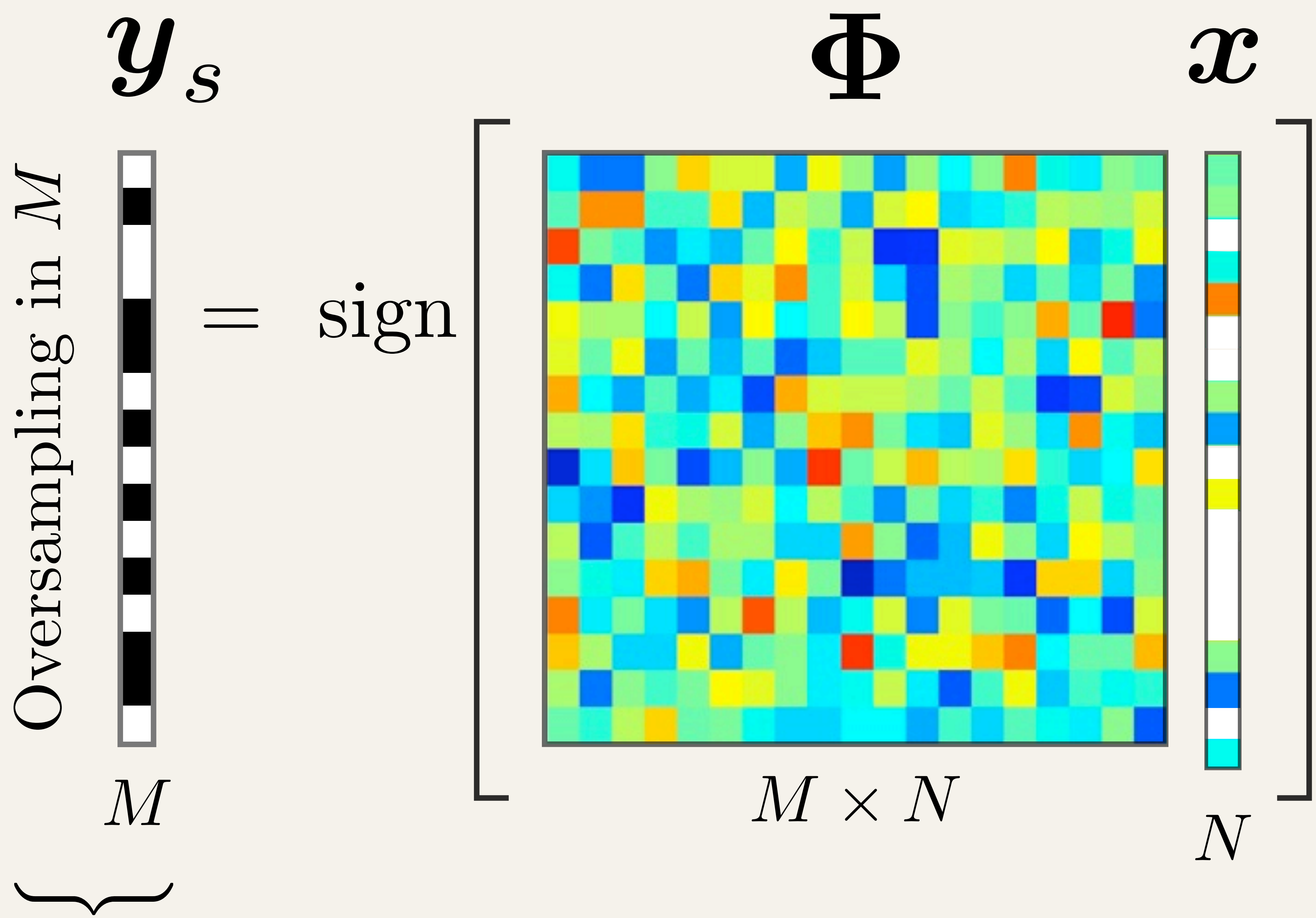
1-bit Compressed Sensing



$$\boldsymbol{y}_s = \operatorname{sign}\left[\boldsymbol{\Phi}\boldsymbol{x}\right]$$

with: $\operatorname{sign} t = \begin{cases} 1 & \text{if } t > 0 \\ -1 & \text{if } t \leqslant 0 \end{cases}$ component-wise

1-bit Compressed Sensing

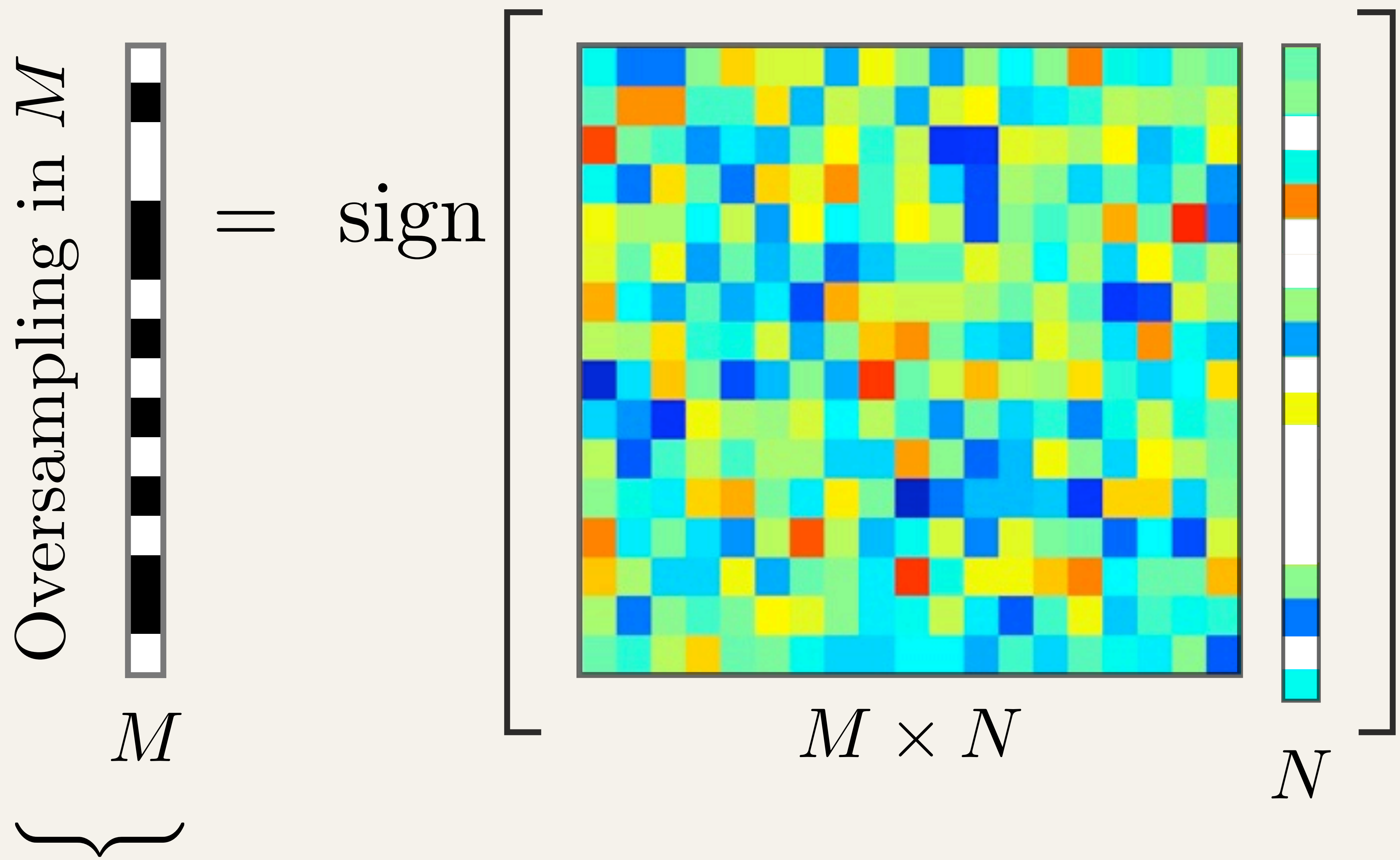


M-bits! But, which information inside $\boldsymbol{y}_s$?

1-bit ~~Compressed~~ Computational Sensing

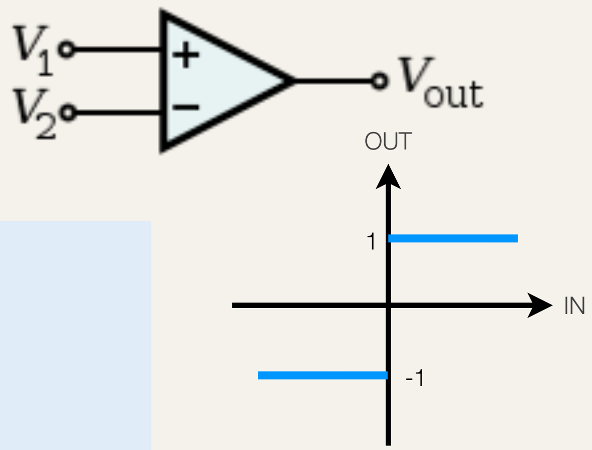
$\boxed{\text{bits matter!}}$

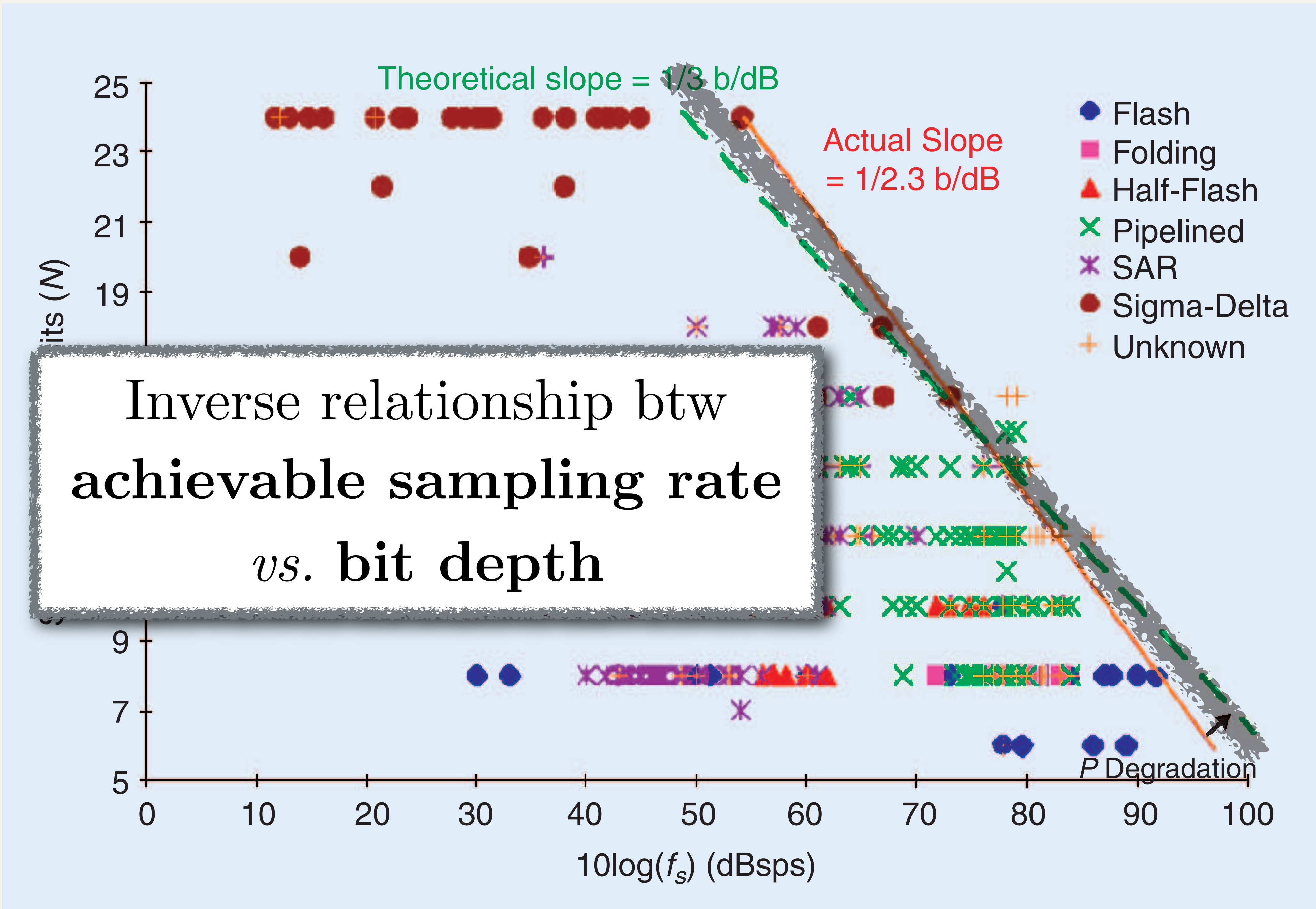
$$\underbrace{\boldsymbol{y}_s}_{M} = \mathrm{sign}\left[\boldsymbol{\Phi}\,\boldsymbol{x}\right]$$



M-bits! But, which information inside $\boldsymbol{y}_s$?

Why 1-bit? Very Fast Quantizers!





Inverse relationship btw
achievable sampling rate
vs. **bit depth**

[FIG1] Stated number of bits versus sampling rate.

[From "Analog-to-digital converters" B. Le, T.W. Rondeau, J.H. Reed, and C.W.Bostian, IEEE Sig. Proc. Magazine, Nov 2005]

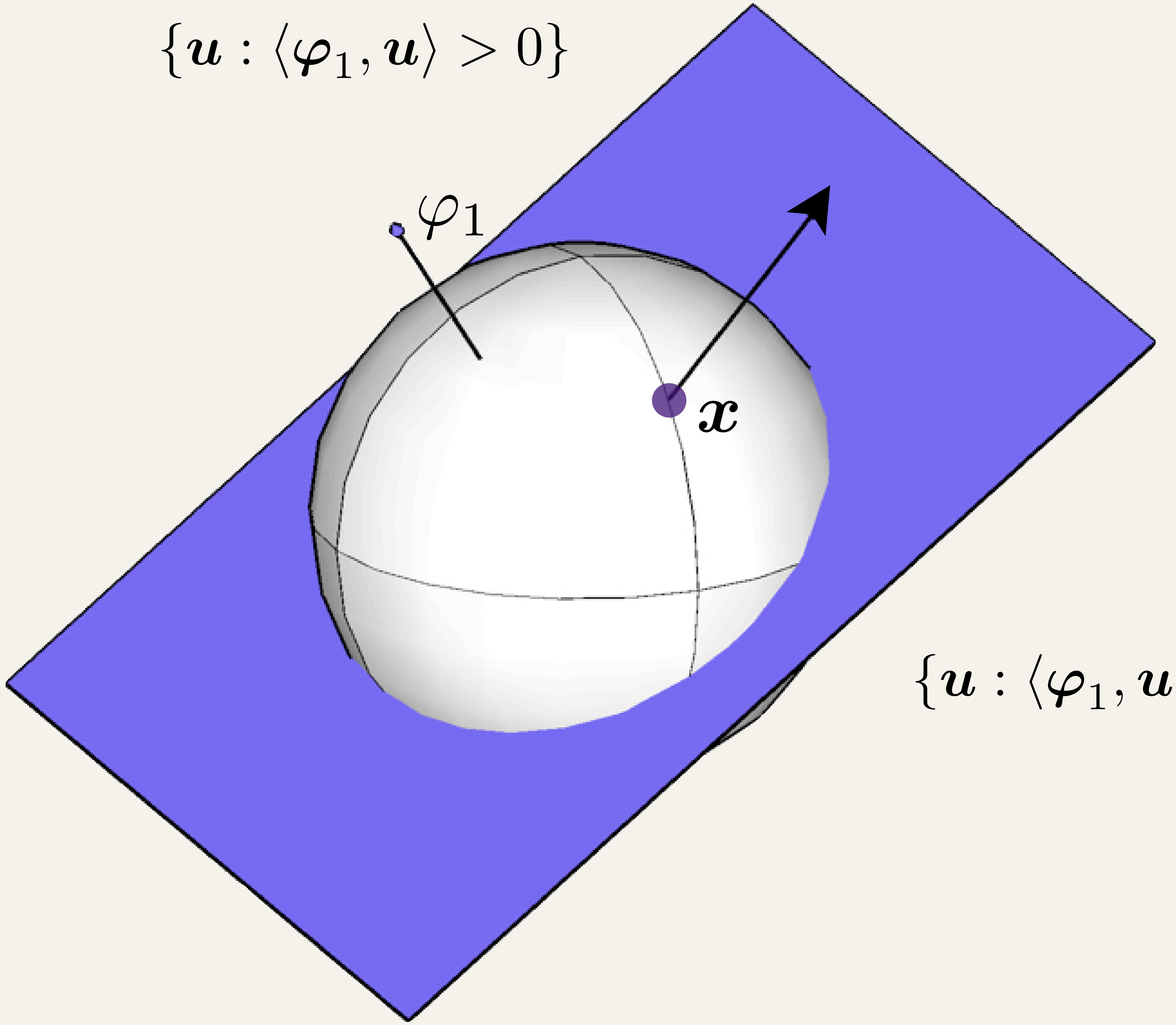
Intuitively ... $\boldsymbol{x}$ on S^2, M vectors $\{\boldsymbol{\varphi}_i : 1 \leqslant i \leqslant M\}$

Intuitively ... $\boldsymbol{x}$ on S^2, M vectors $\{\boldsymbol{\varphi}_i : 1 \leqslant i \leqslant M\}$

$\{\boldsymbol{u} : \langle \boldsymbol{\varphi}_1, \boldsymbol{u}\rangle > 0\}$

1-bit Measurements

$\langle \boldsymbol{\varphi}_1, \boldsymbol{x}\rangle > 0$



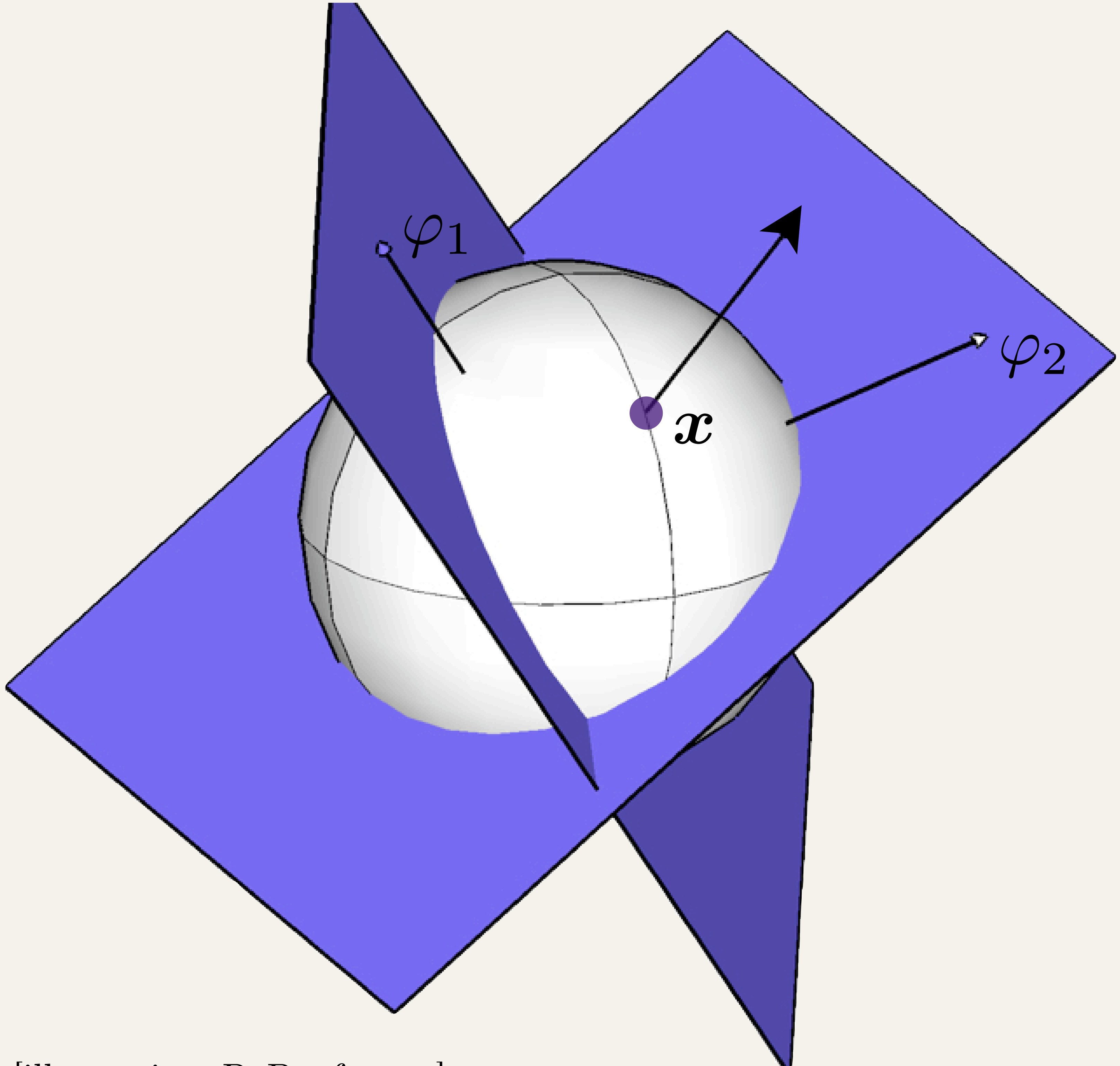
$\{\boldsymbol{u} : \langle \boldsymbol{\varphi}_1, \boldsymbol{u}\rangle \leqslant 0\}$

[illustration: P. Boufounos]

Intuitively ... $\boldsymbol{x}$ on S^2, M vectors $\{\boldsymbol{\varphi}_i : 1 \leqslant i \leqslant M\}$

1-bit Measurements

$\langle \boldsymbol{\varphi}_1, \boldsymbol{x} \rangle > 0$
$\langle \boldsymbol{\varphi}_2, \boldsymbol{x} \rangle > 0$

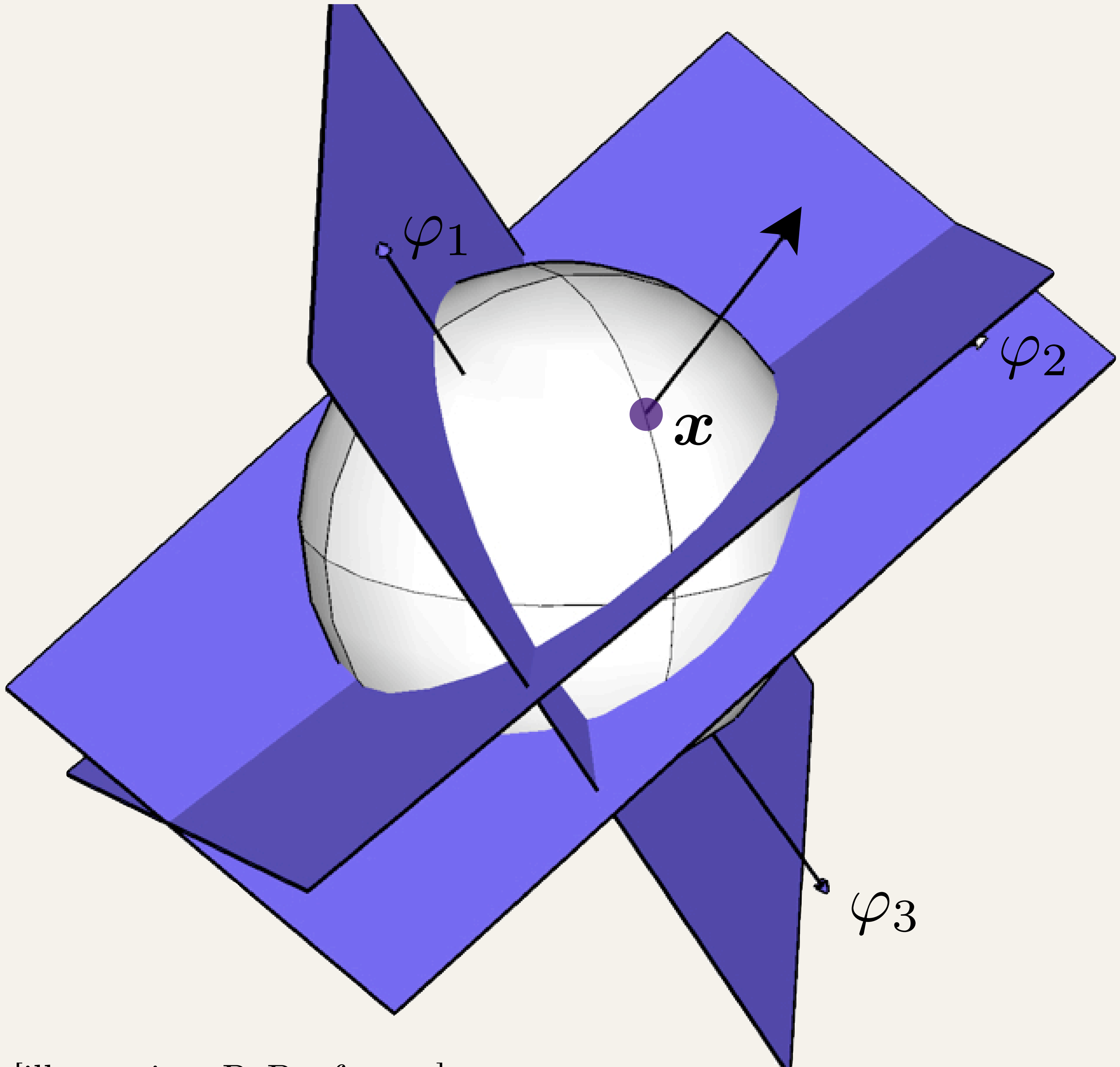


[illustration: P. Boufounos]

Intuitively ... $\boldsymbol{x}$ on S^2, M vectors $\{\boldsymbol{\varphi}_i : 1 \leqslant i \leqslant M\}$

1-bit Measurements

$\langle \boldsymbol{\varphi}_1, \boldsymbol{x} \rangle > 0$
$\langle \boldsymbol{\varphi}_2, \boldsymbol{x} \rangle > 0$
$\langle \boldsymbol{\varphi}_3, \boldsymbol{x} \rangle \leqslant 0$

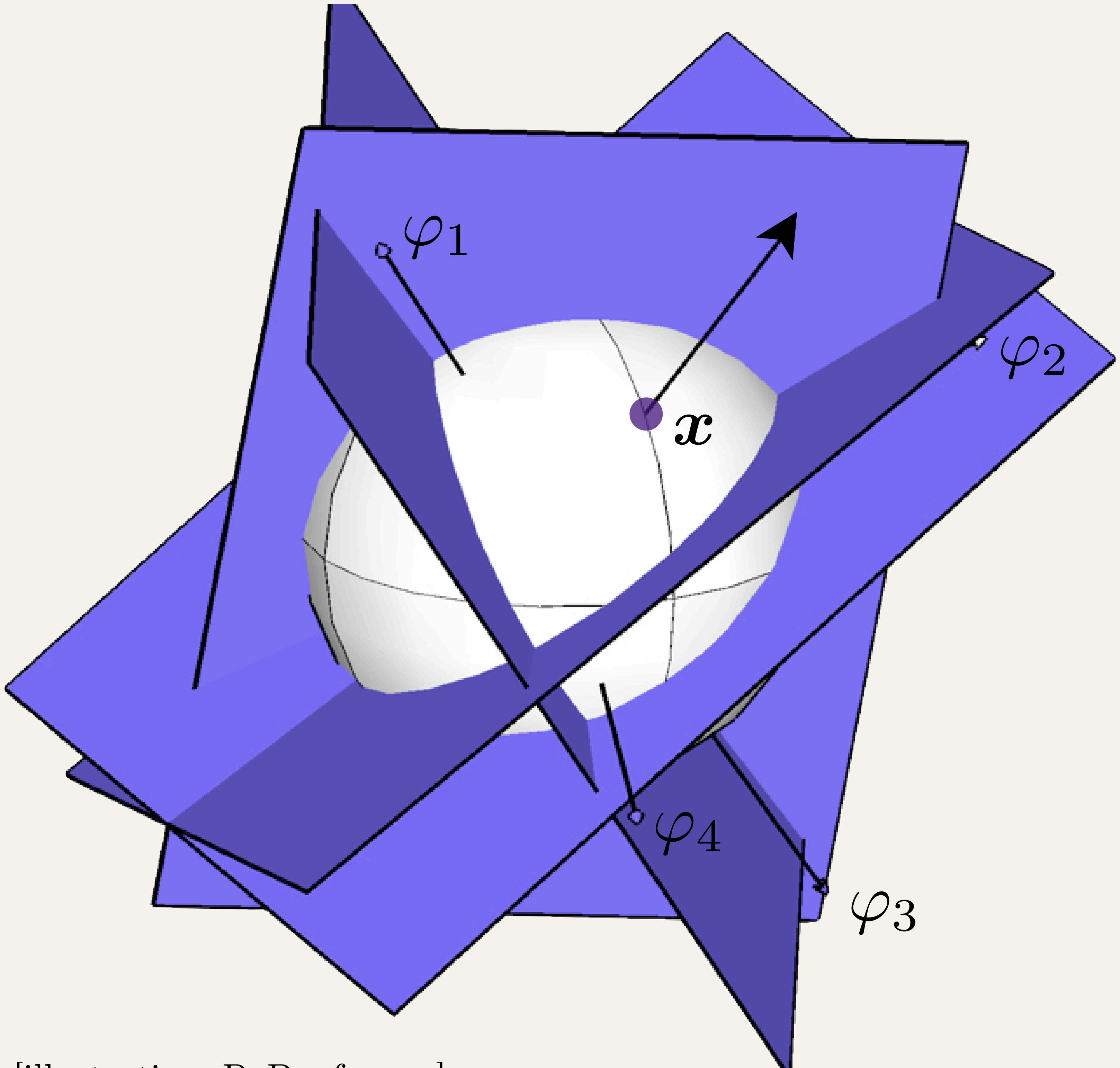


[illustration: P. Boufounos]

Intuitively ... $\boldsymbol{x}$ on S^2, M vectors $\{\boldsymbol{\varphi}_i : 1 \leqslant i \leqslant M\}$

1-bit Measurements

$\langle \boldsymbol{\varphi}_1, \boldsymbol{x} \rangle > 0$
$\langle \boldsymbol{\varphi}_2, \boldsymbol{x} \rangle > 0$
$\langle \boldsymbol{\varphi}_3, \boldsymbol{x} \rangle \leqslant 0$
$\langle \boldsymbol{\varphi}_4, \boldsymbol{x} \rangle > 0$



[illustration: P. Boufounos]

Intuitively ... $\boldsymbol{x}$ on S^2, M vectors $\{\boldsymbol{\varphi}_i : 1 \leqslant i \leqslant M\}$

1-bit Measurements

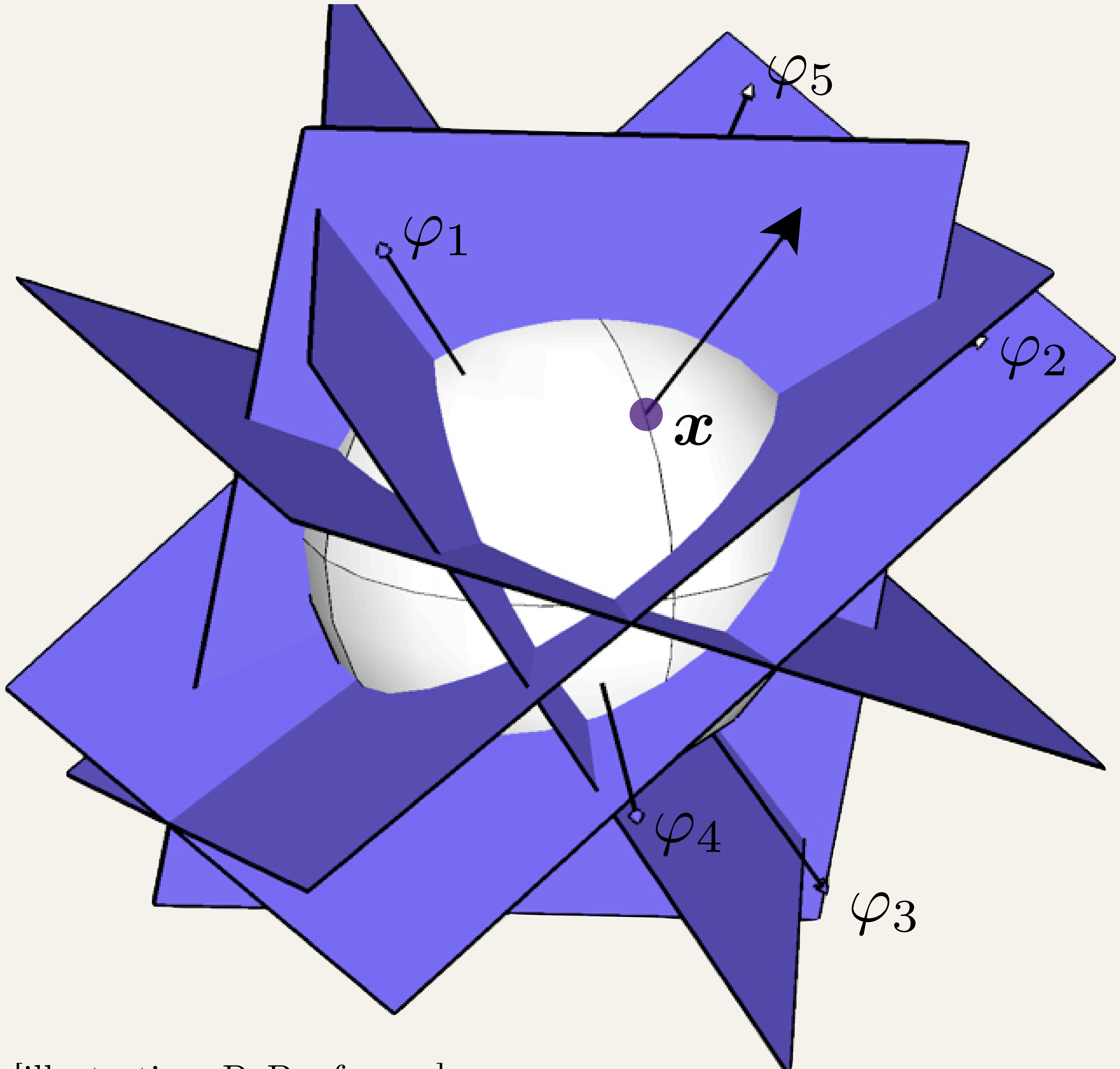
$\langle \boldsymbol{\varphi}_1, \boldsymbol{x} \rangle > 0$

$\langle \boldsymbol{\varphi}_2, \boldsymbol{x} \rangle > 0$

$\langle \boldsymbol{\varphi}_3, \boldsymbol{x} \rangle \leqslant 0$

$\langle \boldsymbol{\varphi}_4, \boldsymbol{x} \rangle > 0$

$\langle \boldsymbol{\varphi}_5, \boldsymbol{x} \rangle > 0$

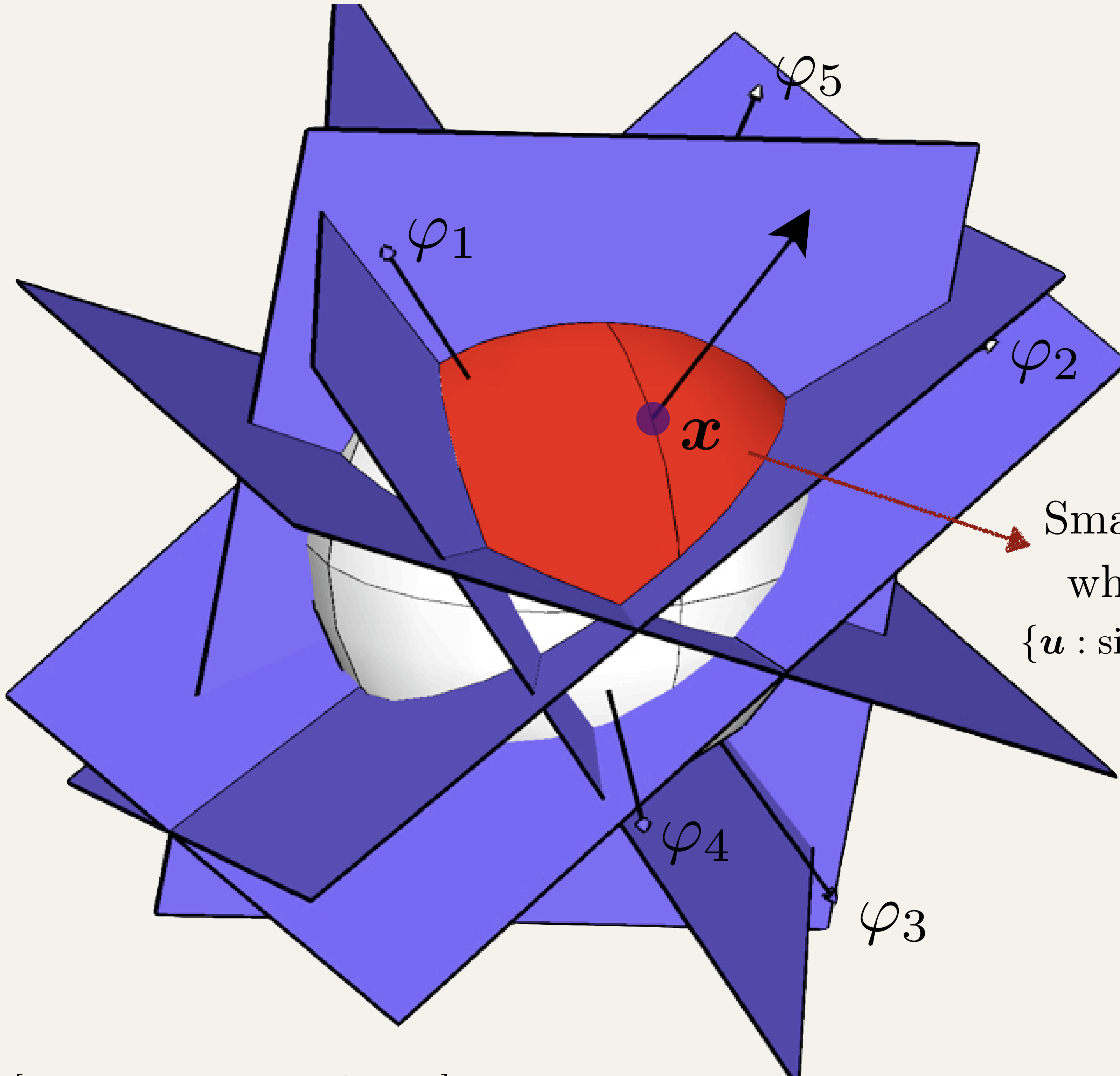


[illustration: P. Boufounos]

Intuitively ... $\boldsymbol{x}$ on S^2, M vectors $\{\boldsymbol{\varphi}_i : 1 \leqslant i \leqslant M\}$

1-bit Measurements

$$\begin{aligned}
&\langle \boldsymbol{\varphi}_1, \boldsymbol{x} \rangle > 0 \\
&\langle \boldsymbol{\varphi}_2, \boldsymbol{x} \rangle > 0 \\
&\langle \boldsymbol{\varphi}_3, \boldsymbol{x} \rangle \leqslant 0 \\
&\langle \boldsymbol{\varphi}_4, \boldsymbol{x} \rangle > 0 \\
&\langle \boldsymbol{\varphi}_5, \boldsymbol{x} \rangle > 0 \\
&\quad\vdots
\end{aligned}$$



Smaller and smaller when M increases

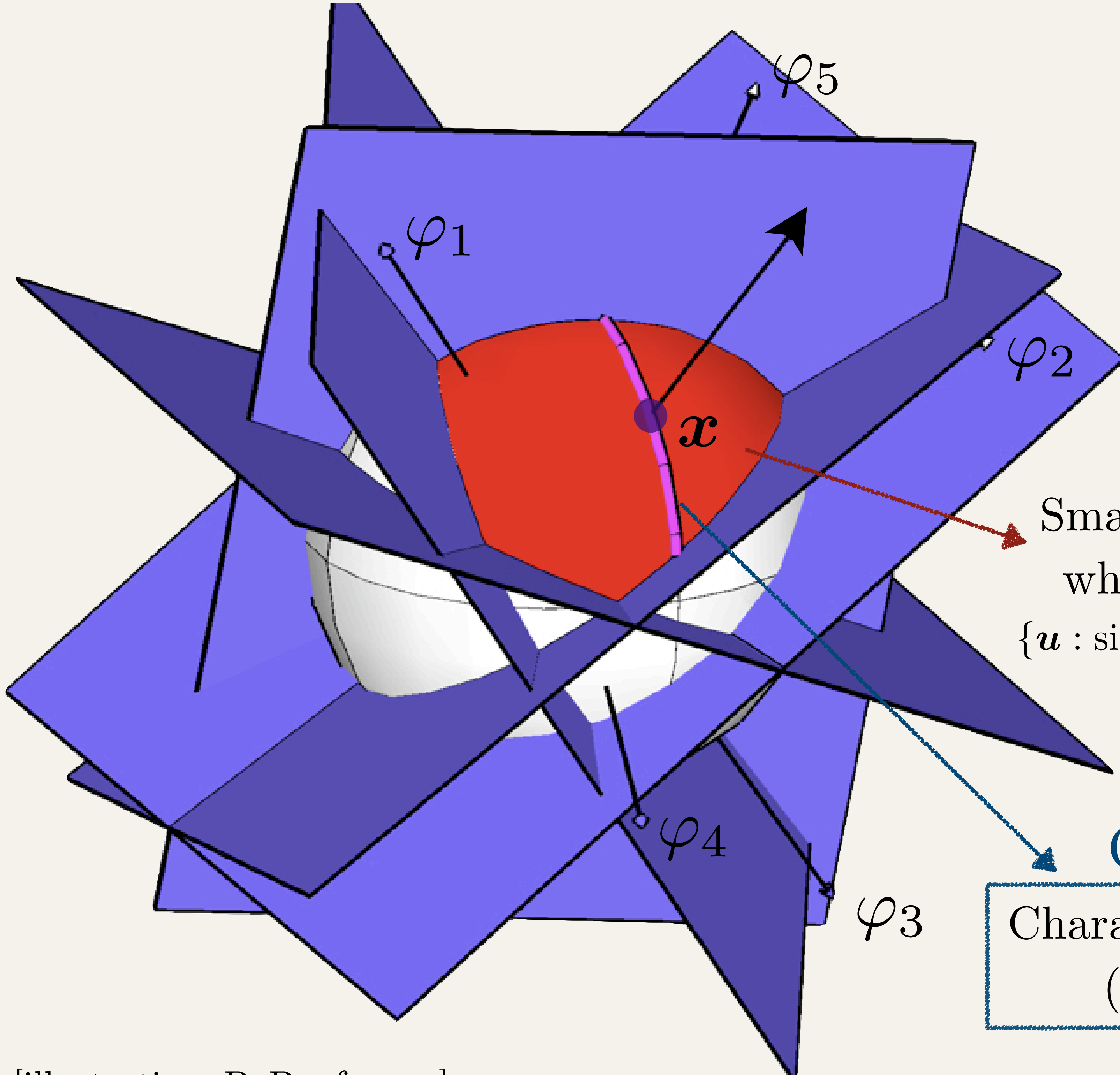
$\{\boldsymbol{u} : \operatorname{sign}(\boldsymbol{\Phi u}) = \operatorname{sign}(\boldsymbol{\Phi x})\}$

[illustration: P. Boufounos]

Intuitively ... $\boldsymbol{x}$ on S^2, M vectors $\{\boldsymbol{\varphi}_i : 1 \leqslant i \leqslant M\}$

1-bit Measurements

$\langle \boldsymbol{\varphi}_1, \boldsymbol{x} \rangle > 0$
$\langle \boldsymbol{\varphi}_2, \boldsymbol{x} \rangle > 0$
$\langle \boldsymbol{\varphi}_3, \boldsymbol{x} \rangle \leqslant 0$
$\langle \boldsymbol{\varphi}_4, \boldsymbol{x} \rangle > 0$
$\langle \boldsymbol{\varphi}_5, \boldsymbol{x} \rangle > 0$
$\vdots$



Smaller and smaller when M increases
$\{\boldsymbol{u} : \mathrm{sign}\,(\boldsymbol{\Phi u}) = \mathrm{sign}\,(\boldsymbol{\Phi x})\}$

Objective

Characterize the width (particular $\boldsymbol{x}$)

[illustration: P. Boufounos]

Binary Stable Embedding

$A(\boldsymbol{x}) := \mathrm{sign}\,(\boldsymbol{\Phi}\boldsymbol{x})$

Let $\boldsymbol{\Phi} \sim \mathcal{N}^{M\times N}(0,1)$, fix $0 \leqslant \eta \leqslant 1$ and $\epsilon > 0$. If

$$M \;\geqslant\; \tfrac{4}{\epsilon^2}\big(K\,\ln(N) + 2K\,\ln(\tfrac{50}{\epsilon}) + \ln(\tfrac{2}{\eta})\big),$$

then $\boldsymbol{\Phi}$ is a BϵSE with $\Pr > 1-\eta$.

$$\boxed{M = O(\epsilon^{-2} K \ln N)}$$

$$\boxed{d_{\rm ang}(\boldsymbol{x},\boldsymbol{s}) - \epsilon \leqslant d_H(A(\boldsymbol{x}),A(\boldsymbol{s})) \leqslant d_{\rm ang}(\boldsymbol{x},\boldsymbol{s}) + \epsilon}$$

Angular distance Hamming distance

Consistent + sparse reconstruction = bounded (angular) error !

$$\epsilon = O\big(\sqrt{K/M}\,\log(NM)\big)$$

Numerical Reconstructions:

* [Boufounos, Baraniuk 2008]

$$\boxed{\boldsymbol{x}^* = \arg\min_{\boldsymbol{u}} \|\boldsymbol{u}\|_1 \quad \text{s.t.} \quad \mathrm{diag}(A(\boldsymbol{x}))\, \Phi \boldsymbol{u} > 0 \quad \text{and} \quad \|\boldsymbol{u}\|_2 = 1}$$

$+$ other iterative methods: Matching Sign pursuit (MSP), Restricted-Step Shrinkage (RSS)

* **B**inary **I**terative **H**ard **T**hresholding (BIHT):

Given $\boldsymbol{y}_s = A(\boldsymbol{x})$ and K, set $l = 0$, $\boldsymbol{x}^0 = 0$: ($\tau > 0$ controls gradient descent)

$$\boxed{\begin{aligned} \boldsymbol{a}^{l+1} &= \boldsymbol{x}^l + \tfrac{\tau}{2} \Phi^T \big(\boldsymbol{y}_s - A(\boldsymbol{x}^l)\big), \\ \boldsymbol{x}^{l+1} &= \eta_K(\boldsymbol{a}^{l+1}), \quad l \leftarrow l+1 \end{aligned}}$$

("gradient" towards consistency)

(proj. K-sparse signal set)

Stop when $d_H(\boldsymbol{y}_s, A(\boldsymbol{x}^{l+1})) = 0$ or $l =$ max. iter.

with $\eta_K(\boldsymbol{u}) =$ best K-term approximation of $\boldsymbol{u}$

minimizes $\mathcal{J}(\boldsymbol{x}) = \|[\mathrm{diag}(\boldsymbol{y}_s)(\Phi\boldsymbol{x})]_-\|_1$, with $(\lambda)_- = (\lambda - |\lambda|)/2$

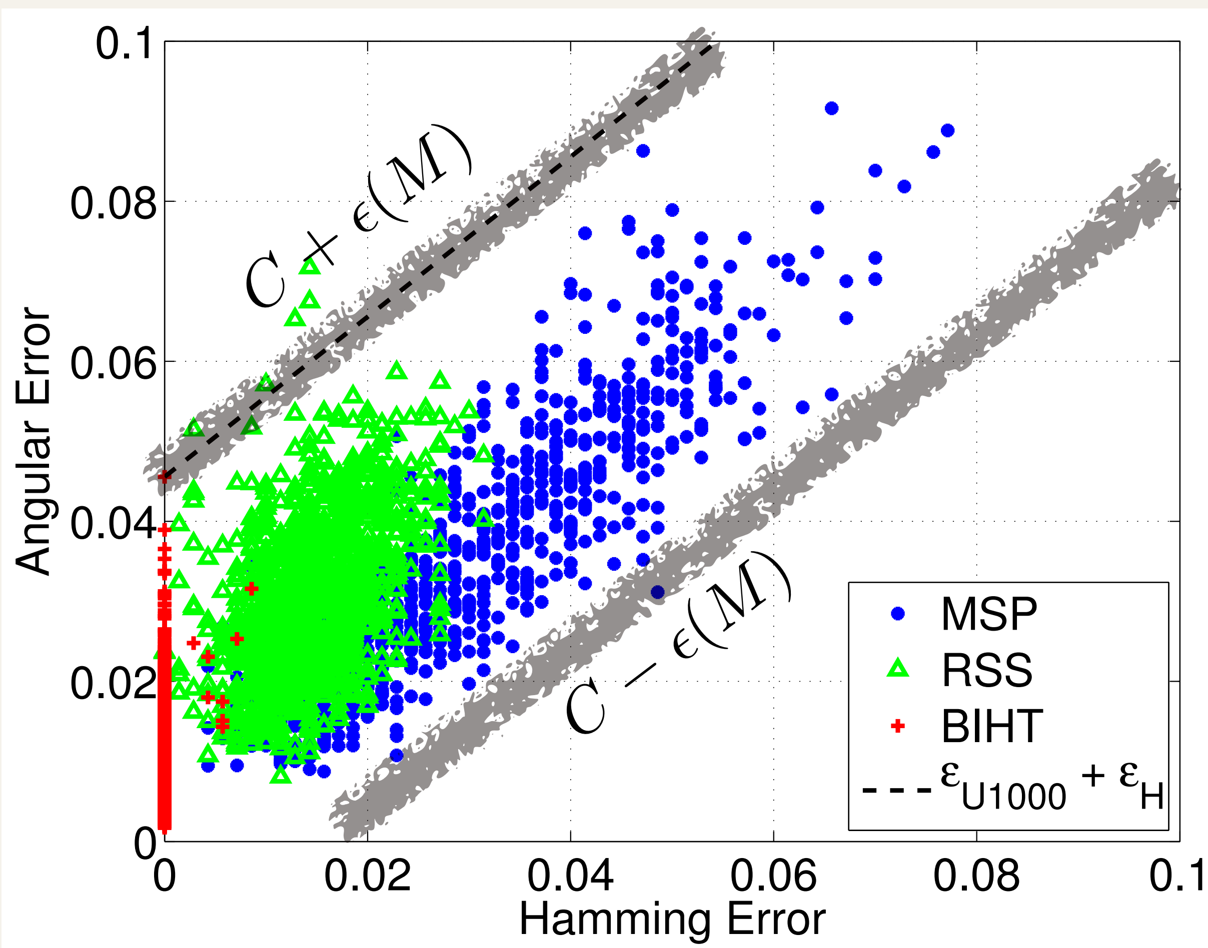
(connections with ML hinge loss, 1-bit classification)

Simulations: Testing BϵSE

$$d_{\text{ang}}(\boldsymbol{x}, \boldsymbol{x}^*) - \epsilon(M) \leqslant d_H(A(\boldsymbol{x}), A(\boldsymbol{x}^*)) \leqslant d_{\text{ang}}(\boldsymbol{x}, \boldsymbol{x}^*) + \epsilon(M)$$

$$\underbrace{d_H(\boldsymbol{x}, \boldsymbol{x}^*)}_{C} - \epsilon(M) \leqslant \underbrace{d_{\text{ang}}(A(\boldsymbol{x}), A(\boldsymbol{x}^*))}_{D} \leqslant \underbrace{d_H(\boldsymbol{x}, \boldsymbol{x}^*)}_{C} + \epsilon(M)$$

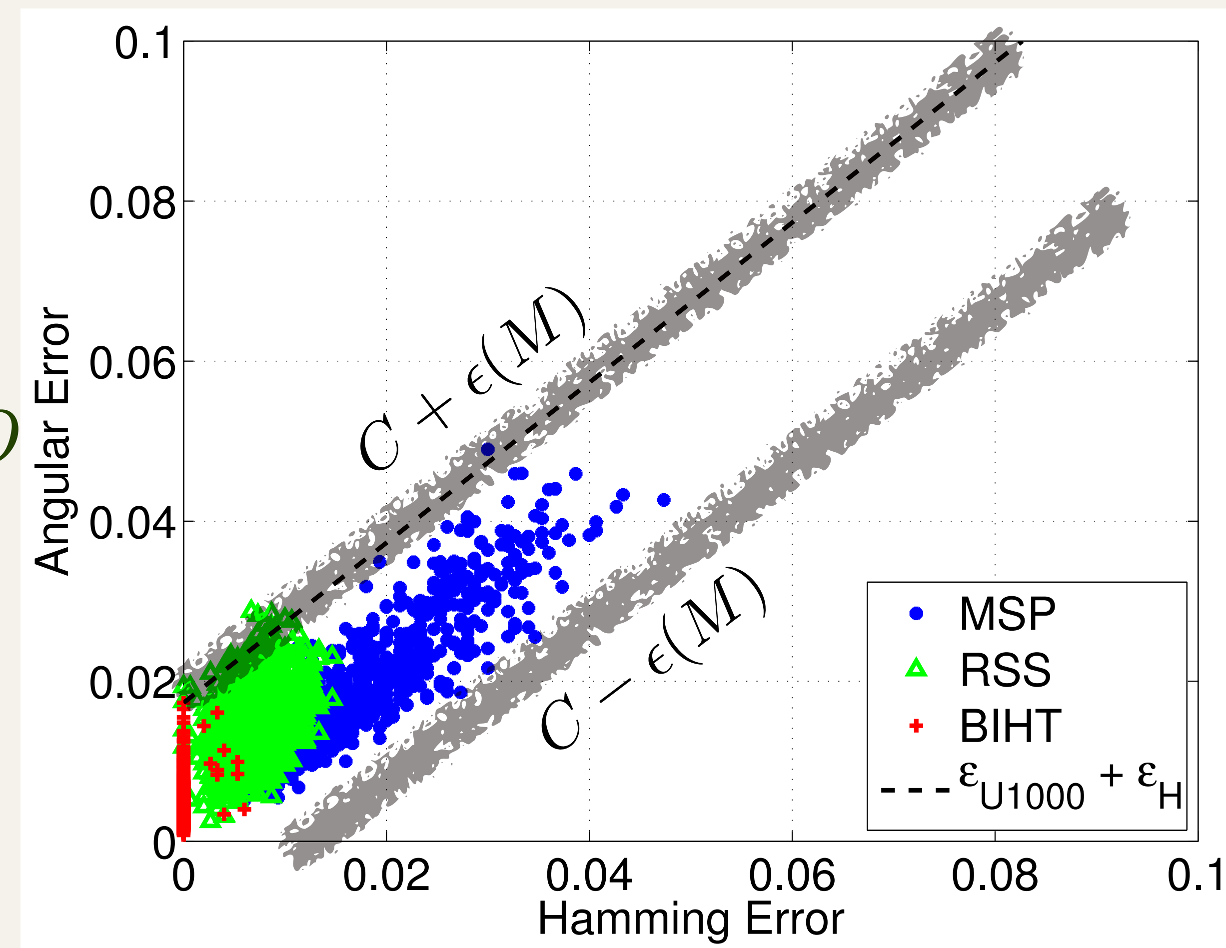
D



C

$M/N = 0.7$

D

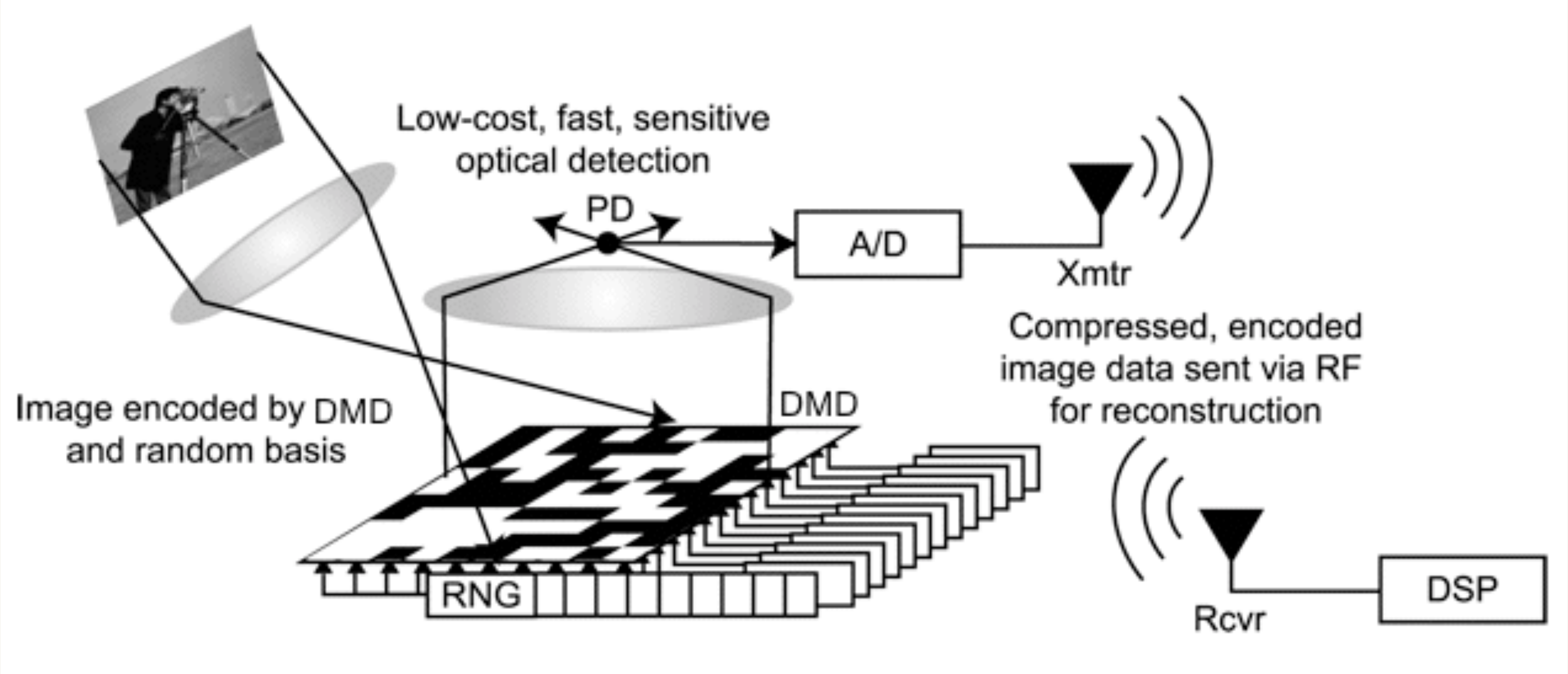


C

$M/N = 1.5$

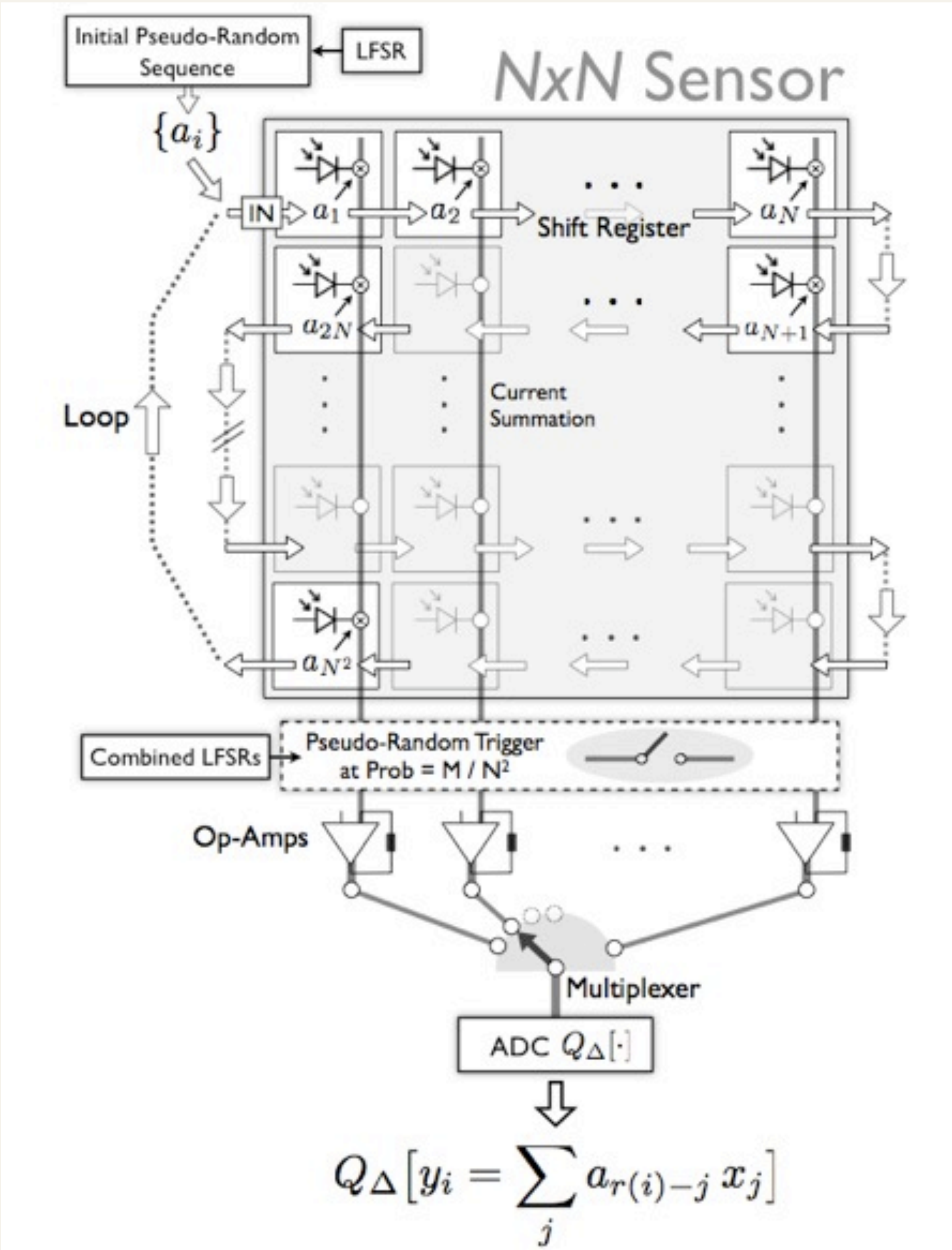
VI. CS ideas for imaging

"Rice One-pixel Camera"



Marco F. Duarte, Mark A. Davenport, Dharmpal Takhar, Jason N. Laska,
Ting Sun, Kevin F. Kelly and Richard G. Baraniuk

CS CMOS Camera

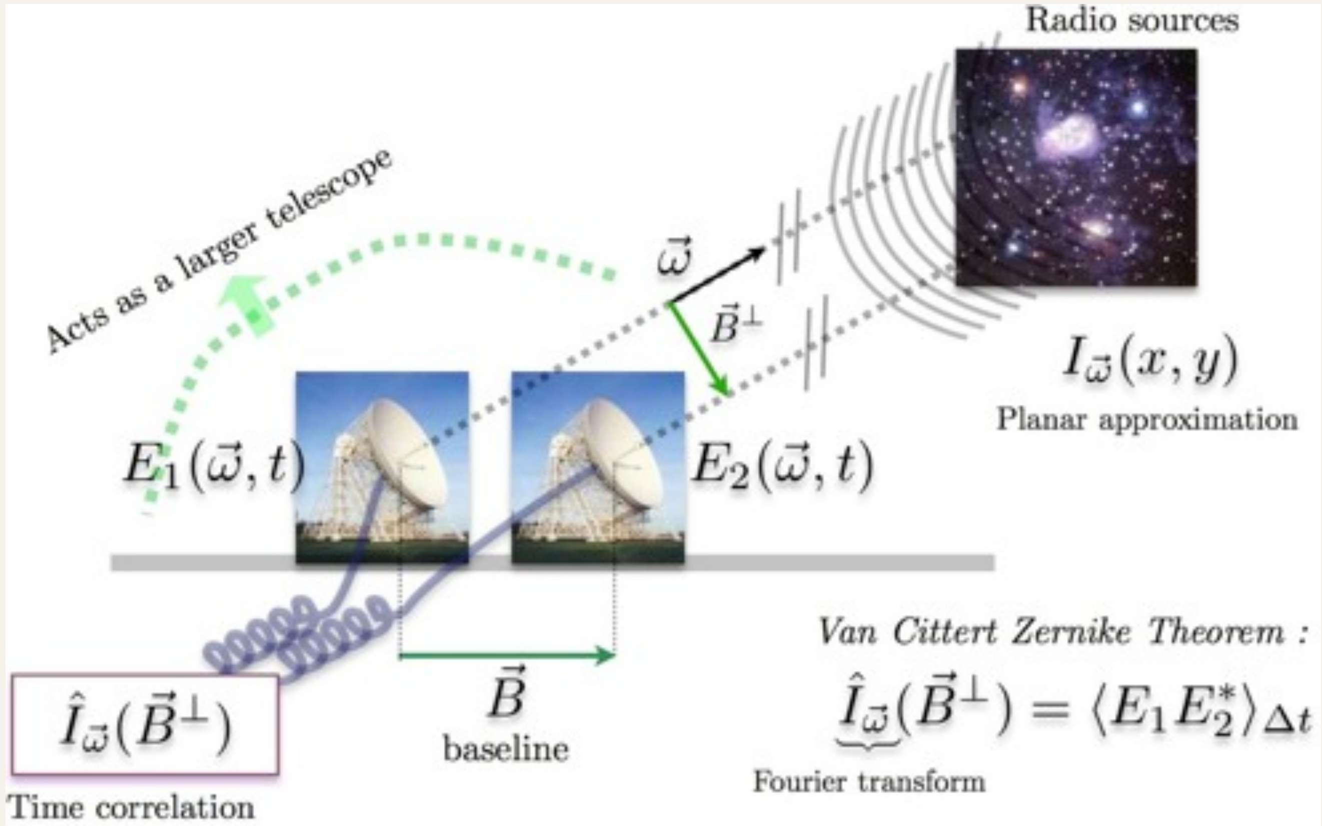


Analog Random Convolution

M=N/3, 11 bit + noise: 27dB

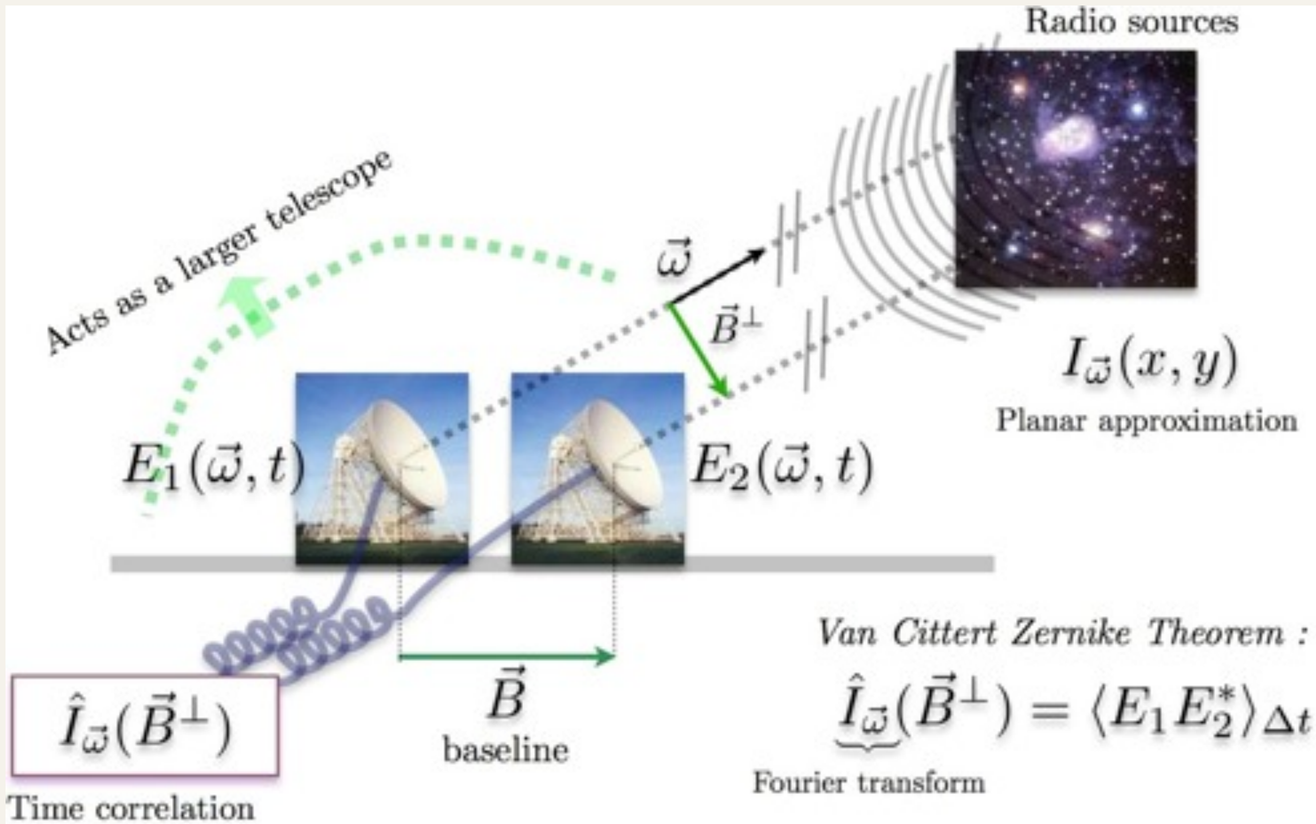
Joint work with: P. Vandergheynst, A. Schmid, Y. Leblebici, EPFL, Suisse

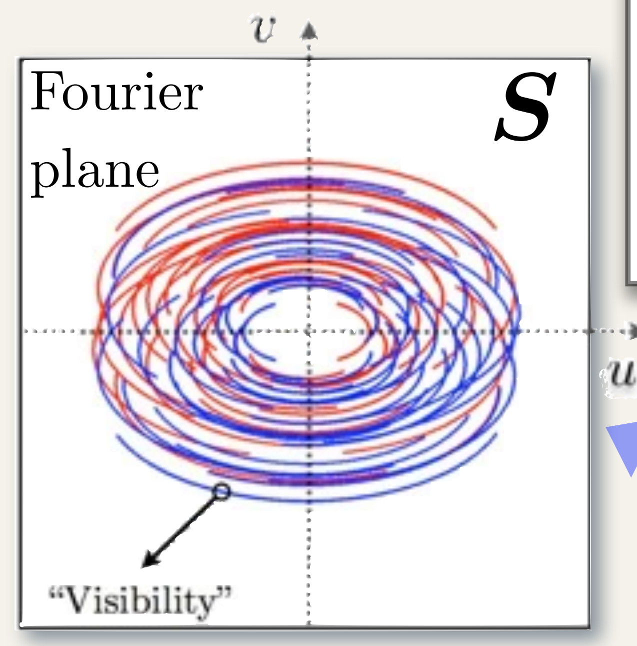
Radio-Interferometry



Joint work with: Y. Wiaux, P. Vandergheynst, EPFL, Suisse. A. Scaife, Cambridge, UK.

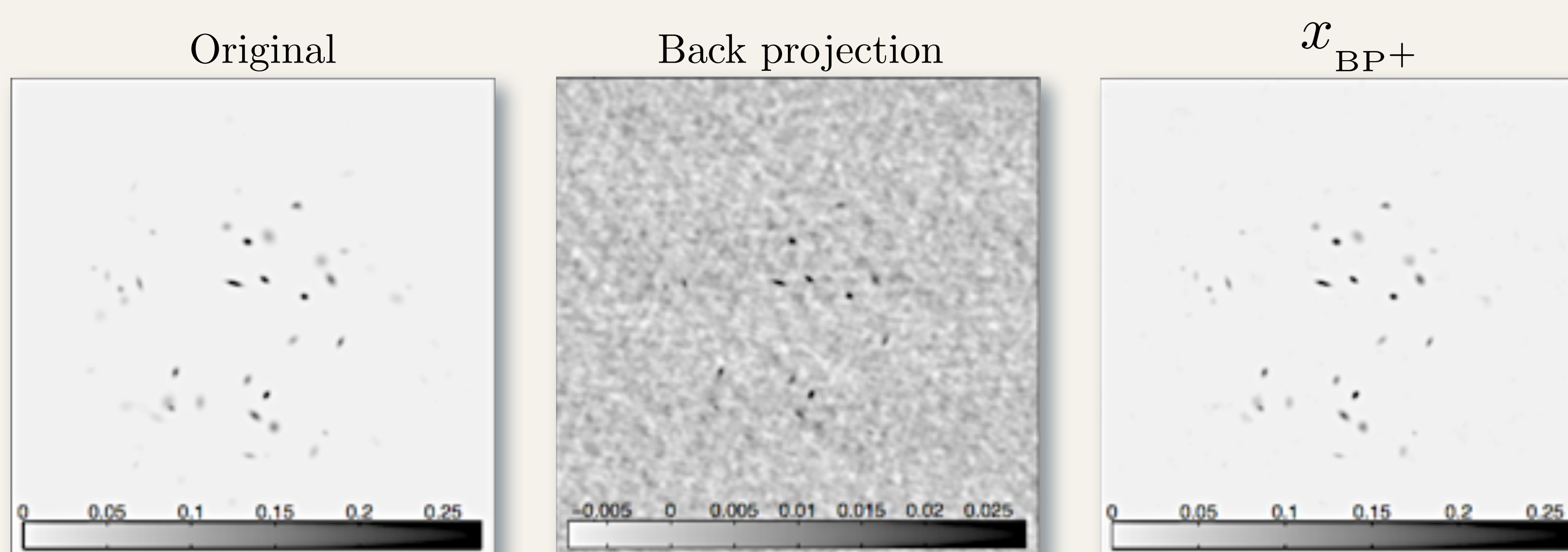
Radio-Interferometry





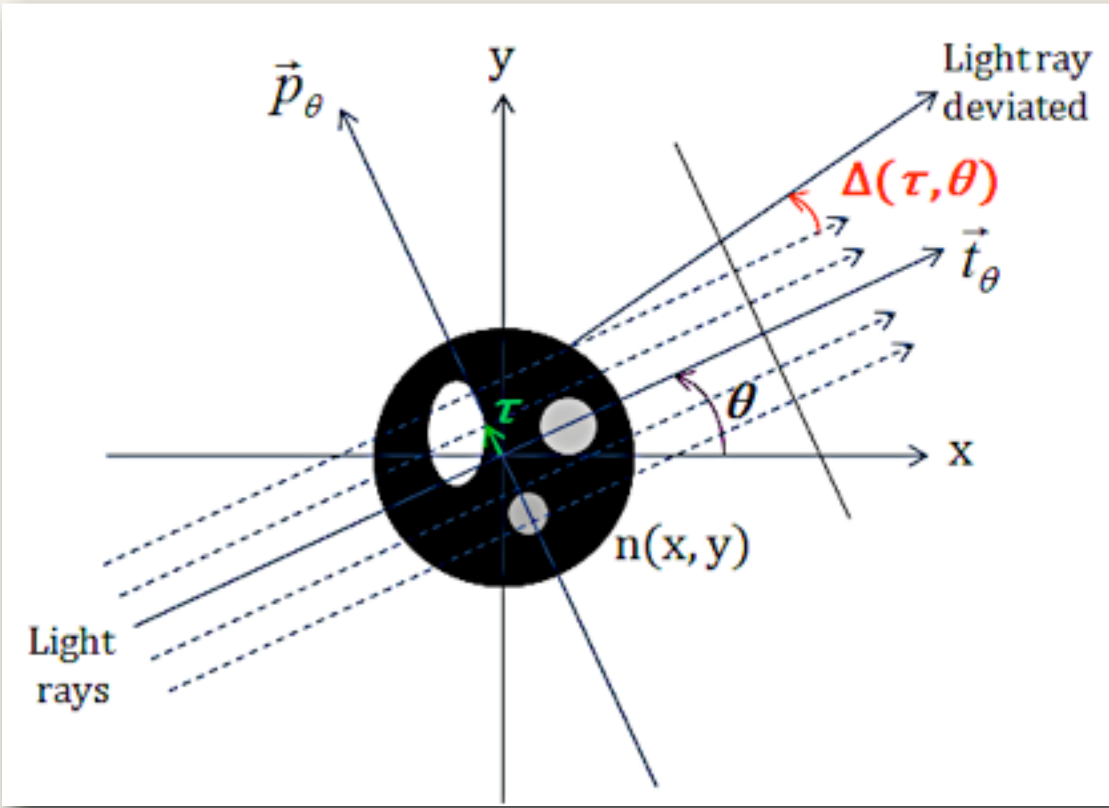
Fourier Sampling!

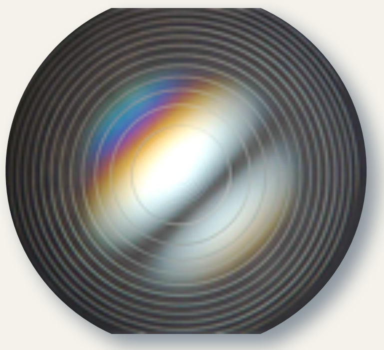
$$y_j = \hat{I}(\boldsymbol{k}_j)$$
$$\boldsymbol{\Phi} = \boldsymbol{S}\boldsymbol{F}$$
$$\boldsymbol{y} = \boldsymbol{\Phi}\boldsymbol{I} + \boldsymbol{\epsilon}$$



Joint work with: Y. Wiaux, P. Vandergheynst, EPFL, Suisse. A. Scaife, Cambridge, UK.

Optical Deflectometry





multifocal lens

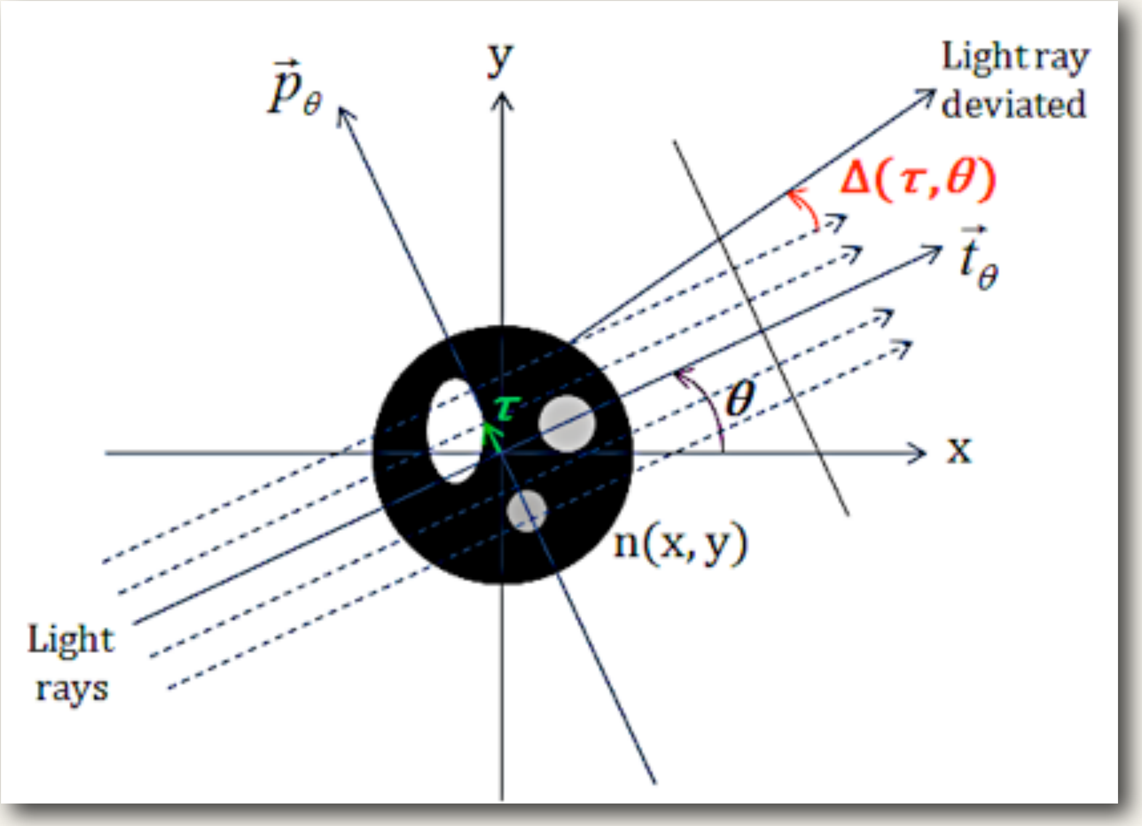
- Problem:

  Reconstructing the refractive index map of transparent materials from light deflection measurements under multiple orientations.

- Interests:
  - (transparent) object surface topology
  - default detection in glass, crystal growth
- Advantages of deflectometry:
  - Insensitive to vibrations (vs. interferometry)
  - Less sensitive to dispersive medium
  - Precision: up to 10nm flatness deviation on 50mm FOV

Joint work with: A. Gonzalez, Ph. Antoine, C. De Vleeschouwer (UCL)

Optical Deflectometry



$$\Delta(\tau,\theta) \simeq \int_{\mathbb{R}^2} \big(\boldsymbol{\nabla} n(\boldsymbol{x})\cdot \boldsymbol{p}_\theta\big)\ \delta(\tau - \boldsymbol{x}\cdot\boldsymbol{p}_\theta)\ \mathrm{d}^2\boldsymbol{x}$$

$=$ Weighted Fourier Sampling

$$y_j = \hat{\Delta}(\omega_j,\theta_j) = i(\mathrm{sign}\,\omega_j)\,\hat{n}(\omega_j,\theta_j)$$

$$\boldsymbol{y} = \boldsymbol{W}\boldsymbol{S}\boldsymbol{F}\,\boldsymbol{n} + \boldsymbol{\epsilon}$$

$\boldsymbol{S}$

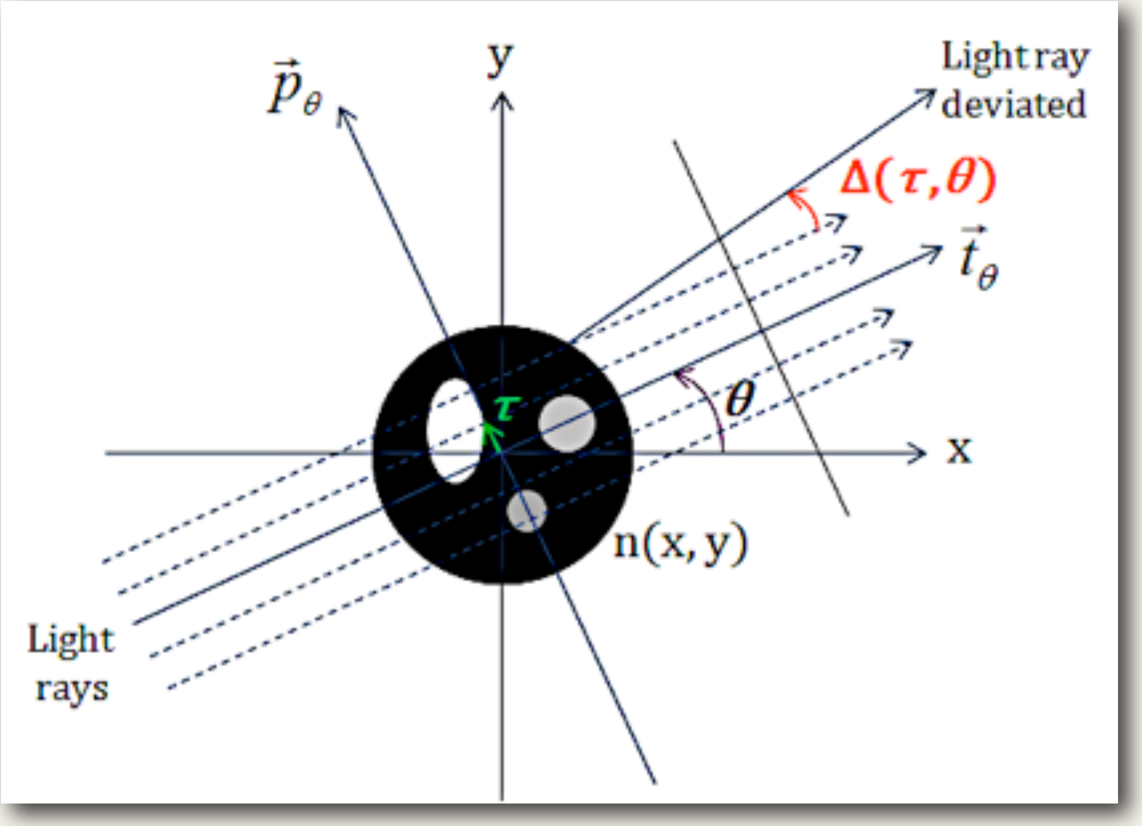
- Problem:

  Reconstructing the refractive index map of transparent materials from light deflection measurements under multiple orientations.

- Interests:
  - (transparent) object surface topology
  - default detection in glass, crystal growth
- Advantages of deflectometry:
  - Insensitive to vibrations (vs. interferometry)
  - Less sensitive to dispersive medium
  - Precision: up to 10nm flatness deviation on 50mm FOV

Joint work with: A. Gonzalez, Ph. Antoine, C. De Vleeschouwer (UCL)

Optical Deflectometry



$$\Delta(\tau,\theta) \simeq \int_{\mathbb{R}^2} (\boldsymbol{\nabla} n(\boldsymbol{x}) \cdot \boldsymbol{p}_\theta)\ \delta(\tau - \boldsymbol{x} \cdot \boldsymbol{p}_\theta)\ \mathrm{d}^2\boldsymbol{x}$$

$=$ Weighted Fourier Sampling

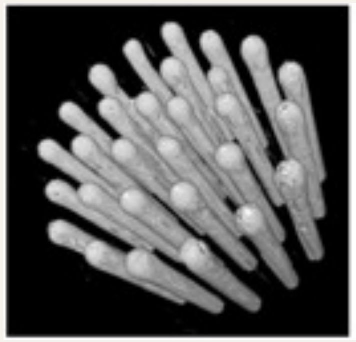
$$y_j = \hat{\Delta}(\omega_j, \theta_j) = i(\mathrm{sign}\,\omega_j)\,\hat{n}(\omega_j, \theta_j)$$

$$\boldsymbol{y} \;=\; \boldsymbol{W}\boldsymbol{S}\boldsymbol{F}\,\boldsymbol{n} + \boldsymbol{\epsilon}$$

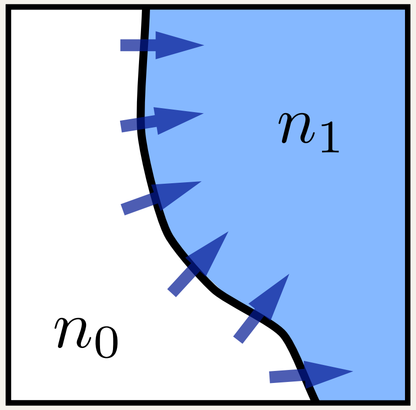
$\boldsymbol{S}$

Sparsity model - TV

$$\mathcal{S}(n) = \|n\|_{TV}$$

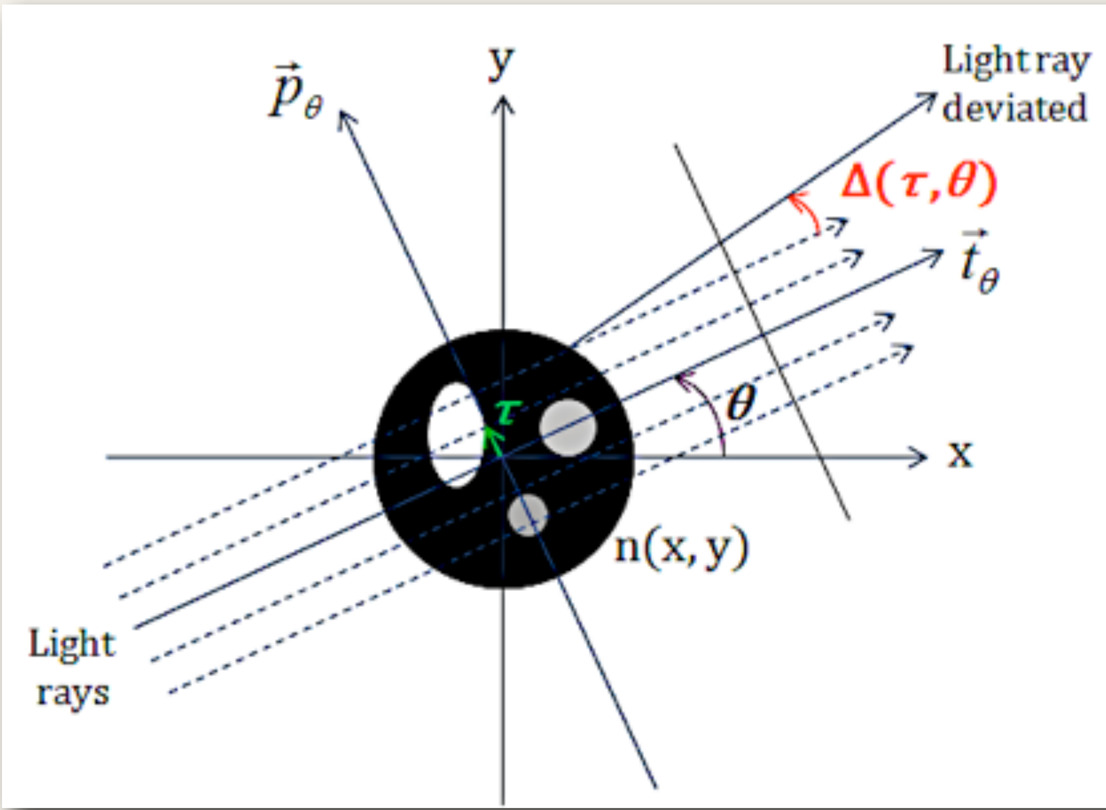


$$= \int_{\mathbb{R}} \|\boldsymbol{\nabla} n(\boldsymbol{x})\|\,\mathrm{d}\boldsymbol{x}$$



Joint work with: A. Gonzalez, Ph. Antoine, C. De Vleeschouwer (UCL)

Optical Deflectometry



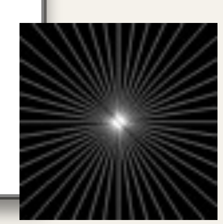
$$\Delta(\tau,\theta) \simeq \int_{\mathbb{R}^2} \big(\boldsymbol{\nabla} n(\boldsymbol{x}) \cdot \boldsymbol{p}_\theta\big)\ \delta(\tau - \boldsymbol{x}\cdot\boldsymbol{p}_\theta)\ \mathrm{d}^2\boldsymbol{x}$$

$= \text{Weighted Fourier Sampling}$

$$y_j = \hat{\Delta}(\omega_j,\theta_j) = i(\mathrm{sign}\,\omega_j)\,\hat{n}(\omega_j,\theta_j)$$

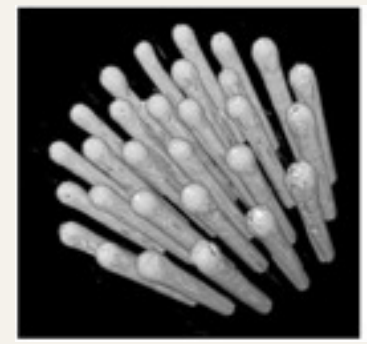
$$\boldsymbol{y} \;=\; \boldsymbol{W}\boldsymbol{S}\boldsymbol{F}\,\boldsymbol{n} + \boldsymbol{\epsilon}$$

$\boldsymbol{S}$

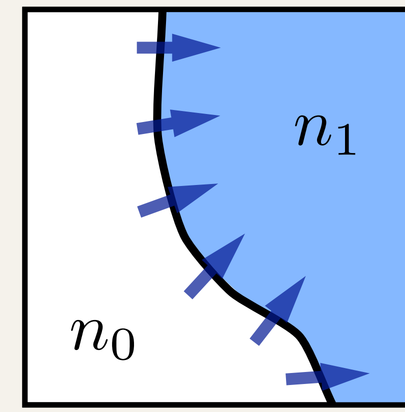


Sparsity model - TV

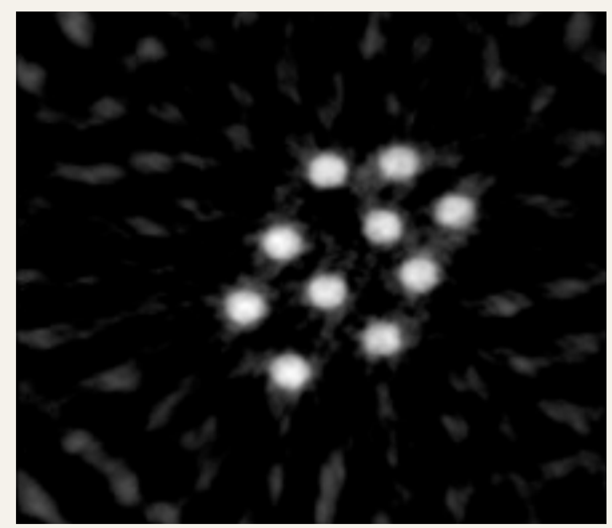
$$\mathcal{S}(n) = \|n\|_{TV}$$



$$= \int_{\mathbb{R}} \|\boldsymbol{\nabla} n(\boldsymbol{x})\|\, \mathrm{d}\boldsymbol{x}$$



$M/N = 10\%$



FBP

TV Rec.

Joint work with: A. Gonzalez, Ph. Antoine, C. De Vleeschouwer (UCL)

Conclusion

Conclusion

- Sparsity prior involves new sensing methods:
  e.g., Compressed Sensing, Compressive Imaging.
- Future:
  - more sensing examples: http://nuit-blanche.blogspot.com
    hyperspectral, network, GPR, Lidar, ... (explosion)
  - better sparsity prior:
    structured, model-based, mixed-norm (Cevher, Bach, ...)
    co-sparsity/analysis model (Gribonval, Nam, Davies, Elad, Candes)
  - non-linear CS?
    1-bit CS is one instance, phase recovery (Candès),
    polychromatic CT, ...

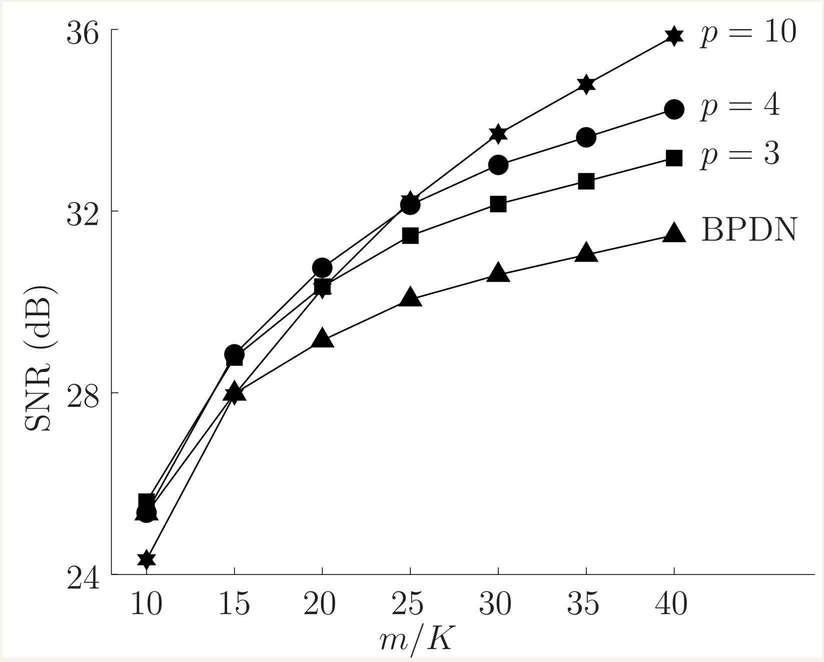
Thank you!

CS Links:

- http://dsp.rice.edu/cs
- http://dsp.rice.edu/1bitCS/
- http://nuit-blanche.blogspot.com/

Appendix

Numerical illustration



M

- ∗ N=1024, K=16, Gaussian Φ , 80 quant. bins
- ∗ 500 K-sparse (canonical basis)
- ∗ Non-zero components follow $\mathcal{N}(0,1)$
- ∗ BPDQ$_p$ solved by proximal optimization and operator splitting (Douglas - Rachford)

http://wiki.epfl.ch/bpdq

Histograms of
$\alpha^{-1}(y - \Phi x_p^*)_i$